Volume holographic storage using the 90° geometry

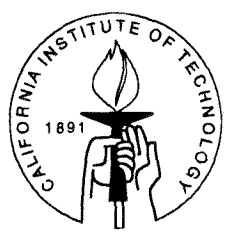
Thesis by

Geoffrey W. Burr

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

1996

(Submitted February 8, 1996)

Acknowledgements

I would like to thank my advisor, Professor Demetri Psaltis, for his advice and guidance over the last four years. During my first few years, I learned much as I struggled to find answers to his questions. Later, I learned to pay attention to the questions he chose to ask, both of me and of others. The opportunity to learn and to work within his group of talented students, and to observe firsthand the high standards of quality control he applies to his talks and papers, have been invaluable. I thank him for keeping me pointed in the right direction, and I apologize for sometimes making it such hard work for him.

I'd like to thank Dr. Sidney Li for teaching me the basics of volume holography and experimental research, and for his example as an all–around researcher and first–class person. I am indebted to Dr. Fai Mok for getting me started on his mirror array project, and for the many long discussions on holographic storage. The sections in Chapter 2 on recording schedules and M/# are modifications and extensions of his ideas. I thank Dr. Kevin Curtis for teaching me about the strength of willpower and of course, for comic relief. I'd like to thank Raanan Miller and his advisor, Dr. Yu–Chong Tai, for involving us in their micromirror projects. Their patience and flexibility were greatly appreciated. Many thanks to George Barbastathis for the long conversations on noise and probability of error, and to our technician, Ya–yun Liu, for helping me fumble through the crystal polishing and high temperature annealing. And I am greatly indebted to Xin An for his close collaboration over the last few months, and his essential help in closing out the 160,000 hologram project.

I'd like to thank all my other friends and colleagues in the Optical Information Processing Group, past and present, for their friendship and support: Dr. Chuanyi Ji, Dr. Subrata Rakshit, Dr. Yong Qiao, Dr. David Marx, Robert Denkewalter, Jean–Jacques Drolet, Ernest Chuang, Michael Levene, Allen Pu, Greg Billock, Dr. Chuan Xie, Xu Wang, Ali Adibi, George Ouyang, and Greg Steckman. I also appreciate the

advice and encouragement I've received at conferences from the extended family of Psaltis graduates, including Drs. John Hong, Kelvin Wagner, Claire Gu, Eung-Gi Paek, Mark Neifeld, and David Brady. In terms of getting things done at Caltech, I am indebted to Lucinda Acosta, Helen Carrier, Su McKinley, and Lyn Hein for favors too numerous to mention.

I was very fortunate to have several non-research activities during my stay at Caltech. For musical distractions, I thank Allen Gross and Delores Bing and the many Caltech and community musicians with whom I was able to play and perform. For athletic distractions, I thank Wendell Jack, Mike Vicic, and Dan Vaughn for help with planning GSC Soccer, and the dozens of soccer players I played with and against.

I owe a tremendous amount to the guidance, support, and advice of my parents, Charles and Phyllis. They taught me that there can be no shame in failure when you have given your absolute best effort. Anything else I might write about their impact on my life would be an understatement.

Finally, I want to thank my wife, Annette Grot, who is also a graduate of the Psaltis group. Her patience and understanding concerning our 300-mile separation over the last year has been amazing, and her love and support unbounded. While she has visited often to bolster my spirits, you'll see in Chapter 6 that she did the etching of our random phase plates as well. I couldn't hope for a better lifetime partner.

Abstract

Volume holographic data storage involves the superposition and independent recall of multiple pages of data within the same volume of a storage medium. These pages, stored as separate holograms, can be accessed by changing the angle of the reference laser beam used to store and retrieve them. Because all the data in a stored page is read out in parallel, the output data rate can be very large. At the same time, large storage capacity is available through the superposition of many data pages.

The topic of this thesis is volume holographic memories using the 90° geometry. This configuration, where signal and reference beams enter orthogonal crystal faces, is attractive for angle multiplexing because of its high angular selectivity. We choose angle multiplexing because it gives us many options for rapid steering of the reference beams. Our goal is to develop read–write holographic memories which achieve high capacity *and* high output data rate.

Our approach, in terms of recording material, is to work with what we have. In our case, the only photorefractive widely available in thicknesses greater than a centimeter is Fe–doped $LiNbO_3$. This material is relatively easy to make with high optical quality, and its performance shows no degradation after repeated record/thermal erase cycles. The disadvantages of $LiNbO_3$:Fe include volatility of storage, which we will treat briefly, and poor dynamic range, which we will discuss extensively in the first part of the thesis. We start in Chapter 2 with a study of dynamic range in holographic storage, in order to determine what is required of a photorefractive crystal. One of the outcomes of this study is a concise metric—which we call the $M/\#$—for measuring the dynamic range performance of a holographic storage system. Chapter 3 discusses the experimental measurement of this $M/\#$ as a function of the oxidation state of $LiNbO_3$:Fe. We find that there exists an optimal oxidation state (for maximum dynamic range performance), and in Chapter 4 we develop a theoretical model which predicts this optimum. In the remainder of Chapter 4, we extend this model to other

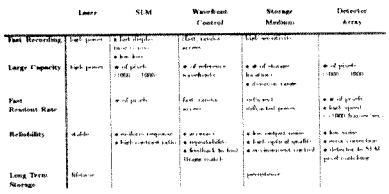
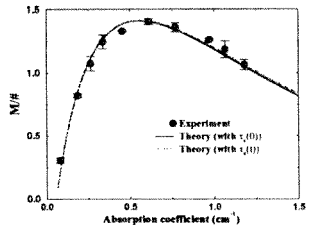
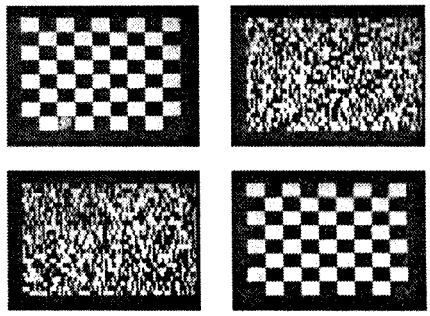
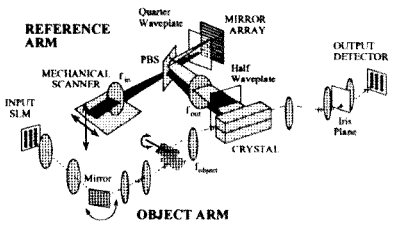
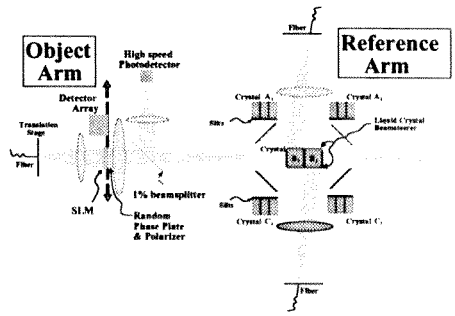
parameters such as crystal size, doping, and modulation depth.

Having squeezed as much performance as possible from our storage material, we turn to the design of a large–scale holographic memory. Our goal is to use angle, fractal, and spatial multiplexing to achieve large capacity—without sacrificing fast access to the stored holograms. In Chapter 5, we discuss our segmented mirror array, and how it makes such a design possible. Then in Chapter 6, we experimentally demonstrate the various features of this memory design. These demonstrations include storage using the mirror array, storage of 1000 holograms using an acousto–optic deflector, storage of 10,000 holograms in the same $\sim$ 1cm^3 volume of $LiNbO_3$, and the demonstration of the 160,000 hologram system with the mirror array and mechanical scanners.

In this last part of the thesis, we consider additional aspects of holographic storage, in preparation for proposing a bigger and better system. In Chapter 7, we discuss systems issues affecting holographic memory design. In this vein, we survey the methods of performing angle–multiplexing, and introduce and demonstrate a new device for angle steering: a silicon bulk–micromachined, magnetically–actuated micromirror. We also discuss time response and noise and error performance of holographic memories. Finally, in Chapter 8 we propose and discuss several next–generation designs for large–scale high–speed holographic memories. This includes a method for nonvolatile readout that combines several previously proposed methods.

What to see in this thesis

An illustrated guide

	Chapter	
Overview of holographic memory	1	
Dynamic range in holographic memory	2	$\eta = \left[\left(\frac{A_0}{\tau_r}\right)\frac{\tau_e}{M}\right]^2 = \left[\frac{M/\#}{M}\right]^2$
M/# vs. α, experiment & theory	3 & 4	
Storage of 10,000 holograms in one location	6	
16 location, 160,000 hologram system	5 & 6	
Next–generation memory design	7 & 8	

Contents

List of Tables

List of Figures

Chapter 1 Overview of holographic memory

Contents

Holography is a method for recording the phase and amplitude of an arbitrary electromagnetic wavefront. It accomplishes this by interfering the unknown (or signal) wavefront with a known reference wavefront. When the interference pattern is imprinted in a storage medium, subsequent illumination with the same reference wavefront reconstructs the signal wavefront.

If the storage medium is thin and the reference wavefront is a plane wave,[1] then the degree of similarity between the recording reference wave and the readout reference wave is not too important. Changes in wavelength or illumination angle affect the angle at which the wavefront is reconstructed, but the diffraction efficiency[2] is not significantly affected. In contrast, the reconstruction of a hologram in a thick storage material depends strongly on the correspondence between the reference and readout wavefronts. We can use this effect to store (and independently recall) multiple signal wavefronts within the same volume. If we have a way to imprint information on the incoming signal wavefront and read it from the reconstructed wavefront, we have a storage technology with high storage density and parallel access.

In this section, we give a a brief historical background to volume holography. We begin with the disciplines upon which volume holography draws: Bragg diffraction, holography, and Fourier optics. We then discuss the development of our understanding of volume holography, through various material–independent theoretical developments. Then a short survey of the storage materials available, and a discussion of some of the advances made in using these materials to store volume holograms.

Later in the chapter, we discuss the motivations behind holographic storage. We introduce the components of a volume holographic data storage system, and the requirements on each. We review the methods available for multiplexing the reference beam,[3] and begin to focus on the topic of this thesis: angle multiplexing in the 90° geometry. Finally, in an appendix, we discuss a simple mathematical formalism for

[1] In a plane wave, only one spatial frequency is represented, which implies that the beam is infinitely large.

[2] the ratio of output power in the reconstructed signal to input power in the readout beam.

[3] at least, those discovered so far!

holography, introduce k–space, and use it to solve for the Bragg selectivity of thick holograms.

1.1 A brief history of holographic storage

As we mentioned, volume holography draws upon three disciplines:

- *Bragg diffraction.* Sir Lawrence Bragg, in pursuit of the atomic arrangements of crystals, invented the "x–ray microscope" [2]. This was a two–step wave–reconstruction process for magnifying and imaging 2–D projections of crystal structure. Bragg developed and used the relationship between the periodicity of the crystal lattice and the direction in which x–rays are scattered [3]. In volume holography, this Bragg condition allows us to know the direction of the scattered light, given the direction of the incident light and the periodicity of the three–dimensional interference pattern.

- *Holography.* In 1948, Dennis Gabor proposed a method of wavefront reconstruction in which one could write ($\gamma\rho\alpha\phi\omega$) the whole ($\delta\lambda o\varsigma$) wavefront [4]. This was an extension of Bragg's two–step process; Gabor's contribution was the preservation of phase information. He hoped to improve the resolution of images from the electron beam microscope. However, he could only get his technique to work at optical frequencies.

  Two developments in the early 1960's led to rapid advances in holography. The first was the availability of a high–power coherent light source: the laser [5]. The second was the development of off–axis holography by Leith and Upatnieks, allowing for spatial separation of the readout and reconstructed wavefronts [6–8]. Because of these developments, holography is now a part of daily life: from magazine covers to anti–counterfeiting devices on credit cards and driver's licenses, from non–destructive testing to analysis of mechanical vibration.

- *Fourier optics.* We use Fourier optics as a tool for understanding volume holography. It allows us to interpret the Bragg condition as a conversion from one

spatial frequency into a second spatial frequency by means of a third (the grating). Using Fourier optics, we decompose our complicated signal wavefront into a summation of infinite plane waves of various spatial frequencies. We can then consider the effect of our volume hologram in terms of easy–to–analyze plane waves, and trust to linear system theory and Fourier optics to reassemble them into the complex wavefront.

The history of the incorporation of linear systems theory and Fourier analysis into optics has several parts [9,10]. The first is the development and refinement of the theory of diffraction, culminating in the Fresnel diffraction integral. The second part is the recognition that the infinite plane wave is an eigenfunction of the diffraction process, just as the sinusoid is an eigenfunction of the linear time–invariant system. This allows the incorporation of Fourier analysis (used by electrical engineers to describe time–dependent electrical waveforms) to describe spatially–dependent optical wavefronts. Along with this comes the wonderful discovery that a lens takes the 2–D Fourier transform of the optical wavefront, converting the spatial–frequency of incoming plane waves to transverse position of focused spots in the focal plane.

Volume holography is the sum of these three disciplines: the volume integration of the Bragg diffraction of plane waves from a thick hologram, followed by assembly into a complex wavefront described by Fourier optics. The first to use volume holography was Yuri Denisyuk, to improve the fidelity of holographic reconstructions [11]. By using a thick substrate, the unwanted conjugate of the signal wave was attenuated by Bragg mismatch.[4] Denisyuk's experiment was also the first use of the reflection geometry (where the wavevector of the grating is perpendicular to the surface of the recording material).

A volume grating attenuates the conjugate wave because the reconstructed signal if affected by the degree of correspondence between the original recording wavefront and the readout wavefront. This Bragg selectivity can be used to store and access

[4]see Appendix 1.6.1.

multiple holograms in the same volume. In 1963, Van Heerden considered the storage density of volume holograms, and found it to be on the order of V/λ^3 [12]. Using $\lambda \approx 1\mu m$, this is 10^{12} bits/cm^3, or the Library of Congress in a volume of 150cm^3 [13]. Needless to say, this sparked some interest.

Previously existing scattering theories could be used to predict the behavior of weak volume gratings [10, 14]. But volume phase gratings can generate near–unity diffraction efficiencies. Kogelnik developed a coupled wave theory for volume holography which predicted diffraction efficiency and Bragg selectivity for thick absorption and phase gratings [15]. Subsequent theories were developed which cover the regime between thick and thin gratings [16–31]. This early history of volume holography is reviewed by Collier et al. [32], Russell [33], and by Solymar and Cooke [31].

A large amount of early research was dedicated to the development of recording materials. Most planar holography was done in photographic film, which is sensitive and has good spatial frequency response. Unfortunately, film requires developing between recording and readout, is difficult to obtain in thick samples, and is generally a write–once material. Dichromated gelatin was another write–once material available to early holographic researchers [32].

In the late 1960's, the photorefractive effect was discovered in ferroelectric inorganic crystals, by researchers attempting to perform second–harmonic generation in lithium niobate ($LiNbO_3$). They found that the index of refraction (and thus their phase–matching condition) would change as the crystal absorbed light [34]. Because this tended to ruin their experiments, the effect was called "optical damage." The photorefractive effect has since been discovered in a number of inorganic crystals, including some semiconductors. We list several with references in Table 1.1. Much research has also been done in amplifying the photorefractive effect by adding dopants [35–47]. In addition, photorefractive polymers have recently been synthesized by combining electro–optic and photoconductive polymers [48–72].

In a photorefractive material, light induces a change in the index of refraction. This occurs in three steps:

Photorefactive Material	Year "discovered"	References
Ferroelectrics		
$LiNbO_3$	1966	[34, 73, 74]
$LiTaO_3$	1966	[34, 74]
$KNbO_3$	1976	[75]
SBN	1969	[76]
$BaTiO_3$	1970	[77]
Sillenites		
BSO, BGO	1976	[78]
Semiconductors		
GaAs	1984	[79, 80]
InP	1984	[80]
Polymers		
bisA:NPDA:DEH	1991	[49]
PVK:TNF:DMNPAA:ECZ	1993	[61]

Table 1.1: Photorefractive materials

- Absorption of photons, exciting electrons (holes[5]) to the conduction (valence) band;
- Charge transport and trapping, forming an electric field between redistributed charges;
- Modulation of the refractive index through the linear (Pockel's) electro–optic effect.

Many theories were developed in the early 1970's to describe this process, particularly the charge transport via drift, diffusion, and other observed phenomena [81–94]. The theoretical model of Kukhtarev [95] was the first to find general acceptance, and subsequent theories usually use it as a starting point [96–104]. In Appendix 4.5, we discuss the Kukhtarev model in some depth. This model shows that a spatial modulation in the light intensity[6] within the crystal will lead to the same modulation in index of refraction. Because of the charge transport and the spatial derivative between charge and electric field, this index modulation is often spatially shifted from the interference pattern.

[5] in $BaTiO_3$ or photorefractive polymers, for instance.

[6] for instance, an interference pattern formed by two intersecting coherent beams.

Before we continue and examine the photorefractive effect as it applies to holographic data storage, we should point out that there are many additional applications for photorefractives. Many of these make use of this phase shift between intensity and index modulations, which can cause light to couple from one of the recording beams into the other. This effect, called two–beam coupling, can be used to amplify a weak signal–bearing beam [105]. The time response of some photorefractive materials can be very fast, leading to real–time and novelty filters[7] [106], phase–conjugation by four–wave mixing[8] [109], sub–harmonic gratings [110,111], and spatial solitons[9] [112]. All of these (especially the last two, which are only a few years old) are current research topics of some interest.

So far, the photorefractive effect sounds wonderful for holographic data storage: we can get a large crystal by cutting a boule of the inorganic material, there's no development procedure, and we get a phase grating which is a copy of the interference pattern. So what's the problem? Unfortunately, the photorefractive effect doesn't stop after we record the grating. Any subsequent illumination (within the absorption band) leads to electrons in the conduction band. These electrons move around under diffusion and drift and fall into traps, destroying the charge pattern which created the phase grating. In addition, some photorefractive materials have significant dark conductivity, so that electrons can get to the conduction band and erase holograms in the absence of light. So volatility of the stored holograms during readout, and persistence (or shelf life) of holograms in the dark became important research issues.

$LiNbO_3$ received much of the early effort: it is (relatively) easy to grow large samples with good optical quality [113], has a high Curie temperature (1100° C [114]) and low dark conductivity, and it exhibits a strong photovoltaic effect [115]. In addition, iron (Fe) was quickly and unambiguously determined to be the best dopant for

[7] in which a stationary pattern is quickly removed from the signal beam by two–beam coupling, so only what is "new" is seen.

[8] A signal beam and one reference create a real–time grating, which is readout by a second, counterpropagating reference beam. The result is a phase–conjugate (or time–reversed) copy of the signal beam [107–109].

[9] Diffraction and Kerr–type self–focusing are balanced, so the optical beam travels through the material without spatial broadening.

$LiNbO_3$[10] [116, 117]. Drawbacks of $LiNbO_3$ include relatively low sensitivity, small electro–optic coefficients, a small spontaneous polarization and large coercive field at room temperature [114]. An impressive body of research on $LiNbO_3$ came out of the RCA Laboratories, from Drs. Amodei, Staebler, Phillips, Alphonse, and their collaborators. This group demonstrated high–quality image storage and retrieval [118], explored the relationship between absorption and the Fe dopant [117,119], developed control over the Fe^{3+} concentration [119], contributed to the early photorefractive theories [81,82,84], and demonstrated storage of more than 500 holograms [120]. The RCA group also demonstrated and explored thermal fixing in $LiNbO_3$ [120–123], in which the volatile pattern of trapped electrons is compensated by ions which are mobilized by elevated temperature. At room temperature the ions are no longer mobile. When the electronic charge pattern is removed with uniform illumination, only the durable ionic copy of the grating is left.[11] During the same period of time, electrical fixing was demonstrated in SBN and attributed to compensation by spatially modulated reversal of ferroelectric domains [125,126].

One of the things that the RCA group was attempting to create in their crystals was rapid erasure of holograms, so that new information could be quickly rewritten [127]. However, the storage of multiple holograms is best aided by rapid recording and slow erasure (as we show in Chapter 2.2.2). A clever method for rapid erasure was demonstrated by Huignard [128]; the idea is to record over the undesired hologram with the same signal beam and a reference beam which is 180° out–of–phase with the original reference.

The closest that holographic storage got to reality during this period was the 500 hologram experiment of the RCA group [120], and the system demonstrated by D'Auria and Huignard [129,130]. In this latter system, 10 holograms were stored at each of 25 locations, accessed by acousto–optic deflectors and thermally fixed with transparent electrodes on the crystals.

[10] best for "single–photon" gratings.

[11] The lifetime of a fixed grating has been reported to range from months to 5 years [124].

1.1.1 Recent interest

Holographic storage has undergone a resurgence in recent years, despite the lack of groundbreaking improvements in the available storage materials [131, 132]. Most of the changes which have led to the reexamination of holographic storage are in the supporting components: coherent light sources, spatial light modulators, and detector arrays. We discuss these components in more detail in the next section. Another factor has been the availability of computer control, which has made it easier to perform large–scale storage experiments. Recent developments in holographic storage have also been driven by the development of new system techniques for holographic storage. These include

- better understanding of material–independent problems such as crosstalk [133–149];
- the invention of new multiplexing techniques. Fractal [150–153] and peristrophic [154,155] multiplexing are augmentations to angle multiplexing which do not use Bragg mismatch, and thus are independent of the usual selectivity considerations. Both peristrophic and shift multiplexing [156] make it possible to increase storage density in thin recording media over traditional multiplexing techniques such as angle or wavelength multiplexing. In addition, hybrid methods such as sparse wavelength angle multiplexing [157] have been developed;
- better understanding of the interaction of the photorefractive effect and multiple hologram storage. This has led to the development of recording schedules [92, 158,159] and the use of $M/\#$ as a figure–of–merit for photorefractive holographic systems [160–163] (see Chapter 2.2).

Material developments: thick materials With the recent proliferation of exciting holographic storage experiments (notably, the storage of 5000 holograms in $LiNbO_3$:Fe by Fai Mok [153]), interest in materials development has returned. $LiNbO_3$:Fe remains as the thick read/write material of choice for system demonstrations, although development continues on SBN [40–42] and $BaTiO_3$ [45–47], and new dopants

for $LiNbO_3$. Electrical fixing of SBN [164–167] and $BaTiO_3$ [168], and thermal fixing in $LiNbO_3$ [169] and other crystals such as $KNbO_3$ [170], $KTa_{1-x}Nb_xO_3$ [171], and $K_{1-y}Li_yTa_{1-x}O_3$ [172] are active research topics. There is also increased interest in the effects of other crystal parameters such as stoichiometry on the holographic storage process [173–175].

An interesting development in photorefractive crystal research is the return of recording with short pulses. This had been demonstrated in the 1970's, and attributed to two–photon absorption [176–184]. However, recent experiments have showed recording of interference patterns formed between IR beams, gated by an intense green pulse [185–191]. The effect is now attributed to shallow traps which are populated by the green light, used for absorption and transport of the IR interference pattern, and which then decay back to the usual deep trap levels [187]. The attractive part is that since, in the absence of the gating light, the crystal does not absorb the IR read beam, no erasure occurs during readout. Unfortunately, the gated pulse itself tends to erase previously recorded gratings, using up dynamic range and reducing the possible storage capacity.

We should note that any grating can be read with a reference beam of a different wavelength. However, an information–bearing signal beam is encoded in gratings with a spread of spatial frequencies. The larger the difference between the recording and readout wavelengths, the smaller the spread of spatial frequencies that can be readout satisfactorily. We refer to this technique of non–volatile hologram access as two–λ readout [192–202], and discuss it briefly in Chapter 7.4.4. In the gating pulse recording technique, the IR readout beam *is* the original reference beam, so the entire signal beam can be reconstructed without problem.

Other viable material candidates have been proposed for holographic storage. Spectral–hole burning materials make it possible to access a small subset of molecules throughout a volume, since the homogeneous linewidth is much smaller (as much as 10^4—10^5 times smaller) than the inhomogeneous linewidth [203–212]. Storage of 6000 images has been demonstrated using spectral hole burning [213]. Because standard volume holography can be applied to each subset of molecules separately, these ma-

terials are capable of 4–D storage (3 spatial dimensions and absorption coefficient). As if that weren't enough, some of these materials are also capable of additional multiplexing using electric field. The main drawback is that low temperature ($<$ 4°K) is required for the narrow homogeneous linewidth and the persistence of the hole–burning.

A second material development is persistent photoconductivity of the DX center in GaAs:Si or other doped semiconductors [214–217]. In this effect, which also requires low temperature (70–100 °K), absorption of a photon frees 2 electrons and causes the Si atom to undergo a configurational change. An energy barrier prevents retrapping of the electron. As a result, recorded gratings are not erased by subsequent illumination. However, the "upper" level can be saturated, meaning that the dynamic range of the material will still be divided between exposed holograms and diffraction efficiency will still fall as $1/M^2$ (where M is the number of exposures). We discuss the $1/M^2$ limit in saturable materials in Chapter 1.3.1. The index grating in DX center materials is formed through persistent photoconductivity of the excited carriers, instead of via the electro–optic effect. These gratings can be thermally erased by raising the temperature until the energy barrier is overcome, returning the atoms to the low energy state.

A final new material being explored for holographic storage is Ge–doped glass [218, 219]. This material is capable of large, persistent index changes. Even larger index changes can be made by preloading the glass with hydrogen before exposure. Drawbacks include the required wavelength range for recording (UV), and low sensitivity.

Material developments: thin materials In contrast to the moderate changes in thick materials, there have been major changes in thin holographic recording materials. An entire new class of write–once photopolymers have been developed. These polymers undergo a light–activated polymerization reaction which modulates the density of the film in proportion to the incident light intensity [220–232]. These materials have high sensitivity, large index changes, require no fixing, and are inexpensive.

Drawbacks include shrinkage during recording,[12] short shelf life, and the inability to stop the photoreaction once it has begun.[13] Another drawback, although difficult to believe, is the high dynamic range. The photopolymers are so sensitive that interpixel gratings between the spatial frequencies of the signal beam can become troublesome [233]. These thin recording materials have become much more attractive with the advent of peristrophic and shift multiplexing. For instance, peristrophic multiplexing has been used to store 1000 multiplexed holograms in the same volume, and to demonstrate an areal density of 10 bits/μm^2 [234,235]. And shift multiplexing has been theoretically shown to be capable of areal densities up to 12bits/μm^2 [236,237].

We mention two other advances in the area of thin materials. The first is the synthesis of photorefractive polymers [48–72]. These are typically guest–host polymers in which several materials have been combined, each with its own role in the photorefractive effect. This field is very young, especially considering that finding the correct proportion of ingredients is mostly trial and error. Researchers have concentrated on achieving single–hologram diffraction efficiency and net two–wave coupling gain, and have recently reached diffraction efficiencies of almost 100% and net gain of 200 cm^{-1} [72]. However, many research challenges remain, including high dark conductivity (lifetimes range from minutes to days), short shelf life, poor optical quality, and the need for extremely large voltages during both recording and readout (> 40V/μm).

The second advance is a system technique for increasing the effective thickness of thin holographic materials. It consists of stacking multiple thin film recording materials between precisely spaced buffer layers [238–242]. The Bragg selectivity of the stack of thickness L (called a Stratified Volume Holographic Element, or SVHOE) is then the same as a bulk storage material of thickness L. Drawbacks include the difficulty in getting good optical quality buffers, and in maintaining the layer thicknesses [240,243]. A 4–element SVHOE has been demonstrated using photorefractive polymers [244].

[12] this makes it hard to Bragg–match all the spatial frequencies in a hologram simultaneously.

[13] All the holograms to be multiplexed in the same volume must be written one after another. Any delay between exposures reduces the achievable dynamic range.

1.1.2 Read–write memories vs. WORM memories

In order to see how thin and thick volume holographic materials can be used for data storage, we say a few words about the memory storage market. Most memory products today fall into three categories:

- *Pre–recorded data*—These include CD–ROMs, audio compact disks, videodisks, and the soon–to–be–released multimedia compact disks. The success of such a memory product depends on infrastructure: the presence of appropriate readers in a large number of homes/businesses. Without a widespread base of customers, content providers are hesitant to release their data on any particular format. Memory standards are then slow to change because of this complex interaction between consumers, manufacturers, and content providers. However, because the audience base for entertainment products is so mind–boggingly large, the payoff for those formats which become standards is well worth the effort.

- *Write–once read–many times*—These include some videodisks and CD–recordable disks. Nominally, one would think that these applications would not depend on infrastructure to the same extent as pre–recorded data. However, even though people record data on WORM media as a permanent record for their own uses, often they would like to exchange or send this data to other users. For this reason, having a WORM memory product which corresponds to a widespread format is very attractive. In addition, many WORM formats are extrapolations of popular pre–recorded formats, meaning that writer/players can be made which are backwards compatible with the pre–recorded format. A WORM format which does not provide this dual–use format is at a severe disadvantage.

  A holographic WORM memory would most likely use a thin recording material, probably photopolymer. In the design of these memories, both total system size and storage density are important. This is because the consumer will probably compare a holographic memory to the other alternatives on the basis of both system and media densities. Other factors of importance might include

storage capacity, readout and recording rates, system cost, and media cost. Adding dual–use capability to a holographic memory writer/player (for instance, building one which can also read compact disks) would certainly add to system cost.

The explosive growth of the Internet should work to the advantage of holographic WORM disks, though. Since users can increasingly exchange data electronically, the physical exchange of data will become less important. It is even conceivable that new computer software might be distributed through the Internet, reducing the use of pre–recorded formats. However, the digital distribution of entertainment (audio or video) in a form which could be easily copied is most unlikely, as it would destroy the profits of the recording and movie industries.

- *Read/Write Memory*—The main reason that there is no read/write holographic memory product is the unfinished quest for the perfect storage medium. Photorefractives are available and reasonably well–studied, but require complicated system solutions to deal with the limited persistence and the volatility of data during readout. Because of the size overhead,[14] holographic memories have a disadvantage when compared to magnetic hard disks for secondary storage. However, for tertiary storage, a terabit holographic memory might compete well against magnetics in terms of cost, and well against disk arrays in terms of access time and readout speed. Applications might include very large databases linked to multiple users, with little need for update but a desire for fast access and rapid output. Because the storage material will most likely remain inside the system at all times, we are not too interested in the crystal storage density. All the consumer will see is the system size.[15]

[14]You need to include the volume of a laser, an SLM, and a CCD before you have stored even one bit.

[15]Not very impressive to promise to store 1 terabit in 10 cubic centimeters of storage material, and then deliver a system the size of a sofa!

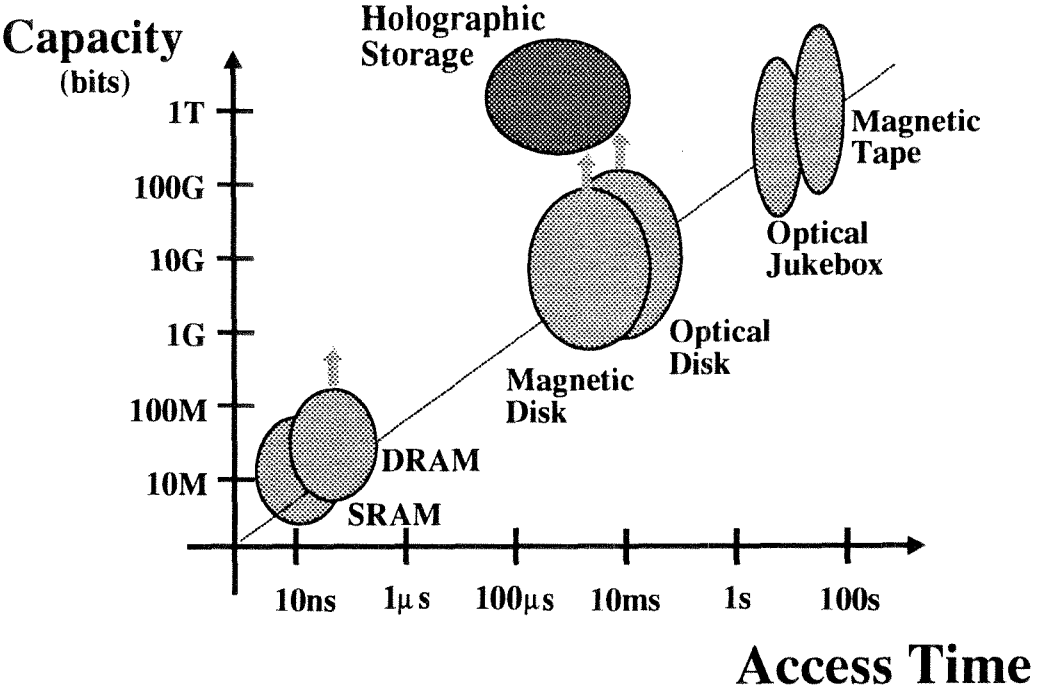


Figure 1.1: Capacity vs. speed: current memory technologies. [244]

1.2 Motivation for holographic memories

Our motivation for investigating volume holographic data storage is the possibility of large capacity and high readout rate. With most memory technologies, increased capacity means decreased readout rate. This is shown in Figure 1.1, where we plot memory capacity vs. memory speed for several storage technologies available today [244]. Also on the plot, we place holographic storage. The indicated regions are general guidelines—one can certainly get systems with different capabilities by spending more, especially in the area of capacity. Since holographic storage is just starting to become commercially available [245], we should take some care not to be over–optimistic when placing it with existing products. We have chosen a capacity of 1 Terabit and a readout rate of 1000 frames/sec. These are achievable goals for a holographic storage system given current technology—the challenge is achieving low cost and system size in the same design. We discuss some of these issues in the following sections and in Chapter 7.2 and 8.

Storage density $\propto V/\lambda^3$ We saw before that volume holography allows us to distribute data throughout the volume and then retrieve it. The original impetus for holographic storage was the possibility of storing one bit per voxel of volume λ^3. This is often called the V/λ^3 result [11]. Here we give an intuitive feel for this limit, and use it to introduce fractal multiplexing in its original context. In Figure 1.2(a), we show a signal beam of dimensions $N \times N$ illuminating a cube (of dimensions $D \times D \times D$). When we try to put this beam through the crystal, the number of resolvable spots is limited by diffraction to $D/\lambda \times D/\lambda$. So N can't be more than D/λ. Likewise, if we are angle–multiplexing the reference beam, we expect our Bragg selectivity to limited to λ/D. (See Section 1.4.1 for more details.) In Figure 1.2(a), we show the focal plane of an angle–multiplexed reference beam. Each dot corresponds to a plane wave of a different spatial frequency. Given a fixed angle range to work with, the number of reference beams we can use is of the order of $N \sim D/\lambda$, giving N^3 connections within the crystal. This makes sense: there is a correspondence between the number of connections and the number of dimensions in the material [150].

But even though we are constrained to N^3 connections, there is no reason why we have to divide them between signal and reference as shown in Figure 1.2(a). If we are interested in making interconnection patterns between two optoelectronic neuron arrays, we might want to have the same number of "nodes" in each array. We can do this by having two arrays of $N^{3/4} \times N^{3/4}$ pixels each [151]. However, we need to be careful where we place these pixels, because of Bragg degeneracy.[16] It turns out that there is a procedure for generating pairs of small grids [150,151] (one grid for the signal beam and one for the reference beam). Larger grids can then be formed from mosaics of small grids, leading to a fractal–like expansion of scale. For this reason, we call these fractal sampling grids [151,152]. In holographic storage, we would still like the signal beam to contain closely packed pixels. We can accomplish this with the sampling grid shown in Figure 1.2(c). The reference beam is composed of discrete

[16]The angle of the readout beam can differ from that of the reference beam in two dimensions. Changes in the plane of the grating (horizontal in this case) cause Bragg mismatch. Changes orthogonal to this plane (vertical in this case) give reference beams which are still Bragg–matched to the grating. We discuss this more in Sections 1.4.3.

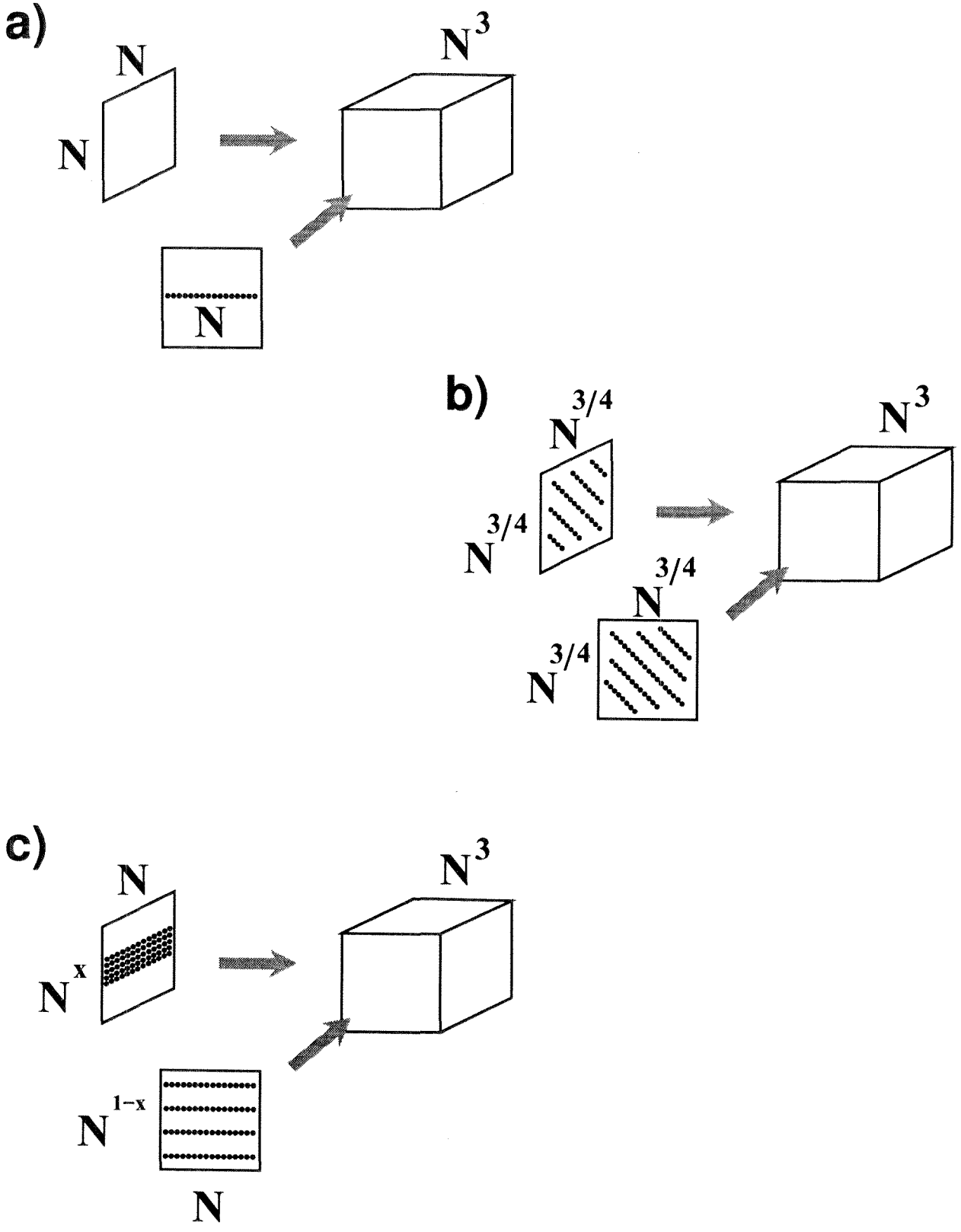


Figure 1.2: Degrees of freedom in holographic storage.

rows of angular locations. The separation between these is made larger than the vertical extent of the signal beam, which guarantees that the number of connections is still N^3. This approach was used by Fai Mok to store 5000 holograms, in 5 fractal rows of 1000 holograms each [152]. We will also use fractal sampling grids (or fractal multiplexing) in Chapters 5 and 6.

The V/λ^3 result would indicate that the 1 terabit memory we discussed above would require less than $1cm^3$ of storage material. So what keeps us from using all of this enormous storage density? We discuss this in the next section and in Chapter 2.2.

1.3 A typical holographic memory system

In Figure 1.3, we show a typical holographic memory. There is an information–bearing signal beam and a unique reference wavefront for each hologram. The signal beam contains an input device for imprinting information on the wavefront. For good reason, this is usually called a Spatial Light Modulator (SLM). A typical SLM is a pixellated device, which creates a spatial ON–OFF pattern on the beam. The modulated signal beam passes through the storage material to the output device: a detector array matched to the SLM pixel spacing. Additional optics, which are not shown here, are used to image the SLM pixel array onto the detector array.

The placement of the storage medium within the signal beam is not fundamentally important. Because holography reconstructs both phase and amplitude, a reconstructed wavefront will continue along the signal path as if it had never been interrupted. In this sense, holography does not perturb the travel of the signal beam from SLM to detector array—it saves the wavefront for replay at some later time. In practice, however, the storage medium is commonly placed at an image plane or a Fourier transform plane. These two planes are where the signal wavefront occupies the smallest transverse area. The choice between image plane and Fourier plane is a complex one involving many factors including density, noise and error performance, total system size, and storage density.

The final element in our holographic storage system is the control over the refer-

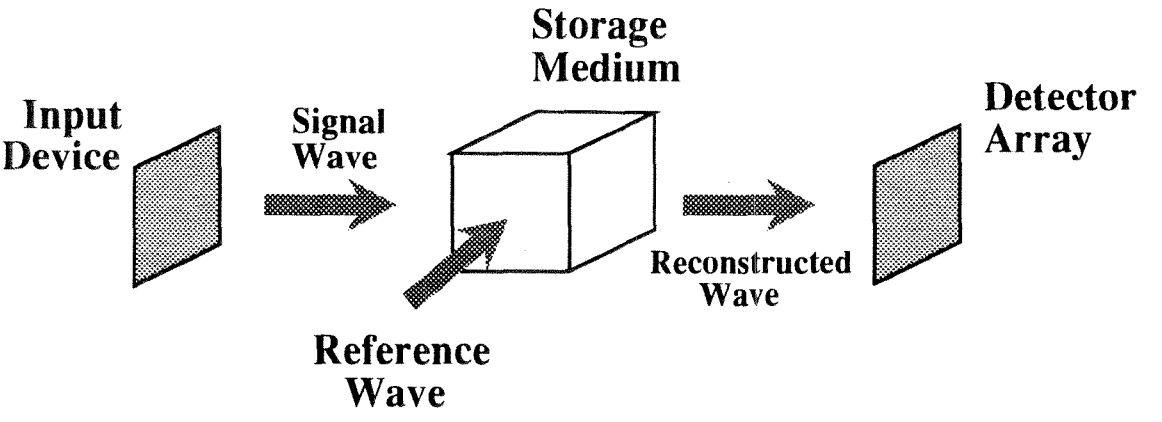


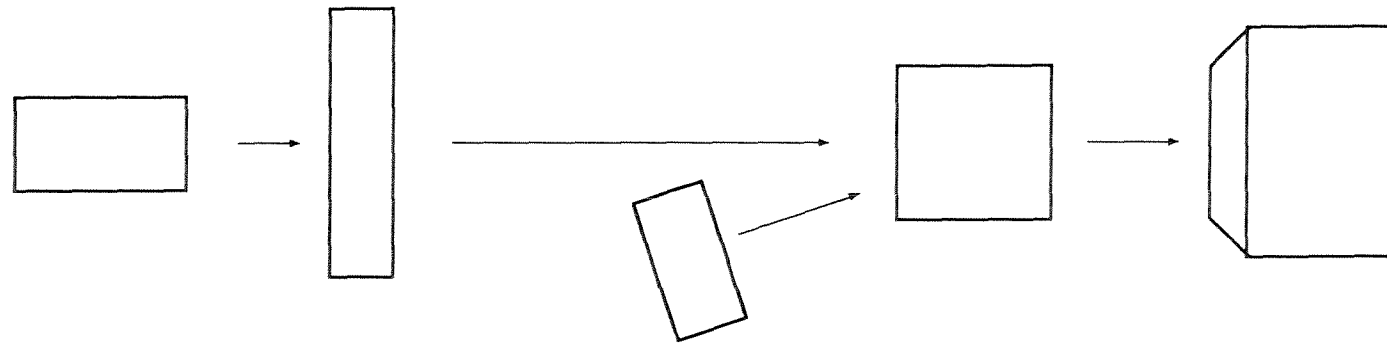
Figure 1.3: A typical holographic memory system.

ence wavefront. This can take many forms, depending on the type of multiplexing used, from tunable lasers to angle deflectors. We survey the currently available multiplexing techniques in Section 1.4. First, however, we discuss the requirements of our holographic storage system, and discuss how they affect the performance specifications of the components.

1.3.1 Component and system requirements

The features that we would like our holographic memory to have include: fast recording, large capacity, fast readout rate, good reliability, and long term storage. In Table 1.2, we list these features vertically and list system components horizontally. At the intersection of a feature row and a component column, we describe how that component affects the desired feature. This is an interesting table, because it can be read in two different ways. If you want to improve a particular feature, reading along the appropriate row lists all the components that are involved. Conversely, reading a column lists the capabilities required of a particular component. Note that this table is intended as a general guide, and one might add (or subtract) entries both in components and in desired features. For instance, in a wavelength–multiplexing system, the wavefront control device might be the laser itself.

In the next few paragraphs, we discuss each component in turn, following the columns of Table 1.2 from left to right.



Desired feature					
Fast Recording	high power	• fast display time (<ms) • low loss	(fast, random access)	sensitivity	
Large Capacity	high power	# of pixels (1000 × 1000)	# of reference wavefronts	• # of storage locations • dynamic range	# of pixels (1000 × 1000)
Fast Readout Rate	high power	# of pixels	fast, random access	sufficient diffracted power	• # of pixels • high speed (>1000 frames/sec)
Reliability/ Fidelity	stable	• uniform response • high contrast ratio	• accuracy • repeatability • feedback to find Bragg match	• low output noise • high optical quality • environment control	• low noise • error correction • detector to SLM pixel matching
Long Term Storage	lifetime			persistence	

Table 1.2: Desired system features, and dependence on component performance.

- *Laser*—In certain circumstances, planar holograms can be recorded with incoherent light [31]. Volume holography, however, requires coherent light. The reference and signal wavefronts for recording usually originate from the same laser. The available laser power is split between the reference beam and the signal beam (which usually has large losses because of the spatial light modulator). The pathlength difference between the two when they intersect is made smaller than the coherence length of the laser. As this condition is violated, the modulation depth of the interference pattern decreases, and the stored holograms have poor diffraction efficiency.

  During recording, any instabilities in laser output can cause the interference fringes to wash out. After recording, the stability requirement is more lax: only the power level of the laser output needs to be stable.[17] However, we would always like high power from the laser. During recording, more power decreases the average recording time; during readout, it boosts the diffracted signal power. The upper limit on the laser power[18] is the point at which the heat due to absorption in the crystal starts to affect the recording and readout dynamics [246].

  Because of these differences in stability and power requirements, we can consider using different lasers for recording and readout, with similar wavelengths (within 1nm or so). We could have one high–power, single–mode, stable laser for recording, and multiple cheap diode lasers for readout. We discuss using multiple same–wavelength lasers in Chapter 7.9.1, as well as the two–λ architecture where the readout wavelength can differ significantly from the recording wavelength.

- *Spatial Light Modulator*—The number of pixels in the SLM affects both the total

[17] Otherwise, the signal power reaching the detector is fluctuating and we are unsure where to place our threshold between dark and bright pixels. See the entry in this section on the detector array.

[18] besides the obvious cost, electrical power, and size considerations.

capacity C and the readout rate $R_{readout}$. We can write these two variables as

$$C = N_p N_h N_l, \tag{1.1}$$

and

$$R_{readout} = N_p R_{frame}, \tag{1.2}$$

where N_p is the total number of pixels in the SLM, N_h is the number of holograms per location, N_l the number of storage locations, and R_{frame} is the number of holograms that can be readout per second. Here we naively assume that N_p, N_h, and N_l are independent variables. This is a plausible assumption for a system which uses thick crystals (like those in this thesis), and capacity is always a more important consideration than density. For systems using thin materials, the line between N_p, N_h, and N_l becomes blurred, and one works to maximize areal storage density. For thick crystals, a large number of bits in each hologram makes holographic storage attractive compared to other technologies, since both capacity and readout rate increase. In Figure 1.1, larger values of N_p move holographic storage away from the line which all other memory technologies share (that is, where capacity and readout rate are inversely related).

So what's the limit on the number of pixels in the SLM? Until recently, it was the fabrication of the SLM. However, SLMs with 640 × 480 pixels are now commercially available [247,248], and SLMs with 1000 × 1000 pixels are being developed [249]. Now the limitation is moving to the imaging system, which must direct the high space–bandwidth product[19] signal through the storage crystal and deliver a focused and properly registered image of the SLM onto the detector array. Registration implies that each and every one of the million pixels in the SLM falls directly upon its counterpart in the detector array.

[19] often abbreviated SBP. See Appendix 1.6.4 for the definition.

We also require good contrast ratio from our SLM, so that we can easily distinguish OFF pixels from ON pixels. If the system SNR is large enough, we could use gray scale on each pixel to encode additional bits. In this case, the dynamic range of the analog modulation of the SLM pixels must also be large. We can then redefine the capacity C as

$$C = b_p N_p N_h N_l, \tag{1.3}$$

where b_p is the number of bits per pixel. Note that the number of gray levels required increases as 2^{b_p}, so more than 3 to 4 bits per pixel is probably unreasonable.

Most of the available large–scale SLMs use liquid crystal technology, and modulate the light passing through a pixel by polarization control. The polarization is either passed unperturbed or rotated orthogonally, depending on the presence or absence of a control voltage. By placing the SLM between appropriately aligned polarizers, this polarization modulation becomes an amplitude modulation. Since the liquid crystal is birefringent, liquid crystal SLMs can also be used to perform phase modulation. This will be of interest to use when we talk about multiplexing the reference beam with phase codes (Section 1.4.5).

We would like to have good efficiency from our SLM, which will depend on the transmissivity of the ON state and the fill factor. Fill factor is effectively the ratio of the width (or height) of the pixel to the inter–pixel spacing. (There is also an areal fill factor: the area of the pixel over the total area dedicated to the pixel). The fill factor is usually less than unity, which creates an interesting effect: the spatial frequency spectrum, after spatial modulation, contains multiple orders. Each of these orders contains the information pattern from the SLM. However, storing more than one of these orders would be inefficient. In addition, it is more difficult to focus the image of the SLM if more than one order is used, because the higher spatial frequency carriers experience different

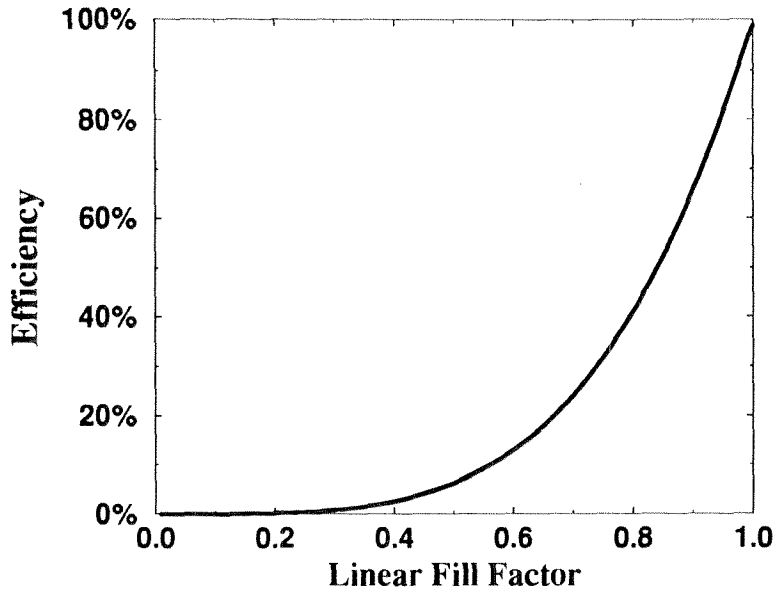


Figure 1.4: Fraction of power in the central diffracted order of the modulated signal beam, as a function of linear fill factor.

aberrations in the optical system.[20] It is common to use the central order, since this contains the most optical power. Light *is* being diffracted into the higher orders though, and the fraction left in the zero order depends on the linear fill factor as shown in Figure 1.4. We show the derivation of this relationship in Appendix 1.6.6.

Two other SLM factors can influence the holographic storage system: the display rate and the uniformity. With photorefractive materials, the rate-determining step in the recording process tends to be material sensitivity, not SLM display rate. If faster recording materials are used, the time required to display a page of information on the SLM might then become a factor. For readout purposes, it is convenient if the ON state of each pixel has the same intensity across the entire display. However, nonuniformity is not a problem, only a headache. We can compensate by using the gray scale response of the SLM (thus losing some efficiency) or by being more clever with our detector array design[21] (thus adding complications and expense).

[20]This depends on the performance of the imaging system.

[21]see Detector Array section on page 27.

In Appendix 1.6.5, we discuss some of additional factors affecting the choice of SLM, including the use of nematic and ferroelectric liquid–crystals, and transmissive and reflective SLM designs.

- *Wavefront Control*—The particular form of wavefront control depends on the multiplexing method. We survey the available methods in Section 1.4. In general, however, the wavefront controller has to be repeatable, so that the identical wavefront used for recording can be replicated. Since minor changes in the Bragg–matching condition are probably unavoidable (from temperature fluctuations or changes in the average index of refraction), we need continuous control over the wavefront (not just a discrete set of wavefronts). The controller should be able to change from one reference wavefront to another rapidly, so that the readout rate is high. If possible, the access time to any wavefront should be random (the time before a new wavefront is present should not depend on the old wavefront). However, non-random access is fine if the worst–case access time is small enough. We discuss the factors that can influence recording and readout rates in Chapter 7.1.

- *Storage Medium*—As we mentioned at the start of the chapter, the storage medium need only modulate its index of refraction in proportion to the intensity of the interference pattern between the signal wavefront and the reference wavefront. Real materials have finite dynamic range, though, which is allocated among the stored holograms. As a result, the diffraction efficiency of M holograms is proportional to $1/M^2$, whether a photorefractive or saturable material is used.

  In Chapter 2.2.1, we derive the $1/M^2$ relation for photorefractives. We show this relation in terms of the recording schedule, and introduce the $M/\#$ as the constant of proportionality between η and M.[22] In the following paragraphs, we give an intuitive explanation of the $1/M^2$ relation for a saturable material.

[22] $\eta \sim (\frac{M/\#}{M})^2$.

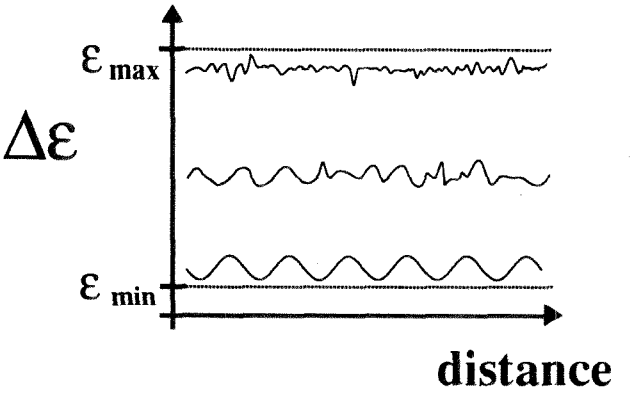


Figure 1.5: Recording of multiple holograms in a saturable material.

$1/M^2$ relation in saturable media First, we note that the hologram exists as a perturbation in the dielectric permittivity (the square root of index of refraction). In this form, it serves as the source term for scattering an incident electromagnetic field (we refer to this as light amplitude). The scattered field can be solved in terms of a volume integral of the incident field and permittivity modulation, as

$$E_{scattered} = \iiint dV\ \Delta\epsilon\ E_{incident}. \tag{1.4}$$

There are some assumptions in this formulation (Born approximation) [9, 13, 250], which are described in Appendix 4.4. For our purposes, we can use this formulation if M is large. We note that a $1/M^2$ division of dynamic range amongst diffraction efficiencies (or scattered intensities) is a $1/M$ division amongst scattered light amplitudes.

Since the material is saturable, we can only change the permittivity in one direction.[23] And this permittivity change is proportional to the total intensity of the interference pattern. For each hologram, we are imprinting both the spatially varying part $(S^*R + SR^*)$ that we care about, as well as the non-spatially varying intensity $(S^2 + R^2)$. In Figure 1.5, we show the permittivity perturbation at three points during storage: early, middle and late. Note that

[23] If, for each hologram, we could choose whether to add to or subtract from the permittivity, we could get $\eta \propto 1/M$. Unfortunately, I don't know of any such materials.

the average level of ϵ increases by the same amount for each hologram,[24] so that the total permittivity change is proportional to M. If we assume that the spatial variations are random from hologram to hologram, the spatially varying part of the total permittivity modulation grows as $\sqrt{M}$. This is true once enough holograms are present (we need enough steps in our random walk). We must stop exposing before the material is saturated, to avoid smoothing out the spatial variations (our holograms). So the best we can do is divide the total permittivity modulation in M parts. As a result, the diffraction efficiency of each hologram falls as $1/M^2$.

There are additional features that we would like our storage material to have, so that the imaging of the complex object beam is not compromised by passing through the material. We would like it to be of high optical quality, and to record holograms linearly, so that the reconstructed holograms are indistinguishable from a transmitted image. Stored holograms should not decay in diffraction efficiency or degrade in fidelity, either while the material sits in the dark (persistence or shelf life) or while data is being accessed. In a photorefractive material, this means we are going to have to add some system solutions to provide this. Finally, since holography is so sensitive to the phase difference between the beams during recording, and to thermal expansion of the hologram after recording, we are going to need have good control over the environment surrounding the optical beams and the storage medium. This may mean a sealed system, isolation from mechanical vibration, and/or temperature control.

- *Detector Array*—The detector array has the task of intercepting the reconstructed wavefronts and extracting the digital information. A charge–coupled device (CCD) is often used. As we mentioned above, the CCD array must be in the image plane of the SLM, with each detector pixel matched to the image of the corresponding SLM pixel. Since diffraction in the optical system will tend to broaden the image of each pixel, even a small fill factor on the SLM will lead

[24] Even if $\Delta\epsilon$ is not a linear function of exposure energy, we can always find a schedule of exposures such that each $\Delta\epsilon$ is identical [31, 251].

to large pixel images at the detector plane. Fortunately, CCDs with unity fill factors are available. These are usually full-frame imagers, so that readout is slow. Video rate imagers require more wiring, so their fill factor is less than unity. The effective fill factor can be returned to unity by using a microlens array to focus light onto the CCD pixels.

The detector array needs to be very fast (on the order of 1000 frames/sec) and very sensitive. Each reconstruction will likely have less than 1μW of power (which is then distributed among up to 500,000 ON pixels).[25] The detector array redigitizes the data, turning the analog measured value (of detected photons or integrated charge) into a binary decision (ON–OFF).[26] The detector array is therefore a logical point in the system to compensate for spatial nonuniformity in the images [252] or to incorporate digital error correction [253–264].

Error correction Error correction algorithms lay their groundwork during recording, adding redundancy bits to each block of data of predetermined size. Upon readout, the algorithm uses the redundant bits to find any bits which are in error, and then correct them. Error codes can be analyzed by the improvement in error rate they provide, the maximum number of errors which can be detected or corrected, the code rate (the fraction of bits which are real data bits), the complexity/size of the VLSI circuitry required, and the power used by this VLSI circuitry [257].

Most error correction routines have generally been developed for 1–D strings of data, and the incorporation of such a code in volume holography was first demonstrated by Heanue et al. [265]. Error correction is capable of reducing error rates enormously—in compact disk players, for instance, the raw bit error rate of 10^{-6} is reduced to an overall bit error rate of 10^{-12} with a code rate of approximately 0.66 [266]. Researchers studying error coding for holographic memories have considered using multiple 1–D codes as well as 2–D codes [263].

[25] More on this in Section 7.1.3.

[26] Or a decision among 2^{b_P} gray levels.

However, the theoretical framework for 2–D codes is not as mature a subject as for 1–D codes. The real challenge for error correction in holographic memories comes in reducing the required circuit density and fabricating it near the detection electronics. One way might be to flip–chip bond detector chip(s) onto the error–correction chip(s). The signals from a small 2–D block of detector pixels could be time–multiplexed on a smaller number of wires to the lower layer, which would perform the error correction on this data and pass on the corrected data. The error correction electronics could even consist of several layers of substrates, each connected to the next lower layer. Needless to say, it is not going to be simple to integrate error correction into the detector array.

So, in conclusion, we require a very sensitive, low noise, megapixel CCD camera with 1kHz frame rate, multiple output channels, and integrated error correction. With the exception of the error correction, each of these features is available in a CCD camera today. The trick is putting them all in the same camera, adding error correction, and doing it all at low cost.

Data format Before moving on to multiplexing methods, a few words about data format. The large Gbit/sec readout rate promised by holographic storage is not going to be manifest to a user who uses only a few bits from each stored hologram before asking to see another. As a result, care needs to be taken to format the data so that a large part of each recalled page is useful to the user(s). This block–oriented formatting is a natural feature of several memory applications, including video on demand (where a finite chunk of compressed video can be readout to a cache memory, decompressed, and sent to the viewer at video rate) or large–scale databases (where the user is either searching the entire database, scanning a index sequentially, or pulling an entry containing text, graphics, audio, and/or video).

1.4 Multiplexing methods

In this section, we discuss some of the available methods for multiplexing the reference beam. The general idea is to create a set of input wavefronts for which the detector array receives light corresponding to only one page of data. These methods can be divided into two general classes:

- *Bragg–matched readout*—All other stored holograms are suppressed by Bragg–mismatch. These methods include angle, wavelength, phase code, and shift multiplexing.

- *Displaced readout*—More than one hologram is reconstructed by the same reference wavefront. However, the scattered waves do not overlap in spatial frequency, so all but one reconstruction can be blocked by an iris. The desired data then continues to the detector array. This characteristic is central to fractal and peristrophic multiplexing.

In addition, we describe spatial multiplexing in this section. This is simply how we refer to storing holograms in another location within the material. This is not strictly multiplexing, since there is no superimposing of holograms. However, it certainly is a valid way of increasing the number of holograms in the storage system.

1.4.1 Angle multiplexing

In angle multiplexing [267] (Figure 1.6), we use a discrete set of reference plane waves which vary in illumination angle. The wavevectors for all of these beams lie in a single plane which includes the central signal wave vector. The diffraction efficiency of a hologram is a strong function of angular change of the reference beam in this plane. Since this angle "selects" the hologram, we refer to this as the selectivity function. When the readout beam is identical to the wavefront used for recording, we are at a central peak in the selectivity function and are "Bragg–matched." Since the volume hologram is the integral of a finite collection of planar holograms, the selectivity function contains nulls where the diffraction efficiency can fall to zero. If

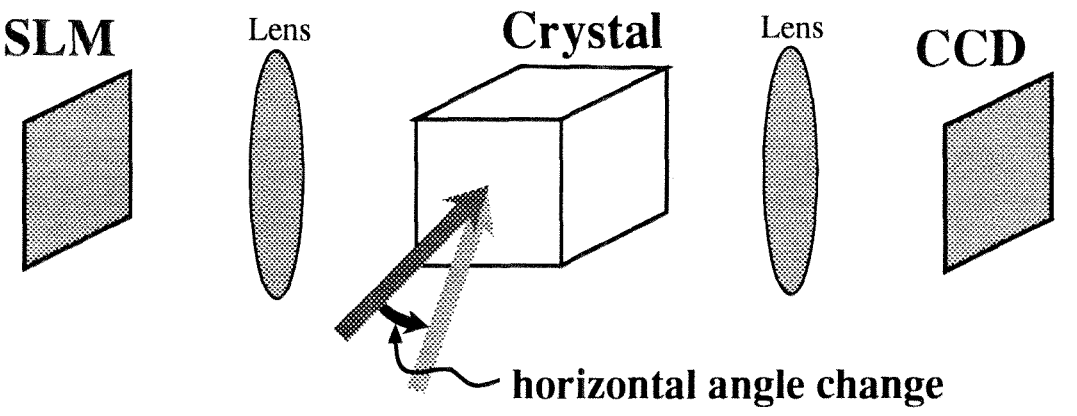


Figure 1.6: Angle multiplexed holographic storage.

each volume element contributes equally, the nulls of the selectivity function go to zero. The angular spacing between the central peak and the first null is referred to as the angular or Bragg selectivity, and can be written as

$$\Delta\theta = \frac{\lambda}{L}\frac{\cos\theta_S}{\sin(\theta_R + \theta_S)}, \tag{1.5}$$

where θ_S and θ_R are the incidence angles of the two beams (see Figure 1.17) and L is the thickness of the volume [268]. This expression holds in the transmission geometry, where signal and reference enter the same face of the crystal, and is derived in Appendix 1.6.3. Note that once the signal wavefront contains multiple spatial frequencies (different θ_S), we have the problem that each different spatial frequency requires a slightly different reference beam for perfect Bragg mismatch. When we use one reference beam for multiple signal beams, we expect to get crosstalk between holograms.

We would like the angular selectivity to be small, so we can store many holograms within some given angular range. The 90° geometry, where reference and signal enter through orthogonal faces, has the smallest angular selectivity [137]. In this thesis, we will be primarily concerned with angle–multiplexing in the 90° geometry. Angle multiplexing can also be performed in reflection geometry, where the reference and signal enter the crystal through opposite faces, although the angular spacing must be

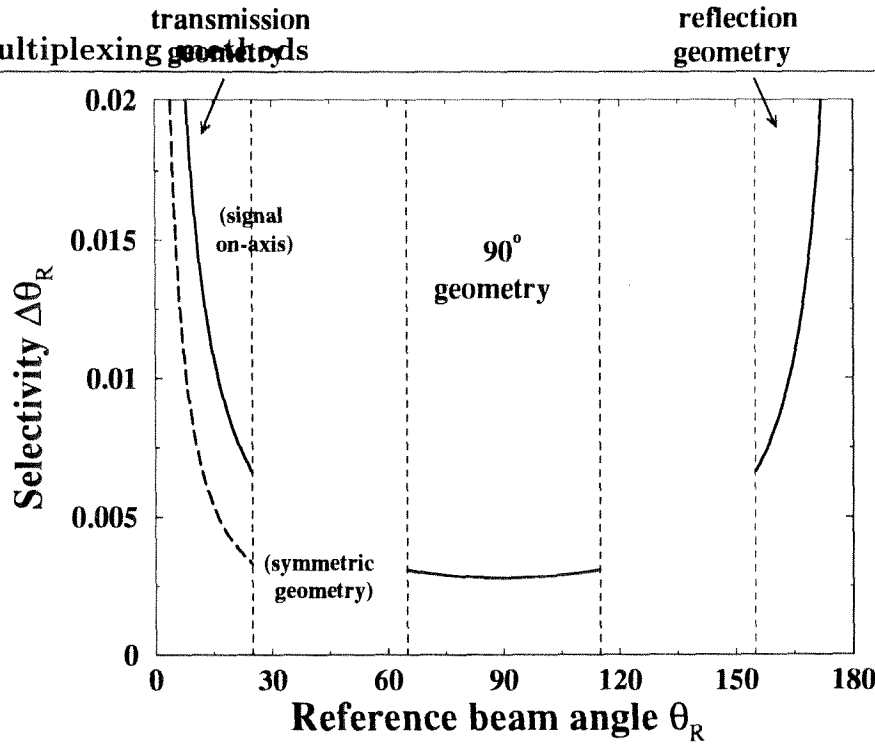


Figure 1.7: Angular selectivity in $LiNbO_3$ as a function of θ_R, for $\theta_S = 0$, $L = 1cm$, and $\lambda = 488$nm. Also shown is the angular selectivity in the transmission geometry when $\theta_S = \theta_R$ (symmetric geometry).

rather large. We plot the angular selectivity as a function of external incidence angle for the three recording geometries in Figure 1.7. The blank spaces correspond to interior angles which cannot be introduced into $LiNbO_3$. Since the ordinary index of refraction is 2.35, the largest possible interior angle is $\sim 25°$. If the grating spacings which fall into these forbidden regions are desired, the crystal can be cut so the faces are not orthogonal.

1.4.2 Spatial multiplexing

As we mentioned above, spatial multiplexing (shown in Figure 1.8) merely recognizes that we are not constrained to superimpose all of the holograms in the same location. Since there is no Bragg–mismatch in planar holography (and no one had discovered peristrophic multiplexing), spatial multiplexing was the only way considered in the late 1960s and early 1970s to store multiple holograms in planar or 2–D media. Many designs and analyses of packing density were considered [269–274]. Two general methods have recently been proposed for accomplishing simultaneous spa-

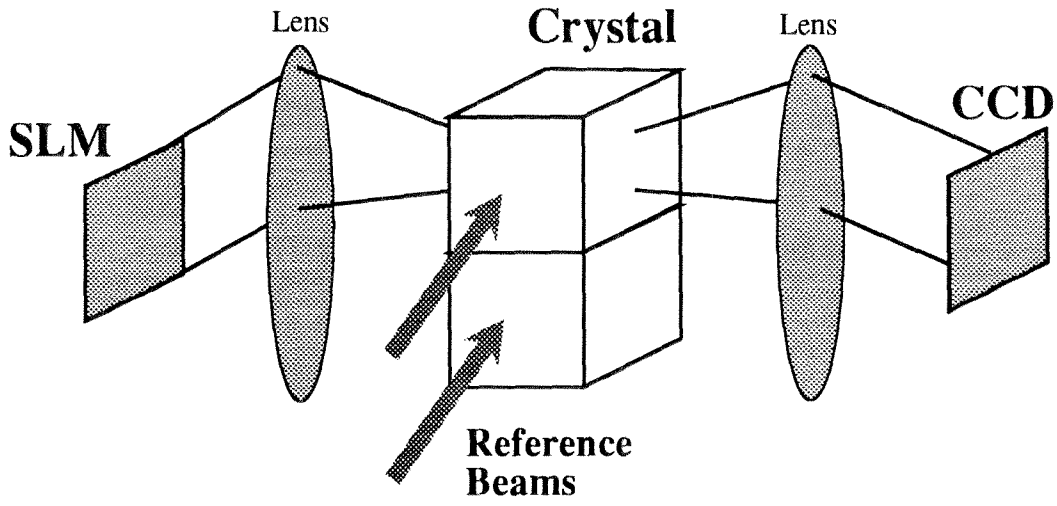


Figure 1.8: Spatially multiplexed holographic storage.

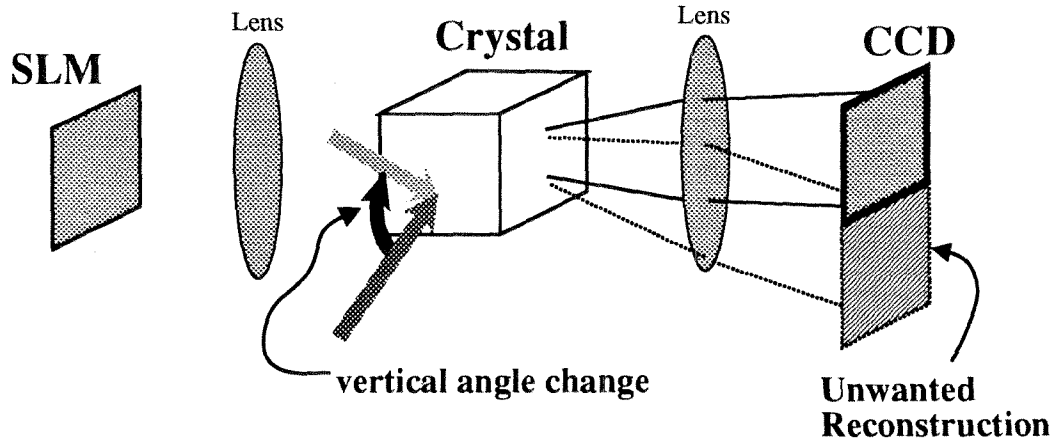


Figure 1.9: Fractal (or out–of–plane) multiplexed holographic storage.

tial and angular multiplexing in 3–D media. The first involves mechanically moving the storage medium to access different spatial locations [130, 275–277]. The second method uses acoustooptic deflectors (AODs) to perform both angle and spatial multiplexing, [128, 129, 278–282] which leads to fast random access at the expense of a more complex optical system. We discuss non–mechanical spatial multiplexing in Chapter 5.1 and Chapter 7.5.

1.4.3 Fractal multiplexing

We have mentioned that the angle–multiplexed reference beam angles used for separate holograms must be distinct, and that this distinction is usually created by angle changes in the plane of interaction of the reference and object beams due to Bragg selectivity. Angle changes orthogonal to this plane of interaction have little effect on the phase buildup of the reconstructed hologram. However, we can still use angle changes in this direction by using the fact that the reconstructed signal wave is entirely contained within a reasonably small solid angle.

In Figure 1.9, we show two input reference beams (represented as vectors). The lower reference beams lies in in the same horizontal plane of interaction as the signal beam. The upper reference beam forms a vertical plane with the lower, orthogonal to the first plane. If one hologram is stored with each of these reference beams, then something strange will happen. Both holograms will be reconstructed with either of the two reference beams. This occurs because the angular selectivity for vertical angle changes is very large. Sometimes the vertical direction is referred to as the degenerate direction, and the double readout effect as Bragg degeneracy [151]. We should also point out that we are using horizontal and vertical only for convenience—we get good Bragg selectivity for reference beam deflections in whichever plane is formed by the reference and signal beams. There is nothing (except that it's convenient in the lab) that says that this plane must be horizontal.

So either reference beam will reconstruct both holograms. However, the undesired or extra reconstruction is not directed along the signal axis. It is tilted vertically by the same angle that separates the two reference beams. If this tilt is larger than the vertical angular bandwidth of the image, the reconstruction of the undesired hologram is completely displaced off the output detector. At this point, the desired hologram can be unambiguously detected on the detector array. By storing holograms with these two reference beams, we have multiplexed two holograms without using Bragg mismatch. As we saw in Section 1.2, this out-of-plane multiplexing was originally conceived in terms of finding non–degenerate 2–D to 2–D interconnection schemes,

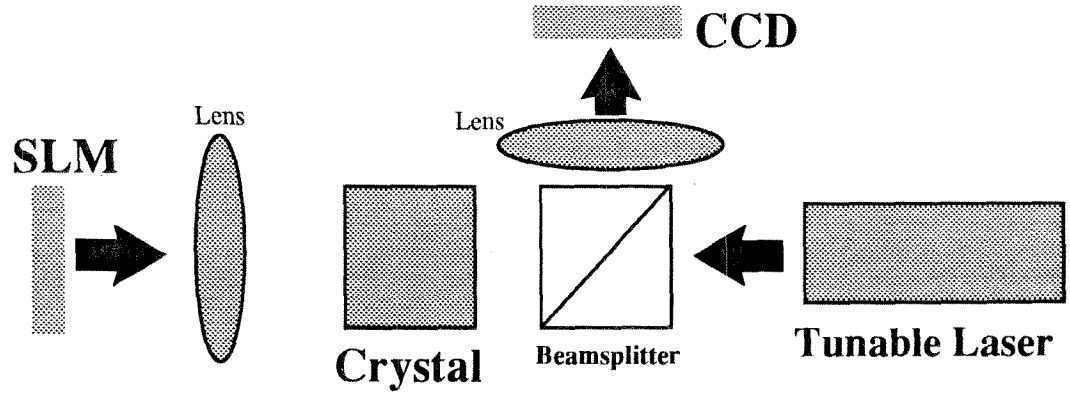


Figure 1.10: Wavelength multiplexed holographic storage.

and the fractal name arose from the scale invariance which these patterns share [150, 151]. Fractal multiplexing has been used together with angle multiplexing to store 5000 [152] and 10,000 [283, 284] holograms at a single location (within the same common 3–D volume of storage material).

1.4.4 Wavelength multiplexing

Wavelength multiplexing [285–294] is very similar to angle multiplexing in that a set of discrete wavefronts is created for which only the desired hologram is reconstructed. However, in wavelength multiplexing, the angle of incidence of the reference beam is unchanged while the wavelength is changed. In parallel with angle–multiplexing, there is a selectivity function with nulls and a wavelength selectivity. This wavelength selectivity is optimal for the reflection geometry, which is shown in Figure 1.10. The problem of multiple reflections in the beamsplitter can be removed if the two beams are displaced slightly. In this case, the reconstruction still "reflects" back through the same face as the incoming reference beam, but is not exactly counterpropagating. (The larger the signal bandwidth, the larger the required angle between reference and signal). We show the wavelength selectivity in the three geometries (transmission, 90°, and reflection) in Figure 1.11. There are currently commercial products which use the high wavelength selectivity in the reflection geometry to make very narrow bandwidth filters [288]. Other wavelength multiplexing applications include correlators [285] and

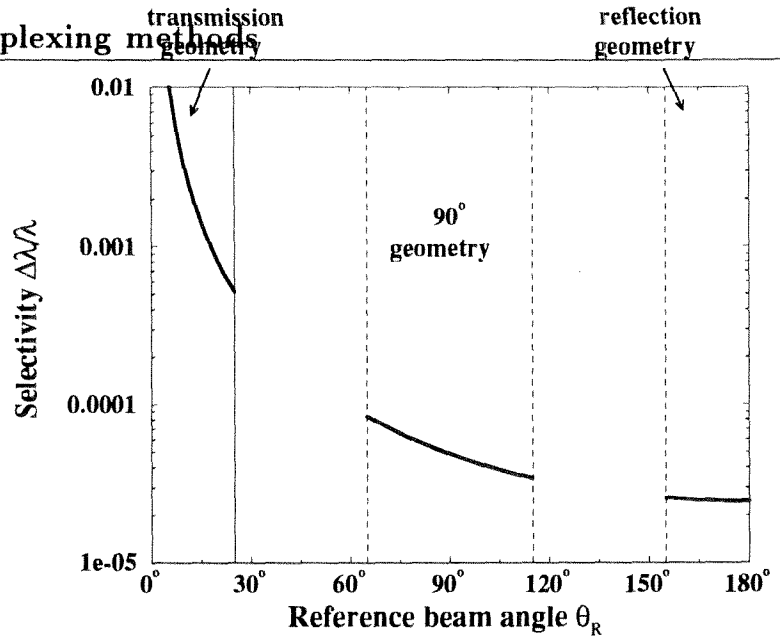


Figure 1.11: Wavelength selectivity in $LiNbO_3$ as a function of θ_R, for $\theta_S = 0$, $L = 1cm$, and $\lambda = 488$nm. Also shown is the wavelength selectivity in the reflection geometry when $\theta_S = \theta_R - 180°$.

dynamic lenses [292].

In a wavelength-multiplexed holographic storage system, there is no need to deflect the reference beam. This can make for a very compact system. Researchers have also promoted wavelength multiplexing because of its crosstalk properties [139, 142, 286]. However, crosstalk in both angle multiplexing and wavelength multiplexing depend on the angular bandwidths used [280, 295], which means that both are subject to increased crosstalk as you try to maximize system performance.

The main drawback to wavelength multiplexing is that the requirements for wavefront control must now be met by the laser source. The laser must be tunable over a broad wavelength range, and capable of rapid tuning to any given wavelength without prolonged instability. There are IR semiconductor lasers which can be tuned over several nanometers in a few hundred nanoseconds [296]. However, storage of 1000 holograms in near–counterpropagating reflection geometry would require 25nm of bandwidth.[27] However, there is also the possibility of using fractal multiplexing techniques in combination with wavelength multiplexing to increase capacity [297].

[27]Assuming storage at the second null, center wavelength near 488nm, and a 1cm thick crystal.

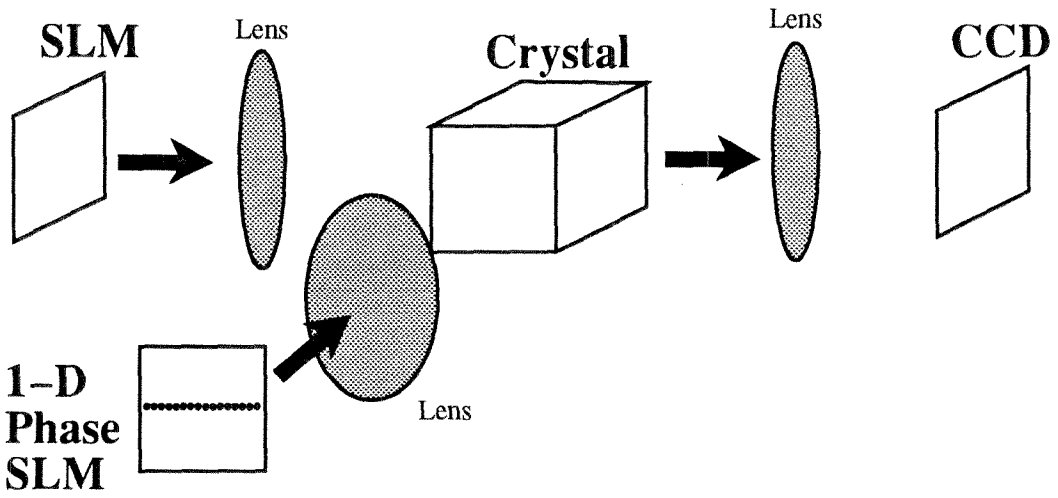


Figure 1.12: Phase code multiplexed holographic storage.

There are additional problems to consider for wavelength multiplexing, including dispersion, variations of recording sensitivity with wavelength [298], and chromatic aberrations in the imaging systems. However, the main reason that wavelength multiplexing is not widely explored for holographic storage is the lack of appropriate tunable sources. If this were to change, wavelength multiplexing would probably be considered as a serious candidate for holographic storage.

1.4.5 Phase code multiplexing

Phase code multiplexing (Figure 1.12) is a generalization of angle multiplexing [299–309]. It resembles angle multiplexing in that a discrete set of reference beams is used—the difference is that now all of the reference beams are used simultaneously. Holograms are multiplexed by imposing orthogonal phase codes on the reference beams. If there are M reference beams, then a set of M orthogonal codes can be used. Usually a few codes are dropped [140], but the trend is that the number of usable codes is close to M. Researchers have studied the use of random codes and the use of deterministic code–generating algorithms such as Walsh–Hadamard codes [310]. In theory, phase code multiplexing has excellent crosstalk performance [140, 295]. However, a serious practical limitation is the inability to precisely modulate the reference beam phases to exactly 0 or π. As these errors increase, the crosstalk performance degrades severely.

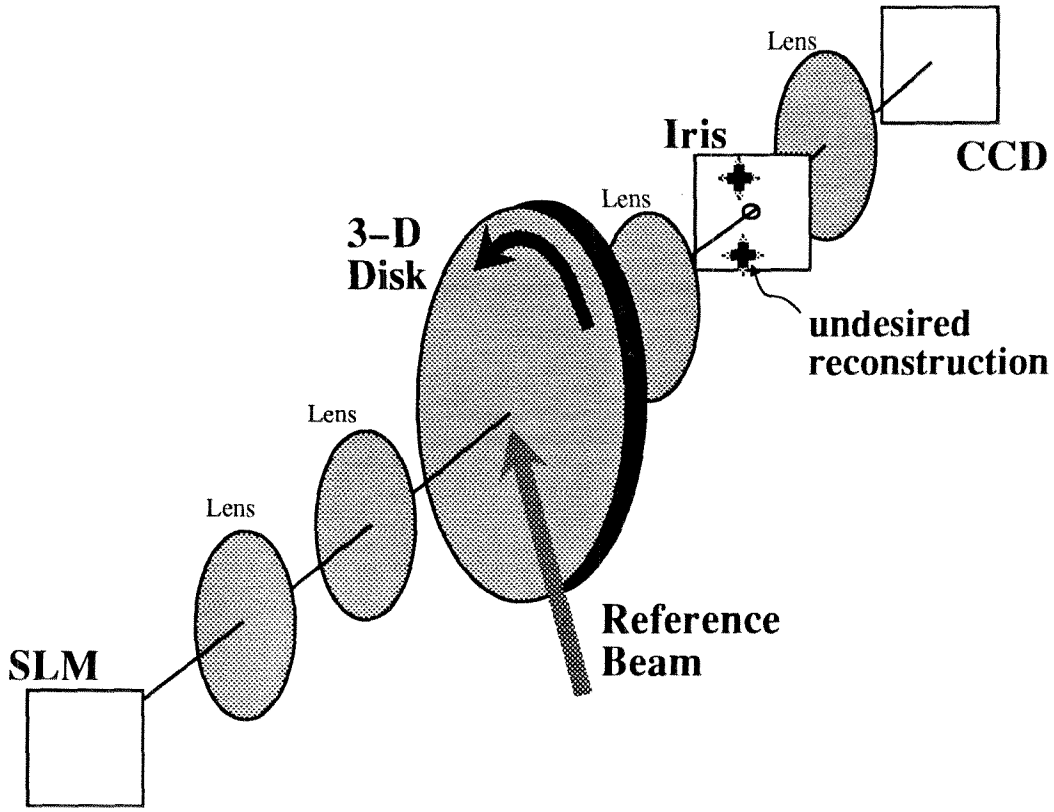


Figure 1.13: Peristrophic multiplexed holographic storage.

The limitation of phase–code multiplexing can be described as the need for a fast 1–D phase SLM with accurate phase control. Since this is our wavefront control, video–rate SLM performance is no longer acceptable. Also, low efficiency of the SLM affects recording rate, readout rate, and total capacity. Despite these problems, researchers have demonstrated the storage of 64 holograms in $BaTiO_3$ [307].

1.4.6 Peristrophic multiplexing

Peristrophic multiplexing [153, 154] (shown in Figure 1.13) is a variant on fractal multiplexing which is extremely attractive for holographic storage in thin materials. The idea is again to displace unwanted holograms, but in peristrophic multiplexing the displacement is generated by rotating the material. As shown in Figure 1.13 for image plane storage, the reconstructed wavefronts from unwanted holograms can be blocked with an iris, allowing unambiguous detection of the desired page of information. The

important point for storage in thin materials is that the size of the iris depends on the SLM pixel spacing, not on the thickness of the recording medium. (If the medium is in the Fourier transform plane, then the size of the iris depends on the size of the SLM, but is still independent of thickness).

Peristrophic multiplexing can be used with angle multiplexing to increase the capacity of thin materials. Since the invention of peristrophic multiplexing in 1994, it has been used to store 1000 image plane holograms [233, 234], to demonstrate areal density of 10bits/μm^2 [233,234], and for storage of correlation templates for a real–time robot guidance system [131, 311]. All of this has been done with 100μm photopolymer which, without peristrophic multiplexing, would capable of storing only 15–20 angle–multiplexed holograms [295].

The main drawback to peristrophic multiplexing is that adding spatial multiplexing over a large 3–D disk requires that the rotational motion be implemented by the reference arm, not the material [130,312]. The record/read head becomes quite complicated. The lens system which implements the angle multiplexing of the reference beam must be able to revolve completely around the normal to the disk in order to implement the peristrophic multiplexing. The next multiplexing method has the advantage that the record/read head is extremely simple.

1.4.7 Shift multiplexing

By using a spherical reference beam, shift multiplexing (Figure 1.14) causes the motion of the recording medium to Bragg–mismatch gratings [155]. Consider a location within the storage crystal. If we are looking at a small enough region and the sphericity is moderate, the reference wavefront resembles a small portion of a plane wave. The wave vector of this plane wave is the normal to the spherical wavefront (in other words, we're thinking about the tangent to the spherical wave at that point). What happens if, after recording, the spherical wave (or the material) is shifted horizontally? Our region is now at a different point on the spherical wavefront, with a normal which points in a different direction. As the material was shifted, the normal

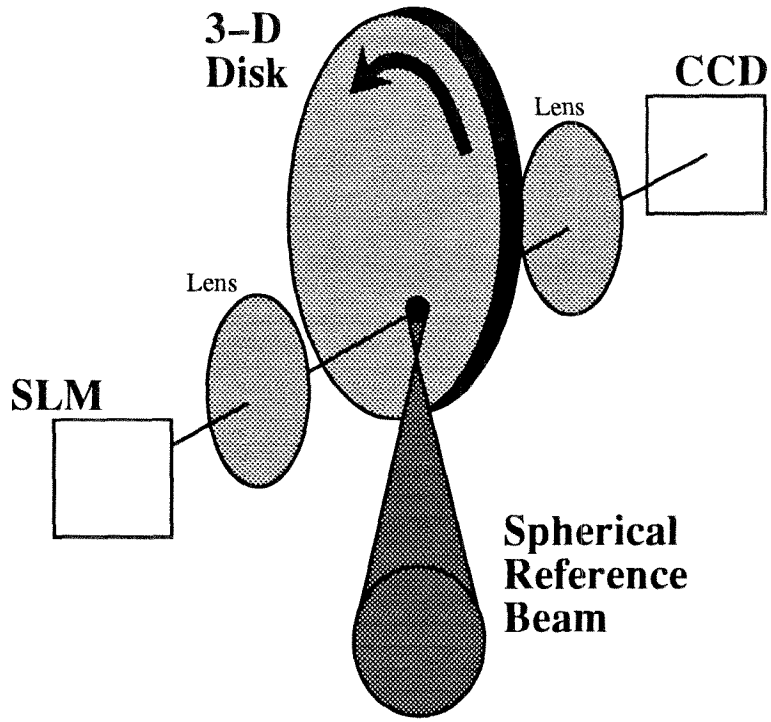


Figure 1.14: Shift multiplexed holographic storage.

to the wavefront at our particular location "rotates". As far as this small region is concerned, it's seeing a plane wave in a different direction! Since this is occurring throughout the volume of the grating (to varying extents, since the spatial derivative of the tangent is not constant), the grating becomes Bragg-mismatched as the shift increases. The diffraction efficiency of the hologram then follows a shift selectivity curve.

Of course, the derivation of the exact shift selectivity cannot be done with a framework of "local plane waves," since the phrase itself is an oxymoron. Our method of considering the normal to the spherical wavefront is suitable only for a handwaving explanation. The exact derivation gives a shift selectivity which has two terms [235, 236]: one which depends on the crystal thickness and illumination angles, and one which depends on the numerical aperture of the spherical wave (this second term serves as a measure of how close the reference beam is to a perfect spherical wave originating from a single point).

Shift multiplexing shares a lot in common with angle multiplexing. In both cases,

we're using a change in angle between reference wavefront and readout wavefront. In one case, the same angle change exists throughout the volume; in the other, the change in angle is induced by translation and varies throughout the volume. From this comparison, one would conclude that, like angle multiplexing, shift multiplexing is unable to get decent results with thin holographic materials. However, there is a fundamental difference. With angle multiplexing, the small number of holograms which can be stored comes from the poor angle selectivity and the limited reference angle bandwidth. With shift multiplexing, the focus of the spherical reference beam can be placed close to the material.[28] Inside the material, the normal to the wavefront will then swing through a huge range of angles across the illuminated spot (which is not a very large distance). Because this large change in angle bandwidth can be larger than even the poor angular selectivity of thin materials, a very small shift can cause Bragg mismatch. As a result, large areal storage densities are possible even in thin recording materials [235, 236]. In addition, shift multiplexing can be performed in the vertical, non–degenerate direction by using fractal techniques.

The second major attraction of shift multiplexing is that the reference arm can be as uncomplicated as the read head in a compact disk player: a single focusing objective. The access to different holograms is accomplished through the motion of the storage medium. The drawbacks to shift multiplexing include the requirement for precise and rapid mechanical positioning, and hologram non–uniformity due to variations in modulation depth and erasure from incompletely overlapping exposures.

1.4.8 Hybrid methods

As one might expect, many holographic storage system designs call for more than one multiplexing method. It is common to use spatial multiplexing at multiple locations, and then two additional methods for superposition in the same location. Combinations include angle and peristrophic multiplexing, or angle and fractal multiplexing (which is what we use in this thesis). Beyond these combinations, several researchers

[28] We still need the reference beam to overlap the signal beam, though.

have developed hybrid methods which combine features from two different methods. We briefly describe two of these hybrid methods.

Spatioangular multiplexing Developed by Tao and Selviah [275,277], this hybrid multiplexing method combines spatial and angular multiplexing. The idea is that the different spatial locations used are partially overlapping. As the reference beam angle is stepped through some angle range, the location used for storage is also slowly stepped across the crystal. By the time the angular range is swept, the crystal has been translated enough that the currently illuminated region has no overlap with the crystal region that was used to store the first hologram. So, one can go back and reuse the first reference beam angle and sweep through the angular range again. The advantages of this method are that it uses the entire crystal volume. Standard spatial multiplexing calls for completely separated storage locations, so that there are always portions of the crystal volume that are never illuminated. The drawbacks of spatioangular multiplexing are that the allocation of reference beam angles is complicated, especially when the spatial movement is over a 2–D grid of locations. Determination of the exposure schedule for uniform holograms is non–trivial, and there is always the question of the holograms at the edges of the spatial array. These holograms exist over a volume of the crystal which does not experience uniform exposure, and the portion of this volume towards the edge of the spatial array will be exposed less than the interior part of the hologram. In an image plane hologram, this leads to spatial nonuniformity across the image. In a Fourier transform hologram, this nonuniformity occurs across the bandwidth of the image, and causes pixel broadening at the detector array. We discuss related issues in Chapter 4.3.1.

Sparse Wavelength and Angle multiplexing A second hybrid multiplexing method is the combination of wavelength and angle multiplexing, termed SWAM (Sparse Wavelength Angle Multiplexing) [156]. The idea is to reduce the angular range required for reference beams by reusing the same reference beam angles at each of a small set of wavelengths. The wavelength spacing is made large enough that

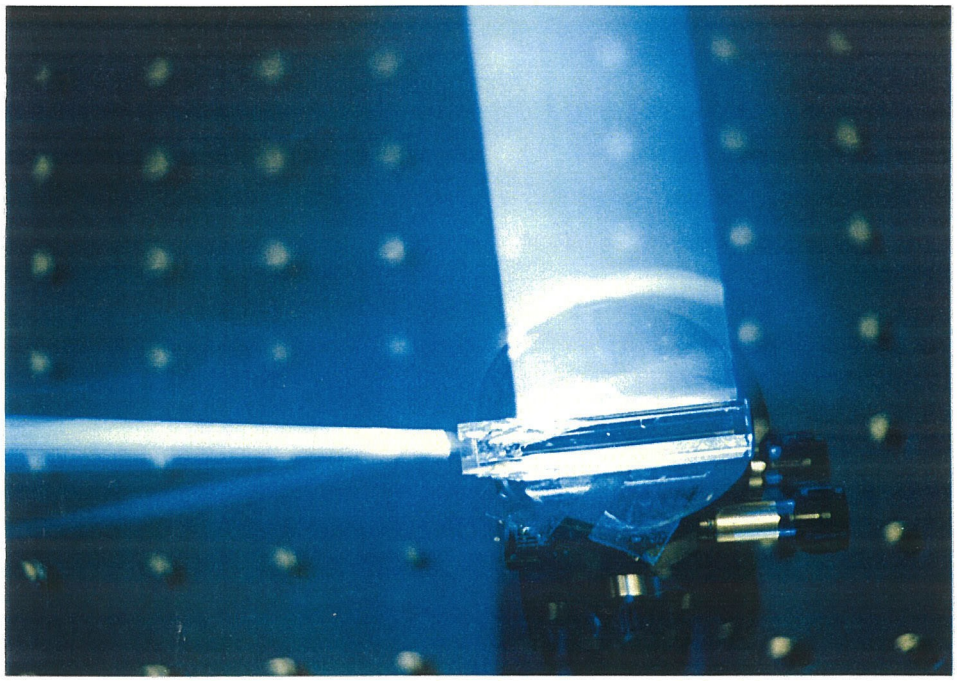


Figure 1.15: The 90° geometry

there is no degeneracy between any of the stored holograms. The main advantage of this method is the reduced horizontal spread of reference beams. This makes the reference arm easier to fabricate and can increase the readout rate if the reference beam angle deflector has non–random access. The drawbacks include the usual problems of wavelength multiplexing, including dispersion, variance in recording sensitivity, and chromatic imaging effects. In addition, while a tunable laser source is not required, multiple lasers of various wavelengths must be incorporated in the system. If the reference beam multiplexing is performed with a laser diode array, however, having a spread of wavelengths within this laser diode array is a natural extension. In addition, two–λ readout techniques can be used so that recording occurs with one wavelength over a large range of reference beam angles, yet readout uses multiple wavelengths and a much smaller range of reference beam angles.

1.5 The 90° geometry

In this thesis, we concentrate on read–write holographic memories using thick recording materials. We have chosen to work with angle–multiplexing because of the variety of available options for reference–beam angle scanning. We use the 90° geometry (Figure 1.15) because it provides high angular selectivity and long interaction lengths.

Since the perfect thick material is still yet undiscovered,[29] we work with the most promising material we have, Fe–doped $LiNbO_3$. In Chapter 2, we study dynamic range in holographic storage to determine what we need out of a photorefractive crystal. In the course of this study, we develop the $M/\#$ as a metric for measuring the dynamic range performance of our system. We then spend Chapters 3 and 4 optimizing the $M/\#$ in $LiNbO_3$:Fe by varying parameters such as oxidation state, crystal size, and doping.

Having squeezed as much performance as possible from our storage material, we turn to the design of a large–scale holographic memory. Our goal is to use angle, fractal, and spatial multiplexing to achieve large capacity—without sacrificing fast access to the stored holograms. In Chapter 5, we discuss how our segmented mirror array makes such a design possible. Then in Chapter 6, we experimentally demonstrate the various features of this memory design. These demonstrations include storage using the mirror array, storage of 1000 holograms using an acousto–optic deflector, storage of 10,000 holograms in the same $\sim 1\text{cm}^3$ volume of $LiNbO_3$, and the demonstration of the full 160,000 hologram system using mechanical scanners. In Chapter 7, we discuss the time response and the noise and error performance of holographic memories. We also survey the methods of performing angle–multiplexing. Finally, in Chapter 8, we propose and discuss several next–generation designs for large–scale high–speed holographic memories.

[29] or maybe it's been discovered and remains unrecognized ...

1.6 Appendices

1.6.1 Mathematical formalism of holography

At this point, we introduce some simple mathematical formalism, and recast what we said in the first few sections. We write the signal and reference as

$$\underbrace{\mathbf{S}}_{\text{amplitude}} \underbrace{e^{j\mathbf{k}_s\cdot\mathbf{r}}}_{\text{spatial frequency}} \quad \text{and} \quad \mathbf{R}\ e^{j\mathbf{k}_r\cdot\mathbf{r}},$$

where $\mathbf{S}$ and $\mathbf{R}$ are amplitudes[30] and $\mathbf{k}_s$ is the spatial frequency (or wave vector or direction, if you like). We set $\mathbf{R}$=1 for simplicity. The hologram is the interference pattern between these two wavefronts, or

$$\left(\mathbf{S}e^{j\mathbf{k}_s\cdot\mathbf{r}} + e^{j\mathbf{k}_r\cdot\mathbf{r}}\right)\left(\mathbf{S}e^{j\mathbf{k}_s\cdot\mathbf{r}} + e^{j\mathbf{k}_r\cdot\mathbf{r}}\right)^*,$$

which we illuminate with the original reference beam to get

$$\left(|\mathbf{S}|^2 + 1\right)e^{j\mathbf{k}_r\cdot\mathbf{r}} + \underbrace{\mathbf{S}^*\, e^{j(2\mathbf{k}_r-\mathbf{k}_s)\cdot\mathbf{r}}}_{3} + \underbrace{\mathbf{S}\, e^{j\mathbf{k}_s\cdot\mathbf{r}}}_{4}.$$

There are four scattered waves, of which we want only the fourth one containing the reconstructed signal $\mathbf{S}$. If the signal and reference are collinear ($\mathbf{k}_r = \mathbf{k}_s$), then all four waves propagate together. With $\mathbf{k}_r \neq \mathbf{k}_s$, then the signal wave goes off in a separate direction than the other three terms.

However, the signal beam is usually a band of spatial frequencies, for instance a complex image of ON and OFF pixels. Now we might find that our reconstructed signal for some $\mathbf{k}_s$ is collinear to the unwanted conjugate (term #3 from above) of some other $\mathbf{k}_s$. If there were a way to attenuate these conjugates, we would get better-looking holograms. This was the motivation for the invention of volume holography by Yuri Denisyuk. By using a thick substrate, he obtained Bragg–mismatch of the conjugate signal wave. Although every element of a volume grating diffracts into the

[30]The intensity in the signal beam, for instance, is proportional to $\mathbf{S}^2$.

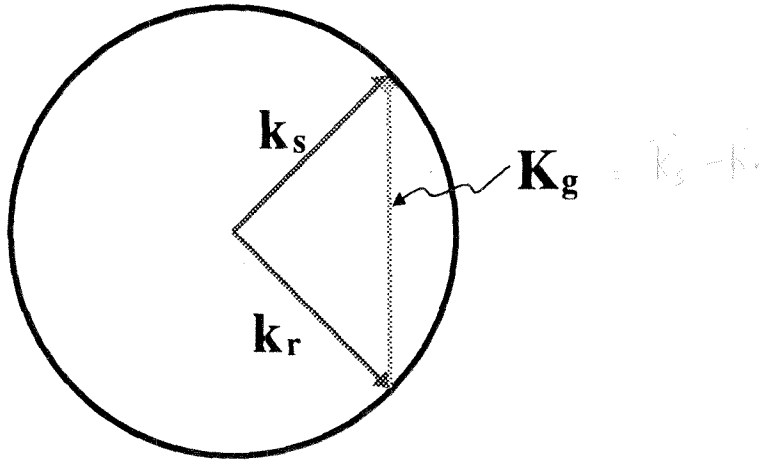


Figure 1.16: K–space

four terms listed above, only the contributions to the signal (term #4) are in phase. Since the conjugate wave is attenuated by phase mismatch, the output images now contain only one copy of the original image.

1.6.2 k–space

This mathematical formulation of wave vectors can be visualized easily in a graphical representation we refer to as "k–space." In this representation, each plane wave $e^{j\mathbf{k}\cdot\mathbf{r}}$ is represented by its wavevector $\mathbf{k}$. Since we are considering monochromatic light, the set of possible plane waves span a sphere of radius $2\pi/\lambda$. For simplicity, we often show a circular 2–D slice through the sphere, as in Figure 1.16. The grating formed between any two plane waves can be represented as the vector which joins the two wavevectors.

Any beam which exists in the laboratory consists of a spread of spatial frequencies. In k–space, this is represented as a cone of vectors. This is of particular importance in theoretical analysis of crosstalk. One last thing to mention before moving on to Bragg selectivity: Since one k–space picture represents every point in space, k–space cannot cannot describe spatial variations. As a result, k–space is not particularly convenient for dealing with spherical beams. Even though spherical waves can be described as an infinite sum of plane waves, k–space contains no facility for representing the

phase of a spatial frequency. But most of the features of spherical waves come out of these phase relationships. As a result, k–space is inappropriate when analyzing shift multiplexing.

1.6.3 Bragg selectivity

We can use k–space to analyze angle or wavelength multiplexing, and in particular, to obtain the Bragg selectivity functions. The trick is to solve for the phase mismatch by relating it to a distance in k–space. To start, we consider the volume integration of the grating via the Born approximation [13] (Chapter 4.4). The transfer function between input spatial frequency $\mathbf{k}_i$ and output spatial frequency $\mathbf{k}_d$ is [268]

$$A(\mathbf{k}_i, \mathbf{k}_d) \propto \iiint d\mathbf{r}' \, \Delta\epsilon_x(\mathbf{r}') \, e^{j\mathbf{k}_i \cdot \mathbf{r}'} \, e^{-j\mathbf{k}_d \cdot \mathbf{r}'}. \tag{1.6}$$

When we expand the spatial variation of the stored grating,

$$\Delta\epsilon_x(\mathbf{r}') = \Delta\tilde{\epsilon}_x(\mathbf{r}') \, e^{j\mathbf{k}_s \cdot \mathbf{r}'} \, e^{-j\mathbf{k}_r \cdot \mathbf{r}'}, \tag{1.7}$$

we get the volume integral of four phase factors, from the reference and signal beams used during recording ($\mathbf{k}_r$ and $\mathbf{k}_s$), the readout beam $\mathbf{k}_i$ and the diffracted beam $\mathbf{k}_d$. Our constraint is that the diffracted beam has to be a solution of the wave equation, or

$$k_{d_x}^2 + k_{d_y}^2 + k_{d_z}^2 = k^2. \tag{1.8}$$

For simplicity, we consider a hologram which is stored in a material which is unbounded in the transverse (x and y) directions, and which is of thickness L in the z direction. This is shown, along with the wavevectors used for recording, in Figure 1.17.

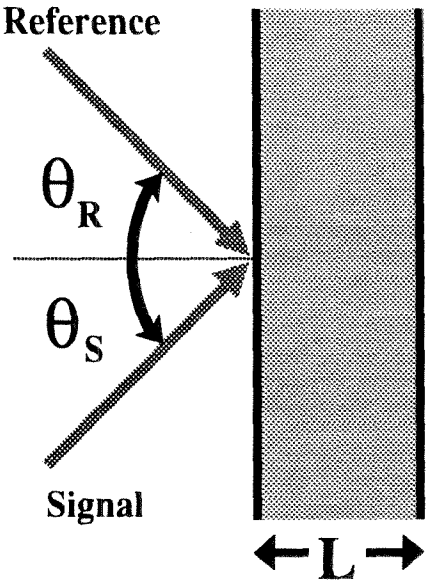


Figure 1.17: Definition of reference and signal angles.

Returning to our integral, we can write $A(\mathbf{k}_i, \mathbf{k}_d)$ as

$$A(\mathbf{k}_i, \mathbf{k}_d) \propto \iiint d\mathbf{r}'\, e^{j\mathbf{k}_s\cdot\mathbf{r}'}\, e^{-j\mathbf{k}_r\cdot\mathbf{r}'}\, e^{j\mathbf{k}_i\cdot\mathbf{r}'}\, e^{-j\mathbf{k}_d\cdot\mathbf{r}'}, \tag{1.9}$$

where we've let $\Delta\tilde{\epsilon}_x(\mathbf{r}')$ be unity for simplicity. Since the material is infinite in the x and y directions, we get delta functions for those two spatial integrals. These two delta functions determine k_{d_x} and k_{d_y}, which in turn constrains k_{d_z}. Then the z integral can be integrated as

$$\int dz' e^{j[(k_{s_z}-k_{r_z})+k_{i_z}-k_{d_z}]z'} = \text{sinc}\left[\frac{1}{2\pi}\left([k_{s_z}+k_{i_z}-k_{r_z}]-k_{d_z}\right)z'\right], \tag{1.10}$$

where each of the $k_{?_z}$ terms is a known quantity. At Bragg–match, $\mathbf{k_r} = \mathbf{k_i}$ and $\mathbf{k_d} = \mathbf{k_s}$ so that we get sinc(0) = 1. As we had expected, we get maximum diffraction efficiency at Bragg–match and can find an incident angle which gives zero diffraction efficiency (when $k_{s_z} - k_{r_z} + k_{i_z} - k_{d_z}$ is a multiple of 2π).

We'd like to be able to solve for the angular change needed for the first null (the

Bragg selectivity). We can do this by using k–space to determine the phase mismatch:

$$k_{s_z} - k_{r_z} + k_{i_z} - k_{d_z},$$

since each of these terms is contained in the z component of one of the vectors in Figure 1.18. $\mathbf{k}_i$ is the reference beam used for readout, and $\mathbf{k}_s - \mathbf{k}_r$ is the stored grating $\mathbf{K}_g$. The vector sum of $\mathbf{K}_g$ and $\mathbf{k}_i$ determines k_{d_x} and k_{d_y} because it sets both

$$k_{s_x} - k_{r_x} + k_{i_x} \tag{1.11}$$

and

$$k_{s_x} - k_{r_x} + k_{i_x}. \tag{1.12}$$

The actual diffracted wavefront is then the solution to the wave equation which shares this same k_{d_x} and k_{d_y}. And the surface of the k–sphere is the set of solutions to the wave equation (with wavelength λ). The line running through the tip of $\mathbf{K}_g + \mathbf{k}_i$ and parallel to the z axis is the set of vectors which satisfy Equations 1.11 and 1.12. Therefore the Bragg mismatch we are looking for is the z difference between the tip of $\mathbf{k}_i + \mathbf{K}_g$ (what the grating would like to diffract) and the k–sphere (what physics requires the grating to diffract). Once we have this Δk_{d_z}, we can just solve for the change in reference beam angle which makes

$$\Delta k_{d_z} L = 2\pi.$$

To do this, we zoom in and look at the region of the k–sphere near the tip of the grating vector (Figure 1.19). We can geometrically solve for the z mismatch between grating tip and k–sphere by breaking it into two parts: the phase mismatch because of change in reference beam angle $\Delta\theta_R$ (which moves the tip of the grating vector), and that due to the change in reconstruction angle $\Delta\theta_S$. Our phase mismatch is then

$$\Delta k_{d_z} = |\Delta k_{s_z}| + |\Delta k_{r_z}|. \tag{1.13}$$

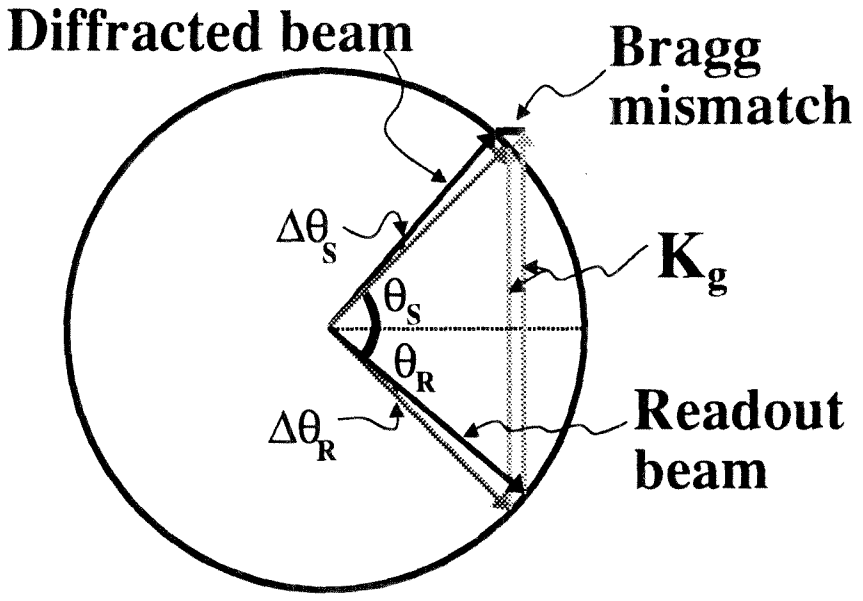


Figure 1.18: K–space for Bragg mismatch in angle multiplexing

To solve for this, we write each wave vector in terms of their incidence angles as

$$\mathbf{k}_r = k\left(-\sin\theta_R\,\hat{x} + \cos\theta_R\,\hat{z}\right) \tag{1.14}$$

$$\mathbf{k}_s = k\left(\sin\theta_S\,\hat{x} + \cos\theta_S\,\hat{z}\right), \tag{1.15}$$

and their derivatives with respect to these angles as

$$\frac{d\mathbf{k}_r}{d\theta_R} = k\left(-\cos\theta_R\,\hat{x} - \sin\theta_R\,\hat{z}\right) \tag{1.16}$$

$$\frac{d\mathbf{k}_s}{d\theta_S} = k\left(\cos\theta_S\,\hat{x} - \sin\theta_S\,\hat{z}\right). \tag{1.17}$$

At this point, we solve for the unknown angle change $\Delta\theta_S$ in terms of the known angle change $\Delta\theta_R$. Using our knowledge that

$$|\,k_{s_x}\,| = |\,k_{r_x}\,|, \tag{1.18}$$

we solve for

$$\Delta\theta_S = \frac{\cos\theta_R}{\cos\theta_S}\,\Delta\theta_R. \tag{1.19}$$

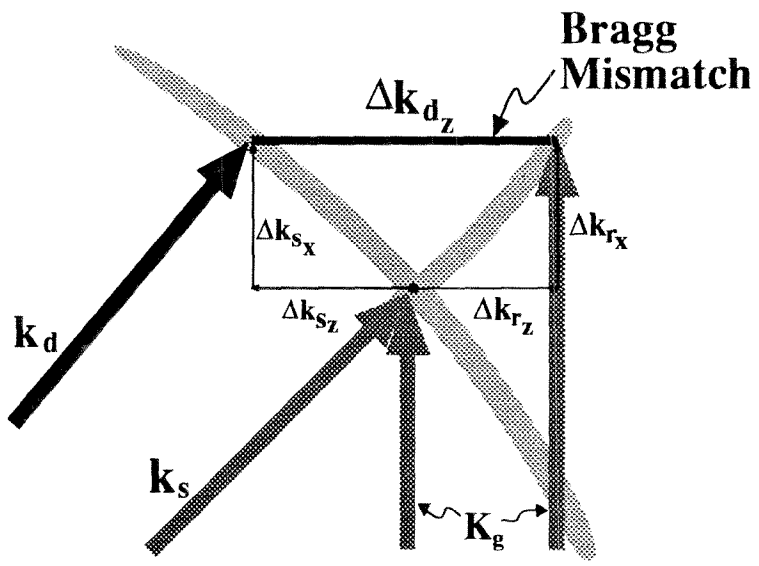


Figure 1.19: Blowup of Figure 1.18

Inserting this into the equations for the z mismatch, we can write

$$\Delta k_{d_z} = k\left(\sin\theta_R + \frac{\sin\theta_S\ \cos\theta_R}{\cos\theta_S}\right)\Delta\theta_R. \tag{1.20}$$

We can then solve for the angular selectivity (the change needed to get to the first Bragg null) as

$$\Delta\theta_R = \frac{\lambda}{L}\frac{\cos\theta_S}{\sin(\theta_R+\theta_S)}. \tag{1.21}$$

As expected, the Bragg selectivity is inversely proportional to the crystal thickness L, or more accurately, the interaction length $L/\cos\theta_S$. Equation 1.21 is the expression we used to plot Figure 1.7.

We can also perform the same procedure to calculate the wavelength selectivity, using the k–space shown in Figure 1.20 and the blow-up in Figure 1.21. We only trace the procedure here. The radius of the k–space sphere is what changes now, not

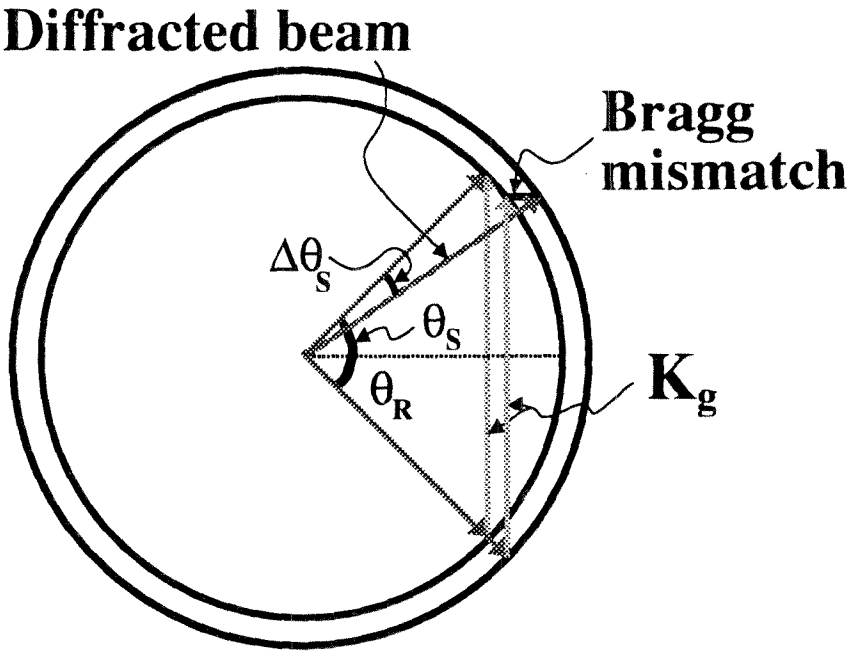


Figure 1.20: K–space for Bragg mismatch in wavelength multiplexing

the direction of the wavevectors.[31] The change in wavevectors is a length change, as

$$\frac{d\mathbf{k}_r}{d\lambda} = -\frac{1}{\lambda}\mathbf{k}_r \tag{1.22}$$

$$\frac{d\mathbf{k}_s}{d\lambda} = -\frac{1}{\lambda}\mathbf{k}_s. \tag{1.23}$$

There are then three components to the Bragg mismatch: the two wavelength terms above and the angle change of the diffracted beam. This third term is exactly analogous to the $\frac{d\mathbf{k}_s}{d\theta_S}$ term from above, and is written here as $\Delta\mathbf{k}_t$. As above, we use the x equation to solve for θ_S, and then solve for the z mismatch. The resulting wavelength selectivity is

$$\frac{d\lambda}{\lambda} = \frac{\lambda}{L}\frac{\cos\theta_S}{1-\cos(\theta_R+\theta_S)}. \tag{1.24}$$

[31]Note that the change shown in Figures 1.20 and 1.21 is a decrease in wavelength.

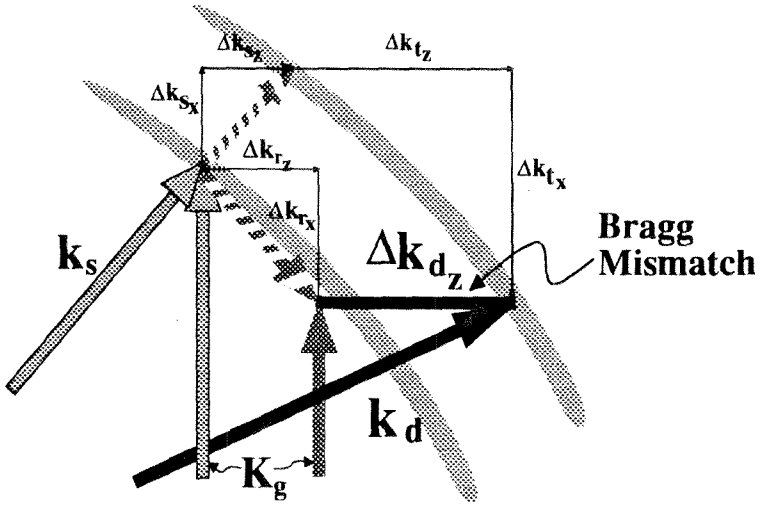


Figure 1.21: Blowup of Figure 1.20

1.6.4 Space–bandwidth product

The space–bandwidth product (SBP) is a measure of the number of resolvable points in an output image [8]. It is a useful term to consider because it is conserved during transmission through an optical system. From the front focal plane to the back focal plane of each lens, the position and angle of a ray change roles. As a result, the product of the maximum position and the maximum angle is constant. A lens can also be said to be capable of supporting a particular space–bandwidth product, or number of resolvable points.

However, SBP is not an all-encompassing measure of the capabilities or limitations of a lens. For instance, an optical beam that is too large will not fit through a lens no matter how small the spatial bandwidth. The same is true of a small beam with large angular range. While space–bandwidth product is useful for general descriptions of an optical system, analysis is best performed with the focal length and aperture lens, both singly and in combination as the f/#.

1.6.5 Liquid–crystal SLMs

Nematic liquid crystals

Liquid crystals are essentially collections of asymmetric organic molecules whose orientation and position become affected by neighboring molecules, leading to interesting bulk properties [313, 314]. Nematics are a class of liquid crystals typified by long–range ordering of molecule orientation, with unordered molecular positioning. Devices are made by sandwiching the liquid crystal material between two closely–spaced substrates. Boundary conditions for molecule orientation are set by chemical or mechanical preparation of the substrates. The OFF state of the device is determined by these boundary conditions—the ON state by an electric field applied across the material (using the substrates as electrodes).

A common boundary condition is twisted nematic, where the two substrates contain orthogonal scratches (or "rubbing") [313, 314]. The polarization of the light is "guided" through a rotation during transit of the liquid crystal cell [314].[32] Most commercial displays in laptops or projection televisions are supertwisted nematic (where the polarization rotation is more than 90°). Another interesting configuration is Hybrid Aligned Nematic or HAN [316], where one substrate has been chemically treated so that the molecules are oriented normal to the surface (called homeotropic alignment), while the other substrate has been rubbed (homogeneous alignment). This configuration is useful because amplitude and phase modulation through a full 2π can be performed with low voltages. [316]

Nematics are fast to respond to an electric field, but slow to return to the rest state when the field is removed. The turn–off time is typically 10–50 milliseconds and depends strongly on the substrate separation. The application of electric field is made complicated by the presence of ionic defects in the liquid crystal material. A DC field will over time separate these charges, creating an opposing space–charge field which will degrade device performance. Fortunately, nematics respond to the square of the electric field, so a DC–balanced square wave can be used for the ON

[32] The output polarization is often elliptical rather than linear. [315]

state [314]. The frequency of the square wave need only be faster than the response time of the ON state, so several kilohertz is fine. This is important when signals are carried to millions of pixels.

The slow turnoff of nematics can be exploited in array designs in which each ON pixel is activated and then left to slowly decay while the accessing electronics continue on through each pixel.[33] This can be done on a pixel column by column basis in a scheme called passive matrix addressing. However, this architecture is difficult to scale to large SLM arrays, and tends to lead to pixel "ghost" effects [317,318]. Passive multiplexing has generally been supplanted by active matrixing, where each pixel has a transistor and capacitor to store the pixel voltage.

Ferroelectric liquid crystals

Ferroelectrics are a subset of a large class of liquid crystals called smectics [314]. These liquid crystals have both orientational and positional ordering, so that the molecules line up in layers of identical orientation. In some materials, the long axis (or director) of the molecules is identical from layer to layer. In others, the director varies from layer to layer, tracing out a helical path (chiral smectics). Ferroelectrics are chiral smectics in which each molecule has a permanent dipole moment, which, as if things weren't complicated enough, is usually *not* parallel to the director [314]. A common device configuration is to space the two substrates so closely that the director is constrained to one of two angles [319]. (You can think of the helix as being squashed: the director can only be on one side of the helix or the other). The electric field operates on the dipole moment to switch the molecules from one state to another. As a result, this device is not capable of gray scale. However, ferroelectric liquid crystal (FLC) devices with true gray-scale modulation are possible with different substrate and electrode configurations [314].

Ferroelectrics are popular [320–324] among researchers making liquid crystal SLMs because the switching can be very fast. In fact, a gray scale display can be made by

[33]Much like an entertainer who sets plates spinning on posts and then runs back and forth keeping each plate from slowing down!

using many ON–OFF cycles during the integration time of the human visual system. Devices are now available which are capable of 1kHz modulation, delivering 8 bits of gray scale during each 33msec period [325]. However, the problem of DC balanced signals is more complicated with ferroelectric liquid crystals, since they respond to both field amplitude and direction [314, 326].

Transmissive SLMs

Transmissive SLMs are complicated to make because the wiring and electronics at each pixel tend to reduce the fill factor. In addition, thin–film transistors are difficult to fabricate with high yield. Despite these difficulties, impressive accomplishments have been made in transmissive liquid crystal displays. Since displays require only video rate performance, most use nematic liquid crystals. Color displays in laptop computers are usually made with RGB triads of pixels (using individual filters). In projection televisions, each of the three primary colors can be routed through a separate SLM, and recombined before projection. Until recently, many of the devices used for optical computing and storage experiments were from projection televisions, with the color filters removed. These have aliasing problems, since the number of pixels in the device is never quite exactly matched to the composite video format. However, it is now possible to purchase gray scale SLMs which are matched to the VGA computer display format [247, 248]. These make it possible to have control over each pixel in the SLM by using the graphic routines available in any PC.

Reflective SLMs

Most SLM researchers are developing reflective devices, as there are several advantages to these devices. For instance, the electronics for each pixel can be buried in the substrate underneath each pixel. As a result, each pixel can have complicated electronics (for "smart" pixels) *and* reasonable fill factor. Techniques have been developed which allow for fabrication of liquid crystal devices on the surface of standard VLSI chips [316, 321]. One of the substrates is a metal modulator pad on the chip, while

the other is a glass substrate coated with ITO.[34] This means that mature technology can be used to inexpensively fabricate the circuitry (through the MOSIS service, for instance) and researchers can concentrate on adding the liquid crystals and improving overall device performance. Reflective SLMs are also useful in that the optical path length (which one would like large, especially if modulating phase) is twice the substrate spacing (which one would like small for speed purposes). Both ferroelectric and nematic reflective SLMs are being developed [316, 321].

The drawback of reflective SLMs is that they are difficult to use in a holographic storage system. There is almost always a beamsplitter involved. If the beamsplitter is non–polarizing, the maximum efficiency[35] is automatically cut to 25%. If the beamsplitter is polarizing, expensive waveplates are required (in addition to the expensive beamsplitter). And there is always the question of finite extinction ratios and residual backreflections. However, the possibility of co-locating SLM and detector pixels on the same chip may make reflective SLMs worth the trouble. By using a conjugate reference beam on readout, we can redirect the reconstructed signal wave back to the original SLM [327, 328]. In this way, any aberrations are retraced in reverse, removing the problem of imaging the SLM through the object arm of the system. Since the system is self–aligning, fabrication costs are lower. In fact, all lenses can be removed from the signal arm and the SLM placed close to the storage crystal, making the system incredibly compact [327, 328]. We discuss the design of such systems in Chapter 7.8.2.

1.6.6 SLMs and fill factor

An SLM with finite fill factor can be described as an aperture with a transmission function of

$$\left[\left(f(x) \times \mathrm{comb}(\frac{x}{A})\right) * \mathrm{rect}(\frac{x}{fA})\right] \times \mathrm{rect}(\frac{x}{NA}), \tag{1.25}$$

[34] Indium Tin Oxide—a popular transparent conductor.

[35] i.e., the efficiency at unity fill factor.

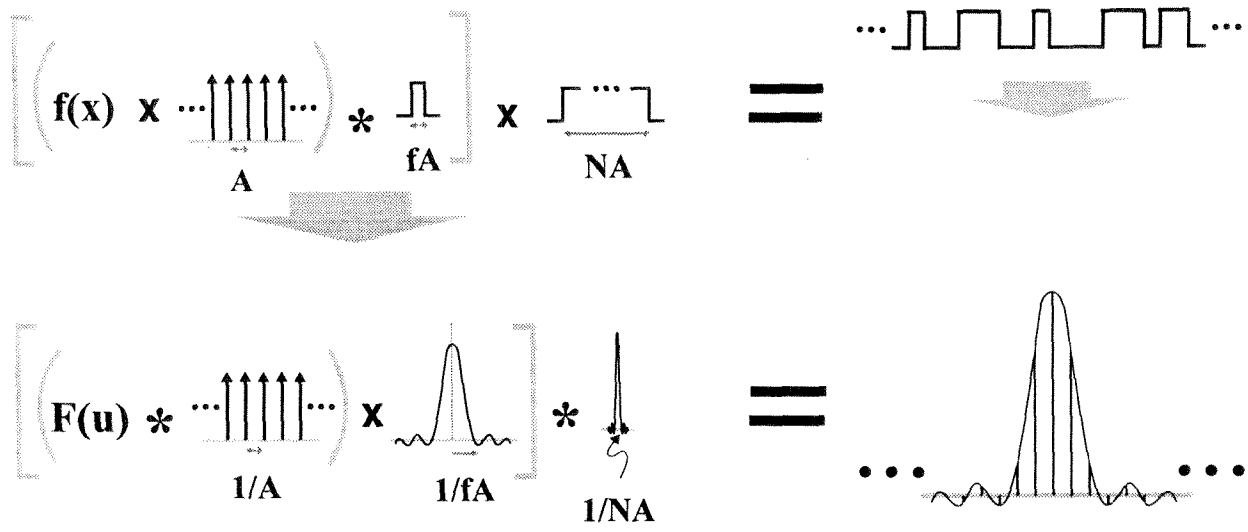


Figure 1.22: A 1–D SLM and its Fourier transform

where $f(x)$ is the function which describes our displayed data, A is the pixel spacing, f is the linear fill factor, N is the number of pixels, and $*$ represents convolution. This is shown in Figure 1.22. We are representing the SLM as one dimensional for simplicity. Note that, because we're sampling $f(x)$ at a (spatial) rate of $1/A$, it is band–limited to $1/(2A)$. The Fourier transform of the SLM (which appears in the focal plane of the first lens) is then

$$\left[\underbrace{(F(u) * \mathrm{comb}(uA))}_{1} \times \underbrace{\mathrm{sinc}(fAu)}_{2}\right] * \underbrace{\mathrm{sinc}(NAu)}_{3}, \tag{1.26}$$

where u is the coordinate in the Fourier transform plane (in terms of the physical coordinate x', $u = x'\lambda/F$). The third term of the aggregate Fourier transform is a very narrow sinc which is present because of the finite extent of the SLM. This will broaden everything very slightly. Term #1 is a set of multiple orders at regular spacing—each containing the complete information about the ON–OFF pattern displayed on the SLM. The orders are weighted by the broad sinc pattern (term #2) which is the Fourier transform of the pixel shape. Note that if the fill factor f is unity, then this broad sinc term is identically zero at each of the sampling points of the comb function.

The total efficiency of the 2–D SLM (besides any non–unity transmissivity of the ON state) is reduced by the light which does not pass through the pixels, and by the loss of light diffracted into the orders. The first effect alone has an efficiency of f^2, the second

$$\frac{1}{1+\left(\Sigma_{i=1}^{\infty}\mathrm{sinc}^2(if)\right)^2}. \tag{1.27}$$

In this equation, the interior square appears because the efficiency is calculated in terms of powers, not amplitudes; the exterior square because the same sum occurs both vertically and horizontally. The unity term on the top (and with the sum) corresponds to the central order. Equation 1.27 was used to plot Figure 1.4.

We can restore the loss due to the opaque areas around the pixels by using two micro–lenslet arrays. These arrays would need to have the same periodicity as the pixels themselves. The first lenslet array would have unity fill–factor, and the second would magnify the incoming light to the size of the SLM pixel and recollimate.[36] At this point, the efficiency (shown in Figure 1.23) depends only on the diffraction to the higher orders.

[36]Without recollimation, the spread of spatial frequencies in the incident light would broaden the Fourier transform pattern.

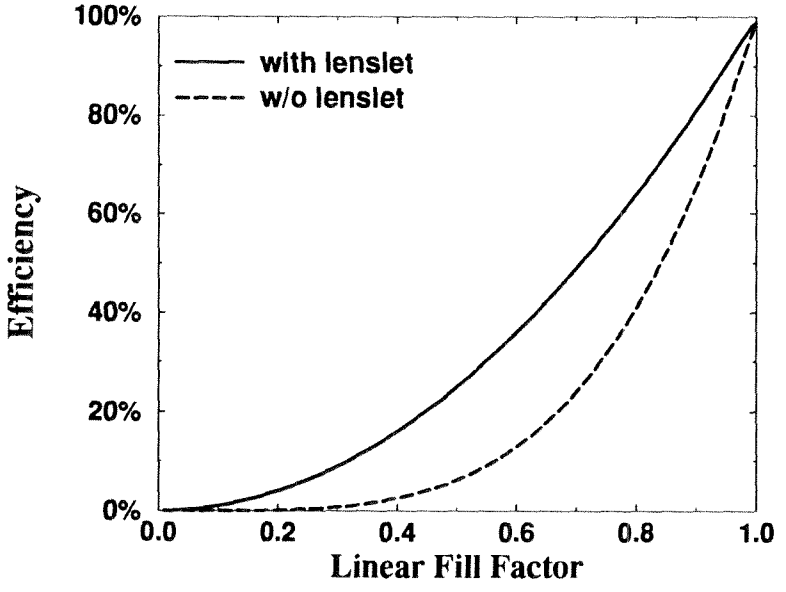


Figure 1.23: SLM efficiency, if a lenslet array is used so that all incident light to pass through the SLM pixels.

Chapter 2 Limitations on storage capacity

Contents

In this thesis, we will be concentrating on angle–multiplexed holographic storage in the 90° geometry. In this architecture, shown in Figure 2.1, the signal beam enters the small end face of a long rectangular crystal bar and propagates the length of the crystal—the wide reference beam enters the orthogonal face and meets the signal at 90°. This configuration is convenient for several reasons:

- the Bragg angle selectivity is small, allowing optimal use of the angular scanning range of the reference beams [137].
- both beams enter normal to the crystal, reducing Fresnel losses.
- the interaction length can be increased simply by extending the length of the crystal, increasing diffraction efficiency as well as Bragg selectivity [282]. For this reason, we often refer to the 90° geometry as the Long Interaction Length Architecture (LILA) [329].
- The path over which fanning of the strong reference beam can occur is set by the width of the signal beam, not by the interaction length. Since scattering of

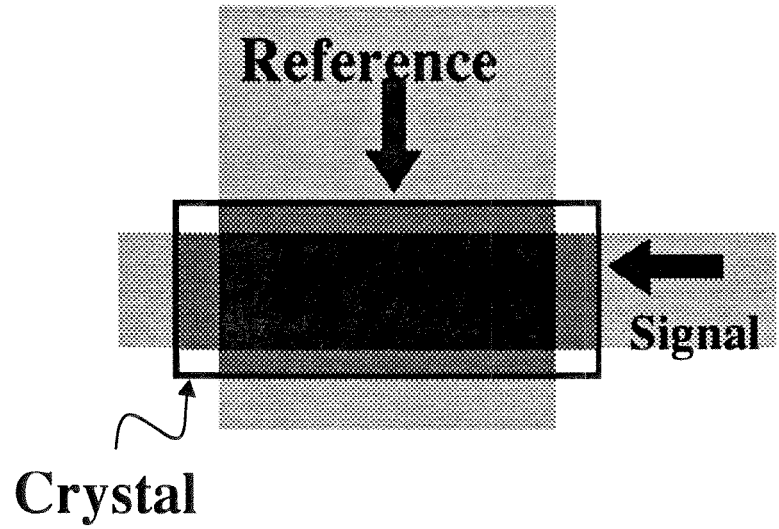


Figure 2.1: The 90° geometry (a) Recording, and (b) Readout

the reference beam is concentrated in the direction of the reference beam, the detector array receives very little.

- the characteristic photorefractive time constant is smaller in the 90° geometry, improving the average recording rate.

There are a few drawbacks to this configuration, including the smaller electrooptic coefficient, the possibility of large absorption losses, and the decreased photorefractive response at the relatively high spatial frequencies being stored. We explore the advantages and drawbacks of the 90° geometry in this chapter and in Chapter 4.

We begin this chapter by examining some factors that limit the capacity of a holographic storage system. Specifically, we're interested in the limitations on the number of holograms that can be stored in a single storage location.[1] We divide these factors into geometric limitations and dynamic range limitations and study each separately. Geometric considerations limit the number of reference beam angles which can be used. On the other hand, finite dynamic range causes the signal power to fall as we store more holograms.[2] We will find that our choice of angle–multiplexing

[1]This would be an incomplete consideration for holographic systems with thin materials, where the number of holograms is strongly linked to the number of pixels per hologram, and overall storage density is the overriding consideration.

[2]In addition, adding more pixels to the SLM without increasing the storage volume will tend reduce the number of photons per pixel.

in the 90° geometry gives us more than enough geometric flexibility in the reference beam. The limitation on the number of holograms is then the dynamic range of the storage medium.

In the latter part of this chapter, we study the dynamic range of photorefractive holographic storage. Although a hologram in a photorefractive material is erased by the exposure of a subsequent hologram, multiple holograms of equal diffraction efficiency can be recorded by using a set of decreasing exposure times [91, 157, 158, 330, 331]. From this recording schedule approach, we find that a single parameter (which we call $M/\#$) can be used to describe the dynamic range performance of our holographic storage system.

2.1 Geometric limitations

The number of distinct reference beams M that we can use for recording holograms is set by the total angular range we have to work with ($\Delta\Theta$), and the minimum angular spacing required for Bragg mismatch ($\delta\theta$), or

$$M \approx \frac{\Delta\Theta}{\delta\theta}. \tag{2.1}$$

The extent of $\Delta\Theta$ is often limited by the angle–deflector, especially in the case of acousto–optic or liquid crystal deflectors. We discuss the available deflection technologies in Chapter 7.5. If the angle–deflector is capable of arbitrarily large deflections, then the angle range is limited by the aperture of the lens system.

A typical lens system, shown in Figure 2.2, is a 4–F system[3] to image the angle change onto the surface of the crystal. In the center of the 4–F system (marked as Plane A in the figure), the reference plane wave focuses to a spot which moves back and forth horizontally as the angle is changed. Each of the lenses in the reference arm needs to be larger than this horizontal movement, which can be written as $f\Delta\Theta$

[3]Where the total system length is twice the sum of the focal lengths. The two lenses are also separated by the sum of the focal lengths. The magnification is the ratio of the two focal lengths, f_2/f_1.

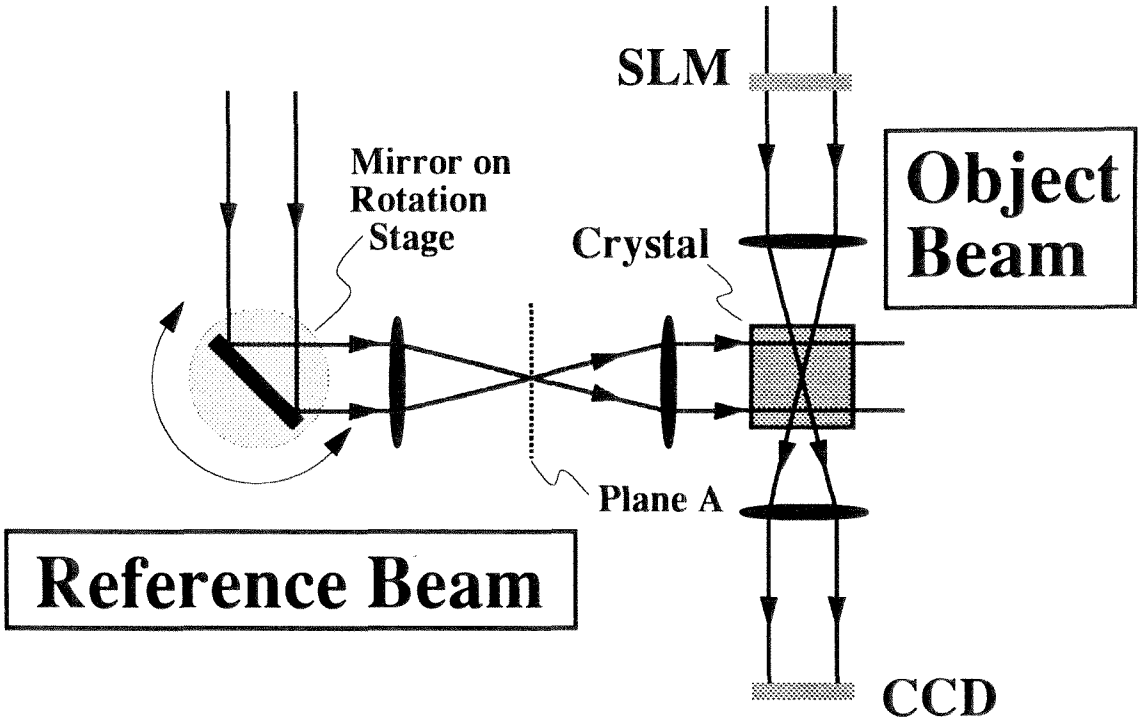


Figure 2.2: Holographic storage in the 90° geometry.

in the paraxial approximation. In addition, the lens diameter must pass the width of the reference beam. This consideration is graphically represented in Figure 2.3, and can be written as

$$A > f\Delta\Theta + L. \tag{2.2}$$

L is the width of the reference beam after collimation and A is the clear aperture of the lens. In practice, the full aperture of the lens cannot be used because of aberrations. If the reference beam passes through the lenses too far from the optical axis, it is no longer truly collimated and the angular selectivity curve of the holograms becomes broadened. Holography is very forgiving of the reference beam, though, and a non–collimated reference beam can be used as long as the same beam is available for readout. So the outside portions of the lens aperture can be used as long as the reference angle spacing is increased for them.[4] If the lens is a two–element achromat, the usable aperture is usually 70–85% of the full aperture—but with a low f/# plano–convex lens, it might be as low as 50–60%. These are rules of thumb based on personal experience—to accurately analyze the geometrical limitations of the signal beam, we would need to be much more precise about aberrations.

Angular selectivity

For the 90° geometry, we can use the expression for angular selectivity which we calculated in Section 1.4.1. However, we redefine the reference beam angle so that the center reference beam angle is at $\theta_R = 0$ (as shown in Figure 2.4) rather than at 90° (as it was in Figure 1.17. With this change, we can write the angular selectivity as

$$\delta\theta = \frac{\lambda}{L}\frac{\cos\theta_S}{\cos\left(\theta_R - \theta_S\right)}. \tag{2.3}$$

[4]A noncollimated reference beam causes a broadening of the selectivity function. The "new" selectivity function is the convolution of the spatial frequency distribution of the reference beam with the "old" selectivity function.

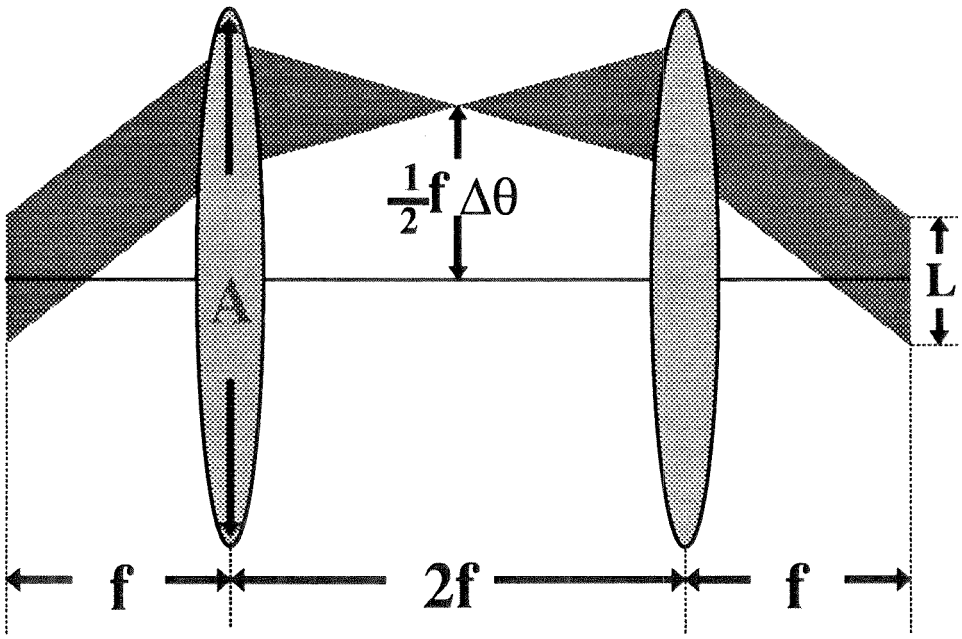


Figure 2.3: Aperture requirements in the reference arm

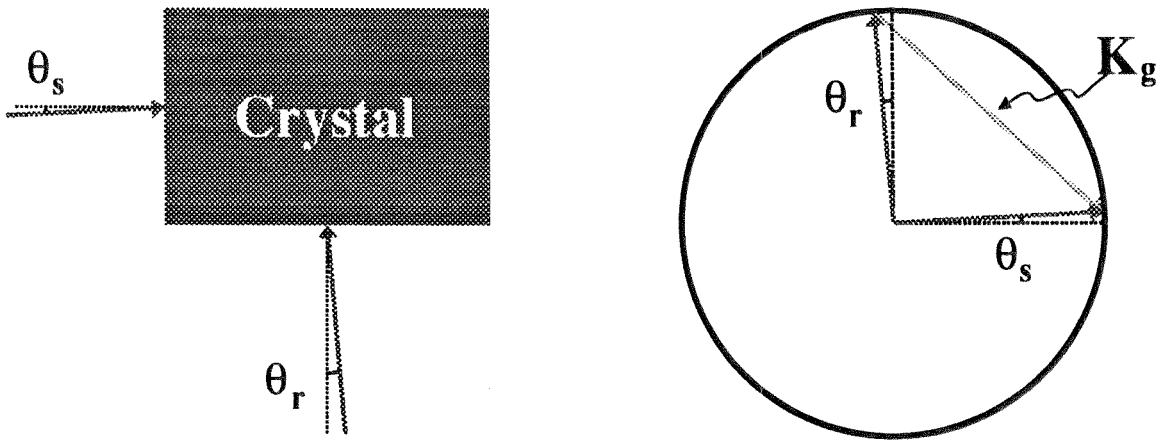


Figure 2.4: Angle definitions and K–space for the 90° geometry

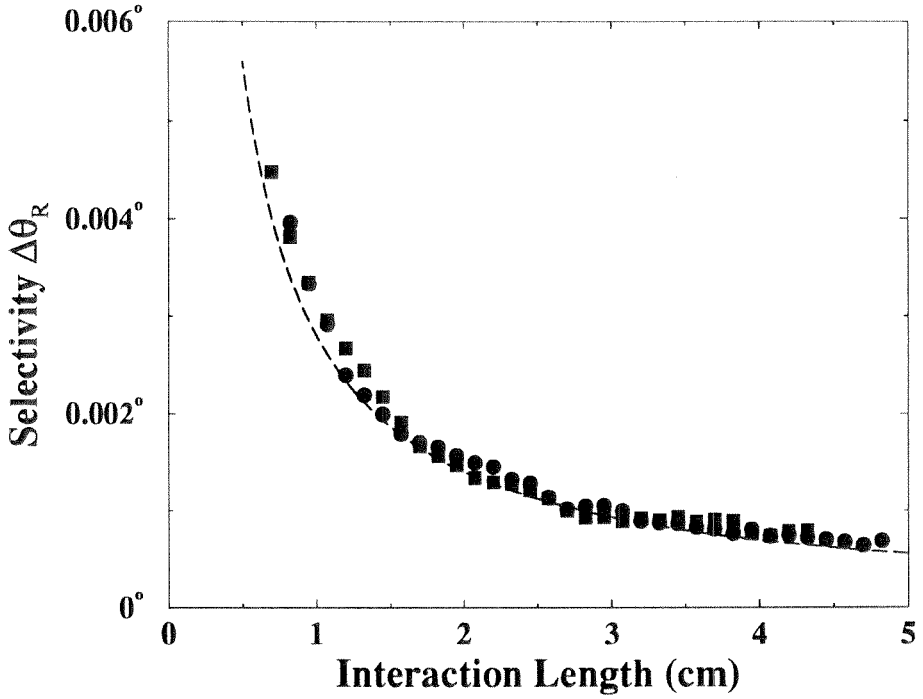


Figure 2.5: Measured angular selectivity in the 90° geometry as a function of interaction length, and theory from Equation 2.3.

As we mentioned, we can enhance the already impressive angular selectivity by increasing the interaction length L. A plot of the experimentally measured angular selectivity at $\theta_R \approx \theta_S \approx 0$ is shown in Figure 2.5, as a function of interaction length L. For the full crystal width of almost 5cm, note that an angle change of 5.6×10^{-4} ° is sufficient to Bragg mismatch a stored grating.

Since Equation 2.3 depends on θ_R, the angular spacing we should use will change across the angular range. This will be a problem if the angular spacing we are using is close to the resolution limit of our angle scanner. We will not have the luxury of using the exact spacing required for every reference angle θ_R. Fortunately, the angular spacing required in the 90° geometry does not change too much over the range of angles which one would like to use. In Figure 2.6, we show the angular selectivity as a function of θ_R for $\theta_S = 0$. Because the angular selectivity is dependent on θ_S, a signal beam containing a spread of horizontal spatial frequencies cannot be Bragg-mismatched over the whole wavefront by any one reference beam. We show this in Figure 2.6 by plotting the angular selectivity of two additional θ_S values. These

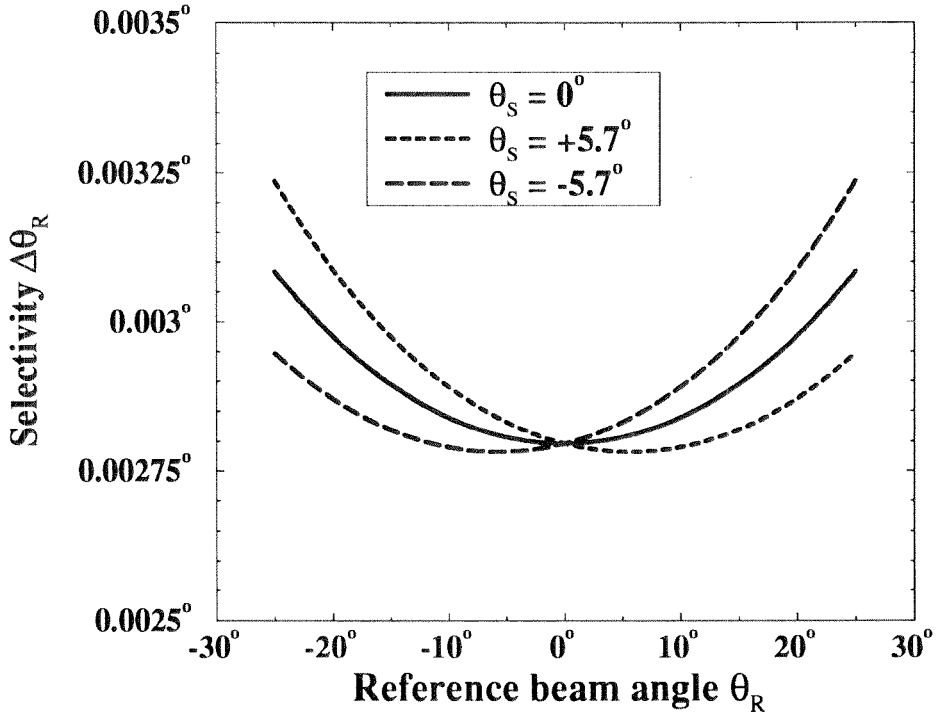


Figure 2.6: Angular selectivity in the 90° geometry

correspond to the angular extent of an 2cm wide SLM viewed through a 20cm focal length lens. Crosstalk analysis considers the sum of this imperfect Bragg mismatch from all of the undesired holograms multiplexed in the same volume.

2.2 Dynamic range limitations

Photorefractive crystals are widely used for holographic storage because incident light modulates their index of refraction [33]. Unfortunately, the same effect tends to erase already stored holograms during exposure of subsequent gratings. In the past decade, several techniques have been developed for storing multiple holograms of equal diffraction efficiency, including incremental [332] and scheduled [157] recording. These techniques involve careful consideration of the recording and erasure process. For instance, the recording schedule is a carefully chosen set of decreasing exposure times. Initial holograms are recorded with a large diffraction efficiency, and are erased by the subsequent exposures to the same diffraction efficiency as the final short exposure. This is shown in Figure 2.7.

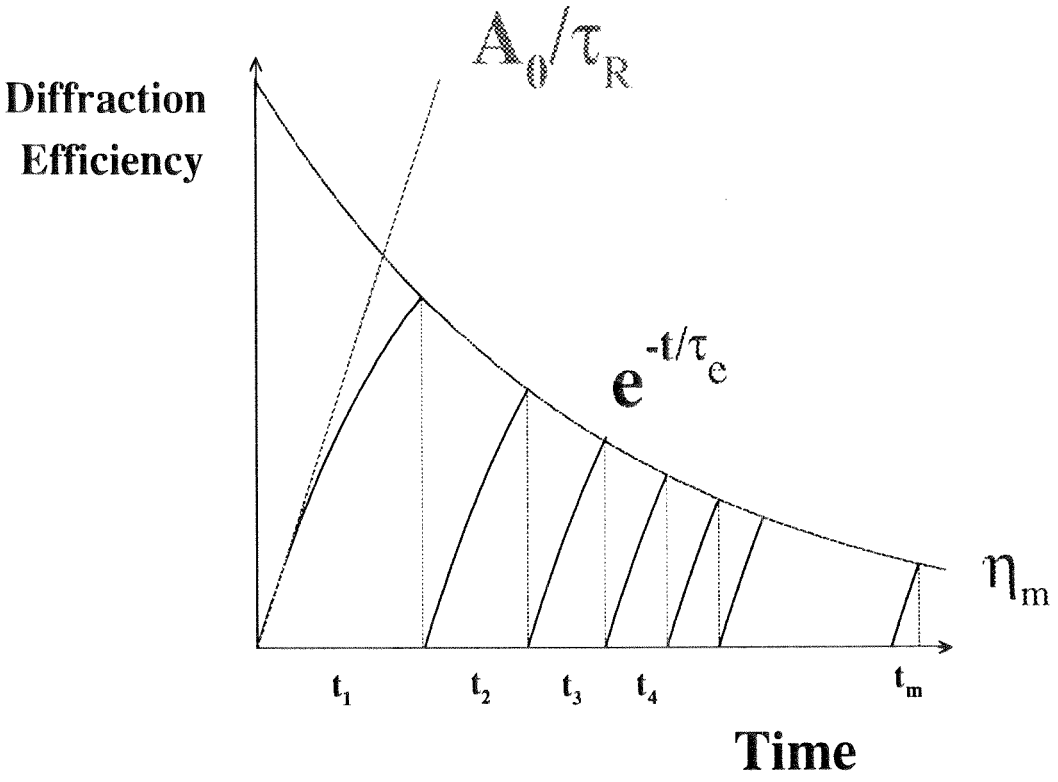


Figure 2.7: Typical recording schedule

These system techniques have provided new insight into the performance characterization of holographic recording systems. For instance, since recording and erasure are balanced, the intensity of the recording beams has no first–order effect on the diffraction efficiency of the resulting holograms (it does affect the recording speed). A second insight is that the first exposure time can be reduced to improve the average recording speed, with minimal loss in diffraction efficiency. But the chief insight from the recording schedule is the relationship between the number of holograms and the diffraction efficiency. For a large number of holograms M, the recording schedule (and the incremental recording procedure) predict that the equalized diffraction efficiency falls as $1/M^2$[[157, 332]. This makes intuitive sense, since the dynamic range available for diffracting the incident electric field is divided in M parts—the conversion to intensity provides the square.

These insights have shown that the parameters developed for comparing photorefractive *materials* are no longer suitable for characterizing holographic memory

systems. For instance, the a relatively easy material parameter to measure is the maximum index change. This involves a long exposure, during which the grating becomes strong enough to affect its own evolution. In contrast, the longest exposure in a recording schedule for 10,000 holograms may be less than 1/1000 of the characteristic erasure time constant. None of the self–interaction effects experienced during a single long exposure have any bearing on the diffraction efficiency expected from multiple weak holograms. As a result, extrapolating from the maximum index change to predict the diffraction efficiency of multiple holograms can be misleading.

A second parameter for comparing photorefractives is the material figure–of–merit, the product of a number of independently measured material parameters, such as index of refraction, dielectric constant, and electro–optic coefficient [333]. Each is difficult to measure, yet the resulting figure–of–merit is useful only for relative comparison of two crystals.

The most useful material parameter is the photorefractive sensitivity. There are many definitions, of which one is the change in diffraction efficiency per unit of exposure energy density per unit crystal length [333]. However, the recording schedule indicates that the equalized diffraction efficiency should be independent of intensity. The discrepancy here is that the photorefractive sensitivity describes only the recording behavior of the crystal. As we mentioned above, the storage of multiple holograms involves both recording and erasure behavior.

2.2.1 Recording schedule

The recording schedule, when used with the correct erasure time, results in holograms of equal diffraction efficiency.[5] But the question of primary interest is what this diffraction efficiency will be? To answer this question, we start with the the derivation of the recording schedule [157]. To calculate the proper set of exposure times, each

[5] I will sometimes refer to the diffraction efficiency of a hologram as its "strength."

hologram is assumed to evolve during recording as

$$A_0(1 - e^{-t/\tau_r}) \approx \left(\frac{A_0}{\tau_r}\right) t \tag{2.4}$$

and decay during erasure (subsequent recording of other holograms) as e^{-t/τ_e}. A_0 is the saturation strength of the hologram, τ_r is the characteristic recording time constant, and τ_e the characteristic erasure time. If τ_r is purely real, then τ_e and τ_r should be identical (See Chapter 4.5.3). In writing Equation 2.4 for the recording behavior, we have assumed that each exposure time is much shorter than τ_r. We are also tacitly assuming that the intensities involved are not large enough to induce nonlinear effects (< 1 W/cm^2 in the case of $LiNbO_3$:Fe). This recording and erasure behavior is supported by experimental observation and the Kukhtarev model for photorefractive dynamics (Chapter 4.5 or Reference [94]). For a large number of holograms M, the final equalized diffraction efficiency is

$$\eta = \left[\left(\frac{A_0}{\tau_r}\right)\frac{\tau_e}{M}\right]^2. \tag{2.5}$$

The original proof of this relation was given in Reference [157]. In the next section, we show how Equation 2.5 follows from the "backwards" recursion algorithm [159] that we use in our laboratory.

Simple "backwards" recursion algorithm

Conventionally, the exposure schedule is calculated by a forward recursion algorithm: an initial time is chosen, and recording times are calculated up to the M^{th} exposure. Knowledge of the recording and erasure time constants is required. Here we introduce a simple "backwards" recursion algorithm, which does not require knowledge of the recording time constant. We summarize the algorithm as follows: given the erasure time constant τ_e and the desired number of holograms M, the last exposure is selected

to be

$$t_M = f\,\frac{\tau_e}{M}, \tag{2.6}$$

where f is some constant ≤ 1. The correct erasure time τ_e can be obtained from measurement (as we describe in Chapter 3.2), or by trial and error. Each previous exposure time is calculated recursively using the expression

$$t_{m-1} = t_m\, e^{t_m/\tau_e}, \tag{2.7}$$

until the first recording time t_1. This exposure will be finite if $f < 1$.

We should verify that each hologram has the same diffraction efficiency at the end of the exposure schedule. Consider exposure m, which was recorded for an exposure time of

$$t_m = t_M\, e^{\sum_{i=m+1}^{M} t_i/\tau_e}. \tag{2.8}$$

The diffraction efficiency (right after exposure m) is

$$\eta_m = \left[\left(\frac{A_0}{\tau_r}\right)\, t_m\right]^2. \tag{2.9}$$

Hologram m is then erased while the subsequent holograms are recorded in the same location. The erasure occurs over a time period of

$$\sum_{i=m+1}^{M} t_i/\tau_e, \tag{2.10}$$

so that the final diffraction efficiency of hologram m is

$$\eta_m = \left[(A_0/\tau_r)\, t_M\right]^2. \tag{2.11}$$

Since this does not depend on m, all the holograms have the same final diffraction

efficiency. A typical set of measured diffraction efficiencies is shown in Figure 3.13. Note that $t_M = f\tau_e/M$ from Equation 2.6, which returns us to Equation 2.5 if $f \approx 1$. We discuss the f parameter in Section 2.2.3.

In the next section, we rewrite Equation 2.5 by incorporating the three parameters (A_0, τ_r, and τ_e) together to form one parameter, which we call $M/\#$. We show how this $M/\#$ parameter succinctly characterizes the performance of a holographic memory system.

2.2.2 M/#–a holographic *system* metric

The M/# is the constant of proportionality between diffraction efficiency and the number of holograms squared. It is a function of the many variables in a holographic recording system. However, as we will show in this section, the M/# of a system can be measured from the recording and erasure of a single hologram [159].

We define $M/\#$ by rewriting Equation 2.5 as

$$M/\# = \left(\frac{A_0}{\tau_r}\right)\tau_e, \tag{2.12}$$

so that

$$\eta = \left[\frac{M/\#}{M}\right]^2. \tag{2.13}$$

One might be tempted to think of the $M/\#$ as the combination of the three variables: A_0, τ_r, and τ_e. However, we saw in the previous section that the first two variables always appear in the ratio (A_0/τ_r), which we refer to as the recording slope. Although it might be useful to know these two separately, we really only need to know their ratio. This is really fortunate, since A_0 and τ_r are extremely difficult to measure separately.[6] This realization simplifies the measurement of the $M/\#$ considerably.

With this interpretation of A_0 and τ_r, the $M/\#$ is the product of the writing

[6]The reason is the same one which complicated the measurement of the maximum index change—A_0 and τ_r measured after a long exposure don't correspond to the value of (A_0/τ_r) which is in effect during the intial exposure of the hologram.

slope (A_0/τ_r) and the erasure time (τ_e). The $M/\#$ confirms that, for multiple strong holograms, a photorefractive crystal needs to have fast recording and slow erasure. As we have mentioned, the $M/\#$ can also be interpreted as the scaling constant for the $1/M^2$ relationship between diffraction efficiency and the number of holograms M. As a result, an improvement in $M/\#$ is a direct improvement in system performance, whether it is applied towards more holograms or higher diffraction efficiency.

Since it is a system measure, the $M/\#$ is a function of many material and system variables. We discuss some of these in Section 2.2.4. In the next section, we show that despite this dependence on numerous factors, the $M/\#$ can be conveniently measured from the growth and decay of a single hologram.

Measuring M/#

The writing slope (A_0/τ_r) is the slope of $\sqrt{\eta}$ as a function of time during hologram formation. This can be obtained from a before–and–after diffraction efficiency measurement of a single short exposure. Several measurements during an exposure (or measurement of multiple holograms of varying exposure) can provide reassurance that the recording process is linear in the range of interest. Note that a probe beam of different wavelength or polarization may give a diffraction efficiency that is substantially different from that expected with the original reference recording beam. The beam used for single hologram measurement should be the same one used for readout in the multiple hologram system being characterized.

Measurement of the erasure time τ_e is conceptually simple, but in practice involves some care. The erasure illumination should be identical to what a hologram will experience during the "remainder" of the recording schedule. This often precludes the use of a third erasure beam, and requires both signal and reference. However, neither reference nor signal should Bragg–match the original grating of interest during erasure, or two–wave mixing may occur and affect the measured diffraction efficiency. Coupling can be avoided either by turning the crystal slightly, or by changing the input angle of the reference beam and imprinting random pages of data on the information–bearing signal beam periodically during erasure. Since the same grating may be

measured at intervals during erasure, the probe beam must not significantly affect the hologram. This can be done with a weakened reference beam, or a probe beam of different wavelength or polarization. In any case, the readout beam should be scanned (in angle or wavelength) across the selectivity curve for each measurement, so that diffraction efficiency is always measured at the Bragg–match condition. After all this, the erasure time τ_e of the grating is then twice the inverse of the slope of $\log \eta$ as a function of time, and the $M/\#$ the product of τ_e and A_0/τ_r.

2.2.3 Practical use of $M/\#$

Expectation vs. reality

This $M/\#$, measured from the recording and erasure of a single hologram, can be used to predict the diffraction efficiency of M holograms using an *ideal* recording schedule. This ideal recording schedule has a final exposure of length τ_e/M, but an the initial exposure which is inordinately long. Any practical initial exposure implies that the final exposure is some fraction, f, of τ_e/M. Since this choice is still a recording schedule, all M holograms will be equalized in diffraction efficiency, but to $\eta = f^2\left(\frac{M/\#}{M}\right)^2$. An equivalent way of thinking of this is that an ideal exposure schedule for $M + X$ holograms has been formulated, but the first X holograms were skipped and recording started with exposure X+1. X, the number of "skipped" holograms, and f, the fraction of $M/\#$ used, are related by

$$f = \frac{M}{M + X}. \tag{2.14}$$

The relationship between X and f is plotted in Figure 2.8 for M=1000 holograms.

The parameter f (or, equivalently, X) is useful because it allows control over the average recording time (for a given intensity). Fortunately, our "backwards" recursion formulation of the recording schedule allows explicit selection of the fraction f. Proper choice of f can reduce the total recording time to practical values with a moderate sacrifice in diffraction efficiency. Figure 2.9 shows the relationship between the average

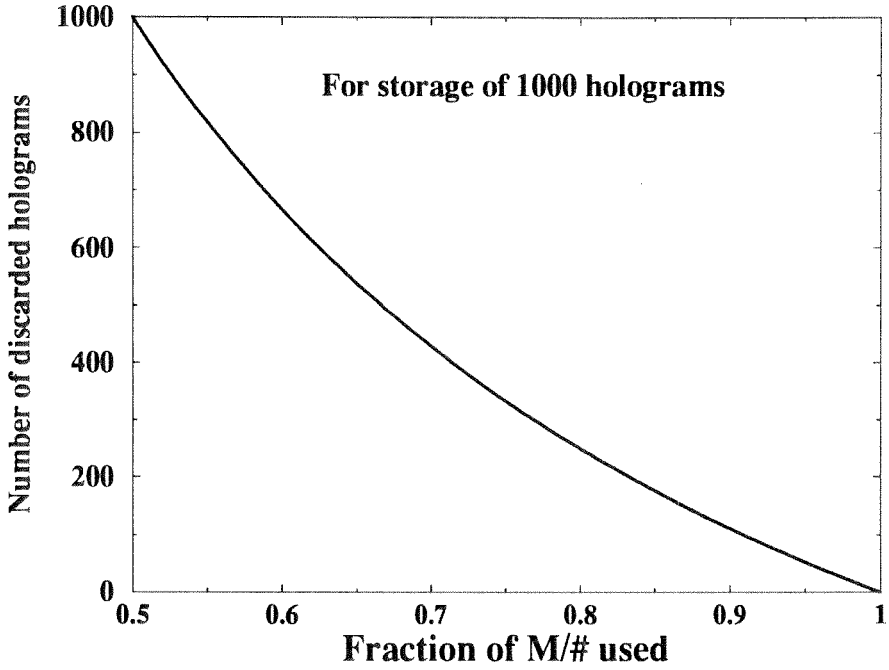


Figure 2.8: Number of holograms discarded from recording schedule, versus fraction f of $M/\#$ used, for storage of 1000 holograms.

recording time and the fraction f for M=1000 holograms. Note that large f values have a high average recording time because of long initial exposures. Additionally, during these first holograms, the plane–wave components of the signal beam can build up interpixel gratings, causing unacceptable distortion of high resolution images.

Effects of using an incorrect time constant

As we said earlier, the recording schedule results in holograms with equal diffraction efficiency. However, the algorithm requires that the user know the correct erasure time constant. If the wrong time constant is used, then the diffraction efficiencies will not be uniform. The erasure time constant can be measured by the procedure detailed above (Section 2.2.2). However, it is sometimes more convenient to find it by trial–and–error. In this case, we need to be able to look at the measured diffraction efficiencies of the recorded holograms, and determine the direction and magnitude of our error in time constant. We might also want to get an idea of what the diffraction efficiency will be once we get all the holograms equalized. We analyze the effects of

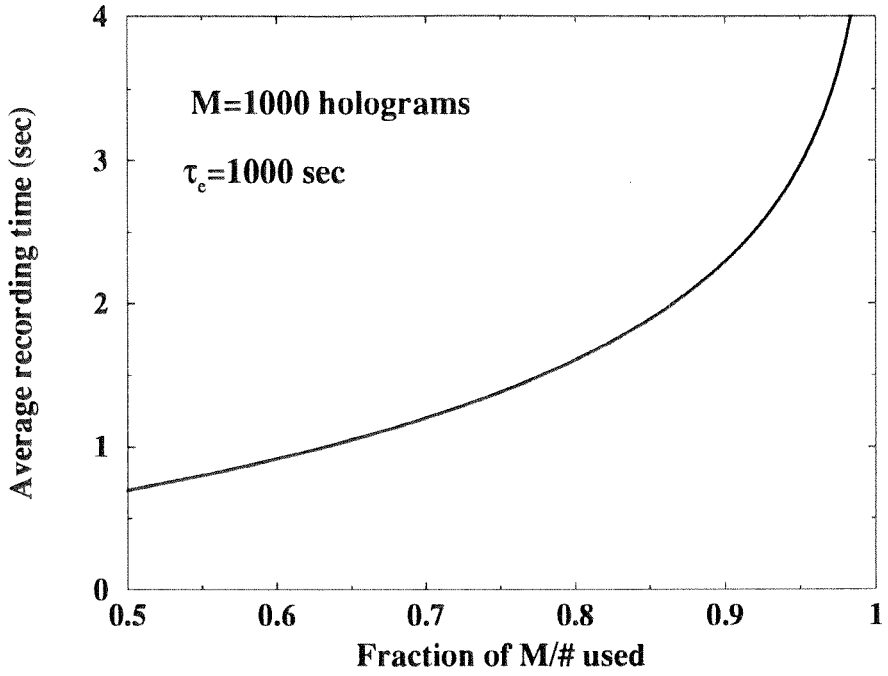


Figure 2.9: Average recording time versus fraction f of $M/\#$ used.

incorrect time constant with our "backwards" recursive description of the recording schedule.

The effects of using the wrong erasure time constant can be divided into two parts: the ratio of the first hologram to the final hologram (a measure of the nonuniformity with the incorrect schedule), and the expected change in diffraction efficiency of the final hologram (between incorrect and correct schedules). Figure 2.10 compares the operation of an incorrect schedule ($\tau_e = \tau_{wrong}$) with the correct schedule ($\tau_e = \tau_{right}$). The scenarios of erring too high (Figure 2.10a) and too low (Figure 2.10b) are both represented. In each figure, we show the first and last holograms. The dashed erasure line shows what we expected to happen—that is, the first hologram erases to the same level to which the final exposure is recorded. The solid erasure line corresponds to what actually happened. In Figure 2.10a, the erasure was slower than we expected, so the first hologram ends up much stronger than the last. In contrast, if the first hologram ends up weaker than the final hologram, then our guess for the erasure time constant was too high. The numerical analysis of the scenarios, however, does not depend on the sign of the error in time constant. For instance, the final exposure time

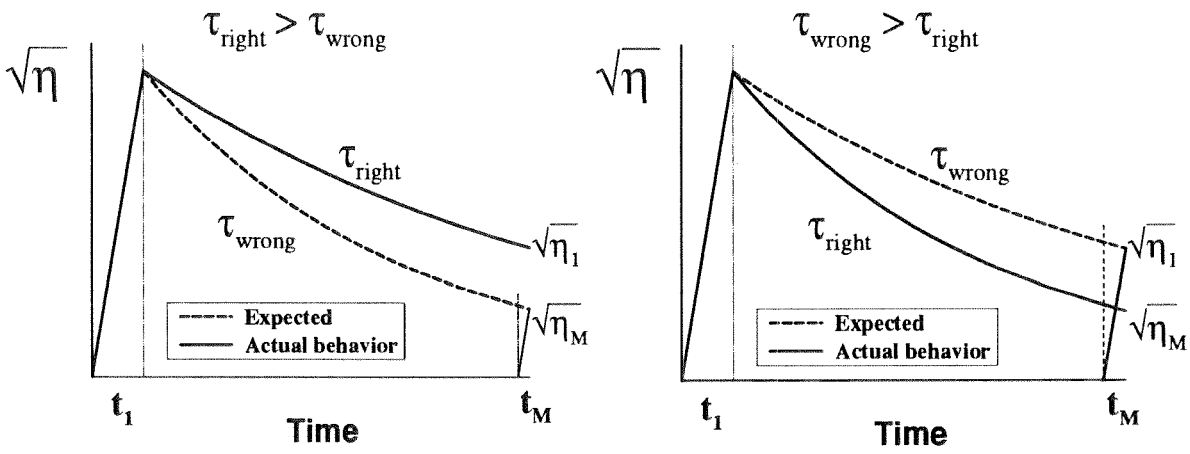


Figure 2.10: Two scenarios for wrong recording schedule: (a) $\tau_{wrong} > \tau_{right}$, (b) $\tau_{right} > \tau_{wrong}$.

in any schedule is proportional to the τ_e used to make the schedule. The diffraction efficiency of the final hologram is proportional to the square of this exposure time. As a result, the ratio of final hologram diffraction efficiency, between correct and incorrect schedules, is $(\tau_{right}/\tau_{wrong})^2$.

Holograms made with an incorrect schedule do not get equalized in diffraction efficiency—we would like to calculate the shape of this nonuniformity. Our incorrect recording schedule, made with τ_{wrong}, calculates

$$t_1 = t_M \, e^{\sum_{i=2}^{M} t_i/\tau_{wrong}} \tag{2.15}$$

with the expectation that $\eta_1 = \eta_M$. However, since the real erasure time constant is τ_{right}, the first hologram erases to the point where

$$\sqrt{\eta_1} = \frac{A_0}{\tau_r} \, t_1 \, e^{-\sum_{i=2}^{M} t/\tau_{right}} . \tag{2.16}$$

Using Equation 2.15, the ratio of the diffraction efficiency from the first hologram to

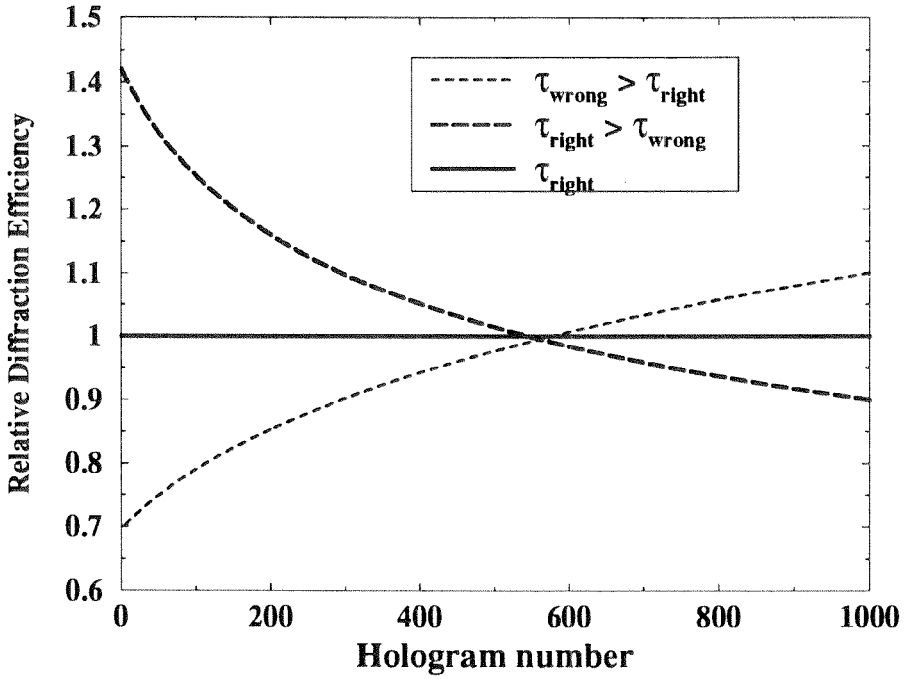


Figure 2.11: Nonuniformity in diffraction efficiency for 1000 holograms with f=0.9 and τ_{right}=1000 seconds: (a) τ_{wrong} = 900 seconds (b) τ_{wrong} = 1100 seconds.

that of the last hologram can be written as

$$\frac{\eta_1}{\eta_M} = \left(\frac{t_1}{t_M}\right)^{2(1-\tau_{wrong}/\tau_{right})}. \tag{2.17}$$

The same equation applies to hologram m with substitution of t_m for t_1. In Figure 2.11, we use Equation 2.17 to show the effect on η_m of a 10% error in τ_e is shown for M = 1000 and f = 0.9. As we had intuitively expected from Figure 2.10, underguessing the time constant makes the first hologram stronger than the last, and vice–versa for overguessing. Note that the value of t_1/t_M depends strongly on the choice of f, with weak dependence on M. In Figure 2.12, we show the ratio t_1/t_M as a function of M for several values of f.

Let's suppose that we have written M holograms using an incorrect recording schedule. How do we use our measured diffraction efficiency to find the right time constant? Once we know the t_1, t_M, and τ_{wrong} that were used, we solve for τ_{right}

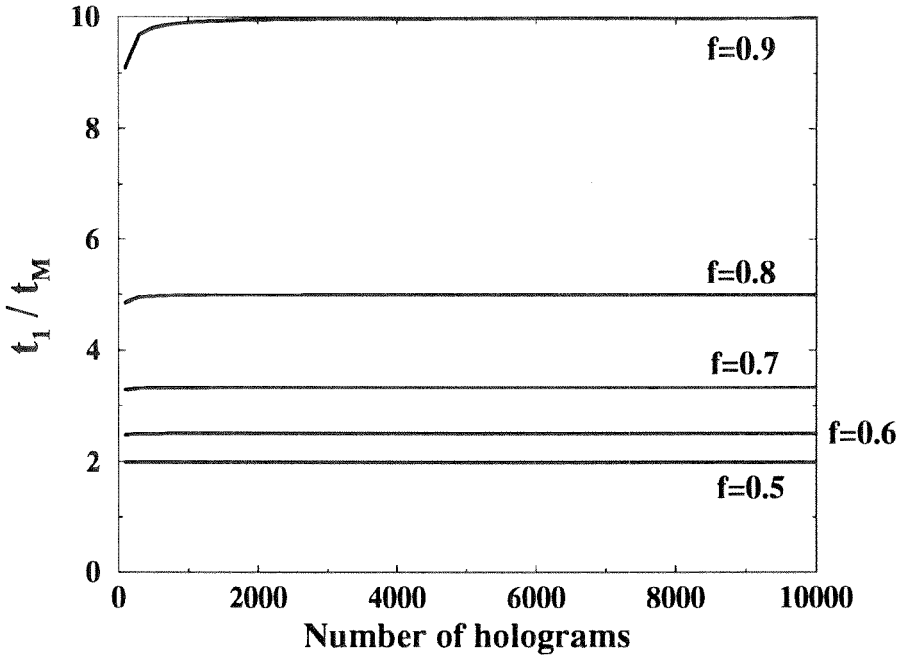


Figure 2.12: Dependence of t_1/t_M on M, for several values of fractional $M/\#$.

from Equation 2.17,

$$\frac{1}{\tau_{right}} = \frac{1}{\tau_{wrong}} \left(1 - \frac{\ln \frac{\eta_1}{\eta_M}}{2 \ln \frac{t_1}{t_M}} \right). \tag{2.18}$$

In practice, since there is occasionally random fluctuation in measured diffraction efficiencies, it is usually best to analyze the diffraction efficiency of several holograms and use the average when guessing the new value of τ_{right}.

Recording schedules for non–exponential behavior There are several cases where the recording and erasure process cannot be described with a single constant. This can occur when the recording slope varies from hologram to hologram, or because holograms (and thus subsequent erasure) are only partially overlapping in volume. The first effect occurs for peristrophic multiplexing in photorefractives, because the projection of grating vectors on the c–axis changes during the recording schedule. Variance in recording slope can also occur because of absorption, as when the signal beam is moved horizontally in the 90° geometry. Non–exponential erasure can occur

through absorption effects or by partial overlapping in shift multiplexing, peristrophic multiplexing, and spatioangular multiplexing. In addition, researchers have observed non–exponential erasure through the interaction of multiple trapping species in the photorefractive effect [96–103]. We discuss these effects in the context of our $M/\#$ theory in Chapter 4.

We will take the same general approach no matter what causes the non–exponential behavior. The trick is to re–examine our "backwards" recursion formula,

$$t_m = \underbrace{t_M}_{1} \underbrace{e^{\sum_{i=m+1}^{M} t_i/\tau_{wrong}}}_{2}. \tag{2.19}$$

This formula calculates each exposure time by knowing how much diffraction efficiency we will get from the exposure (Part #1), and how much we will lose during the remainder of the schedule (Part #2). We can write this in a general form as

$$\sqrt{\eta_m} = \sqrt{\eta_M} \times \left[\text{Loss due to erasure during time } \sum_{i=m+1}^{M} t_i\right]. \tag{2.20}$$

If we write the recording slope for hologram m as $(A_0/\tau_r)_m$, and the erasure at hologram m over the subsequent time t as $l_m(t)$, then we can rewrite the recursion algorithm as

$$t_m = \left[t_M \frac{\left(\frac{A_0}{\tau_r}\right)_M}{\left(\frac{A_0}{\tau_r}\right)_m}\right] l_m\left(\sum_{i=m+1}^{M} t_i\right). \tag{2.21}$$

All we need is the recording slope of each hologram, and the erasure behavior during subsequent exposures. We will use this approach in Chapter 4 when analyzing the effects of moving the signal beam horizontally in the 90° geometry (similar to spatioangular multiplexing in some ways).

2.2.4 Factors that affect M/#

In this section, we describe some of the factors that can affect the $M/\#$ (or dynamic range) of our holographic storage system. We have divided these into two parts: those affecting the growth and decay of the local space–charge field, and those affecting the volume sum of the electromagnetic field scattered from this perturbation. This separation foreshadows the approach we will take in Chapter 4 when making theoretical predictions of the $M/\#$.

Factors affecting the space–charge field

Modulation depth (during recording) The modulation depth or fringe visibility of the interference pattern (often written as m), is the ratio of spatial modulation that we care about $(\mathbf{S} \cdot \mathbf{R}^*)$ to the background spatially uniform light $(S^2 + R^2)$. We will see in Chapter 4 that the interaction of the modulation depth in the A_0 term and the total intensity in τ_R would seem to make the diffraction efficiency of a recorded hologram proportional to $(SR)^2$, not to the square of modulation depth m^2. However, diffraction efficiency is also proportional to the square of the exposure time. In a recording schedule, we have selected this exposure time to be proportional to τ_e, which contains the total intensity $(S^2 + R^2)$. As a result, the diffraction efficiency of a hologram in the recording schedule is still proportional to m^2.

In a crystal with a significant absorption coefficient, the modulation depth is a spatially varying quantity. This will have an effect on the spatial frequency spectrum of our diffracted wave (Chapter 4.3.1), and will influence the choice of external modulation depth needed to obtain maximum diffraction efficiency (Chapter 4.2).

Crystal doping and Oxidation State The photorefractive effect in $LiNbO_3$ can be enhanced by doping with iron [115]. We will see in Chapter 4 that the occupancy of the Fe^{2+} state and the Fe^{3+} state are both important in the recording and erasure dynamics of the photorefractive crystal. Since iron is not generally found to be present in any other state [334], we can describe these two occupancies by the total Fe doping

and the oxidation state (quantitatively described by the ratio Fe^{2+}/Fe^{3+}). The total Fe doping is determined during crystal growth. The oxidation state, though, can be varied by annealing [118]. We will show experimental results of holographic recording as a function of oxidation state in Chapter 3, and then compare these experiments to a theoretical model in Chapter 4. To motivate these experiments, the last section in this chapter discusses the effect that we expect the oxidation state to have on the dynamic range of our crystal (Section 2.2.5).

We should note that crystal doping level and dark conductivity have been empirically observed to be related. This will be important when we use our theoretical model to examine the effect of doping on $M/\#$, since we will need to include dark conductivity to see the tradeoffs involved in choosing the doping level.

Wavelength during recording The wavelength used for recording is important since the photosensitivity of a photorefractive crystal is wavelength–dependent. Recording should be done at a wavelength inside the absorption band of the crystal and/or its dopants. It is also important that a large fraction of the absorbed photons result in a carrier in the conduction band (or valence band), which depends on the "oscillator strength" of the absorber transition [118].

Grating period The spatial transport of charges in the photorefractive effect generally occurs as the movement of a vast number of electrons over very small distances between excitation and retrapping. The photorefractive time constant is therefore affected by the interplay between the mean travel distance of the electrons and the grating spacing. As the grating spacing gets much larger than the mean travel distance of the charges, diffusion becomes less effective (because the local charge gradient is smaller) and the space–charge field does not become as large. In contrast, when the grating period is very small, the space–charge electric field becomes trap–limited.[7] In general, space–charge field falls off for both very large and very small grating periods.

[7]There just aren't enough trapping sites to support a large charge gradient, which limits the space–charge field.

Externally applied field/Photovoltaic effect The presence of an externally applied field adds drift to the charge transport processes. Likewise, a bulk photovoltaic effect is present in some photorefractive crystals (including $LiNbO_3$), which adds another charge transport process (through preferential charge excitation along the c–axis). Both of these affect the magnitude of the space–charge field, as well as its spatial frequency response (see Chapter 4).

Setup stability/laser coherence length These factors reflect the loss of dynamic range due to partial coherence between the reference and signal beams, or due to disturbances in the interference pattern during recording. A change in path length in one of the two arms will cause the interference pattern to shift inside the hologram. If this occurs during recording, the hologram will have several out–of–phase contributions, leading to a serious degradation in diffraction efficiency. Unlike the other factors in this section, we do not include stability in our $M/\#$ theory in Chapter 4.

Factors affecting the volume diffraction

Size and shape of interaction volume The diffraction efficiency of the hologram is affected by the interaction length over which the diffracted wave is integrated. In the 90° geometry, both the interaction length and the width of the signal beam are important. Intuitively, this is because the hologram scatters an incident intensity into an output intensity, while the diffraction efficiency is a ratio of beam powers. The grating volume also affects the diffraction efficiency through absorption, since the absorption loss increases with the distance traveled in the crystal.

Electro–optic coefficients The reconstructed wavefront is the scattering from an index (or permittivity) modulation, not the space–charge field itself. In photorefractive inorganic crystals, the linear electro–optic effect creates an index modulation from the space–charge field. Selection of a material and a geometry with a large electro–optic coefficient is an excellent way to get large diffraction efficiencies.

Bulk absorption coefficient As we mentioned, the absorption loss in the crystal reduces the externally observed diffraction efficiency. For this reason, the size of the index modulation (Δn) is not a good measure of the dynamic range of a material. Unfortunately, we can't just remove the absorption, because the photosensitivity of the material depends on the absorption. In $LiNbO_3$:Fe, the link between photosensitivity and absorption coefficient is the Fe^{2+} concentration [118]. We discuss this in the next section.

Ability to re–Bragg–match holograms This describes any mismatch between the readout wave and the reference wave used for recording. If the wavefront controller[8] is reasonably good (or if we have continuous control over the wavefront and feedback from the diffracted beam), this should not be a problem. However, if a mechanical stage with backlash or limited resolution is used, or the hologram is read with a phase–conjugate reference beam (Chapter 8.3), wavefront recall might be an issue.

Wavelength during readout The wavelength during readout can differ from the wavelength of light used for recording if the incident angle is changed appropriately. However, there is the issue of the limited signal bandwidth which can be acceptably reconstructed (Chapter 7.4.1). There is an additional consideration as well: as we mentioned in Chapter 1, the volume integration of the hologram occurs not over some distance, but over some number of wavelengths (the angular selectivity depends on the ratio of L to λ). This same integration over interaction length affects diffraction efficiency. As a result, moving to a longer wavelength reduces diffraction efficiency by reducing the effective interaction length.

2.2.5 Oxidation State

One way to improve the $M/\#$ is to alter the oxidation state of the $LiNbO_3$ crystal. It has long been established that Fe is present in $LiNbO_3$ only as Fe^{2+} or Fe^{3+} [335],

[8] See Chapter 1.3.1.

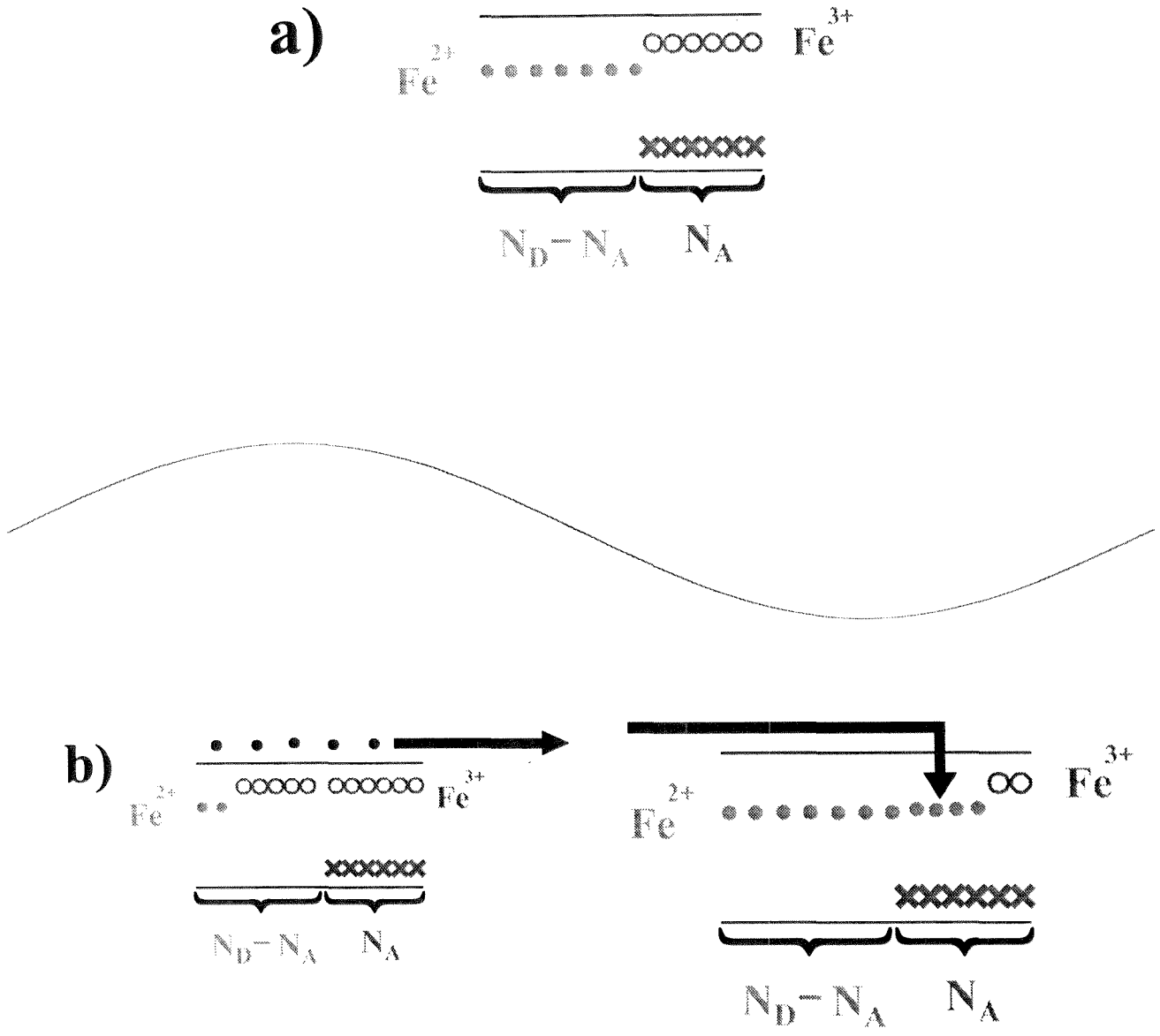


Figure 2.13: Band diagrams for $LiNbO_3$:Fe: (a) before recording, (b) at two spatial locations in the interference pattern during recording.

which take the roles of photon absorber and electron trap, respectively, in the photorefractive effect. This is shown in Figure 2.13a, where we show a stylized band diagram of $LiNbO_3$:Fe. Note that the number of Fe^{3+} is determined by the occupancy of a deep acceptor level N_A through overall charge neutrality. In $LiNbO_3$:Fe, this deep level is believed to represent the filling of oxygen vacancies.[9] The absence of Fe^{4+} is supported by the lack of a contribution to the absorption spectrum by the Fe^{3+} state, as well as by independent measurements such as Mossbauer spectroscopy [118,335].

The oxidation state of the Fe impurities in the crystal controls the ratio of Fe^{2+} to Fe^{3+}, and thus the recording and erasure characteristics [118]. In Figure 2.13b, we show the operation of the photorefractive effect by visualizing the dopant occupancies at two spatial locations in the illuminating interference pattern. We can see that if the Fe dopants were all in one state, then no space–charge field can develop. In one case, there would be no absorption of light. In the other, there would be no traps available for spatial separation of absorbed electrons. Based on this, we expect the saturation (steady–state) space–charge field to fall for very large and very small Fe^{2+}/Fe^{3+} ratios. The saturation space–charge field is shown (normalized to one) in Figure 2.14.

However, we mentioned above that there is an interaction between the absorption coefficient and the photosensitivity. This manifests itself when we realize that the bulk absorption coefficient should depend on the Fe^{2+} concentration. In fact, there is a linear relationship between these two, as was empirically observed for absorption of extraordinary polarized light at 450nm by Phillips et al [118]. Their data, shown in Figure 2.15, was obtained by matching absorption measurements with the Fe^{2+} concentration obtained from EPR and atomic absorption spectroscopy measurements.

The effect of this increase in absorption coefficient with increasing (Fe^{2+}/Fe^{3+}) ratio is an overall loss in diffraction efficiency $e^{-\alpha L}$. For crystals of centimeter thicknesses, this loss becomes intractably large for absorption coefficients over $2cm^{-1}$ or so. If we examine this same range of oxidation states in Figure 2.14, we find that it corresponds to small values of (Fe^{2+}/Fe^{3+}), for which the space–charge field is mono-

[9]There are plenty of oxygen vacancies in congruent $LiNbO_3$ because it is only 48.6% Li.

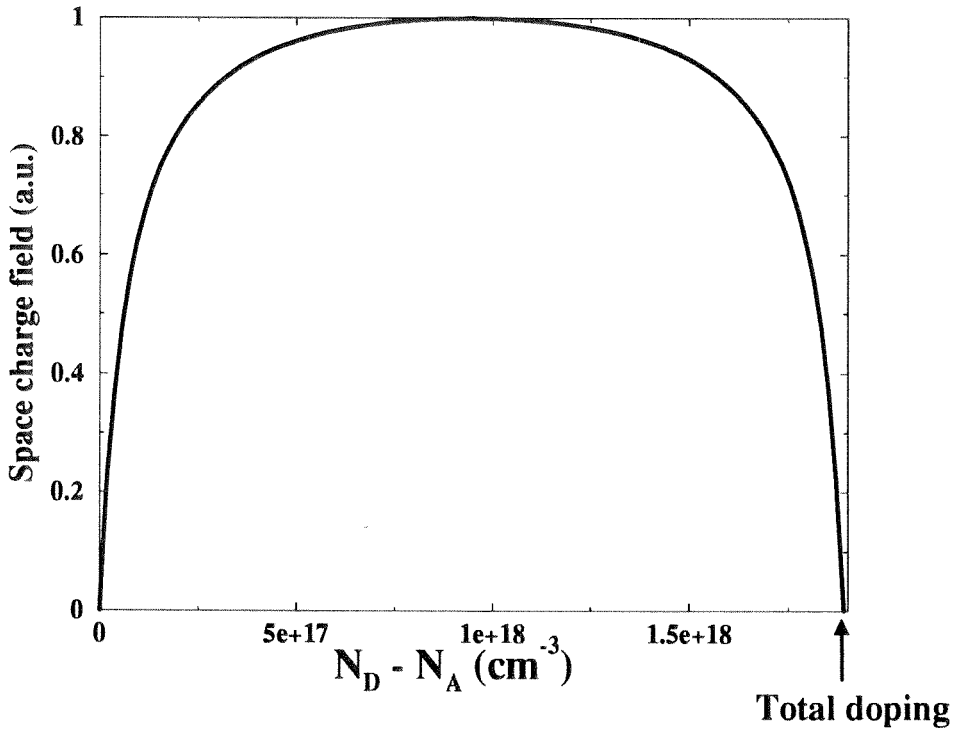


Figure 2.14: Effect of oxidation state on space–charge field.

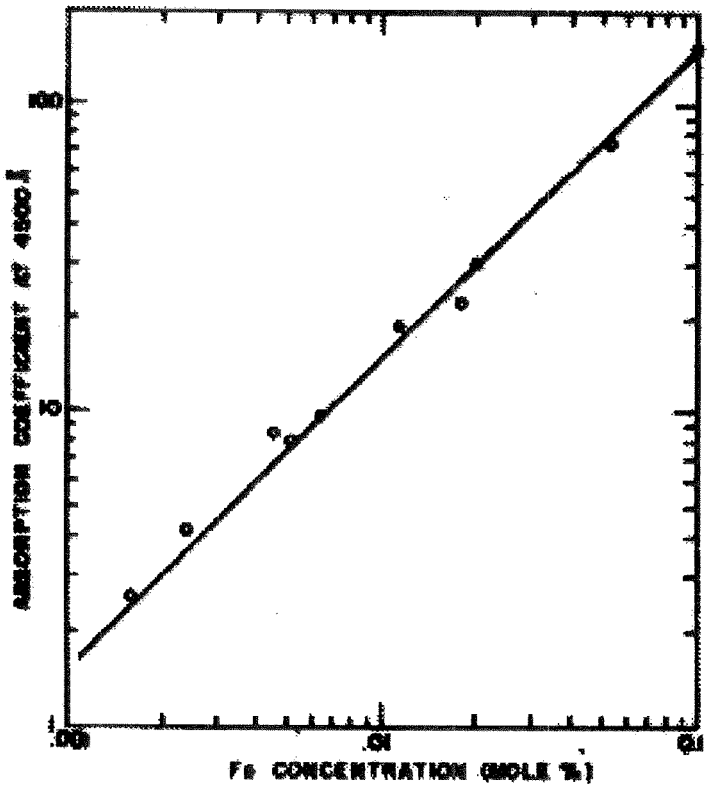


Figure 2.15: Relationship between Fe^{2+} concentration and absorption coefficient at 450nm (from Reference [118]). Data for fully reduced crystals of various doping levels.

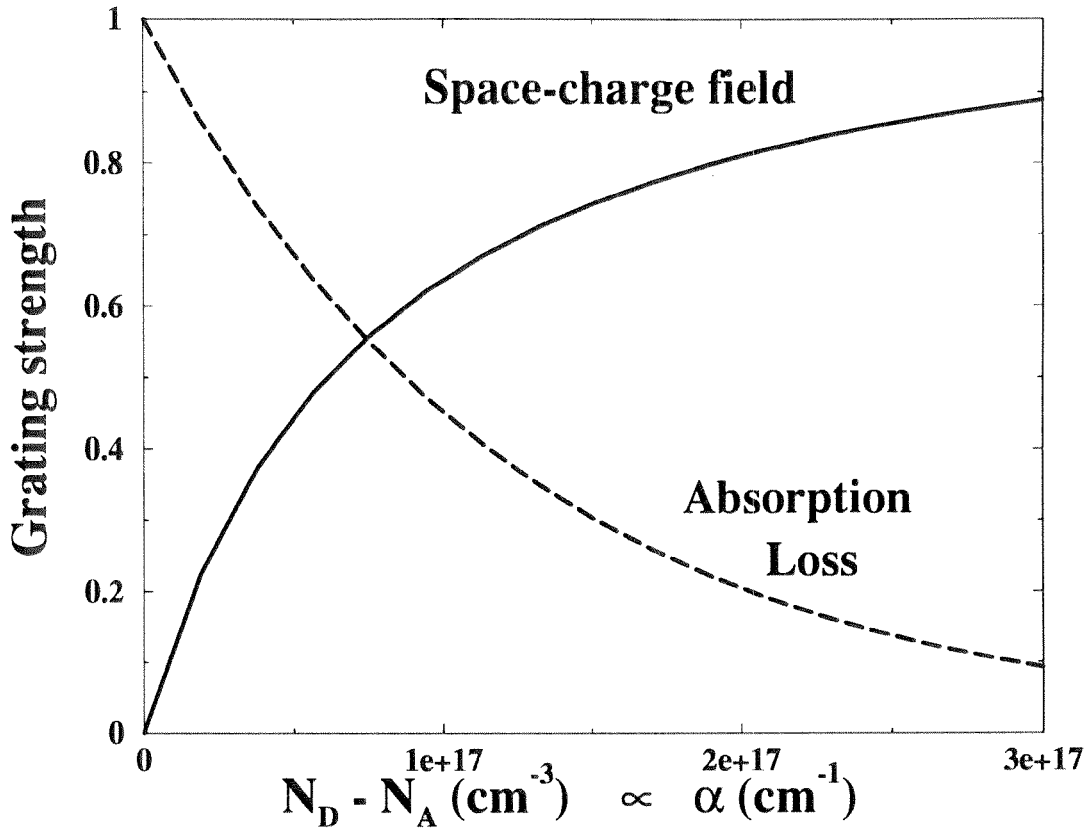


Figure 2.16: Counteracting effects of the dependence of space–charge field and absorption loss on oxidation state.

tonically increasing.[10] We replot the normalized saturation space–charge field within this range in Figure 2.16, along with a normalized absorption loss for a 2cm long crystal. Because diffraction efficiency is affected by both of these curves, we expect to find an optimal oxidation state: one which maximizes the dynamic range. We can conveniently gauge this dynamic range by measuring the $M/\#$ from the recording and erasure of a single hologram.

To find and use this optimum, we need to be able to change the oxidation state in $LiNbO_3$:Fe. We do this by high temperature annealing in argon, with oxygen partial pressure or temperature as the control variable [118]. We keep track of the oxidation state through the absorption coefficient. From Figure 2.15, the optical absorption at the center wavelength of the Fe^{2+} absorption band (450–475nm) is a

[10]This is true for doping levels over 0.01%, which is the smallest doping level that is generally considered (or available) for use in holographic storage.

direct measure of the Fe^{2+} concentration [118,334]. In the next chapter, we describe the experimental modification of the oxidation state of $LiNbO_3$:Fe crystals by high–temperature annealing. We measure the M/# as a function of oxidation state and show that there is an peak, as expected. In Chapter 4, we develop a theoretical model which predicts the position of this optimal absorption coefficient and the shape of the M/# dependence.

Chapter 3 M/#: Experiments

Contents

In this chapter, we discuss the experimental measurement of $M/\#$ as a function of $LiNbO_3$:Fe oxidation state. We performed the experimental measurements on a 1cm × 1cm × 2cm crystal from the Shanghai Institute of Nonmetals, doped 0.01% Fe. The crystal was nominally congruent, with the **c**–axis at 45° to the 4 vertical faces, and parallel to the 2 horizontal faces (1cm × 2cm).[1] In the first part of the chapter, we describe the oxidation and reduction procedure we used, and in the latter part, we describe the optical setup and $M/\#$ measurement procedure. In an appendix, we briefly describe the polishing setup which we used on some of our $LiNbO_3$ crystals.

Note: The main purpose of this chapter is to provide enough detail that the reader can repeat these experiments or convince themselves that our results are valid. If you would like to skip the gory details, there are three main results to see:

[1] a standard "90° geometry cut."

1. The experimentally measured $M/\#$ as a function of absorption coefficient, in Figure 3.11 on page 112. This graph shows that there is an optimal oxidation state for the 90° geometry in $LiNbO_3$:Fe.

2. The experimentally measured diffraction efficiency as a function of the number of stored holograms, in Figure 3.14 on page 116. This graph demonstrates the $1/M^2$ dependence of diffraction efficiency, and shows the effect of the choice of f when constructing a recording schedule.[2]

3. The comparison between the $M/\#$ measured from recording/erasure of a single hologram, and the $M/\#$ achieved during storage of multiple holograms. This is shown in Table 3.1 on page 116, and verifies that the $M/\#$ measured from a single hologram is a good predictor of the diffraction efficiency of multiple holograms.

3.1 Oxidation and Reduction of $LiNbO_3$:Fe

3.1.1 High–temperature annealing

Manipulation of the oxidation state was described by Phillips *et al* in Reference [118]. In this paper, the RCA group showed how to control the Fe^{3+} concentration in $LiNbO_3$:Fe by annealing in the presence of argon. They concluded that vacuum annealing or annealing with H_2 damaged their crystals by inducing significant dark conductivity.[3] Most of their work was with thin crystals, so they controlled the degree of reduction by annealing temperature (in the range 800—1150°C) followed by rapid cooling, or by oxygen partial pressure. By using pure argon with a temperature of 1150°C, they were able to induce a maximum Fe^{2+}/Fe^{3+} ratio of 9. In a related experiment, they reported reduction ratios >20 by packing the $LiNbO_3$ crystals in powdered $LiCO_3$. They did not perform many experiments with crystals thicker than a millimeter, but did report that with slow cooling (required to keep a thick crys-

[2]See Chapter 2.2.1 for a definition of this f parameter.
[3]We will discuss this below.

tal from cracking) the degree of reduction was less than that achieved with rapid cooling from the same temperature. In conclusion, they measured the dependence of absorption coefficient on Fe^{2+} concentration which we showed in Figure 2.15.

Our approach was to try to repeat these experiments for thick crystals, inducing absorption changes from near zero to $2cm^{-1}$, and then measure $M/\#$. We used oxygen partial pressure to control the degree of oxidation, and annealing time to control the degree of reduction. Our initial annealing experiments resulted in a rapid degradation of crystal surfaces. We then made three corrections to our system, which corrected the problems:[4]

- We replaced the quartz tube in the furnace (see description below).
- We replaced the quartz "boat" which supports the crystal while in the furnace. Our first boat contained two parallel cylindrical rods, and the crystals would often settle between them and become scratched. We describe the second boat below.
- We moved from a temperature range of $\sim$ 850°C to 950°C. Our motivation was to avoid a second–phase separation which takes place at temperatures between 750°C and 850°C [336].[5] This transition is from $LiNbO_3$ to $LiNb_3O_8$, and has been observed by some as an internal cloudiness in the crystal [337, 338]. We suspect that some of our observed surface degradation may have been this second–phase separation.

We describe the setup of our high–temperature furnace (with the changes described above) in the next section. We then describe the oxidation and reduction procedure used.

Furnace setup —We used a Thermco Ana-Lock Furnace, Controller Series 201 with a temperature range of 400-1400°C. There were three heating zones along the 1.08m length of the ceramic heater of inside diameter 9cm. The quartz tube, of

[4] or at least reduced the rate of surface degradation remarkably.

[5] The exact temperature of the transition is extremely sensitive to crystal stoichiometry.

length 1.25m and inside diameter 7.5cm, rested inside the ceramic heater without any thermal packing or insulation. The gases were routed to the furnace via 1/4 inch Tygon tubing attached to the quartz tube with a hose clamp. No special sealants or covers were used.

The boat used was a 0.8cm thick quartz plate of size 15cm × 5cm. A hole through the plate at one end allowed remote manipulation. The plate contained 6 quartz bumps of size 2.5cm × 0.75cm each, separated by 1.5cm, formed by melting 1mm thick slabs onto the boat. We typically placed our 2cm long crystal across the first two bumps, so that the center of the crystal was 48cm from the front edge of the ceramic heater.

The control settings for the three zones were set so that the temperature was uniform over the length of the boat. In Figure 3.1, we show a temperature profile, measured by a thermocouple with the front end uncapped, as a function of distance along the furnace. The temperature was measured at the bottom of the tube.[6] From Figure 3.1, the temperature at the center of the crystal had a gradient of $< 0.02°C/mm$.

We used 99.999% pure Argon[7] and 99.993% pure Oxygen[8]. The cylinders were always opened to a gauge pressure of $20lbs/in^2$. The flow was controlled by a Matheson 603 (E500) flowmeter on each line, and the two gasses combined with a "T" connector before routing to the furnace. Early on, we realized that the oxygen flowmeter did not have enough resolution for the small oxygen partial pressures we wanted to use. A second flowmeter of finer resolution[9] was attached in parallel. Since we did not have access to a calibration chart for the second flowmeter, we performed a rough calibration against the first flowmeter. However, some of the flow rates we used in our experiments were of such small values on the second flowmeter that the accuracy of an extrapolated flowrate is uncertain.

[6]The center of the crystal during annealing was 2cm from the bottom of the tube.

[7]$O_2 < 1ppm$, $H_2O < 1ppm$.

[8]$H_2O < 2ppm$.

[9]The metal float of the second flowmeter saturated when the metal float of the first flowmeter read 23 out of 150.

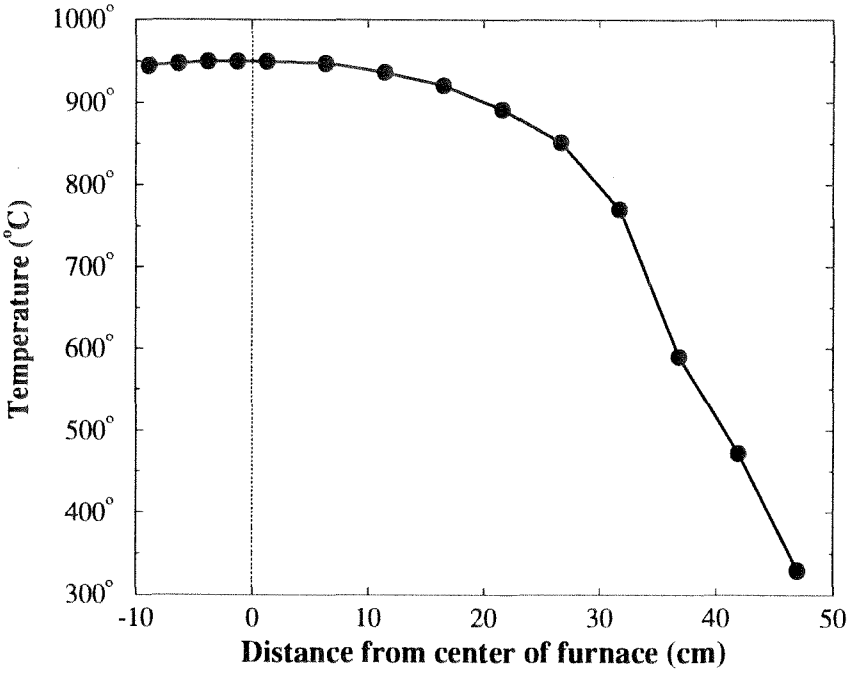


Figure 3.1: Measured temperature profile of our furnace.

Annealing procedure —Before a crystal was annealed for the first time (and periodically during a set of experiments), the top and bottom surfaces were roughened with 15μm grit. The aim here was to increase the surface area available for oxygen in–diffusion or out–gassing. Before each anneal, the crystal was cleaned with acetone and methanol, then placed in ultrasound for 10–15 minutes each in consecutive baths of TCE, acetone, and methanol. Between each ultrasound cycle, we rinsed the crystal and beaker in deionized water. The crystal was dried in N_2, inspected for waterspots, and placed on the boat at the entrance of the furnace.

At this point, the interior of the furnace would be approximately 500°C and argon gas had been flowing for more than 15 minutes. The crystal was pushed into the center of the "cold" furnace and the temperature control turned to 950°. The rate of temperature rise is shown in Figure 3.2. If we were performing an oxidation of the crystal, the oxygen was not introduced to the gas mix until after the furnace reached high temperature (we usually waited 1 hour). We empirically observed that oxidation was taking place much more quickly than reduction, and felt that the oxidation process needed to be precisely timed. At the end of the time period we

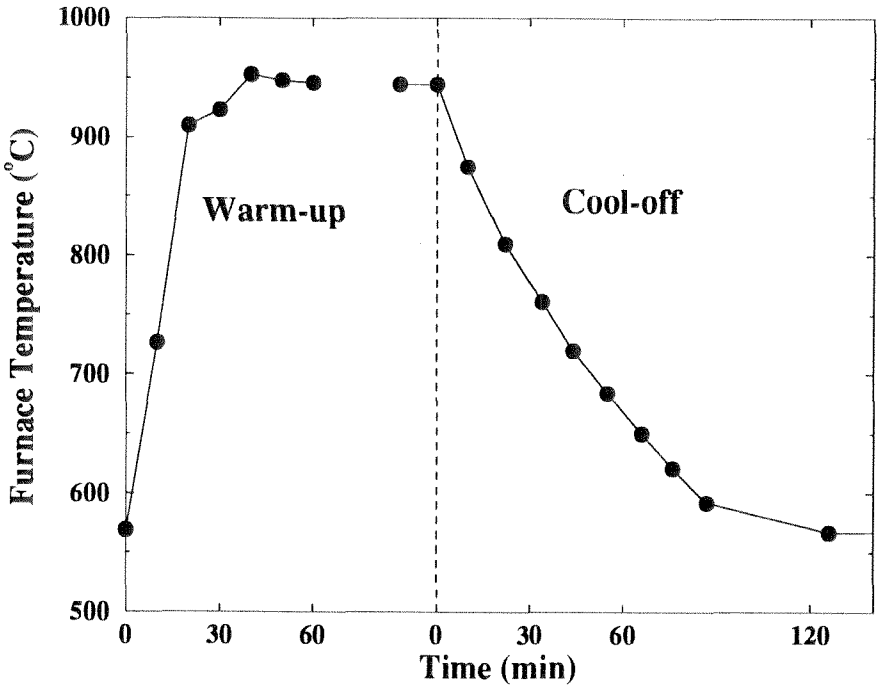


Figure 3.2: Warm–up and cool–off at the center of the furnace, measured with the front end uncapped.

planned to use, we turned the temperature control down and the oxygen flow off simultaneously. The only time we left the oxygen on during cooling was when we were using 50/50 argon/oxygen for full oxidation of the crystal. If the annealing step was a reduction, then no oxygen was involved at all. The typical cool–off rate at the center of the furnace is shown in Figure 3.2. To speed the cooling of the crystal, we opened the front cap of the tube and pulled the crystal boat out slowly A typical cooling rate was one 5cm pull every 5–10 minutes. Once the crystal was out of the high temperature part (noted by the color of the furnace), we left it to cool slowly before measuring absorption.

We describe our absorption measurements in Section 3.1.2. In the next section, we describe our empirical observations of absorption coefficient vs. our control parameters of time and oxygen partial pressure.

Control of the absorption coefficient The purpose of our oxidation/reduction procedure was to obtain different oxidation states of the same $LiNbO_3$:Fe crystal for

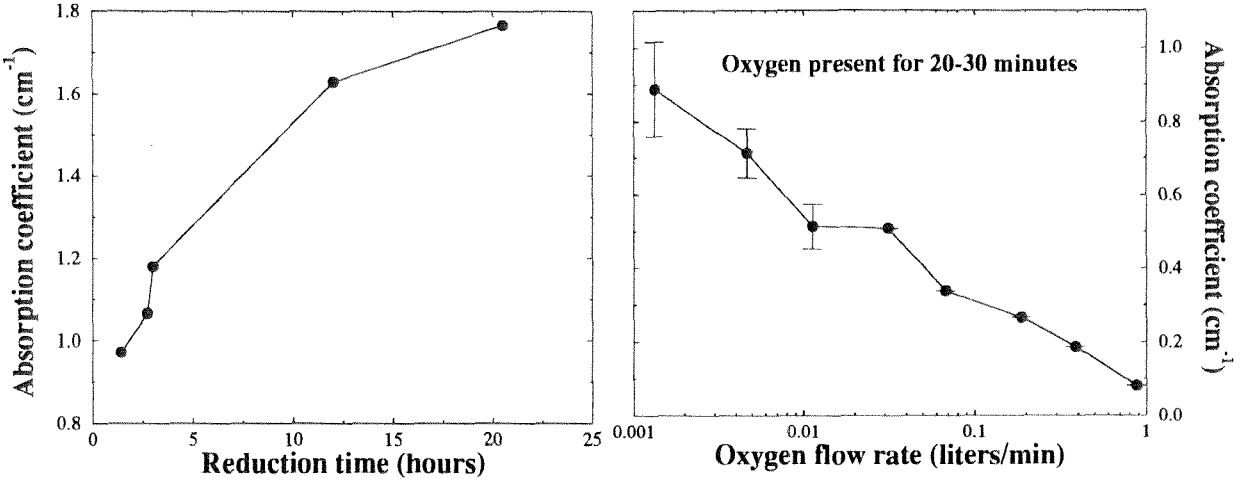


Figure 3.3: Absorption coefficient vs. control parameters

measurement of the absorption coefficient and $M/\#$. The type of control we needed over the oxidation/reduction process was over the direction of a relative change in absorption coefficient, and magnitude of that change. Development of a correspondence between absolute absorption coefficient and the annealing process was of secondary interest. As a result, we did not make any attempt to return to a known state before performing each anneal—we moved from one absorption coefficient to the next in order to fill in the gaps in our $M/\#$ measurement.

In this section, we try to use the data we have for absorption coefficients and the anneal time and/or oxygen partial pressure which we used to obtain them. Figure 3.3 shows the induced absorption coefficients as a function of the control parameters. The left–hand plot shows reduction, as a function of time, while the right–hand plot shows oxidation, with oxygen flow rate as the control parameter. Each oxidation run shown in this plot lasted 20 minutes.[10] Figure 3.4 shows photographs of the crystals at different oxidation states: the top photograph corresponds to a highly reduced state, while the bottom is the same two crystals after full oxidation (Yes, they really are there but you might have to look closely!).

Since we used time as a control variable, there was some question about possible

[10]the oxygen was on for this time, even though warm–up and cool–down took longer.

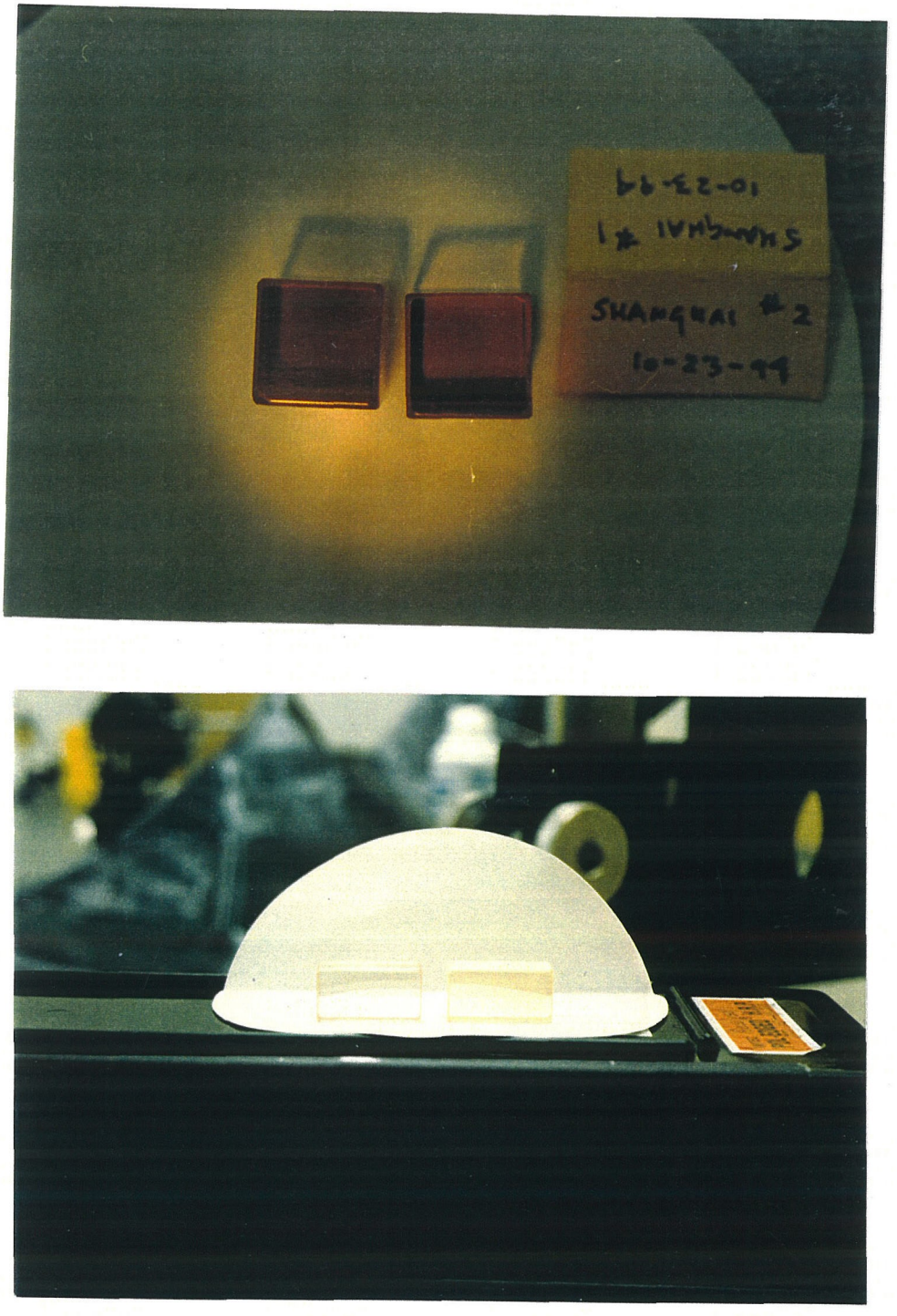


Figure 3.4: Photographs of two crystals when (a) highly reduced, and (b) fully oxidized.

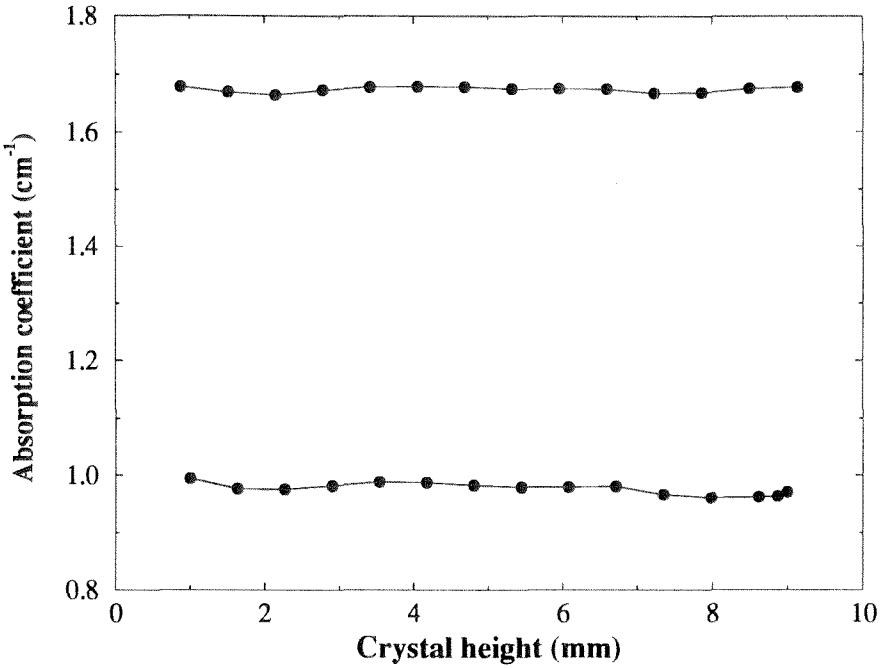


Figure 3.5: Measured spatial absorption profile: (a) after reduction, (b) after oxidation

spatial variations in absorption coefficient along the diffusion direction. In our case, we might expect a slow variation along the vertical axis of the crystal—perpendicular to the c–axis and spanning the two unpolished faces of high surface area. We did not visually observe any significant spatial variation in absorption coefficient after any of the annealing steps. To be absolutely certain, we measured the absorption coefficient along the vertical axis of the crystal after one reduction step, and after one oxidation step. These absorption profiles, shown in Figure 3.5, were measured with an illuminating spot of transverse size 1 mm.

After reduction treatments longer than 12 hours, crystals exhibited an irreversible absorption increase. Similar "useless" absorption was observed in Reference and attributed to reduction from Nb^{5+} to Nb^{4+}. We observed this additional absorption when the crystals were subsequently fully oxidized—the resulting absorption coefficient was much larger than when previously fully oxidized. In Figure 3.6, we show two absorption plots for full oxidation of the same crystal, before and after a long reduction. In addition, before the crystal was overreduced, the match between our

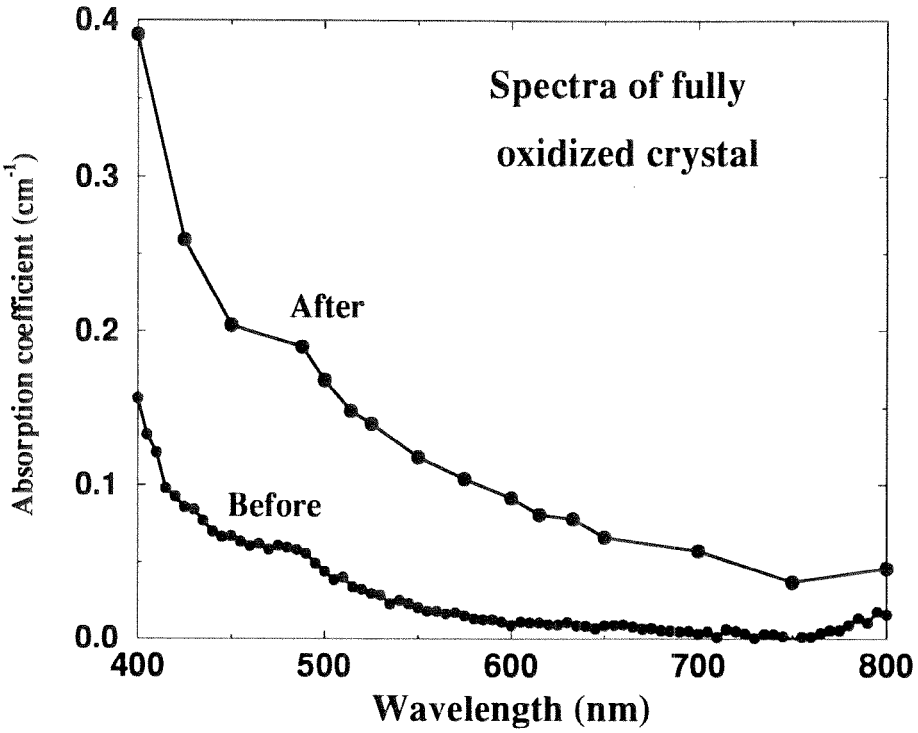


Figure 3.6: Absorption spectra of the same crystal, demonstrating useless absorption induced by long reduction anneals.

$M/\#$ theory (Chapter 4) and the experimental data was quite good, without any offset in absorption coefficient. After being overreduced, the match between experiment and theory was still excellent, but only after offsetting the theoretical plot by a small amount of useless absorption (See Figure 4.6).

3.1.2 Absorption measurements

After each annealing step, we measured the absorption coefficient in incoherent ordinarily polarized light. We used a broadband white light source *(Dolan–Jenner Industries FIBER–LITE Model 180)* and a *Monochrometer Industries 1681A* monochrometer. For each measurement, the background spectrum of the light source, and the spectrum with the crystal present were measured. The detector was a UDT Model 370. For stability, the light source was turned on an hour before any measurements were taken. For some measurements, we used a computer controlled stepper motor to turn the crank of the monochrometer. Often, however, the computer was not

available, so we took the absorption measurements by hand at a few selected wavelengths.[11]

Since the crystals we were annealing were not AR coated,[12] we needed to compensate for Fresnel reflections in the measured absorption data. We assumed that the before–and–after ratio of powers was composed of multiple contributions from internal reflection at near normal incidence. We write the measured ratio Z as

$$\begin{aligned} Z &= AT^2\left(1 + A^2R^2 + A^4R^4 + \cdots\right) \\ &= \frac{AT^2}{1 - A^2R^2}, \end{aligned} \tag{3.1}$$

where A is the loss due to absorption over one transit of the crystal, and R and T are the Fresnel reflection and transmission coefficients for normal incidence. When we solve for the absorption coefficient, we obtain

$$A \equiv e^{-\alpha L} = \frac{1}{2}\left(-\frac{T^2}{R^2Z} + \sqrt{\frac{T^2}{R^2Z} + \frac{4}{R^2}}\right), \tag{3.2}$$

where L is the length of the crystal (1.91cm for the crystal used in our $M/\#$ experiment), and α is the desired absorption coefficient. We obtained the Fresnel coefficients from the ordinary index of refraction for $LiNbO_3$[13] and the formulas $R = ((n_o - 1)/(n_o + 1))^2$ and $R + T = 1$.

The absorption spectra obtained during our $M/\#$ experiment are shown in Figure 3.7. Note that the ratio of the absorption coefficient at 633nm to the absorption coefficient at 488nm is relatively constant. This may be of interest to us when we discuss the two–λ architecture in Chapter 7.4.1.

[11] 400, 450, 488, 500, 514, 525, 550, 600, 615, 633, 650, 700, 750, and 800nm.

[12] Not many thin film coatings can survive 950°C!

[13] We used a spline fit to interpolate the index of refraction data from Reference [339].

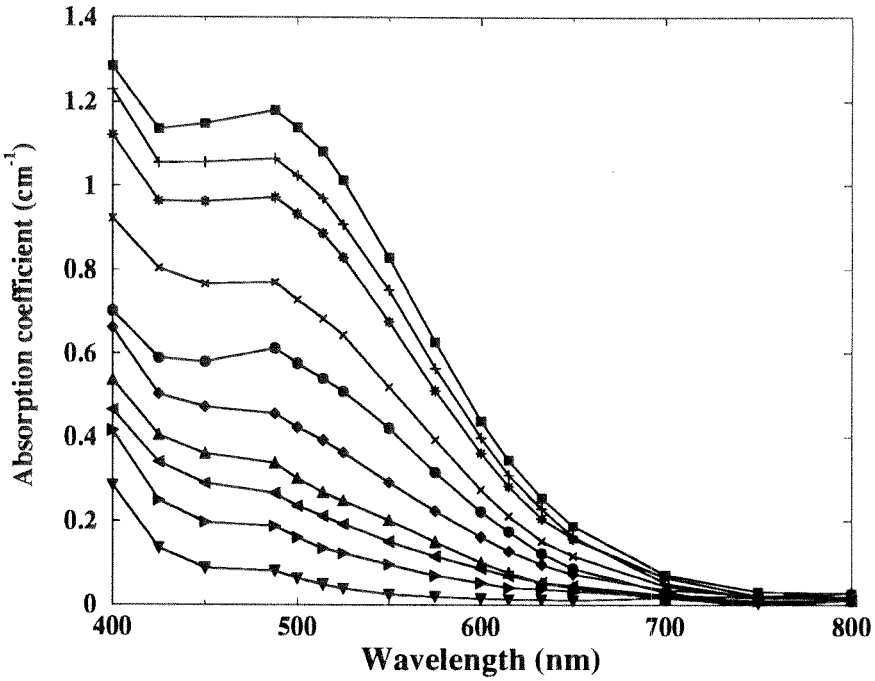


Figure 3.7: Absorption spectra of various oxidation states induced in our test crystal during oxidation/reduction experiment. The crystal is described at the beginning of the chapter.

3.2 Experimental measurements of M/#

For each oxidation state induced by the high–temperature annealing process described in the previous section, we obtained several measurements of the $M/\#$ from single hologram recording/erasure. In addition, we measured the $M/\#$ of a "control" crystal,[14] in order to gauge any fluctuations caused by the test setup. Data was only discarded when an mistake was made in the experimental setup (inadvertent closing of an iris, or burnout of a shutter solenoid, for instance) or after a significant mode–hop of our Ar laser. Between separate $M/\#$ measurements of the same oxidation state, we thermally erased the crystal at 200° for one hour.

In this section, we describe the optical setup of our $M/\#$ measurement and the experimental procedure. We then show the measured $M/\#$ and erasure time constant of our test crystal. In order to gauge the accuracy of our experimental measurements,

[14]An as–grown 1.5cm × 2cm × 1cm crystal, doped 0.01% Fe, from Deltronics Crystal. The absorption coefficient of this crystal was 0.17cm^{-1} during the entire experiment.

we show the error bars on our experiment as well as the measured $M/\#$ of our control crystal during the same period of time. Since we tried a number of optical setups which did not give repeatable results, we describe some of the things which were discarded during our preliminary experiments. In the final part of the chapter, we describe how we adapted our measurement setup for multiple hologram recording and compare single–hologram $M/\#$ measurements to the diffraction efficiency of multiple holograms.

3.2.1 Experimental apparatus

The experimental setup used for our $M/\#$ measurements is shown in Figure 3.8. The light source was a Coherent Innova 300 argon laser, used at 488nm with an intracavity etalon. Since the internal power monitor was affected by backreflections from our optics, the laser stabilization routines provided by the laser controller were turned off. Stabilization was performed by software monitoring of the power level reported by the laser. The laser beam was collimated (at lower left in the figure) and split into three coplanar beams: Read, Write, and Signal. The Signal and Write beams typically contained 62mW/cm^2, while the Read beam was much weaker (1.17mW/cm^2). All beams were collimated and vertically polarized at the crystal. The Signal beam was expanded to overfill the square aperture (marked **W**) before illuminating the crystal end face. The read and write beams were routed to a mirror mounted on an accurate rotation stage (Newport 495 rotation stage with a PMC-200P controller). The angle between the two beams was approximately 13.6° and the mirror could be turned to direct either of the Write and Read beams down the optical axis of the 4–F system shown.[15] This 4–F system had a magnification ratio of 10/3, so that the smallest angular change possible with the mirror rotation stage would be much smaller than the angular selectivity of the grating.

An aperture (marked **H**) was placed in front of the crystal to constrain the height

[15]The other beam, deflected outside the first lens aperture, was blocked from continuing. An additional detector, not shown in the figure, was positioned to monitor the power level of the Read beam when the mirror was at the angle needed to deflect the Write beam.

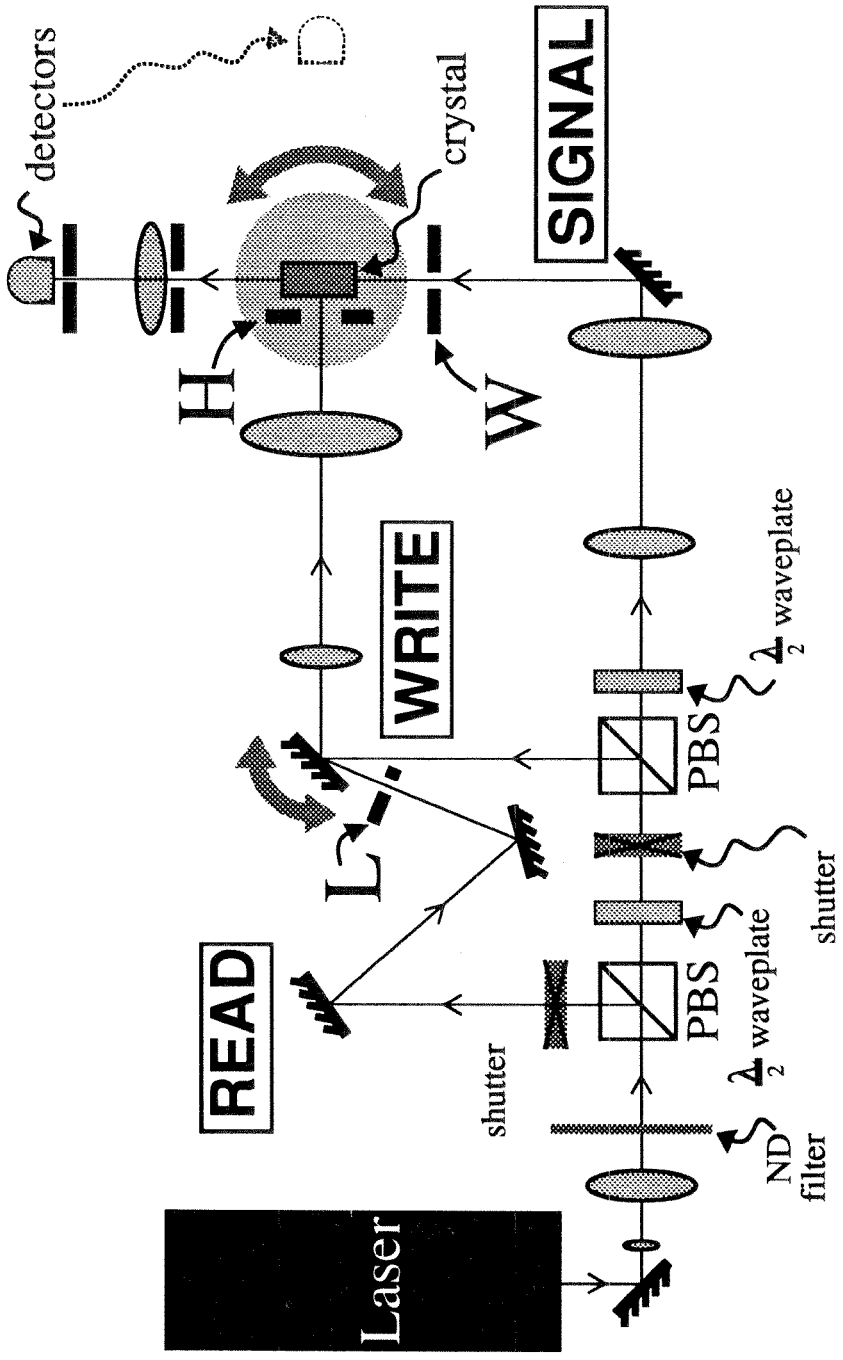


Figure 3.8: $M/\#$ measurement setup

and width of the Write beam. The horizontal aperture of the Read beam was controlled by the aperture marked **L**, placed close to the rotating mirror so that it was nearly imaged onto the crystal surface. The width, height, and length of the grating volume used for readout were then set by the apertures marked **W**, **H**, and **L**. The advantage of this setup was that no aperture had to be exactly imaged onto another. The width of the recorded grating, as set by the horizontal extent of the **H** aperture, was slightly larger than the portion used for reading, The height of the Signal beam was made larger than the height of the Write beam, so the overlap of these two beams was guaranteed. The only alignment question was the collinearity of the read and write beams. This was set by placing a beamsplitter on the crystal stage and obtaining interference fringes between the Signal beam and each of the Read and Write beams (when deflected down the center of the 4–F system). Once set, none of these optics were disturbed during the $M/\#$ measurements. The size of the grating volume used was $L = z_2 - z_1 = 0.83$cm, $W = x_2 - x_1 = 0.46$cm, and $\mathbf{H} = .25$cm.

The readout optics consisted of two irises, a focusing lens, and a UDT Model 370 detector placed just beyond the focal plane. Since one of the crystals had a larger wedge between the signal faces than the other, the irises were repositioned and resized between each measurement. The readout optics were covered to avoid any background light, and electrical tape was used to block scattered light from the crystal edges or the rest of the optics. The entire apparatus was surrounded and covered by cardboard to avoid air currents, and placed on an optical table to avoid mechanical vibration. The repeatability of the recording measurements showed the success of this approach.

3.2.2 Experimental procedure

Before each measurement, the beam powers were measured using the two detectors shown in Figure 3.8. Since the Write and Signal beam powers were over the saturation limit of our detectors, an ND–20 filter was placed in front of the first beamsplitter. The weak Read beam was measured with and without the ND filter each time, and the

ratio used to calculate the intensities of the Write and Signal beams. It is possible that the ND filter rotated the polarization slightly, affecting the accuracy of our power and external modulation depth measurements. However, our primary goal was to obtain the same experimental conditions for each measurement. For this, we used a rotating post assembly to ensure that the same section of the ND filter was used each time.

The crystal was placed on a low–resolution rotation stage (New Focus Picomotor 8401) on an orthogonal pair of manual translation stages. Repeatable positioning of the crystals was obtained through the following procedure:

- The crystal was placed on the stage with the alignment mark[16] in the chosen direction.
- The corner of the crystal was aligned with the edges of the Read and Signal beams. This was made easier by positioning the crystal so that a thin sliver of light passed by the edge of the crystal, and then translating it with the linear stages by a set amount (0.4mm in this case).
- The crystal was rotated until normal to the Signal beam by observing its back-reflection.
- The linear stages were used to translate the crystal by 2mm so the Signal and Read beams no longer intersected the edges of the crystal.
- The crystal stage was rotated clockwise by 2.5° degrees by computer control.

Several holograms were stored during each experiment (that is, between each thermal erasure step). Typically, 5 holograms were recorded to measure the writing slope, and 3 holograms for measuring erasure. Between each hologram, the crystal stage was rotated counter-clockwise by approximately 1.35°. Over the course of these measurements, the grating vector moved inside the crystal a total of 2.3° with respect to the c–axis. Recording measurements were made by measuring the diffraction efficiency before and after an exposure of 4 seconds. A typical diffraction efficiency varied between 6×10^{-6} and 6×10^{-4}. Erasure measurements were performed on holograms

[16]Delineating the direction of the c–axis.

recorded with identical exposures. The hologram sat in the dark for 50 seconds, and the crystal stage rotated by 0.36° before starting the erasure measurements. Between each diffraction efficiency measurement, the Signal and Write beams were turned on for 50 seconds. The Write beam was deflected by 3×10^{-4} ° every second to avoid building up a strong grating during erasure.

The diffraction efficiency was measured with the weak Read beam and with the auto–ranging feature of the UDT power meter disabled. Three scanning routines of varying angular spacing were used in order to find the selectivity peak. The coarsest scanning routine spanned an external angle of .06° at the crystal[17] in 40 steps. We used this very coarse routine for the first measurement of the recording behavior and the first measurement of the erasure behavior.[18] For each point in the measurement, 10 measurements of the UDT analog output signal were made with an A/D converter. The measurements were averaged and converted to power with a previously calibrated linear relationship. Between each measurement in a scan, the Newport controller was jogged by the appropriate angle step. A scan of 20 points took <4 seconds. If the Bragg condition was determined to be outside or near the edge of the scan range, the mirror was repositioned and the scan repeated. Likewise, if the peak A/D voltage fell outside an acceptable range, the computer reset the dynamic range setting of the UDT power meter accordingly and rescanned the Bragg peak. Once each scan found the peak, the computer repositioned the mirror for the next finer scan. The regular coarse scan consisted of 20 steps spanning .018°; the fine–level scan of 20 steps spanning .006°. Three consecutive fine–level scans were averaged to obtain the diffracted power—the standard deviation among these measurements was typically 1–3% of the average.

The results of a typical measurement are shown in Figure 3.9. In Figure 3.9a, we show the measured A_0/τ_r, obtained by dividing the square root of the measured diffraction efficiency by the exposure time. The standard deviation among these measurements was typically < 1%. Figure 3.9b shows the exponential decay of the

[17]for comparison, the measured angular selectivity of the hologram was ~ .01°.

[18]Since the crystal was rotated slightly, the readout beam needed to be deflected accordingly.

holograms for three consecutive holograms on a log scale, and Figure 3.9c shows the exponential time constant τ_e obtained by linear regression of Figure 3.9b. Since the time constant varied by a factor of 2 during the experiment, the regression was always performed with the data points between zero and one time constant. For instance, for the data shown we used 8 points for regression; when the time constant was ~ 240 seconds, we used 6 points. The standard deviation of the measured time constants in the same run was 2–3%. Figure 3.9d shows the decay in A_0/τ_r observed during the erasure experiments. We attribute this decay to the buildup of a compensating electric field by the photovoltaic effect. Because of the uncertainty of the effect of the photovoltaic field on erasure, we used only the first erasure time to calculate the $M/\#$.

Since the Write and Signal beams were of equal intensity, the recorded grating had a large modulation depth. This becomes an issue in particular when we want to compare the data to a theoretical prediction based on a linearization of the Kukhtarev model (Chapter 4.5). Typically, if the modulation depth is too high, then higher orders (harmonics of the spatial modulation) must be included in the analysis. Fortunately for us, these higher orders develop more slowly than the first–order grating. Figure 3.10 shows that the measured recording slope (A_0/τ_r) was a linear function of the external modulation depth.[19] This result shows that higher order Fourier components did not become significant during the short exposure times used (typically $< 0.02\ \tau_e^{ext}$), even at high modulation depth.

Figure 3.11(a) shows the measured erasure time constant (τ_e) as a function of absorption coefficient. This plot satisfies our intuitive expectation that increased absorption will speed up the holographic storage and erasure processes. In Figure 3.11(b), we show the measured $M/\#$.[20] As expected, there is an optimal absorption coefficient, which occurs for this crystal and configuration near 0.5cm^{-1}. We show the measured $M/\#$ in the control crystal during the same period of time (Figure 3.12).

[19]We were able to control the modulation depth in our setup by rotating the half–wave plate in the Signal beam, shown in Figure 3.8. This changed the value of $\mathbf{S} \cdot \mathbf{R}$ in the modulation depth without changing the total intensity $(S^2 + R^2)$.

[20]obtained by multiplying the measured values of (A_0/τ_r) and τ_e.

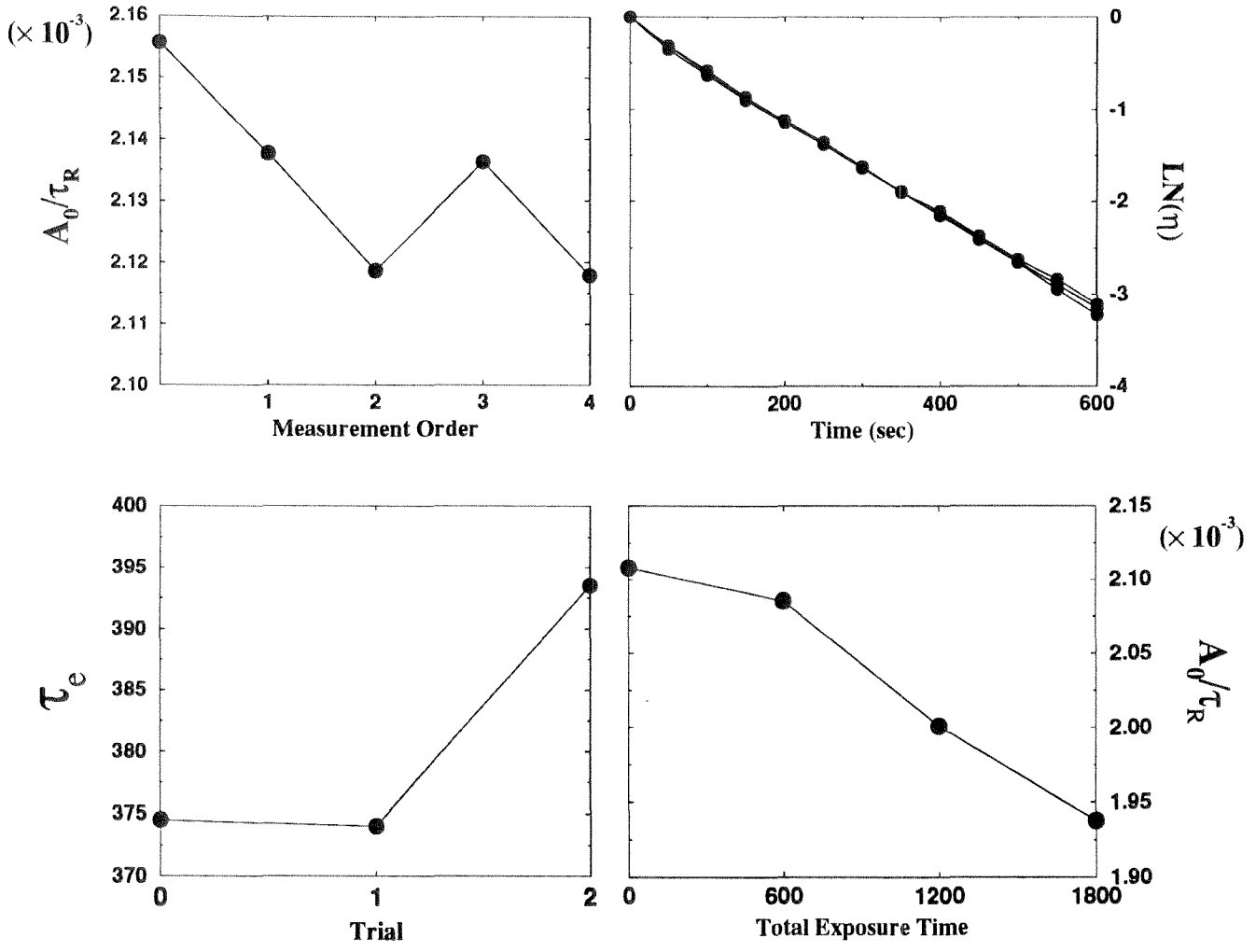


Figure 3.9: Typical measured data: (a) (A_0/τ_r) before erasure measurements, (b) erasure curves, (c) exponential erasure time, and (d) (A_0/τ_r) during erasure measurements.

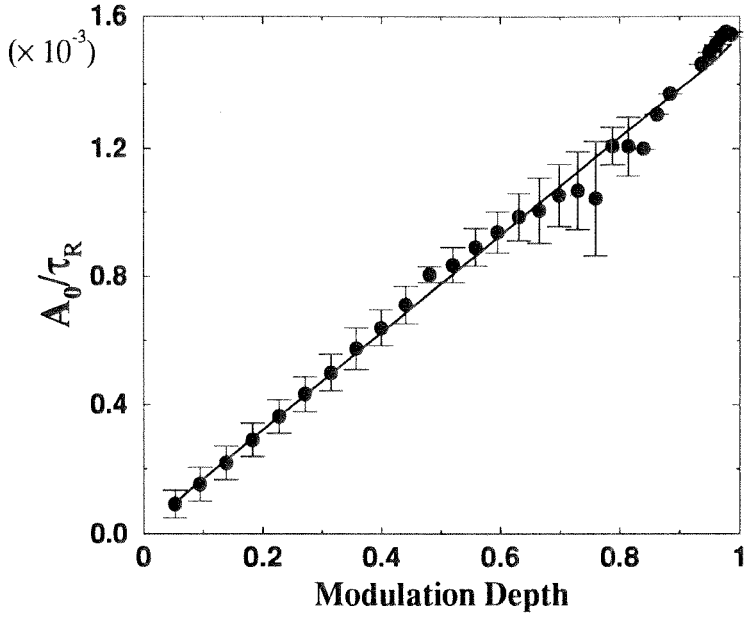


Figure 3.10: (A_0/τ_r), measured versus external modulation depth

This plot and the size of the error bars in Figure 3.11 show that our $M/\#$ measurement apparatus achieved acceptable repeatability. We have annealed three other 90° geometry 0.01% Fe–doped crystals over the same range of absorptions and seen similar behavior in time constant and M/# from each.

3.2.3 What was changed to achieve repeatability

In this section, we briefly list some of the things we tried in our $M/\#$ setup which did not give repeatable results, and why they did not work.

- *Fixed Read beam*—Initially, we did not scan the Read beam. Instead, we had a beamsplitter on the high resolution mechanical stage to combine the two beams. We would write a hologram between the Read beam and Signal beam, and scan the beamsplitter until the Write beam was Bragg–matched to it. Then we would rotate the crystal slightly, and use the Read beam to monitor holograms written between the Signal and Write beams. This scheme had two disadvantages. First, since the beamsplitter had to be non-polarizing, we wasted a lot of light. Secondly, the erasure curves varied wildly from one experiment to the next.

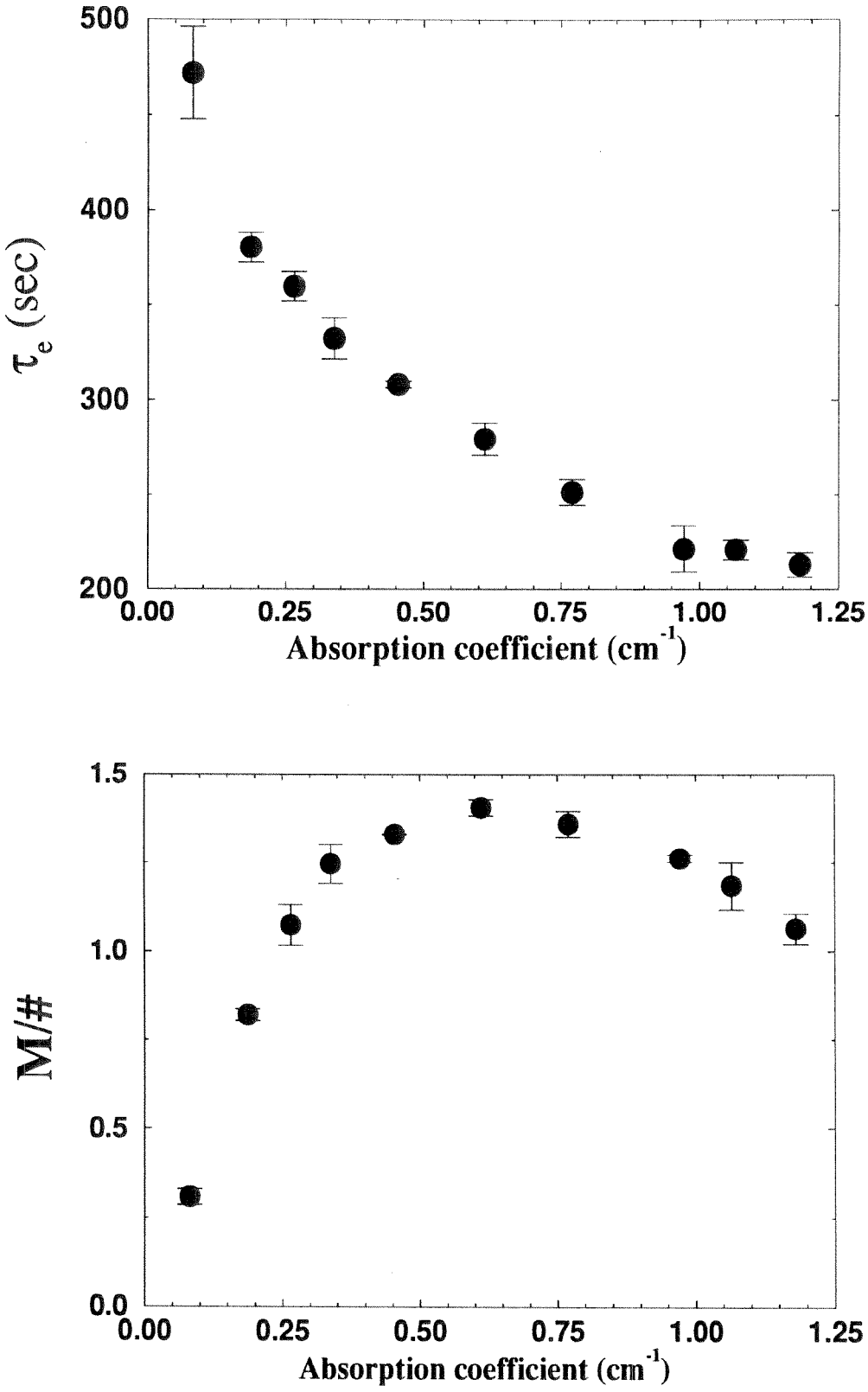


Figure 3.11: Measured $M/\#$ and τ_e^{ext} as a function of absorption coefficient

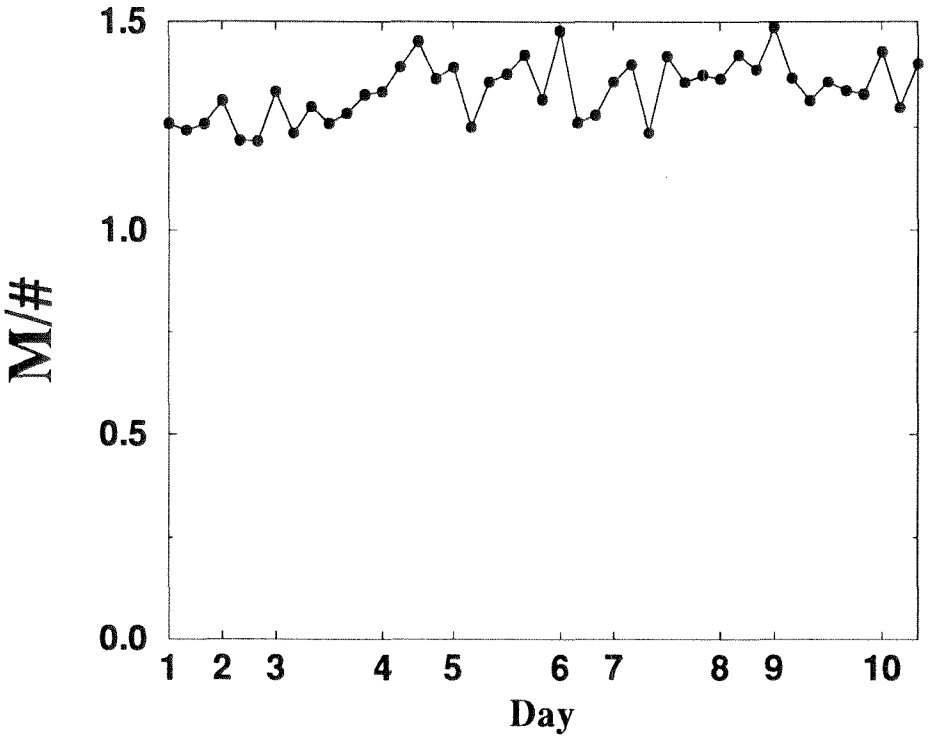


Figure 3.12: Measured $M/\#$ of the control crystal during the experiment.

We believe that this was caused by a changing Bragg–matching condition. The strong Write beam, illuminating the crystal during erasure, was building up a field across the illuminated region which changed the bulk index of refraction. This in turn caused the same readout beam to become Bragg–mismatched with the hologram through Snell's law at the crystal surface. The change to a scanning mirror solved both problems, although it complicated the measurement of the holograms as we have described above.

- *Positioning the crystal*—Our first approach was to position the crystal by observing where the beam entered the crystal face. While the recording slopes were consistent within each measurement, the variance between measurement sets (where the crystal was removed for thermal erasure) was quite large. The two linear stages and the careful repositioning procedure removed this problem.

- *Circular beams*—We originally had just one circular aperture in front of the first beamsplitter. However, since the grating volume was a strange sum of

"intersecting cylinders," it was not appropriate data to compare to our theoretical model. In addition, we worried about imperfect alignment of the two intersecting beams, and about the nonuniform Gaussian profile of the expanded beams. By saving the power that had been wasted in the non-polarizing beamsplitter, we were able to expand the beam size. Switching to the rectangular apertures (that we described above) removed our reservations about comparing the experimental data with our theoretical model.

3.2.4 Recording of multiple holograms

In this section, we explore the question: How well does the M/#, when measured from the measurement of the recording and erasure of a single hologram, describe the diffraction efficiency when recording multiple holograms? In order to compare both single hologram and multiple hologram experiments on the same setup, we modified our M/# setup for recording of multiple holograms.

As we mentioned above, the M/# setup was designed so that the minimum angle change of our mechanical deflector would be many times smaller than the angular width of the holograms. Because of this, the total usable angular range was limited to $<5°$, and only 400 reference beam angles were available (for storage at the second sinc null). However, we were able to measure the average diffraction efficiency for 1000 holograms by recording most of the holograms at a spacing much smaller than this. Every 50^{th} hologram was separated from the others by more than 8 sinc nulls, giving 20 reliable diffraction efficiency measurements which spanned the recording schedule.

We stored holograms in the control crystal from the single–hologram experiment. To construct the recording schedules for M multiple holograms, we used an erasure time constant (τ_e) of 540 seconds. This was within the range of time constants we had measured with the crystal, slightly below the average.[21]

[21]This time constant gave uniformly strong holograms. Since the recording time constant was probably decreasing slightly due to the buildup of photovoltaic screening field (see Figure 3.9), the slightly lower time constant compensated for this.

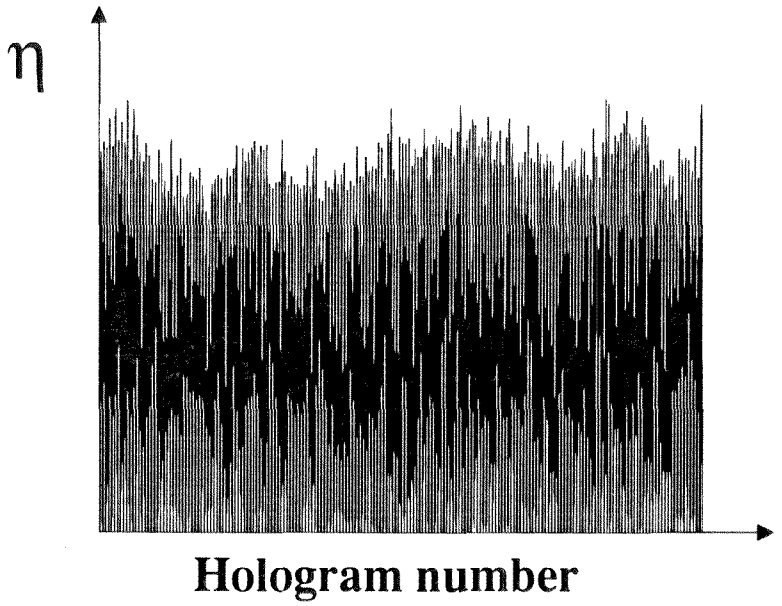


Figure 3.13: Equalized diffraction efficiency of 200 holograms.

From the discussion in Chapter 2.2.3, we know that the choice of f affects the diffraction efficiency of the holograms. As a result, when we record multiple holograms, the measured diffraction efficiencies lie on a $(C/M)^2$ curve, where C is the product of the fraction f and the $M/\#$ predicted by the single–hologram measurement. We stored 100, 200, 400, and 1000 holograms using schedules built with $f = 0.6$ and $f = 0.8$. After recording a set of multiple holograms, the diffraction efficiency was measured by sampling a subset of 20 well–distributed holograms. A sample comb function (the selectivity functions of the recorded holograms) for 200 holograms is shown in Figure 3.13. The average measured diffraction efficiency is shown as a function of the number of holograms in Figure 3.14. The data from each value of f is shown with a $1/M^2$ fit. In the table below, we compare the $M/\#$ from the single–hologram measurement with the values of C/f from the fits to the multiple hologram data.

The close agreement of these numbers shows that the straightforward single–hologram $M/\#$ measurement is indeed an accurate measure of the performance of a multiple hologram holographic memory system.

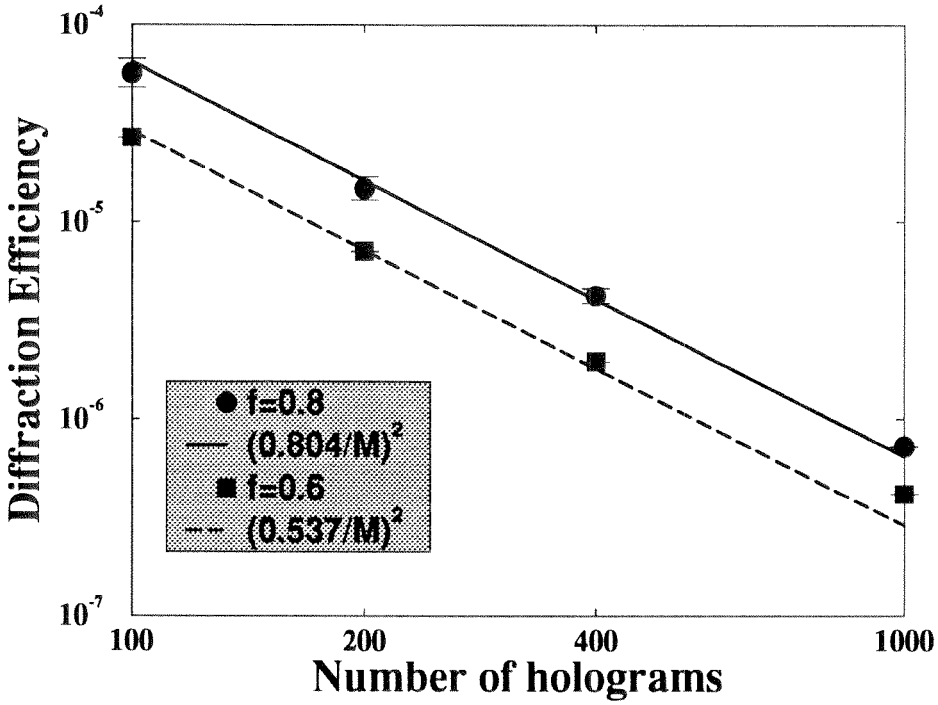


Figure 3.14: Diffraction efficiency versus number of holograms stored, for $f = 0.8$ and $f = 0.6$. Final recording time for M holograms is $t_{last} = f\tau_e/M$, and $\tau_e = 540$ sec.

Experiment		M/#
One hologram recording/erasure		$1.368 \pm .061$
Multiple holograms	$f = 0.8$	1.256
Multiple holograms	$f = 0.6$	1.492

Table 3.1: Experimentally measured M/#

3.3 Appendix: Polishing of $LiNbO_3$:Fe

Our aim in this section is to provide a written record of the polishing procedure which we developed for possible future use by members of the Psaltis group. If you are not one of this number, please feel free to turn to the beginning of the next chapter.

We developed a three-part polishing procedure to remove the surface degradations we had initially caused with our annealing procedure. We altered the annealing procedure enough that the surfaces are no longer degraded, but several crystals had been affected by this point. The polishing procedure does not approach a professional finish in most respects, but was acceptable for all of our purposes except transmission of zero–blemish high–resolution images.

The first and most difficult step in our procedure is to mount the crystal in a brass jig. We used a cylindrical jig of diameter 63mm, height 21mm, and weight 467 grams, with a rectangular hole of dimensions 15mm $\times$ 25mm cut through the center (along the cylinder height). On the bottom (working) surface of the jig were mounted 4 rectangular equal-width fragments of $LiNbO_3$, which surrounded and overlapped the edge of the hole. These fragments were mounted to the jig with low-temperature (150–200°C) wax, and lapped to make a flat surface. The crystal to be polished is then placed in the hole between these fragments so that they formed a nearly continuous flat polishing surface. Since the hole in the jig is larger than the crystal, the only support for our crystal was wax.

This makes the mounting procedure a little tricky: the jig has to be heated to melt the wax, then placed over the crystal on a flat surface. This puts the desired crystal and the fragments in the same plane. Then the fragments have to be pushed horizontally against the desired crystal while the jig is still upside down on the flat surface. We often needed several heating and cooling cycles to get the desired coplanarity between fragments and the crystal.[22]

The second step is to lapp the ensemble crystal/fragment surface with 5μm grit in DI water on a piece of flat glass. A small circular motion is sufficient with no

[22] A good test is to drag one's fingernail across the crystal/fragment junction in both directions and have it not catch.

additional pressure (beyond the weight of the jig). If the coplanarity of the surface is not good, then the areas of fresh exposure will be only in the corners and getting a uniform "grey" surface will take a long time. If this process takes too long, one can remount the crystal or move to larger grit. After using larger grit, one should lapp with the 5μm grit for at least 10 more minutes.

The second step we used is polishing with 1μm alpha alumina grit using a Strasbaugh Model 68K lapping machine and a Buehler Microcloth polishing cloth. We rotated the polishing surface at 30 revolutions/minute and set the pump to 20 (units unknown). Our jig was placed in a metal ring of inside diameter 63mm, and the ring and jig together in one of the three rings of the lapping machine.[23] Between 12 and 16 minutes of polishing was enough—any more led to overpolishing of the corners. A microscope in darkfield mode can be used to inspect the crystal surface for any scratches.

The final stage is polishing with 0.06μm Extec colloidal polishing suspension on Ecomet 3 polishing machine with another Buehler Microcloth polishing cloth. We used the same metal ring in a custom–built holder, and a overhead flask with stopcock for addition of the polishing solution. The plate was spun at a revolution rate of 80–90 revolutions/minute. From 20 to 40 minutes of polishing was enough for a good surface. As the polishing time was increased, the edges of the bordering fragments became rounded without affecting the crystal. The drawback to a long polishing time is an increase in lapping the bordering crystals when polishing the next surface. This will quickly deplete the thickness of the rectangular fragments and hasten their replacement.

The surface finishes we were able to achieve were quite free of observable blemishes. With the bordering fragment crystals, the final polishing time could be extended so that the final flatness was quite good.[24] However, we had no facility for measuring or controlling the parallelism of opposite faces. This was actually an advantage of

[23]The large ring should be removed from the polishing surface after use as it rusts very quickly.

[24]We had no quantitative way of measuring the flatness. We usually looked for any curvature in the reflected images of straight line objects. In addition, we could qualitatively observe the quality of high resolution images after transmission through the crystal.

sorts, as the image backreflection from uncoated faces is quite noticeable in hologram reconstructions if the faces are exactly parallel.

Chapter 4 $M/\#$: Theoretical model

Contents

In Chapter 3, we experimentally demonstrated that there exists an optimal oxidation state of $LiNbO_3$:Fe for 90° geometry holographic storage. But, as we mentioned in Section 2.2.4, the $M/\#$ depends on both many additional factors, including material properties and the system geometry used. We don't particularly want to repeat the experiments of Chapter 3 for each factor when we're trying to improve $M/\#$. It would be preferable to have an analytical tool which we could use to maximize the $M/\#$ for any holographic storage system.

In this chapter, we develop a theoretical model which predicts the $M/\#$ in $LiNbO_3$:Fe as a function of a number of factors, including absorption coefficient, doping level, photovoltaic field, crystal length, the volume of the grating and its location within the crystal. We compare the experimental data from Chapter 3 with the predictions of the model. Having established the validity of the model, we then use it to explore the effects of other factors (besides oxidation state) on the $M/\#$. Finally, we expand the theory to cover complications within the 90° geometry.

Note: The main purpose of this chapter is to provide enough details so that readers may convince themselves that our theoretical approach is valid. If you would like to skip the gory details, there is one main result to see:

1. The correspondence between the experimentally measured $M/\#$ from Chapter 3 and the predictions of the theoretical model, shown in Figure 4.6 on page 149. This result shows that our model has validity.

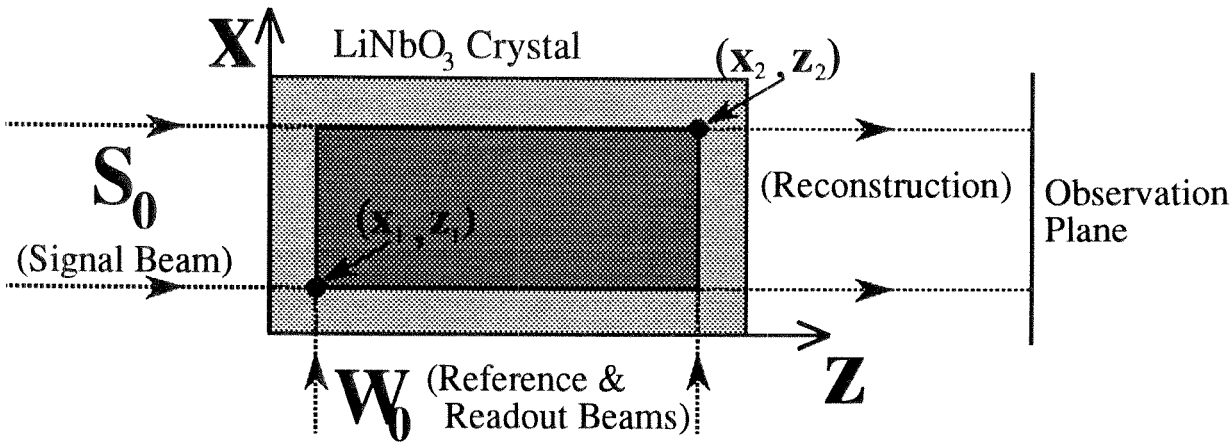


Figure 4.1: The 90° geometry

4.1 The 90° geometry

Our model is essentially the determination of the diffraction efficiency expected from a hologram in the 90° geometry (shown in Figure 4.1). Initially, we'll assume that both the signal S_0 and reference W_0 are plane waves, and that these beams are orthogonal to each other and to their respective entrance faces. The **c**–axis of the $LiNbO_3$ crystal is assumed to be at 45° to the crystal faces, and parallel to the grating vector. Since we are interested in recording multiple holograms, we can expect that each hologram will be relatively weak. For such gratings, the local permittivity modulation $\Delta\epsilon$ of a recorded hologram depends only on the local intensities $S_0(x, y, z)$ and $W_0(x, y, z)$. However, the externally observed diffraction efficiency of the hologram is **not** just the volume integral of these local index modulations. The output power diffracted by the hologram is attenuated by absorption losses. This depletion occurs to the reference beam before it reaches the local permittivity modulation, and to the diffracted light on its way out of the crystal (Figure 4.2). However, the reference beam is not considered to be depleted by the diffraction process itself, since the holograms are assumed to be weak.

The $M/\#$ was defined in Chapter 2.2.2 as the combination of two separate terms: the writing slope (A_0/τ_r), and the erasure time τ_e. Our theoretical model computes

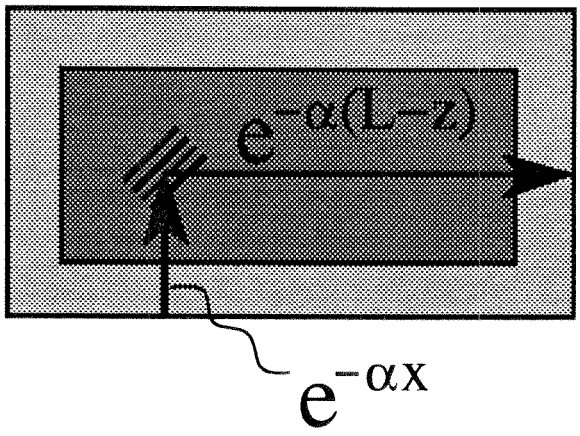


Figure 4.2: Contributions to the externally observed diffraction efficiency

the $M/\#$ by combining these two terms, each calculated separately. We make a distinction between the internal/local version of the variables and the external version. For each of these two important variables, we evaluate the local version and then integrate over the volume of the hologram, using the Born approximation for weak Bragg–matched holograms. As mentioned above, the diffraction from a local element of the grating is weighted by the additional absorption encountered during readout. After integration, we obtain the external writing slope $(A_0/\tau_r)^{ext}$ and erasure time τ_e^{ext} , corresponding to the measurements one can make in the laboratory.

4.1.1 The Born approximation

The electric field at some distance from a scattering volume can be obtained from the first–order Born approximation [13, 268], under the assumption that there is no multiple scattering. Generally this assumes that the permittivity modulation in the volume is relatively small, and is sometimes referred to as the undepleted pump approximation. If this assumption is not valid (i.e., the gratings are strong enough that multiple scattering is not negligible), then the more rigorous coupled–wave must be used. For weak gratings in isotropic or weakly anisotropic materials like $LiNbO_3$, the first–order Born approximation and the coupled wave formalism give identical results (Chapter 2 of Reference [250]).

Using the Born approximation, we can write the scattered field at a location (x, y)

$\mathbf{r}$ x, y	location at which scattered field is observed
S_x	polarization contribution
$\mathbf{r}'$	location inside the crystal
R_0	amplitude of the incident field
$\mathbf{k}_i$	wave vector of the incident field (the reference beam)
$\mathbf{k}_d$	wave vector of the diffracted field
$\Delta\epsilon_x(\mathbf{r}')$	local permittivity modulation

Table 4.1: Variables from Equation 4.1

external to the crystal [268] as

$$E_{d_x}(x,y) = S_x \iint dk_{d_x} dk_{d_y} A(\mathbf{k}_i, \mathbf{k}_d) e^{j\mathbf{k}_d \cdot \mathbf{r}}, \tag{4.1}$$

where

$$A(\mathbf{k}_i, \mathbf{k}_d) = \iiint d\mathbf{r}' \frac{\Delta\epsilon_x(\mathbf{r}') R_0 e^{j\mathbf{k}_i \cdot \mathbf{r}'} e^{-j\mathbf{k}_d \cdot \mathbf{r}'}}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}}, \tag{4.2}$$

and the other variables are defined in Table 4.1. $A(\mathbf{k}_i, \mathbf{k}_d)$ is the transfer function between the input reference beam $\mathbf{k_i}$ and the output diffracted beam $\mathbf{k_d}$. $E_{d_x}(x,y)$ is the resultant scattered field polarized in the x direction. The derivation of Equation 4.1 from the wave equation is outlined in Appendix 4.4. The equations for $E_{d_y}(x,y)$ and $E_{d_x}(x,y)$ are equivalent, with the only changed terms being S_x and $\Delta\epsilon_x$. These two terms correspond to the effect of input polarization on the scattering polarization, and on the choice of electro–optic coefficients, respectively.

There are two ways of picturing $A(\mathbf{k}_i, \mathbf{k}_d)$, which we quickly outline for future reference before moving on with Equation 4.1. First, we note that $A(k_i, k_d)$ is a transfer function between the incident and scattered wavevectors. Secondly, since Equation 4.1 is the Fourier transform of $A(k_i, k_d)$ in the k_d coordinate, an additional Fourier transform would return us to $A(\mathbf{k}_i, -\mathbf{k}_d)$. For instance, we could focus the scattered light with an ideal lens and observe the amplitude in the focal plane. In this case, k_{d_x} and k_{d_y} would correspond to transverse spatial coordinates in the focal plane

of our lens.[1] $\mathbf{k}_i$ still corresponds to the incident wavevector of the readout beam, and reflects the degree of Bragg–match to the recorded grating. In drawing this analogy, we have ignored the diffraction which arises from the finite aperture of the lens, which is valid if the lens is much larger than the exit aperture of the crystal.

Returning to Equation 4.1, we consider it in the context of our desire to calculate the output diffraction efficiency. First, we note that it does not include the absorption losses we wanted to incorporate. We should be able to include these in the integral without any loss of generality, since the scattering process is linear.[2] We include this absorptive depletion as an additional term $S_\alpha(\mathbf{r}')$.

Secondly, since we are interested in the diffraction efficiency when using a Bragg–matched reference beam, we would like to expand

$$\Delta\epsilon_x(\mathbf{r}') = \Delta\tilde{\epsilon}_x(\mathbf{r}')e^{j\mathbf{k}_s\cdot\mathbf{r}'}e^{-j\mathbf{k}_j\cdot\mathbf{r}'}, \tag{4.3}$$

where $\mathbf{k}_s$ and $\mathbf{k}_j$ are the wavevectors of the signal and reference beams, respectively, used for recording. Then we can cancel all the phase terms by setting $\mathbf{k}_i = \mathbf{k}_j$ and $\mathbf{k}_d = \mathbf{k}_s$. The first of these corresponds to the Bragg–matching condition; the second means that the integral now represents only the light diffracted in the direction of the original signal plane wave. It is important to note that this integral is *not* a direct measure of the amount of light scattered by the grating, an important point which we will need to address before comparing the results of the integral to experimental measurements.

In addition, because we are considering a normally incident signal beam, we can set

$$k_{d_x} = k_{d_y} = 0, \tag{4.4}$$

and because we are using the 90° geometry, we can assume that both the incident and diffracted fields will be vertically polarized. (Otherwise, the polarizations of the

[1] Recall that k_{d_z} is known once both k_{d_x} and k_{d_y} are given.

[2] In other words, the scattering efficiency is independent of the amplitude of the incident field.

incident and scattered fields would be orthogonal). In this case, we are solving for $E_{d_y}(x,y)$ using

$$S_x = -\omega^2 \mu_0 = -\frac{k_0^2}{\epsilon_0}, \tag{4.5}$$

where ω is the frequency of oscillation, μ_0 is the vacuum permeability, ϵ_0 the vacuum permittivity, and k_0 the magnitude of the wavevector in vacuum [268]. In a photorefractive crystal, we write the permittivity modulation in terms of the photogenerated space–charge field, as

$$\Delta\epsilon(r') = \epsilon_0 \left[-n_o^2 \Delta(\frac{1}{n^2})_{x,y,z} n_o^2 \right] = -\epsilon_0 \left(n_o^4 \, r_{13} \, E_1(x,y,z) \right), \tag{4.6}$$

where n_o is the ordinary index of refraction for $LiNbO_3$, r_{13} is the appropriate electro–optic coefficient [Yariv/Yeh], and $E_1(x,y,z)$ is the magnitude of the photorefractive space–charge field. Note again that the space–charge field is assumed to be parallel to the **c**–axis of the crystal.

When we incorporate all of these substitutions into our integral, we get

$$\begin{aligned} E_{d_x}(x,y) &= \frac{k_0^2 n_0^4 r_{13}}{2jk} e^{jkz} R_0 \iiint S_\alpha(x,y,z) E_1(x,y,z) dx dy dz \\ &= \frac{k_0}{2} n_0^3 r_{13} R_0 \iiint S_\alpha(x,y,z) E_1(x,y,z) dx dy dz, \end{aligned} \tag{4.7}$$

where $k = k_0 n_o$ is the magnitude of the wavevector inside the crystal. As expected, the scattered field has no dependence on x or y since it is a plane wave traveling along the z dimension. We have dropped the phase term $(-je^{jkz})$ for simplicity.

At this point, we have adapted the Born approximation for our purposes, and have a solution for the diffracted field in terms of the amplitude of the incident field, the local space–charge field, and the bulk absorption coefficient. To complete our model, we still need to

- convert from scattered field to diffraction efficiency,
- solve for the local space–charge field $E_1(x,y,z)$ using the Kukhtarev equations,

- and use the resulting equation to compute $(A_0/\tau_r)^{ext}$ and τ_e^{ext} separately.

4.1.2 From scattered field to diffraction efficiency

There are two factors which we consider here. The first is the loss due to Fresnel reflection coefficients when the readout beam enters the crystal, and when the diffracted beam leaves the crystal. Since both beams are normally incident, the reflection coefficient R is simply [2]:

$$R = \left(\frac{n_o - 1}{n_o + 1}\right)^2. \tag{4.8}$$

This is the fraction of power reflected, so that the fraction of the amplitude which is transmitted through the interface is $\sqrt{1-R}$. Since Fresnel reflection occurs at both entrance and exit, the total surviving fraction is $1 - R$.

The second factor we consider is the conversion of the scattered field to output power. This is made complicated because, as we mentioned above, the integral in Equation 4.7 corresponds only to light diffracted in exactly the same direction as the original signal plane wave. In other words, the field at any point (x, y) is not only the $E_{d_x}(x, y)$ which we calculate from Equation 4.7.[3] Instead, it is the coherent sum of a spectrum of spatial frequencies—each corresponding to Equation 4.7 with a slightly different $\mathbf{k}_d$. However, we know the shape of the spatial frequency spectrum of the diffracted light since we know the transverse aperture of the grating. At this point, we will assume that the spatial frequency spectrum is the 2–D Fourier transform of this transverse aperture. This is equivalent to assuming that the diffracted wavefront is nearly spatially uniform across the aperture. We will check the validity of this assumption in Section 4.3.1.

At this point, we constrain the crystal to be rectangular, with dimensions of length L (in the z direction), width W (in the x direction), and height H. The

[3]To avoid confusion, I will subsequently write the result of Equation 4.7 as $E_{d_x}|_{k_d=0}$.

spatial frequency spectrum is then

$$A_{out}\,\mathrm{sinc}(uW)\,\mathrm{sinc}(vH), \tag{4.9}$$

where u, v are the spatial frequency variables and the sinc function is defined as

$$\mathrm{sinc}(t) \equiv \frac{\sin(\pi t)}{\pi t}$$

is the Fourier transform of the "rect," or top–hat, function. Note that A_{out} is not the scattered field $E_{d_x}(x, y)$, but is instead its Fourier transform: the value of the $A(\mathbf{k}_i, \mathbf{k}_d)$ for $\mathbf{k}_i = \mathbf{k}_j = 0$ and $\mathbf{k}_d = \mathbf{k}_s = 0$. When we put a detector at the exit face of the crystal and measure the output power, we get

$$P_{measured} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[A_{out}\,\mathrm{sinc}(uW)\,\mathrm{sinc}(vH)\right]^2\,du\,dv. \tag{4.10}$$

We can assume that most of the spatial frequencies are represented if the detector is larger than the exit face of the crystal and located not very far away. In practice, the spatial frequency spread is so small that the diffracted light can be collected on the far side of a lens with no significant loss of higher spatial frequencies.

We can use Parseval's relation [8] to write the measured power as[4]

$$P_{out} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{E_{d_x}}{WH}\,\mathrm{rect}(\frac{x}{W})\,\mathrm{rect}(\frac{y}{H})\right]^2\,dx\,dy = \frac{|\,E_{d_x}\,|^2_{k_d=0}}{WH}. \tag{4.11}$$

This is what we were looking for: a relationship between measured power P_{out} and our integral $|\,E_{d_x}\,|_{k_d=0}$. If we need, we can write $I_{out} = |\,E_{d_x}\,|_{\forall k_d}$ in terms of $|\,E_{d_x}\,|_{k_d=0}$. The output intensity I_{out} can be written as P_{out}/WH so that

$$I_{out} \equiv |\,E_{d_x}\,|^2_{\forall k_d} = \frac{|\,E_{d_x}\,|^2_{k_d=0}}{W^2H^2}. \tag{4.12}$$

[4]We can use $I = E^2$ as opposed to $I = \frac{1}{2}\epsilon_0 c^2 E^2$ because we're only interested in the ratio of output to input power. Any constants we include at this point will cancel out when we take the ratio, so we can keep things simple and ignore them.

To get to diffraction efficiency, we represent the input power as:

$$P_{in} = R_0^2 LH. \tag{4.13}$$

As we saw in Chapter 2.2.2, the writing slope (A_0/τ_r) and erasure time τ_e are defined in terms of the square root of the diffraction efficiency. Using Equations 4.7, 4.11, and 4.13, we can now solve for $\sqrt{\eta}$ as

$$\begin{aligned}\sqrt{\eta} &\equiv \sqrt{\frac{P_{out}}{P_{in}}} = \frac{k_0 n_0^3 r_{13}}{2H\sqrt{WL}} \frac{4n_0}{(n_0+1)^2} \\ &\quad \times \iiint S_\alpha(x,y,z) E_1(x,y,z) dx dy dz \qquad (4.14) \\ &= \frac{k_0 n_0^3 r_{13}}{2\sqrt{(x_2-x_1)(z_2-z_1)}} \frac{4n_0}{(n_0+1)^2} \iint S_\alpha(x,z) E_1(x,z) dx\, dz. \qquad (4.15)\end{aligned}$$

Here we have recognized two things: first, that since both beams and the **c** axis are in the xz plane of the crystal, there is no variation of the integrand in the vertical dimension. So we can simplify the integral by dropping the y dimension. Secondly, the effects we have been describing in terms of H and L depend on the area of the grating, which is not necessarily the area of the crystal. Therefore, we replace H and L with $(x_2 - x_1)$ and $(z_2 - z_1)$, respectively, where x_1, x_2, z_1, z_2 define the region of the crystal used for recording,

At this point, we should check to see if Equation 4.15 satisfies our intuition regarding the dependence of η on the crystal dimensions. We note that, if absorption is ignored, the term $S_\alpha(x,z)$ in Equation 4.15 disappears, and $E_1(x,z)$ can be placed outside the integral. As a result,

$$\eta \propto LW. \tag{4.16}$$

We can convince ourselves that this is correct by considering identical cubic grating volumes, as shown in Figure 4.3. If there is a single cube, then a readout beam of power P gives an output power of ηP (Figure 4.3a. If we put two cubes next to each other (doubling W—as in Figure 4.3b), the same input power P is seen by each cube,

since there is no beam depletion and we are currently disregarding absorption. Each cube diffracts ηP, doubling the total diffraction efficiency. If two cubes are placed in order to double L (Figure 4.3c), then the input power to each cube is $P/2$, and each cube diffracts an amplitude of $\sqrt{\eta P/2}$. Since the gratings and incident wavefronts are identical, the diffracted plane waves from each cube are in phase and the amplitudes should be added before determining the output power. The diffracted power is then $2\eta P$, and the diffraction efficiency is proportional to the interaction length L. Finally, we try stacking two identical cubes vertically (Figure 4.3d). The input power to each cube is $P/2$, and each diffracts $\eta P/2$. However, these two cubes just double the size of the diffracted beam, and the output power is ηP. As expected, the diffraction efficiency is independent of the crystal height. And our intuition about each crystal dimension is in complete accordance with the predictions of Equation 4.15.

We now have an equation for diffraction efficiency (or its square root) in terms of the space–charge field and the bulk absorption coefficient. In the next section, we write out the integrand of Equation 4.15 in terms of the absorption coefficient, the incident amplitudes used for recording, and the photorefractive parameters of the crystal.

4.1.3 The local photorefractive grating

The photorefractive effect is the creation of an phase grating through photon absorption, charge transport, and trapping. This can occur in inorganic crystals as well as in appropriately doped polymers. If the intensity pattern is spatially modulated, then the trapped charge density will be identically modulated.[5] The resulting electric field then creates an index grating through the linear (Pockel's) electro–optic effect. The Kukhtarev equations (see Appendix 4.5) detail the operation of the photorefractive effect through

- the rate equation for absorber(donor) density,
- the continuity equation relating charge densities and current density,

[5] although there is often a phase shift between the two.

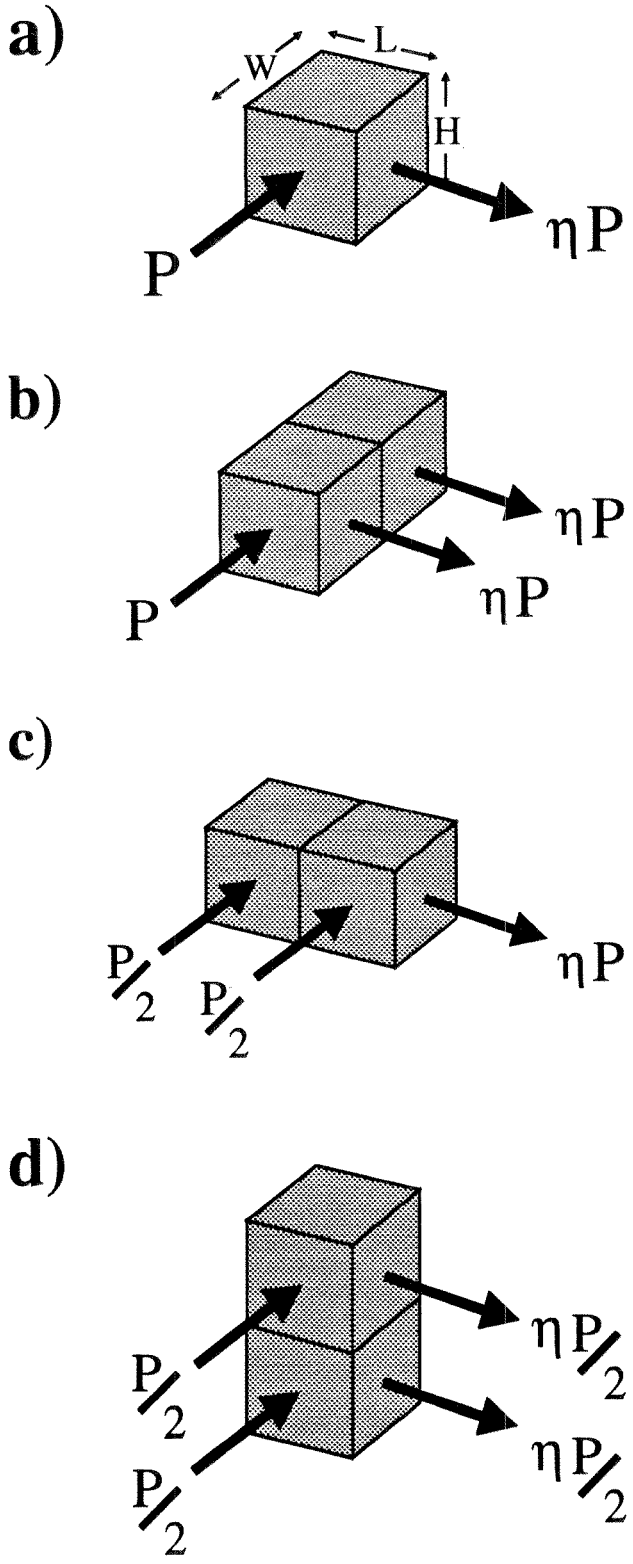


Figure 4.3: Dependence of diffraction efficiency on crystal dimensions

- the current density in terms of drift, diffusion, and a photovoltaic component,
- and Poisson's equation to relate charge density to space–charge field.

The photovoltaic effect mentioned here is a photogenerated current parallel to the c–axis of the crystal, due to preferential electron excitation in this direction. The effect is present in ferroelectric crystals such as $LiNbO_3$ because of asymmetries in the crystal lattice along the poling (**c** axis) direction. In $LiNbO_3$, this effect is quite significant and can be the dominant contribution to charge transport for some recording geometries. It is called the photovoltaic effect because a voltage can be measured across an open–circuited crystal, formed by transported charges that get trapped at the surfaces.

For our analysis, we follow the Kukhtarev equations as presented in Reference [340], with a photovoltaic current proportional to the local absorber concentration,

$$J_{ph} = p(N_D - N_D^+)I. \tag{4.17}$$

This is in contrast to earlier forms of the Kukhtarev equations which included a photovoltaic current proportional to the bulk absorption, $J_{ph} = \kappa\alpha I$.[6] This early form of the photovoltaic current cannot fully explain the large phase shifts and the spatial frequency response measured in crystals with resorting to some type of non–local photovoltaic model [341]. However, as detailed in Reference [342], incorporation of Equation 4.17 into the Kukhtarev equations resolve these difficulties.

The Kukhtarev equations can be linearized and solved for two different cases: the initial evolution of the grating, and the grating at steady–state. Linearization assumes that the modulation depth of the interference pattern is small. If this is not true, then additional Fourier components (the higher harmonics of the sinusoidal interference pattern) must be included in the analysis. In Appendix 4.5, we linearize the Kukhtarev equations using just the first Fourier component. The result for the

[6]The constant, κ, is often called the Glass coefficient [114].

Variable		Definition
N_D	total Fe doping	
N_A	initial Fe^{3+} concentration	
τ_{di}	Dielectric relaxation time	$\frac{\epsilon}{q\mu}\frac{\gamma_R N_A}{sI_0(N_D-N_A)}$
E_D	Diffusion field	k_BTK/q
E_q	Saturation space charge field	$\frac{qN_A(N_D-N_A)}{\epsilon K N_D}$
E_μ	Drift field	$\gamma_R N_A/\mu K$
E_{0ph}	Photovoltaic field	$p\gamma_R N_A/q\mu s$

Table 4.2: List of parameters used in Equations 4.19—4.22

initial evolution of the space–charge field is

$$E_1(x,z) = m(x,z)\, E_{sc}\, (1 - e^{-t/\tau_l(x,z)}\, e^{-j\omega_l(x,z)t}), \tag{4.18}$$

where

$$m(x,z) = \frac{(S_0 e^{-(\alpha/2)z})(W_0 e^{-(\alpha/2)x})}{S_0^2 e^{-\alpha z} + W_0^2 e^{-\alpha x}} = \frac{(S_0 e^{-(\alpha/2)z})(W_0 e^{-(\alpha/2)x})}{I_0(x,z)} \tag{4.19}$$

$$E_{sc} = E_q \sqrt{\frac{E_{0ph}^2 + E_D^2}{[(N_A/N_D)E_{0ph}]^2 + (E_D + E_q)^2}} \tag{4.20}$$

$$\tau_l(x,z) = \tau_{di}(x,z)\frac{1 + (E_D/E_\mu)}{1 + (E_D/E_q)} = \frac{\tau_x}{I_0(x,z)} \tag{4.21}$$

$$\omega_l(x,z) = \frac{1}{\tau_{di}(x,z)}\,\frac{N_A E_{0ph}}{N_D E_q}\,\frac{1}{1 + (E_D/E_\mu)}, \tag{4.22}$$

and the parameters N_D, N_A, τ_{di}, E_D, E_q, E_μ, and E_{0ph} are defined in Table 4.2. S_0 and W_0 are the signal and reference amplitudes, $I_0(x,z)$ is the local intensity, and α is the (intensity) absorption coefficient. In the context of our oxidation and reduction experiment, the high–temperature annealing process changes the occupancy of the N_A level through the diffusion of oxygen into or out of the crystal. This change of oxidation state affects terms containing N_A or the absorption coefficient α (proportional to $N_D - N_A$). Note that the total Fe doping N_D remains unchanged during annealing.

We now use Equations 4.19—4.22 to evaluate the local version of A_0/τ_r. In the next section, we will use the integral evaluated above to obtain the external versions of both A_0/τ_r and τ_e.

For recording exposures much shorter than τ_l, E_1 evolves locally as [7]

$$m(x,z)\frac{E_{sc}}{\tau_l(x,z)}(1+j\tau_l(x,z)\omega_l(x,z)). \tag{4.23}$$

To calculate the local contribution to the output power, the efficiency of the scattering from each point is reduced by the absorption of the readout beam before diffraction ($e^{-(\alpha/2)x}$), and of the diffracted beam after diffraction ($e^{-(\alpha/2)(L-z)}$). Thus our absorption loss term is

$$S_\alpha(x,z) = e^{-(\alpha/2)x}e^{-(\alpha/2)(L-z)}. \tag{4.24}$$

Combining this together, we arrive at[8]

$$\left(\frac{A_0}{\tau_r}\right)^{local}(x,z) \equiv \frac{S_\alpha(x,z)E_1(x,z)}{t_0} \tag{4.25}$$

$$= E_{sc}(1+j\omega_l\tau_l)\frac{m(x,z)}{\tau_l(x,z)}e^{-\frac{\alpha}{2}x}e^{-\frac{\alpha}{2}(L-z)}$$

$$= E_{sc}(1+j\omega_l\tau_l)e^{-\frac{\alpha}{2}L}\frac{S_0W_0}{\tau_x}e^{-\alpha x}. \tag{4.26}$$

There are several interesting aspects here. First, the only term which depends on x or z is the $e^{-\alpha x}$ term, so that the grating strength decays in the x direction and is constant in z. This may have an impact on the assumption we made in the previous section concerning the spatial uniformity of the diffracted beam. Secondly, note that any dependence on the total irradiance has canceled out of Equation 4.26. The recording slope depends only on the product of the two recording amplitudes (S_0, W_0), not the modulation depth m as is usually assumed.

[7] This is the time derivative of Equation 4.18.

[8] t_0 is the exposure time during which the scattering amplitude $S_\alpha E_1$ is formed.

4.1.4 The externally observed diffraction efficiency

Because we are planning on using an intensity detector outside the crystal, we are interested in the diffracted power. According to Equation 4.26, the complex time constant causes the grating to continuously shift in phase while it being recorded, an effect which cannot be observed with a square–law detector. However, both the in–phase and out–of–phase components of the grating contribute to the magnitude of the diffracted signal. As a result, the presence of the imaginary time constant increases A_0/τ_r during the initial evolution of the grating, beyond what the recording slope would be with a purely real time constant. The saturation (or steady–state) space–charge field is unaffected by the imaginary part of the time constant (see Figure 4.4). But since we are recording multiple holograms, the exposure times will always be much smaller than the time constant and the imaginary time constant improves A_0/τ_r and thus $M/\#$. So we take the absolute value of the photorefractive parameters in Equation 4.26:[9]

$$|E_{sc}(1 + j\omega_l\tau_l)| = E_q\frac{\sqrt{E_{0ph}^2 + E_D^2}}{E_q + E_D}. \tag{4.27}$$

The effective writing slope seen by an external observer can then be written by integrating the local writing slope (Equation 4.26) using Equation 4.15:

$$\begin{aligned}\left(\frac{A_0}{\tau_r}\right)^{ext} &= \frac{k_0 n_0^3 r_{13}}{2\sqrt{(z_2 - z_1)(x_2 - x_1)}}\frac{4n_0}{(n_0+1)^2}E_q\frac{\sqrt{E_{0ph}^2 + E_D^2}}{E_q + E_D}e^{-\frac{\alpha}{2}L} \\ &\quad \times \frac{S_0 W_0}{\tau_x}\int_{z_1}^{z_2}\int_{x_1}^{x_2} e^{-\alpha x}\,dx\,dz \\ &= \frac{1}{2}k_0 n_0^3 r_{13}\frac{4n_0}{(n_0+1)^2}\sqrt{\frac{z_2 - z_1}{x_2 - x_1}}E_q\frac{\sqrt{E_{0ph}^2 + E_D^2}}{E_q + E_D} \\ &\quad \times \frac{S_0 W_0}{\tau_x}e^{-\frac{\alpha}{2}L}\frac{e^{-\alpha x_1} - e^{-\alpha x_2}}{\alpha},\end{aligned} \tag{4.28}$$

where τ_x was defined in Equation 4.20.

[9] Of course, the proper thing to do is to take the absolute value of everything, including the integral. However, the $\omega_l\tau_l$ product is the only imaginary term.

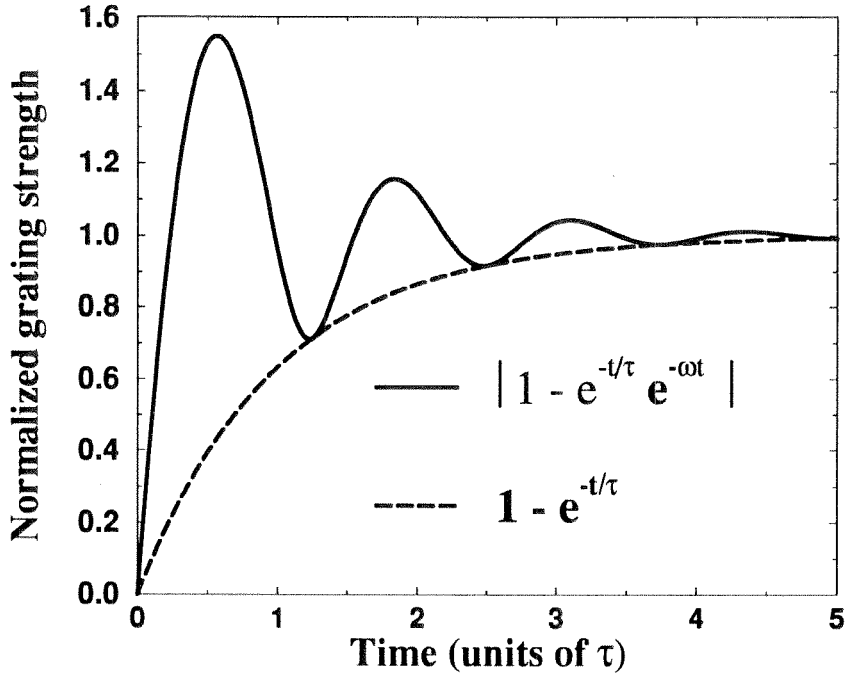


Figure 4.4: Effect of the imaginary part of the time constant on grating evolution

4.1.5 The erasure time constant

Upon erasure, either upon readout or during storage of subsequent holograms, the local permittivity evolves as $e^{-t/\tau_l(x,z)}e^{-j\omega_l(x,z)t}$. Each grating element is decaying and shifting in phase at a rate set by the local erasure intensity. As a result, the overall decay rate of the entire hologram (τ_e^{ext}) is a function of time. Intuitively, this is because the local erasure constant is smallest for the grating elements first encountered by the erasing beams. As these sections are quickly erased and contribute less to the diffracted output, the average time constant increases.[10] For moderate absorption coefficients, this effect is negligible over the duration of a standard recording schedule (often only 3-5 time constants in length). We prove this at the end of this section.

For now, we assume that τ_e^{ext} at time $t = 0$ is an accurate description of the hologram decay during the entire recording schedule. To solve for this erasure time

[10]Of course, calling it an erasure time *constant* at this point is, at best, an oxymoron.

constant, we write the evolution of the grating as

$$\sqrt{\eta}(t) \propto \int_{z_1}^{z_2}\int_{x_1}^{x_2} e^{-\alpha x} e^{-\frac{t}{\tau_x} I_0(x,z)} dx dz. \tag{4.29}$$

Here we have made the additional assumption that the effect of the imaginary part of the time constant can be neglected. The external erasure time constant is then

$$\frac{1}{\tau_e^{ext}\,|_{t=0}} \equiv \frac{\frac{-d}{dt}(\sqrt{\eta}\,|_{t=0})}{\sqrt{\eta}\,|_{t=0}} = \frac{1}{\tau_x}\left[\frac{W_0^2}{2}(e^{-\alpha x_1} + e^{-\alpha x_2}) + \frac{S_0^2}{\alpha}\frac{e^{-\alpha z_1} - e^{-\alpha z_2}}{z_2 - z_1}\right]. \tag{4.30}$$

The rest of this section is concerned with increasingly more accurate solutions of the erasure time constant. If this is not interesting to you, go ahead and skip to the next section where we solve for the $M/\#$ using the $t = 0$ approximation.

In order to obtain the time–varying erasure time constant, we can repeat the above analysis for some time $t = t_0 \neq 0$. As expected, the solution is more complicated:

$$\begin{aligned}
\frac{1}{\tau_e^{ext}\,|_{t=t_0}} &= \frac{W_0^2}{\tau_x}\frac{\int_{x_1}^{x_2} dx\; e^{-2\alpha x} e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x}}}{\int_{x_1}^{x_2} dx\; e^{-\alpha x} e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x}}} + \frac{S_0^2}{\tau_x}\frac{\int_{z_1}^{z_2} dz\; e^{-\alpha x} e^{-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z}}}{\int_{z_1}^{z_2} dz\; e^{-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z}}} \\
&= \frac{W_0^2}{\tau_x}\left(\frac{(e^{-\alpha x_2} + \frac{\tau_x}{W_0^2 t_0}) e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x_2}} - (e^{-\alpha x_1} + \frac{\tau_x}{W_0^2 t_0}) e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x_1}}}{e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x_2}} - e^{-\frac{t_0}{\tau_x} W_0^2 e^{-\alpha x_1}}}\right) \\
&\quad + \frac{1}{t_0}\left(\frac{e^{-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z_2}} - e^{-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z_1}}}{\mathrm{Ei}\left[-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z_1}\right] - \mathrm{Ei}\left[-\frac{t_0}{\tau_x} S_0^2 e^{-\alpha z_1}\right]}\right),
\end{aligned} \tag{4.31}$$

where $\mathrm{Ei}(z)$ is the exponential integral,[11]

$$\mathrm{Ei}(z) \equiv -\int_{-z}^{\infty} \frac{e^{-t}}{t} dt.$$

To show that the approximations we have made are valid, we show in Figure 4.5 the change in $M/\#$ expected if a recording schedule for 10,000 holograms is recalculated for the time–dependent erasure time constant. The decay of holograms is assumed to be non–exponential, and given at any point in time by Equation 4.32. As expected,

[11] A standard tabulated function.

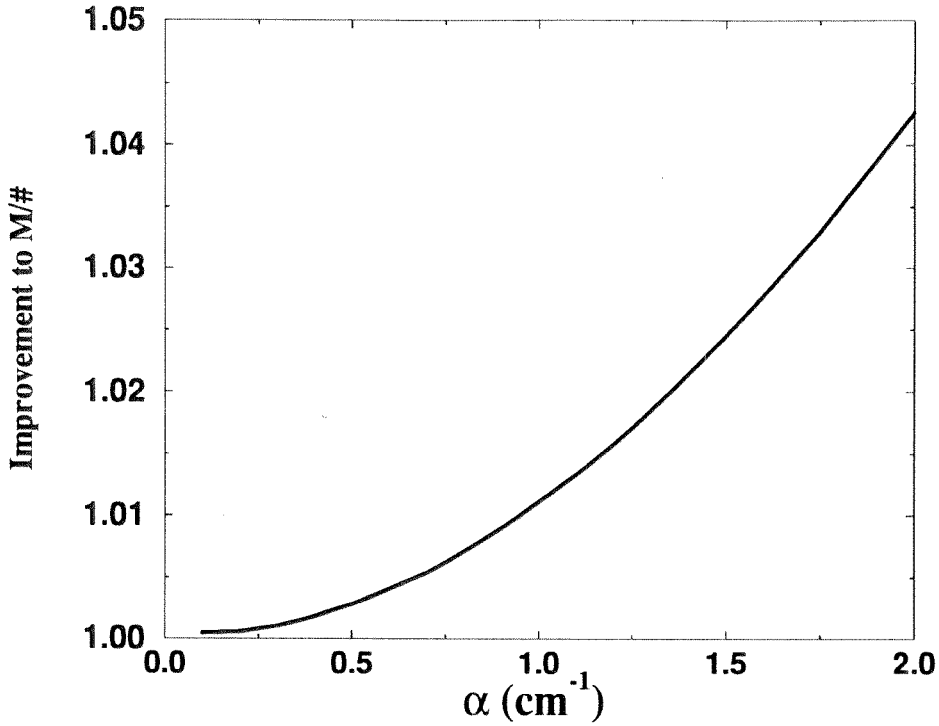


Figure 4.5: Change in $M/\#$ expected from the time–dependent erasure time constant. This shows that the approximation in Equation 4.30 is valid.

the difference is fairly unimportant even for moderately large absorption coefficients.

Unfortunately, the exact solution—where we incorporate the imaginary part of the time constant—can only be solved numerically. The equation to be numerically integrated is

$$\begin{aligned}\frac{1}{\tau_e^{ext}\,|_{t=t_0}} &= W_0^2\left(\frac{1}{\tau_x}+j\omega_x\right)\frac{\int_{x_1}^{x_2} dx\; e^{-2\alpha x}e^{-t_0\left(\frac{1}{\tau_x}+j\omega_x\right)W_0^2 e^{-\alpha x}}}{\int_{x_1}^{x_2} dx\; e^{-\alpha x}e^{-t_0\left(\frac{1}{\tau_x}+j\omega_x\right)W_0^2 e^{-\alpha x}}} \\ &\quad + S_0^2\left(\frac{1}{\tau_x}+j\omega_x\right)\frac{\int_{z_1}^{z_2} dz\; e^{-\alpha x}e^{-t_0\left(\frac{1}{\tau_x}+j\omega_x\right)S_0^2 e^{-\alpha z}}}{\int_{z_1}^{z_2} dz\; e^{-t_0\left(\frac{1}{\tau_x}+j\omega_x\right)S_0^2 e^{-\alpha z}}} \qquad (4.32)\end{aligned}$$

4.1.6 The $M/\#$

Our model for $M/\#$ is a combination of the $(A_0/\tau_r)^{ext}$ from Equation 4.28 and the erasure time constant at time $t = 0$ (Equation 4.30). The resulting analytical expression is

Variable		Definition
k_0	Wave vector	$\frac{2\pi}{\lambda}$
n_o	Ordinary refractive index	
r_{13}	Electro-optic coefficient	
S_0	Signal amplitude during recording	
W_0	Reference amplitude during recording	
E_q	Saturation space charge field	$\frac{qN_A(N_D-N_A)}{\epsilon K N_D}$
E_{0ph}	Photovoltaic field	$p\gamma_R N_A/q\mu s$
E_D	Diffusion field	$k_B T K/q$
x_1, x_2, z_1, z_2	Region of the crystal used for recording	
α	Absorption coefficient $\propto (N_D - N_A)$	
L	Length of the crystal in the z direction	
N_D	total Fe doping	
N_A	initial Fe^{3+} concentration	

Table 4.3: List of parameters used in Equation 4.33

$$M/\# = \frac{1}{2}k_0 n_0^3 r_{13} \frac{4n_0}{(n_0+1)^2}\frac{S_0}{W_0}E_q \frac{\sqrt{E_{0ph}^2+E_D^2}}{E_q+E_D}\sqrt{\frac{z_2-z_1}{x_2-x_1}} \times \frac{e^{-\frac{\alpha}{2}L}\left(e^{-\alpha x_1}-e^{-\alpha x_2}\right)}{\left(\frac{\alpha}{2}\right)\left(e^{-\alpha x_1}+e^{-\alpha x_2}\right)+\left(\frac{S_0}{W_0}\right)^2\frac{e^{-\alpha z_1}-e^{-\alpha z_2}}{z_2-z_1}}, \quad (4.33)$$

and the variables are listed again in Table 4.3 for convenience.

The two dominant terms in this equation are E_q, which increases with absorption (via $N_D - N_A$), and $e^{-(\alpha/2)L}$, which decreases with absorption. Because of these two competing terms, we expect to find an absorption coefficient which maximizes $M/\#$ for any given crystal and geometry. To predict this maximum accurately, we need the appropriate numerical values. This is the subject of the next section.

4.1.7 Numerical evaluation of the model

In this section, we assign numerical values to the variables found in our theoretical model for $M/\#$. Some are simple to find, some are difficult, and for some, the values reported in the literature cover a large range of values. We begin with the easy ones

Variable		Value	Reference
ϵ	Dielectric constant	2.6×10^{-12} F/cm	[119]
μ	Electron mobility	16 cm^2/V sec	[126]
n_o	Ordinary index of refraction	2.3489	[343]
r_{13}	Electro–optic coefficient	10×10^{-10} cm/V	[343]

Table 4.4: Material parameters of $LiNbO_3$

and move to the hard ones. We will be reporting all distances in centimeters, but otherwise we stick to the MKS system. In Table 4.4, we list the easily found and verified parameters for $LiNbO_3$ along with the reference.

The relation between the molar Fe doping level and the corresponding doping density is interesting to derive, so we do so here. Given the atomic weights of Lithium (6.941 amu), Niobium (92.91 amu), and Oxygen (16 amu), we can calculate that there are 147.851 amu per $LiNbO_3$ unit [344].[12] Given that the density of $LiNbO_3$ is 4.628×10^{-3} kg/cm^3 [343], and that there are 1.6606×10^{-27} kg per amu [344], there are 1.885×10^{22} $LiNbO_3$ units per cm^3. This is the same relation given in [118]. Since our experiments were performed with a wavelength of 488nm, the period of the grating in the material is

$$\Lambda = \frac{1}{\sqrt{2}} \frac{\lambda}{n_o} = 1.4663 \times 10^{-5} \text{cm}. \tag{4.34}$$

The relationship between absorption and $N_D - N_A$ is a crucial part of our theoretical model, since it is the link between the photorefractive behavior and the bulk absorption coefficient. Luckily, the question of what states of Fe were present in $LiNbO_3$ was a research subject in the mid–1970's, and several papers explore the relationship between the Fe^{2+} concentration and the absorption around 3.6eV. In Reference [118], this relationship is experimentally measured for extraordinary polarized light of 450nm. Our experiment is in ordinarily polarized light at 488nm, but we

[12] We can't call this a unit cell because there are actually 2 Li, 2 Nb, and 6 O_2 in the crystallographic unit cell. However, this is a story for another chapter!

can use the quoted relationship of

$$\alpha = 7.936 \times 10^{-18}\,\mathrm{cm}^2\,[N_D - N_A] \tag{4.35}$$

as our starting point.

The next question is how to relate this absorption coefficient to s, the photoexcitation cross–section. As defined in the Kukhtarev equations (see Appendix 4.5), this term relates the change of ionized donor density to the available unionized donor density and the incident intensity, or

$$\begin{aligned}
\frac{\partial N_D^+}{\partial t} &= sI(N_D - N_A) \\
\frac{\partial}{\partial t}\left(\frac{\text{Empty traps}}{\text{Volume}}\right) &= sI\left(\frac{\text{Filled traps}}{\text{Volume}}\right) \\
&= \frac{\partial}{\partial z}\left(\frac{\text{Photons}}{\text{Time}}\,\underbrace{\frac{\text{Photons captured}}{\text{Photons available}}}_{f}\right)\frac{1}{\text{Area}} \\
&= f\frac{\partial}{\partial z}\left(\frac{\text{Photons}}{\text{Time}}\right)\frac{1}{\text{Area}}.
\end{aligned} \tag{4.36}$$

Here we have extracted f, the oscillator strength of the Fe^{2+} transition. This factor is essentially the fraction of photons absorbed which result in an electron in the conduction band. Values reported for the oscillator strength of the Fe^{2+} transition range from 0.02 [342] to 0.05 [118]. Returning to the above expression, can use the definition of intensity

$$I = \underbrace{\left(\frac{\text{Energy}}{\text{photon}}\right)}_{h\nu}\frac{\text{Photons}}{\text{Time}}\frac{1}{\text{Area}}, \tag{4.37}$$

and substitute into the previous expression. This finds us with

$$f\frac{\partial}{\partial z}\left(\frac{I}{h\nu}\right) = sI(N_D - N_A), \tag{4.38}$$

σ_{ph}/I	$(N_D - N_A)/N_A$	s/γ_R	Reference
3×10^{-12} cm/V^2	2.2	5.33×10^5 sec/J cm	[345]
1–2 $\times 10^{-14}$ cm/V^2	0.01	3.72–7.4 $\times 10^5$ sec/J cm	[342]

Table 4.5: Various photoconductivity measurements

which we can rewrite as

$$\frac{\partial I}{\partial z} = \underbrace{\left[h\nu \frac{s}{f}(N_D - N_A)\right]}_{\alpha} I, \tag{4.39}$$

which we recognize as the definition of the absorption coefficient. If we use the relationship between $N_D - N_A$ and α from Equation 4.35, and the appropriate value of $h\nu$ (4.074×10^{-19} J), we find that we can solve for the ratio s/f:

$$\frac{s}{f} = 19.482 \text{ cm}^2/\text{ J}. \tag{4.40}$$

If we use the value of $f \approx 0.05$ from Reference [118], then

$$s \approx 0.97 \text{ cm}^2/\text{J}. \tag{4.41}$$

Fortunately, s does not often appear by itself in our equations, but usually appears with γ_R, the carrier recombination rate. We can use reported measurements of photoconductivity in $LiNbO_3$:Fe to get the ratio s/γ_R, since

$$\sigma_{ph} = \frac{q\mu s}{\gamma_R} \frac{N_D - N_A}{N_A} I. \tag{4.42}$$

Measurements are usually reported in terms of σ_{ph}/I as a function of $\frac{N_D - N_A}{N_A}$. Two such measurements are shown in Table 4.5, where we list the measured photoconductivity, the ratio of Fe^{2+} to Fe^{3+}, and the resulting ratio of s/γ_R. Here we have used $q = 1.6 \times 10^{-19}$ C, and the mobility value reported above ($\mu = 16$ cm^2/V sec).

The photovoltaic effect

This is a material parameter which can be measured in a number of different ways and reported in a number of different forms. Worst of all, the reported values (when you convert them all to the same representation) span a range of more than an order of magnitude. We take these topics up in order: the types of measurements, the different variables for representing the photovoltaic effect, and the range of reported values.

Since the photovoltaic effect tends to create an open–circuit voltage across the crystal, one can illuminate the crystal and measure this voltage [346]. Some researchers avoid the problem of ohmic contacts by placing two crystals in electrical contact: one is used to generate the photovoltage while the other is used as an electro–optic modulator. The measured phase–shift of the modulator is proportional to the voltage across the crystal. This result is usually reported in terms of the Glass constant, κ.

A second method for measuring the magnitude of the photovoltaic effect is the phase shift between the phase grating and the interference pattern used for recording. This phase shift can be measured with two–beam coupling. The measurements are commonly reported in terms of a photovoltaic transport length.

A third method is to apply an external DC electric field across the crystal which opposes the **c**–axis. Then one increases the applied field until the photovoltaic effect is entirely compensated. Of course, this point is difficult to determine since the index change becomes too small to measure accurately. However, one can take measurements for several field amplitudes and extrapolate to the point of exact compensation. Since electric fields larger than the effective photovoltaic field will again give measurable index changes, one can also extrapolate down from larger field values, creating a "V". This measurement is usually reported as an equivalent electric field, corresponding to the E_{0ph} value which we use in our theory.

The conversion between transport length, Glass constant, and equivalent electric field is relatively simple. First, we represent equivalent electric field in terms of p,

the photovoltaic parameter we use in our version of the Kukhtarev equations (Appendix 4.5). Then we relate transport length, l_{ph}, and Glass constant, κ, in terms of p.

As defined in Table 4.4, the (equivalent) photovoltaic field is

$$E_{0ph} \equiv \frac{p\gamma_R N_A}{q\mu s}. \tag{4.43}$$

If we use the first value from Table 4.5 for s/γ_R, we obtain E_{0ph} in terms of p as

$$E_{0ph} = \left(7.33 \times 10^{11}\mathrm{V}^2/\mathrm{cm}\right) pN_A, \tag{4.44}$$

where the units of p are [cm^3/Volt]. Note that the equivalent photovoltaic field depends on the occupancy of the Fe^{3+} level in the crystal. This means that we need to know N_A to be able to compare a measured photovoltaic field to the Glass constant or transport length. Secondly, it implies that we can increase the doping level[13] to get a larger photovoltaic field.

Now we would like to relate the other measured parameters to our photovoltaic parameter p. The Glass constant and p are both defined in terms of the photovoltaic current density,

$$J_{ph} = pI(N_D - N_A) = \kappa\alpha I. \tag{4.45}$$

This quickly leads to

$$p = \left(7.936 \times 10^{-18}\mathrm{cm}^2\right)\kappa, \tag{4.46}$$

where the units of κ are [cm/V] and we have used Equation 4.35 to convert from α to $N_D - N_A$. To obtain the photovoltaic transport length, we use the mobility and the equivalent photovoltaic field to write a "velocity". We then break this velocity

[13] assuming that we keep p and N_A/N_D constant.

Value measured	Corresponding E_{0ph} (in V/cm)	Reference
$E_{0ph} \sim 10^4 - 10^5$ V/cm for $N_A = 10^{19}\,\mathrm{cm}^{-3}$	$(10^{-15} - 10^{-14})\,N_A$	[123]
$\kappa = 3 \times 10^{-9} - 4.8 \times 10^{-9}$ cm/V	$(1.85 \times 10^{-14} - 2.95 \times 10^{-14})\,N_A$	[114]
$l_{ph} = 3\,\mathrm{nm}$	$3.6 \times 10^{-14} N_A$	[342]
$l_{ph} = 13\,\mathrm{nm} - 200\,\mathrm{nm}$	$(1.6 \times 10^{-13} - 2.4 \times 10^{-12})\,N_A$	[342,347,348]

Table 4.6: Measurements of the photovoltaic effect

into a distance and a time, as

$$\mu E_{0ph} = \underbrace{\frac{p}{qs}}_{l_{ph}} \underbrace{\gamma_R N_A}_{1/t}. \tag{4.47}$$

We can then write l_{ph} in terms of p, as

$$l_{ph} = \left(6.44 \times 10^{18}\mathrm{V/cm}^2\right) p. \tag{4.48}$$

We should take this number with a grain of salt, since our value for s is not that reliable. It should, however, be good enough for our purposes, since we just want a ballpark estimate of the measured photovoltaic field.[14]

In Table 4.6, we record some of the measurements of the photovoltaic constants which are in the literature. For each, we record the value (or values) reported and the corresponding photovoltaic field. We include the transport lengths from References [342,347,348] even though these are so large as to be almost absurd.

So there is a large range of reported values in the literature for the magnitude of the photovoltaic effect, from

$$E_{0ph} = 10^{-15} - 3 \times 10^{-14}, \tag{4.49}$$

where we are being pretty conservative on the high end. In our laboratory, we also

[14] I'm going to stop calling it the photovoltaic field, even though it's not exactly equivalent to an external applied field (See Equation 4.144).

see a large variance in the magnitude of E_{0ph},[15] especially between crystals from different manufacturers. Fortunately, the magnitude of the photovoltaic effect tends to scale our $M/\#$ theory, without changing the shape of the curve. So the absorption coefficient needed to get the maximum $M/\#$ does not change with photovoltaic constant.

Now, we move ahead and compare our measured data to the theory by using a value of E_{0ph} which brings our theoretical prediction to the same scale as the measured data points. In Section 4.2.1, we will explore the effect of E_{0ph} on our $M/\#$ theory.[16]

4.1.8 Comparison with experiment

In a perfect world, we would know precisely all of the material parameters of our particular $LiNbO_3$:Fe crystal, and comparison between experiment and theory would be a simple matter. Instead, we have three classes of parameters:

1. Values we used in our experiment. We can accurately measure these, and have a good idea of their uncertainty. These include L, x_1, x_2, z_1, z_2, S_0, W_0, λ, T, and α. In this class, we can also include universal constants such as q and k_B.

2. Values which are obtained from the manufacturer, or from the literature by measurements on other $LiNbO_3$ crystals. These are nominally a single value, but we don't have a good measure of the uncertainty.[17] These include N_D, r_{13}, n_o, ϵ, and μ.

3. Values for which the literature gives a range of values, or values which were measured under different conditions than our experiment. Examples of the former include γ_R, s, and E_{0ph}. The relationship between α and $N_D - N_A$ from Reference [118] seems solid, except that it was measured for extraordinary

[15] We do not measure E_{0ph} directly. Instead we measure $M/\#$ and backcalculate the size of E_{0ph} using the model. While one can argue that this assumes the correctness of our model, certainly sizable differences in $M/\#$ between otherwise identical crystals indicates that something strange is going on.

[16] This is where we will prove that E_{0ph} does not significantly affect the shape of the curve or the location of the optimal absorption coefficient.

[17] error bars, number of significant figures, etc.

Variable	Modeled as
k_BT/q	0.0259 V
L	1.91 cm
$x_1 = z_1$	0.2 cm
x_2	0.66 cm
z_2	1.03 cm
λ	488 nm
$\frac{S_0}{W_0}$	1.0
W_0^2	62.5 mW/cm^2
N_D	1.89×10^{18} cm^{-3}
E_D	1.11×10^4 V/cm^2
$\frac{4n_o}{(n_o+1)^2}\frac{1}{2}kn_o^3r_{13}$	6.742×10^{-4} cm^{-1}
E_q	1.41×10^{-13} V cm^2 $\frac{N_A}{N_D}(N_D - N_A)$
E_μ	1.46×10^{-7} V sec/cm $\gamma_R N_A$
τ_{di}	1.033×10^6 sec/cm^3 $\frac{\gamma_R}{s}\frac{N_A}{I_0(N_D-N_A)}$

Table 4.7: List of parameters used. Fields are in terms of [Volts/cm], concentrations in [cm^3]. Parameters are functions of α in [cm^{-1}]

polarization at 450nm. We don't know exactly what to expect in ordinarily polarized light at 488nm, except not to be surprised when the behavior is different. Another item which falls into this class is the precise amount of useless absorption which we induced by excessive reduction prior to our experiment (see Chapter 3.1.1).[18]

Our approach here is to take the values from the first and second classes as given. Then we fit the two experimental curves (one for τ_e^{ext} and one for $M/\#$) using the four values from the third class: γ_R/s, E_{0ph}, $\alpha/(N_D - N_A)$, and $\alpha_{useless}$. Once we have the best fit, we check that these four values are not outside the range of reported values. We show two tables, one for the values we use directly (Table 4.7), and one for the four values used for fitting (Table 4.8). In the latter table, we describe the effect on each curve, and list the value which gave the best overall fit to both curves.

Let's check that the fitted values are within the appropriate ranges, and then look at the plots of experiment vs. theory. The value we use for s/γ_R is satisfying close

[18]You could think of this as a non-zero intercept on the relation between α and $N_D - N_A$.

Variable	Effect on $M/\#$	Effect on τ_e^{ext}	Value used
$\frac{s}{\gamma_R}$	None	Scales vertically	5.45×10^5 sec/J cm
E_{0ph}	Scales vertically	None	1.09×10^{-14} V cm$^2 N_A$
$\alpha_{useless}$	Shifts horizontally	Shifts horizontally	0.05 cm^{-1}
$\frac{\alpha}{N_D - N_A}$	Changes shape slightly at large α	Changes shape	6.62×10^{-18} cm^2

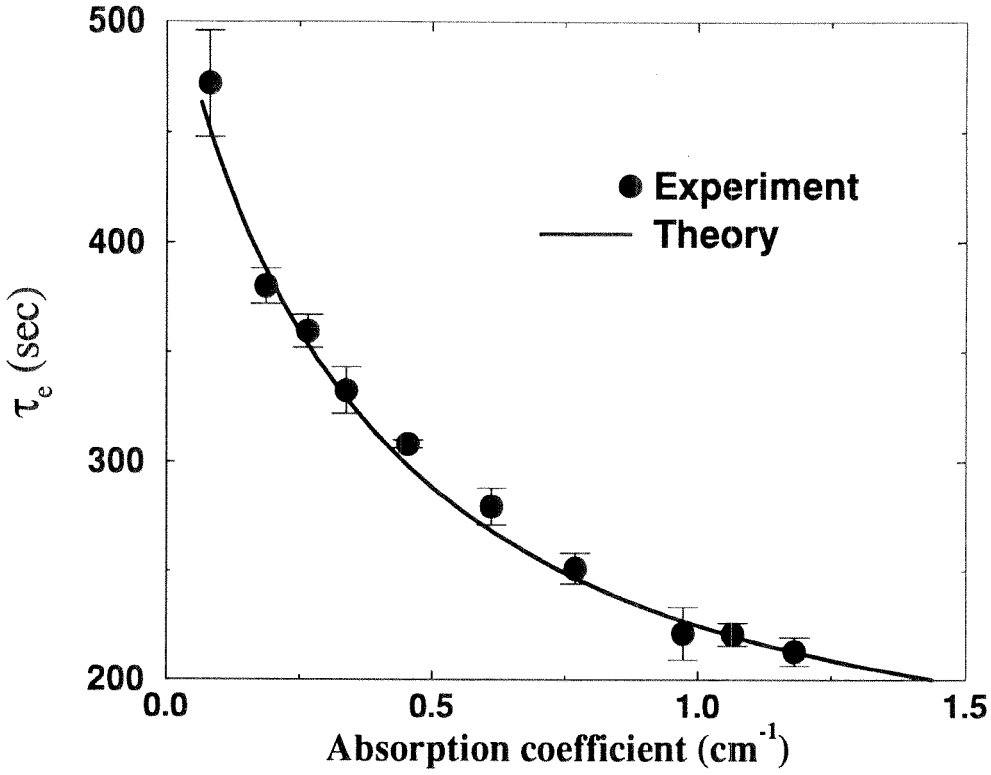
Table 4.8: List of uncertain parameters, and their fitted values.

to the values in Table 4.5. Likewise, the value for E_{0ph} is well within the range given in Equation 4.49. The value of $\alpha_{useless}$, for which only the $M/\#$ curve was used for fitting, is smaller than the lowest measured absorption coefficient (0.08cm^{-1}). The final parameter is the relationship between absorption at 488nm (α), and the Fe^{2+} occupancy ($N_D - N_A$). We used only the τ_e^{ext} curve to fit this, since the $M/\#$ is barely affected by small changes in $\alpha/(N_D - N_A)$.

After all this, we show in Figure 4.6 the prediction of the theoretical model for $M/\#$ and τ_e^{ext} with the experimentally measured data.[19] The agreement between the measured data and theory is fairly striking. The dotted line in the $M/\#$ vs. absorption plot corresponds to the $M/\#$ from a schedule computed with the time–dependent τ_e from Section 4.1.5. The approximate theory is quite accurate, even for large values of absorption coefficient.

What does this result mean? It means that we can use our theoretical model with a fair amount of confidence that its results have some validity. For instance, let's assume we're ordering a crystal from a previously untried manufacturer. Because of the uncertainty in E_{0ph}, we don't know exactly what $M/\#$ we're going to when we record holograms. However, this uncertainty only scales the $M/\#$, without affecting the shape of the dependence on absorption. We do know that if we use our theoretical model to select an absorption coefficient, we will be getting the maximum $M/\#$ available with that crystal. There is no need to repeat the oxidation/reduction experiments from Chapter 3. In fact, since the annealing process is best done before polishing and coating, we could ask the manufacturer to deliver the crystal with the appropriate absorption coefficient.

[19] see Chapter 3.2 for details of the measurement procedure.



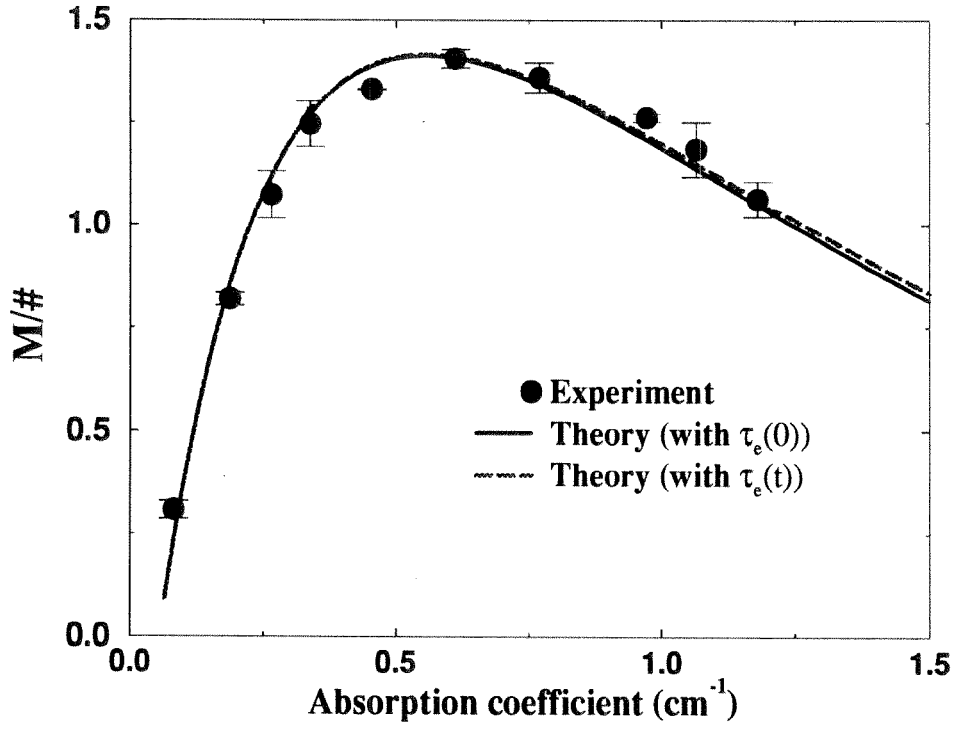


Figure 4.6: $M/\#$: experiment vs. theory

The crystal manufacturer has a moderately large tolerance when setting the absorption coefficient, since Figure 4.6 shows that the $M/\#$ is not a rapidly changing function of α near the maximum. However, if you're going to miss the peak absorption coefficient, there are two good reasons to err on the high side. First, the penalty in $M/\#$ is smaller since the slope levels out on the high side of the maximum. Secondly, a larger absorption coefficient gives faster recording (lower average recording time). We explore this tradeoff between $M/\#$ and average recording time in Section 4.3.2.

4.2 Predictions of the model

In this section, we use the theoretical model to predict the behavior of the $M/\#$ as a function of some of the other variables that we have under our control. For instance, we can control the external amplitude ratio, S/W, to find an optimum $M/\#$ as shown in Figure 4.7. In a sense, this is expected, since the rule of thumb is that holograms are proportional to modulation depth. We said in Section 4.1.3 that the holograms are proportional to the product of the signal and reference amplitudes, or

$$\frac{A_0}{\tau_r} \propto S_0 W_0. \tag{4.50}$$

Since τ_e is itself inversely proportional to the total intensity, the $M/\#$ is then proportional to modulation depth,

$$M/\# \propto \frac{S_0 W_0}{S_0^2 + W_0^2}. \tag{4.51}$$

In the following paragraphs, we describe the behavior of the $M/\#$ as a function of photovoltaic field, crystal doping level, crystal length, and crystal width. For each variable, we plot the best available $M/\#$, optimizing for both absorption coefficient and modulation depth. The default parameters are

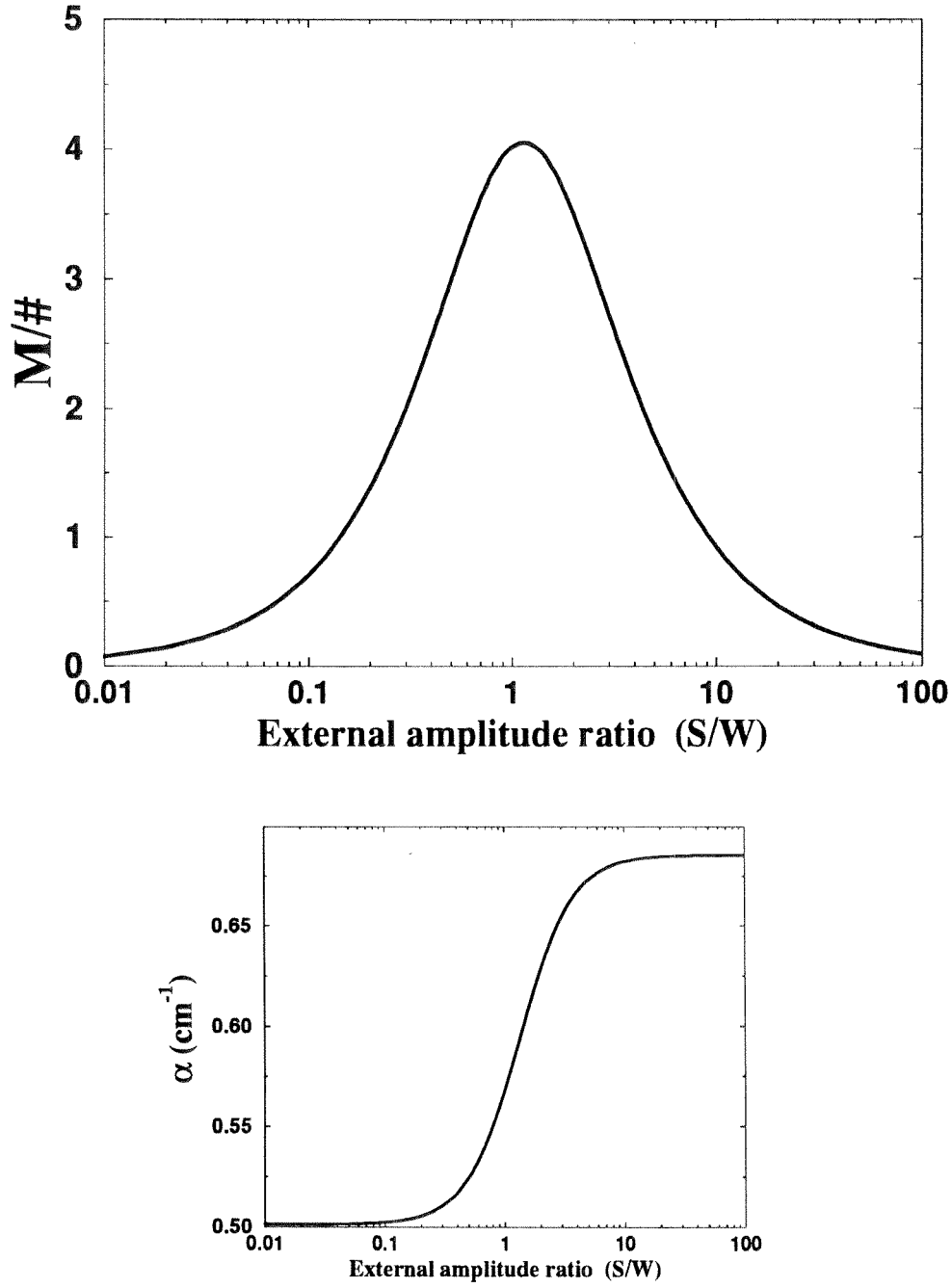


Figure 4.7: $M/\#$ vs. external amplitude ratio

Photovoltaic field:	$1.145 \times 10^{-14} \times N_A$ V/cm
Crystal doping:	0.015 Fe doping (mol %)
Crystal length:	2 cm
Crystal width:	1 cm

These values correspond to the crystal which we used for our 160,000 hologram demonstration (Chapter 6.4). Note that we assume that the usable region of the crystal is smaller than the dimension of the crystal by 1mm on each edge. Thus, a crystal length of 2cm implies that the grating is written in an interaction region of 1.8cm.

4.2.1 Photovoltaic field

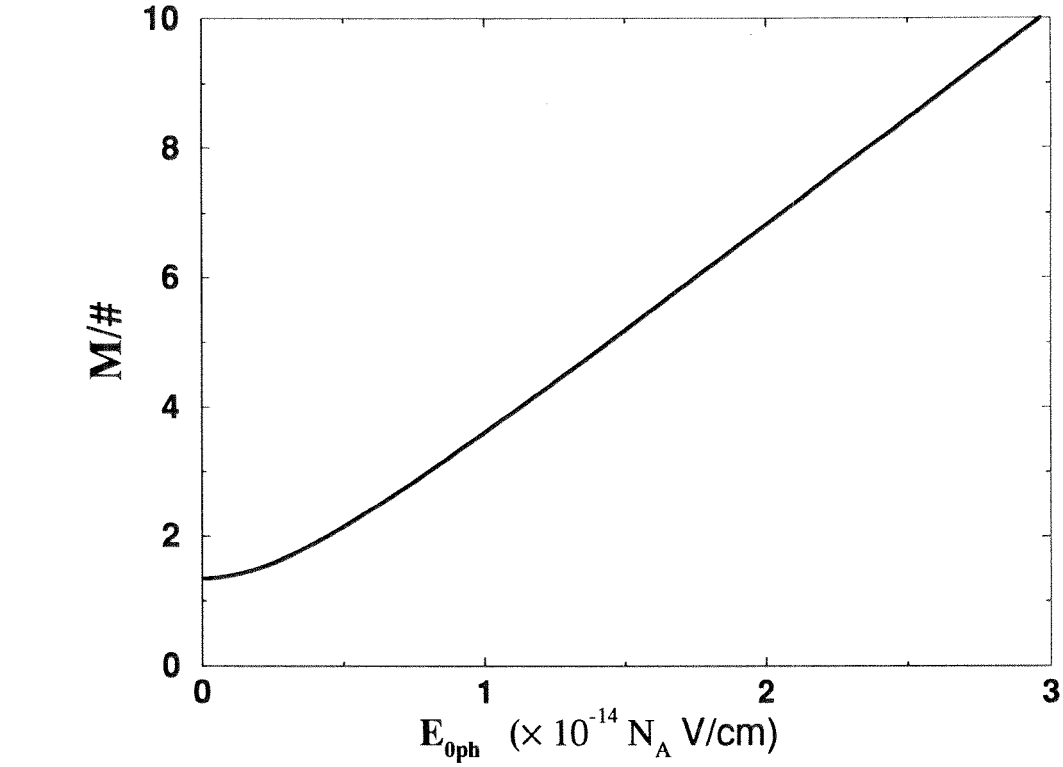
As shown in Figure 4.8, the $M/\#$ is a strong function of the photovoltaic field. The range shown is the same one extracted from the literature in Section 4.1.7. Note that the diffusion field has a noticeable effect at low photovoltaic fields, but is definitely a secondary influence for moderate photovoltaic fields. As a result, if we are uncertain about the value of E_{0ph} that we will have in any crystal, we are uncertain about the $M/\#$ that we will obtain. However, note that the value of the optimum absorption coefficient is effectively independent of the photovoltaic field. This is good news: we can still use our model to pick an absorption coefficient and be assured that we are getting the optimal $M/\#$, despite any uncertainty concerning the size of the photovoltaic field.

4.2.2 Fe doping level

At first glance, Figure 4.9 would appear to be quite exciting. We should be able to solve all our dynamic range problems just by doping the crystal more heavily. However, Figure 4.9 does not quite reflect the whole story. There are two additional items to consider:

- **Photovoltaic field**—In plotting Figure 4.9, we have assumed that the photovoltaic constant[20] does not change with doping level. We do not know whether

[20]The number we multiply by N_A to get the effective photovoltaic field.



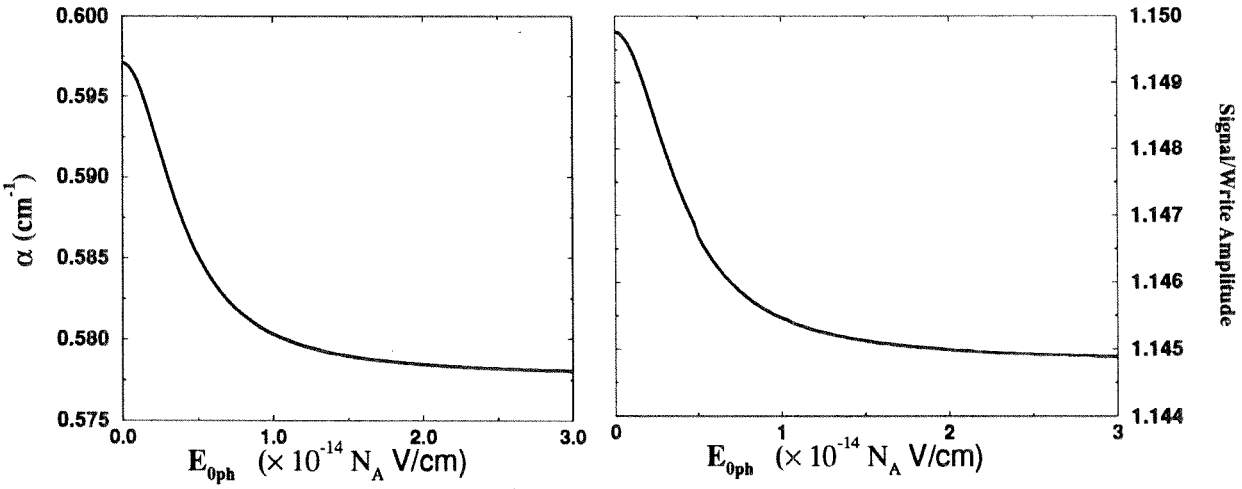


Figure 4.8: $M/\#$ vs. Photovoltaic field

this is true—if the photovoltaic constant is linked to the doping concentration in some way, then the curve shown in Figure 4.9 will not be correct.

- **Dark conductivity**—It has generally been observed that the dark conductivity is much larger for high (> 0.05 %) doping levels. So for our gains in diffraction efficiency, we are sacrificing dark storage time. If a system designer is primarily concerned with erasure during illumination, a small loss of dark storage time may be acceptable.

Note that the optimal absorption coefficient is again relatively independent of the doping level.

4.2.3 Crystal length

When we first studied the Long Interaction Length Architecture, we postulated that we could always improve the dynamic range by increasing the length of the crystal. Our explanation was that we could keep the $e^{-\alpha L}$ loss from increasing by decreasing the absorption coefficient. Although this would slow down the crystal, we would get more diffraction efficiency because we were integrating the volume hologram over a longer interaction length. In Figure 4.10, we plot the absorption loss for several crystal lengths. We can maintain the same vertical position (absorption loss) on each of the various dotted lines by moving to smaller α values as L increases. However, as we move to smaller absorption coefficients, we see from the solid line that the saturation space–charge field begins to decrease.[21] We enter a realm of diminishing returns: we increase the length L, decrease the absorption α, but don't get much more $M/\#$ because the photorefractive dynamics are being affected as well.

The effect of crystal length L on $M/\#$ is shown in Figure 4.11. As we expected, the improvement in $M/\#$ tails off at large L, yet the optimal absorption coefficient continues to drop with increased crystal length. At this point, we can think about choosing a tradeoff between dynamic range (represented by $M/\#$) and average recording time (inversely proportional to α). See Section 4.3.2 for some thoughts on this.

[21] We become absorber–limited, as opposed to trap–limited.

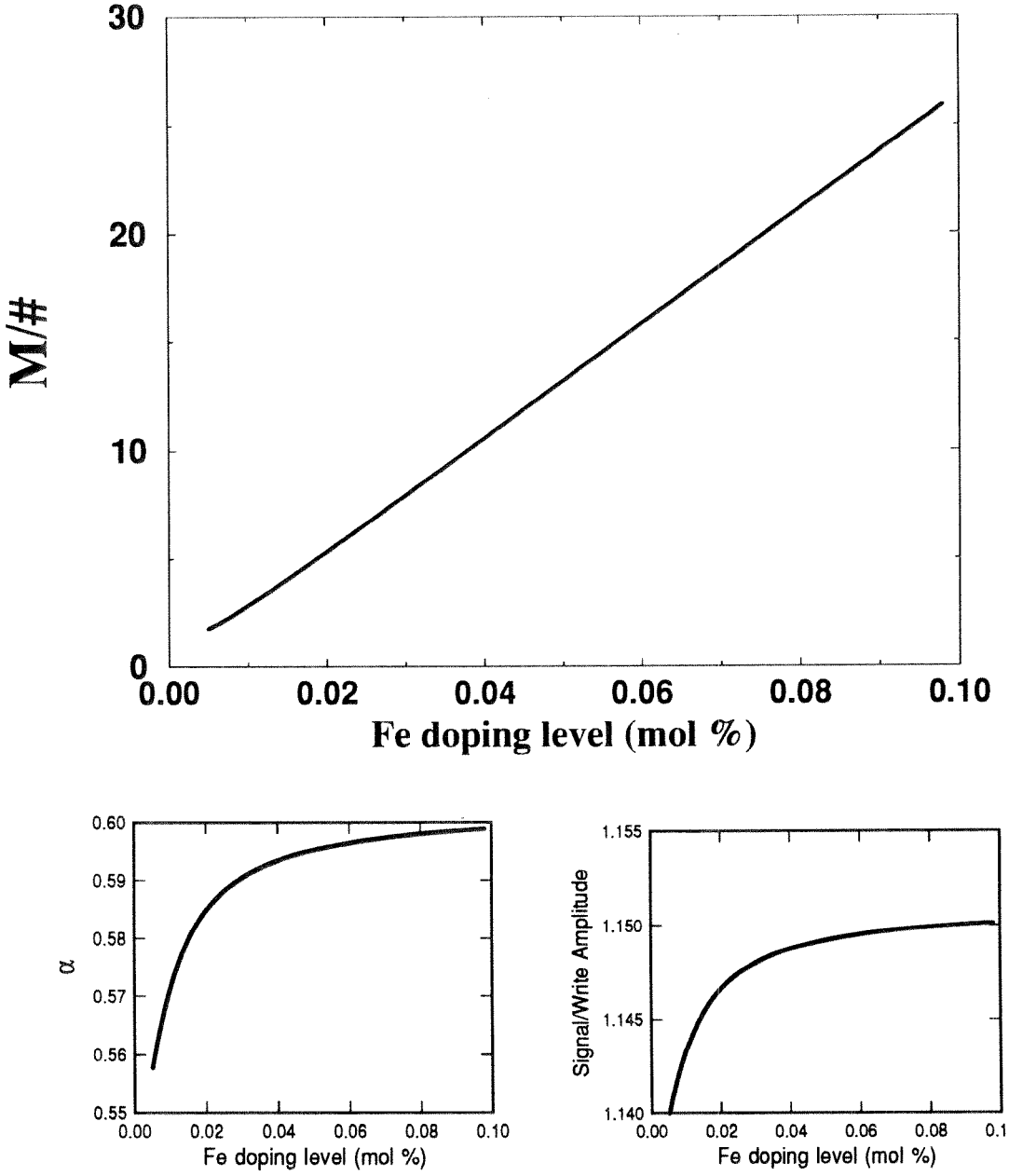


Figure 4.9: $M/\#$ vs. Fe doping level

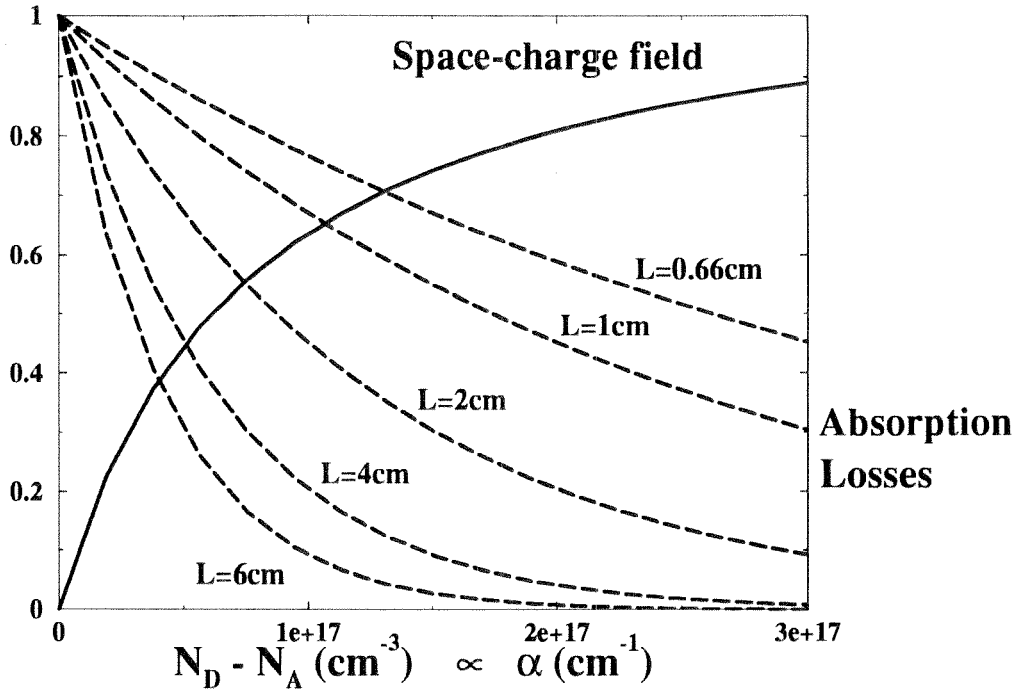


Figure 4.10: Space–charge field and absorption loss for various length crystals, as a function of absorption coefficient

4.2.4 Volume of crystal used

We saw in Section 4.1.2 that diffraction efficiency is proportional to the cross–sectional area of the hologram: interaction length $\times$ signal width. As with the crystal length, we expect that the increase in $M/\#$ with crystal width W will saturate, and that the optimal absorption will fall as we try to obtain more dynamic range. This is shown in Figure 4.12. Note that since most SLMs are symmetric, an increase in signal width is accompanied by an increase in signal height. The height change does not improve diffraction efficiency, but reduces the number of storage locations we can put in any given volume of crystal. There is a tradeoff between M/# and the vertical height of the storage location within the crystal, which we explore now.

The reason this tradeoff exists is that the width and height of the signal beam are related (the horizontal and vertical SLM pixel spacings are usually quite similar). Because of this, every time we change the magnification ratio or focal length to make the signal beam wider in the crystal, we are also making it larger vertically and have

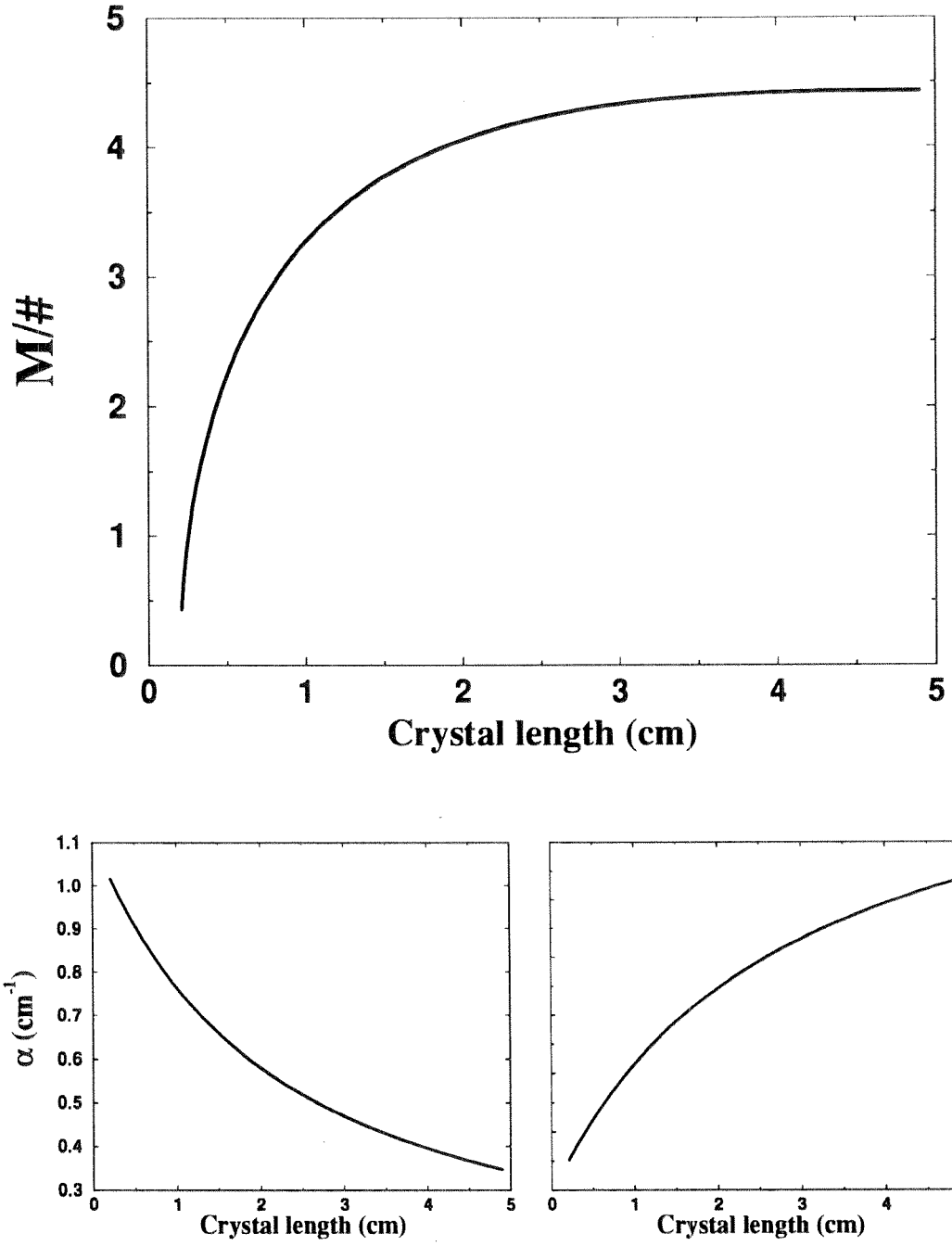


Figure 4.11: $M/\#$ vs. crystal length

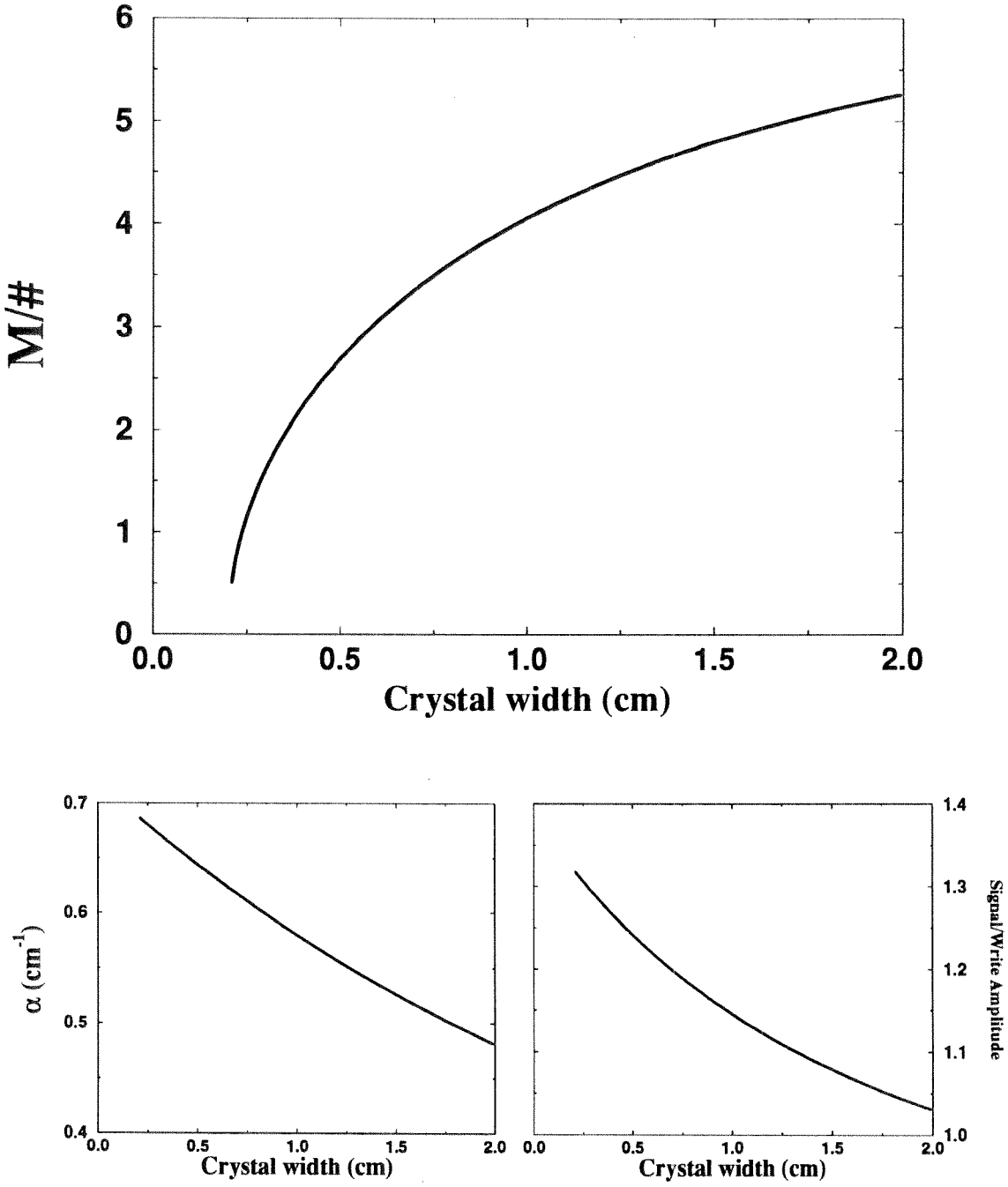


Figure 4.12: $M/\#$ vs. width of Signal beam

fewer storage locations in any given height of crystal. For instance, we might assume that we have 4cm of crystal to work with. To examine our tradeoff, we assume that each vertical storage location must be separated from the next by 1mm. This is plausible as long as the SLM image is not so large that the signal beam dramatically changes size within a centimeter of the Fourier transform plane. Let's also assume that we will use our $M/\#$ to store as many holograms as possible as long as the diffraction efficiency is above the target for our photon budget. We won't talk about this until Chapter 7.1.3, but the gist is that our target diffraction efficiency is about 3×10^{-6}.

With these assumptions, the maximum number of holograms in a 4cm crystal is the lower curve in Figure 4.15. It reaches a maximum of 20,000 holograms for a signal width of about 1mm (the aspect ratio of the pixel spacing is assumed to be 1:1). This is a pretty dismal prospect, considering that we are going to use 4cm of crystal to store 16 locations in later chapters. However, we have not used much of the horizontal width of the crystal. It turns out that we can get a lot more holograms if we move the signal beam back and forth horizontally within each storage location. This is what we have been doing in practice (see Chapters 5 & 6), although with focusing signal beams instead of the plane waves we are using here. This theoretical approach is perfectly valid, though, for signal beams with a pixel–matched random phase plate at the SLM, since each pixel can be treated as a "plane wave" and the signal intensity is roughly spatially uniform.

To analyze the effects of moving the signal beam horizontally within a storage location, we need to know how the $M/\#$ varies with horizontal position. In Figure 4.13, we show the $M/\#$ for a 5mm wide plane wave, as it is moved horizontally away from the reference beam. The horizontal axis gives the distance the reference beam travels in the crystal before encountering the edge of the signal beam. As expected, as the signal beam moves to the far side of the crystal, the $M/\#$ decreases because of the additional absorption encountered by the reference beam. This is not a question of modulation depth, because for each point shown in Figure 4.13, we have selected the

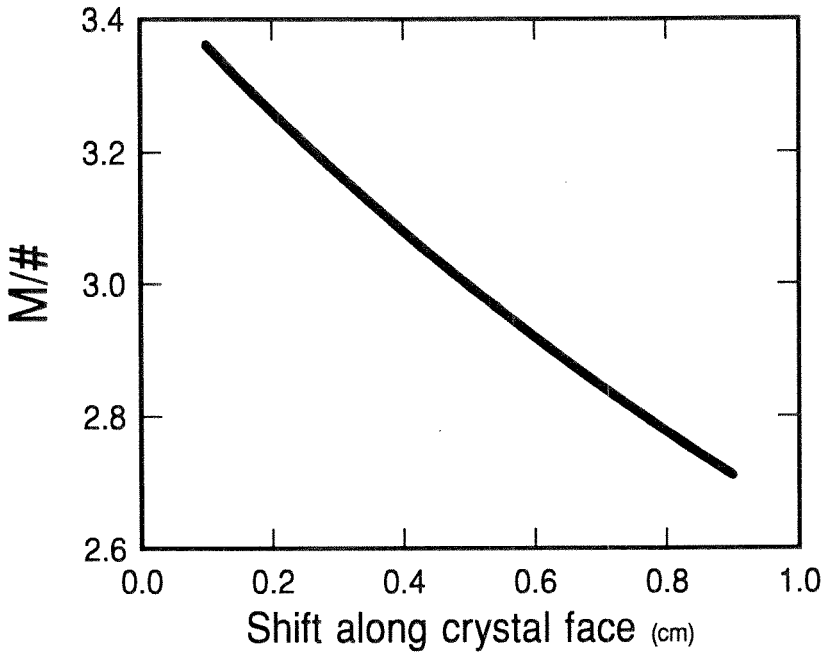


Figure 4.13: $M/\#$ vs. position of Signal beam

external ratio of beam intensities which gives the best $M/\#$.[22]

Now we can see the effects of moving a signal beam. The lower curve in Figure 4.14(a) is the same as that in Figure 4.12, plotted on logarithmic axes. In comparison, we show the $M/\#$ if the signal beam is moved horizontally throughout the storage location during the sequence of exposures. To make the calculation easier, we used the empirical observation from Figure 4.13 that the $M/\#$ as a function of shift along x is reasonably linear. We then average the $M/\#$s for the nearest and farthest position. The reason the $M/\#$ gets so large is that the average signal intensity during erasure is reduced by the ratio:

$$\frac{\text{Signal width}}{\text{Total crystal width - 0.2cm}},$$

where the 0.2cm represents a safety margin which keeps the signal beam from being clipped by the edge of the crystal. Because we get to use all the signal beam intensity for recording but experience much less of it during erasure, we find that it makes

[22]This means that, in order to achieve the best $M/\#$ in practice, we would need to tune the reference and signal powers as the recording schedule proceeds.

sense to

1. Make the signal beam as narrow as possible.

2. Make the signal intensity much larger than the reference beam.

We should note that this analysis is going to break down for signal beams which are narrower than 1mm, because the signal beam which reaches this size at the Fourier transform plane is arriving with very steep angles. In addition, the external ratio of intensities will call for large signal intensities. Since the loss in the signal arm is usually much greater than the reference arm, the practical realization of such intensity ratios would make the total intensity extremely low and extend the average recording time to unreasonable values. So getting an $M/\#$ larger than 10 by using moving signal beams is probably not a practical option.

Another point to note is that the last few holograms written do not experience the spatial averaging effect of the erasure. For instance, holograms written at the back side of the crystal but early in the schedule are initially weak because there isn't much reference beam power there. However, during the recording schedule, this same hologram will also erase more slowly and the final diffraction efficiency will not be a strong function of position. However, holograms written on the final pass of the signal beam through the crystal (say, the last 100–200 holograms written) *will* have their diffraction efficiencies be strong functions of position. For these last few holograms, one can plan to schedule their exposures to even out their diffraction efficiencies.

So if we return to our tradeoff between $M/\#$ and signal width, but use the entire width of the crystal to store holograms, we find that we can get many more holograms in our 4cm tall crystal (Figure 4.15). The center curve in Figure 4.15 corresponds to an SLM with a 1:1 aspect ratio in pixel spacing, while the top curve is for a 2:1 aspect ratio, so that the signal height at the Fourier transform plane is half the width. Again, we take the caveat that we can't really expect to fit the Fourier transform of real SLMs through a cylinder at the Fourier transform plane which is 2cm long and smaller in vertical height than 1mm. However, we certainly note the trend to more storage locations and smaller signal beams. A conservative goal is 100,000–200,000

holograms in the 4cm crystal.[23] And these predictions are for the parameters which correspond to the crystal we have in the lab—$M/\#$s as large as 12 have been reported, and even larger values might be expected with higher doping and/or changes in poling procedure or stoichiometry.

4.2.5 Miscellaneous

AR coating

Without anti–reflective coating, Fresnel reflection coefficient for normal incidence of $LiNbO_3$ is 0.8371. One effect of an improvement in reflectivity is an increase in the total amount of power present during recording. While this will make the storage process faster, it does not improve $M/\#$. In addition, however, more readout beam power makes it into the crystal to be diffracted, and more of the diffracted power exits the crystal towards the detector array. The factor by which the $M/\#$ improves is

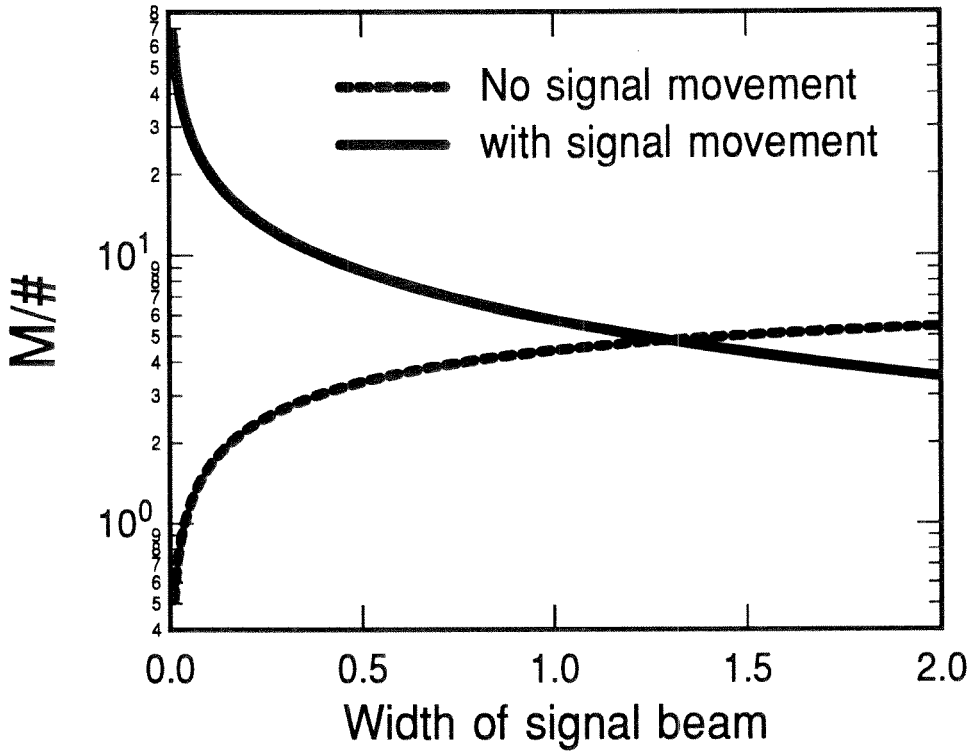
$$\frac{\text{Power transmission of AR surface}}{.8371},$$

and cannot be greater than 119% of the original $M/\#$.

Phase–conjugate readout

The only difference between phase conjugate readout and our normal readout is the absorption of the readout beam on the way into the crystal and of the diffracted light on the way out of the crystal. All of the other absorption effects during recording remain identical. For phase–conjugate readout, the readout beam enters from the opposite face of the crystal (being the complex conjugate of the original reference) .

[23]Note that in our experiments in Chapter 6.4, we are using a signal beam which focuses strongly in the crystal rather than a uniform signal beam, because the best phase plate we made in our laboratory has phase regions of 8×8 SLM pixels. However, we have observed that we get much stronger holograms, reduced interpixel noise, and reduced photovoltaic distortion by moving our signal beams horizontally throughout the storage location.



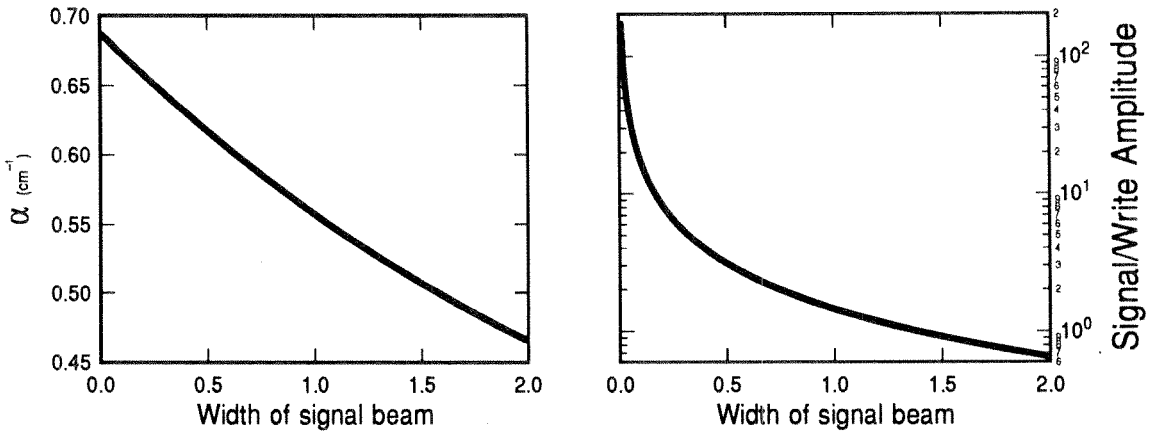


Figure 4.14: $M/\#$ vs. width of Signal beam, with and without horizontal movement of the signal beam

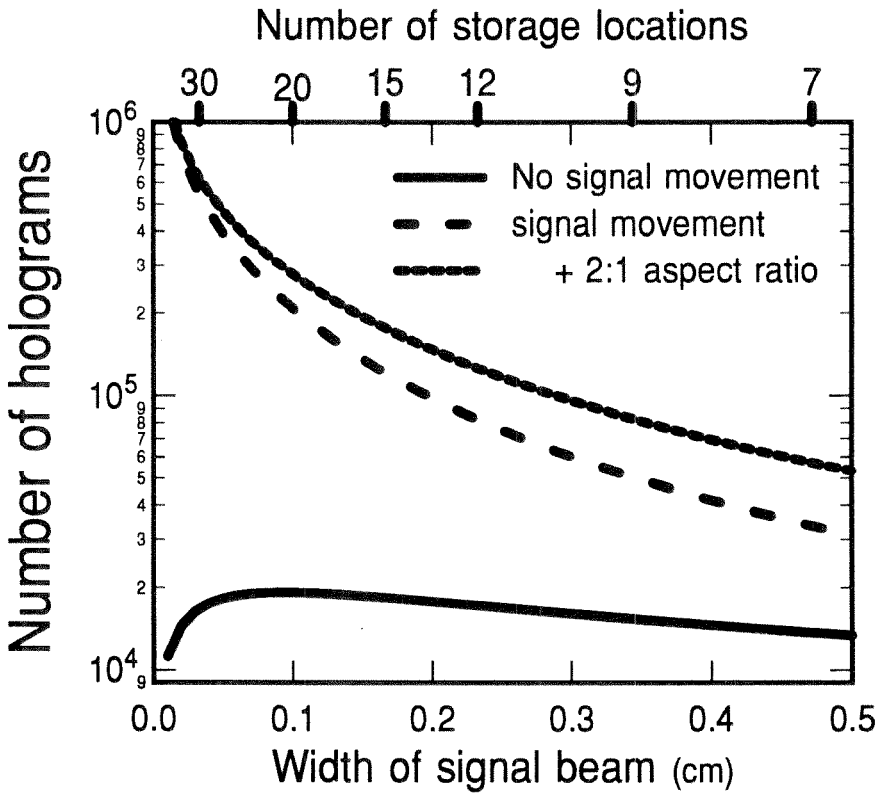


Figure 4.15: Number of holograms that can be stored in 4cm height of crystal vs. width of Signal beam. The crystal is assumed to be 1.5cm wide and there is a 1mm buffer between each storage location.

Then, for each point (x, y), instead of a contribution of

$$e^{-\alpha x} e^{-\alpha(L-z)},$$

there is a contribution

$$e^{-\alpha(X-x)} e^{-\alpha z},$$

where X is the width of the crystal. Since this change occurs in the integrand before spatial integration, we need to reintegrate. The result is in a simple ratio between the phase conjugate $M/\#$ and the normal $M/\#$:

$$\frac{M/\#_{\text{phase conjugate}}}{M/\#_{\text{normal}}} = \frac{e^{-(\alpha/2)X}}{e^{-(\alpha/2)L}} \quad \frac{x_2 - x_1}{z_2 - z_1} \quad \frac{e^{-\alpha z_1} - e^{-\alpha z_2}}{e^{-\alpha x_1} - e^{-\alpha x_2}}. \tag{4.52}$$

For the range of absorptions and crystal dimensions which we usually use, this factor ranges from 0.95 to 1.1. We do not show any plots since they would not be especially revealing. The gist of this discussion is that, theoretically speaking, the diffraction efficiency of holograms should be similar no matter whether the original or phase–conjugate reference is used. Any loss of diffraction efficiency is mostly due to mismatch between the spatial frequency profile of the two readout beams.

There are some minor philosophical points. If you are going to use the original reference, you should place the signal beam close to the face through which the reference will enter (making x small). If the reference beam is not as wide as the crystal, there is no preference to which portion of the crystal is illuminated (along the z direction). This is a statement of the fact that we have to suffer the $e^{-\alpha L}$ loss, whether we do it before or after storage.

For phase–conjugate readout, the situation is reversed. The placement of the signal beam along its entrance face (x) is now completely arbitrary. And we would getter higher $M/\#$ if the reference beam intersects the signal beam soon after the signal enters the crystal (low z).

4.3 Additional complications

4.3.1 Non–uniform grating profile

In the beginning of the chapter, when we needed to know the distribution of the output spatial frequencies, we assumed that the strength of the grating was uniform throughout the volume. In this section, we want to show that the nonuniform grating profile is not going to cause a problem. There are two things to think about: effects on the spatial frequency output, and on the Bragg selectivity. Our initial assumption of spatial uniformity would result in a sinc function for both; but now we expect the nulls of the sinc functions to move away from the center and to be "filled in" (to no longer be identically zero). When we say spatial frequency response, we are referring to the impulse response—what the hologram does to each plane wave we try to store.

We return to our expression for the Born approximation, just as we did at the beginning of our $M/\#$ derivation:

$$S_x \iint dk_{d_x} dk_{d_y} \left[\iiint \frac{\Delta\epsilon(\mathbf{r}') e^{j(\mathbf{k}_i \cdot \mathbf{r}')} e^{-j(\mathbf{k}_d \cdot \mathbf{r}')} d\mathbf{r}'}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}} \right] e^{j(\mathbf{k}_d \cdot \mathbf{r})}. \tag{4.53}$$

This time, however, when we substitute the permittivity modulation (given by Equation 4.3), we keep the various phase terms. The full expression becomes

$$S_x \iint dk_{d_x} dk_{d_y} \left[\iiint \frac{\Delta\tilde{\epsilon}_x(\mathbf{r}') e^{j((\mathbf{k}_i - \mathbf{k}_j) \cdot \mathbf{r}')} e^{j((\mathbf{k}_s - \mathbf{k}_d) \cdot \mathbf{r}')} d\mathbf{r}'}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}} \right] e^{j(\mathbf{k}_d \cdot \mathbf{r})}. \tag{4.54}$$

As we did before, we can drop the y dependence in the spatial integral because no terms change along that direction. We can follow all of the steps that we performed in the initial derivation, carrying along the phase terms that represent variation in the readout beam and diffracted direction. We note that the original signal ($\mathbf{k}_s$) and reference ($\mathbf{k}_j$) were on–axis. We can substitute this in, leaving only the change in the direction of the readout beam ($\mathbf{k}_i$), and change in the direction in which we observe

the diffracted light ($\mathbf{k}_d$). Mathematically, this works out as

$$e^{j((\mathbf{k}_i-\mathbf{k}_j)\cdot\mathbf{r}')} \approx e^{jkx'} \left(e^{-jkx'}\ e^{-jk_{i_z}z'}\right) \tag{4.55}$$

$$e^{j((\mathbf{k}_s-\mathbf{k}_d)\cdot\mathbf{r}')} \approx e^{jkz'} \left(e^{-jkz'}\ e^{-jk_{d_x}x'}\right). \tag{4.56}$$

Here we have assumed that there are only small changes of the reference beam away from normal incidence (that is, along the x axis), and of the diffracted beam away from normal exit (that is, along the z axis). This allows us to decouple the effects of Bragg mismatch from those of variation in spatial frequency. For the same reason (only small angle changes), we will assume that the denominator $\sqrt{k^2 - k_{d_x}^2}$ is approximately k.

We are interested in the grating profile as a function of time, so we rewrite the time dependence from Equation 4.29 as

$$\sqrt{\eta}(t) \propto \iint dk_{d_x} dk_{d_y} \left[\int_{z_1}^{z_2}\int_{x_1}^{x_2} e^{-\alpha x} e^{-\frac{t}{\tau_x} I_0(x,z)} e^{jk_{i_z}z'} e^{jk_{d_x}x'} dx'dz'\right]. \tag{4.57}$$

This is the scattered field at the exit of the crystal. Since we are interested in the spatial frequency response of the diffracted light, we only need the term within the brackets. This can be written as

$$\begin{aligned}\sqrt{\eta}(t) \propto & \int_{x_1}^{x_2} e^{-\alpha x'} e^{-(\frac{t}{\tau_x} W_0^2 e^{-\alpha x'})} e^{jk_{d_x}x'} dx' \\ & \times \int_{z_1}^{z_2} e^{-(\frac{t}{\tau_x} S_0^2 e^{-\alpha z'})} e^{jk_{i_z}z'} dz'.\end{aligned} \tag{4.58}$$

A quick check that we're on the right track before we go on. If the absorption coefficient were zero, then each of these integrals is the Fourier transform of a rect or top–hat function. As expected, we get a sinc dependence for both Bragg selectivity (k_{i_z}) and horizontal output spatial frequency (k_{d_x}). (Recall that the vertical output spatial frequency will continue to follow a sinc in k_{d_y} even for non–zero absorption).

Now let's see the effect of the absorption coefficient. To keep things simple, we look at time $t = 0$. The first 3–D plot in Figure 4.16 shows the x, z distribution of the grating strength at $t = 0$ (the integrand of Equation 4.58). There is no

z dependence at all, which leads us to expect that the Bragg selectivity is still a sinc. The x dependence is a decaying exponential, so the Fourier transform should be the convolution of a sinc with a Lorentzian.[24] What happens, though, as the grating continues to be erased by the spatially varying reference and signal light? The regions near the entrance faces of these two beams get erased faster than the portions deeper in the crystal (See Figure 4.16). In the z dimension, this creates a non–uniformity which affects the Bragg selectivity. In the x direction, the erasure actually counteracts the original ($t = 0$) nonuniformity in grating strength, so that the grating strength is roughly uniform in x at some set time $t > 0$. In Figure 4.16, this occurs near $t = 2\tau_e$. After this point, erasure continues and the grating strength becomes nonuniform again.

In summary, then, we expect the Bragg selectivity to start at $t = 0$ as a sinc and get progressively worse. In contrast, the output (horizontal) impulse response in spatial frequency is non–sinc–like at time $t = 0$, and passes through near sinc–like behavior as the erasure continues. These effects are shown in Figure 4.17. For $\alpha = 0.5$, the dependence at $t = 3\tau_e$ still has troughs and peaks, in roughly the same places as the sinc. For a larger absorption coefficient (lower left portion of Figure 4.17), the troughs and peaks in the distributions become vestigial. Crosstalk between holograms and horizontally broadened pixels will result. Figure 4.17 shows some good reasons why we don't want to use high absorption coefficients or long exposure schedules (corresponding to large f parameters from our treatment in Chapter 2).

4.3.2 Getting large $M/\#$s and low average recording time

We mentioned earlier that the $M/\#$ does not indicate how fast the holograms are going to record. To decrease the average recording time, we can increase the intensity of light during recording. We do not get a larger $M/\#$, but we do get our holograms faster. We can show that we spend the same amount of energy either way, because

[24]The Fourier transform of $e^{-|x|}$ is a Lorentzian: $2/(1 + x^2)$.

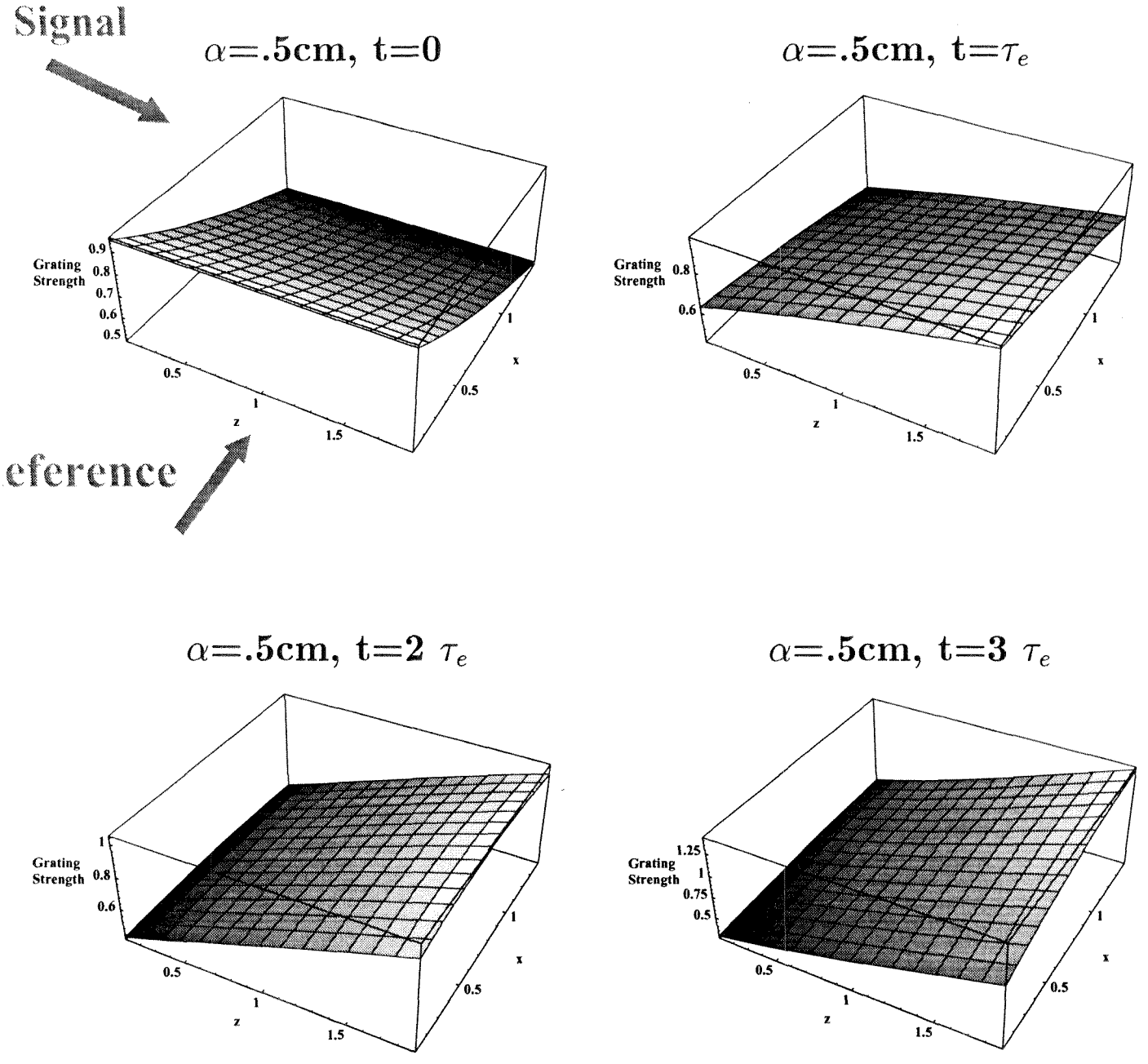


Figure 4.16: Grating strength inside crystal as a function of exposure time.

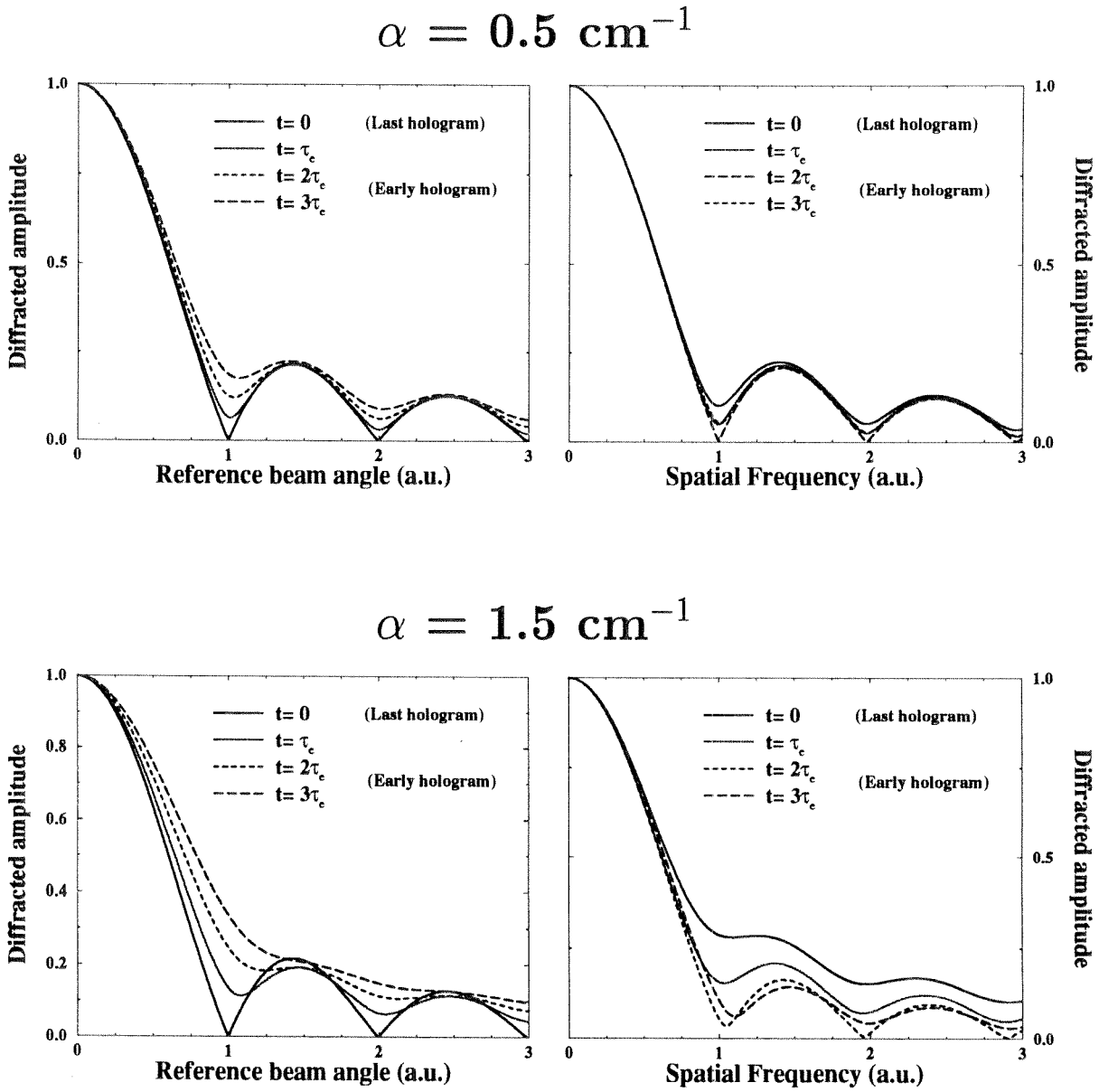


Figure 4.17: Bragg selectivity and spatial frequency response (horizontal) as a function of exposure and absorption coefficient.

we have linear recording.

Let's say that the dimensions of the illuminated volume are $\Delta x \times \Delta y \times \Delta z$, and that we have chosen an f value which sets the total recording exposure to about $2\times$ the erasure time constant τ_e. The total power (neglecting losses in the optics) is then

$$S_0^2(\Delta y\ \Delta x) + W_0^2(\Delta y\ \Delta z), \tag{4.59}$$

and the total recording time

$$2\tau_e = 2\frac{\tau_x}{I_0}. \tag{4.60}$$

Again, τ_x corresponds to all the terms in the expression for erasure time constant except for light intensity. By multiplying τ_e and the total power, we have an expression for the total energy. Since this expression contains S_0^2 and W_0^2 in both numerator and denominator, an increase in total intensity does not affect it.

We might want to choose some tradeoff between $M/\#$ and amount of energy expended, expressing the grating strength gained per unit of energy as

$$\frac{M/\#}{\text{Joule of energy}} = \frac{M/\#}{2\tau_x\ \Delta x\ \Delta y \left[\dfrac{S_0^2 + \dfrac{\Delta z}{\Delta x} W_0^2}{S_0^2 + W_0^2}\right]}. \tag{4.61}$$

The energy efficiency is poor if τ_x is large—this corresponds to low α. This makes intuitive sense: if the crystal is transparent, then we are sending all the light straight through with getting much of it used for hologram storage. A plot of $M/\#$ achieved per unit of input energy is shown in Figure 4.18, for a crystal corresponding to the parameters used in Section 4.2. The important change to notice (besides the scale), is that the peak of the curve is at an absorption coefficient of 1cm^{-1} instead of 0.5cm^{-1}. At this higher absorption, although we can't get an $M/\#$ that is as quite as large, we can write holograms more efficiently from an input energy standpoint.

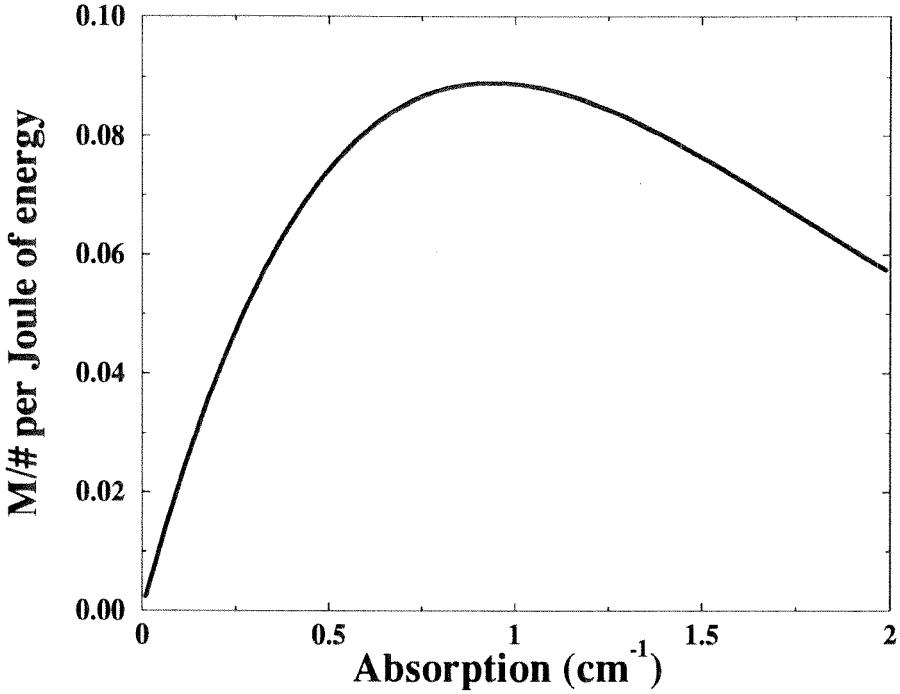


Figure 4.18: $M/\#$ achieved per unit of input energy, for 0.015% Fe doping.

4.3.3 Non–plane wave signal beams

Up to now, we have been discussing the recording of holograms between two plane waves. This is a valid treatment for information–bearing signal beams if we have a random phase plate. Here we discuss what occurs when we record holograms with information–bearing signal beams and no random phase plate.

For Fourier plane storage, each pixel in the SLM is a plane wave at the Fourier transform plane.[25] However, the total intensity profile resembles a converging cone which has a very small transverse area at the Fourier transform plane. This does not bother us when we record plane waves, because the total intensity falls out of the expression for A_0/τ_r.

If the signal S is a sum of many plane waves, then we can separate them and treat

[25]Of course, since the pixel is not infinitesimally small, its Fourier transform is not really a plane wave but instead a broad beam with a sinc envelope. If this envelope is varying slowly enough, the approximation is valid.

them separately. The modulation depth is

$$m = \frac{SW}{S^2 + W^2} = \frac{\Sigma S_i\, W}{(\Sigma S_i)^2 + W^2} = \Sigma \frac{S_i\, W}{(\Sigma S_i)^2 + W^2}, \tag{4.62}$$

where each S_i corresponds to a plane wave of different spatial frequency. The expression in the numerator is very similar to what we had before: the interference between two plane waves. The denominator is where the effect of the focusing takes place, since the complex plane waves are added before squaring. Here's where the random phase plate would make a difference: the term $(\Sigma S_i)^2$ becomes (ΣS_i^2). Either way, though, this term for total intensity cancels out for the recording behavior when we divide the modulation depth by the recording time constant τ_r. In essence, the plane wave from each pixel is uniformly recorded at the appropriate position in the crystal. You might be wondering about the portions of the signal beam which are located away from the high intensity DC spot. Well, in the same way that Fourier analysis says that there are many signal plane waves there that just happen to add to zero, there are also many holograms being recorded there that also just happen to add to zero. Since the angles involved are small outside the crystal and get smaller inside, we will treat each plane wave as if it were traveling along the z axis. The only difference between the various columns of pixels will be a horizontal shifting within the crystal depending on where the crystal is relative to the Fourier transform plane.

Specifically, the signal plane wave is always the same size, which depends on the size of the SLM pixel and the focal length of the Fourier transform lens. In our system, we used a focal length of f_{object} = 200mm, SLM pixels of 45.5μm vertically and 56.25μm horizontally, and a 0.666 magnification stage. This means that the plane wave at the Fourier transform plane was 2.6mm wide (=$2\Delta x$) and 3.2mm vertically (=$2\Delta y$). These are the values we use for the volume integration of A_0/τ_r. But we need to know how large to make the reference beam, and where the signal beam is within the crystal horizontally.

To do this, we model the converging signal beam. Most of the power is carried in the DC spatial frequency, which is H=18mm wide by V=13.33mm tall at the Fourier

transform lens and almost a point at the Fourier transform plane. We represent the distance between the focal plane and where the front face of the crystal is put as r_0. However, because of refraction, the presence of the crystal moves the focal plane back to a distance $z_0 = n_o r_0$ (See Figure 4.19). With this formulation, we can write down how large the signal beam is at the crystal surface as

$$v = | V\frac{r_0}{f} | \tag{4.63}$$

$$h = | H\frac{r_0}{f} | . \tag{4.64}$$

The horizontal extent of the signal plane wave within the crystal from

$$x_1 = x_m + x_s \mid h\frac{r_0}{f} \mid -\Delta x \tag{4.65}$$

to

$$x_2 = x_m + x_s \mid h\frac{r_0}{f} \mid +\Delta x. \tag{4.66}$$

x_s is the horizontal column within the SLM and can range from $-H$ to $+H$, and x_m is the middle of the converging or diverging signal beam. We need to make this large enough that we don't clip the edge of the crystal, or

$$x_m = \Delta x + \mid h\frac{r_0}{f} \mid +0.1\text{cm}. \tag{4.67}$$

Here we give ourselves one millimeter safety margin at the edge of the crystal.

Now we need to know what sort of intensities we are working with. Since we want to be able to analyze a variety of configurations, we will take the approach that we can divide a constant amount of power between the two beams with a polarizing beamsplitter and a half waveplate. The amount of power available at the crystal depends on the efficiency of the delivery optics. In our 160,000 hologram system described in Section 6.4, we had about 800mW available at the beamsplitter. The power efficiency of the reference beam was 3.75% (the loss comes mainly from the

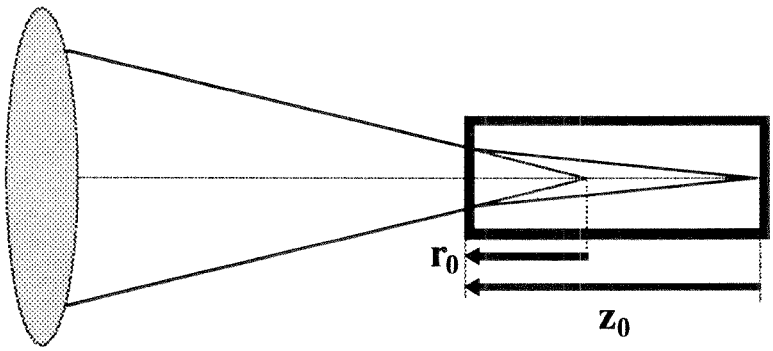


Figure 4.19: Converging signal beam.

mirror array and the 4 cylindrical lenses), and .0275% in the signal beam. The loss in the signal beam comes from the non–unity fill factor of the SLM, the half ON–half OFF nature of the displayed images, and from the sacrifice of light in order to have uniformity across the signal beam.

The area of the signal beam plane wave is 4 Δx Δy. This is what we use for considering the recording behavior, with the implicit assumption that if there are N ON pixels, each signal plane wave has $1/N$ of the total signal amplitude. For erasure, we consider that most of the signal power is in the converging cone—which at a distance z inside the crystal, has a cross–sectional area of

$$4vh(\frac{z_0 - z}{z_0})^2. \tag{4.68}$$

To keep the signal intensity from becoming infinite, we limit the value of $z_0 - z$ to 0.05cm when it would otherwise fall below this. The size of the reference beam is determined by the maximum vertical extent of the signal beam. The beam is always as wide as the crystal face (L), and the height is the larger of v and Δy.

To compute the $M/\#$, we use the same expression for A_0/τ_R from Equation 4.28 with the values for S and W that we just derived. We solve for τ_e with Equation 4.29 and

$$\tau_e = \frac{2t_0}{\log\left(\frac{\eta(0)}{\eta(t_0)}\right)}. \tag{4.69}$$

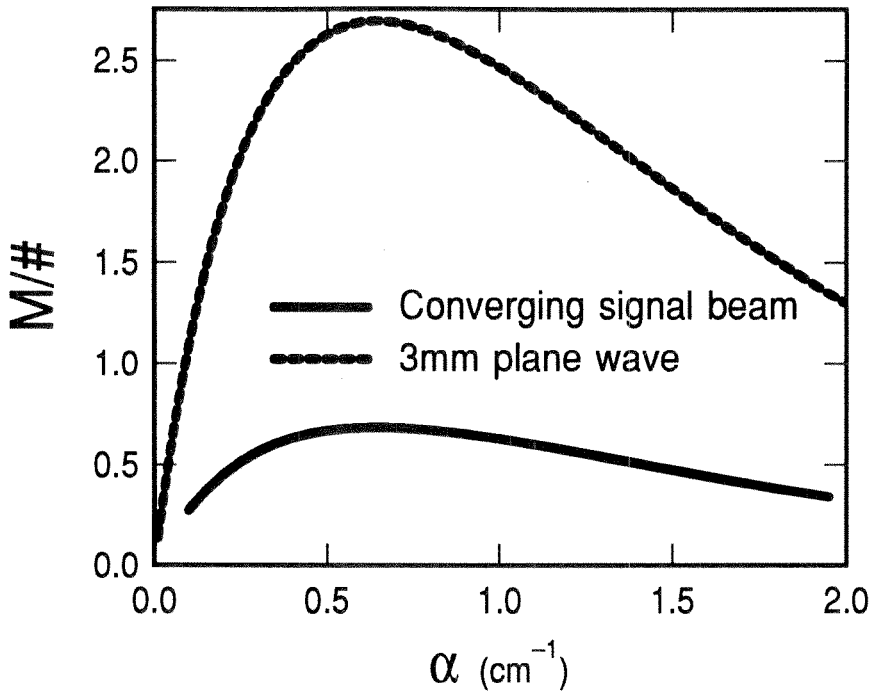


Figure 4.20: $M/\#$ vs α, for Fourier plane storage

We have several variables that we can look at. These include

Variable	Description	Default
α	Absorption coefficient	0.5cm^{-1}
r_0	Distance between the crystal and the Fourier transform plane	1cm
$\phi_{waveplate}$	Waveplate angle (division of power between reference and signal)	25°
x_s	Pixel column on the SLM	0 (center)

The first that we look at is absorption. In Figure 4.20, we plot the $M/\#$ as a function of α for Fourier plane storage. For comparison, we also show the $M/\#$ vs α for the same crystal if we use a plane wave which is 3mm wide. Note that using a smaller volume for storage affects the $M\#$ strongly.

The second thing that we look at is the $M/\#$ as a function of where the crystal is placed relative to the Fourier plane, as shown in Figure 4.21. Here we notice that the $M/\#$ gets progressively worse as the cross–sectional area of the signal beam decreases. The lack of smoothness in the region near the center is an artifact of the

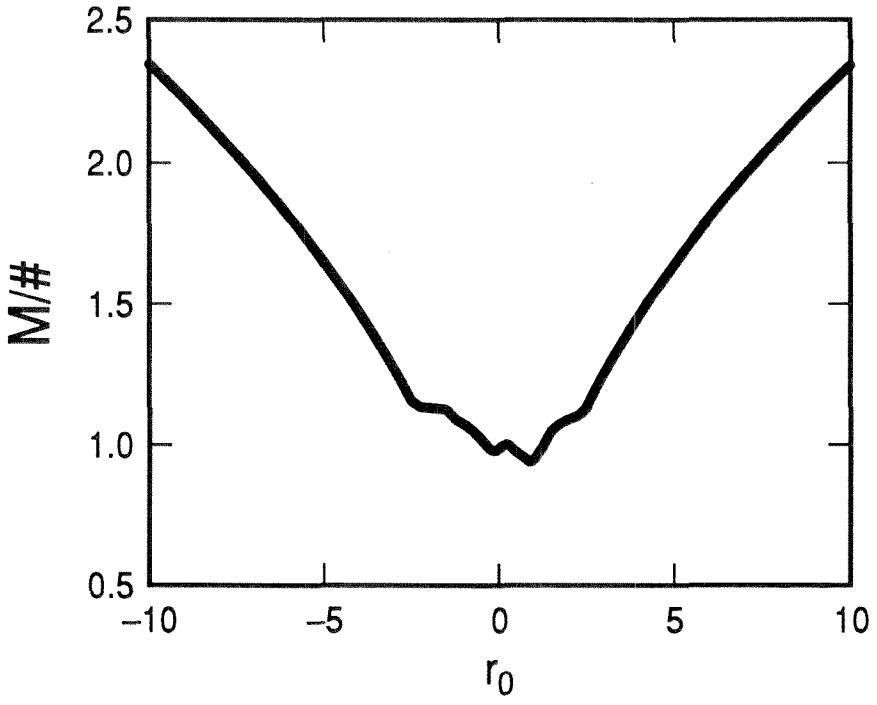


Figure 4.21: $M/\#$ vs distance between the front surface of the crystal and the Fourier plane.

spatial limiting of the signal beam to no smaller than 0.5mm. However, our purpose here is to see what sort of trend is to be expected. We also solved for the $M/\#$ as a function of the horizontal pixel position in the SLM, but since the variation turned out to be quite small ($< 1\%$), we don't show the plot.

Modulation depth in the laboratory

Another consideration is the choice of modulation depth. The cross–sectional area in the signal beam changes dramatically with the placement of the crystal relative to the Fourier transform plane. As a result, it is not always easy to know how to divide the limited input power between the reference and signal beams. We express this division as a rotation angle on the appropriate half–waveplate, where 0° corresponds to diverting all the power to the reference beam, and 45° to the signal beam. For the particular power efficiencies that we had in our system, the writing time constant A_0/τ_R was roughly symmetric with this waveplate angle (See Figure 4.22). However, since the time constant is affected by the focused signal beam, we find that we need

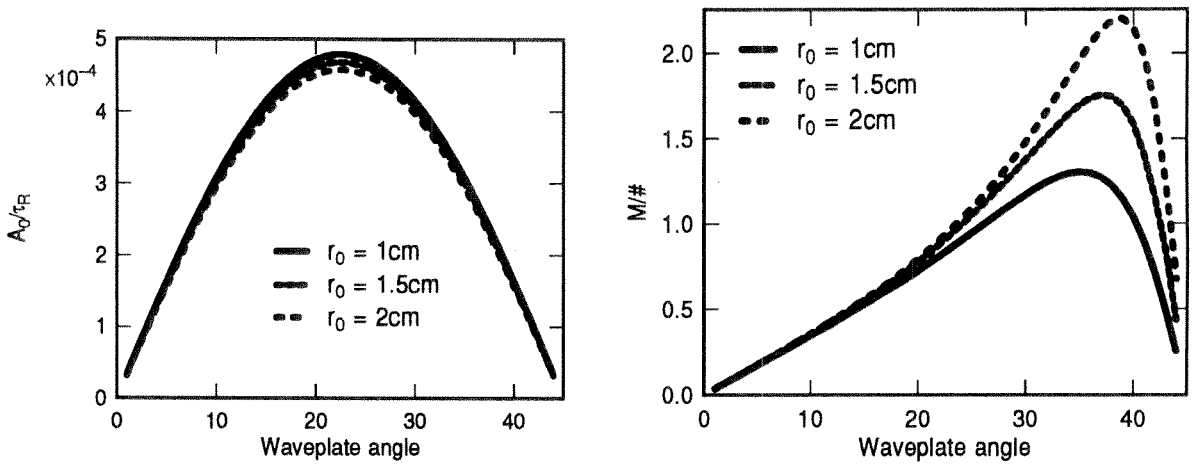


Figure 4.22: A_0/τ_r and $M/\#$ vs. half–wave plate angle in the laboratory, for Fourier plane storage.

to put more power into the signal beam in order to get the best $M/\#$. The position of this peak is a function of r_0, as shown in Figure 4.22. As we observed before, the practical limitation on this waveplate is the speed of recording. As we increase the division of power in favor of the signal beam, we find that most of our power is being wasted in the signal arm and the average recording speed becomes untenably long. In our experiment, we used an equivalent waveplate angle of 25°.

4.3.4 Off–axis reference and signal beams

In this section, we consider the $M/\#$ variations that occur for light beams that are not normally incident on the crystal. For simplicity, we return to considering plane waves. We would like to include any effect which arises from the variation in Fresnel reflection/transmission coefficients, modulation depth (including power and the dot product between polarizations), electro–optic coefficient, dielectric constant, and reduction in effective photovoltaic field. This last contribution recognizes that the photovoltaic current continues to flow along the c–axis, while the K_g vector of the grating may not be exactly parallel to the c–axis.

It turns out that almost all of these effects can be represented as dot products between a few vectors:

$\mathbf{K}_g$	Grating vector
$\hat{\mathbf{c}}$	c–axis
$\mathbf{k}_r$	Reference wave vector
$\hat{\mathbf{e}}_r$	Reference polarization vector
$\mathbf{k}_s$	Signal wave vector
$\hat{\mathbf{e}}_s$	Signal polarization vector

In fact, we can write most of the terms that we said we wanted to know in terms of these quantities. Here we write the new version of the term along with the value we used for the normally incident beams:

Term	Old	New
Fresnel reflection	$\frac{4n_o}{(n_o+1)^2}$	$\frac{SW(\hat{\mathbf{e}}_r\cdot\hat{\mathbf{e}}_s)}{S^2+W^2}$
Modulation depth	$\frac{S_0W_0}{S_0^2+W_0^2}$	
Electro–optic coefficient	r_{13}	$\hat{\mathbf{e}}_s\cdot(\mathbf{r_T}\cdot\aleph(K_g))\cdot\hat{\mathbf{e}}_r$
Dielectric constant	ϵ_z	$\epsilon_z\cos^2\phi_G+\epsilon_x\sin^2\phi_G$
Photovoltaic field	E_{0ph}	$E_{0ph}\cos\phi_G$

Here ϕ_G is the angle between the c–axis and the grating vector K_g, $\mathbf{r_T}$ is the $3\times 3\times 3$ electro–optic tensor [250], and the $\aleph()$ operation normalizes a vector to unit magnitude. Note that we have combined the Fresnel reflection and the modulation depth together, solving for the amount of light amplitude (S, W) which makes it inside the crystal as well as their polarizations.

So what are the appropriate Fresnel reflection coefficients? Since the crystal is birefringent, we need to be careful to separate the transmitted light into extraordinary and ordinary polarizations. It turns out that the rigorous solution is not even contained in Reference [9]! Here we detail an approximate solution and check it by summing the transmitted and reflected power against the input power. This is not really necessary for the small variations away from normal incidence that we are interested in here; however, we would like to use the same techniques to look at the transmission geometry. Then we will want to be able to deal with large input angles and arbitrary rotations of the c–axis. So we might as well do it the right way now.

Fresnel reflection coefficients We are not looking for an analytical solution, but more an algorithm which we can then input to *Mathematica* or some other number-cruncher. This is important because we cannot even find the wave vector of the extraordinary light inside the crystal without running into a transcendental equation.

We start with an incident wave vector $\mathbf{k}_i$, which has three components (k_x, k_y, k_z). We'll assume that it is traveling mostly along the positive z axis, corresponding to our signal beams. The algorithm for the reference beam will be identical except that the x and z axes will switch roles. The plane of incidence includes all the beams of interest and the surface normal, as shown in Figure 4.23. Note that the ordinary beam is refracted more than the extraordinary, because the ordinary index of refraction for $LiNbO_3$ is larger. The incidence angle θ_i can be written as

$$(\theta_i = \tan^{-1} \frac{\sqrt{k_{x_{air}}^2 + k_{y_{air}}^2}}{k_{z_{air}}}. \tag{4.70}$$

For the ordinary beam, the refraction process can be simply represented as multiplying k_z by n_o, giving an ordinary beam inside the crystal of

$$\mathbf{k}_o \equiv (k_{x_o}, k_{y_o}, k_{z_o}) = n_o \left(\frac{k_{x_i}}{n_o}, \frac{k_{y_i}}{n_o}, \sqrt{k^2 - \left(\frac{k_{x_i}}{n_o}\right)^2 - \left(\frac{k_{y_i}}{n_o}\right)^2} \right), \tag{4.71}$$

From this, the angle θ_o is obtained in analogy with Equation 4.70. The polarization vector of the ordinary light can be obtained from

$$\hat{\mathbf{e}}_\mathbf{o} = \aleph(\mathbf{k}_o \times \hat{\mathbf{c}}), \tag{4.72}$$

and is independent of the input polarization state.

For the extraordinary beam, the wave vector inside the crystal is the solution of the transcendental equation:

$$(k_{x_e}, k_{y_e}, k_{z_e}) = \left(\frac{k_{x_i}}{n_e(\delta)}, \frac{k_{y_i}}{n_e(\delta)}, \sqrt{k^2 - (\frac{k_{x_i}}{n_e(\delta)})^2 - (\frac{k_{y_i}}{n_e(\delta)})^2} \right), \tag{4.73}$$

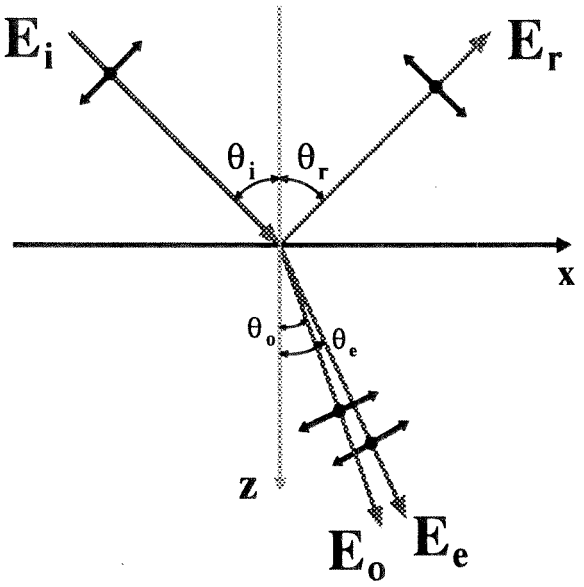


Figure 4.23: Input beams for calculation of Fresnel reflection coefficients.

where

$$\delta = \cos^{-1}\left(\aleph(\mathbf{k}_e \cdot \hat{\mathbf{c}})\right), \tag{4.74}$$

is the angle between the extraordinary wavevector and the c–axis. The difficulty is that the exact value of index of refraction is itself a function of angle [349]:

$$\frac{1}{n_e(\delta)} = \frac{\cos^2\delta}{n_o^2} + \frac{\sin^2\delta}{n_e^2}. \tag{4.75}$$

With the solution of Equation 4.73 in hand, though, we can solve for the extraordinary polarization as

$$\hat{\mathbf{e}}_e = \aleph\left(\mathbf{k}_e \times \mathbf{k}_e \times \hat{\mathbf{c}}\right). \tag{4.76}$$

Now we would like to represent these polarization states by an angle $\phi_?$ as shown in Figure 4.24, where ? represents either e or o polarization. $\phi_?$ is zero when the polarization is in the plane of intersection, and can be found by

$$\phi_? = \sin^{-1}\left(\hat{\mathbf{e}}_i \cdot \aleph(\mathbf{k}_? \times \hat{\mathbf{z}})\right), \tag{4.77}$$

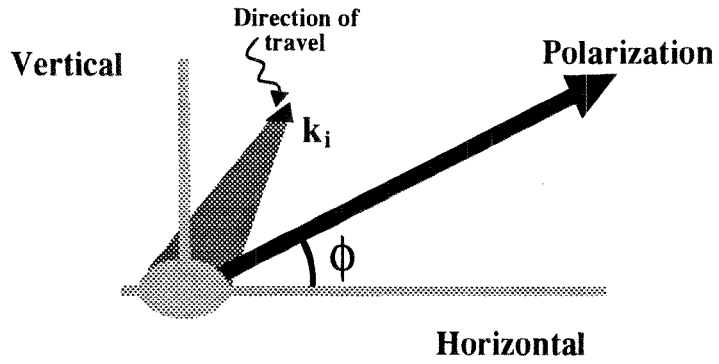


Figure 4.24: 3–D representation of polarization angle

where $\hat{\mathbf{z}}$ is the unit vector in the z direction. Note that if $\mathbf{k}_?$ is nearly parallel to $\hat{\mathbf{z}}$, then the cross–product should be replaced by $\hat{\mathbf{y}}$.

Now we have numbers for all of the terms in Figure 4.23. The only unknowns are the output electric fields (E_r, E_o, E_e) and the polarization state of the reflected light (ϕ_r). At the boundary, we match the tangential components of $E_?$ and the tangential components of $H_? = B_?/\mu$, resulting in

$$E_i^v + E_r^v = E_o^v - E_e^v \tag{4.78}$$

$$\frac{n_{ext}}{\mu_{ext}}\left(E_i^v - E_r^v\right)\cos\theta_i = \frac{n_o}{\mu_{int}}E_o^v\cos\theta_o - \frac{n_e}{\mu_{int}}E_e^v\cos\theta_e \tag{4.79}$$

$$E_i^h\cos\theta_i - E_r^h\cos\theta_i = E_o^h\cos\theta_o + E_e^h\cos\theta_e \tag{4.80}$$

$$\frac{n_{ext}}{\mu_{ext}}\left(E_i^h + E_r^h\right) = \frac{n_o}{\mu_{int}}E_o^h + \frac{n_e}{\mu_{int}}E_e^h, \tag{4.81}$$

where

$$\begin{aligned} E_?^h &= \cos\phi_? E_? \\ E_?^v &= \sin\phi_? E_?. \end{aligned} \tag{4.82}$$

These four equations can be solved for the four unknown variables: E_r, E_o, E_e, and ϕ_r. In writing these equations, we have made the approximation that $\hat{\mathbf{e}}$ (or polarization vector) for each wave is mutually perpendicular to the magnetic field and to the wave vector. In a birefringent material, this is not strictly true: the $\hat{\mathbf{d}}$ vector is mutually

perpendicular, and the $\hat{\mathbf{e}}$ vector "leans" slightly in the direction of the wavevector. For simplicity, we have ignored this difference between $\hat{\mathbf{d}}$ and $\hat{\mathbf{e}}$, so this treatment is not wholly rigorous.

Note that the $E_?^v$ are perpendicular to the incidence plane, yet are not guaranteed to be vertical (along the y axis). In essence, the plane shown in Figure 4.23 needs to be rotated about the normal until it contains the input wavevector (the axis labeled as x in Figure 4.24 then lies somewhere in the x–y plane

We can check the accuracy of these Fresnel coefficients by summing the reflectance and the two transmittances and verifying that they sum to one. These three terms are

$$T_o = E_o^2\, n_o\, \frac{\cos\theta_o}{\cos\theta_i} \tag{4.83}$$

$$T_e = E_e^2\, n_e(\delta)\, \frac{\cos\theta_e}{\cos\theta_i} \tag{4.84}$$

$$\begin{aligned} R = {} & (E_o \sin\phi_o - E_e \sin\phi_e - \sin\phi_i)^2 + \\ & (n_o E_o \cos\phi_o + n_e(\delta) E_e \cos\phi_e - \cos\phi_i)^2 . \end{aligned} \tag{4.85}$$

Once we trust that the refraction process is correctly modeled, we can proceed to use the angles and vectors we have derived. The effective electro–optic coefficient is simply

$$r_{eff} = \left| \left(\overline{\overline{rot}} \cdot \hat{\mathbf{e}}_{\mathbf{s_o}}\right) \cdot \left(\overline{\overline{r}} \cdot \left(\overline{\overline{rot}} \cdot \hat{\mathbf{u}}\right)\right) \cdot \left(\overline{\overline{rot}} \cdot \hat{\mathbf{e}}_{\mathbf{w_o}}\right) \right|, \tag{4.86}$$

where $\hat{\mathbf{u}}$ is a unit vector along the grating vector, $\overline{\overline{r}}$ is the tensor which describes the electro–optic effect in $LiNbO_3$,

$$\begin{matrix} r_{13} & 0 & 0 \\ -r_{22} & 0 & 0 \\ 0 & -r_{22} & r_{42} \end{matrix}$$

$$\begin{matrix} 0 & r_{13} & 0 \\ 0 & r_{22} & r_{42} \\ -r_{22} & 0 & 0 \end{matrix}$$

$$\begin{matrix} 0 & 0 & r_{33} \\ 0 & r_{42} & 0 \\ r_{42} & 0 & 0 \end{matrix}$$

and $\overline{\overline{rot}}$ is a rotation matrix which rotates vectors from the coordinate system of the beams (where z is the direction along which the signal propagates) to the crystal coordinate system (where z is the direction along the **c**–axis. For the 45° cut, this is

$$\frac{1}{2}\begin{pmatrix} \sqrt{2} & 0 & \sqrt{2} \\ 0 & 2 & 0 \\ -\sqrt{2} & 0 & \sqrt{2} \end{pmatrix}$$

if the **b**–axis of the crystal and the y axis are assumed to be identical. The other assumption is that the **a**–axis corresponds to y, but there is little difference in the results, so I do not write that rotation matrix here. The angle ϕ_G between the **c**–axis and the grating vector is defined as

$$\phi_G = \cos^{-1} \hat{\mathbf{c}} \cdot \hat{\mathbf{u}}, \tag{4.87}$$

and affects the effective dielectric constant and photovoltaic field.

Because both beams split into ordinary and extraordinary, there exist the possibility for several gratings to be formed [350]. These will have varying modulation depths, both because of the amount of power in each eigenpolarization, but also because of

the dot–product between the polarizations. For the 90° geometry with normally incident beams, the ordinary–ordinary grating is the strongest grating. We explore the variation of this grating as a function of changes in several incidence angles:

Angle		Origin	Expected range
Reference	Horizontal	Angle–multiplexing	-10° to 10°
	Vertical	Fractal–multiplexing	-16° to 16°
Signal	Horizontal	SLM pixel columns	-5° to 5°
	Vertical	SLM pixel rows	-5° to 5°

These effects are shown in Figures 4.25 and 4.26. Figure 4.25 shows the variation in $M/\#$ as the external incidence angle of the reference beam varies both horizontally and vertically. This corresponds to the variations encountered during angle and fractal multiplexing over a large range of angles. Again, these correspond to holograms recorded between plane waves. The asymmetry observed for negative fractal angles comes from a small (5%) contribution of the r_{22} electro–optic coefficient. The asymmetry is reversed if we assume that the a–axis of the crystal corresponds to the y dimension. Figure 4.26 shows the variation in $M/\#$ as the external incidence angle of the signal beam changes, which corresponds to different SLM pixels (assuming Fourier plane storage).

The other gratings (besides that formed between the two ordinary beams) are not large enough to have a significant effect—the largest is the grating written between the ordinary signal beam and the extraordinary reference beam. The only reason this can occur at all is because these polarizations are not exactly orthogonal. For the angle ranges mentioned above, the largest modulation depth of the ordinary—extraordinary grating is 3% of the ordinary–ordinary grating.

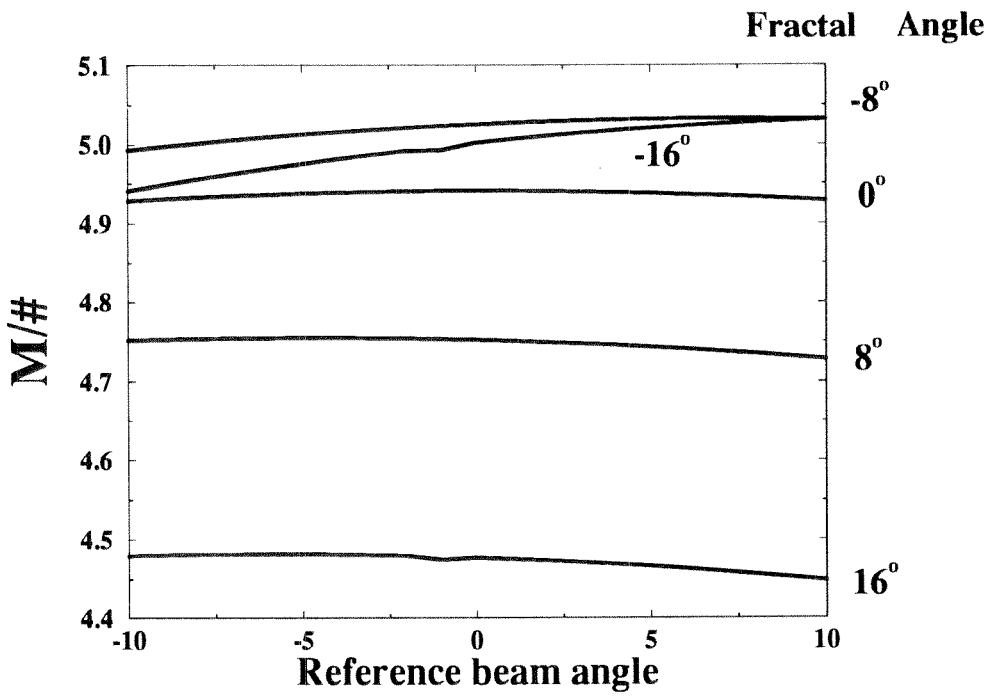


Figure 4.25: $M/\#$ vs. (horizontal) reference beam angle, for various fractal (vertical) incidence angles. Signal beam is normally incident.

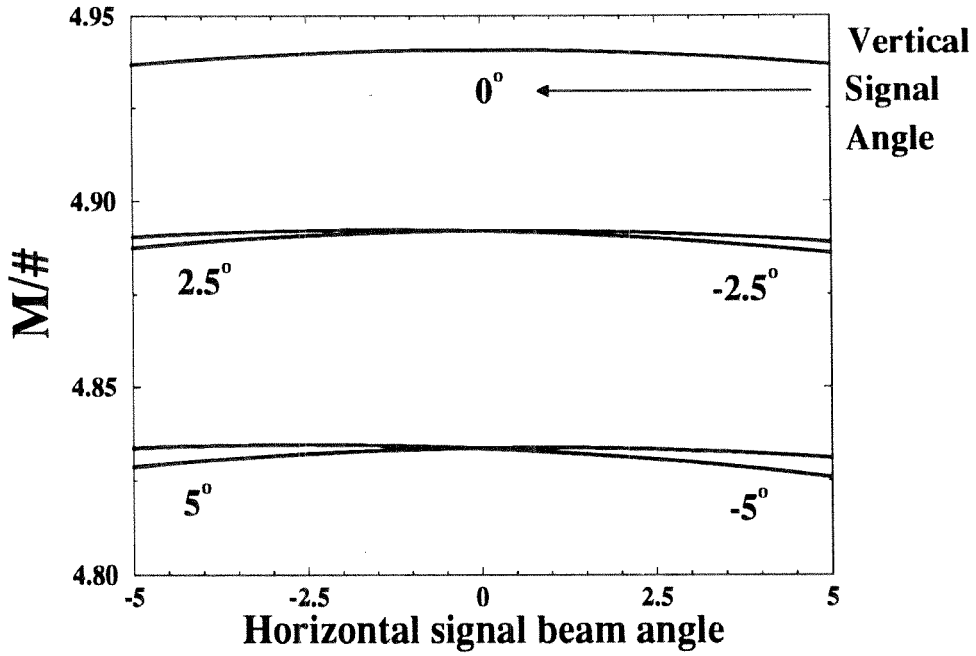


Figure 4.26: $M/\#$ vs. horizontal signal beam angle, for various vertical signal beam angles. Reference beam is normally incident.

4.4 Appendix: Derivation of the Born approximation

In this section, we derive the form of the Born approximation which we used in our $M/\#$ theory (Equation 4.1). Here we show the derivation from Prof. Psaltis' class notes [268], which start from Maxwell's equations. We choose to repeat this derivation[26] since the reference is not widely available outside Caltech, yet the result is a fundamental starting point for the results shown in this chapter. To save space, we take up the derivation at the wave equation

4.4.1 The wave equation

Our starting point is a volume containing a real, scalar permittivity modulation $\Delta\epsilon(\mathbf{r}')$, but with no charge, no conductivity, and no current ($\rho = \sigma = J = 0$). We will also assume that the volume is isotropic, with real, scalar permittivity ϵ_0 and permeability μ_0. $\mathbf{r}'$ refers to coordinates inside the crystal while $\mathbf{r}$ to coordinates outside the crystal (where the scattered field is observed). Our purpose is to determine the scattered field at point $\mathbf{r}$ given an incident reference plane wave, $e^{j\mathbf{k}_i\cdot\mathbf{r}'}$ with some polarization $\hat{e}$.

The wave equation for electric field, which follows directly from Maxwell's equations and a vector identity, is

$$\nabla^2\mathbf{E} + \omega^2\mu_0\epsilon\mathbf{E} = \nabla(\nabla\cdot\mathbf{E}), \tag{4.88}$$

where

$$\epsilon = \epsilon_0 + \Delta\epsilon(\mathbf{r}'). \tag{4.89}$$

[26] I stress here that this appendix is not original work by the thesis author.

Because $\nabla \cdot \mathbf{D} = 0$, we can write

$$\nabla \cdot \mathbf{E} = -\frac{\nabla\epsilon \cdot \mathbf{E}}{\epsilon} \approx -\frac{\nabla\epsilon \cdot \mathbf{E}}{\epsilon_0}. \tag{4.90}$$

Here we have assumed that the modulation on the permittivity, which we can write as

$$\Delta\epsilon(\mathbf{r}') = \Delta\tilde{\epsilon}(\mathbf{r}')e^{j\mathbf{K}_g\cdot\mathbf{r}}, \tag{4.91}$$

is relatively small in amplitude and bandwidth.[27] As a result,

$$\frac{\nabla\epsilon \cdot \mathbf{E}}{\epsilon_0} \approx \frac{j\mathbf{K}_g \cdot \mathbf{E}}{\epsilon_0}\Delta\epsilon, \tag{4.92}$$

and the wave equation becomes

$$\nabla^2\mathbf{E} + \omega^2\mu_0\epsilon_0\mathbf{E} = -\frac{j\mathbf{K}_g \cdot \mathbf{E}}{\epsilon_0}\Delta\epsilon - \omega^2\mu_0\Delta\epsilon\mathbf{E}. \tag{4.93}$$

The next step is to decompose the total electric field into two parts: the incident reference beam $\mathbf{E}_i$, and the scattered field $\mathbf{E}_d$. These two terms are substituted into Equation 4.93. Simplifications to this wave equation can be made by recognizing that the scattered field is expected to be much smaller than the incident field ($\mathbf{E}_d \ll \mathbf{E}_i$), and that the incident reference wave is already a solution to the homogeneous wave equation:

$$\nabla^2\mathbf{E}_i + \omega^2\mu_0\epsilon_0\mathbf{E}_i = 0. \tag{4.94}$$

After dropping terms which are second–order in $\Delta\epsilon$ and reusing the small modulation assumption, we arrive at

$$\nabla^2\mathbf{E}_d + \omega^2\mu_0\epsilon_0\mathbf{E}_d = \mathbf{S}\ \Delta\epsilon\ e^{j\mathbf{k}_i\cdot\mathbf{r}'}, \tag{4.95}$$

[27]The largest spatial frequency component in $\Delta\tilde{\epsilon}(\mathbf{r}')$ is much smaller than $\mathbf{K}_g$.

where

$$\mathbf{S} = \frac{\mathbf{K}_g \cdot \hat{e}_i}{\epsilon_0}(\mathbf{k}_i + \mathbf{K}_g) - \omega^2 \mu_0 \hat{\mathbf{e}}_i. \tag{4.96}$$

$\mathbf{S}$ is a vector which determines the polarization of the diffracted light given the incident polarization $\hat{\mathbf{e}}_i$.

Equation 4.95 represents the contribution to the scattered field from any point within the scattering volume. Since the Born approximation is linear (as we pointed out in Section 4.1.1), the total scattered field is the superposition of these local contributions,

$$E_{d_x}(x,y) = \iiint \Delta\epsilon(\mathbf{r}')\ S_x\ e^{j\mathbf{k}_i\cdot\mathbf{r}'} h(\mathbf{r}',\mathbf{r}) d\mathbf{r}', \tag{4.97}$$

where $h(\mathbf{r}',\mathbf{r})$ is the impulse response for the scattering, which we expect will be a spherical wave.

4.4.2 The impulse response

We need to find the h which satisfies the Helmholtz equation:

$$\nabla^2 h + \omega^2 \mu_0 \epsilon_0 h = \delta(x-x')\,\delta(y-y')\,\delta(z-z'). \tag{4.98}$$

We start by writing this equation in terms of the transfer function

$$h(x,y,z) = \frac{1}{4\pi^2} \iint H(k_{d_x}, k_{d_y}, z) e^{j(k_{d_x}+k_{d_y})} dk_{d_x} dk_{d_y}. \tag{4.99}$$

Substituting this into Equation 4.98, we get

$$\begin{aligned} \iint \left[-(k_{d_x}^2 + k_{d_y}^2)H + \frac{\partial^2 H}{\partial z^2} + k^2 H \right] e^{j(k_{d_x}x + k_{d_y}y)} dk_{d_x} dk_{d_y} \\ = \delta(z-z') \iint e^{j[k_{d_x}(x-x') + k_{d_y}(y-y')]} dk_{d_x} dk_{d_y}. \end{aligned} \tag{4.100}$$

This equation will be satisfied if the Fourier transforms of both sides are equal, or

$$\frac{\partial^2 H}{\partial z^2} + (k^2 - k_{d_x}^2 + k_{d_y}^2)H = \delta(z - z')e^{-j(k_{d_x}x' + k_{d_y}y')}. \tag{4.101}$$

The solution to this differential equation is

$$H(k_{d_x}, k_{d_y}, z) = \begin{cases} \alpha e^{+j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}\,(z-z')} & z > z' \\ \alpha e^{-j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}\,(z-z')} & z < z' \end{cases}. \tag{4.102}$$

As expected, this is a spherical wave. The value of α can be obtained by examining the integral of Equation 4.101 over a very small interval near $z = z'$,

$$\left.\frac{\partial H}{\partial z}\right|_{z=z'+\Delta z} - \left.\frac{\partial H}{\partial z}\right|_{z=z'-\Delta z} + (k^2 - k_{d_x}^2 + k_{d_y}^2)\int_{z'-\Delta z}^{z'+\Delta z} H dz = e^{-j(k_{d_x}x' + k_{d_y}y')}. \tag{4.103}$$

Each of the first two terms of this equation give $j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}$, and the third term approaches zero as $\Delta z \to 0$, so that

$$\alpha = \frac{e^{-j(k_{d_x}x' + k_{d_y}y')}}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}}. \tag{4.104}$$

The transfer function is then

$$H(k_{d_x}, k_{d_y}, z) = \frac{e^{-j(k_{d_x}x' + k_{d_y}y')}}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}} e^{-j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}\,(z-z')} \tag{4.105}$$

and our final expression for the scattered field is

$$S_x \iint dk_{d_x} dk_{d_y} \underbrace{\left[\iiint \frac{\Delta\epsilon(\mathbf{r}')e^{j(\mathbf{k}_i\cdot\mathbf{r}')}e^{-j(\mathbf{k}_d\cdot\mathbf{r}')}d\mathbf{r}'}{2j\sqrt{k^2 - k_{d_x}^2 - k_{d_y}^2}}\right]}_{A(\mathbf{k}_i,\mathbf{k}_d)} e^{j(\mathbf{k}_d\cdot\mathbf{r})}. \tag{4.106}$$

The integral inside the square brackets corresponds to the $A(k_i, k_d)$ term that we use in Section 4.1.1. The only difference between Equation 4.106 and Equation 4.1 is the

incorporation of R_0, the amplitude of the incident field, which was assumed to be unity in this derivation.

4.5 Appendix: Solution of the Kukhtarev equations

In this section, we derive the expression for the first spatial harmonic of the space–charge field during its initial evolution.[28] Many theoretical models for the behavior of the photorefractive effect were introduced in the early to mid–1970's [80–93], but the Kukhtarev equations have been universally accepted as the general model for the operation of the photorefractive effect [94]. Many papers have been published which expand on the basic set of equations [95–103]. In particular, we refer to three almost simultaneously published papers [340,342,351] in which the local photovoltaic current is made proportional to the local absorber concentration, instead of to the bulk absorption coefficient. As mentioned in Section 4.1.3, this minor detail changes the solution of the equations markedly, permitting correct prediction of the size of the phase shifts and the spatial frequency response which had been observed in $LiNbO_3$:Fe [342]. Here we follow the notation of Reference [340].

We repeat the derivation of the space–charge field here for several reasons: First, we should write down the Kukhtarev equations somewhere, and in Section 4.1.3 they would have distracted us from our derivation of the $M/\#$. Secondly, it gives us some context for the assignment of numerical values in Section 4.1.7. Finally, the exact derivation including the photovoltaic effect is difficult to find. In Reference [340], only the initial equations and the final result were shown. The derivation itself was described only sketchily (it would have been a lengthy distraction in that paper). We show the full derivation here, using Reference [340] to set our starting point and to verify our results. The details of the derivation are the work of the thesis author, though (unlike Appendix 4.4).

[28] In other words, the regime where the space–charge field is too small to affect its own development.

4.5.1 The Kukhtarev equations

As mentioned in Section 4.1.3, there are four Kukhtarev equations:

$$\text{Rate equation for donor density} \qquad \frac{\partial N_D^+}{\partial t} = (sI + \beta)(N_D - N_D^+) - \gamma_R N_D^+ n, \tag{4.107}$$

$$\text{Continuity equation} \qquad \frac{\partial n}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{q}\frac{\partial J}{\partial x}, \tag{4.108}$$

$$\text{Current density} \qquad J = q\mu n E + k_B T \mu \frac{\partial n}{\partial x} - p(N_D - N_D^+)I, \tag{4.109}$$

$$\text{Poisson's equation} \qquad \epsilon \frac{\partial E}{\partial x} = q(N_D^+ - n - N_A), \tag{4.110}$$

and we have listed the definitions of the assorted variables in Table 4.9. We will consider these equations in one spatial dimension, x,[29] along the **c**–axis of the photorefractive crystal. This is fine with us because we are interested in gratings which vary mostly along the **c**–axis. To simplify matters, we ignore the effects of the thermal generation, β, and assume that the material is isotropic.[30]

Our next step is to linearize the equations, by writing many of the variables in terms of a DC component and a sinusoidal component of the same spatial frequency, K, as the intensity interference pattern. So if the intensity pattern illuminating the crystal is

$$I = I_0 + I_1 e^{jKx}, \tag{4.111}$$

[29] This is not the same x we were using in Section 4.1.3.

[30] The operation of the Kukhtarev's equations in an anisotropic medium is considered in Reference [250].

N_D	Density of dopants (Total Fe concentration)
N_D^+	Density of ionized dopants (Fe^{3+} concentration)
n	Density of electrons in the conduction band
N_A	Compensative acceptors[31]
I	Optical intensity in the crystal
E	Total electric field
J	Total current density
β	Thermal generation rate
s	Photoexcitation cross–section
γ_R	Carrier recombination rate
p	Photovoltaic constant[32]
μ	Carrier mobility[33]
ϵ	Dielectric constant of the crystal
q	Charge of the electron
k_B	Boltzmann's constant
T	Temperature (in Kelvin)
t	Time
x	Distance along the **c**–axis of the crystal

Table 4.9: Variables from Equations 4.107—4.110

then we expect the following variables to have DC and first–order components:

$$E = E_0 + E_1 e^{jKx}, \tag{4.112}$$

$$N_D^+ = N_{D0}^+ + N_{D1}^+ e^{jKx}, \tag{4.113}$$

$$n = n_0 + n_1 e^{jKx}, \tag{4.114}$$

$$J = J_0 + J_1 e^{jKx}. \tag{4.115}$$

Note that E_0 can either be an externally applied electric field (not common in the 90° geometry), or an internal field which builds up over the illuminated region during operation of the photovoltaic field.

When we insert these substitutions into the Kukhtarev equations, we can separate

[31]Note that this species does not participate in the photorefractive effect—it just provides charge neutrality. This is the variable we can change with oxidation and reduction of $LiNbO_3$:Fe.

[32]This is **not** the same as the Glass constant, κ. See Section 4.1.7 for the relation between p and κ.

[33]This is rather low in photorefractive crystals, which is what makes the whole thing work.

the equations into two sets: DC terms and first–order terms. There will also be some second–order terms (in e^{2jKx}) which we will drop. The plan is to solve the simpler DC equations and then use the results to simplify the more involved first–order equations.

4.5.2 The DC response

To start, we note that the change of the DC ionized donor level (ΔN_{D0}^{+}) is identical to the change in the DC density of conduction electrons (Δn_0). This is because the DC current density (J_0) has no spatial dependence. But, because $\gamma_R N_{D0}^{+}$ is large, an equilibrium between n_0 and N_{D0}^{+} is reached very quickly for

$$n_0 = \Delta(N_{D0}^{+}) \ll N_{D0}^{+}. \tag{4.116}$$

Let's show this in a little more detail. The rate equation for the DC conduction electrons is

$$\begin{aligned} \frac{\partial n_0}{\partial t} &= \frac{\partial N_{D0}^{+}}{\partial t} && (4.117) \\ &= sI_0(N_D - N_{D0}^{+}) - \gamma_R N_{D0}^{+} n_0, \\ &= \gamma_R N_{D0}^{+}\left[\frac{sI_0(N_D - N_{D0}^{+})}{\gamma_R N_{D0}^{+}} - n_0\right], && (4.118) \end{aligned}$$

so that the solution for n_0 is

$$n_0 = \frac{sI_0(N_D - N_{D0}^{+})}{\gamma_R N_{D0}^{+}}\left(1 - e^{-t(\gamma_R N_{D0}^{+})}\right). \tag{4.119}$$

Now we use the fact that $\gamma_R N_{D0}^{+}$ is very large, both in an absolute sense and relative to the excitation rate $sI_0(N_D - N_{D0}^{+})$. So n_0 reaches a very small value relative to N_{D0}^{+} and it reaches it rapidly. Now we can go to the DC version of Poisson's equation and make some simplifications. Since the DC electric field has no spatial variation, we expect that there will be charge balance between n_0, N_A, and N_{D0}^{+}. But we just

concluded that n_0 is very small, so we can approximate

$$N_{D0}^{+} \approx N_A, \tag{4.120}$$

and rewrite the DC conduction electron density as

$$n_0 = \frac{sI_0(N_D - N_A)}{\gamma_R N_A}. \tag{4.121}$$

Although we aren't going to use it right now, we should note that there is a DC current density,

$$J_0 = q\mu n_0 E_0 - p(N_D - N_A)I_0. \tag{4.122}$$

This doesn't seem particularly interesting, until you recall that this current density exists only in the illuminated portion of the crystal. When transported electrons get to the boundary of this illuminated region, they are trapped there and aren't reexcited. Likewise, more and more of the donor sites at the other boundary are ionized and no electrons recaptured, creating a net positive charge. As the illumination continues, this boundary charge layer creates a growing electric field, which is essentially an applied DC electric field. As expected, this electric field opposes the transport of charge which creates it, leading to an equilibrium condition. This will affect the creation of holograms, the $M/\#$, and the fidelity of our reconstructed holograms as detailed in Section 4.3.1.

4.5.3 The spatially modulated response

Now we return to the Kukhtarev equations and consider the first–order terms, using the additional knowledge from the DC solution. We will continue to write n_0 for brevity, remembering that we know exactly what it is (Equation 4.121). The resulting

equations are

$$\frac{\partial N_{D1}^{+}}{\partial t} = sI_1(N_D - N_A) - sI_0 N_{D1}^{+} - \gamma_R N_{D1}^{+} n_0 - \gamma_R N_A n_1, \quad (4.123)$$

$$\frac{\partial n_1}{\partial t} = \frac{\partial N_{D1}^{+}}{\partial t} + j\frac{K}{q} J_1 \quad (4.124)$$

$$J_1 = q\mu n_0 E_1 + q\mu n_1 E_0 + jk_B T\mu K n_1 - p(N_D - N_A)I_1 + pN_{D1}^{+} I_0 \quad (4.125)$$

$$j\epsilon K E_1 = q(N_{D1}^{+} - n_1), \quad (4.126)$$

where we have dropped the spatial variation (e^{jKx}) from both sides of the equation after taking spatial derivatives. We now insert Equations 4.126 and 4.125 into Equation 4.124 to get two equations:

$$\frac{\partial N_{D1}^{+}}{\partial t} = (N_D - N_A)sI_1 - N_{D1}^{+}(sI_0 + \gamma_R n_0) - \gamma_R N_A n_1, \quad (4.127)$$

$$\begin{aligned} \frac{\partial n_1}{\partial t} - \frac{\partial N_{D1}^{+}}{\partial t} &= j\frac{K}{q}\left[q\mu n_0\left(\frac{jq}{\epsilon K}(n_1 - N_{D1}^{+})\right) + q\mu n_1 E_0 + jk_B T\mu K n_1\right. \\ &\quad \left. - p(N_D - N_A)I_1 + pN_{D1}^{+} I_0,\right] \\ &= n_1\left[jK\mu E_0 - \frac{\mu k_B T K^2}{q} - \frac{q\mu n_0}{\epsilon}\right] + N_{D1}^{+}\left[\frac{q\mu n_0}{\epsilon} + j\frac{pKI_0}{q}\right] \\ &\quad - j\frac{pK}{q}(N_D - N_A)I_1. \end{aligned} \quad (4.128)$$

To solve these two equations, we substitute one into the other and drop the second–order derivative to get a first–order differential equation. First we redefine the above two equations so we can see what's going on without getting lost in all the terms:

$$\begin{aligned} \frac{\partial X}{\partial t} &= A - BX - CY \\ &= A - (B + C)X - C(Y - X) \end{aligned} \quad (4.129)$$

$$\begin{aligned} \frac{\partial (Y - X)}{\partial t} &= DY + EX + F \\ &= D(Y - X) + (E + D)X - F, \end{aligned} \quad (4.130)$$

where our two variables are

$$\begin{aligned} X &= N_{D1}^{+} \\ Y - X &= n_1 - N_{D1}^{+}, \end{aligned} \tag{4.131}$$

and our constant coefficients are

$$\begin{aligned} A &= (N_D - N_A) s I_1 \\ B &= s I_0 + \gamma_R n_0 \\ C &= \gamma_R N_A \\ D &= jK\mu E_0 - \frac{\mu k_B T K^2}{q} - \frac{q\mu n_0}{\epsilon} \\ E &= \frac{q\mu n_0}{\epsilon} + j\frac{pKI_0}{q} \\ F &= j\frac{pK}{q}(N_D - N_A) I_1. \end{aligned} \tag{4.132}$$

We construct a single differential equation by rewriting Equation 4.130 in terms of X,

$$X = \frac{1}{E+D}\left[\frac{\partial(Y-X)}{\partial t} - D(Y-X) + F\right] \tag{4.133}$$

and substituting into Equation 4.129. After dropping the second–order derivative,[35] the differential equation becomes first–order:

$$(B+C-D)\frac{\partial Y - X}{\partial t} + [CE - BD](Y-X) = A(E+D) - F(B+C). \tag{4.134}$$

This seems like an apt variable to solve for, because Equation 4.126 relates the first–order electric field (which is what we wanted) to this variable we have called $Y - X$

[35] We will discuss below what we have assumed by saying that $\frac{\partial^2(Y-X)}{\partial t^2} \approx 0$.

by

$$E_1 = \frac{-jq}{\epsilon K}(Y - X) \tag{4.135}$$

The solution of our differential equation in terms of electric field is then

$$E_1 = E_1 \mid_{t\to\infty} \left[1 - e^{-t/\tau'}\right], \tag{4.136}$$

where

$$E_1 \mid_{t\to\infty} = \frac{-jq}{\epsilon K} \frac{A(E+D) - F(B+C)}{CE - BD} \tag{4.137}$$

$$\tau' = \frac{B + C - D}{CE - BD}, \tag{4.138}$$

and τ' can be complex. To get these two items, all we need to do is substitute our definitions for A–F (Equation 4.132) and for n_0 (Equation 4.121) back in and simplify. In doing so, we will also cancel some negligible terms, making use of the fact that

$$\gamma_R N_A \gg s I_0 \tag{4.139}$$

and

$$N_A \gg n_0. \tag{4.140}$$

Let's deal with $E_1 \mid_{t\to\infty}$ and τ' in turn. The expression for the saturation space–charge field[36] becomes

$$E_1 \mid_{t\to\infty} = \frac{jq}{\epsilon K} \frac{I_1}{I_0} \frac{\frac{k_B T}{q} K + jE_0 + \frac{j}{q\mu s} p\gamma_R N_A}{\frac{q}{\epsilon K} + \frac{1}{N_D - N_A}\left[\frac{N_D}{N_A}\left(\frac{k_B T}{q} K + jE_0\right) + \frac{j}{q\mu s}\gamma_R N_A\right]}. \tag{4.141}$$

We need to make some definitions in order to make this mess a little more readable.

[36]This is not necessarily what the space–charge field will actually be *at* saturation, but is instead where it is headed (at time $t \sim 0$) under the assumptions we have made.

If we define

$$\begin{aligned} m &= \frac{I_1}{I_0} \\ E_D &= \frac{k_B T}{q} K \\ E_q &= \frac{q N_A (N_D - N_A)}{\epsilon K N_D} \\ E_{0ph} &= \frac{p \gamma_R N_A}{q \mu s}, \end{aligned} \tag{4.142}$$

then we can rewrite Equation 4.141 as

$$E_1 \mid_{t \to \infty} = E_q \frac{j E_D - (E_0 + E_{ph0})}{(E_q + E_D) + j \left(E_0 + \frac{N_A}{N_D} E_{0ph} \right)} \tag{4.143}$$

or[37]

$$|E_1|_{t \to \infty} = E_q \sqrt{\frac{E_D^2 + (E_0 + E_{ph0})^2}{(E_q + E_D)^2 + \left(E_0 + \frac{N_A}{N_D} E_{0ph} \right)^2}}. \tag{4.144}$$

The expression for time constant can be similarly expressed as

$$\frac{1}{\tau'} = \frac{1}{\tau_{di}} \frac{1 + \frac{E_D}{E_q} + j \left(\frac{E_0}{E_q} + \frac{N_A}{N_D} \frac{E_{0ph}}{E_q} \right)}{1 + \frac{E_D}{E_\mu} + j \frac{E_0}{E_\mu}}, \tag{4.145}$$

where we have used the shorthand variables defined above as well as two new ones:

$$\begin{aligned} \tau_{di} &= \frac{\epsilon}{q \mu} \frac{\gamma_R N_A}{s I_0 (N_D - N_A)} \\ E_\mu &= \frac{\gamma_R N_A}{\mu K}. \end{aligned} \tag{4.146}$$

[37] We only give the amplitude of the space–charge field here. We omit the phase shift term, which is very important if you want to consider two–beam coupling. You can find the value of this phase shift in Reference [340].

Since this time constant is complex, we can break it into real and imaginary parts, as

$$\frac{1}{\tau} \equiv \Re\left\{\frac{1}{\tau'}\right\} = \frac{1}{\tau_{di}} \frac{\left(1+\frac{E_D}{E_q}\right)\left(1+\frac{E_D}{E_\mu}\right) - \frac{E_0}{E_\mu E_q}\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right)}{\left(1+\frac{E_D}{E_\mu}\right)^2 + \left(\frac{E_0}{E_\mu}\right)^2}, \tag{4.147}$$

and

$$\omega \equiv \Im\left\{\frac{1}{\tau'}\right\} = \frac{1}{\tau_{di}} \frac{\frac{E_0}{E_q} - \frac{E_0}{E_\mu} + \frac{N_A}{N_D}\frac{E_{0ph}}{E_q}\left(1+\frac{E_D}{E_\mu}\right)}{\left(1+\frac{E_D}{E_\mu}\right)^2 + \left(\frac{E_0}{E_\mu}\right)^2}. \tag{4.148}$$

The only difference between Equations 4.144 and 4.145 displayed here and Equations 4.20–4.22 are that the DC electric field has been set to zero in the latter.

The final matter for this appendix is the validity of the assumption that the term $\frac{\partial^2 E_1}{\partial t^2}$ can be ignored. To justify this, we should check that the second derivative of our solution to the space–charge field is rather small.[38] This is equivalent to showing that the first derivative is mostly constant. We write again

$$E_1 = E_1 \Big|_{t\to\infty} \left[1 - e^{-t/\tau'}\right], \tag{4.149}$$

and expand the time dependent portion using the Taylor series for e^x. Dropping the constant term, we get

$$\begin{aligned} E_1 &\propto \left[1 - \left(1 - \frac{t}{\tau}(1+j\omega\tau) + \frac{t^2}{2!\tau^2}(1+j\omega\tau)^2 - \ldots\right)\right] \\ &\propto \frac{t}{\tau}(1+j\omega\tau) - \frac{t^2}{2!\tau^2}(1+j\omega\tau)^2 + \ldots \end{aligned} \tag{4.150}$$

Since ω and τ are of the same order of magnitude, the magnitude of the first derivative of E_1 is

$$\left|\frac{\partial E_1}{\partial t}\right| = \frac{\sqrt{2}}{\tau} - \frac{2t}{\tau^2} + (\text{terms involving } \frac{t^2}{\tau^3}) \tag{4.151}$$

[38] You can think of this in terms of a perturbative analysis: Will it make any difference if we plug our solution back in and solve for another "order"?

As long as $t \ll \tau$, the contribution of the second derivative is negligible. The first derivative is mostly constant, and we are justified in ignoring the second derivative.

Chapter 5 Holographic Random Access Memory: Design

Contents

5.1 Motivation

As we discussed in Chapter 1.4.2, spatial multiplexing can be used to increase the storage capacity of a holographic memory. Spatial multiplexing is quite different from the other multiplexing techniques because it's not limited by dynamic range—we're using a different part of the storage volume.[1] However, if we are using mechanical motion to bring the desired part of a large storage medium to where we can read it, our access time is going to be limited. In order to have random–access to stored holograms, we would like to access our holograms without translating the storage media.

The first to demonstrate non–mechanical spatial multiplexing was Huignard *et al* in the early 1970's [128, 129]. Their system used two cascaded AOD deflectors

[1] Spatioangular multiplexing blurs this distinction, making incremental use of new storage volume.

to supply the same 2–D spatial translation to both the reference and object arms. A reflection grating caused the beams to intersect at the crystal surface. Angle multiplexing at each location was performed by a mechanical rotating beam–shearing element placed in the reference arm. Using this arrangement, they were able to store 10 holograms at each of 25 locations. Recently, Hong et al. have demonstrated a compact system (< 520 cm^2 of optical tabletop) using 3 AODs, achieving random access to up to 1000 holograms at each of 20 vertically spaced locations [276]. In this system, the object beam cannot be deflected, so the crystal is mechanically translated for storage.

In general, non–mechanical wavefront controllers are angle scanners. As a result, we need a system which can perform angle and spatial multiplexing with angle scanners. One solution is the segmented mirror array [279, 282, 283] which we discuss in this chapter. This device requires a pair of crossed angle deflectors (such as AODs), and can perform angle multiplexing at a 2–D grid of locations [279], or angle and fractal multiplexing at a 1–D grid of locations [352, 353]. In this chapter, we outline the operation of the mirror array, and use it in the design of a large–scale random–access holographic memory.

5.2 Mirror Array

5.2.1 Theory of operation

The motivation behind the mirror array is the desire to use one angle deflector for angle multiplexing, and a co–located and orthogonal angle deflector for spatial multiplexing. The best place to start explaining how it works is the standard angle–multiplexing reference arm shown in Figure 5.1(a). The deflection created by an angle scanner is imaged to the surface of the crystal with a 4–F system. An angle change of the reference beam at the scanner becomes a position change in the center of the 4–F system. Likewise, an angle change in this center plane becomes a position change at the surface of the crystal—exactly what we wanted. We just have to figure

out how to induce an angle change in this center plane. Unfortunately, we can't put an angle deflector in this plane without seriously limiting the deflection range of the first deflector.

To see how the mirror array solves this dilemma, we note that we can place a fixed mirror in the center focal plane without affecting the operation of the 4–F system (Figure 5.1(b)). The reference beam is a focussed spot at the surface of the mirror.[2] To angle–multiplex holograms with the reference beam, the focused spot traces a horizontal path across the surface of the fixed mirror. If we were to remove all of the mirror surface above and below this horizontal path, leaving only a thin mirror strip, the operation of the system would be unchanged. We can then place additional mirror strips in this newly vacated region, and orient each with its own fixed deflection angle. The result is a vertically stacked array of long thin mirror strips which we use to create our desired angle change in the center plane.

Figures 5.2(a) and (b) diagram the operation of such a mirror array—Figure 5.2(a) shows selection of output location by the vertical angle scanner (AOD), while Figure 5.2(b) shows angle multiplexing at a spot by the horizontal AOD. The deflection angle of the vertical AOD determines which mirror strip will be illuminated. Each mirror strip is tilted in both dimensions in order to redirect incident light to one of the storage locations in the crystal. The horizontal angle of incidence is determined by the horizontal position of the focussed spot on the mirror strip(Figure 5.2(b)). In this way, the 2–D angle scanner selects the position and incidence angle of the reference beam at the crystal surface. In an actual system, we need to use a beamsplitter so that the surface of the mirror array can be exactly in the center focal plane. This allows us to minimize the vertical size of each mirror strip, yet still avoid crosstalk to other storage locations. Note also that the same mirror array can be used to combine spatial multiplexing with either wavelength or phase-code multiplexing.

Since current AOD technology provides SBP[3] on the order of 1000 or so, storage of up to 10,000 angle–multiplexed holograms at each location might be problematic.

[2]Yes, the plane of focus and the mirror surface are not exactly identical.

[3]See Chapter 1.6.4.

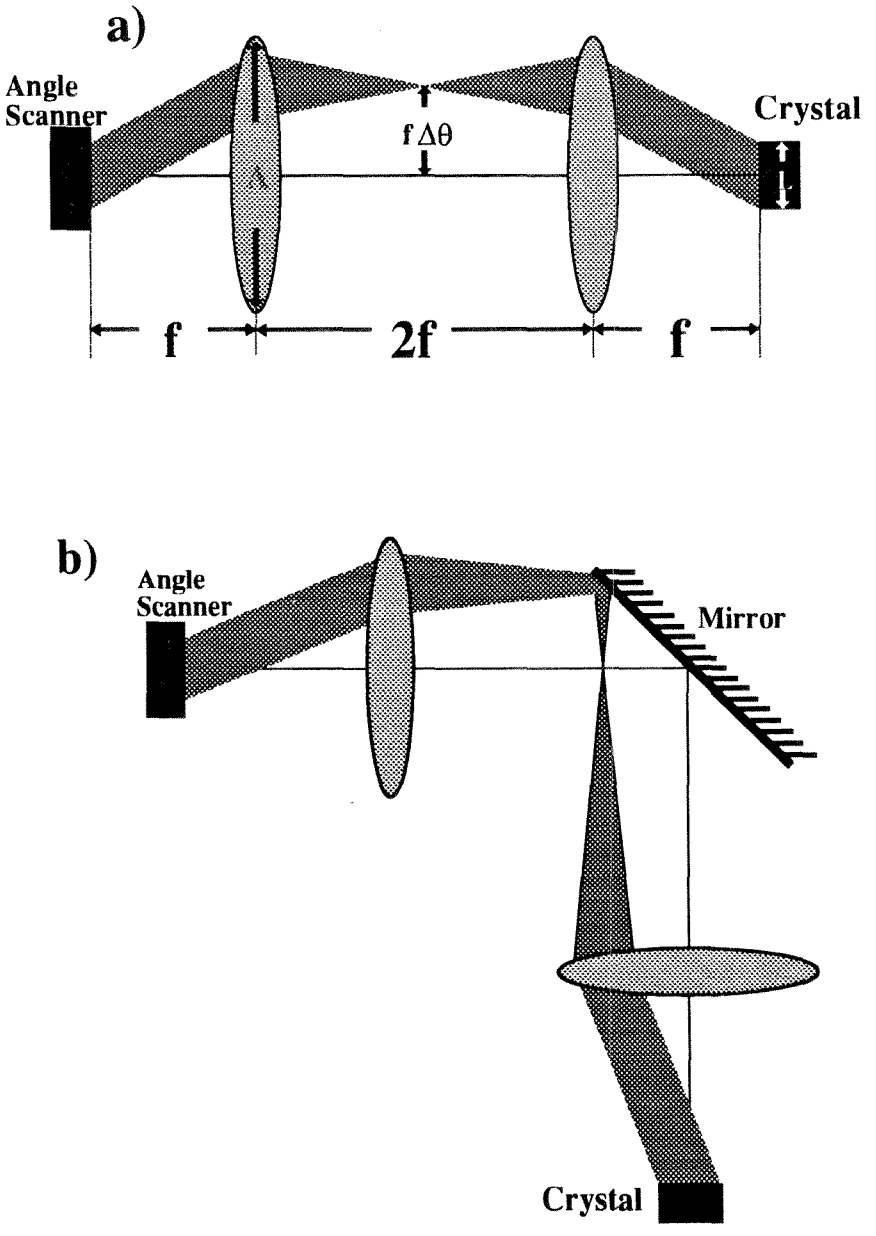


Figure 5.1: Angle–multiplexed holographic memory using an angle scanner.

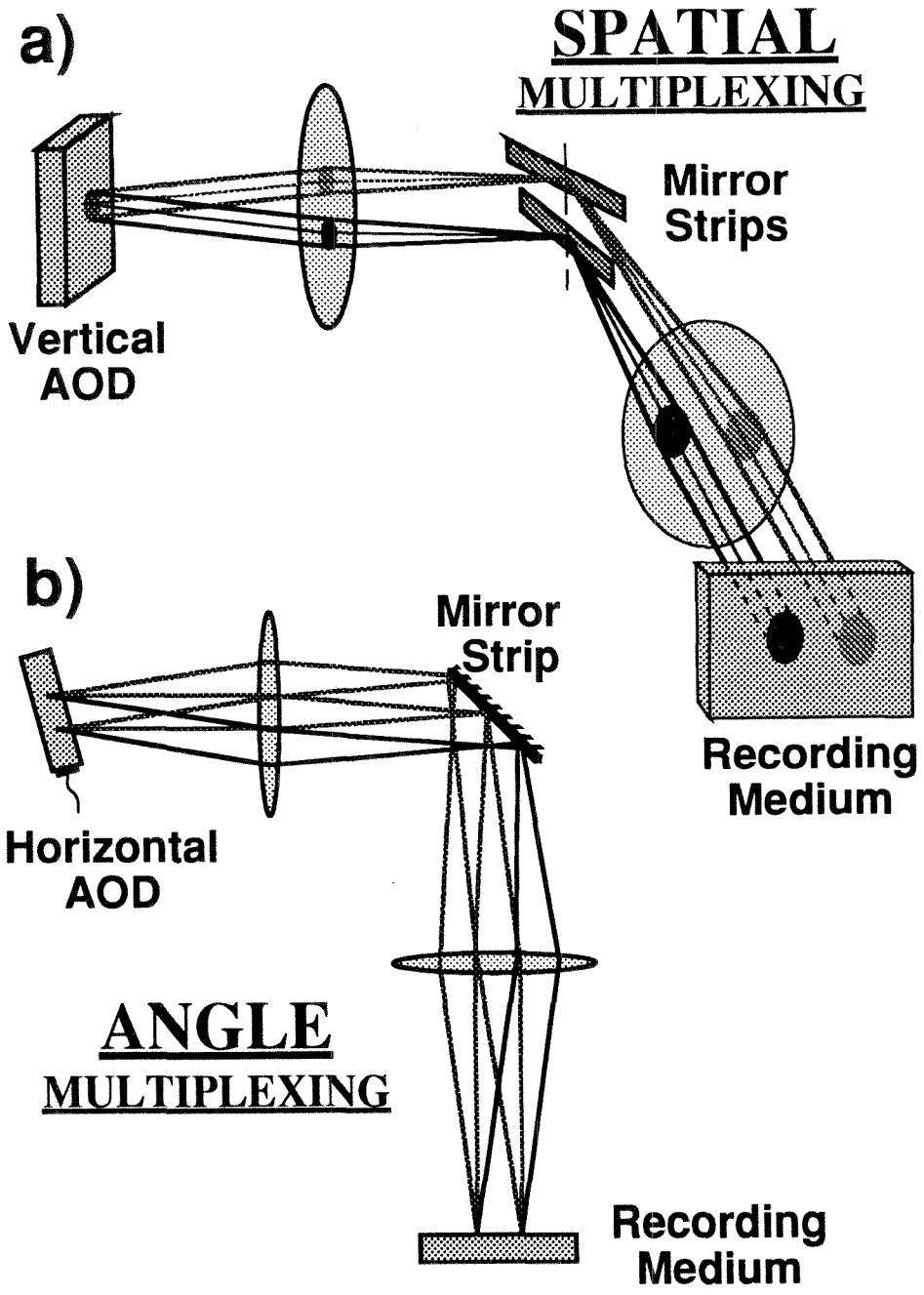


Figure 5.2: Operation of mirror array segments. (a) Selection of spatial location by vertical beam deflection. (b) Selection of incident reference beam angle by horizontal beam deflection.

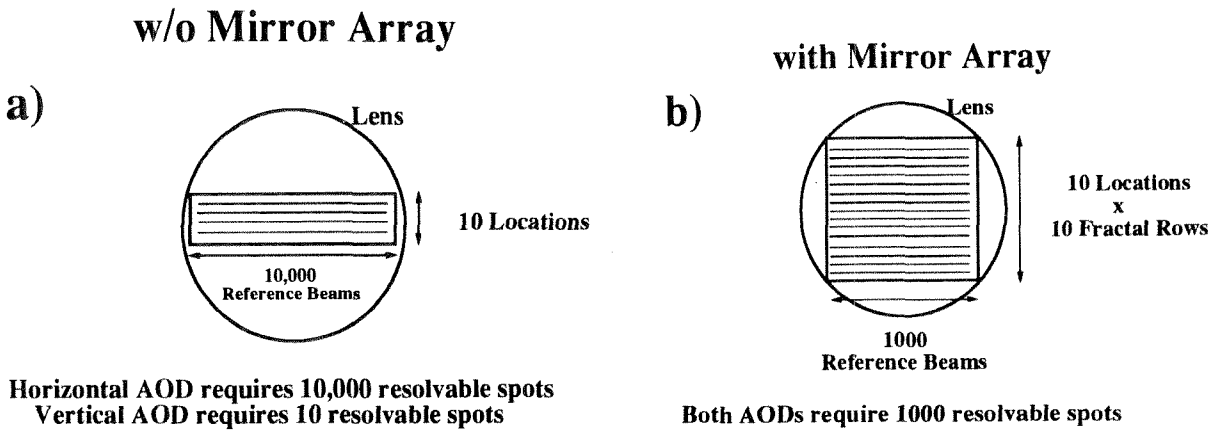


Figure 5.3: Space–bandwidth loading without and with mirror array

The horizontal AOD is overloaded by the large number of angles required, while the vertical AOD is underutilized. The SBP requirement also affects the size of the lens apertures needed in the reference arm. This is graphically shown in Figure 5.3(a), where we show the SBP load by indicating the portion of the lens aperture which is used. One solution for this asymmetric SBP loading is to use multiple mirror facets for each location. This reduces the number of angularly multiplexed holograms required per facet, dividing the SBP load evenly between the horizontal and vertical deflectors (Figure 5.3(b)). In using multiple mirror facets for a location, we are performing fractal multiplexing.

Figure 5.4 shows how fractal multiplexing[4] works in the context of the mirror array. Mirror strips 1 and 2 perform the expected spatial multiplexing at locations A and B. Mirror strip 3 is oriented parallel to strip 2, but displaced vertically. Beams deflected to strip 3 also arrive at storage location A, but with a different vertical incidence angle than those arriving via strip 2. If the vertical separation between strips 2 and 3 is larger than the vertical bandwidth of the images being stored, then each can be used to store and retrieve holograms.

[4]See Chapter 1.4.3 for more about fractal multiplexing.

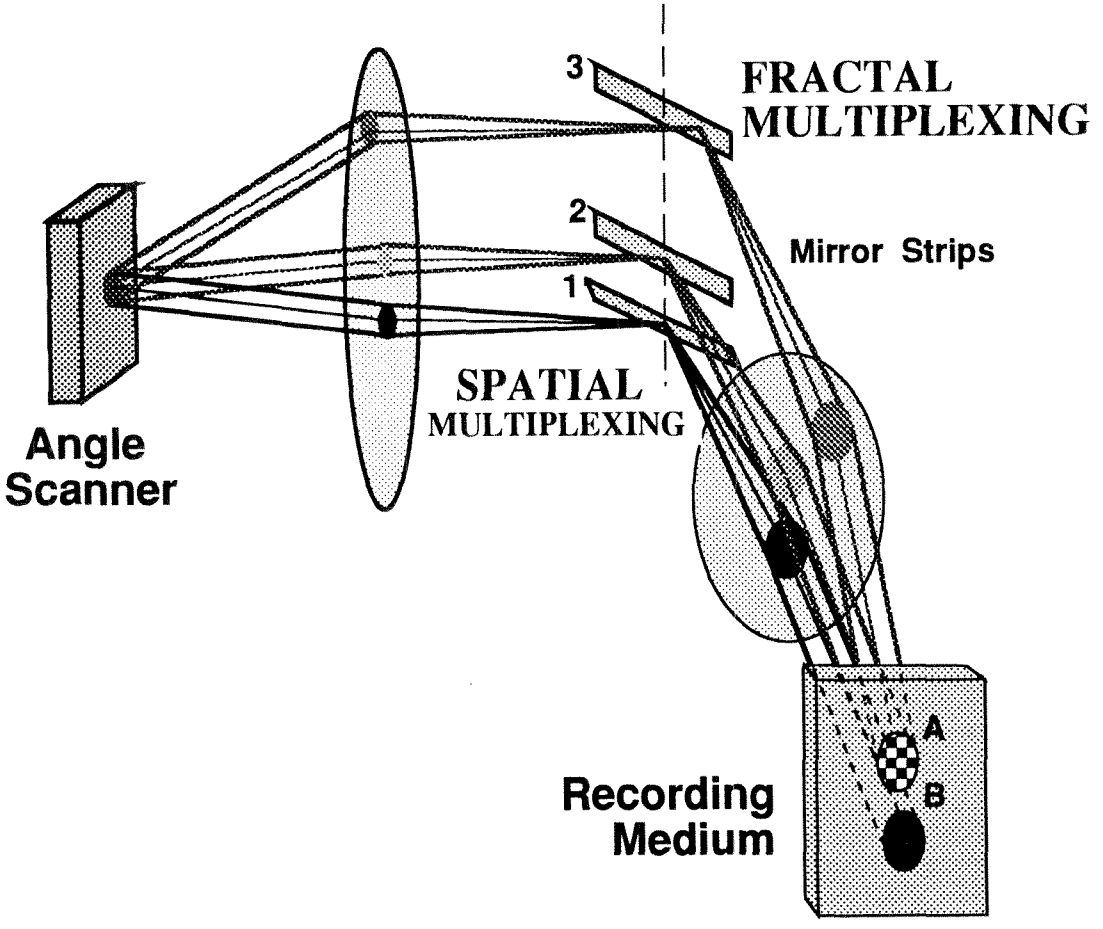


Figure 5.4: Implementation of fractal multiplexing with the mirror array.

5.2.2 First–generation device

The schematic of the mirror array is shown in Figure 5.5. There are 256 mirror strips—one for each location on the crystal. Each mirror strip is 150μm wide and 75mm long. The mirror array is composed of 16 "tiles," each of which has 16 mirror strips cut as grooves. The change in angle between tiles and between groove faces is 0.5°. The angular change between tiles is orthogonal to the increment between grooves, allowing tilting of individual facets in both directions. This mirror array was designed for spatial multiplexing over a 2–D grid of 16 $\times$ 16 locations [279]. We can use the same mirror array for fractal multiplexing at each of 16 locations by removing the tilt between the tiles. Each tile contains a set of 16 vertical tilts for 16 locations—and each location has 16 mirror strips "pointed" at it, one from each tile. The separation between identically tilted mirror strips is 16$\times$ the strip width, or 2.4mm.

A prototype mirror array was fabricated with standard blazed grating technology. This technique involves cutting grooves with a diamond tip in a gold surface on a brass substrate.[5] The groove angles are controlled by the tilt of the tip with respect to the substrate, while the groove width is set by the dimensions of the tip. An error in programming the diamond tip angles during fabrication of our prototype caused the actual change in angle between mirror strips to be 1°. The mirror array (each containing 16 mirror strips) is held together by two rods which pass through the tiles. Two holes were drilled on each tile for these rods—one in the same "reference" position, and one at a different offset on each tile. As a result, the rods hold the mirror strips like a fan with an angle change between each tile of 0.5°. We could make all the tiles parallel by pulling the second rod out and retightening the first—this is what we did to enable the fractal multiplexing over a 1–D grid. We show a photograph of the finished mirror array (aligned for the 2–D grid) in Figure 5.6.

[5] In retrospect, silver would have been a better choice since the reflectivity of gold is relatively poor in blue light.

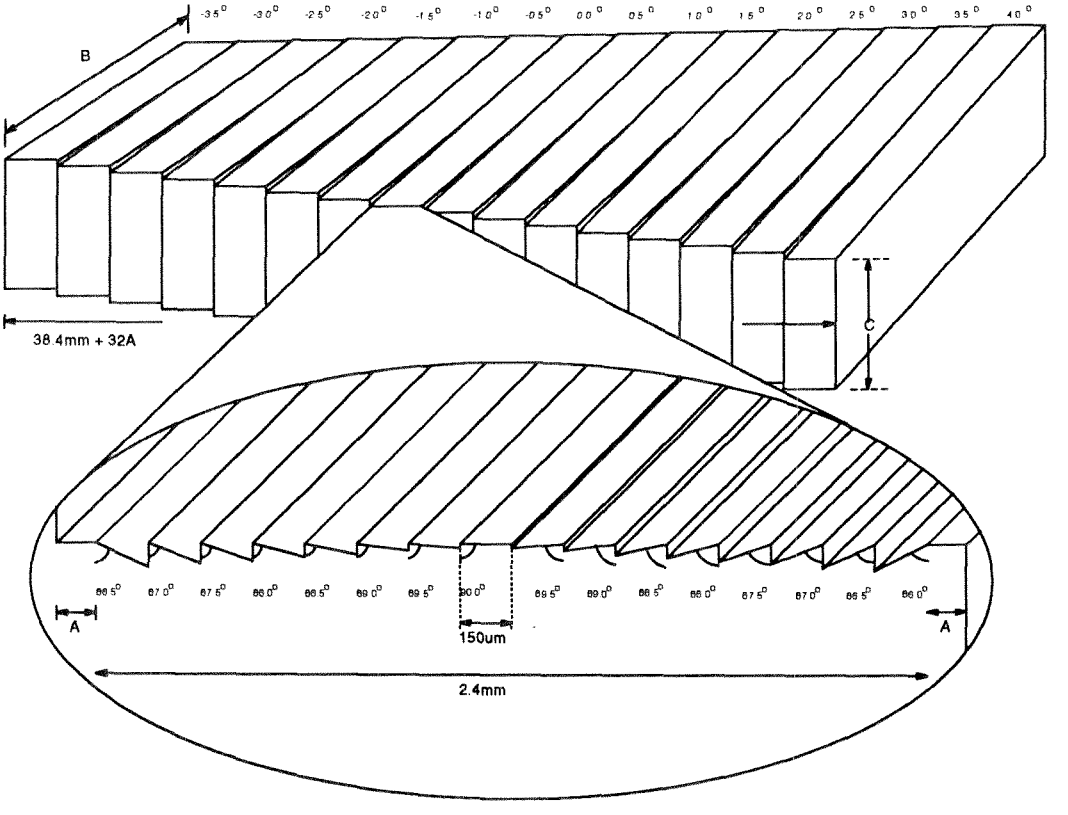


Figure 5.5: Mirror Array schematic.

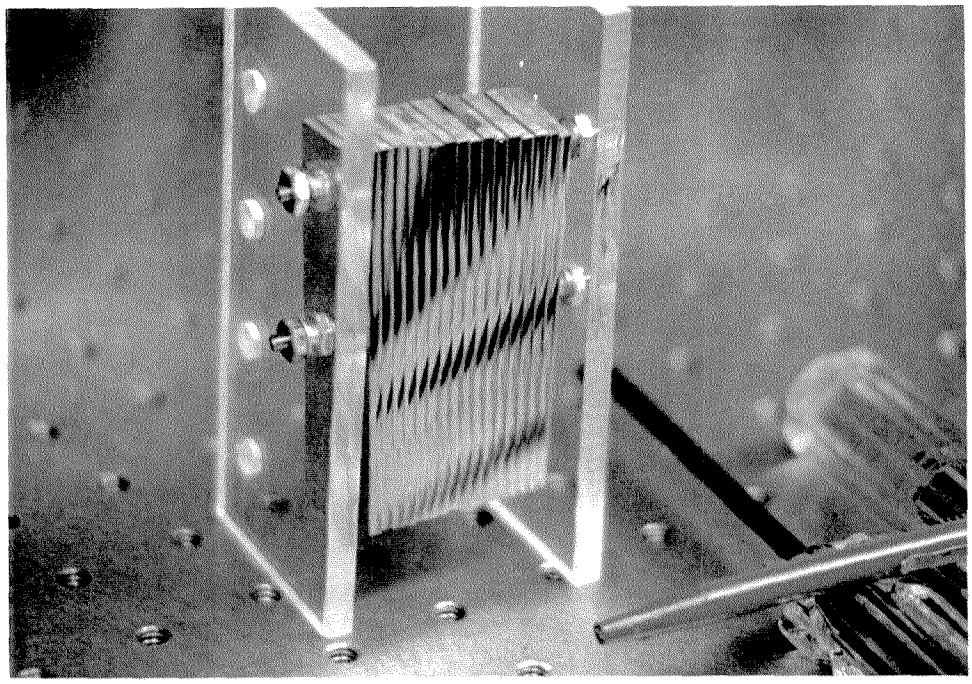


Figure 5.6: Mirror Array: Photograph

5.3 Memory Design

5.3.1 160,000 hologram system

We describe a page–formatted random–access holographic memory designed to store 160,000 holograms. The segmented mirror array allows rapid access to any of the stored holograms with a non–mechanical angle scanner. The memory consists of 16 vertically spaced locations, each containing 10,000 holograms. Each location is organized as 10 fractal–multiplexed rows of 1000 angularly–multiplexed holograms each. This assignment requires a horizontal SBP of 1000 and a vertical SBP of 160, comfortably within the capabilities of currently available AODs. The total storage capacity is 160 Gbits[6] and a random access time determined by the 2–D angle scanner. If this is a pair of crossed AODs, the access time can be less than 100μsec.

The system design is shown in Figure 5.7. A laser beam is split in two parts and

[6] assuming 1000 × 1000 pixels in the SLM.

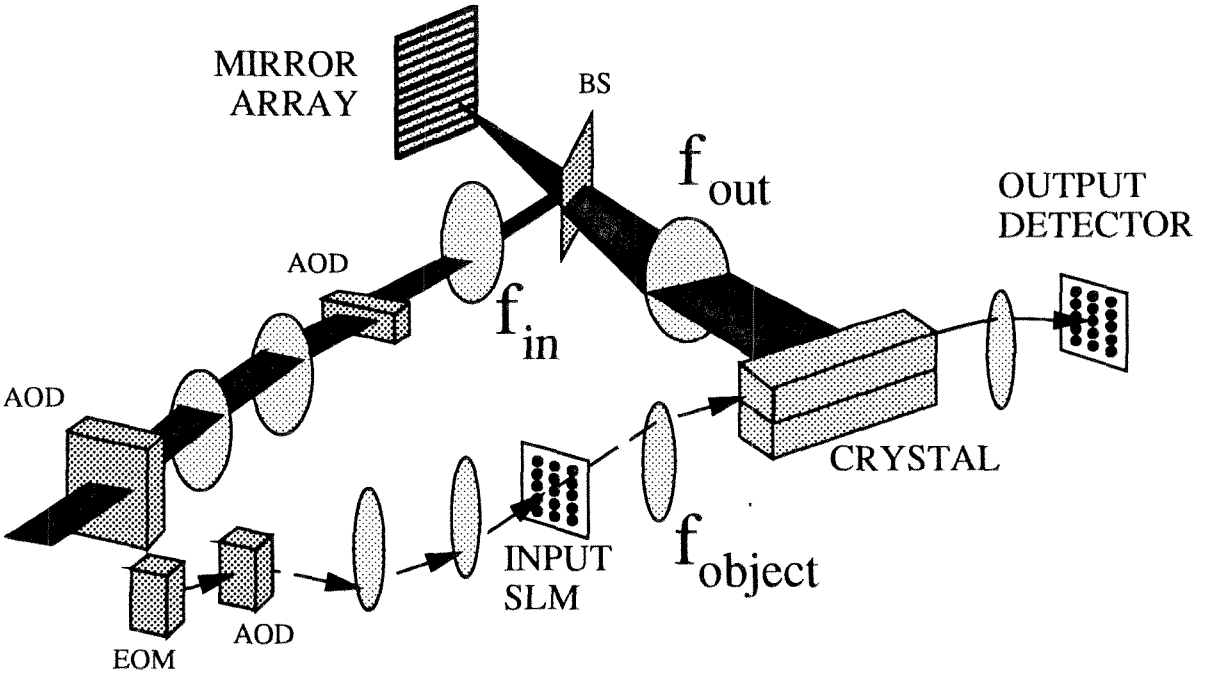


Figure 5.7: 160,000 hologram system

then brought together at a storage location within a stack of photorefractive crystals. The object beam is imprinted with the information displayed upon the input SLM, while the reference beam is a plane wave. A segmented mirror array and two crossed acousto–optic deflectors (AODs) allow the reference arm of the system to control both the position and angle of incidence of the reference beam. Another AOD is used to deflect the object beam to control the position of the information-bearing object beam on the crystals. The Doppler shift introduced by the AODs is compensated in the object arm by an electro–optic modulator (EOM), so that the interference pattern is stationary during storage.

We demonstrate the various elements of this design in Chapter 6. This includes

- storage using the mirror array,
- storage of 1000 holograms using an AOD,
- storage of 10,000 holograms at a single location,
- storage at each of the 16 locations using the mirror array, and

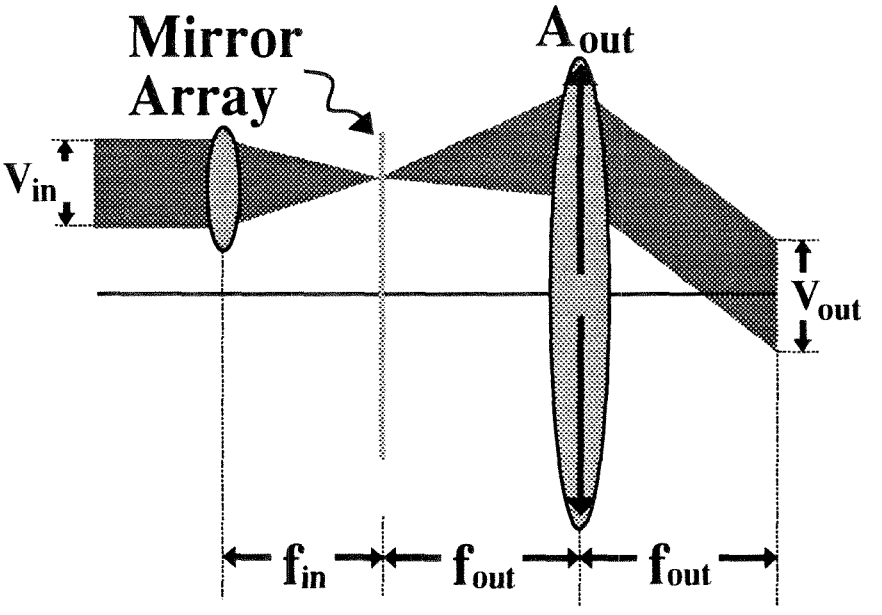


Figure 5.8: Reference beam with mirror array

- demonstration of 10,000 holograms at the top, center, and bottom locations.

In most of our experimental demonstrations, we use a mechanical scanner to deflect the focussed spot over the surface of the mirror array.

In the remainder of this chapter, we discuss the lens design issues involved with bringing the signal and reference beams to the crystal, and the reconstructed holograms to the detector array. We concentrate on the lens constraints involved with a mechanically accessed setup. We describe the design we have built in our laboratory, using the first–generation mirror array and commercially available lenses and discuss some of the drawbacks of this design.

5.3.2 Lens constraints

We start our analysis with the reference beam. As shown in Figure 5.8, a focussed spot illuminates the mirror array surface and is then directed through a lens to the crystal. Between the mirror array and the output lens, the beam passes through a circular quarter–wave plate and cube beamsplitter (not shown for simplicity). Figure 5.8 introduces several variables to describe the beam, the apertures, and the mirror array.

These include

A_{out}	Aperture of the output lens
f_{in}, f_{out}	Focal lengths of the two lenses
V_{in}, V_{out}	Vertical dimension of the plane wave
H_{in}, H_{out}	Horizontal dimension of the plane wave
d_{facet}	Mirror strip width (150μm in our device)
$\Delta\theta$	Vertical angle change between neighboring mirror strips (1° in our device)

Note that since we are scanning the input lens mechanically in the laboratory, we don't need to worry about its aperture.

We will first deal with the spatial and fractal multiplexing, and then see how the apertures of the output lens, beamsplitter, and waveplate affect the system. The output beam dimensions are set by the choice of focal lengths and the input beam dimensions as follows:

$$\begin{aligned} H_{out} &= \frac{f_{out}}{f_{in}} H_{in} \\ V_{out} &= \frac{f_{out}}{f_{in}} V_{in}. \end{aligned} \tag{5.1}$$

The input beam height V_{in} can't be too small, or the focussed spot will be larger than the mirror strip width. This constraint can be written as

$$d_{facet} > \frac{\lambda}{f_{in}} V_{in}. \tag{5.2}$$

The vertical spacing between storage locations must be larger than the beam dimension in order to keep the storage locations distinct, as shown in Figure 5.9. This calls for

$$V_{out} < 2\Delta\theta f_{out}. \tag{5.3}$$

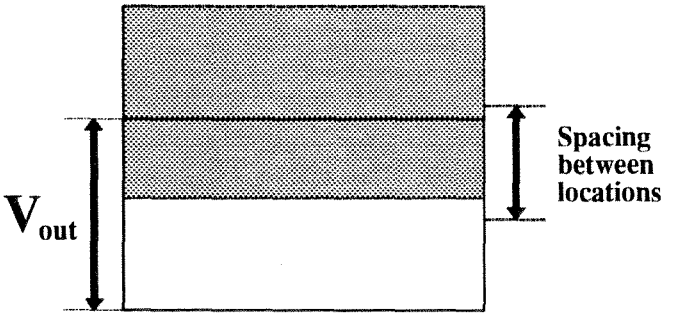


Figure 5.9: Isolation of spatially multiplexed locations

Now we need to deal with the interaction between the signal beam and the reference beam. The vertical extent of the stack of 16 locations is given by 16× the right side of Equation 5.3. The signal arm must be able to deliver the information–bearing beam to the crystal locations within this stack, and then return the reconstructions to the same on–axis array detector, all without noticeable distortion. We could write an expression here for the aperture of the object beam lens, f_{object}, but it turns out the off–axis imaging performance of this lens is a more stringent requirement. We will take this up in the next section.

We need to keep the reconstructions from the various fractal rows from overlapping at the detector array. In order to use two identically tilted mirror strips to store holograms, their vertical separation (expressed in terms of angle at the crystal surface) must be larger than the vertical bandwidth of the image. We can write this as

$$\frac{m \times \text{SLM height}}{f_{object}} < n \frac{16 \times d_{facet}}{f_{out}}, \tag{5.4}$$

where f_{object} the focal length of the Fourier transform lens in the object beam, m is the magnification in the object beam,[7] and n is the number of mirror tiles between mirror strips used for fractal multiplexing. As shown in Figure 5.10, when Equation 5.4 is not satisfied then multiple output images (from different fractal rows) overlap at the array detector.

[7]In other words, ($m\times$ SLM height) is the height of the image of the SLM at the front focal plane of lens f_{object}. See Figure 6.14.

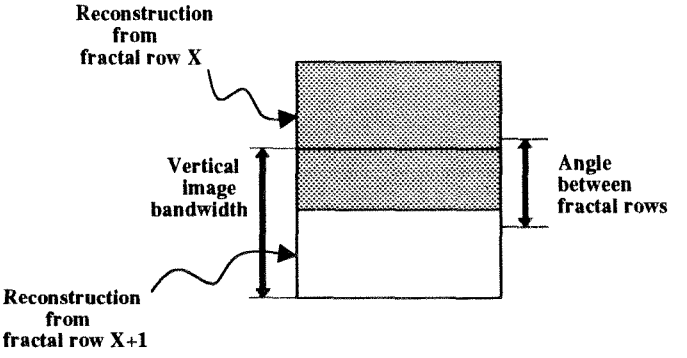


Figure 5.10: Separation of images from different fractal rows

The final consideration on the signal beam is that the height of the reference beam overlaps the Fourier transform of the information–bearing beam. If this condition is not met, then the hologram will be an incomplete copy of the signal beam. If the crystal is at the exact Fourier transform plane and there is incomplete overlap, then the output will be spatially filtered; if the crystal is not in this plane, portions of the SLM image may be missing. The object beam requirement can be written as

$$V_{out} > \frac{\lambda f_{object}}{\delta} = \frac{\lambda f_{object}(\#\text{ of SLM pixels})}{m(\text{SLM height})}, \tag{5.5}$$

where we have used the relationship between the number of SLM pixels, the SLM size at the image plane of lens f_{object}, and δ. Note that Equation 5.5 would imply that f_{object} ought to be small, but Equation 5.4 indicates that f_{object} should be large. There is a tradeoff between the number of fractal rows and the vertical bandwidth of the signal beam. This is equivalent to saying that the total vertical bandwidth of the object beam is fixed—no matter how you decide to divide it into SLM pixels and fractal rows, the product of the number of pixels and the number of fractal rows is unchanged.

Aperture limitations between the mirror array and the crystal can be caused by the output lens, the cube beamsplitter, or the quarter–wave plate (Figure 5.11). If the apertures are too small vertically, then the beam leaving a particular mirror strip does not make it to the crystal. Too small horizontally, and the angular multiplexing range is limited. In general, the two effects are coupled, and we begin to lose angular

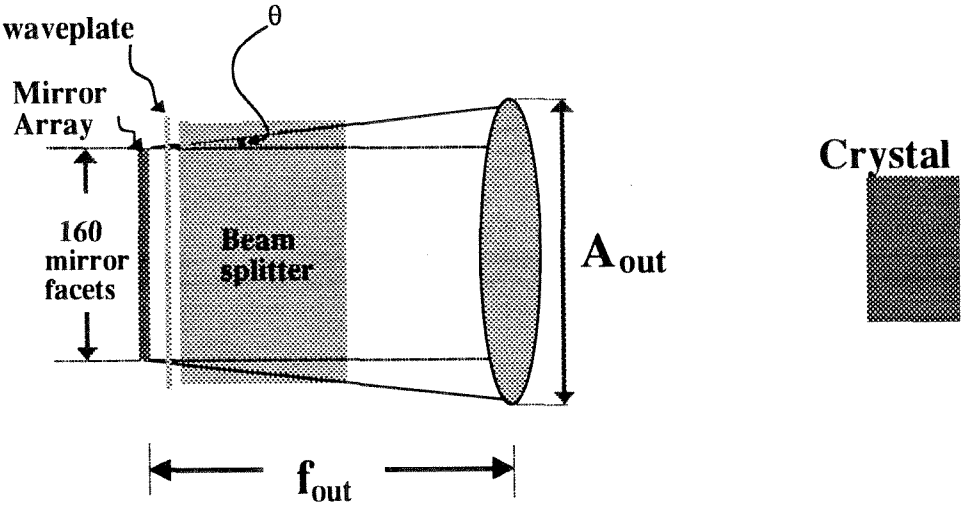


Figure 5.11: Apertures between the mirror array and the crystal

multiplexing range towards the top and bottom of the mirror array.

When determining if the beam will pass the aperture, there are three reasons that the reference beam might approach the edge of an aperture. The first is the initial position of the focussed point on the mirror strip. Vertically this is the mirror strip position, horizontally the multiplexing angle. The second contribution is the angular deflection caused by the mirror array, and the third is the non–zero size of the output beam at the aperture. We can write these as

$$\text{Vertical half-aperture} \;>\; y_{focus} + 2\,d\,\tan(k\,\delta\theta) + V_{out}\,\frac{d}{f_{out}} \tag{5.6}$$

$$\text{Horizontal half-aperture} \;>\; x_{focus} + 0 + H_{out}\,\frac{d}{f_{out}}, \tag{5.7}$$

where d is the distance of the aperture from the mirror array, x_{focus} & y_{focus} give the position of the focussed spot on the mirror array, and $k\,\delta\theta$ is the deflection angle of the mirror strip being illuminated (k= -8, -7, . . . , 6, 7). We can use these equations to map out which mirror strips are usable, and how much horizontal movement we can make on each mirror strip. We can express the result in terms of angle–multiplexing "zeroes." Note that Equations 5.6 and 5.7 are probably more restrictive than they could be. We consider only the range of horizontal angles for which the entire width of the reference beam passes the aperture. However, we can still use reference beams

which are partially occluded by the apertures. Since the interaction length is smaller, we will need to space holograms farther apart (in angle) at these large reference beam angles. But we should be able to get more angle–multiplexed holograms per mirror strip, if this should be necessary.

In the case of a circular aperture such as a waveplate or lens, the vertical aperture and horizontal aperture are not independent. The horizontal aperture of such an aperture will be a function of the y position at which the beam strikes, as

$$\text{Horizontal half-aperture} = \sqrt{\text{Radius}^2 - y^2}, \tag{5.8}$$

where the value of y is given by the right–hand side of Equation 5.6.[8]

Now we are in good shape: we know where the apertures are (d) and how big they are. Given a mirror strip (y_{focus} and k), we can determine whether it can direct reference beams to the crystal, and how much angle multiplexing range we will have with it. We will take up this issue in the next section.

5.3.3 Design of the mechanically accessed system

Our approach to designing a lens system is to pick a set of off–the–shelf lenses which satisfy Equations 5.1–5.5. Since we are mostly interested in demonstrating storage of holograms, we are willing to sacrifice a little on Equation 5.5 and lose some vertical spatial frequencies from our image. If we need, we can reallocate the pixels lost in this dimension to the horizontal dimension, since we have a lot of tolerance in this direction.[9] We then check our design to see if the object beam deflected to the top and bottom locations can be returned to the on–axis CCD detector without distortion. Because this requires precise knowledge of the aberration performance of the particular lenses to be used, we perform this step empirically. Then, once we have a set of lenses which allow access to 16 locations, we determine how many reference beam angles we have at each location.

[8] Of course, if this y value is outside the vertical aperture, then there's not much point in discussing how much horizontal deflection we can do.

[9] This assumes that we can redesign our SLM as desired.

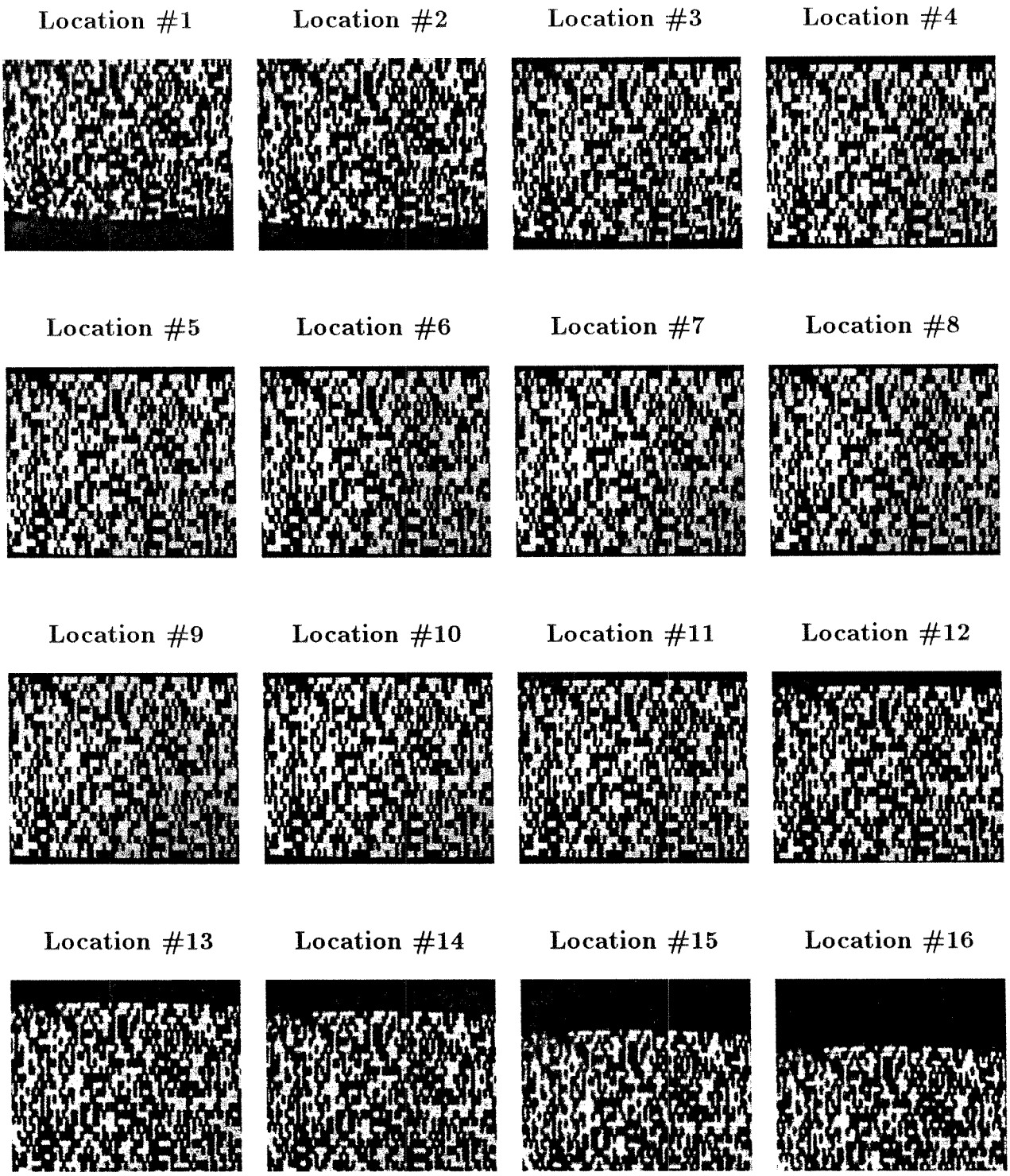


Figure 5.12: Storage of holograms in 16 locations: poor imaging system

Variable	Design #1	Design #2
H_{out}	2cm	2cm
# SLM pixels, vertical	440	440
height of SLM	2cm	2cm
f_{out}	12cm	7cm
A_{out}	10cm	7cm
vertical spacing between each location	.42cm	.24cm
total vertical extent of the 16 locations	6.7cm	3.9cm
SLM magnification	0.333	0.67
f_{object}	30cm	20cm
A_{object}	8.2cm	7cm
vertical extent of the object beam Fourier transform	0.97cm	0.32cm

Table 5.1: Design variables for 160,000 hologram system

We show the parameters of our first design in the center column of Table 5.1. The first few lines show variables which were constrained by other parts of the holographic system. These include our desire to use a 2cm wide crystal, the dimensions of the Epson 480 $\times$ 440 SLM, and the parameters of the prototype mirror array. Equation 5.1 and 5.3 give the required dimensions of the beam which is input to lens f_{in}, and Equation 5.2 is readily satisfied by reference beams which are larger than 0.5mm or so. Equation 5.3 gives the vertical extent of each storage location. This parameter is the crucial sticking point: will each location be too small to capture the spatial frequencies of the image, or will the stack of locations be too tall to permit undistorted imaging of the SLM onto the CCD for all locations?

The magnification of the SLM influences three things in the object beam. We want the make the SLM image appear small[10] so we can use many fractal rows, and also so we can fit the SLM image through the aperture. However, we want the SLM image to appear large so its Fourier transform is small. We usually proceed as follows: we decide to use every n mirror strips as fractal rows. We then calculate the minimum required magnification. For instance, we would need to magnify the SLM by a factor of 0.3 in order to use every mirror strip. If we use every other mirror strip, we could

[10] i.e., demagnify it more from the SLM to the front focal plane of lens f_{object}, or $m < 1$.

magnify by a factor of 0.6. In view of the problem discussed in the next paragraph, we settled on a magnification factor of 0.333 so that the SLM image would be smaller when it passed through aperture A_{object}. As you can see in Table 5.1, this choice made the Fourier transform quite a bit larger than the spot size, causing loss of some spatial frequencies.

This first design used $f_{object} = 300mm$, a computer–optimized achromat lens of 82mm clear aperture. A matched lens sat beyond the exit of the crystal to complete the 4–F system. These lenses had excellent performance over most of the clear aperture. However, the stack height of 6.7cm, consisting of >80% of the vertical clear aperture, proved to be too large for these lenses. This can be seen in Figure 5.12, where we show holographic reconstructions for storage in 16 locations. The center 10–12 locations or so show good imaging—however, the locations which use the extreme edges of the lens show unacceptable aberration and image loss.

At this point, we had three choices to correct this 16 location problem. We could have chosen to:

1. Obtain custom designed lenses of large aperture.

2. Shrink the size of the stack of 16 locations by reducing the focal length of the output lens f_{out}. We would also get the same effect through a redesign of the mirror array tilt angles, but we preferred to use our existing mirror array.

3. Use phase–conjugate readout to limit the clear aperture required of the signal beam lenses.

In our experiment, we chose to use Option 2. We describe the design here, and show the experimental results in Chapter 6.4.

We reduced the focal length of the reference beam lens L_1 from 120mm to 70mm, as shown in Table 5.1. With this change, we can use a pair of f=200mm, 70mm aperture achromats in the object arm. The advantage of moving to a shorter f_{object} is that the size of the Fourier transform spot decreases, which we really need because the vertical spacing between each location has also shrunk. Since the aperture A_{object}

was no longer a problem, we could use a magnification ratio that just allows us to use every other mirror strip (n=2). From Table 5.1, the Fourier transform of the object beam is still larger than the spacing between locations, but the discrepancy is much smaller.

Given these lenses and apertures, we can determine how many of the 256 mirror strips are accessible to us, and how much angle multiplexing bandwidth we have on each. We will assume that we'll be using a 2 inch cubic beamsplitter, a 2 inch circular quarter–wave plate with 1cm of clearance between the waveplate and the mirror array, and a Fresnel lens with 7cm focal length and a 7cm square aperture. The quarter–waveplate has a more involved effect than the others, since it alone has a circular aperture. We show the maximum number of angle–multiplexing nulls in Figure 5.13 if the waveplate were the only limiting factor. As expected, the angle multiplexing range is smaller for the mirror strips which lie at the edges of the mirror. The curves are not symmetric because of the tilt of the mirror facets—at one end of the mirror array, light is directed back towards the optical axis where the waveplate is wider. At the other extreme of the mirror array, light is directed away from the optical axis and fewer angle multiplexing nulls are available. We have assumed that the angular selectivity is

$$\frac{\lambda}{L} = 2.44 \times 10^{-5} = \frac{1.7\mu m}{70mm}.$$

As we discussed in Chapter 2.1, the theoretical angular selectivity increases as the reference beam moves away from normal incidence. In practice, however, the measured angular selectivity is usually broadened considerably by the non–plane wave nature of the reference beam. This is especially true when using a Fresnel lens, as the reference beam may be composed of multiple spherical waves. However, angular selectivity continues to operate, as long as the radius of curvature of the spherical wavefront is much larger than the interaction length of the hologram [147]. In terms of our analysis of the capabilities of our reference arm, we need to space the holograms further apart within the same horizontal deflection range. The non–plane wave reference beam reduces the number of angle–multiplexing nulls.

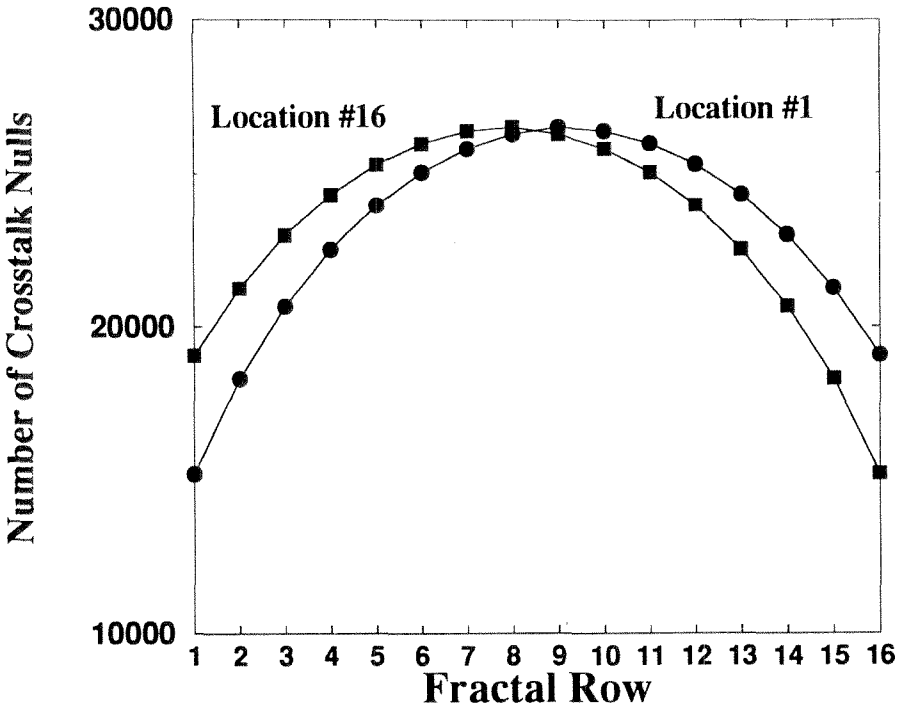


Figure 5.13: Maximum number of angle–multiplexing nulls allowed by the aperture of the quarter–waveplate

So far, we have only discussed the limitation on the reference beam imposed by the aperture of the quarter–waveplate. The beamsplitter and Fresnel lens impose an additional limitation on both vertical deflection and horizontal angle–multiplexing bandwidth. However, since these two apertures are square, the horizontal and vertical effects are not coupled. The 2 inch beamplitter limits the angle–multiplexing to 15,100 nulls, and the Fresnel lens to 17,650 nulls. In addition, beams from some of the mirror strips do not clear the aperture of the beamsplitter. Only the extreme locations: 1, 2, 15, and 16 are affected. The bottom 4 mirror strips which are pointed at location 1 are blocked, as are the top 4 mirror strips pointed at location 16. This makes sense, as the light leaving these mirror strips starts far from the optical axis and is deflected farther by the tilt of the mirror strip. No surprise that the beams never get out of the far end of the beamsplitter. For locations 2 and 15, the two extreme mirror strips cannot be used.

Even though the reference beams pass through all of the apertures, they still may

be unusable. If the vertical dimension of the beam is shrunk by off–axis aberrations, it may fail to overlap all of the information the signal beam. If the horizontal dimension shrinks, then we can still use the reference beam if we compensate for the reduced diffraction efficiency and angular selectivity.

Chapter 6 Holographic Random Access Memory: Experiments

Contents

In this chapter, we experimentally demonstrate the various elements which compose our holographic random access memory. We start by demonstrating spatially multiplexed storage in 8 locations using the mirror array. We then demonstrate storage of 1000 angle-multiplexed holograms using an AOD in the reference arm with an EOM in the signal arm. This corresponds to one fractal row of our 160,000 hologram system. Returning to mechanical scanners, we store 10,000 holograms in a single location ($\sim 1\text{cm}^3$) using the image plane and then the Fresnel plane geometries. We then demonstrate storage in 16 locations using the mirror array, and use the full

system to store 10,000 holograms at the top, center, and bottom locations. Finally, we conclude with some related experiments which we performed, including using the mirror array system for simultaneous memory readout and real–time face correlation, and thermal fixing of multiple holograms.

6.1 Storage using the mirror array

The 90° geometry—with reference and signal beams entering orthogonal crystal faces—was used to angularly multiplex up to 500 holograms at each of 8 spatially multiplexed locations in a $LiNbO_3$ crystal. The segmented mirror array and a 2–D mechanical scanner were used to perform both angular and spatial multiplexing.

The experimental setup is shown in Figure 6.1. The recording medium was a 0.01% Fe–doped $LiNbO_3$ bar of dimensions 8mm × 8mm × 50mm, cut for the 90° geometry. Holograms are recorded with the signal beam propagating down the long axis of the bar and the reference entering from the side. This configuration is convenient: since the signal beam is present at all spatial locations, there is no need to deflect the object beam. Input images were presented on a fixed input SLM and reconstructions observed on a fixed 2–D CCD array at the other end of the bar, no matter which location was being accessed. In this experiment, the SLM was a transparency mounted on a rotational stage.

The reference arm consisted of a mechanical scanner, a segmented mirror array, and several lenses. The scanner focused the input beam to a spot with a single lens and reflected the light off a mirror oriented at 45° with respect to the optical axis of the lens. The position of the focused spot could be scanned in both dimensions by moving the lens and mirror assembly with two computer–controlled actuators. Note that horizontal translation was achieved by moving the assembly in the axial direction, so there was no limit on movement in this dimension. A periscope arrangement can be used to achieve the same effect in both dimensions, but was not necessary in this experiment. A pair of lenses imaged the focused spot onto the mirror array. As described in Chapter 5.2.1, the 2–D movement of the focused spot on the mirror

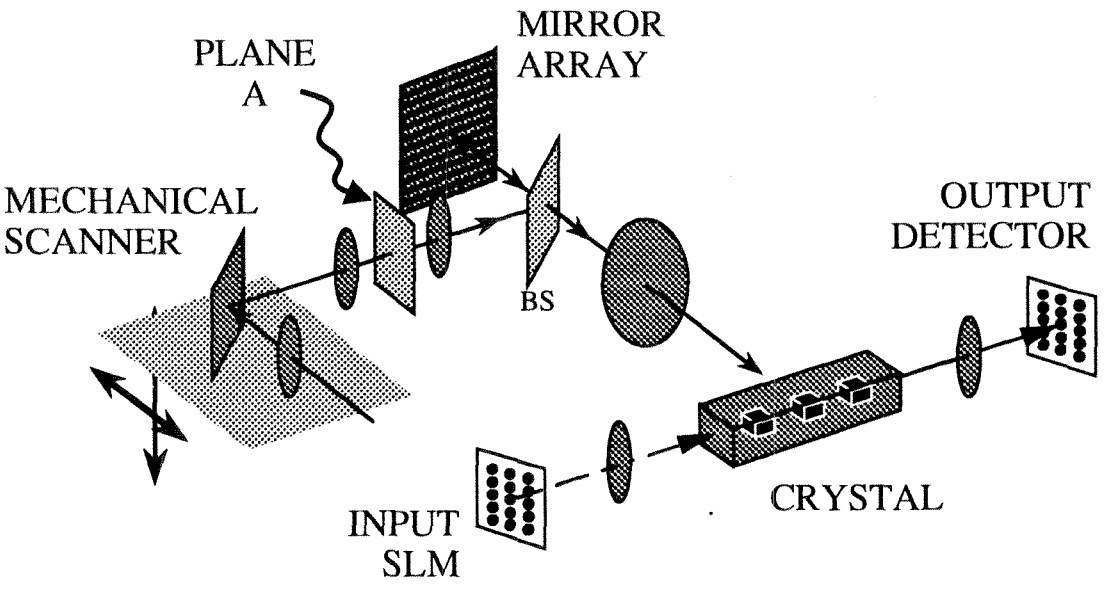


Figure 6.1: Spatially–multiplexed storage using the mirror array

array implements both spatial and angular multiplexing at the crystal.

First, 500 holograms were stored at a single location with a 6mm diameter reference beam. An exposure schedule [157] was used to equalize the diffraction efficiencies of the holograms. The longest exposure was 12.2 seconds and the shortest 1.65 seconds. The total recording intensity incident on the crystal was 130 mW/cm^2, the ratio of the reference to signal beam intensity incident on the crystal was 6.25, and the average diffraction efficiency 10^{-8}. Several reconstructions are shown in Figure 6.2. The original transparency is shown at upper left, and was rotated 1° between each exposure.

Next, holograms were recorded at each of 8 spatially multiplexed spots (each reference beam was 4mm in diameter) along the length of the crystal, using 8 of the 256 mirror strips on the mirror array. At locations near the center of the crystal, 500 holograms were stored at each location. Towards the ends of the bar, the limited aperture of the beamsplitter restricted us to 350 holograms per location. The average diffraction efficiency was again 10^{-8}. Several reconstructions are shown in Figure 6.3. Note that, in the reconstructions of the holograms stored towards the ends of the bar, the top and bottom edges are missing. This occurred because the reference beam

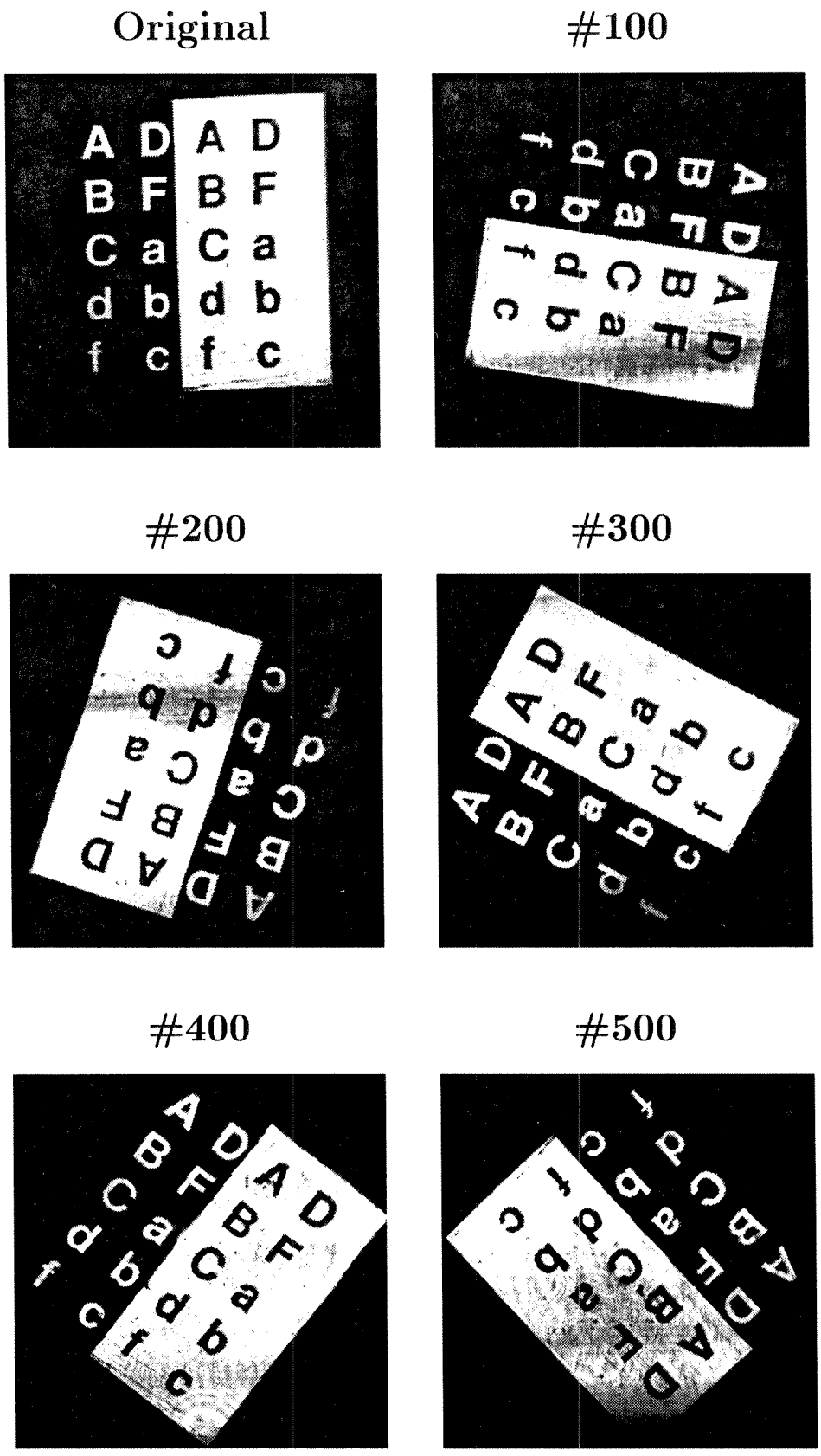


Figure 6.2: Example reconstructions: storage of 500 holograms at one spot.

becomes oval-shaped at these outermost storage locations, due to lens aberrations and the inclination of the mirror strip out of the focal plane as the focused spot moves off-axis.

In our experiment, the limit on the number of holograms per location was the number of distinct reference angles we could provide to each location, rather than crystal dynamic range. This limit was imposed by the beamsplitter aperture, not the mirror array. We only used 28mm (about 37%) of the horizontal extent of the mirror array. Therefore, an optical system with larger apertures can access 1000 holograms per spot, or more.[1] In this system, there were additional losses not normally encountered in a standard angle multiplexing holographic setup, bringing signal levels closer to the fixed detector noise floor and reducing dynamic range. These losses include the optical loss at the nonpolarizing beamsplitter, the low reflectivity (40%) of the gold surface of the mirror array, and absorption in the long crystal. Absorption allows approximately 5% of the light illuminating one of the end faces to be transmitted out the other end.

One final aspect to be considered is the presence of the object beam at all locations, including those where holograms have already been written. Holograms at the location which is first used for storage will be erased by the object beam during exposure of all subsequent holograms. This does not cause much difficulty in the system described above for two reasons. One reason is that the external ratio of reference to signal intensity was large (approximately 6.25). In addition, the storage locations can be filled starting from the end farthest from the entry of the object beam. In this way, the holograms exposed to the object beam for the longest time are buffered from its effects by the absorption of the long crystal. For these two reasons, very little erasure of the holograms occurs once the reference beam moves to the next storage location.

In the experiment described above, we used a mechanical scanner to move a focused spot and a pair of lenses to image this moving spot onto the surface of the mirror array. In the center of this pair of lenses (plane **A** in Figure 6.1), there exists

[1] In Section 6.4.2, we store 2500 holograms with one mirror strip.

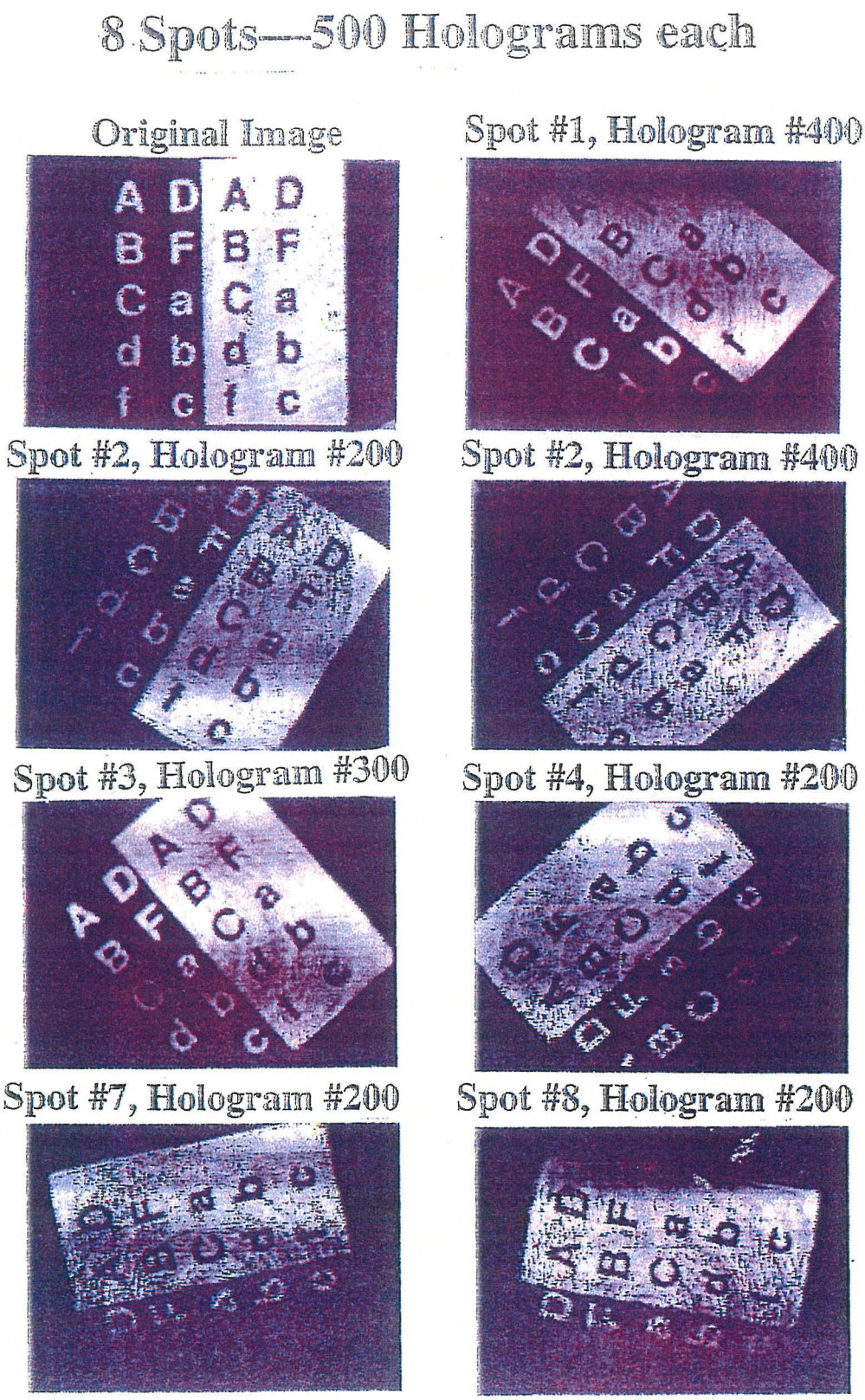


Figure 6.3: Example reconstructions: storage at 8 locations.

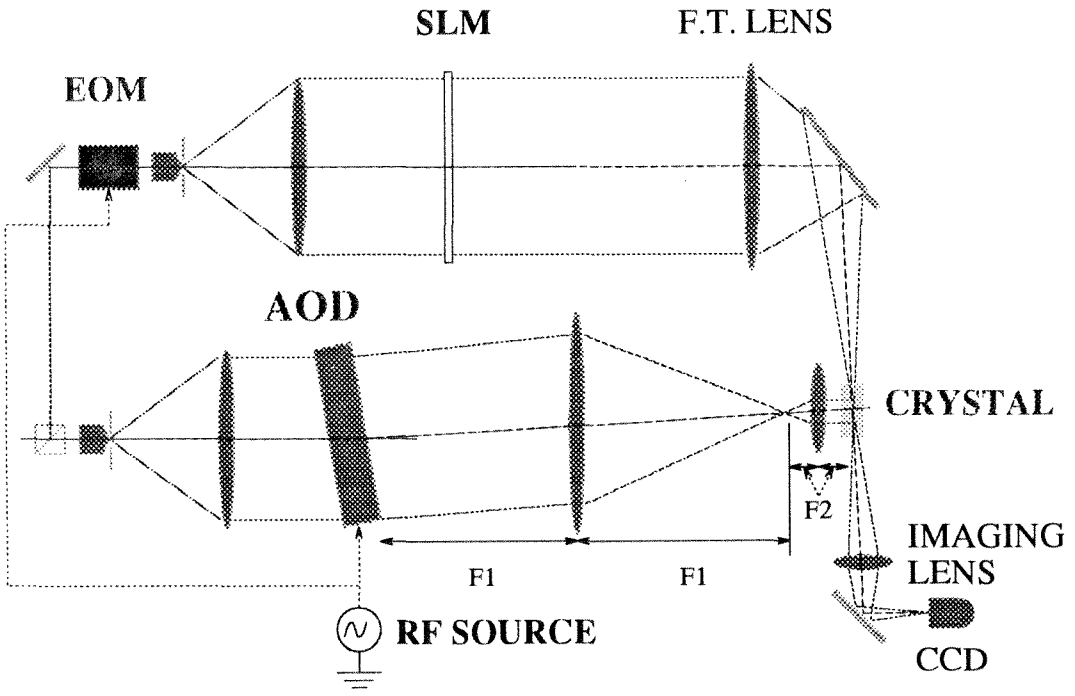


Figure 6.4: Storage of 1000 holograms using an AOD: experimental setup.

a plane wave centered on the optical axis with a vertical and horizontal incidence angle. We can create an identical plane wave by using a 2–D non–mechanical angle scanner in this plane **A**. One such example is a pair of crossed AOD cells with cylindrical accessing lenses [129, 354]. The Doppler shift added by these cells can be removed by an electro–optic modulator (EOM). In this way, the mirror array can provide spatially multiplexed holographic storage without mechanical movement. In the next section, we demonstrate holographic storage using one AOD and an EOM for frequency compensation.

6.2 Storage using an AOD

We used the 90° geometry to demonstrate storage and readout of holograms using an AOD and a compensating EOM [355]. The experimental setup is shown in Figure 6.4. The reference arm is a 4–F system which images the angle change from the AOD onto the crystal. A block in the center Fourier transform plane spatially filters the DC or undeflected light. We used a Crystal Tech: AOD4050–2 TeO_2 slow shear

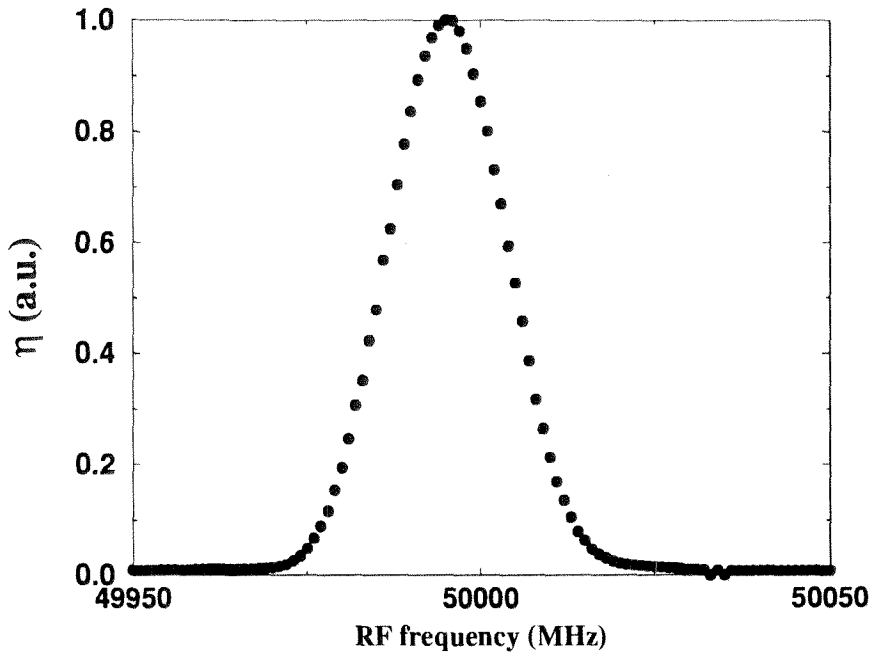


Figure 6.5: Measured and expected angle selectivity, in terms of RF frequency.

AOD with a center frequency of 50MHz and an acoustic velocity of 617 m/sec. The angle deflection is proportional to the driving RF frequency. For the full frequency range from 33Mhz to 66Mhz, we obtained a deflection of 1.5°. The 5:1 4–F system demagnified the 4cm illuminated aperture of the AOD to a 0.8cm wide beam at the crystal, and expanded our angle deflection range to 7.5°. The measured and expected angle selectivity, represented in terms of RF frequency, are shown in Figure 6.5. Since the reference beam was apodized by the AOD aperture, the selectivity function no longer has visible nulls. We stored 1000 holograms by using a inter–hologram spacing of 30kHz. By using a quarter–wave plate both before and after the AOD, we were able to achieve an peak efficiency of 86% at the center frequency. However, in order to have uniform efficiency across the tuning range, we detuned these waveplates to obtain a uniform response of 22%. The main purpose of doing this was to simplify the exposure schedule as much as possible.

The signal beam contained a New Focus Model 4002 broadband electro–optic modulator using Mg–doped $LiNbO_3$. This device operates on the "raw" laser beam and outputs an undeflected phase–modulated beam containing many spectral orders.

The signal beam was then expanded, illuminated the SLM, and directed to the crystal. The crystal was placed just in front of a Fourier transform plane of the SLM image (Fresnel plane storage). In this experiment, the SLM consisted of a slide which was rotated several degrees between each exposure.

The phase–modulated beam can be described as [349]:

$$e^{j\omega t}e^{j\delta\cos\Omega t} = e^{j\omega t}\left(J_0(\delta) + J_1(\delta)e^{\pm j\Omega t} + J_2(\delta)e^{\pm 2j\Omega t} + \cdots\right). \tag{6.1}$$

Only one of these terms ($e^{j(\omega+\Omega)t}$) corresponds to the Doppler–shifted reference beam, so efficiency and modulation depth are serious concerns in this system. The parameters that we can control are δ and the input polarization. We are also free to use any output polarization, since we can use a half–wave plate to return the signal beam to the vertical polarization we need for the 90° geometry. Our optimization goal is not obvious, as we need to increase the amount of light modulated by the correct order, as well as minimize the amount of light remaining in the unmodulated zero order. The exact determination of the configuration depends on the relative importance of these two parameters. In our experiment, we achieved 45% modulation into the first–order with 10% of the power remaining in the zero order. Note that there exists, corresponding to the useful plus one order, an equally efficient yet useless minus one order. The waste light reduces modulation depth and $M/\#$. In the AOD system, we obtained an $M/\#$ of 0.28; in contrast, the same crystal[2] achieved an $M/\#$ of 0.8 for holograms stored without an AOD. As a result, the diffraction efficiency of the AOD–stored 1000 holograms was $\sim 8 \times 10^{-8}$.

The reconstructed holograms were measured with a Photometrics Imagepoint CCD array containing an 8–bit digitizer—some reconstructions are shown in Figure 6.6. The detector pixel map covering the image (310,000 pixels in all) was divided into regions which were expected to be bright (or dark). In this process, "edge" pixels near a dark/bright transition were discarded. The two resulting histograms are representations of the probability density functions (PDF) for storage of binary data

[2] A 1 cm × 1 cm × 2 cm $LiNbO_3$ crystal, with 0.01% Fe doping.

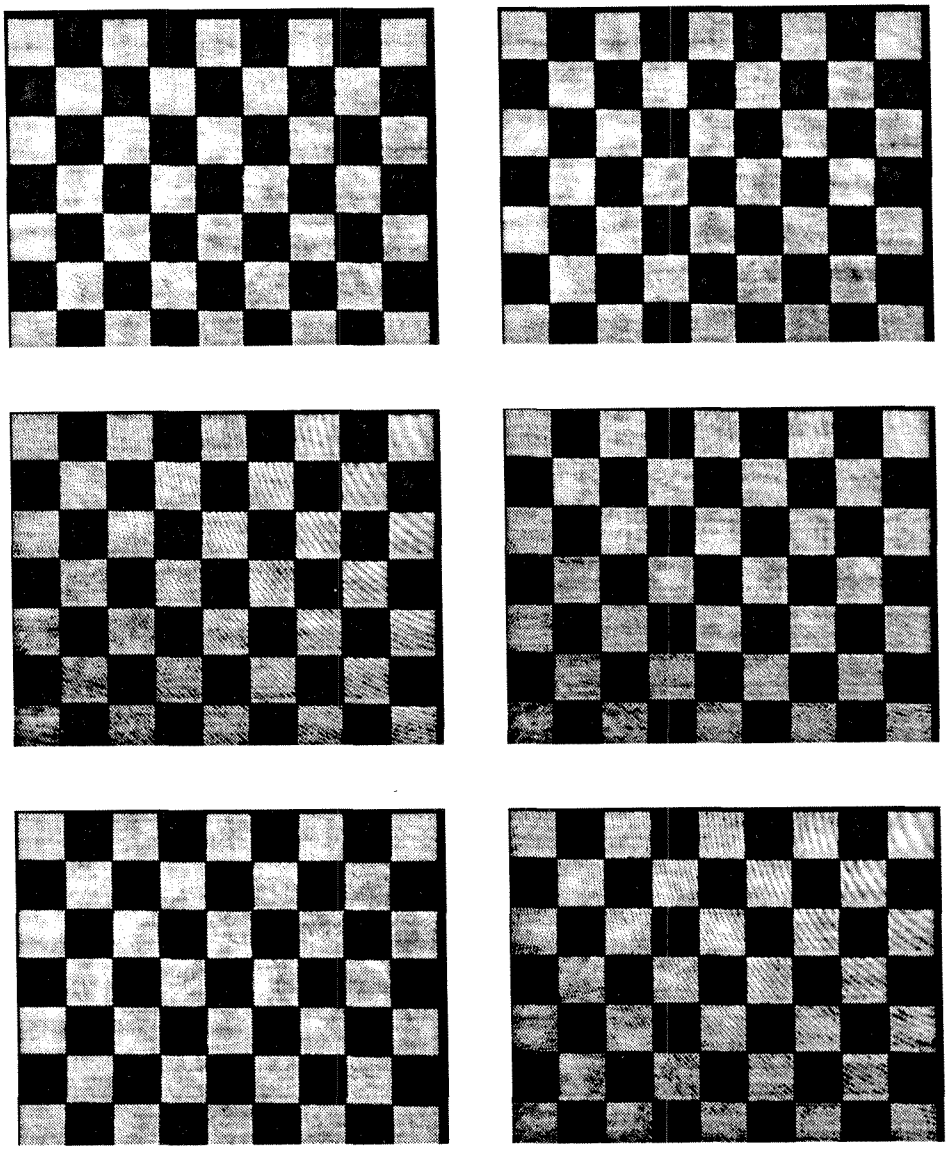


Figure 6.6: Storage of 1,000 holograms using an AOD: reconstructions.

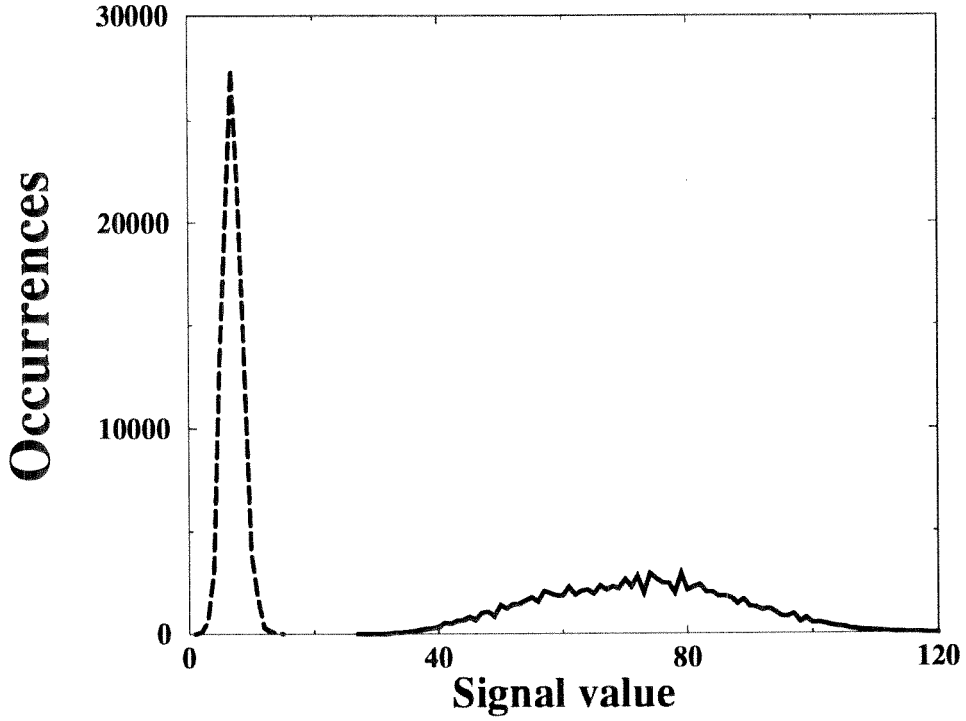


Figure 6.7: Probability density functions for ON and OFF pixels.

in this system. The two PDFs are shown in Figure 6.7 for a sample reconstruction—approximately 200,000 pixels from the image are represented in all. An optimal threshold was empirically determined, giving a measured raw probability of error of 10^{-5}. We could reduce the measured error to zero by using one threshold for the center portion of the image, and a different one for the edges. In Section 6.3.2, we will use a position–dependent threshold which we obtain from a second, all–pixels–ON hologram. We will refer to this process as compensation of the illumination profile.

6.3 Storage of 10,000 holograms

6.3.1 10,000 image plane holograms

Using the system shown in Figure 6.8, we stored 10,000 image plane holograms in an 8×8×50 mm bar of 0.01% Fe–doped $LiNbO_3$ [329]. The **c**–axis was again cut for the 90° geometry, at 45% to the vertical faces. The reference beam was 45mm × 5mm in area, and the object beam 4mm × 5.4mm for a total interaction volume of 0.972 cm^3.

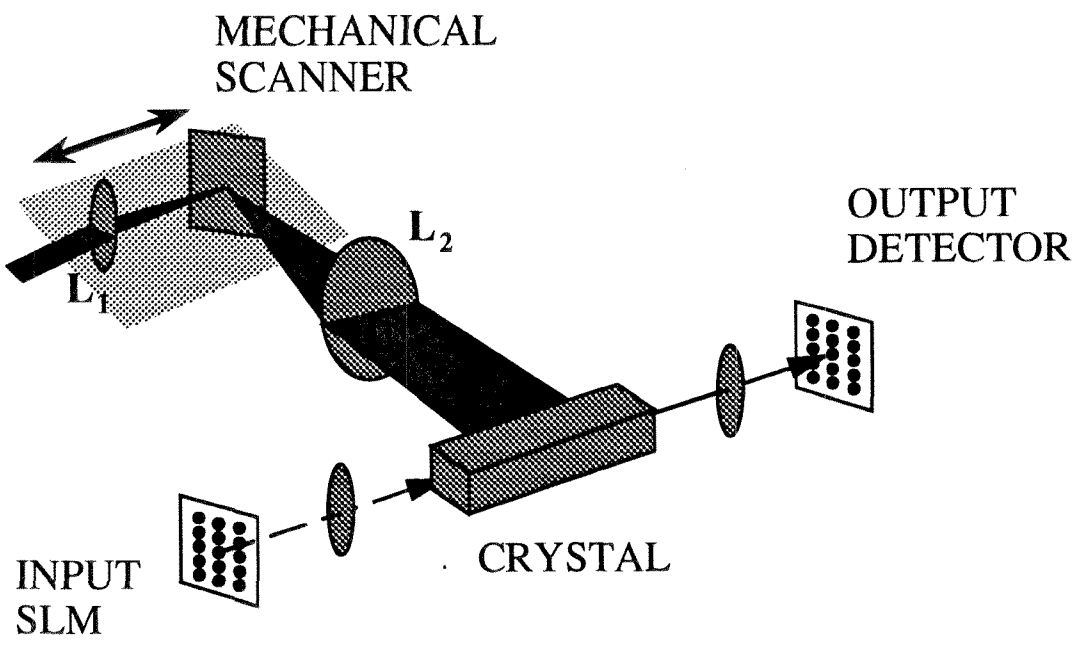


Figure 6.8: Storage of 10,000 image plane holograms: experimental setup

Images were presented on a liquid–crystal SLM from a projection television. The SLM was demagnified by a factor of 5 and imaged to a plane located approximately halfway along the long dimension of the crystal. Each hologram contained 480 × 440 pixels, so that assuming one bit per pixel, 2.11 GBits were stored. Using a final exposure of 0.21 seconds and an erasure time constant of 2250 seconds, a standard recording schedule was computed. This led to a initial exposure of 3.1 seconds and a total exposure time of 102 minutes. The average exposure time was 0.61 seconds. The actual time required to complete recording was significantly longer since we included extra delays between each exposure for the mechanical scanner and video input to stabilize.

A 4–F system was used to magnify and image the reference beams onto the long face of the crystal, with a mirror at 45° near the center focal plane to fold the optical axis by 90° (See Figure 6.8). Lens L_1 and the mirror were fixed relative to each other and mounted on a linear actuator capable of 4 inches of travel parallel to the incoming reference beam, while lens L_2 and the crystal were mounted on the optical table. By moving the linear actuator back and forth, the focused spot at the folding mirror was

translated horizontally in the back focal plane of lens L_2. This translation appeared at the crystal as a change in the horizontal incidence angle of the reference beam, enabling angular multiplexing.

The experimentally measured selectivity of 1.1×10^{-5} radians (all angles external to the crystal) agreed with theoretical expectations for the 90° geometry and a 45mm interaction length. Although the presence of images broadened the effective selectivity by a factor of 2 or so, a spacing of 4×10^{-5} radians was sufficient to suppress holographic readout by about 30dB. To implement this spacing with a lens L_2 of focal length 150mm, the linear actuator was moved 6μm between holograms for a total travel of 60mm.

The input images during the experiment included gray–scale cartoon images as well as a benchmark chessboard pattern. Several reconstructions of this chessboard pattern are shown in Figure 6.9. You can see the effects of non–uniformity of the signal and reference beams, in the loss of fidelity at the corners of the output images. The hologram number appears in the upper left corner of each hologram. As can be seen, the time constant used was slightly too high, and the early holograms decayed more than expected. The average diffraction efficiency of the chessboard images was measured to be about 5×10^{-9}, so the power in the reconstructed holograms was within the same order of magnitude as the background light scattered by the crystal. Fanning of the reference beam did not appear to affect storage.

At this point, we can store 10,000 holograms at each location. Incorporating the SLM limits mentioned in Chapter 1.3.1, the total achievable storage capacity at one location is between 1 and 10 gigabits. In order to increase the total capacity beyond this limit, we can use spatial multiplexing. If we continue with image plane storage however, we will need to move the array detector to access holograms from different locations.[3] However, by storing Fourier plane holograms, we can detect the reconstructed holograms from multiple locations without moving the detector array. We can obtain the same effect by storing in a plane slightly away from the Fourier plane, which we usually refer to as Fresnel plane storage.

[3]Or somehow steer all the reconstructed signal waves into the same path.

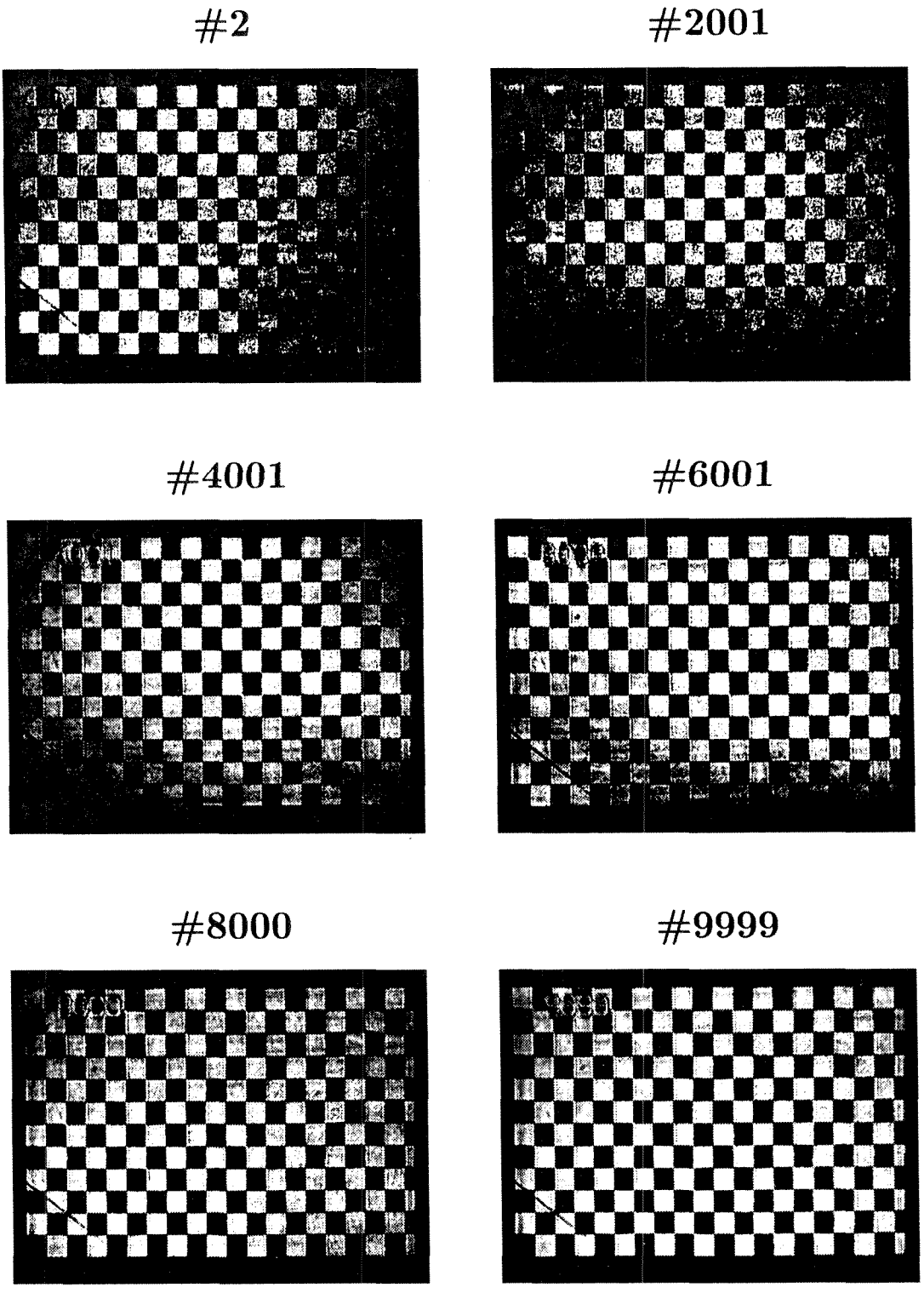


Figure 6.9: Storage of 10,000 image plane holograms: reconstructions.

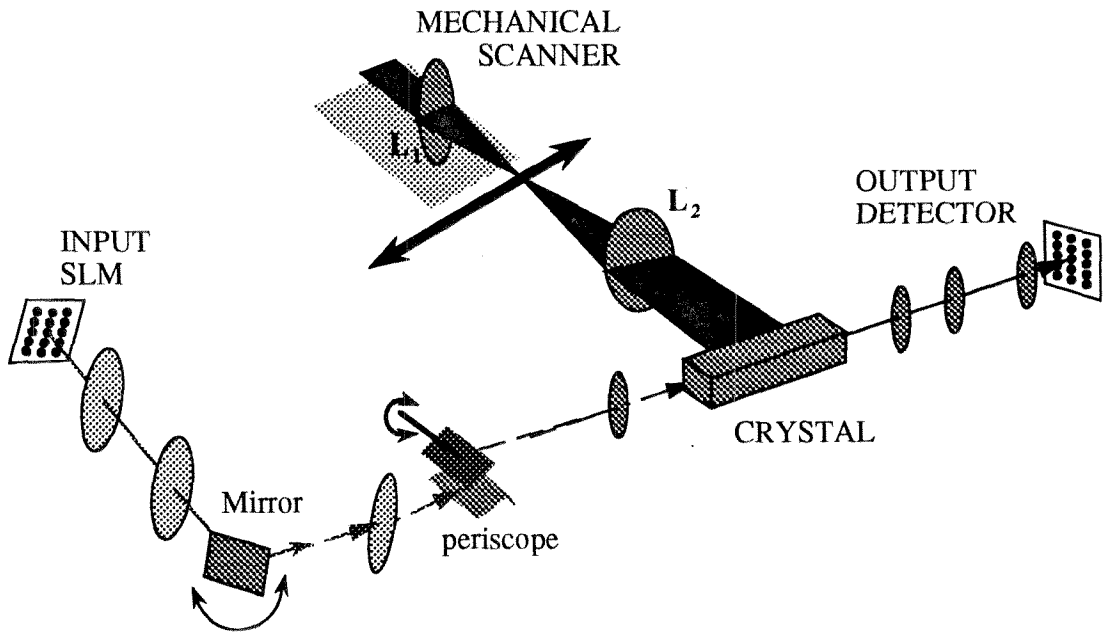


Figure 6.10: Storage of 10,000 Fresnel plane holograms: experimental setup

6.3.2 10,000 Fresnel plane holograms

In this section, we describe the storage of 10,000 holograms in one storage location of a $LiNbO_3$:Fe crystal using near–Fourier plane storage. Figure 6.10 shows the experimental setup. The reference arm contains an XY mechanical scanner which moves the focused reference beam horizontally for angle multiplexing, and vertically for fractal multiplexing. You can think of this setup as corresponding to one location of the 160,000 hologram system, where multiple identically–tilted mirror strips access the same location. Not shown in the drawing are two cylindrical lenses which magnify the horizontal dimension of the reference beam, and two mirrors in a periscope arrangement to convey this beam onto the mechanical scanner. Lenses L_1 and L_2 had focal lengths of 80mm and 120mm, respectively. The reference beam spot size was elliptical, about 20mm wide and 6mm high, with an area of 0.95 cm^2.

The object beam is directed to the proper location after the image information has already been imposed on the beam. The image presented on the liquid–crystal SLM is demagnified by 3× and imaged to a mirror mounted on a rotation stage. This horizontal deviation is not required in the theoretical design of our large–scale system, but it has an important role in its practical realization. One of the difficulties of

storage in the LILA geometry is that as the amount of fixed pattern energy[4] increases, images become increasingly distorted. If we deflect the signal beam horizontally, these noise gratings do not build up as fast. In addition, although these holograms are all stored at the same vertical location, we have used additional crystal volume for storage, which will us increase the diffraction efficiency of the holograms. We discussed this issue and how it affects the recording schedule and the dynamic range in Chapter 4.2.4.

The image plane in the center of this horizontally rotating stage is imaged via a 4F system to a pair of mirrors in a periscope arrangement. These two mirrors are used in the full system to deflect the signal beam vertically— in this experiment, they remained stationary. On the far side of lens L_3 (focal length 300mm), the horizontal angle determines where the Fourier transform of the displayed information arrives. Since a random phase plate was not used, the crystal was not placed in the Fourier transform plane, but was displaced beyond it by 80mm. At this point, the DC portion of the expanding image was approximately 1.77mm high $\times$ 2.4mm wide. Three lenses after the crystal filter out scattered light and magnify the reconstructed images onto a Photometrics Imagepoint cooled scientific CCD. The advantage of Fourier transform storage becomes apparent at this point, since a reconstruction from any spot in the crystal can be imaged onto the single detector array.

10,000 holograms were stored in a 0.01% Fe–doped crystal bar of dimensions 10×10×20 mm. The c–axis was at 45° to the vertical faces. The images were displayed on a 640×480 pixel VGA monitor, and sampled for the 480×440 pixel SLM.[5] Both random bit patterns and a standard chessboard pattern were stored. The last exposure time was 0.26 second and the erasure time constant was 3800 seconds. The initial exposure lasted 7 seconds, the total exposure time was 134.6 minutes, and the average exposure time 0.81 seconds. Four fractal rows were used for storage with 2500 holograms stored on each. The vertical spacing between fractal rows was 5mm (1.4°); the horizontal spacing between holograms was 15μm (.004°). The angles listed are

[4]from dust, fixed patterns on the SLM display, or common features among the presented images.

[5]This sampling process causes aliasing at the SLM pixels. As a result, we do not have pixel–based control of the SLM.

the external angles at the crystal face. The vertical angular bandwidth of the images was 1.15°, so the fractal spacing completely displaced unwanted reconstructions off the detector array. Several reconstructions are shown in Figure 6.11. These images were taken with exposure times between 0.5 and 1 second.[6] The average diffraction efficiency was approximately 5×10^{-9}. The average power in the reconstructions (which were already half dark) was 2.5 times the background scatter (measured before storage).

6.3.3 Compensation of the background illumination profile

To characterize the noise performance in this angle multiplexed memory, we normalized the reconstruction by an overall illumination profile [252]. This allows us to significantly suppress deterministic sources of errors, such as beam nonuniformity and dust particles on the optical components. Such deterministic error sources, common in a research–grade system, can in principle be eliminated by careful engineering. The performance we obtain after normalization provides an estimate for the performance that is expected from a prototype system.

As in the noise analysis described in Section 6.2, several reconstructions of the chessboard images were captured with the detector array and digitized. The edge pixels were discarded, leaving two sets of detector pixels: ON pixels and OFF pixels. Each detector pixel is treated as a separate sample—there is no spatial averaging. The histogram of each set of detector pixels is a PDF. We can obtain several parameters from these two PDFs to describe the error performance of the holographic memory:

- the measured bit error rate or P_e. We determine the best measured P_e by empirical selection of the threshold between ON and OFF.

- the signal–to–noise ratio. We define SNR as

$$SNR = \frac{\mu_1 - \mu_0}{\sqrt{\sigma_1^2 + \sigma_0^2}}. \tag{6.2}$$

[6] We could distinguish ON and OFF pixels with shorter exposures, but to have accuracy in measuring bit–error rate, we always tried to use the full dynamic range (8 bits) of the camera.

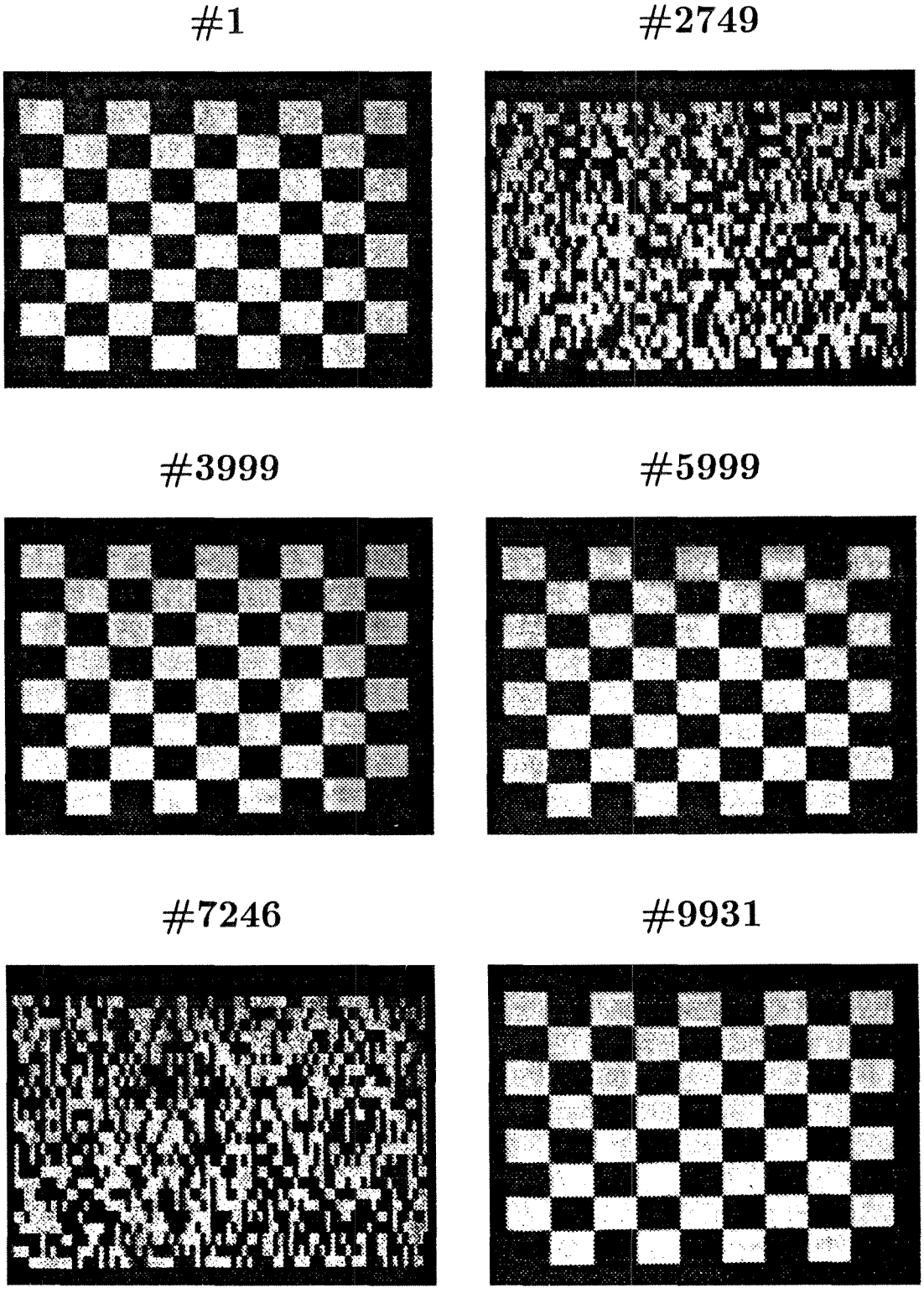


Figure 6.11: Storage of 10,000 Fresnel plane holograms: reconstructions.

The SNR is a useful indicator of the error performance: a large SNR indicates a low bit–error rate. We discuss the relationship between SNR and estimated bit–error rate in Chapter 7.2.1.

- the estimated bit–error rate. We estimate the shape of the PDFs by assuming that they follow a particular distribution, such as the Gaussian distribution. This is often necessary if the error rate is smaller than the inverse of the number of samples.

The normalization procedure is pictured in Figure 6.12. On the top of the figure, we show the uncompensated reconstruction and a cross–section. In the center is the reconstruction of a hologram stored with all pixels ON. This blank page hologram is stored near the data hologram so that it has the same spatial profile. We then divide the data hologram by the blank page hologram. The profile is spatially smoothed and normalized so that the compensated data pattern has amplitude values in the same range as the original data. We show the compensated reconstruction and a cross–section at the bottom of Figure 6.12.

In order to see what noise was added by the holographic storage process, we analyzed the SNR of the images for several additional conditions. We captured images without the crystal present, for transmission of the signal beam through the crystal (no hologram), for storage of one hologram, and storage of 10,000 holograms. Each chessboard image was also compensated by a blank page image as described above. The SNR for these various conditions is shown in Figure 6.13. The SNR after compensation is the upper curve; before compensation, the lower curve. Note that there is a degradation of SNR as the crystal is introduced and more holograms are stored, both with and without normalization. However, for the normalized images, there is a more pronounced difference between the SNR of the various test images and the SNR obtained when the 10,000 holograms are stored. This implies that the normalization procedure is more effective for imaging than for holographic reconstructions.

The SNR obtained in the absence of the crystal is principally limited by residual nonuniformity in the illumination and the SLM. The introduction of the crystal

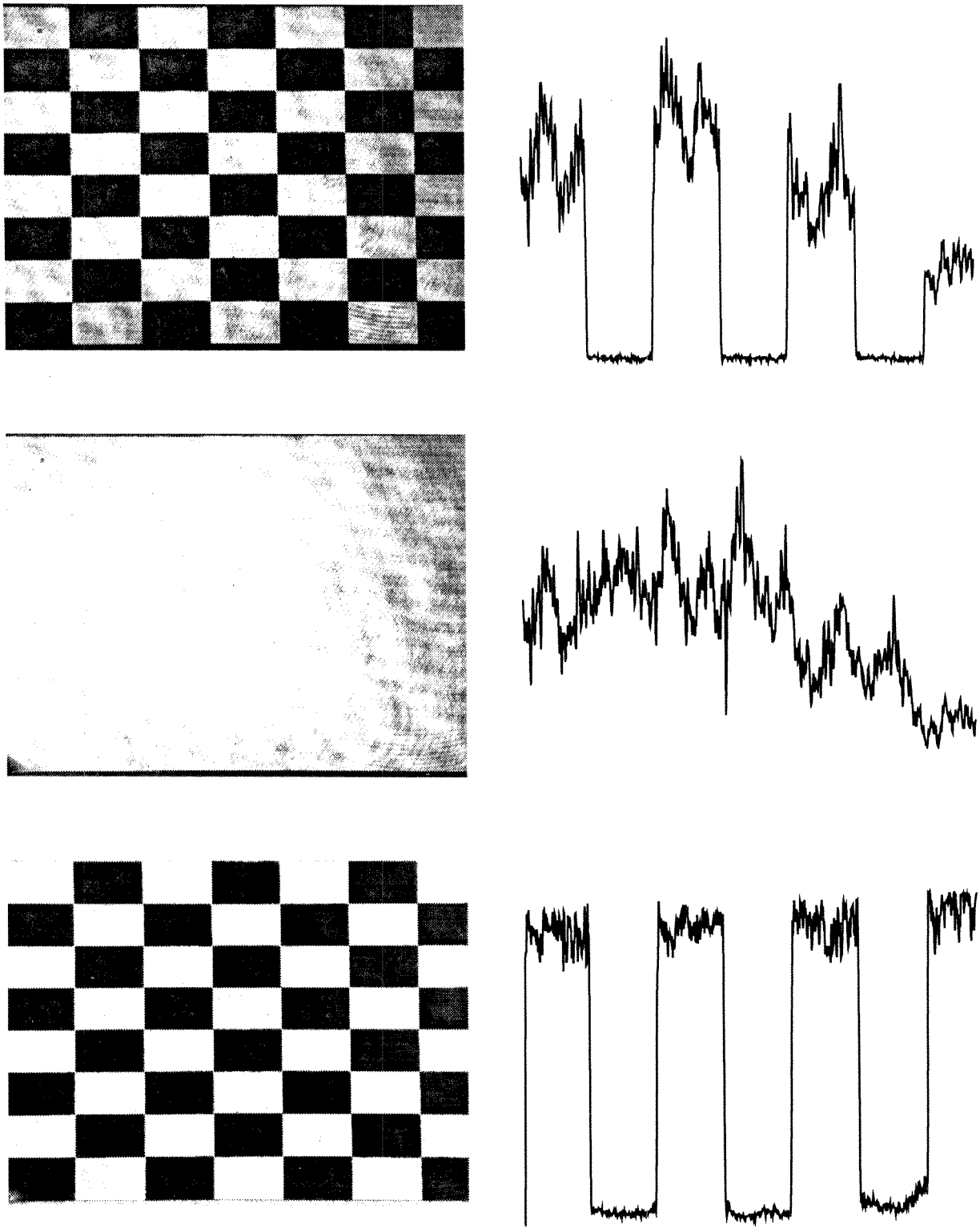


Figure 6.12: Normalization by the background illumination profile.

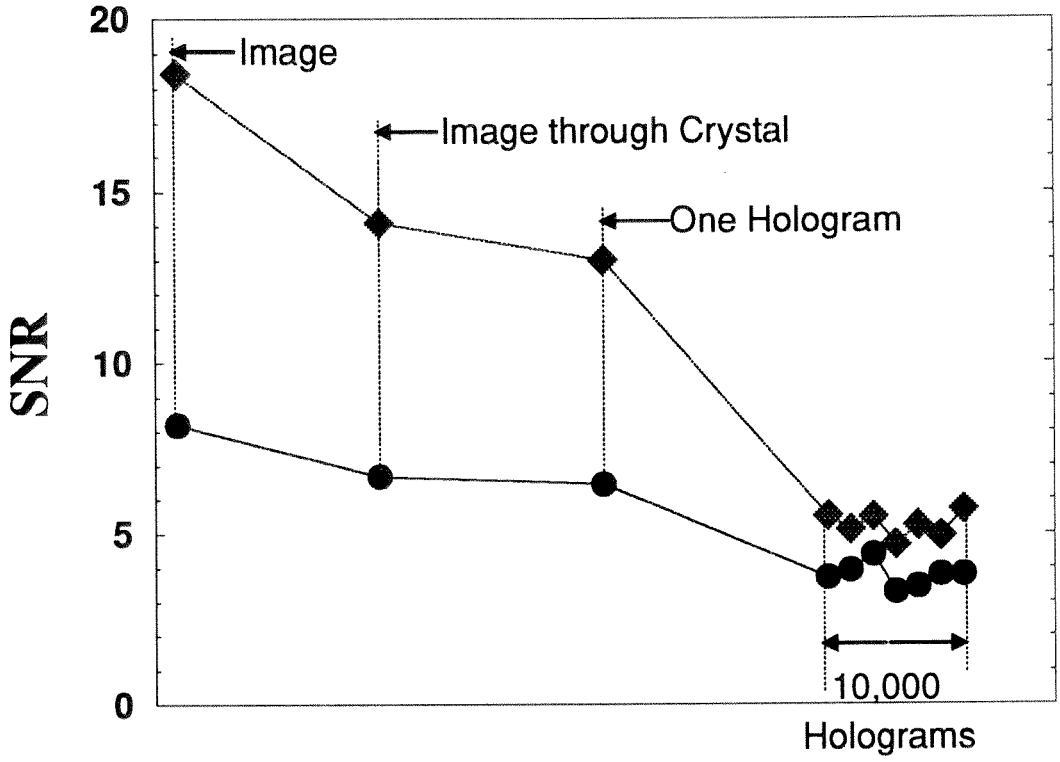


Figure 6.13: Evolution of the system SNR

reduces the SNR because of surface defects and scatter noise in the uncoated crystal. Note that introduction of the crystal is the largest source of SNR loss in the system. The small reduction in SNR when a single hologram is stored is attributed to the nonuniformity of the reference beam and the spatially varying modulation depth. We have significantly reduced this variation in modulation depth by recording the holograms away from the Fourier plane. The final SNR, after storage of 10,000 holograms, is lower than the SNR from a single hologram. The reason is not loss of signal strength, because we made the single hologram measurement with the same diffraction efficiency as each of the 10,000 holograms. Instead we attribute the lower SNR to three factors: Crosstalk (adjacent holograms were recorded at the 3^{rd} null of the angular selectivity curve), development of interpixel and other noise gratings over the long exposure sequence, and nonuniform erasure of the recorded holograms due to absorption in the crystal and the movement of the signal beam.

6.4 Demonstration of 160,000 hologram system

6.4.1 Demonstration of storage in 16 locations

In this section, we demonstrate the full 160,000 hologram system. First we demonstrate storage in each of the 16 spatially multiplexed locations. The experimental setup is shown in Figure 6.14. It is similar to the setup for the storage of 10,000 holograms, with the additional incorporation of the mirror array. The same 2–D mechanical scanner with periscope is used in the reference arm. The focused spot is directed onto the surface of the mirror array with a large polarizing beamsplitter cube. A quarter–wave plate in front of the mirror array makes both incoming and outgoing polarizations linear and orthogonal to each other. A half–wave plate in front of the crystal rotates the polarization to vertical, which we need for the 90° geometry. The overall efficiency of the beamsplitter/waveplate/mirror array combination is less than 25%.

We use the periscope in the image arm (described in the previous section) to perform vertical deflection of the signal for storage in the spatially multiplexed locations. The image plane in the center of the horizontally rotating stage is imaged to this pair of mirrors via a 4F system. The lower mirror is fixed and deflects the image by 90° to the upper mirror. The upper mirror is on a rotating stage,[7] and returns the object beam to a near–horizontal path. The center of rotation of the upper mirror is in the second image plane of the object arm, and deflection originates from the optical axis of the lens f_{object}. At this point, the object beam has deflected by an arbitrary 2–D angle. On the far side of lens f_{object}, this angle determines where the Fourier transform of the displayed information arrives.

We used an f_{object}=200mm achromat lens of 70mm aperture, and an identical lens after the crystal to complete the 4–F system. Using a .015% Fe–doped crystal of dimensions 2cm×1.5cm×4cm, we were able to store holograms in 16 locations without loss of information, as we show in Figure 6.15. In Figure 6.16, we show SNR as a function of location for a single hologram in each location, and 1000 holograms in

[7] the rotation axis is horizontal.

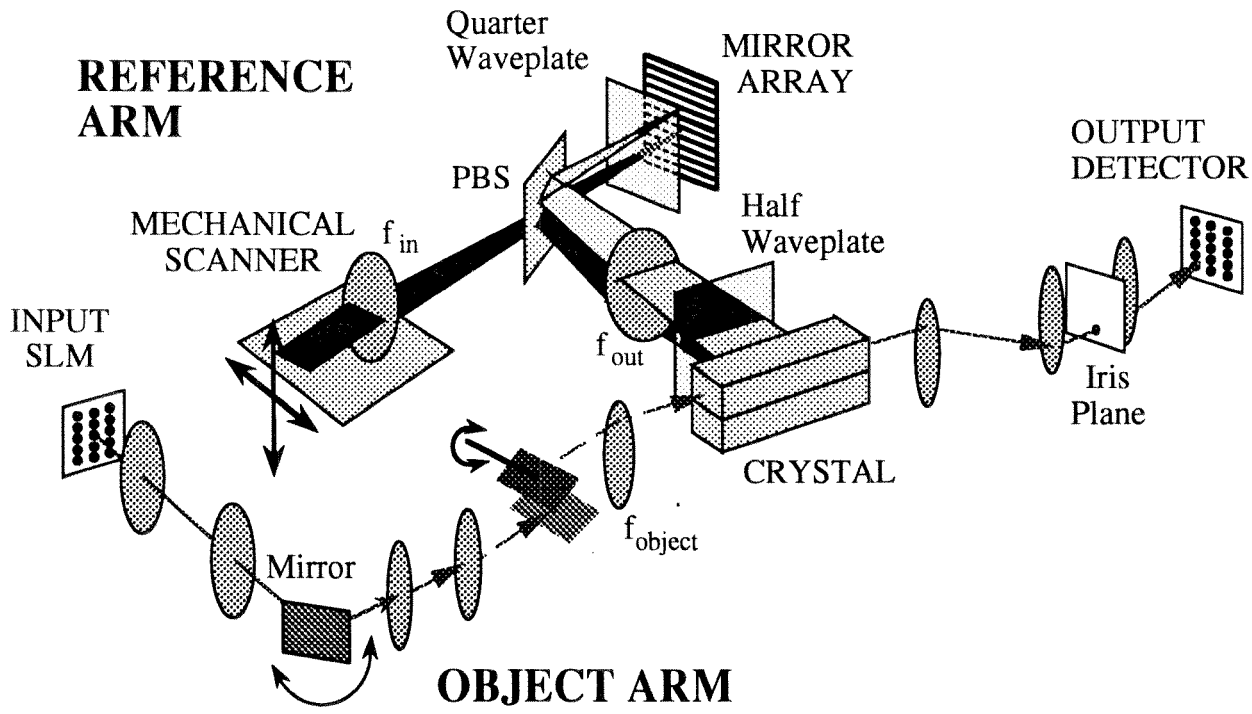


Figure 6.14: Experimental 160,000 hologram system.

each location. The method for SNR analysis is described with the 30,000 hologram experiment below.

The top and bottom locations required a little refocusing and redirecting of the object beam for best focus and to maintain the reconstructions in the center of the camera.[8] The need for this sort of adjustment implies that the distance between the lenses in the 4–F systems were not exactly correct, and that the rotation axes of the two deflecting mirrors were not positioned exactly in image planes of the SLM. We did not implement any refocusing with the array detector, but instead by translating lens f_{object} before storage. The readout optics then remained fixed and there was no effect on the readout speed. The lens f_{object} was moved within a range of 5mm for focusing purposes, and the vertical deflection mirror in our object beam periscope moved 1.6mm to center the images. Note that a slight rotation is seen of the vertically deflected images. This implies that the vertical deflection mirror also deflected the image horizontally a little. Another way of saying this is that the beams traveling in the vertical periscope did not lie in a single vertical plane.

6.4.2 Demonstration of storage of 30,000 holograms

In order to demonstrate the full system, we used the same setup from Figure 6.14. In the previous section, we demonstrated the ability to deflect the signal beam to each of the storage locations and to get the reconstructions back to the array detector. Here we demonstrate storage of 10,000 holograms in the top and bottom locations, plus the center location for good measure.

Preparation of the reference arm —For this experiment, we needed to be able to use several mirror strips to implement the multiple fractal rows. However, some of the mirror strips that we would have liked to use were not available. The reason is that the mirror array was constructed as a stack of tiles which were then bolted together. In order to keep the total height of the mirror array from becoming too

[8] See below for more description of the camera, and the size of the images relative to the active area of the camera.

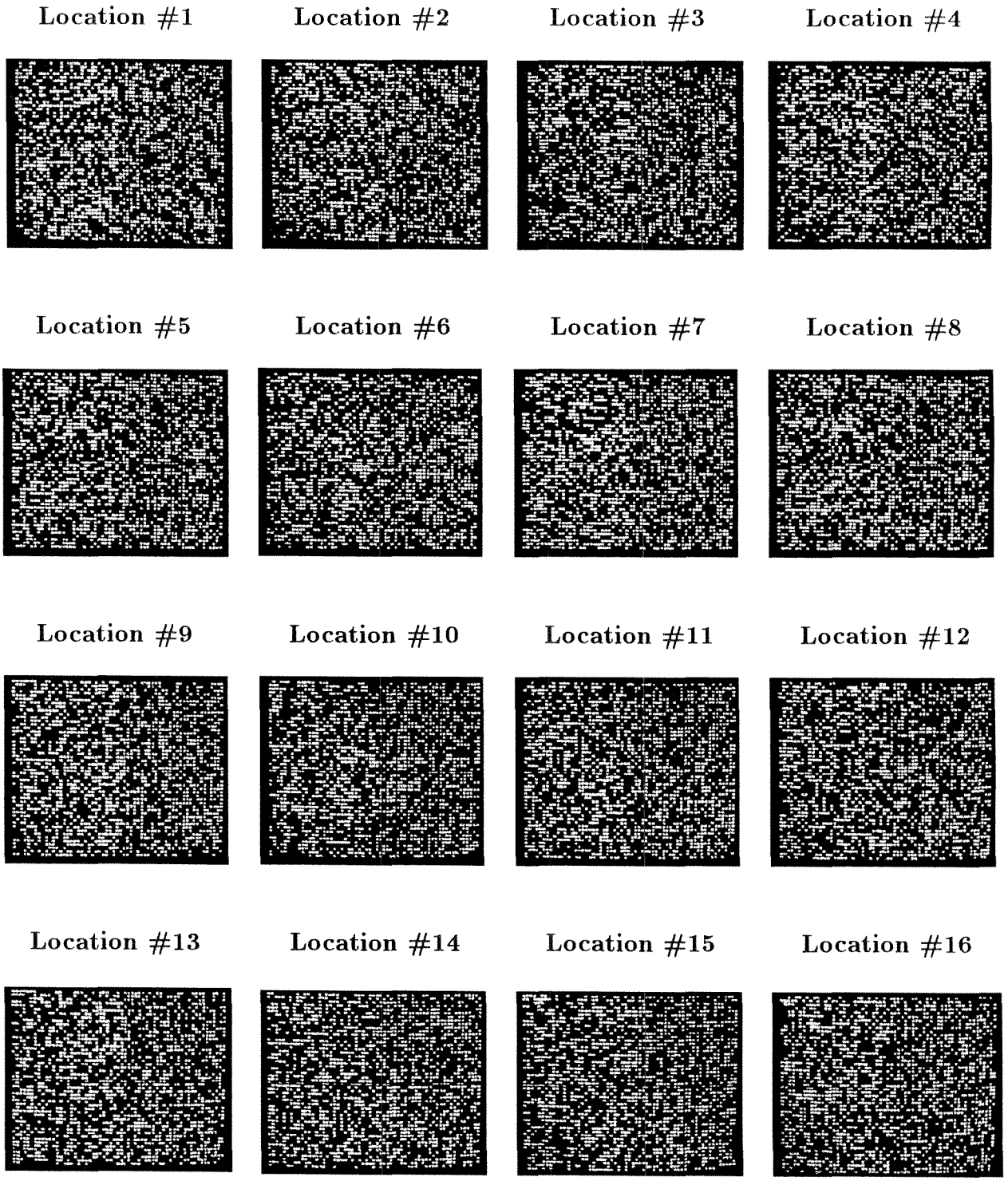


Figure 6.15: Reconstructions from 1,000 holograms at each of 16 locations

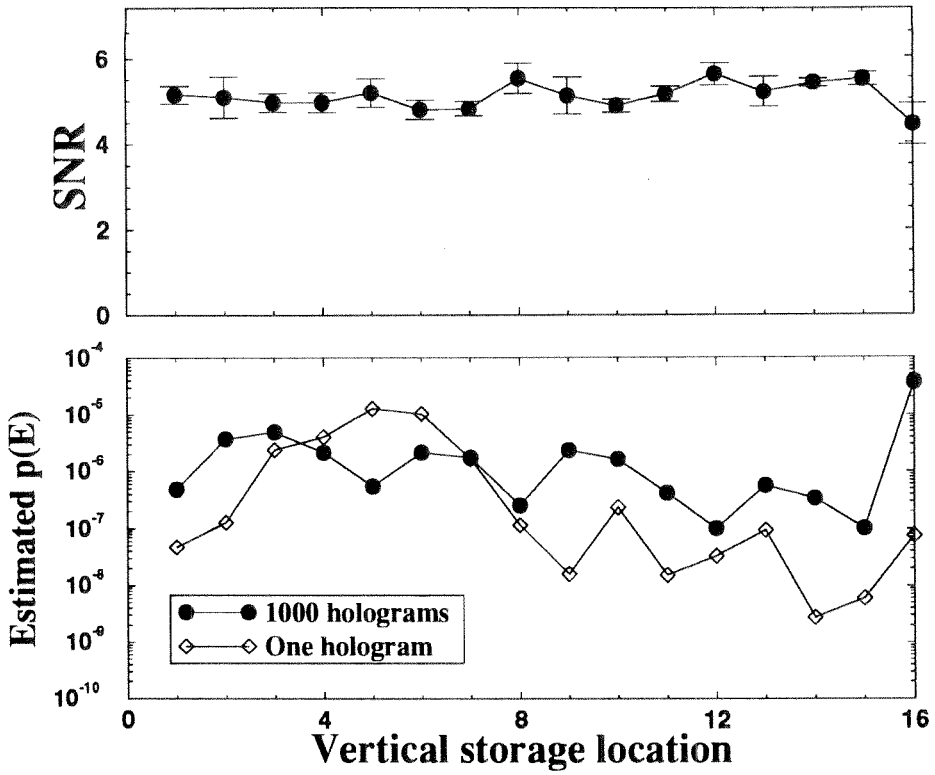


Figure 6.16: SNR as a function of hologram location.

large, not very much tolerance was left between the outside mirror strips and the edge of the tile, which meant that when the tiles were assembled, some of these edge mirror strips became damaged or distorted. In some places this appeared as small blemishes; in others, an entire mirror strip would be unusable. These outside mirror strips, unfortunately, are exactly those which direct reference beams to the top and bottom location.

A second difficulty was the passage of the deflected reference beams from the mirror array to the crystal. As we mentioned in Chapter 5.3.3, the limited aperture of the lens f_{out} affects the number of fractal rows that we can use, as well as the range of multiplexing angles available in each fractal row. Each location has 16 mirror strips assigned to it, evenly distributed across the mirror array. As we would expect from our discussion in Chapter 5.3.2, reference beams from the lower mirror strips work well if they are deflected towards the top location. This makes sense because they pass through the center of lens f_{out}, on the way from the bottom of the mirror array

to the top of the crystal stack. Of course, because of the vertical size of the SLM, we must skip every other mirror strip when assigning fractal rows.[9] As we use higher mirror strips, the reference beam passes through lens f_{out} farther from the optical axis and aberrations begin to distort the reference beam.

If we had used a plano–convex lens as lens f_{out}, we would have only had three fractal rows for angle multiplexing at the top and bottom locations. In order to reach our target of 10,000 holograms per location, we would have needed to store 3,333 holograms in each fractal row. However, the aberrations of the plano–convex lens also affect the horizontal deflection of the reference beam, adding complications to the spacing of the holograms. We were able to get around these difficulties by using a Fresnel lens of 70mm focal length and 70mm square aperture. The advantages were that our horizontal angle multiplexing range and the number of fractal rows increased dramatically. Even though the reference beam no longer resembled a single plane wave, we did not see any dramatic difference in angular selectivity. However, we did observe that we were unable to remove the crystal, reposition it in place, and find previously written holograms. This implies that the holograms correlate with the complex reference beam, and that any small shift in the crystal placement causes poor cross–correlation. As a result, we were unable to try fixing holograms that were written with the Fresnel lens.

With the Fresnel lens, however, we were able to use 4 fractal rows with 2500 holograms each, spaced at 10μm of linear translation.[10] For comparison, the theoretical angle selectivity for a 1.8mm interaction length is 2.71×10^{-5} radians, or 1.9μm on the mirror array. The experimental angular selectivity we measured was around 5–6μm of movement, with vestigial secondary peaks. In justification of our angular spacing, we show the Bragg mismatching of one of 10,000 holograms in Figure 6.17. This sequence of images covers approximately half of the distance between the Bragg–matching condition and the next hologram (5μm). In Figure 6.18, we show the mean

[9] If we were not to do this, then we would get three holograms overlapping at the detector array: one centered, one in the upper half of the screen, and one in the lower half of the screen. We discuss this briefly in Section 5.3.3.

[10] Translation of the focused spot on the mirror array.

of the ON pixel regions, the optimum threshold, the SNR, and probability of error for these reconstructions. We describe the complete SNR analysis procedure later in this section. These two figures show that crosstalk was not a significant factor in our experiment.

Preparation of the object arm One of the main differences between the 10,000 hologram experiment described in Section 6.3.2 and this experiment was the vertical extent of the storage location. For the 10,000 holograms, we used a reference beam that was larger than the signal beam in order to guarantee overlap. (Of course, not too large or we would have wasted optical power and made our diffraction efficiencies seem poor). For the demonstration of the 160,000 hologram system, the height of the reference beam needed to be close to the vertical spacing between locations, or about 3mm. We said in Chapter 5 that we could overlap the object beam using only 3mm, but we had to store at the Fourier plane. Our previous 10,000 hologram experiments had all been in the Fresnel plane because we did not have our own random phase plate to spread out the DC (zero spatial frequency) power of the object beam. So to demonstrate the 160,000 hologram system, we fabricated a random phase plate.

We discuss the fabrication of the phase plate in Section 6.5.3. Here we discuss how we integrated the random phase plate into our system. We found that even transition widths of 6μm between the regions of 0 and π phase showed up as dark lines on the detected images. For example, in Figure 6.19, we show a blank screen imaged through the system, as well as one of the 10,000 holograms: a grey–scale image of faces. The random pattern of the phase plate is clearly visible. This particular phase plate contained one phase region for every 8×8 SLM pixels. We also tried 2×2 and 4×4 phase plates, but chose to use an 8×8 phase plate because the transition widths were the smallest, there was no speckle in the transmitted images, and we could hide the phase plate lines through careful choice of display patterns.

We were able to keep the lines of the random phase plate transitions from showing up in our reconstructions by turning off the pixels on the SLM in a series of cross–hatched lines which exactly overlapped the phase plate transitions. After this, we had

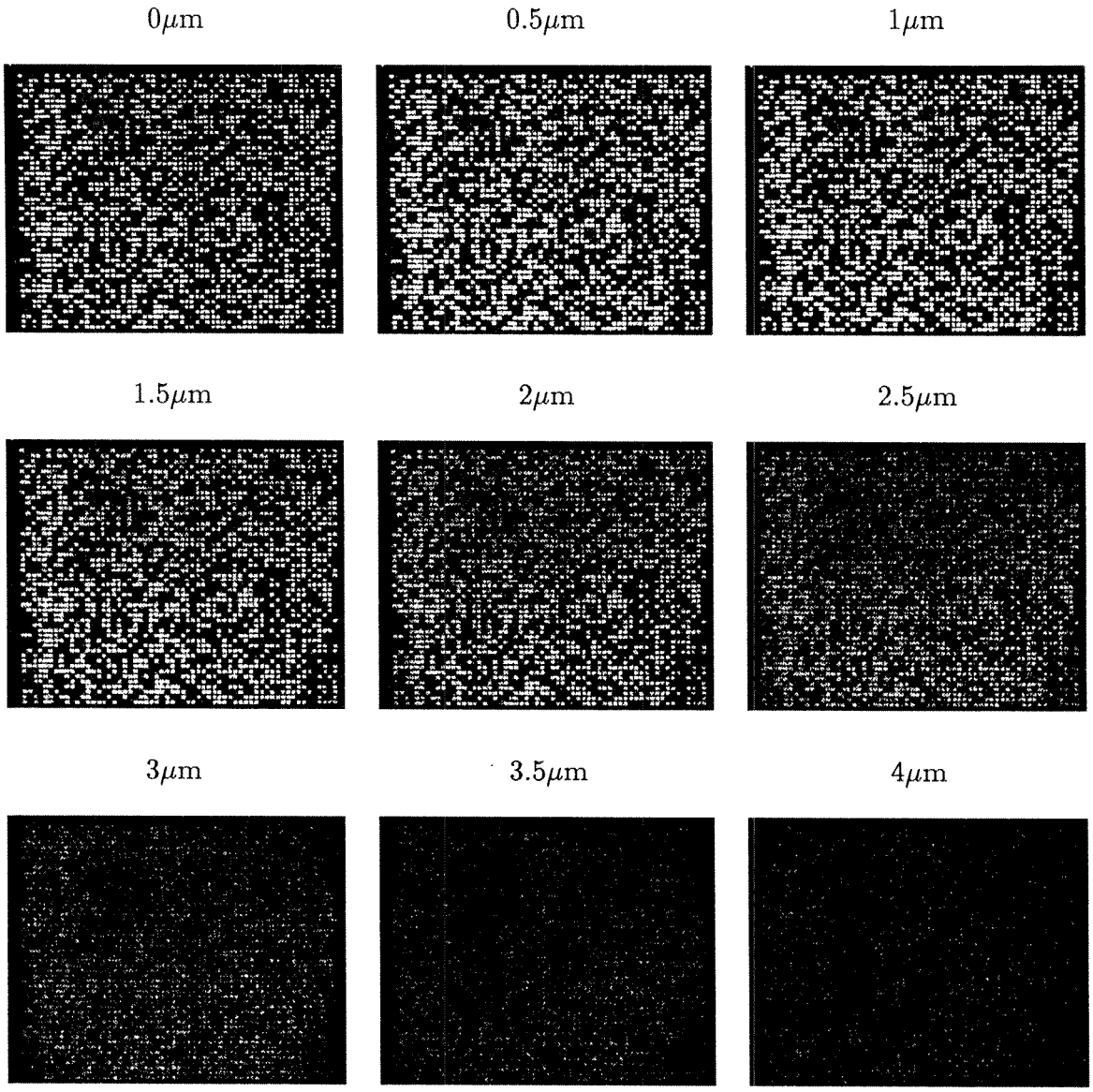


Figure 6.17: Bragg–mismatch of one of 10,000 holograms by translation of the focussed spot on the mirror array.

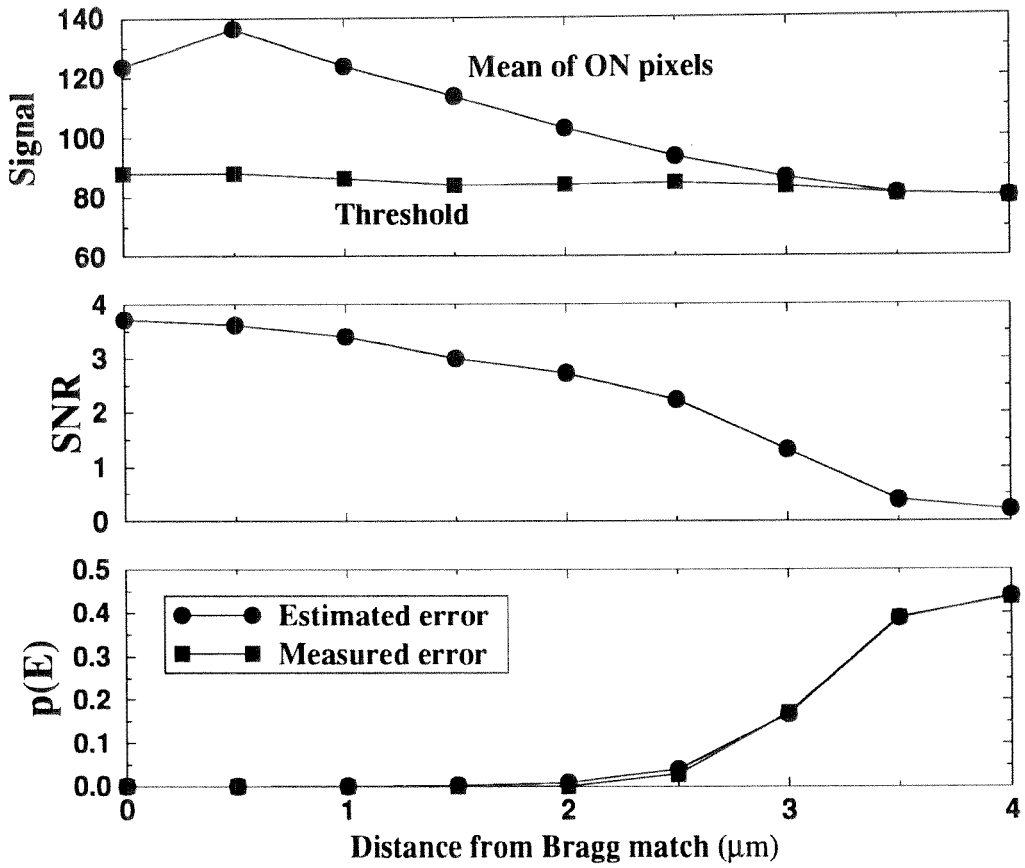


Figure 6.18: Analysis of images from Bragg–mismatching sequence shown in Figure 6.17. Plotted as a function of shift in focused spot (angle mismatch) are the mean of the ON pixel regions and value of the optimal threshold, the SNR, and measured and estimated probabilities of error.

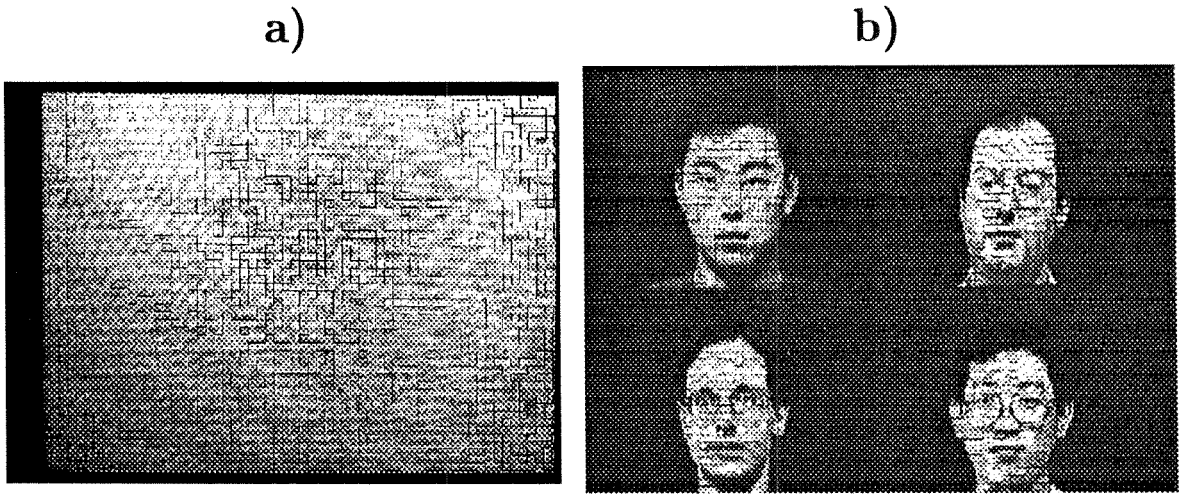


Figure 6.19: Images showing lines from random phase plate: a) image through crystal, b) one of 10,000 holograms at top location.

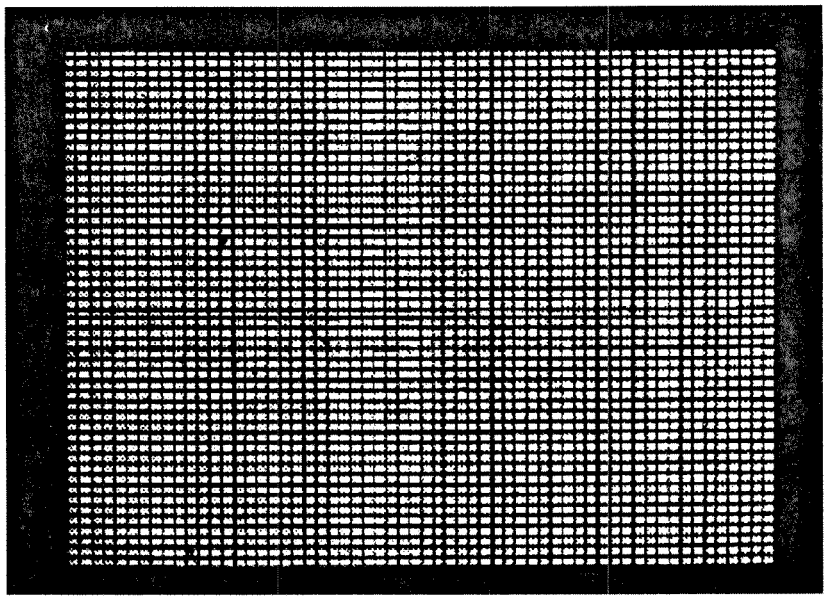


Figure 6.20: Set of 60 × 50 pixel regions used to avoid random phase plate lines.

a set of 60×50 pixel regions as shown in Figure 6.20.[11] The phase plates were made with masks that were slightly too small—the most noticeable effect of this is that one of the rows near the center of the pattern is thinner than the rest. The mismatch in the horizontal direction is not as noticeable since we had more flexible control over the columns than over the rows. We had 640 vertical columns to play with on the VGA screen, with only 200 rows. The pattern that we display on this VGA monitor gets mapped to a 480×400 windows on the 480×440 SLM and then imaged onto the center of the 60×55 phase plate.

We placed the random phase plate in the first image plane after the SLM, as shown in Figure 6.21. We put the object beam through a spatial filter and a polarizer before illuminating the random phase plate. If we had placed the polarizer after the phase plate, light would still strike the random phase plate transitions. Even though we had set those SLM pixels OFF, light was only modulated in polarization and would still illuminate the phase plate transitions. This light would have scattered

[11]Note that 50 vertical pixels show up instead the 55 that might be expected, since 440/8 is 55. The reason is that the display projector does not use vertical aliasing for the 640×200 VGA display mode we were using. The top 20 and bottom 20 rows of the display remain blank. Perhaps we could have gotten around this by programming in C instead of in BASICA.

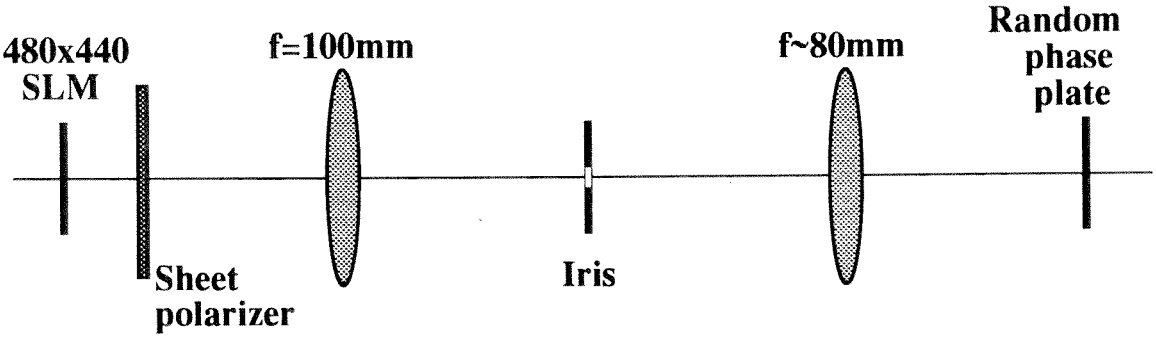


Figure 6.21: Object arm, shown between the SLM and the random phase plate.

throughout the object beam, causing distortions in the stored holograms. By placing the polarizer before the phase plate, we removed this fixed pattern noise light. The disadvantage to this was that we had to use a sheet polarizer, which was less efficient and introduced interference fringes in the object beam. In the past, we had used a polarizing beamsplitter at a later point in the system, placed on its side so that the unwanted light was deflected vertically (the extinction ratio in a PBS is better for transmitted light). Unfortunately, the 1 inch beamsplitter was too small to place before the random phase plate.[12].

The random phase plate and the SLM were each mounted on rotational stages which were then mounted on a 3–axis translation stages. The remainder of the object beam was as shown in Figure 6.14, where the plane marked as "Input SLM" corresponds to the plane where the demagnified SLM image overlapped the random phase plate. The lens f_{object} was of focal length 200mm and aperture 70mm, as was the lens immediately after the crystal. The reconstructed images were imaged and further demagnified onto the detector array with a 4–F system composed of an $f = 80$mm, 50mm aperture achromat and a f/1.6 camera lens "focused" on infinity (effective back focal length: 50mm). The camera was a STAR I cooled CCD camera, with 576×384 detector pixels. Each pixel was 23μm square, with unity fill factor. The reconstructed holograms we show below are 540×360 detector pixels in size, which includes a small

[12]We could have placed the polarizing beamsplitter near the Fourier transform plane. In this case, we would have experienced a lot of nonuniform amplitude modulation through "linear polarization rotation tilt" [356,357].

border around the holographic data.

Description of the experiment —We stored 30,000 holograms in a $LiNbO_3$:Fe crystal of dimensions 1.5cm × 2cm × 4cm. We used locations #1, #9, and #16, corresponding to the bottom, center, and top locations of our 160,000 hologram system. The crystal was fabricated from a 0.015% Fe–doped $LiNbO_3$ boule donated by Crystal Technologies, and was not anti–reflection coated. Because a 2 inch half–waveplate was not available, we used a 1 inch half–waveplate and translated it vertically to access the different locations. The edges of the crystal were masked off with electrical tape, and the entire object beam after the crystal sealed from all light save the exit aperture of the crystal. An iris was mounted on a two–axis translation stage for movement in the plane marked "IRIS Plane" in Figure 6.14. This was translated to the appropriate height for reconstruction of holograms from a particular location. The camera received a larger amount of scattered light if this iris was not in place. Although the holograms were still easily observable, we closed down the iris to cut down on the scattered light. Although a moving iris is not a practical solution in a fast memory, a fixed array of slits, aligned with the Fourier transforms of the various locations, might be a reasonable compromise for a practical system.

Each set of 10,000 holograms was stored with 4 fractal rows of 2500 holograms each. Within each fractal row, all holograms were spaced by 10μm. For most fractal rows, the translation of the focused spot spanned a horizontal translation range of -12.5mm to +12.5mm, where 0mm corresponds to the horizontal optical axis of the Fresnel lens. A few mirror strips, however, were used from -15mm to +10mm to avoid damaged spots on the mirror strips. The limit on the angle multiplexing range in our experiment was the aperture of the 1 inch waveplate (which was placed 2mm from the crystal surface). With a 2in waveplate, we could have used an angle range from -25mm to +25mm. Since the angle spacing between holograms was comfortably large, we had more than enough reference beams for 20,000 or even 25,000 holograms per location.

Dynamic range, however, was a different matter. We reduced the background

scattering by eliminating all stray light from the object beam path for reconstruction. In order to use all of the horizontal extent of the crystal within each vertical location, we deflected the object beam horizontally to spread the 10,000 holograms throughout the crystal. We had hoped that this deflection would avoid any distortion caused by overexposure at the focused DC spot of the image. Without the random phase plate, however, we observed an unacceptable amount of distortion and stray light in the regions of OFF pixels. With the random phase plate, even though it only contained one phase pixel per every 8 SLM pixels, the power in the DC was dispersed enough that the distortion disappeared. We deflected the signal horizontally across the width of the crystal, completing 50 complete cycles during the storage of 10,000 holograms. Since the reference beam was slightly larger than the size of the focused spot, we deflected the object beam vertically within the storage location by a small random amount along with each horizontal movement. This vertical deflection of the object beam was constrained to be within reasonably tight limits, so that the object and reference achieved good overlap for all holograms.

We observed no strong trend in diffraction efficiency between holograms written with object beams close to the reference beam and those written on the far side of the crystal. This would be expected for most of the holograms in the schedule, since holograms remote from the entrance face of the reference beam would have poorer modulation depth yet slower erasure from the reference beam. We would expect that only the holograms written in the last pass of the signal beam through the crystal would be unequal in diffraction efficiency, since for these holograms, little or no erasure occurred between recording and readout. This nonuniformity can be easily counteracted by minor adjustment of the recording times for these last few holograms.

Analysis of the reconstructions We captured reconstructions from the center, top, and bottom locations, using a 1 second exposure with the STAR camera. These included several chessboard images, the Caltech logo, and a mosaic of faces. Most reconstructions, however, were random data pages based on the 60×50 grid described above. Of these images, there were 23 reconstructions from the center location, 22

from the bottom, and 28 from the top. Many of these reconstructions are shown in Figures 6.22, 6.23, and 6.24. One original image—a random page imaged through the crystal—is shown in Figure 6.22 for comparison. The characteristics of these data images are plotted in Figure 6.25, including the means of the ON and OFF pixel regions, the optimal threshold, the SNR, and estimated probability of error of the reconstructed holograms. The holograms chosen for analysis were well distributed among the 30,000 holograms.

In the experiments we described previously, we analyzed SNR by treating each detector pixel within a region as a separate sample. By using large regions of ON and OFF pixels and throwing out edge pixels, we simplified the assignment of locations and were assured that the SLM did not introduce electrical crosstalk noise. With the lines introduced by the phase plate, however, we could no longer use large regions. For this experiment, therefore, we decided to average all of the detector pixels within each pixel region to create sample points for analysis. This gives us, on average, 1500 ON pixel regions and 1500 OFF regions. The effects of SLM nonuniformity across a pixel region get averaged out. However, we still needed to assign the ON and OFF regions—to draw a box around each region. We cannot use boxes which fill the detector area, though, because the ON regions are smaller than the spacing (the fill factor is less than 1). A further complication is a variance in the size of the pixel regions—caused by the mismatch between the phase plate and the SLM image. Finally, a slight rotation and bending of the images occurred at the extreme locations. It is not immediately observable in the images we show in Figures 6.23, but you can see it if you look for it. We might hope to use a "grid" of pixel regions which is aligned with the rows and columns of detector pixels. However, even a mild rotation of the SLM image causes this grid to fail to register with the pixel regions. In practice, the rotation could be avoided by a routine for rapid SNR analysis in the laboratory, perhaps configured to automatically translate and rotate the SLM and detector until they are registered to each other.

Despite the slight distortions of our images, we wanted a quantitative measure of our holograms. However, given a reconstructed image from a particular location,

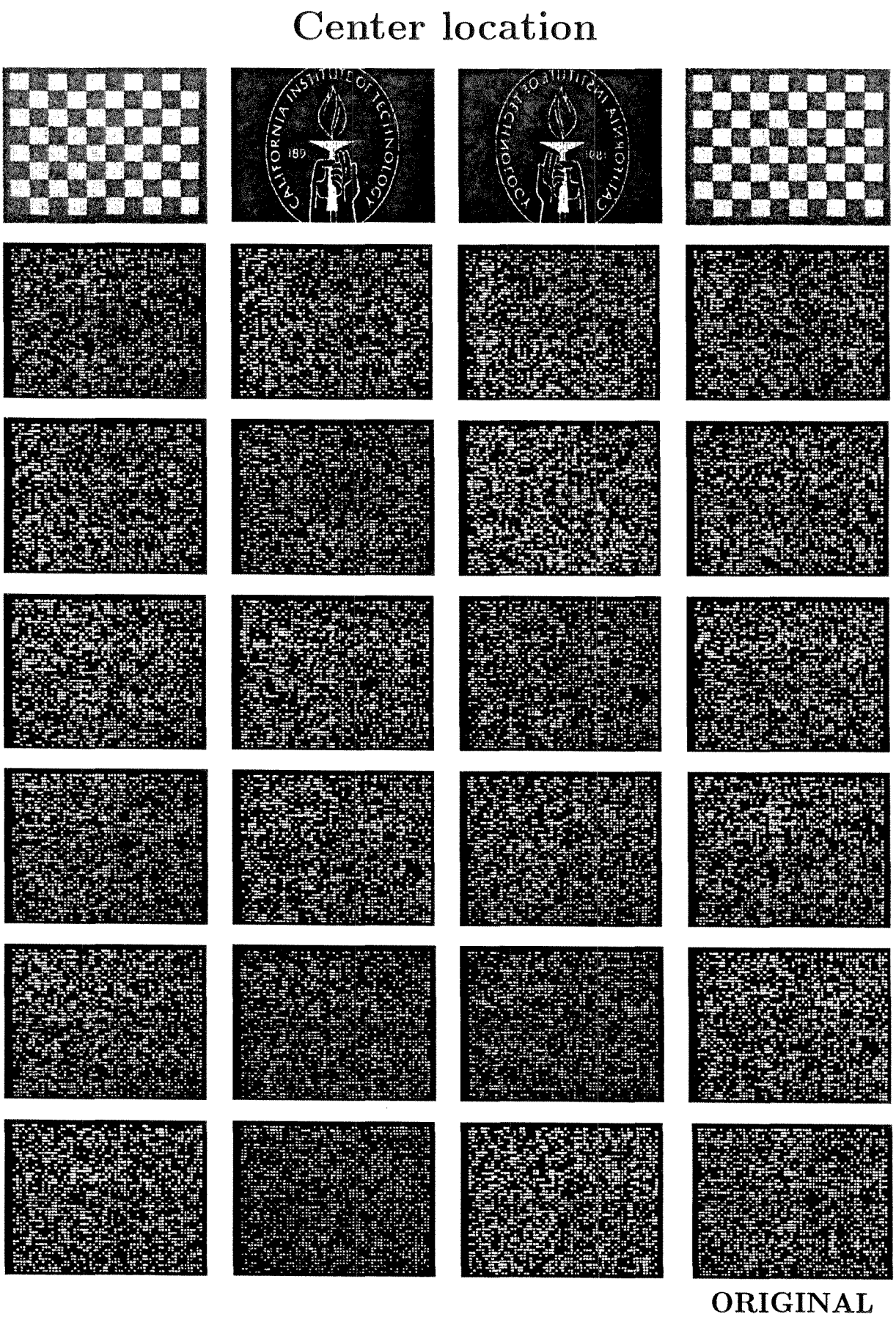


Figure 6.22: Storage of 10,000 holograms at center location (#9)

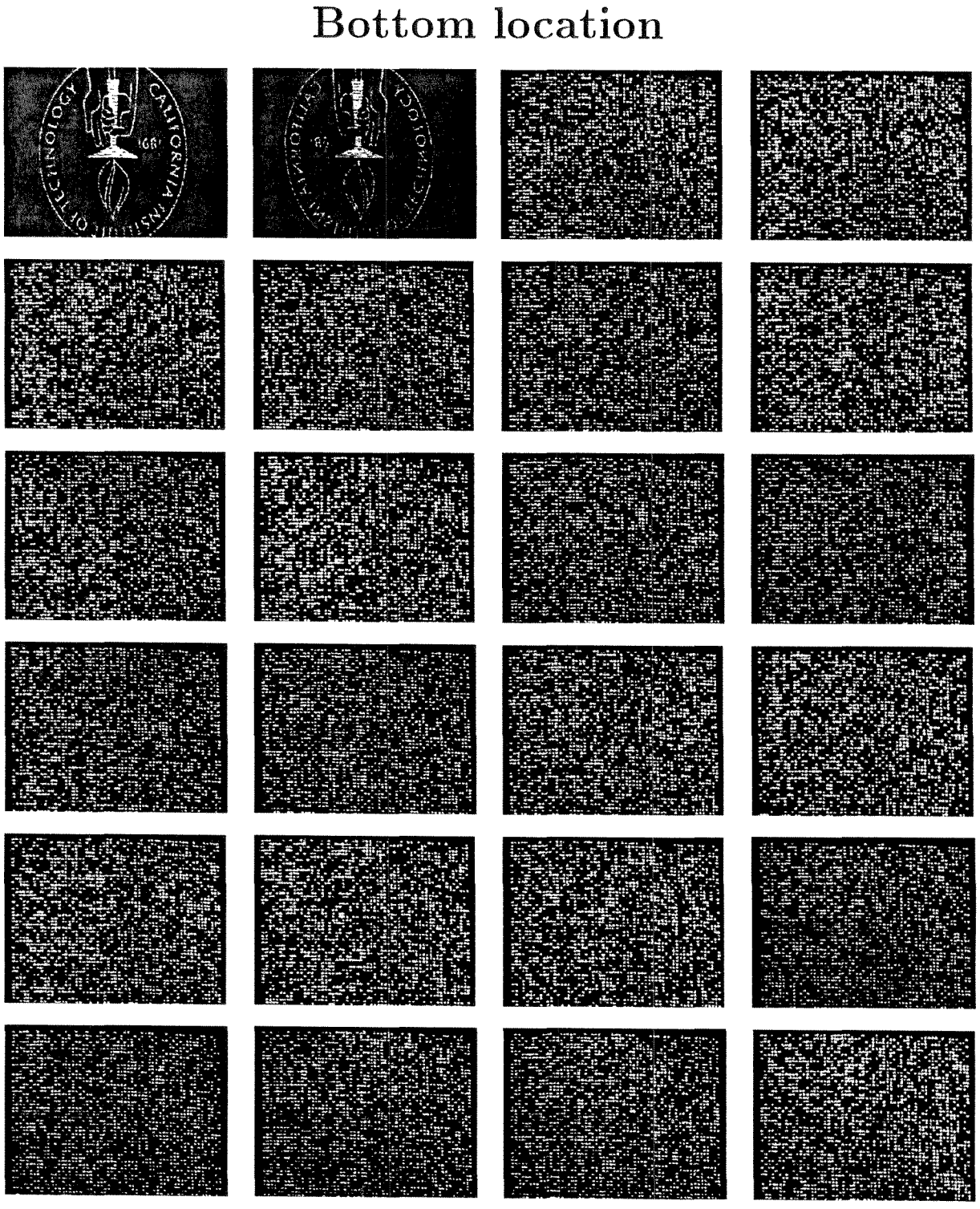


Figure 6.23: Storage of 10,000 holograms at bottom location (#1)

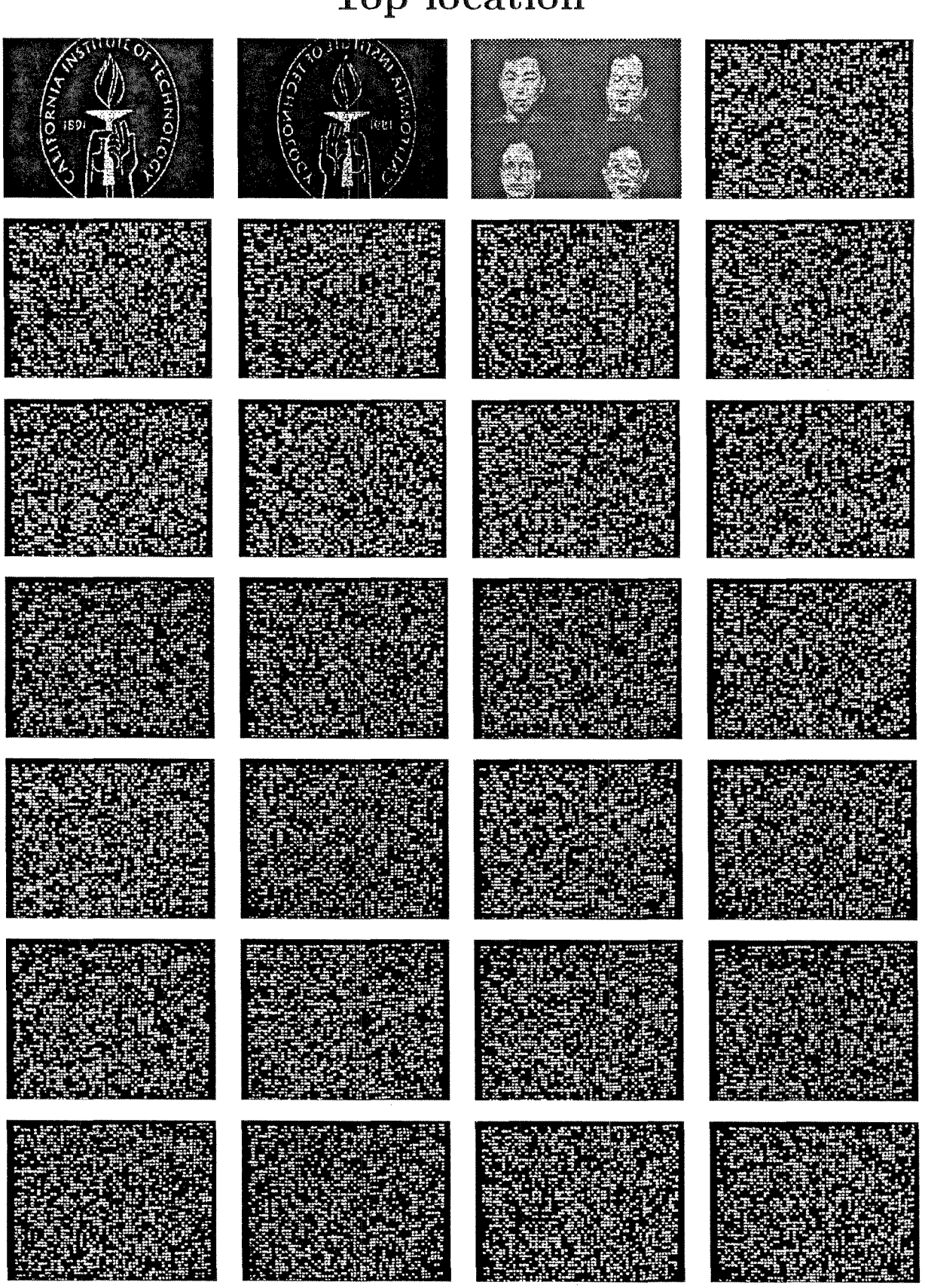


Figure 6.24: Storage of 10,000 holograms at top location (#16)

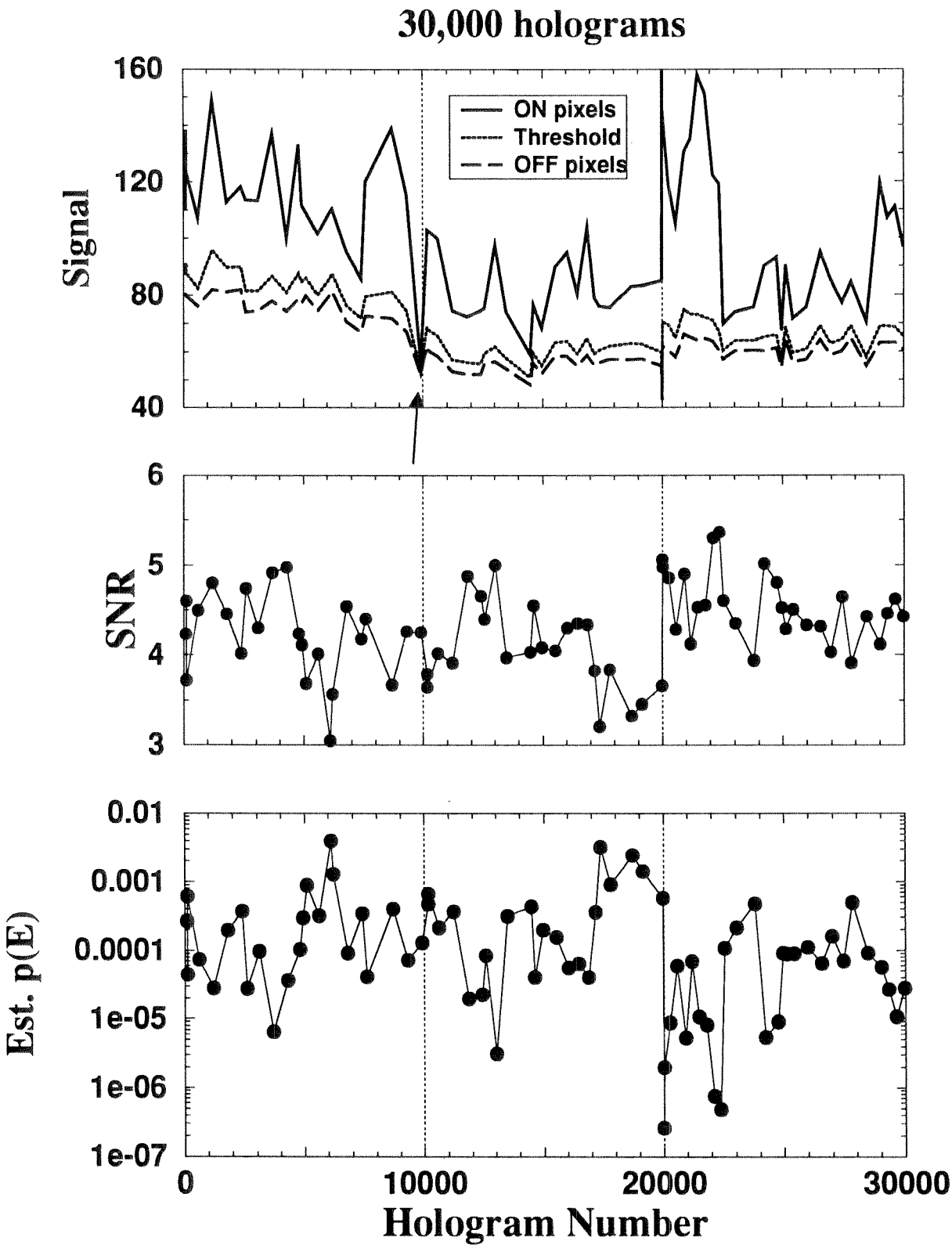


Figure 6.25: Analysis of data-bearing images from Figures 6.22–6.24. Plotted as a function of hologram number are the means of the ON and OFF pixel regions, value of the optimal threshold, SNR, and both measured and estimated probability of error.

each of the 3000 possible ON pixel regions shows up at some unknown box of detector pixels. The relative spacing between each box, and the relative width of the boxes does not change. For example, we might know that column #25 always has narrower boxes than the other columns, or that the spacing between row #16 and its neighbors is always a bit more than between the other rows. The only thing which changed was where these columns and rows of boxes appear, and whether they might be bent a little. We devised a computer analysis method in which we manually aligned a grid pattern from one of the reconstructed images, and then used it to analyze all the other reconstructions from the same location. We first manually aligned 42 boxes in a sparse pattern as shown in Figure 6.26, matching these against one of the captured reconstructions. In terms of the grid of 3000 boxes, these boxes corresponded to the intersections of the 1^{st}, 10^{th}, 20^{th}, 30^{th}, 40^{th}, 50^{th}, and 60^{th} columns with the 1^{st}, 10^{th}, 20^{th}, 30^{th}, 40^{th}, and 50^{th} rows. After these were set, we interpolated the full pattern of 3000 boxes using our knowledge of the relative spacing and widths of the various rows and columns. The center of each box was determined by interpolation; the width and height set to 70% of the distance to the nearest neighbor box. Each box coordinate is then rounded to the nearest integer. At this point, we have a grid of 3000 boxes—an example of one grid from the top location is shown in Figure 6.27.

The next step was to divide this appropriately into the classes of ON and OFF pixels. We did not want to choose the ON and OFF pixels from the reconstructions, since we wanted to know that we were comparing the results against what had been stored. Instead, before the experiment began, we created 100 random patterns to be displayed on the SLM and saved the 3000 bit sequence associated with each.[13] Then, when we went to analyze each reconstruction, we first determined which of the 100 pages were represented by visual inspection and comparison with our small database of bit sequences. This gave us the correct assignment of the grid of 3000 boxes into the classes of ON and OFF regions.

[13]We could not have practically performed the experiment with random images constructed as each hologram was stored, since the process of drawing a 640×200 screen pixel–by–pixel in BASICA takes 30–40 seconds. In comparison, the display of a pre–stored image with the BLOAD command takes less than half a second.

We found that the reconstructions tended to be shifted by a few detector pixels from each other because they were stored at different locations in the crystal. As we mentioned above, the deflecting mirrors were not in exactly the image plane, so we got small movements of the image plane along with larger movements of the focused image beam in the Fourier plane. For each reconstruction, we shifted the entire grid by a few detector pixels in either direction to find the maximum SNR. We tried to use the same set of 3000 boxes for all of the reconstructions from a particular location, keeping the interbox spacings identical. For instance, all of the reconstructions in Figure 6.23 were analyzed with one grid pattern. For the other locations, we occasionally needed to realign the sparse grid of 42 boxes and then reinterpolate. If we did not realign the grid, the SNR would tend to drop to between 2.5 and 3. We do not believe that this change in interbox spacing was due to changes in the aberrations encountered, since reconstructions from the center location needed realignment as often as those from the top location. Instead, the change in the grid patterns probably stemmed from the slightly varying optical path lengths encountered as the signal beam was moved back and forth within the storage location. This occurred even though the crystal was professionally polished. The solution to this problem is phase–conjugate readout, which we discuss in Chapter 8.3.

Note that one of the reconstructions in Figure 6.25 (marked with an arrow) appears to be much weaker than the others. This particular reconstruction was captured with a 0.2 second exposure. It is important to note that the SNR and probability of error are unaffected by the shorter exposure time. This means that we could have used this shorter exposure for all of the holograms. The reason we did not is that we wanted enough dynamic range in the reconstructions to be able to display the images in PostScript greyscale. In terms of a practical system, the application of the threshold should take place in analog at or close to the pixel. This reduces the I/O output of the pixel to exactly the number of information bits and reduces the quantization error of the thresholding operation to the resolution of the analog comparator.

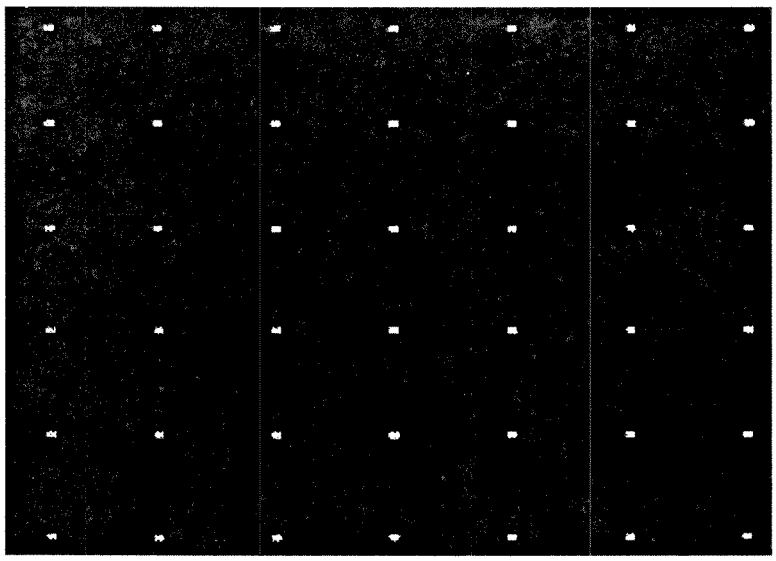


Figure 6.26: Sparse grid of 42 pixel regions aligned manually and then used for interpolation.

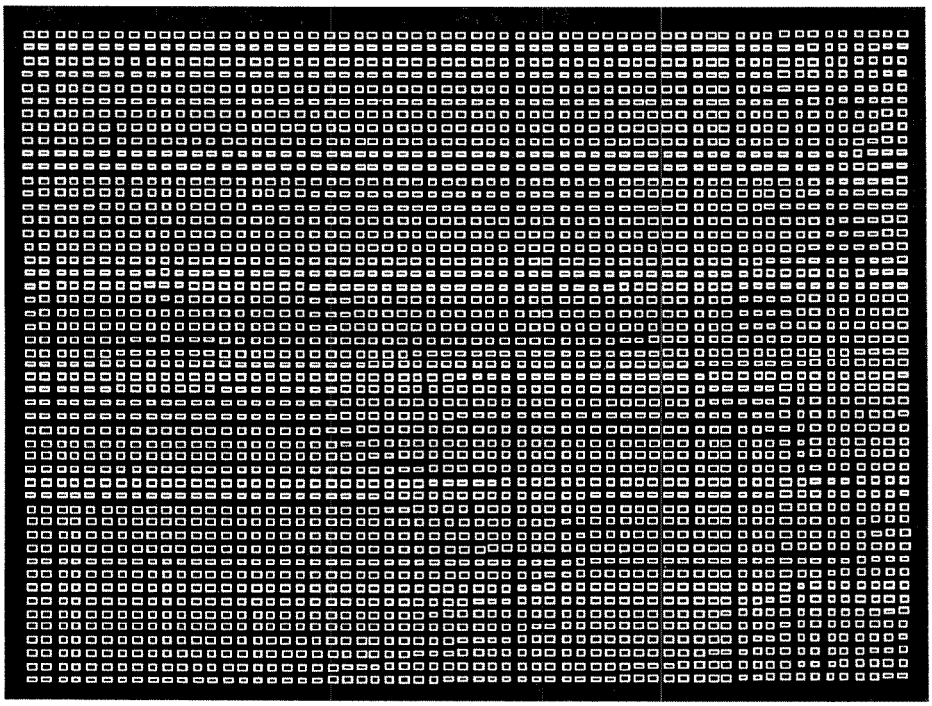


Figure 6.27: Grid of 3000 pixel regions formed by interpolating the sparse grid pattern.

6.5 Miscellaneous experiments

6.5.1 Simultaneous memory and correlator

One of the other applications that we demonstrated with our memory setup was the implementation of real–time face recognition [250], where stored holograms are used as correlation templates. Holograms are read out with the signal beam, reconstructing the appropriate reference beams. The strength of each reference beam is proportional to the correlation between the image used originally for storage, and the image being displayed in real–time on the SLM. The reference beams can be focused onto a detector, forming correlation peaks. Because the holograms have Bragg–degeneracy in the vertical direction, the system has vertical shift invariance. In other words, the input face can be shifted vertically and the correlation peak shifts along with it. If the input face is shifted horizontally, then the correlation peak disappears through Bragg mismatch.

We used a setup similar to the 160,000 hologram setup for this experiment. As shown in Figure 6.28, we used the mirror array to direct the reference beams and our standard object beam for deflecting the SLM signal. We mounted the SLM sideways so most of the natural side–to–side face motion would be vertical motion in the system, for which we had shift invariance. Our first experiment was the demonstration of memory storage in one location and face–recognition at a neighboring location. First, we stored holograms in the upper portion of our crystal.[14] The Fourier transform plane of the object beam was in front of the crystal. After these holograms were stored, we changed the incoming portion of the object beam of the system to prepare for correlation.[15]

In the object beam, we placed a DC block in the Fourier transform plane of the SLM. This spatially filters the images presented to the crystal, both during storage and during real–time correlation. Without a DC block, it is difficult for the system

[14]The crystal was a 23mm × 23mm × 13mm 90° geometry crystal cut from a 0.015% Fe–doped $LiNbO_3$ boule donated by Crystal Technologies.

[15]The memory holograms don't care if the front end of the object beam is still there or not, as long as the remainder of the readout arm is untouched.

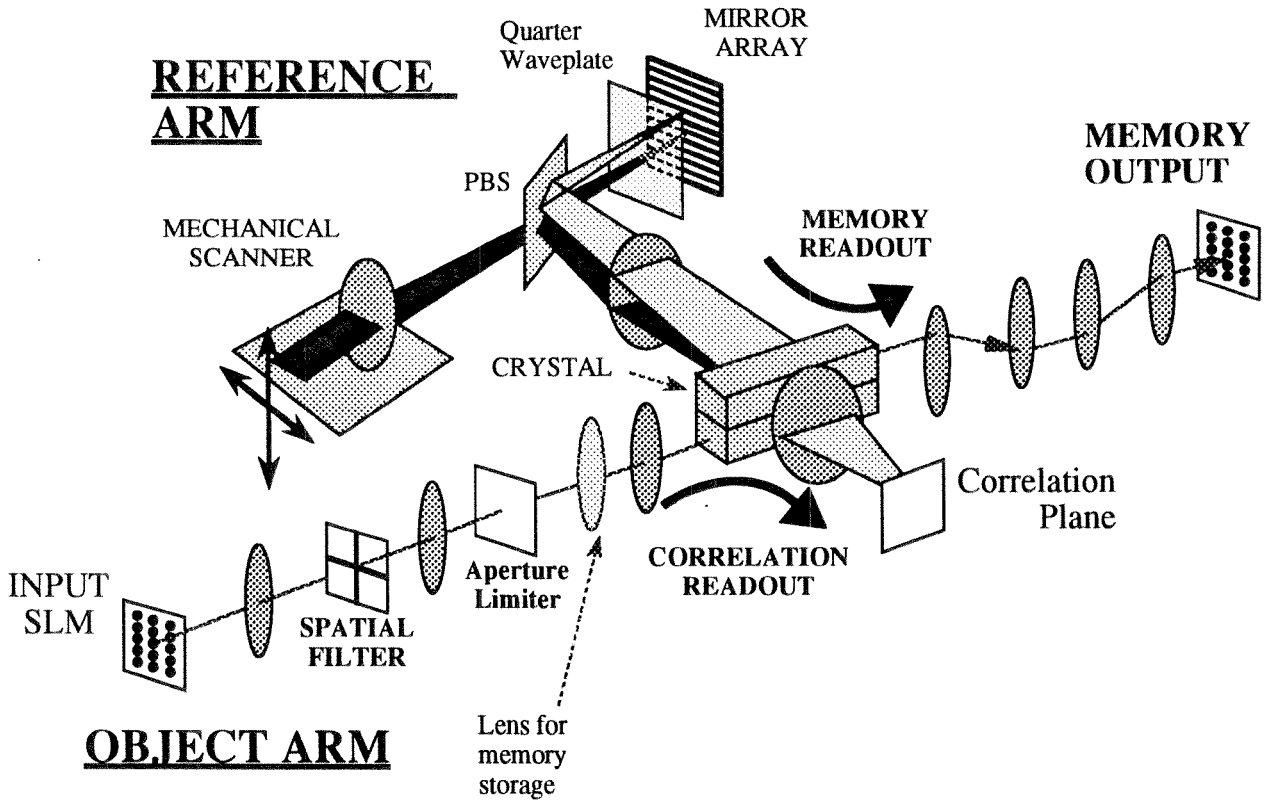


Figure 6.28: Simultaneous memory and correlator: experimental setup

to distinguish one image from another, because each image contains so much DC energy. Although the edge–enhanced images contain much less energy, the discrimination performance is significantly improved. We placed a rectangular aperture in a subsequent image plane to block the edge–enhanced borders of the picture and further improve the discrimination performance. Finally, we replaced lens L_3 with a shorter focal length lens. Although this makes it impossible to have the SLM image plane coincide with the back focal plane of lens L_3, we can place the Fourier transform of the edge–enhanced image directly within the crystal. If the crystal is not exactly at this Fourier transform plane, then the shift invariance is reduced. There was not much that we needed to do in the reference arm, because the mirror array allows us to store at a different location simply by translating the mechanical scanner.

We used several fractal rows to store correlation templates. In this first demonstration, we had a correlation database of 60 separate images each of 4 people. We used one fractal row for each person. To read out the correlation peaks, we placed a f/1.6 camera lens at the side face of the crystal as shown in Figure 6.28 and a CCD detector at its focal plane. Note that the correlation plane is the image of the mirror array surface—therefore we expect that the different fractal rows will be separated by some distance (corresponding to the mirror strips for the 15 other locations). We show this in Figure 6.29, where there are 6 "lines" or fractal rows of correlation peaks. We'll keep referring to these as fractal rows even though they appear as columns in these pictures.[16] We were able to readout the memory and the correlation peaks at the same time by using the mirror array to deflect the reference beam back to the top location. In Figure 6.30, we show several examples of the output of the two CCD detectors: one portion of the picture shows the correlation plane, while the other shows the reconstructed memory. Note that since the correlation templates were taken from a sequential video sequence, an input face often correlates with several similar templates.

Having demonstrated memory and correlation, we concentrating on increasing

[16]Figure 6.29 was generated by an experiment which accidentally stored the same template for all the reference beams.

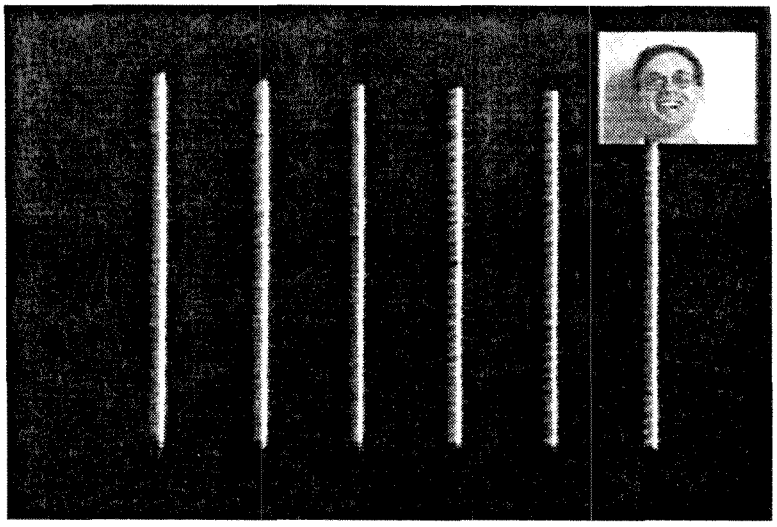


Figure 6.29: The correlation plane

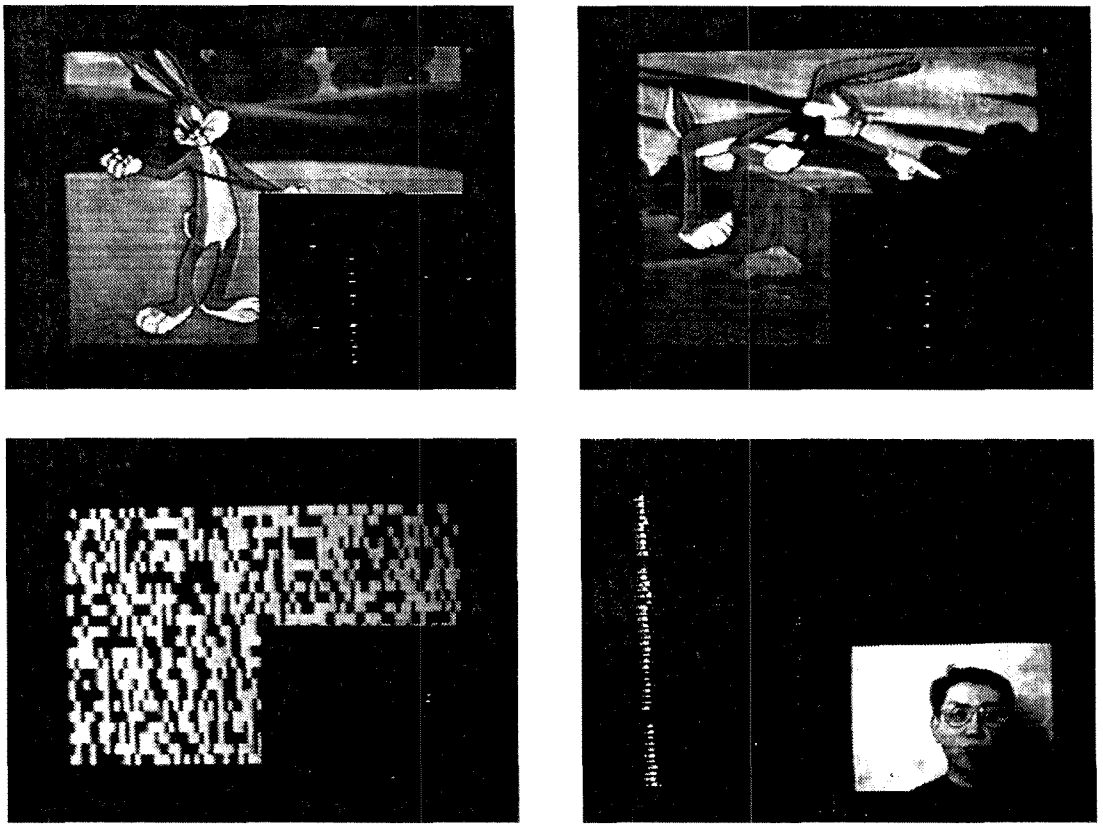


Figure 6.30: Simultaneous memory readout and real–time correlation.

the number of stored templates. We were not really limited by any geometric limits, despite the fact that the correlation plane detector does not cover the full image of the mirror array. The reason is that we have no need to distinguish individual correlation peaks within the same fractal row. We can then place them closer together than the angular selectivity. The limit to the number of templates is mostly dynamic range. We have an advantage here over memory readout, because the reconstructed wavefront is detected at its high–intensity focus. However, we're also at a disadvantage because the edge–enhanced input image input to the crystal contains very little power, on the order of 1-10μW.

We used the nonlinear filtering capabilities of the Imagepoint camera to enhance the correlation peaks, which essentially picks a threshold below which we cannot distinguish the peak. We were able to store 1200 correlation templates, arranged in 6 fractal rows of 200 templates each. As above, each fractal row corresponded to templates from a single individual. We show several examples of real–time correlation in Figure 6.31. In these images, the corner image is the input image and the remainder of the frame is the correlation plane.

6.5.2 Thermal fixing of multiple holograms

We used the 160,000 hologram setup (Section 6.4) to demonstrate fixing. In order to be able to find the holograms after replacement of the crystal, we used a glass plano–convex lens of 70mm focal length, 60mm aperture in the reference arm. We recorded 1000 holograms at the same 10μm reference beam spacing. After recording the holograms, we observed a few reconstructions. We removed the crystal from the apparatus and immediately replaced to observe the degradations caused by imperfect crystal positioning. We observed that, even with the glass lens, the uniformity of the holograms was seriously affected by the simple act of removing and replacing the crystal. This is probably an effect of the poor plane wave quality of the reference beam after 4 cylindrical lenses and could be corrected by the use of diffractive optics for beam shaping.

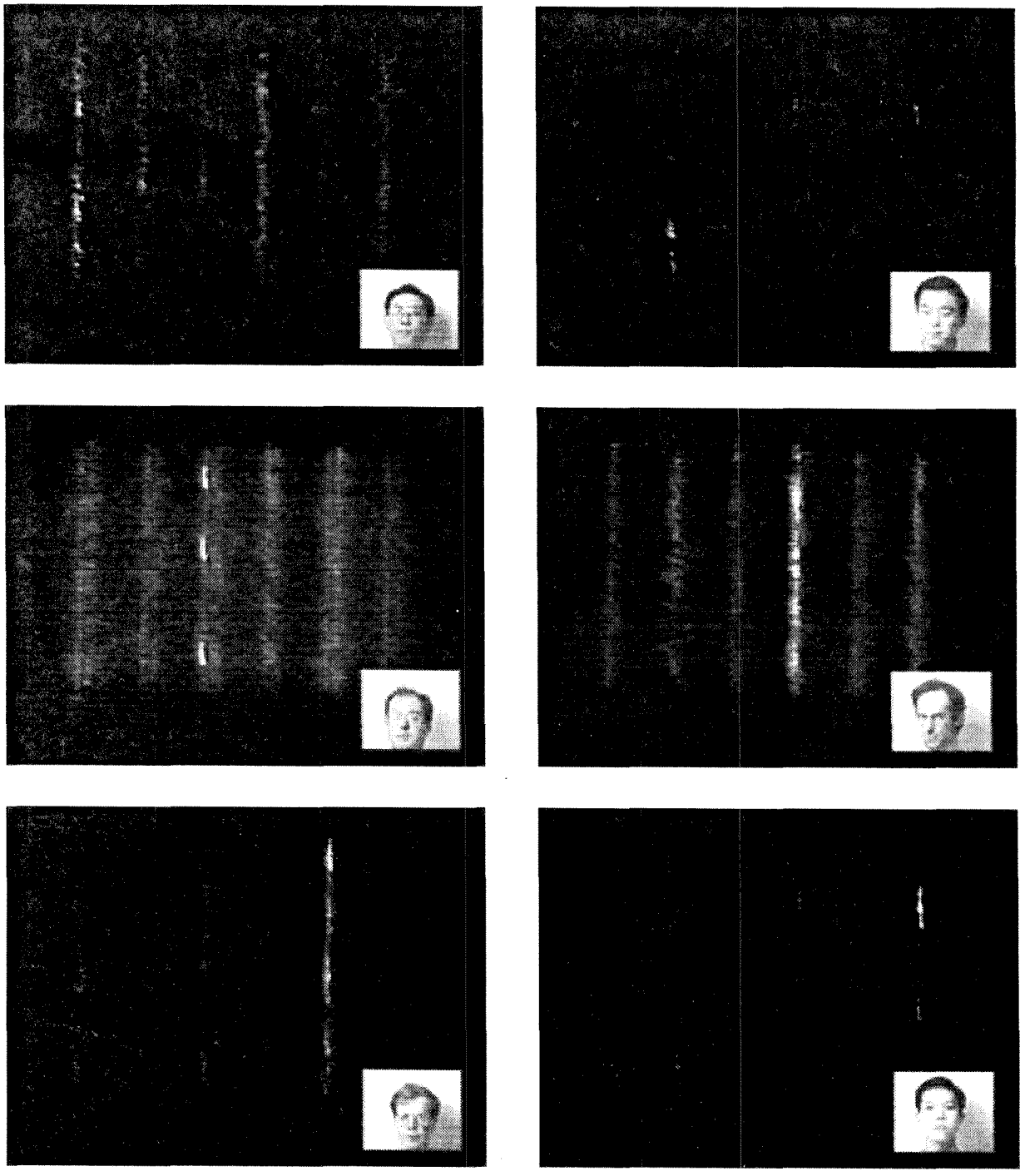


Figure 6.31: Demonstration of a real–time face correlator with 1200 face images

After removing the crystal again, we placed it in an oven for 30 minutes. The oven had been preheated to 120° C. After cooling, the crystal was replaced in the setup and illuminated. As expected, no holograms were immediately observed as the ionic gratings compensated the charge gratings Two experiments were performed: in the first, we illuminated the crystal (erasing the charge gratings) with UV illumination for 210 minutes. We observed that approximately 27% of the original diffraction efficiency was retained by the fixed holograms. The observed degradations of the images did not differ much from those seen after the first (non–fixed) replacement of the crystal. In the second experiment of 1000 holograms, we illuminated the crystal with the 30mW reference beam. The reference beam was swept through the full range continuously in order to avoid any fanning buildup which might occur. After 210 minutes, approximately 30% of the original diffraction efficiency was retained by the fixed holograms. Figure 6.32 shows reconstructions of fixed holograms from both of these experiments. The top 6 images correspond to the UV erasure experiment, and the bottom 6 to the reference beam erasure experiment.

No study of the SNR is shown because we found that it depended too strongly on the accuracy of the repositioning. When we found holograms which were unsatisfactory in fidelity, we could reposition the crystal and improve the uniformity and diffraction efficiency substantially. However, we have no way of knowing if the exact same holograms were improving, or whether we were looking at different holograms each time. In order to study the SNR of fixed holograms carefully, one must find a way to either fix in situ, or to reposition the crystal with very high accuracy.

6.5.3 Random–phase plate

This is the procedure we used to fabricate the random phase plates. Xin An and Geoff Burr made the masks used for lithography, including the determination of the required size. Annette Grot developed and performed the processing procedure.

The procedure has essentially three parts: Put a layer of photoresist onto a clean glass substrate, expose portions of the photoresist through a contact mask, and then

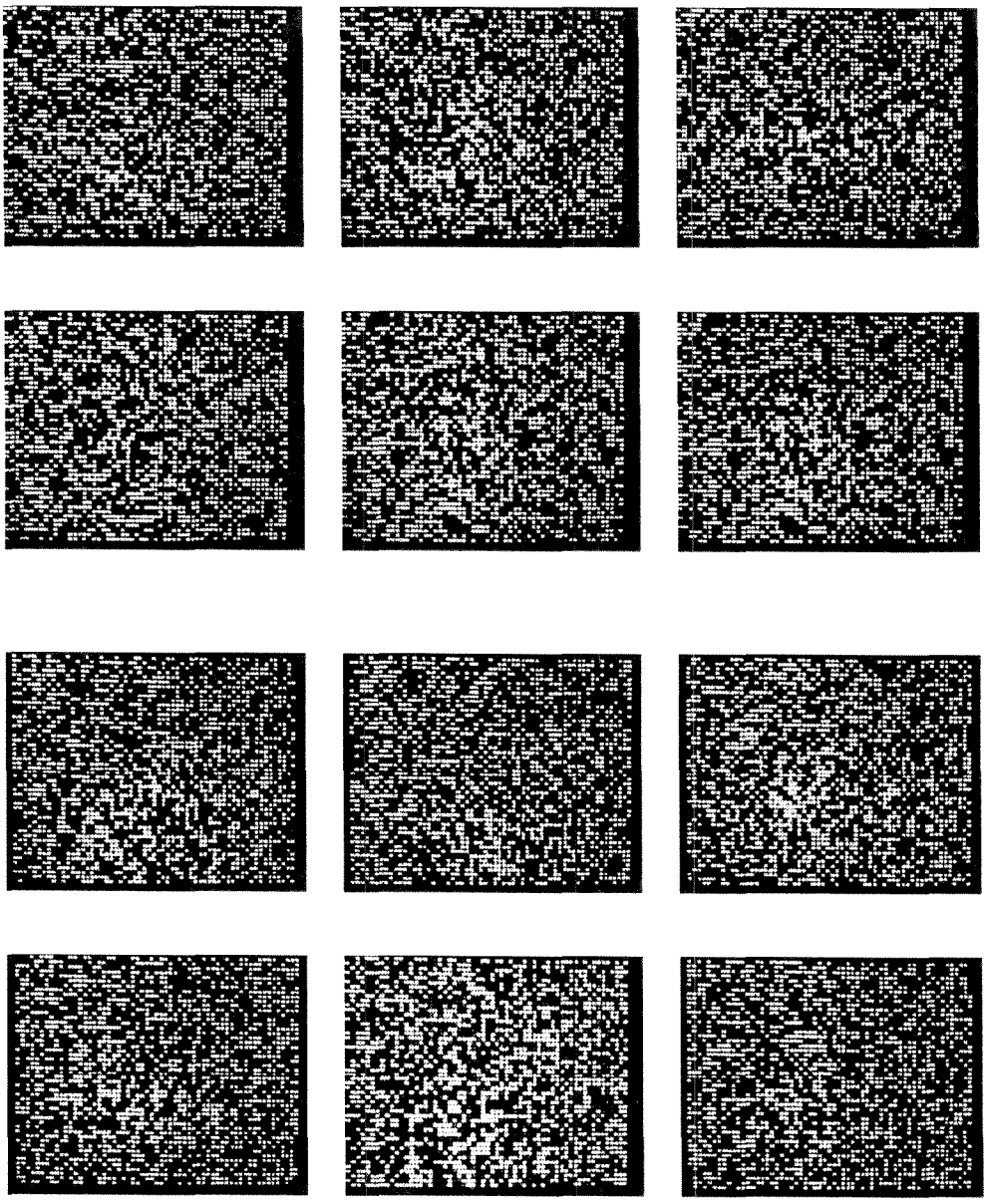


Figure 6.32: Storage of 1,000 fixed holograms. Shown are reconstructions from two separate experiments.

etch the glass in these regions with acid. The substrate can be cleaned with soap and water and then chromic sulfuric acid, rinsed, blown dry with N_2, and dried on a hot plate. Photoresist primer is spun on the substrate at 2000 rpm for 10 sec, followed by photoresist at 4500 rpm for 40 seconds. The photoresist is baked on a hot plate at 90°C for 90 seconds to remove some of the solvents.

At this point, the photoresist is ready to be exposed. We determined the size of the mask to use by placing a photoresist–coated glass plate in our optical system in the plane where the random phase plate was to go (Figure 6.21). The pattern of the fully illuminated SLM was then imprinted into the photoresist. After developing the photoresist, we used a Sloan Dektak II surface profilometer to determine the required spacing to within 1 part in 100. The resulting vertical spacing was 29.4μm; the horizontal spacing was 35.7μm. However, we needed to have these numbers accurate to better than 1 part in 500. To further zero in on the precise values, we had a mask commercially[17] printed on transparent film with a 3386 dot–per–inch printer. We then checked the alignment of the exposed photoresist (representing the image of the SLM) against the mask. This led to another cycle of mask–printing.

Once we got close to the desired mask, we etched the patterns into the glass plates as follows. A contact mask aligner (Karl Suss MJB3) was used to expose a photoresist–coated glass plate. The plate was developed for 2 minutes and then baked on the hot plate for 2 minutes. The quality of the photoresist pattern was checked under a microscope before etching. The baking of the photoresist retards undercutting of the acid from regions of low phase (where we want the glass to be etched) into regions of high phase (where we want the glass to be protected by photoresist). This undercutting was the primary source of soft transitions between the two phase levels. The etchant we used was 1:10 hydrofluoric acid: water, and the substrate was immersed for approximately 50 seconds to reach the target etch depth of 5000Å.[18]

The best transition width that we were able to fabricate was 6μm, as shown in

[17]The cost per page was <$25.

[18]We wanted $(n-1)d$ to be approximately $\lambda/2$, where λ was 488nm and $n \sim 1.5$.

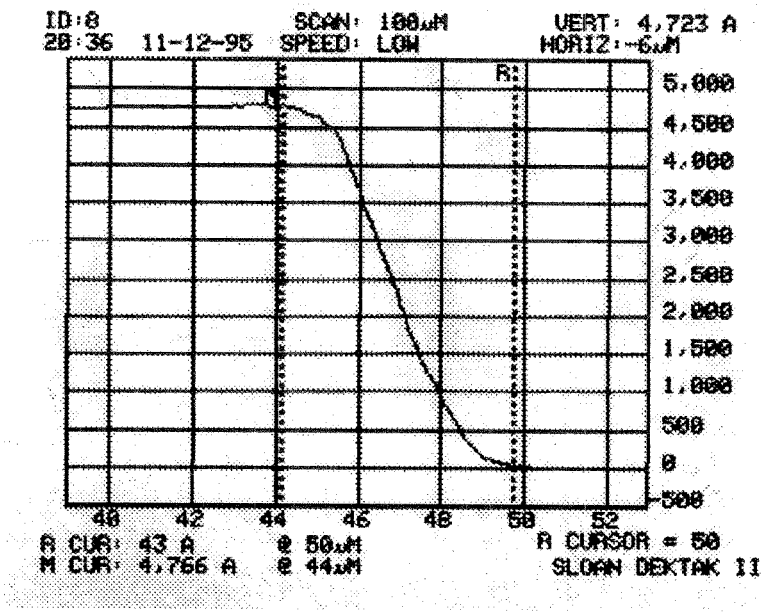


Figure 6.33: Surface profile of transition from π to 0 phase.

a Dektak trace in Figure 6.33. Note that a slight underetching of the entire pattern helped keep the transition width low—conversely, for slight overetching, the transition width became much larger. The transition shown in Figure 6.33 was measured on the 8×8 phase plate we used in our 30,000 hologram experiment. The 8×8 refers to the number of SLM pixels within each phase region. We fabricated several 4×4 and 2×2 phase plates, with transition widths ranging from 7–13μm.

When we placed the phase plates in the optical system, aligned them to the SLM image and observed the resulting image with the detector array, we found that we were unable to make the transition lines entirely disappear. We had originally expected that, if these lines were appearing to mismatch in the size of the mask, we would be able to translate the phase mask and observe movement in the visible phase plate lines. This would take place as some phase plate transitions moved into the gaps between SLM pixels while others moved out of these gaps. Instead, what we observed was that the visibility of the phase plate lines depended mostly on radial distance from the center of the SLM. If the lines in the center disappeared, those at the edge were quite visible. We also found that we could reverse this situation by refocusing the phase plate, making the transition lines in the center noticeable and those at

the edge invisible. In practice, we focused the phase plate to where the phase plate transitions were moderately visible in both the center and at the edges, and invisible in a ring about the center. This is the condition that is shown in Figure 6.19(a).

When we went through the process of determining which pixels on the display needed to be turned off in order to hide the phase plate transitions, we discovered that the phase plate was approximately 2 SLM pixels too small. By the time we got another mask made in preparation for making new phase plates, our experiments with the first were progressing so well that we were reluctant to make a change.

Chapter 7 System considerations

Contents

In this chapter, we review some of the system considerations in preparation for proposing a next–generation design which builds on the mirror array–based design demonstrated in Chapter 6. We begin by considering the time response of a holographic memory and reviewing the available techniques for non–volatile hologram readout. We review the available technologies for angle–deflection and discuss a novel device fabricated with bulk micromachining. We then propose system techniques for rapidly finding the Bragg–match condition through feedback on the reconstructed wavefront. Finally, we consider various ways of bringing signal and reference beams to a large number of storage locations.

7.1 Time Response

7.1.1 Recording

Because of the use of the recording schedule, the rate at which data is recorded into a holographic storage system is not constant. Despite this, we can address the factors which affect recording by considering an average recording rate. This average recording rate depends on the f factor from Chapter 2.2.3 and the erasure time τ_e. In turn, the erasure time is inversely dependent on the total intensity. If we would

like a fast recording rate, we need to have a fair amount of power and need to keep f moderate (around 0.6–0.9).

One of the problems with increased optical power is that the erasure rate increases commensurately with the increase in recording rate. This will cause us problems after our data is stored and we want to read it out without degradation. We describe the available non–volatile readout solutions in Section 7.4.

7.1.2 Access–limited Readout

If the signal power of each reconstructed wavefront is sufficient, then the readout rate will be limited by the access time of the wavefront controller. If the controller is a mechanical device, then the access time to a new wavefront in turn depends on the distance traveled. So we do not get random access to any hologram. If the worst–case access time is acceptable, there is no problem. We can get true random access by using non–mechanical deflectors such as acousto–optic deflectors or liquid crystal beamsteerers. We discuss some of the available mechanical and non–mechanical angle scanners in Section 7.5.

What happens if we have a mechanical scanner that has some vibration or settling at the end of its travel period—how will this affect our reconstructed hologram? If the vibration were fast enough, we might envision having trouble obtaining Bragg–match over the volume of the hologram. However, as long as the coherence length (corresponding to the frequency of the mechanical vibration) is longer than the interaction length of the grating, there should be no problem. For a 1cm thick crystal, the vibration rate needs to approach 30 GHz before there is a serious effect. On the low end, if the mechanical vibration rate is slower than the integration time of the detector, we won't be getting an average over several cycles of the vibration.[1] For any vibration frequency between these limits, we need only consider the magnitude of the oscillation of the reference beam angle around the Bragg–match condition.

We would like to know how much diffraction efficiency we lose because of this

[1] In this case, the measured diffraction efficiency will depend on where the integration period falls within the sinusoidal vibration, relative to the Bragg–match condition.

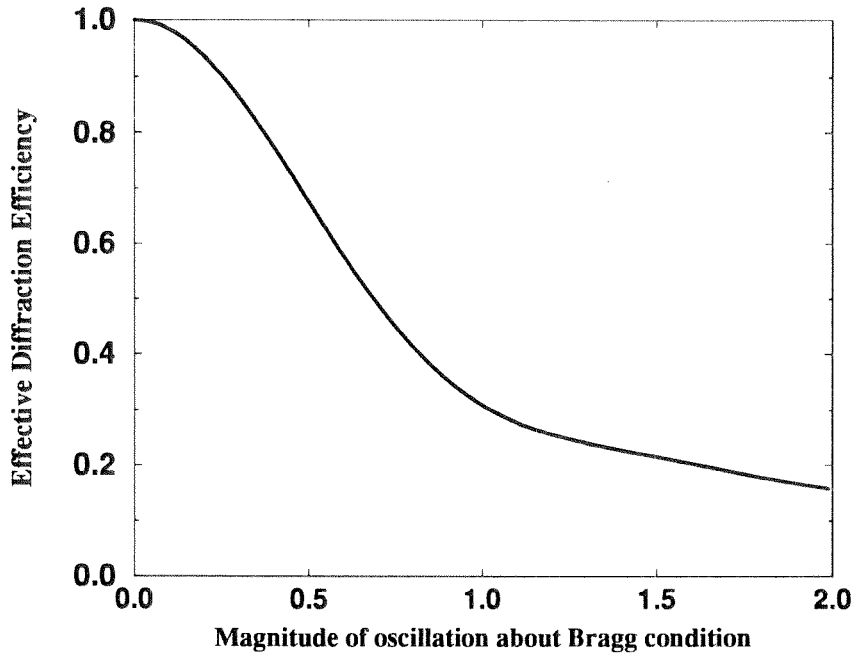


Figure 7.1: Time–averaged diffraction efficiency, relative to diffraction efficiency at Bragg match, as a function of the amplitude of a rapid oscillation in reference beam angle.

oscillation about the Bragg condition. For simplicity, we normalize the magnitude of the oscillation to the first–null angular selectivity. The time–averaged diffraction efficiency is the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{sinc}^2\left(x \sin t\right) dt. \tag{7.1}$$

We plot this diffraction efficiency as a function of x in Figure 7.1. An oscillation which varies about the Bragg condition by ± 0.25 of the Bragg selectivity gives us 90% of the Bragg–matched diffraction efficiency.

However, in addition to lost diffraction efficiency, we should consider the loss in SNR due to crosstalk. As we integrate our signal wavefront while swinging through reference beam angles which are not exactly at the Bragg–condition, we can expect to get more contribution from the undesired neighboring holograms. This will be much more troublesome than the mild loss of diffraction efficiency. We show in Figure 7.2, the measured SNR as a function of the Bragg–matching condition for

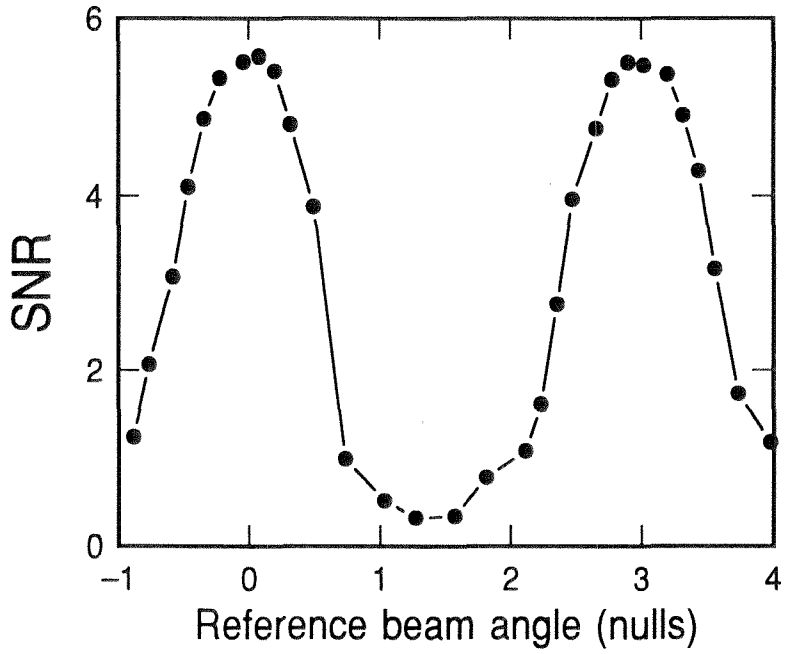


Figure 7.2: Experimentally measured SNR of holograms as a function of Bragg mismatch, for 2 of 100 holograms.

two neighboring holograms stored at approximately the third null. This data was taken from 2 of 100 holograms stored at the center location of the 160,000 hologram system (Chapter 6.4). The dotted line represents the point at which we switched from matching the reconstructed images with Hologram X to Hologram X+1. The effect of an oscillation in Bragg condition would be seen as an averaging over some span of the plotted curve. Note that the drop in SNR would be much more precipitous for holograms stored at the first null.

7.1.3 Photon–limited Readout

If the signal power is not sufficient, then access time is not the limit on the readout rate. Instead, the rate–determining step is the time which the reference beam must remain Bragg–matched to the hologram, so the detector array can integrate the incoming signal photons. We can write the number of photogenerated electrons as

$$N_e = \eta_{tr}\,\eta_q\,\frac{\eta_h\,\eta_{im}\,P_{in}}{h\nu}\,\frac{1}{r_{ON}\,N_p}\,t_{int}, \tag{7.2}$$

Variable	Definition	Value
N_e	Number of signal electrons	1000
η_{tr}	Electron transfer efficiency	0.9
η_q	Quantum efficiency	0.9
η_h	Holographic diffraction efficiency	—
η_{im}	Efficiency of readout optics	0.9
P_{in}	Readout power	100 mW
$h\nu$	Power per photon	4.073×10^{-19} J
$r_{ON}N_p$	Number of ON pixels	0.5×10^6
t_{int}	Integration time	.001 sec

Table 7.1: Variables from Equation 7.2.

where the variables are defined in Table 7.1. We are going to take the approach that we have a target number of photoelectrons (after readout of the detector array) which we want to reach for each ON pixel. We then solve Equation 7.2 for the required holographic diffraction efficiency. Note that η_{tr} is the worst–case transfer efficiency from CCD pixel to external electronics, and η_{im} accounts for any losses in the imaging system between the crystal exit and the detector array.[2] With the numbers from Table 7.1, we can solve for the required holographic diffraction efficiency as

$$\eta_h = 2.8 \times 10^{-6}. \tag{7.3}$$

Since we know that the diffraction efficiency is related to the number of holograms M by the $M/\#$, we can use Equation 7.3 to solve for the minimum required $M/\#$. This approach, shown graphically in Figure 7.3, allows us to relate this target diffraction efficiency to the $M/\#$ results from Chapter 4. However, having a crystal with the minimum $M/\#$ is not going to be sufficient if we have not completely solved the readout volatility problem. If we are planning on using holograms while they are slowly erasing, then the cushion between our $M/\#$ and the minimum shown in Figure 7.3 represents the amount of time between recording cycles. We discuss this in more detail in Section 8.1.

[2]See Section 7.7 concerning real–time monitoring of diffraction efficiency.

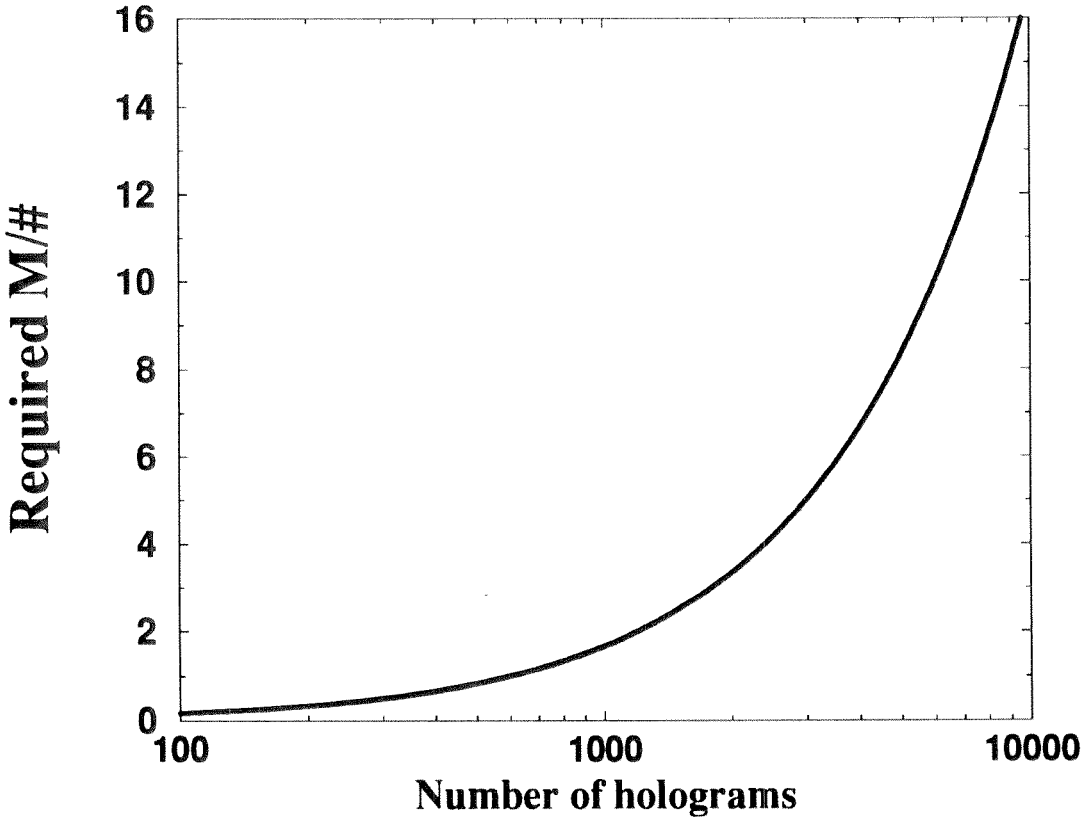


Figure 7.3: Minimum $M/\#$ as a function of M for 1 millisecond integration

What are our options for reducing the minimum required diffraction efficiency? We could increase the readout power or reduce the number of required photoelectrons. With more readout power, we are more susceptible to light-induced signal degradation such as absorptive heating and photovoltaic screening field buildup [246]. If we reduce the number of photoelectrons, then we need to worry about shot noise. We can consider the shot noise from the arrival of the photons or the electrons. In the next section, we consider the general noise statistics of the holographic storage process, and describe some of the noise sources.

7.2 Noise statistics and Error performance

7.2.1 SNR vs. Probability of error

We use the Signal–to–Noise Ratio (SNR) as a measure of the error performance of our experiments, defined as

$$SNR \equiv \frac{\mu_1 - \mu_0}{\sqrt{\sigma_1^2 + \sigma_0^2}}. \tag{7.4}$$

However, we would really like to have the estimated probability of error, P_e, perhaps expressed as a bit–error–rate. What sort of connection do we have between them? Well, as always it depends strongly on the PDFs involved. A typical assumption is a Gaussian dependence, so that the probability of error can be estimated from

$$P_e = \frac{1}{2}\mathrm{erfc}\,\frac{\mu_1 - \theta}{\sqrt{2}\sigma_1} + \frac{1}{2}\mathrm{erfc}\,\frac{\theta - \mu_0}{\sqrt{2}\sigma_0}, \tag{7.5}$$

where the complementary error function is defined as

$$\mathrm{erfc}(z) \equiv \frac{2}{\pi}\int_z^\infty e^{-t^2}\,dt. \tag{7.6}$$

The optimal threshold θ is the point at which the two gaussians meet, or the

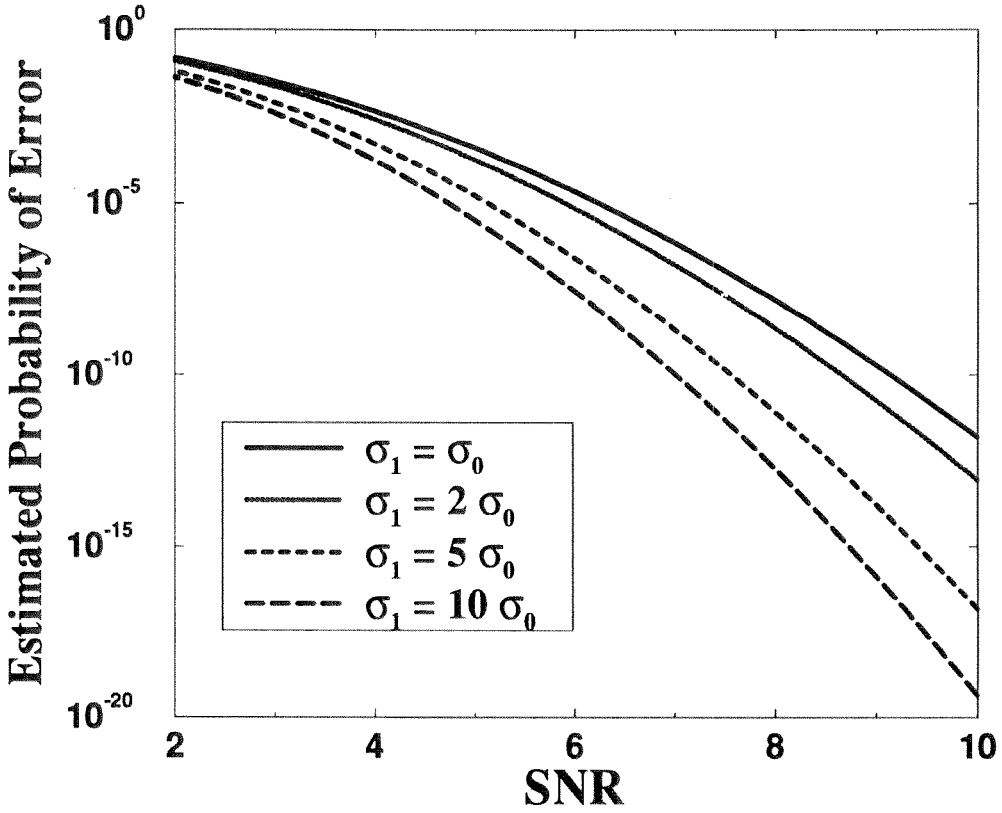


Figure 7.4: Estimated probability of error vs. SNR, for various ratios of σ_1/σ_0.

solution to

$$\frac{1}{2}\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2}\right)\theta^2+\left(\frac{\mu_0}{\sigma_0^2}-\frac{\mu_1}{\sigma_1^2}\right)\theta+\left[\frac{1}{2}\left(\frac{\mu_1}{\sigma_1}\right)^2-\frac{1}{2}\left(\frac{\mu_0}{\sigma_0}\right)^2+\log\frac{\sigma_1}{\sigma_0}\right]=0. \tag{7.7}$$

The correct root of this equation will be the one which falls between μ_1 and μ_0. Once we have these equations, we can see what the relationship between SNR and P_e is for Gaussian distributions. We plot this relationship in Figure 7.4. The plot was generated parametrically, with μ_{ON} as the independent parameter. Note that as one of the standard deviations (σ_{ON} say) becomes larger than the other, the same SNR now corresponds to lower probability of error. Luckily, this is the case for holographic memory, as you can see from looking at Figure 6.7.

7.3 Noise sources

Noise source	Type of noise	Solution
Input Noise		
SLM Noise	Deterministic	precompensate
Input beam nonuniformity	Deterministic	precompensate
Fixed pattern noise	Random	minimize (clean optics)
Storage Noise		
Nonuniform gratings		
(nonuniformity among pixels)	Deterministic	precompensate
(nonuniformity among spatial frequencies)	Random	limit design (# of pixels, holograms)
Fringe motion during storage	Random	minimize (stability control)
Media nonuniformity	Random	minimize (good crystals)
Unwanted gratings	Random	minimize (scattered light, backreflections)
Readout Noise		
Laser fluctuations	Random	minimize (active power control, adaptive thresholding)
Change in Bragg condition	Random	minimize (environment control, angle & wavelength tuning)
Reference beam pointing errors	Random	minimize (monitor diffraction efficiency)
Pixel registration errors	Random	minimize (phase conjugate readout, limit # of pixels)
CCD nonuniformity	Deterministic	precompensate
Detector noise	Random	minimize (cool detector, low–noise design)
Scatter noise	Random	minimize (block edges, AR coatings, iris in readout arm)
Crosstalk noise	Random	minimize (large angular spacing)

Table 7.2: Noise sources in holographic storage and possible solutions.

7.3.1 Input noise

This is noise that is already present on the spatially modulated beam when it reaches the crystal.

SLM Noise This input noise can be introduced by the SLM, in the form of a nonuniform response across the pixels. It is even conceivable that there will be dead pixels in any particular device. However, SLM noise is a deterministic noise source. When a particular SLM is put into a system, we can determine exactly what nonuniformity is introduced and then compensate for it. If the SLM has grey scale, then we decrease the grey level of those pixels which are emphasized by the nonuniform response. This reduces efficiency and requires a fair amount of dynamic range from the SLM. Another possibility for compensating nonuniformity is incremental recording—as the hologram is exposed, we begin to turn off pixels that we know will be strongly recorded (due to the nonuniformity). By controlling the exact exposure for each pixel, we can make the resulting holograms uniform in readout power. In this case, the same nonuniformity is present for each hologram, so the relative timing of the pixel turnoff is the same for each exposure. A third option is to compensate at the detector, as we did in our experiments in Chapter 6.3.3. In this scheme, some of the capacity of the memory is used to store the non–uniform profile of the ON pixel levels. When this hologram is reconstructed, the detector uses the incoming information to set thresholds at each pixel, either individually or after some local smoothing. The disadvantage here is that we have made the detector array more complicated.

In the case of dead pixels within a particular SLM, we can avoid errors by simply avoiding those pixels.

Input beam nonuniformity The spatially modulated information beam may also have nonuniform ON pixels because the light beam illuminating the SLM is nonuniform. There is a tradeoff between the amount of light which is wasted in beam expansion, and the uniformity of the center area which strikes the SLM. If the light is nonuniform, this is again a deterministic noise source which we can precompensate

either at the SLM, or during exposure, or at the detector array.

Fixed pattern noise This refers to the diffraction rings and spots that appear on a image transmitted through an optical system, because of the presence of dust at points in the system. The farther the dust is from an image plane, the more pixels that are affected. Dust in an image plane tends to affect only a few pixels, but often to the point of complete obscuration. The solution here is to use good optics which are kept clean.

7.3.2 Storage noise

This refers to noise which is added to the signal beam while it is recorded. Included here is the failure of the hologram to record the signal beam exactly, as well as the inclusion of additional holograms which will be reconstructed simultaneously.

Storage nonuniformity This noise source is the failure of the recording process to preserve the hologram. In the appendix to Chapter 1, we wrote that the hologram is proportional to SR^*, so that illumination with the same reference R returns the exact input signal S. However, there were several assumptions here:

- Absorption effects— In Chapter 1, we made an assumption that the $\mid R \mid^2$ in the reconstructed signal falls out easily. However, in the presence of absorption, we cannot guarantee that $\mid R \mid^2$ (which will be part of our reconstructed signal) is the same across the signal beam.

- Nonlinearity— A second assumption was that the material records in the same way independent of light intensity. However, most materials are linear only within limited range of light intensities. If the storage material is placed at or near the Fourier transform plane (without a random phase plate), then the intensities can be out of this range and the fidelity of the hologram will suffer.

- Media nonuniformity— The storage media may not have the same response throughout the volume because of media defects or gradients in doping density

or oxidation state. In addition, across a signal beam containing many spatial frequencies, the photorefractive response will vary as the period of the stored gratings changes. The same effect will occur from hologram to hologram in angle multiplexing, as the reference beam changes angle.

To what extent can we eliminate these noise sources? Well, the first is to avoid storing at or near a focused Fourier transform plane. This is really not so bad, because in Chapter 4.3.3 we saw the dynamic range is lower for this configuration anyway. The second answer is to get crystals which are defect–free and uniform. Any nonuniformity across pixels which shows up in every hologram, or shows up as a deterministic function of reference beam,[3] can be precompensated just as we did above. This would include effects from absorption as well as photorefractive response.

The other effects are things that affect the spatial frequency impulse response of the grating, and in turn the point spread function of the images at the detector array.[4] That is, we stored a plane wave which was originally uniform across some aperture. Since this appeared as a top–hat function at the F.T. plane, the point–spread function was a sinc. Now that we have recorded the hologram, some non-uniform profile has been imposed (multiplied) on the grating. At the image plane, the point–spread function is now convolved with the Fourier transform of your nonuniformity and is wider. Now pixels begin to "bleed" into each other more than they did from just the diffraction through the original aperture. The solution to this is in design: the pixels must be far enough apart that, after coherent summation, the ON pixels of reconstructed holograms do not significantly spill over into any neighboring pixels. This may set a design limit on the fill factor and/or spacing of detector/SLM pixels. Note that the optimal configuration may not have the same pixel layout in both horizontal and vertical dimensions

Fringe motion during storage Holograms are stored interference patterns. Since available storage materials require long exposures, the interference pattern must be

[3]Meaning that every time we record with a reference beam at some known exterior angle, we get such-and-such a nonuniformity.

[4]Assuming Fourier plane storage.

held stable during the entire exposure. Any change in the period or phase of the interference pattern during an exposure will cause holograms to wash out. This loss of dynamic range brings the diffraction efficiency of holograms closer to the noise level, increasing error.

Interference patterns can change from laser drift, mechanical vibration of components, air currents in the optical setup. Just about anything which changes the optical path difference between the reference and object beam is bad news. The solution is to enclose the setup away from environmental exposure, both in terms of air currents and mechanical vibration. The first is generally straightforward (and desirable from the point of view of clean optics); the second is implemented easily in the lab with large optical tables sitting on pneumatic legs. Obtaining the same performance in a small package will be an engineering challenge.

An alternate solution is to use active feedback on the pathlengths with either an AOD or piezoelectric crystal [358,359]. The pathlength is varied sinusoidally (a very small amount) and the electrical control signal at the same frequency filtered out of a detected optical signal. This requires the ability to interfere the reference and signal in such a way that the fringes of the interference pattern are wider than the aperture of the photodetector. Practically, this implies that both reference and signal are plane waves when they intersect, and is tricky for Fourier transform plane storage since the signal beam is not a plane wave. One way it might be done is shown in Figure 7.5. Here we divert part of the signal beam that was going to be wasted anyway, and direct it as a sort of skew ray through the system. The test beam focuses at the center of the SLM. A distance of 2F away at the the Fourier transform plane of the SLM, the test beam is a plane wave located at some distance from the intersection of signal and reference. If a second test beam is passed through the reference optics to meet this beam, then an interference pattern can be formed. Here the test beams have encountered all of the optical elements common to the recording setup, with the addition of the mirror and lens in the object beam, and the exclusion of the crystal mount itself.

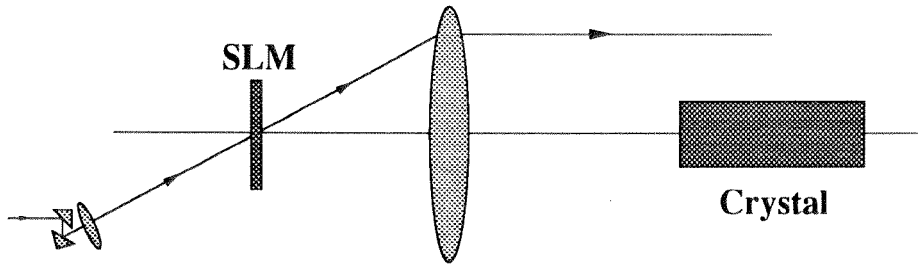


Figure 7.5: Generation of a plane wave at the Fourier transform plane for use in active fringe control.

Unwanted gratings This refers to the third contribution to the storage noise: unwanted gratings. The source for these gratings can be scattered light (from fixed pattern noise, fanning, or backreflections). Fanning in particular refers to the amplification of scattered light gratings through continued two–beam coupling between the strong reference and the initial scattered light. Empirically, the 90° geometry is less affected by fanning noise than the transmission geometry.

Unwanted gratings can also be written between the various plane wave components of the signal beam. This is one way to think about the distortion which takes place when the crystal is illuminated with a high power signal beam. We show the effects of continuous illumination of the crystal with a >3 mW/cm^2 signal beam in Figure 7.6. This beam is only a little more powerful than the signal beams we typically use for storage experiments. In only a few minutes the image fidelity has been seriously degraded. If we change the signal beam every few seconds by modulating it with new images, the effect builds up much more slowly. Any pattern noise which is common to all or many of the images will tend to be affected more rapidly, which is why we often see the Fresnel rings from dust particles becoming more prominent in later holograms.

7.3.3 Readout noise

OK, now the holograms are stored and it's time to get them out again. What could go wrong? Well, there's:

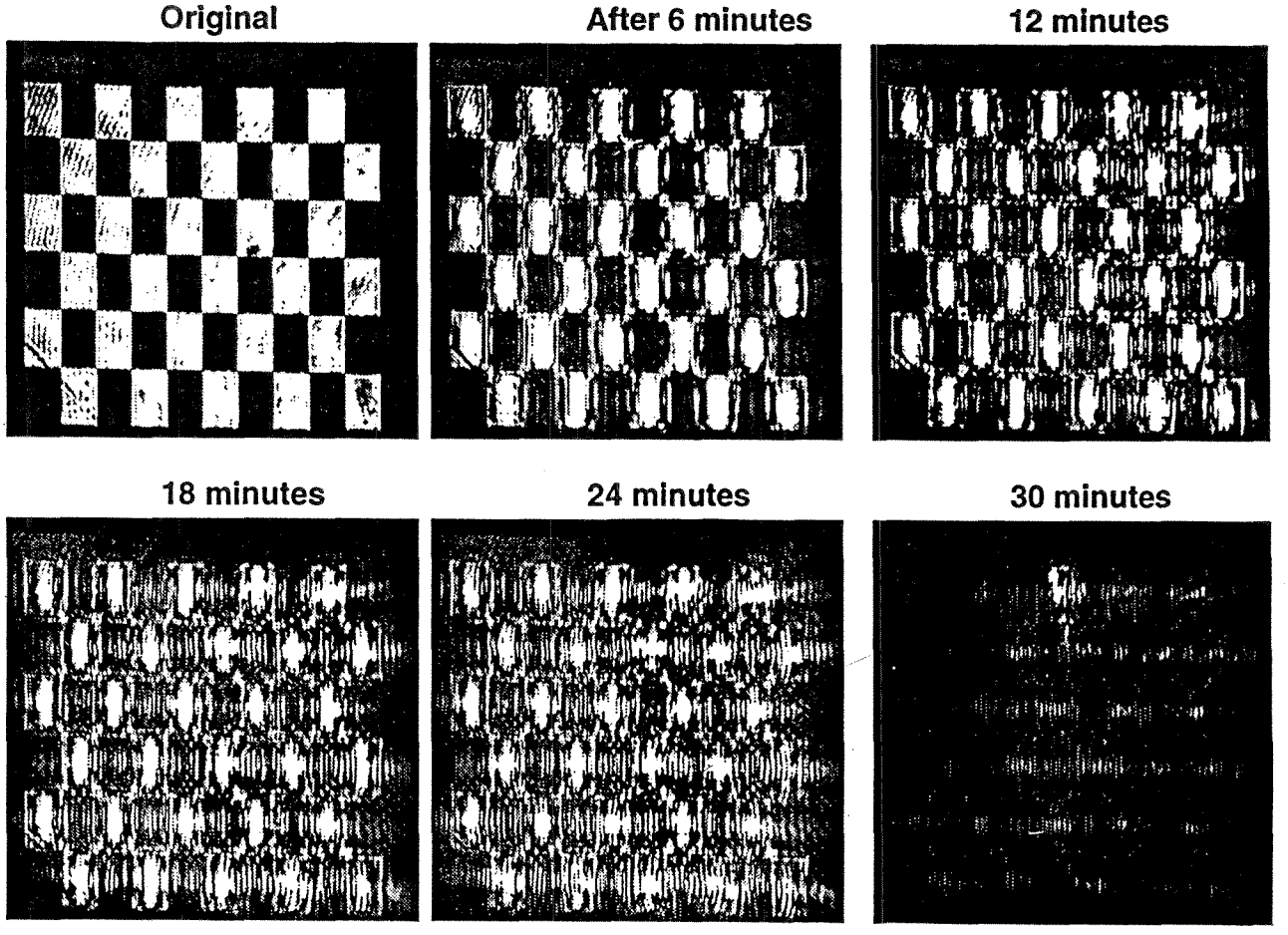


Figure 7.6: Degradation of signal beam by recording of strong interpixel gratings.

Signal power fluctuations You have to decide on a threshold to use at the CCD pixels, to say that "If I get more than x photoelectrons in this bin, I call this an ON pixel." This threshold depends on the signal power in the reconstruction (as well as the sources of nonuniformity which we discussed already). If the total laser power is fluctuating, this may affect the effectiveness of the threshold and through this the bit–error rate. Akin to this are the previously mentioned effects which change the diffraction efficiencies between holograms.

The solution to this sort of problem can include either careful control over the diffraction efficiencies and the laser power, or some sort of manipulation of the thresholding operation. This latter can be the choice of an adaptive threshold which is pixel–based, or of an encoding algorithm which can deal with variations in signal power fluctuation (such as the differential encoding used in Reference [265]).

Reference beam pointing errors This refers to errors in returning the wavefront controller to the same wavefront that was used to store the holograms. This can have effects on the signal strength and on the pixel registration. If the reference beam, for instance, is mismatched in horizontal incidence angle upon readout, then the diffraction efficiencies of the holograms will suffer. If the vertical incidence angle is mismatched, we would seem to be okay, because the hologram is still Bragg–matched and we are getting output power. However, the change in vertical incidence angle is going to show up on the output signal wavefront, which (in the Fourier transform geometry) will cause the image to shift on the detector array.[5]

One solution to the Bragg–matching problem is to use feedback on the power of the reconstructed signal and tweak the horizontal reference beam angle until the output power is maximized. However, since the output power is not a strong function of vertical incidence angle, an equivalent technique might not work for errors in this dimension. In this case, the quality of pixel matching would need to be tested by having a known pattern on the image, such as a bright border around the image. In

[5] As the horizontal incidence angle changes, the image is shifting on the detector as well as losing power due to Bragg–mismatch. If the angular selectivity is poor, then the loss of registration may become more important than the loss of signal power.

this case, the readout speed and the usable number of pixels will certainly suffer.

Pixel registration errors We continue our discussion of pixel registration and include other sources. Rotation of the crystal away from the position in which it was recorded will cause the image of pixels to move relative to the grid of detector pixels. An equivalent effect can occur if the signal passes through a phase profile (perhaps a linear phase ramp) that was not present when the object beam was lined up.

This is a serious problem which doesn't have a quick and easy solution. To a certain extent, one can minimize the long–term aspects by putting some positional control on the CCD camera. Periodically, an automatic procedure aligns a known hologram with the camera. This will take care of slow gradual drifts which are bound to occur.

If the alignment drift is not gradual, we may have to limit the number of pixels in our design, throwing out some camera pixels as buffers between our bit regions, and averaging others to deal with the fact of life that we do not have arbitrary control over the size of the camera pixels. A solution to the problem of pixel registration, initially and during prolonged hologram readout, will be a crucial step towards the practical realization of holographic memories.

CCD nonuniformity Nonuniformity in the CCD pixels has much the same effect as nonuniformity of the SLM pixels: if it's mild, we can precompensate. If the CCD pixel is badly nonuniform, we can decide not to use it at all.

Detector noise Here we include electrical noise generated by the detector array, as well as the effects of shot noise on the input photons. Since the number of photons per pixel per integration period will probably be small, then the magnitude of the shot noise may be a consideration. In addition, the noise of the camera will start to become a consideration as the readout rates begin to increase and the electrical bandwidth becomes larger.

Scatter noise Scattered light can be a serious consideration in holographic memories because the diffraction efficiency of the signals are so low. We can minimize this by using a large angle between reference and signal. One of the best features of the 90° geometry is its incredibly low scatter—one can get the scatter of the reference beam to below effective diffraction efficiencies of 10^{-10}. Highly polished optical surfaces with good anti–reflection coatings are important, as is the masking of crystal corners and edges from illumination by the strong reference beam. The best way to protect the camera from scattered light is to make the reconstructed signal pass through an iris on the way to the camera. This passes only the scattered light which has spatial frequencies near the signal beam and reduces the scatter level dramatically. In a system with storage at multiple locations, we can either use a series of slits which are always open, or use fast liquid crystal phase retarders and a polarizer to block the light from the slits not in use.

Crosstalk noise Crosstalk noise is a subtopic of scattered light in a sense, since it refers to the residual power diffracted (or scattered) from all of the holograms except the one that is Bragg–matched. As we discussed in Chapter 1, Bragg mismatch of these other volume holograms tends to attenuate these other holograms, since they cannot build up in phase. Since the crystal is finite in space, the diffraction efficiency response as a function of Bragg angle has nulls in it. If the grating response is uniform through the volume, then these nulls really go to zero. However, the exact position of the null (that is, the angle to which you detune the reference to make the hologram go away) is a function of the signal spatial frequency. If each hologram has many spatial frequencies (like any information–bearing image), then each of these requires a slightly different Bragg mismatch angle. The overall result, in theory, is that some part of the hologram is almost always visible. All of these differences occur in the horizontal plane, though, so that we should expect to see entire columns of pixels disappear together.

Theoretical treatments of crosstalk has been a topic of intense research recently [103, 134–142, 144, 145, 147–149, 360–366]. In some papers, crosstalk was being used

to compare multiplexing methods. The general trend is that crosstalk noise depends on spatial bandwidths. In the case of angle multiplexing, the theoretical worst–case crosstalk noise is the product of the reference bandwidth and the signal bandwidth. [137, 139, 228, 280] For wavelength multiplexing, it is the square of the signal bandwidth. [139, 228, 280] So the more pixels one tries to pack in (horizontally), the more crosstalk noise. In addition, in angle multiplexing, the more holograms you try to add, the more crosstalk noise.

So much for theory—what really happens in the lab? Well, empirically it is rare to see Bragg mismatched images with the theoretically predicted horizontal sinc modulation across reconstructed images. Instead, the entire hologram tends to fade out together (see Figure 6.17). This may be due to signal cone angles which are not wide enough for the effects to show, or to intensity profiles on the reference beam which ruin the top–hat spatial profile of the diffracted intensity. Also empirically, we have never noticed a strong trend in crosstalk noise as a function of reference beam angle.[6] In general, we usually consider crosstalk to be not a serious problem (except for the micromirror experiments where the curvature of the device tends to introduce many spatial frequencies into the reference beam). Most crosstalk effects in practice appear to come from the first one or two neighboring holograms on both sides.

We used our 160,000 hologram system to study crosstalk noise as a function of the number of holograms stored. During storage of 100, 200, 500, 1000, 2000, 5000, and 10,000 holograms, the 50^{th} location was skipped and left vacant. The power appearing at the detector when the reference beam was returned to this angle was measured, and the histogram of data values computed for the same detector area that would be occupied by the reconstructed SLM images. The camera returns a 12–bit value for each pixel—a 1 second exposure was used for all measurements. Since this is a representation of the dark or zero level of the output hologram, the desired histogram is a very sharp peak located at very low pixel values. As can be seen in Figure 7.7, the

[6]This is where the expected increase in crosstalk noise as a function of the number of holograms comes from: the worst case reference beam is always supposed to be the one farthest from normal incidence, and it is assumed that the more holograms you try to store, the larger this worst–case angle.

crosstalk noise approaches these desired characteristics as more and more holograms are stored. By the time 10,000 holograms are stored, the crosstalk noise is very near the original background scatter measured before any holograms are stored. This trend is backed by recent theoretical crosstalk studies, [137, 139, 228, 280] which show that crosstalk noise decays along with the signal strength as more holograms are stored. If we were to store varying numbers of holograms at the same angle spacings, but have the same worst–case angle for each, we should find that the SNR due to crosstalk would be independent of the signal strength.

In our experiment, we used larger angle ranges (noisier worst–case angles) when storing more holograms. So the strength of each stored hologram decays as $1/M^2$ while the normalized worst–case crosstalk noise grows as M. We find that, because of dynamic range considerations, crosstalk noise becomes less worrisome for weak holograms because the undesired holograms are weaker too. Instead, other noise sources such as scatter or detector noise will be the limitation. In practice, the desired solution might be to increase the strength of all the holograms to the point where crosstalk noise dominates all the other noise sources—then we can be sure that we ae getting the best SNR.

7.4 Non–volatile readout of holograms

Our goal is to design a holographic storage system with existing read/write materials. We have chosen to work with photorefractives, and in particular, Fe–doped $LiNbO_3$. As we have mentioned, holograms in this material will erase as they are readout. Therefore, we need to design system solutions which either slow down this erasure, or which recopy our stored data before it is erased. In this section, we survey the methods which have been proposed and demonstrated for non–volatile readout.

7.4.1 Readout with a second wavelength

Since the photorefractive effect occurs through absorption of photons, we might try using a photosensitive wavelength for recording and a different one for readout. If

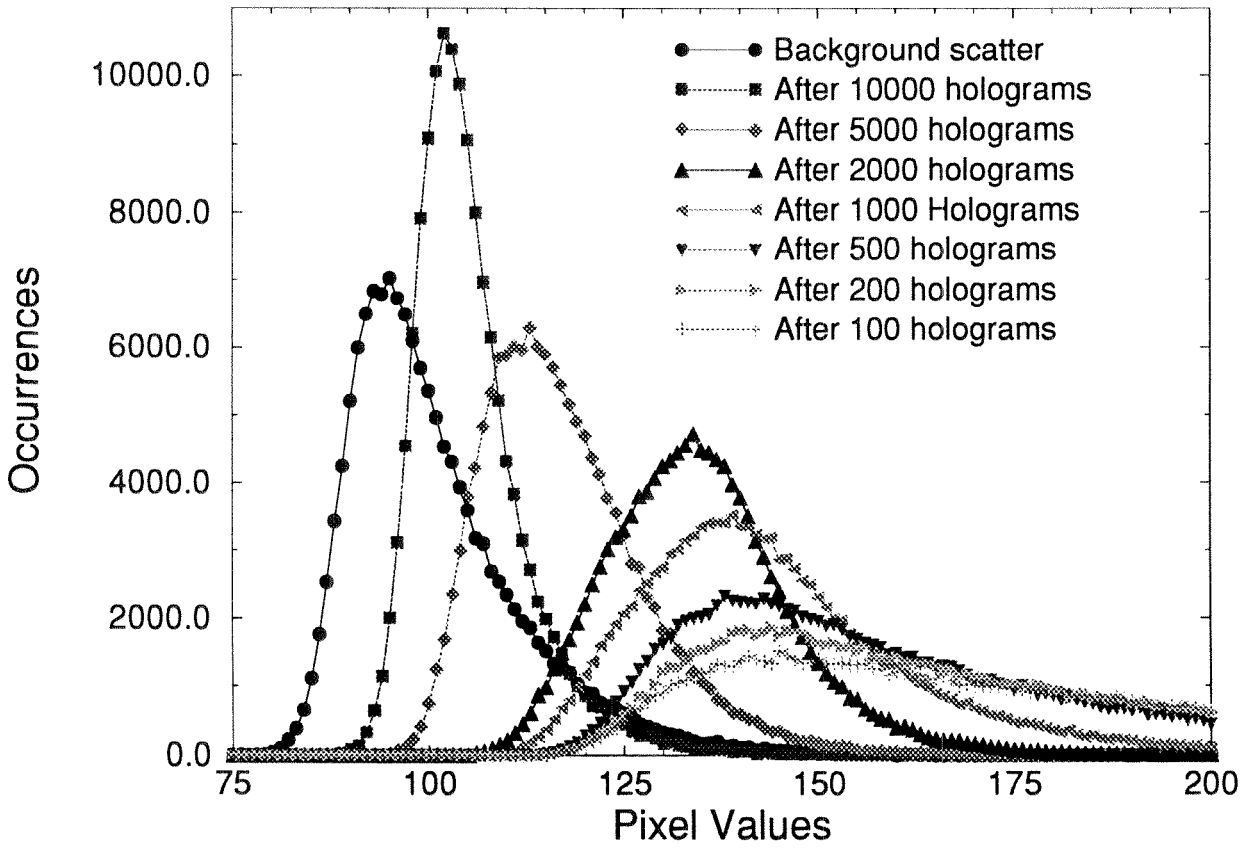


Figure 7.7: Background scatter and crosstalk noise histograms

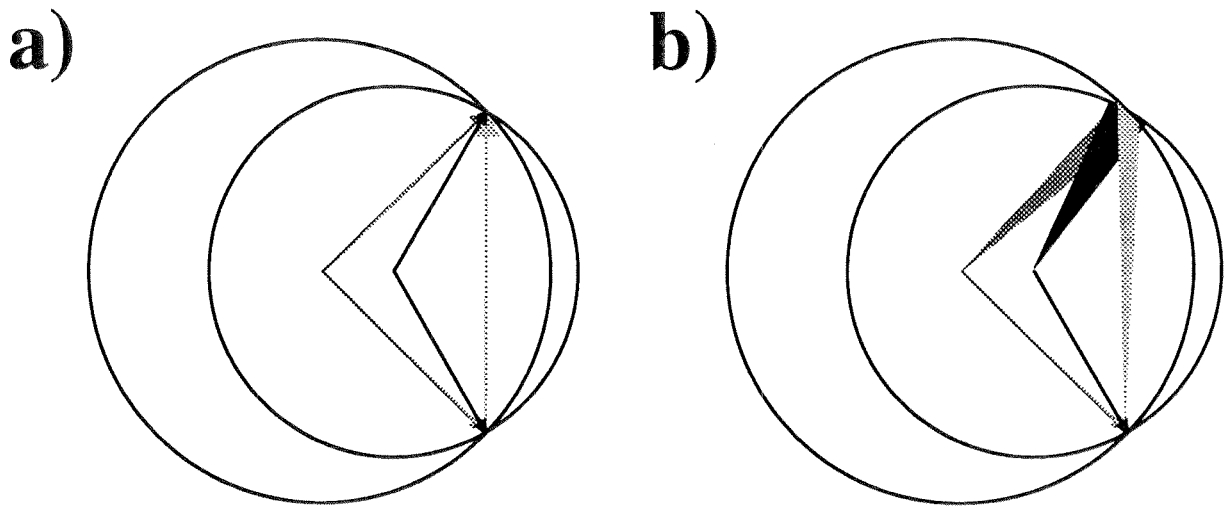


Figure 7.8: K–space diagrams for readout at a second wavelength.

the absorption at this second wavelength is lower, then the gratings will erase at a commensurately slower wavelength. However, it still remains to Bragg–match the grating with the second wavelength. We show in Figure 7.8(a) the K–space diagram for this operation. We can Bragg–match one grating with different wavelengths, but both the readout and reconstruction angle will change.

However, we normally store a signal wavefront which contains information and thus exists in some band of spatial frequencies. This band is represented in Figure 7.8(b) by an arc on the surface of the K–circle. We see that we cannot simultaneously Bragg–match all of the spatial frequencies in the stored hologram with the second wavelength. If the holograms are stored in the Fourier plane, this Bragg–mismatch washes out the (left and right) sides of the image. In the image plane, the image becomes low–pass filtered and the horizontal edges of the pixels start to smear. The effect can be minimized by reducing the wavelength difference, or by decreasing the angular selectivity.[7]

Despite the need for poor angular selectivity, 1000 holograms have been stored at 488nm and reconstructed with 633nm light by using a combination of angular and peristrophic multiplexing [233, 234]. However, the improvement in erasure time by

[7]This implies smaller interaction angles or smaller interaction length.

moving from 488nm and 633nm is limited, because there is still significant absorption at 633nm. During our oxidation/reduction experiment in Chapter 3, the ratio of the absorption coefficient at 633nm to the absorption coefficient at 488nm was approximately 0.2 for all oxidation states. This 2–wavelength readout technique would become more attractive if a dopant can be found with a step–function–like spectral response.

The 90° geometry and two–lambda readout are a poor match because of the high angular selectivity in our geometry. It seems that we would need to limit the horizontal signal bandwidth by a considerable amount in order to avoid signal degradation. If we wanted to build a rapid–access memory using two–lambda readout, we would need to use transmission geometry. This could be done with techniques similar to our scanning point source. However, in order to keep the scope of our discussion from getting any larger than it already is, we will not consider two–lambda readout in this thesis.

7.4.2 Fixing

The second non–volatile readout technique we discuss is fixing. Since we have chosen to work with $LiNbO_3$:Fe, we concentrate on thermal fixing. However, we will also say a few words about electrical fixing in SBN [124, 125, 163–166] and $BaTiO_3$ [167]. In thermal fixing, we electrically compensate the trapped electrons of a written grating with mobile H^+ ions [367]. The ions are mobilized by elevating the crystal temperature. If the temperature and the heating time is chosen judiciously,[8] then the ionic compensation occurs more rapidly than the thermal erasure of the trapped electrons [120–122, 168, 367–374]. Immediately after fixing, there is no overall space–charge field and thus no hologram. However, once we illuminate the crystal and randomize the stored electrons, only the ionic grating remains. At room temperature, the lifetime of the fixed hologram ranges from months to 5 years [123]. Note that the decay of the hologram is now independent of illumination. There is a tradeoff

[8]Temperatures in the range of 100–150°, and times on the order of 5–30 minutes have been reported.

between the lifetime and the fixing efficiency, which seems to depend on the number of mobile ions [168].

In electrical fixing, the trapped electron hologram is compensated by partial reversal of the local ferroelectric domains. The domain reversal is induced by a negative high-voltage pulse along the **c**–axis. However, as with thermal fixing, the grating cannot be observed because the overall space–charge field is zero. A positive voltage pulse "reveals" the grating, allowing it to be observed and erased. Unlike thermal fixing, the gratings cannot be fixed and then read out without worry. The set of holograms stored in a single location must be revealed, read out, and then re–fixed. The reason for this is that the compensation between the polarization grating and the charge grating is nearly complete.

7.4.3 Copying

Optical

In copying, holograms are sequentially reconstructed, refreshed, and stored anew [375–387]. With an optical copying loop, the holograms are refreshed by reconstructing them onto a gain element such as a liquid–crystal light valve. Alignment of the reconstructed wavefront back into the original signal beam is extremely critical. Note that the holograms are being rewritten back into the same storage location. Care must be taken to start the copying loop when the holograms are strong—the last hologram to be copied will be erasing during the entire copying procedure, and must still be above the threshold for correct detection at the end of the copying procedure. This tends to limit the number of holograms which can be stored in each location.

Electronic

In contrast, electronic copying is rather simple: holograms are reconstructed onto the array detector, the detected bit pattern is routed to the SLM and displayed, and the data stored in a separate location. The reconstruction process takes no longer than when holograms are accessed for output to the user. Since the new holograms take

much longer than this to record, readouts for copying purposes can be interspersed with normal operation of the memory. In addition, after all the holograms are copied from a crystal, the crystal can be thermally erased to remove any residual photovoltaic fields. We discuss electronic copying further in Section 8.1.

7.4.4 Two–photon recording

One materials solution to the volatility problem has arisen from research in the storage of holograms using red or infrared light. The original impetus was to move the photosensitivity of the photorefractive crystals towards the range of available laser diodes by using appropriate dopants. However, researchers have demonstrated storage of interference patterns formed between IR beams in Fe–doped $LiNbO_3$, but only when gated by high–intensity green pulses [184–190]. The IR writing beams can either be pulsed or DC. This development could become an attractive non–destructive readout technique, since the IR read beam is not strongly absorbed and can Bragg–match all the spatial frequencies of the grating. The main drawback is that the intense green pulse erases previously written holograms and reduces $M/\#$. The exact effect on $M/\#$ and thus the number of stored holograms per location is still an open question.

7.5 Beam deflection devices

In this section, we survey the available angle deflection devices for a large–scale high–speed holographic memory. Although phase–code and wavelength multiplexing are also possible alternatives for such a system, we limit the scope of our discussion to angle–multiplexed storage. However, many of the issues we discuss concerning the signal beam would apply to a phase–code or wavelength multiplexed holographic storage system.

7.5.1 Acousto–optic deflectors

Acousto–optic deflectors work by Bragg diffraction from a moving phase grating. The grating is introduced to the acoustic medium by driving a transducer at some RF frequency. The strain disturbance travels through the material with a given acoustic velocity, modulating the index of refraction through a strain–optic tensor (much like the electro–optic tensor). The period of the phase grating Λ is related to the acoustic velocity v and driving frequency f by

$$\Lambda = \frac{v}{f}. \tag{7.8}$$

As a result, deflection angle is a function of the RF frequency around the center frequency f_0, as

$$\Delta\theta = \frac{\lambda}{nv\cos\theta_B}\,\Delta f, \tag{7.9}$$

where λ is the wavelength of light, n the index of refraction, and θ_B is the angle at which light is incident (for Bragg–matched readout). Because the periodicity of the grating is small and the material thickness not too large, an AOD usually operates in the Raman–Nath regime and multiple orders can be observed. For this reason, the device can only be used for RF frequencies from $\frac{2}{3}f_0$ to $\frac{4}{3}f_0$.[9]

The transducer is usually reasonably small (in transverse area) so the acoustic phase grating has some spread in spatial frequency. This allows the device to diffract the same input beam into various output angles with changes in the grating period. If the acoustic phase grating had a small spread of spatial frequencies, the device would quickly become Bragg–mismatched as the RF frequency as changed. To get good efficiency, we would then need to change the input angle as well, making the device essentially useless. The residual effects of this Bragg–mismatching causes AODs to have some spatial frequency response, with high efficiency at the center frequency and lower efficiency at the edges of the usable range.

[9] Consider the overlap between second order at 2/3 f_0 and the first order at 4/3 f_0.

Once we know this frequency response, we can record uniform holograms. For instance, we can increase the exposure time for less efficient deflection angles, so that the diffracted signal power of each hologram is equal. In terms of our work in Chapter 2, we would make our recording schedule as

$$t_m = \left[t_M \frac{\left(\frac{A_0}{\tau_r}\right)_M}{\left(\frac{A_0}{\tau_r}\right)_m} \sqrt{\frac{P_{in}|_M}{P_{in}|_m}} \right] l_m \left(\sum_{i=m+1}^{M} t_i \right). \tag{7.10}$$

We can obtain the ratio $P_{in}|_M / P_{in}|_m$ from the spatial frequency response of the AOD. Note that this approach will work for any other angle deflector whose efficiency depends on deflection angle.

Because the phase grating is moving, the diffracted beam is Doppler shifted by value of the RF driving frequency. On the other hand, the moving nature of the grating allows quick update of the phase grating—the access time is the time needed to get the new RF frequency to the transducer and across the illuminated aperture of the AOD. This is usually on the order of 10-100μsec. The size of the illuminated aperture also affects the SBP of the AOD through the size of the resolvable spot. The SBP is the number of these resolvable spots which fit inside the angular tuning range of the device as set by Equation 7.9, or [388]:

$$SBP \equiv \frac{\Delta\theta}{\delta\theta} = \frac{\pi\,\omega_0}{2v\,\cos\theta_B} = \tau\,\Delta f. \tag{7.11}$$

A typical SBP value for a commercial AOD is 1000–1100.

The main advantage of AODs is the rapid access. Disadvantages include the small deflection range, low efficiency, nonuniformity, power consumption, and the need for a 4–F system for each AOD. We should note that despite the need for a 4–F system or imaging lens with an AOD, a compact system can still be designed. A team at Rockwell has recently built a system which occupies 520cm^2 and contains 3 AODs: 2 in the reference arm and one in the object arm [276].

7.5.2 Rotational actuators

We can use a mechanical actuator to rotate a mirror to deflect the reference beam. With this approach, we require a 4–F system to image the angle change onto the surface of the crystal. We could put the angle deflector close to the crystal, but then we will be getting some combination of angle and spatial multiplexing. Because the stage is driven by a motor with some given resolution, there is always a tradeoff between the resolution and the maximum rotation speed. We can get around this by using a small rotation stage with fine resolution and limited angle range on top of a larger rotation stage with coarse resolution, high speed, and large deflection range. However, since the beam to be deflected is 1–2 centimeters wide, the mirror on the rotation stage must also be large. This adds weight to the stage and limits deflection speed. For these reasons, a rotational actuator does not seem to be the best option for high–speed angle deflection.

7.5.3 Linear actuators

Most of the experimental demonstrations that we showed in Chapter 6 were done by using a linear actuator to perform angle multiplexing. However, our linear actuator moves the weight of a lens and two periscope mirrors—too much to think about rapid motion. However, all we really need to do is move a focussed spot or point source. We could translate a small laser diode or the end of a fiber, adding very little weight to the stage. Another advantage of a linear actuator is that we only need one lens. In contrast, the rotational actuator requires three: one for collimation, and two to make up the 4–F system. We can use the same trick with the linear actuator that we used on the rotational actuator: a fine–resolution stage mounted on a fast, coarse motion stage. A possible choice for the fine–resolution stage might be a piezoelectric device, capable of 100 μm translations with submicron accuracy [389]. Once we are close to the desired reference beam, we could use the piezoelectric transducer to find the Bragg condition rapidly using feedback on the power of the reconstructed hologram.

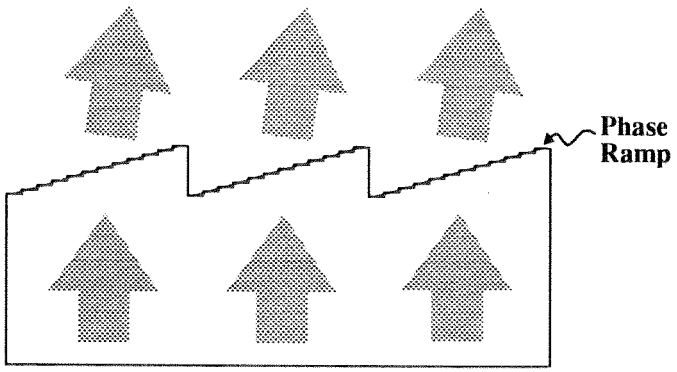


Figure 7.9: Operation of a liquid–crystal beam steerer.

7.5.4 Liquid crystal beam steerers

A liquid crystal beam–steerer is a programmable linear phase ramp (Figure 7.9). As with other LC devices, the liquid crystal is sandwiched between two substrates. One of the substrates is covered with closely spaced parallel electrodes. The boundary conditions are configured so that the voltage on the electrode can control the index of refraction, and thus the phase of the transmitted light. The voltages increase from one electrode to the next, creating a phase ramp. When the phase reaches 2π, the phase wraps around to 0, forming a blazed grating.

The finite size of the electrodes causes the linear phase ramp of a blazed grating to be approximated by discrete steps (see Figure 7.9). The effect of these steps on the output spatial frequencies are shown in Figure 7.10.[10] The Fourier transform contains a comb function which can be shifted back and forth relative to the sinc envelope. This shifting of the Fourier transform is the angle deflection created by the device, and can be controlled by the number of electrodes involved in one 2π cycle of the phase ramp. There is no theoretical limit on the lowest non–zero deflection, since we can use the entire aperture of the device to implement a small fraction of one 2π cycle. In practice, the number of bits of dynamic range in the D/A converter used to set the electrode voltages limits the spatial frequencies on the low end.

The highest spatial frequency calls for every other electrode to have π phase shift. However, this is a square wave phase grating, which has two equal first diffraction or-

[10]This analysis is similar to what we did for the SLM in Chapter 1.6.6.

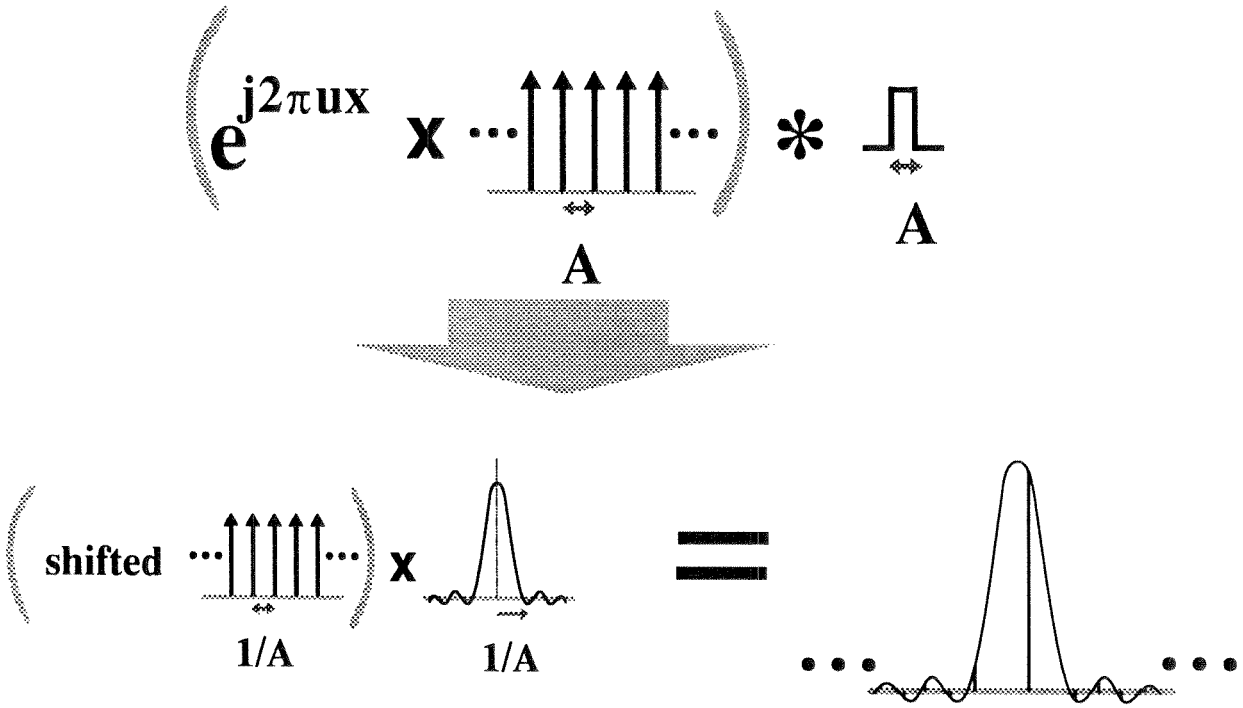


Figure 7.10: A liquid–crystal beam steerer and its Fourier transform.

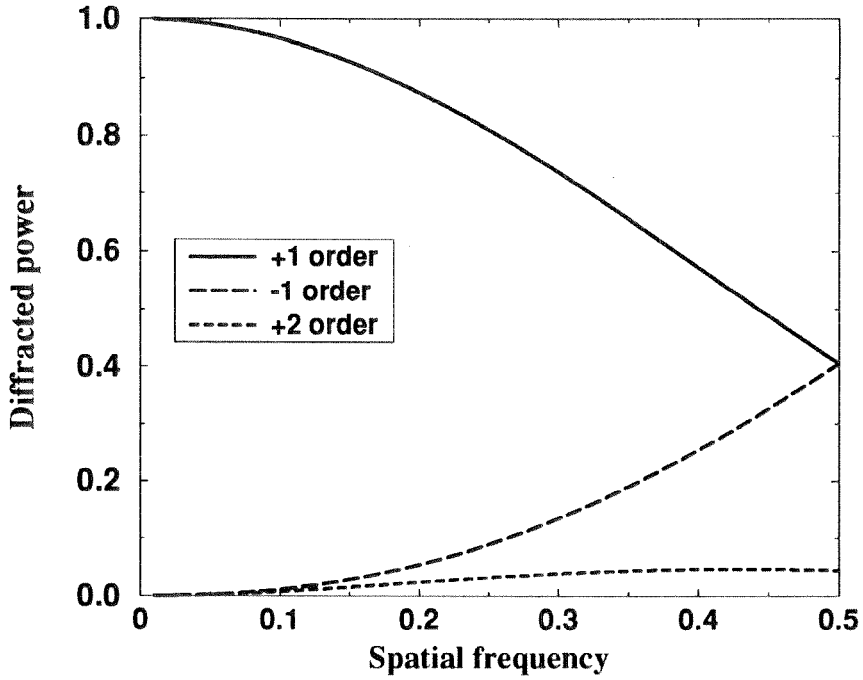


Figure 7.11: Spatial frequency response of a liquid–crystal beam steerer.

ders. As we use fewer and fewer electrodes per ramp cycle, we find that the amplitude of both the -1 and +2 order increases. You can envision this in Figure 7.10 as the comb function slides behind the sinc function. One spike in the comb function sits in the main lobe of the sinc. As this central spike moves away from the center of the sinc function (corresponding to larger deflection angles), its height decreases (signifying reduced efficiency). At the same time, one of its neighbor (the -1 order) enters the central lobe and its other neighbor (the +2 order) enters the first sidelobe. Both of these imply higher crosstalk. In Figure 7.11, we show the power in the desired +1 order and these two largest undesired orders as a function of the spatial frequency. The right–hand edge of the plot corresponds to the square wave phase grating, where the +1 and -1 orders are equal. If we were to direct the deflected beam through a 4–F system, we can filter out most of the undesired light with an opaque block in the center focal plane. However, we would really like to place the liquid crystal beamsteerer against the surface of the crystal. In this case, we had better stay away from deflections which call for fewer than 5–10 electrodes per 2π phase cycle.

There are two additional considerations which we should discuss before going on.

The first is undiffracted light, which is caused by incomplete coverage of the phase range between 0 to 2π. We could intentionally leave all of the light in this zero order by setting all of the electrodes to the same phase. However, we should not store a hologram at this (zero) spatial frequency because it will always have light reading it out. This will be a consideration if we store holograms with a liquid crystal beamsteerer. The second consideration is that there will be some smoothing of the phase ramp by fringing fields in the gaps between the electrodes. This is both a good and bad thing: the smoothing helps suppress the sidelobes of the sinc from Figure 7.10 and increase diffraction efficiency. Unfortunately, the same smoothing occurs at the discrete jump at the boundary of each phase ramp cycle. This creates a short and steep phase ramp which diffracts light away from our desired order.

At this point, we have a device which is capable of very fine resolution for low deflection angles, but poor resolution at high deflection angles. For instance, let's assume that our device is 1cm wide and has electrodes spaced by 1μm. If the interaction length of the crystal is also 1cm, the angular selectivity is roughly 5×10^{-5} radians. The angle difference between a phase ramp cycle of 10 electrodes and one of 11 electrodes is 4.5×10^{-3} radians, or 91 hologram zeroes. This not only gives us no room to tune to the Bragg angle as environmental conditions change, it wastes a lot of angles which could be used to store additional holograms.

One solution is to place two beam–steerers in contact: one to deflect the beam by a large angle, and the next to provide fine steering about this carrier frequency. We might be tempted to make this second device with wider electrodes, since we don't need it to have large deflection angles. However, if we were to do this, we would be creating an additional headache we don't need. The higher orders of the steerer with coarse electrodes will be well within the tuning range of the fine device. We could easily avoid this by making the second device with the same electrode spacing as the first—even if we don't use the largest angles, we are placing the higher orders outside of the combined tuning range.

We still have the same problem that we have a discrete set of angles. Our ability to tune to the Bragg–matching condition may be insufficient at large deflection angles.

We solve this problem in our proposed system by using a separate method to get continuous angle tuning over a small angle range. We will use the the liquid crystal beam steerer to get close, and our continuous tuner to zero in on the exact Bragg–match angle. This simplifies considerably the requirements on our liquid crystal beam steerer. There will still be a tradeoff, however, between the angle range of our continuous tuner and the largest deflection angle of the beam–steerer which we can use.

7.6 MicroMirrors

In this section, we describe compact angle deflectors which has been developed by the Caltech micromachining group of Dr. Yu–Chong Tai and his student Raanan Miller.

7.6.1 Passive MicroMirror

The first device is a silicon micro–flap, consisting of a 4mm $\times$ 4mm, 40 μm thick plate connected to the wafer through two "S"–shaped in–line springs (Figure 7.12). It is fabricated through the combination of bulk micromachining[11] and magnetic thin–film processing. The bulk micromachining provides rigid structures and a flat finished single–crystal surface. Through magnetic thin–film processing, a thin permalloy layer is electroplated onto the micro–flap. The permalloy layer causes the flap to try to align itself with an external magnetic field. The restoring force of the spring opposes this motion, giving a one–to–one relation between the deflection angle and the magnitude of the magnetic field. A typical plot of deflection angle as a function of magnetic field is shown in Figure 7.13, and several fabricated mirror strips of varying spring constants are shown in operation in Figure 7.14. A permalloy mesh structure on the top surface of the plate overhangs the edge and constrains the flap to deflections above the surface of the wafer. The entire flap is coated with a thin layer aluminum

[11]Bulk micromachining is the fabrication of structures by removing material from a single crystal silicon wafer. This is in contrast to surface micromachining, in which structures are formed by adding polycrystalline silicon layers to the wafer.

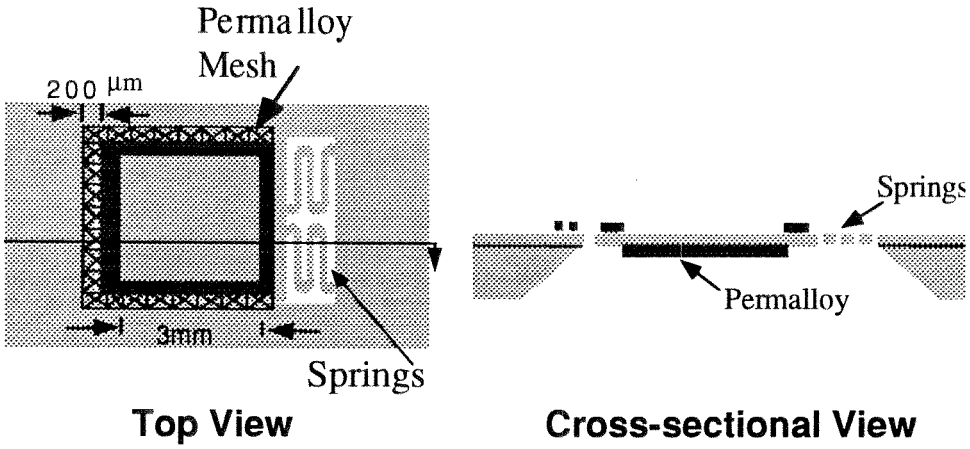


Figure 7.12: Micromirror design.

to increase reflectivity.

Operation and Control

Before we could use the micromirror for angle deflection in holographic storage, we needed to have computer control over the deflection angle. We chose to mount the micromirror on the end of a solenoid and drive it by applying voltage to the solenoid. The solenoid had a fairly constant series resistance, so an input voltage provided the current required to generate a magnetic field. We used a LM317 voltage regulator and a 12 volt battery to control the driving voltage, with a set of 16 switched resistors as shown in Figure 7.15. The input lines could be driven by TTL output logic, allowing computer control of the solenoid voltage. Each resistor varied from its neighbors by a factor of 2, providing voltage control over a range from 2.2 to 6.5 volts. The parallel switching of resistors creates an inverse relationship between the 16–bit control word and the voltage.

Experiments

We performed angle–multiplexing experiments with the micromirror using the setup shown in Figure 7.16. The setup of the object beam is a standard Fresnel plane geometry using a liquid–crystal SLM. The reference beam consists of a near 4–F system

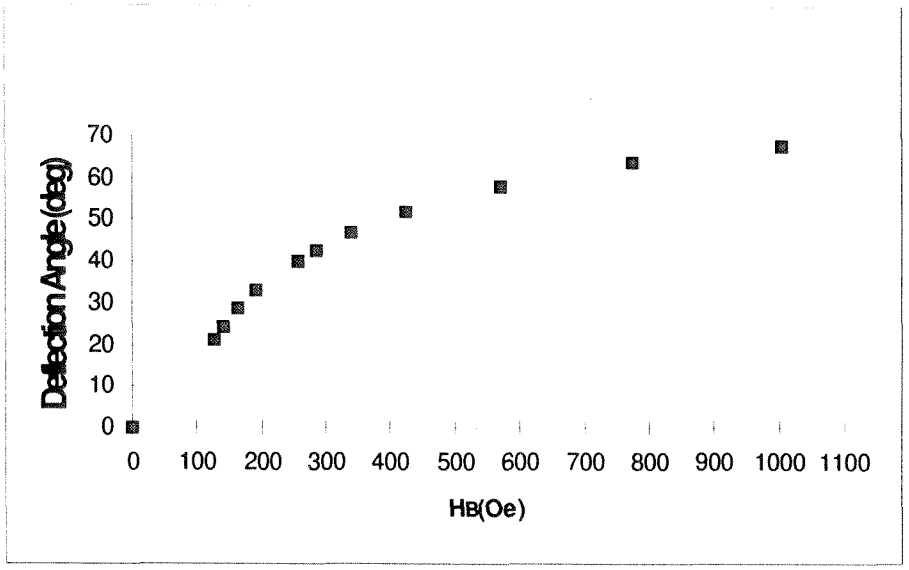


Figure 7.13: Deflection angle as a function of magnetic field.

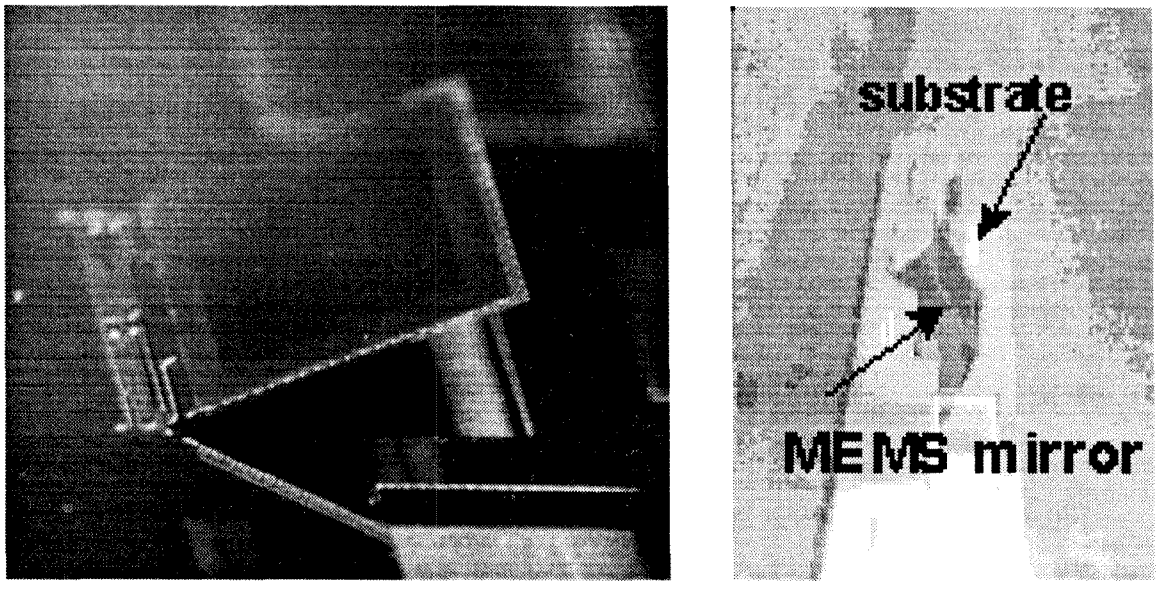


Figure 7.14: Silicon micromirrors fabricated with bulk micromachining.

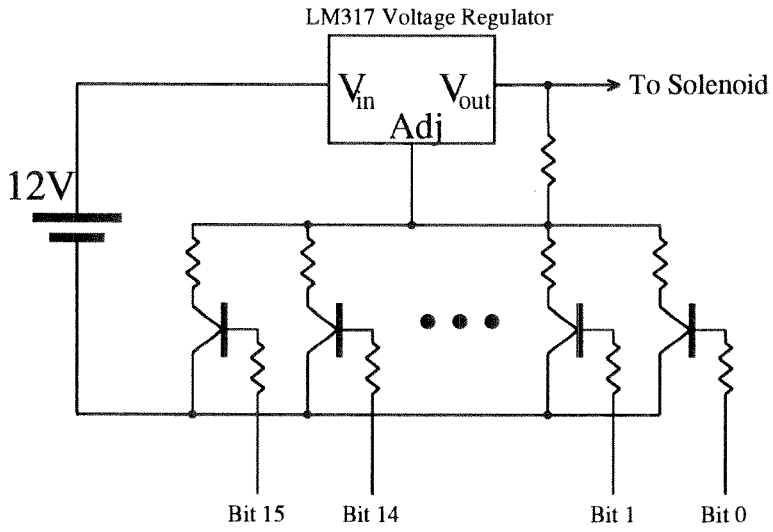


Figure 7.15: Electrical circuit used for driving solenoid

to direct the deflected beam to the crystal. We used a first–generation micromirror which had a convex surface (see next section on Future work). We illuminated the micromirror with a raw (uncollimated) laser beam and used the micromirror curvature to expand the beam. The 2–lens system was spaced slightly farther apart than the sum of the focal lengths, to get a reasonably collimated beam at the crystal. We used an 80mm focal length lens followed by a 300mm focal length lens to get roughly 3.6× magnification. The crystal was placed near the Fourier transform plane of the second lens. At this position, the angle deflection caused some spatial translation. However, this was not a problem as the illuminated spot was 14mm wide compared to the crystal width of 23mm. We illuminated the micromirror with 65mW; the power in the reference beam at the crystal was 40mW.

In preliminary experiments with the micromirror, we had observed that it was affected by mechanical and air vibrations. We decided to suppress air currents around the device by completely surrounding it with a 1cm thick wall of PlayDoh, capped by cardboard. We show this in Figure 7.16 by a thick line surrounding the solenoid, the iris for the input lens, and the 80mm lens. This turned out to be quite impervious to air currents. With this improvement, the micromirror stayed Bragg–matched to a hologram for several minutes without any feedback.[12] In the system we used for

[12]The deflection angle changed slowly due to thermal effects—see the discussion on Future Work.

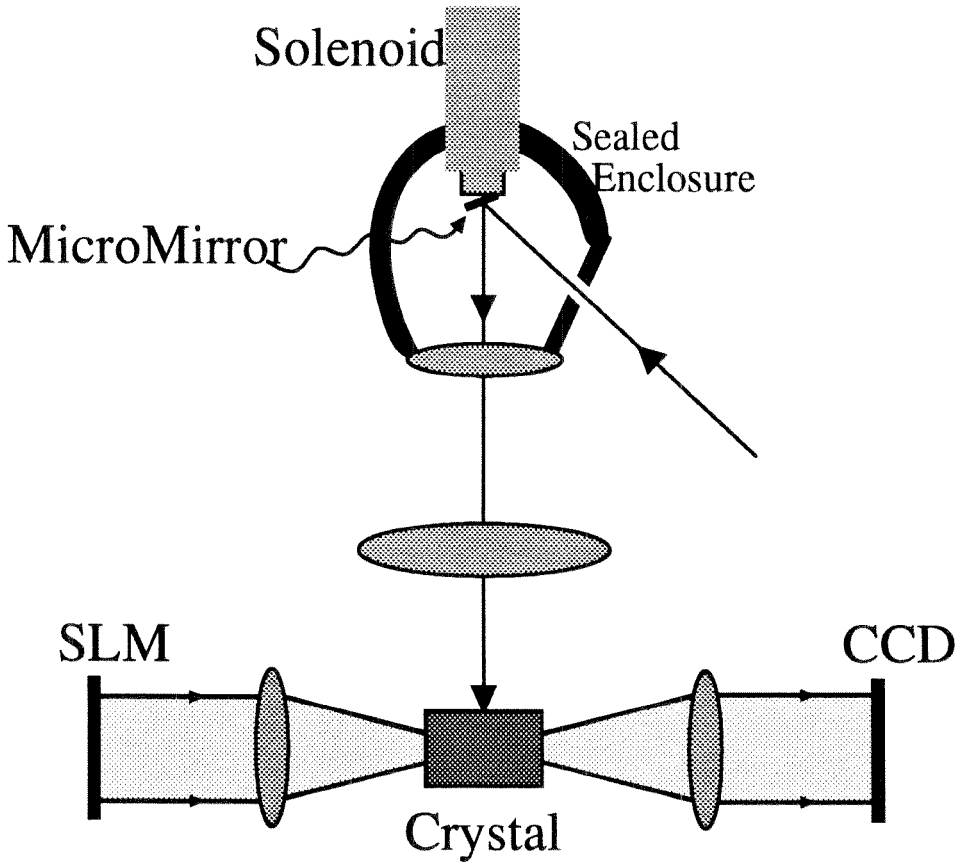


Figure 7.16: Angle–multiplexing using the micromirrors: experimental setup

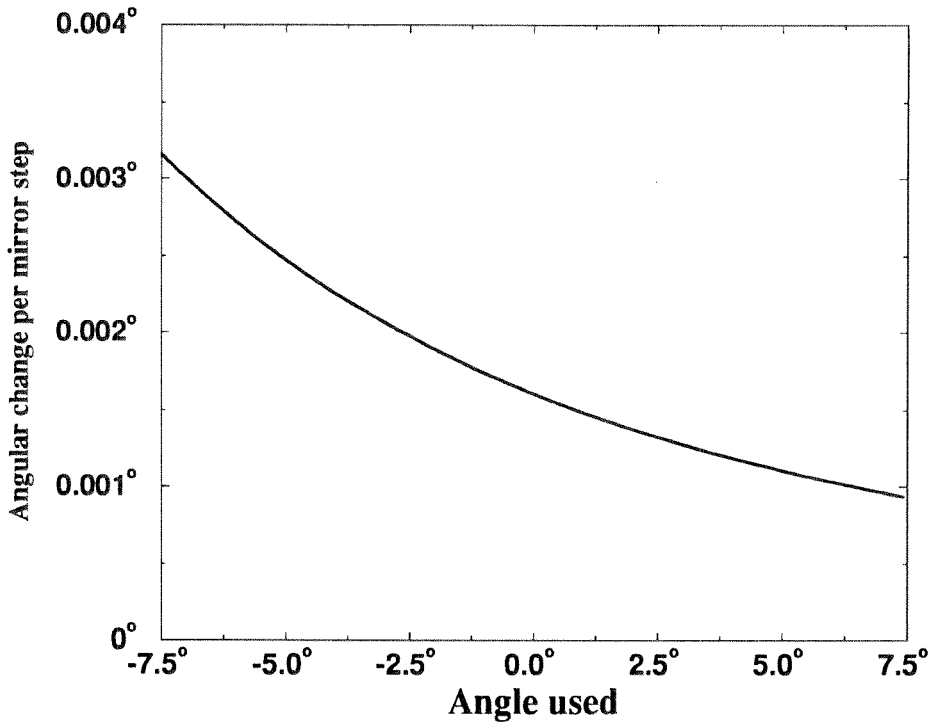


Figure 7.17: Angular resolution of micromirror control circuit.

this experiment, the limitation on the number of holograms was the aperture of the optical system and not the stability of the micromirrors.

We had no trouble deflecting the micromirror to the very edges of the first, f =80mm lens (its aperture was 50mm). However, beams passing through the edge became distorted at the crystal because of aberration. In general, the illuminated spot at the crystal became much larger, decreasing the average intensity and making holograms stored at these angles quite weak. To achieve our target of 100 holograms, we chose to use a scanning range of 15°. Over this angular range, the angle change per one–bit change of the control word was not constant, as shown in Figure 7.17. However, our worst case was more than 45 angle settings per hologram spacing.

We measured an angular selectivity of approximately .05°. In contrast, for an interaction length of 15mm, we would expect the angle selectivity to be .01°. We attribute the difference to the non–plane–wave nature of the reference beam at the crystal. We used an angle spacing of .15° to store 100 holograms. We show several reconstructions in the upper half of Figure 7.18. The diffraction efficiency was in the

range of 10^{-6}–10^{-7}.

In the lower half of Figure 7.18, we show several reconstructions from storage of 100 holograms in a slightly different configuration. For this experiment, we placed the micromirror very close to the crystal surface, eliminating the two lenses in the 4–F system. For this experiment, we placed the crystal right in the Fourier transform plane of the object beam and did not shutter the reference beam at all during recording. In order to preserve dynamic range, we could not take much time between exposures. As a result, we moved the object beam horizontally only a few times during the exposure of the 100 holograms. Note the degradation between the first high quality holograms and the later ones.

Time Response

Due to hysteresis in the electromagnet, the control setting needed to Bragg–match holograms was affected by the immediate past history of the electromagnet. As a result, we could not demonstrate true random–access to the holograms. However, we were able to measure time response by repeated jumping back and forth between pairs of holograms. We videotaped the reconstructed signal and counted the number of video frames between good reconstructions. We plot the resulting transit time as a function of the angular deflection in Figure 7.19. As expected, the transit time was a strong function of the angle change. For small deflections (less than 0.7°), a one–step jump could be used and the micromirror would settle in less than 33 milliseconds. For larger angle deflections, a one–step jump caused unacceptable ringing. However, we were able to improve on this performance by dividing the larger jump into a series of smaller jumps, with intial "acceleration" and final "deceleration" of the micromirror. With this approach, a deflection of 10.7° could be performed in 200 milliseconds. In Figure 7.20, we show a 2.8° deflection when performed as a one–step jump and a 1.4° when divided into a rapid series of 6 smaller jumps. Several consecutive video frames are shown.

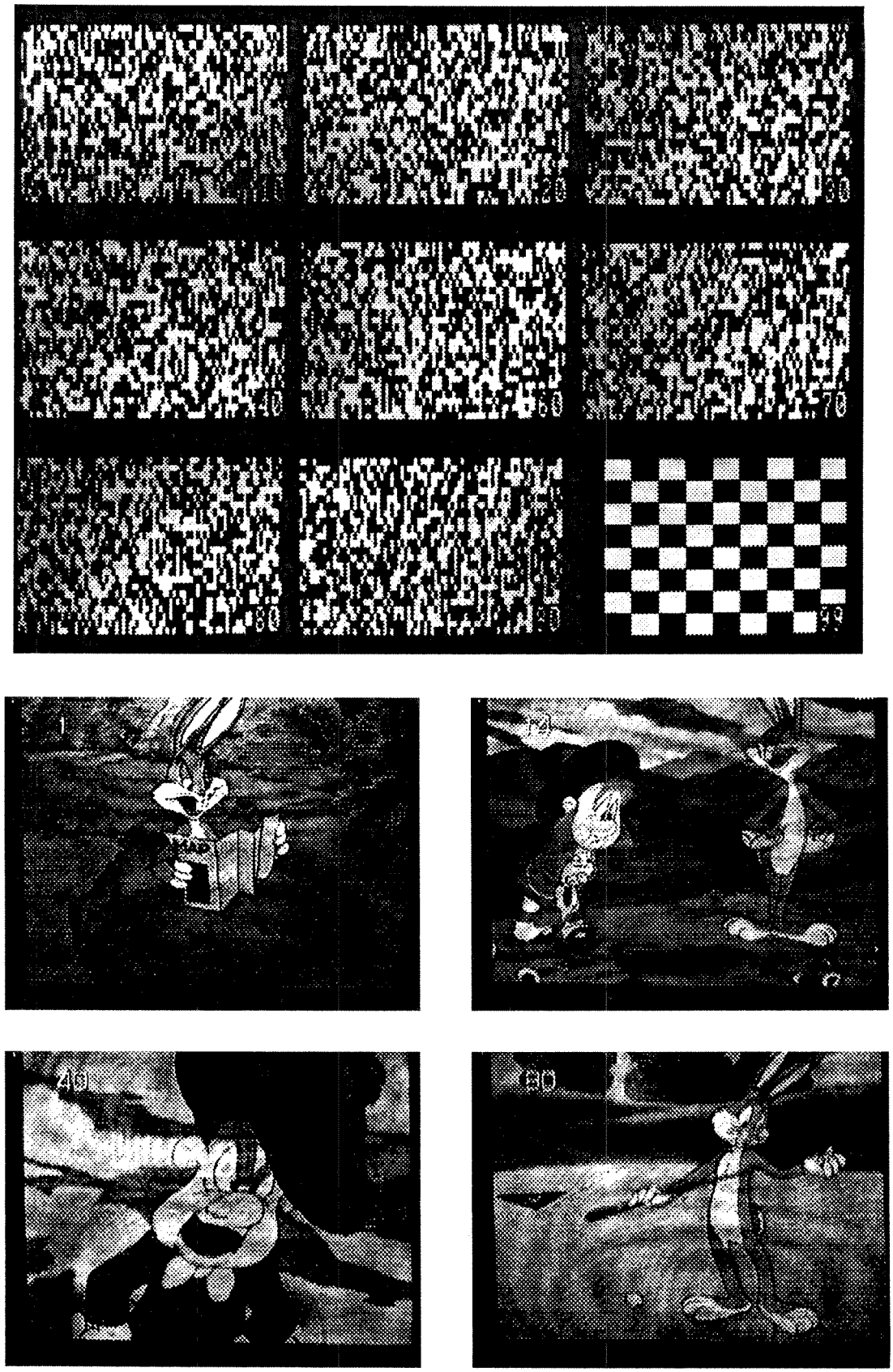


Figure 7.18: Storage of holograms with the passive micromirror. The top corresponds to the experiment with the 4F lens system—the bottom to where the micromirror is

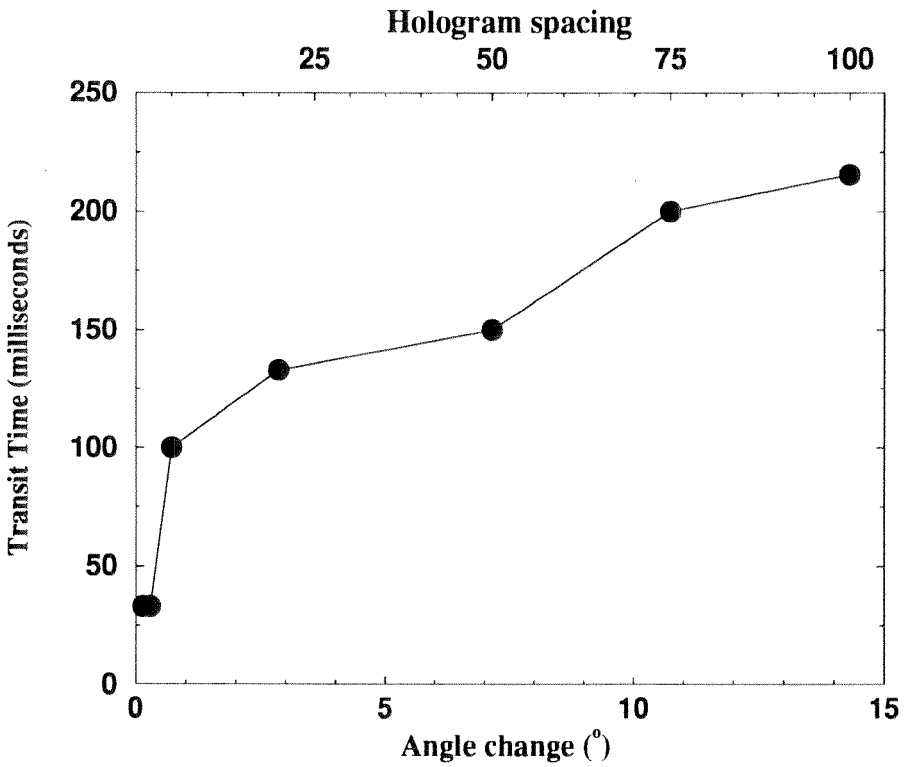


Figure 7.19: Travel time for angular deflections of the micromirror.

Future work

The first–generation micromirrors we used have the permalloy layer deposited on the underside. Stress in this layer creates an overall convex curvature of the flap. In contrast, when micromirrors are electroplated with permalloy on the top, the curvature becomes concave. In addition, both the curvature and the deflection angle are strong functions of temperature. This causes difficulty when the micromirror is illuminated at high intensity, as shown in Figure 7.21. In each column, we shown video frames taken at various points after illumination of a micromirror from room temperature. Shown is the reflected laser beam spot, observed on a wall at a distance of 2.2m from the micromirror. The left–hand column corresponds to a micromirror without aluminum coating; the right–hand column to 6000Å of Al. The aluminum coating reduces the magnitude of the changes in curvature and deflection, but does not eliminate them. In our holographic storage experiment, we used a micromirror with 7000Å of Al, and further bypassed the thermal problem by placing the reference

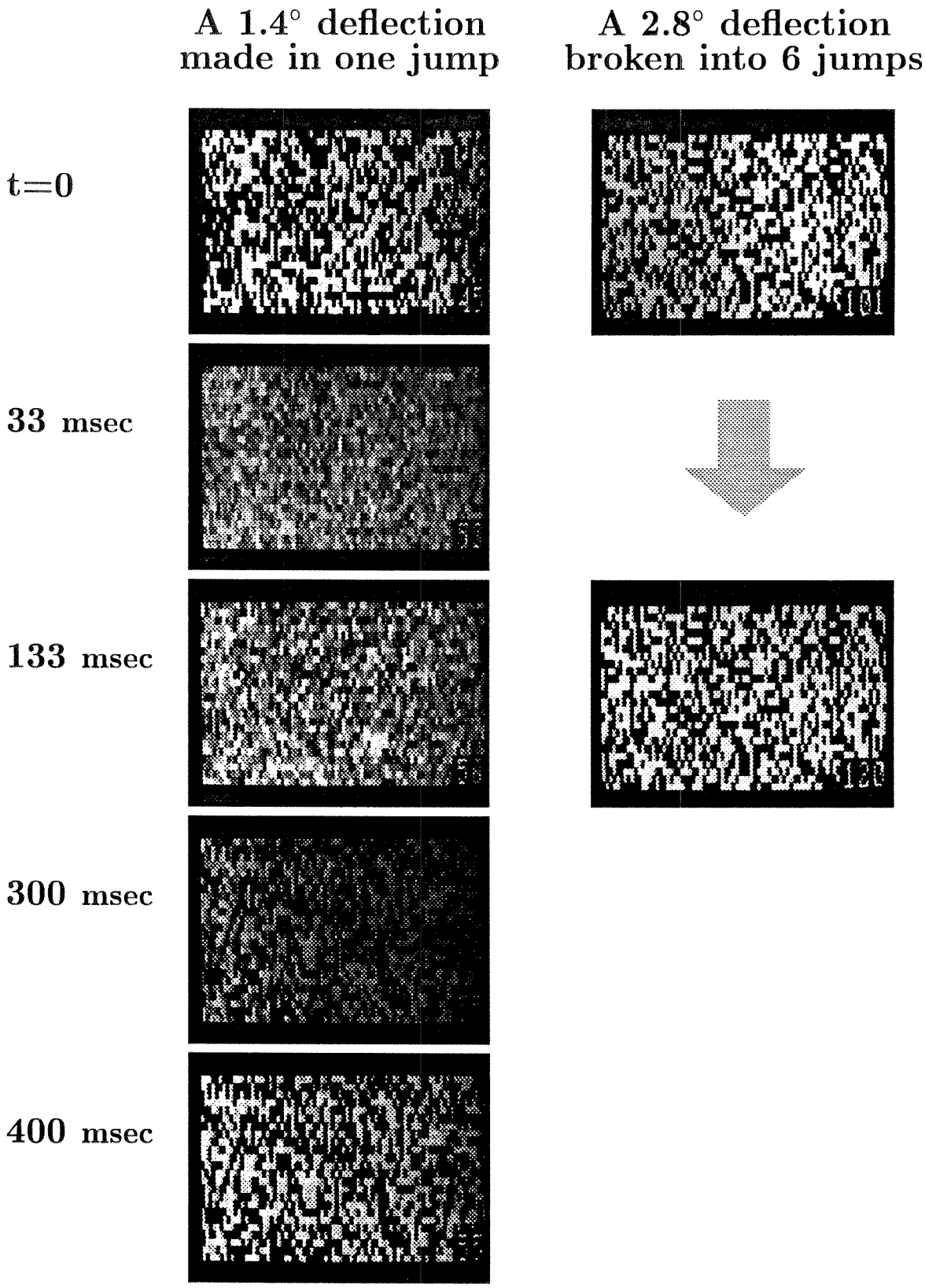


Figure 7.20: Improvement of micromirror settling time by dividing a deflection into several steps.

beam shutter between the micromirror and the crystal. The power incident on the micromirror during the experiment only changed with fluctuations in the output power of the laser.

Future generations of the micromirror are planned with the same amount of permalloy on both top and bottom. The hope is that the stress–induced and temperature–induced changes caused by the two layers will be equal and opposite.

7.6.2 Active MicroMirrors

Dr. Tai's group has also been working on actively–driven flaps, in which an electromagnet is fabricated on the flap surface. The design is shown in Figure 7.22, and a photograph of several working flaps in Figure 7.23. The 30–turn coil is made in a single plane with copper, and connected to the two bonding pads through copper plated on the springs. The electrical connection to the center of the coil is made with a via to a second layer of copper. The hole in the center of the copper coil is filled by electroplated permalloy. The copper coil is electrically isolated by a thin layer of photoresist, on top of which is the 11μm permalloy layer. The role of the permalloy layer is to allow the flap to be biased by an external magnetic field—without this layer, the flap would pivot about the plane of the substrate. When the microflap was designed, it was hoped that the presence of the permalloy within the coil structure would increase the effectiveness of the on–flap electromagnet; however, experiments seem to show that the permalloy only affects the coil's magnetic performance when the external field is extremely weak.

The active flap still requires the presence of a magnetic field. However, this can be provided by a small permanent magnet which biases the microflap to a 45° deflection. The microflap can then be deflected away from this starting point in either direction by running a small current through the coil. By eliminating the electromagnet, these active flaps provide very compact deflection. The amount of current required is less than 50mA. Since the resistance of the coil is $\sim$20Ω, the electrical power requirement is much lower than that of the passive flaps. The disparity between the power dissipation

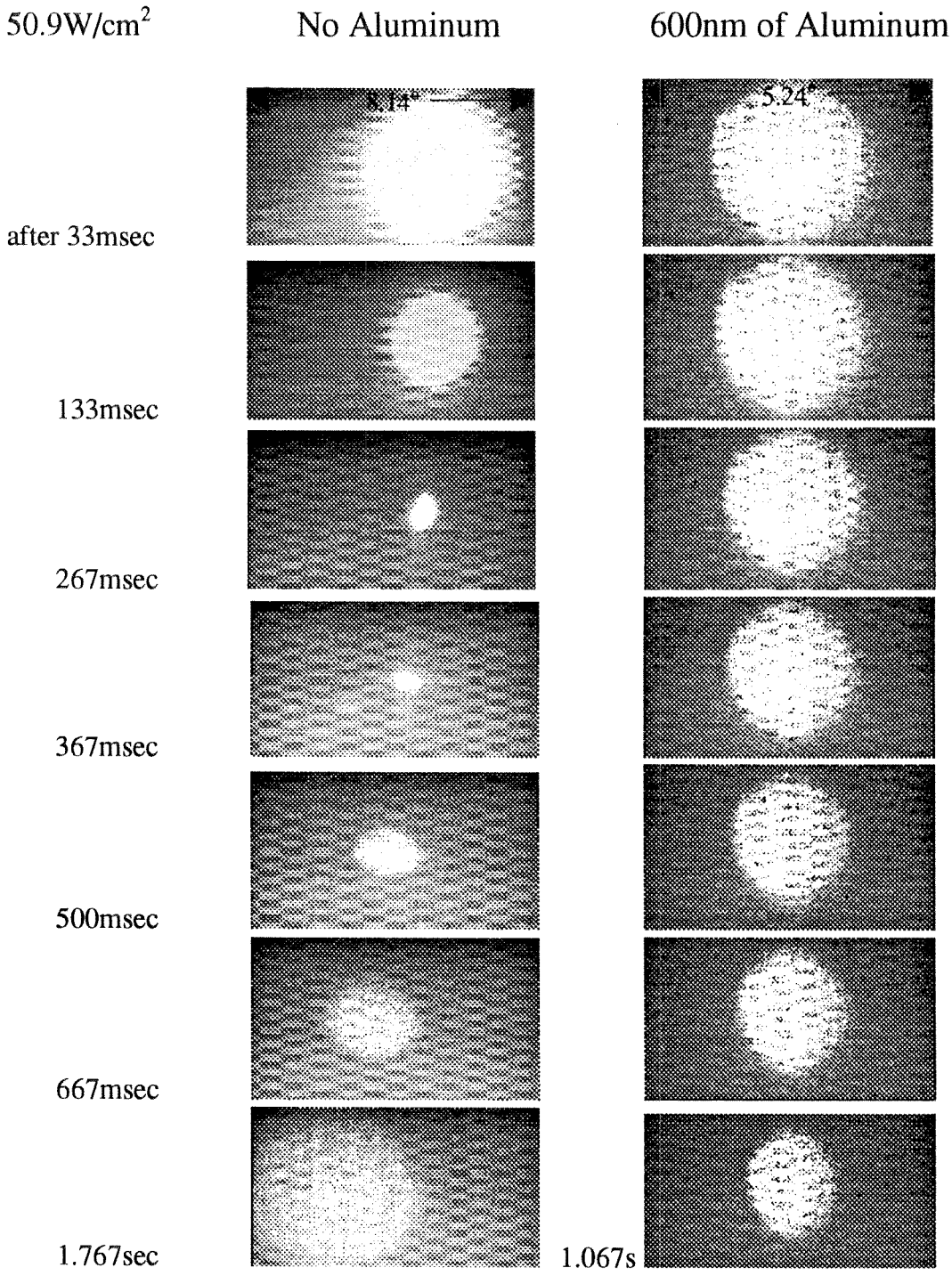


Figure 7.21: Illumination of the micromirror from room temperature: (a) without Al coating, (b) with Al coating.

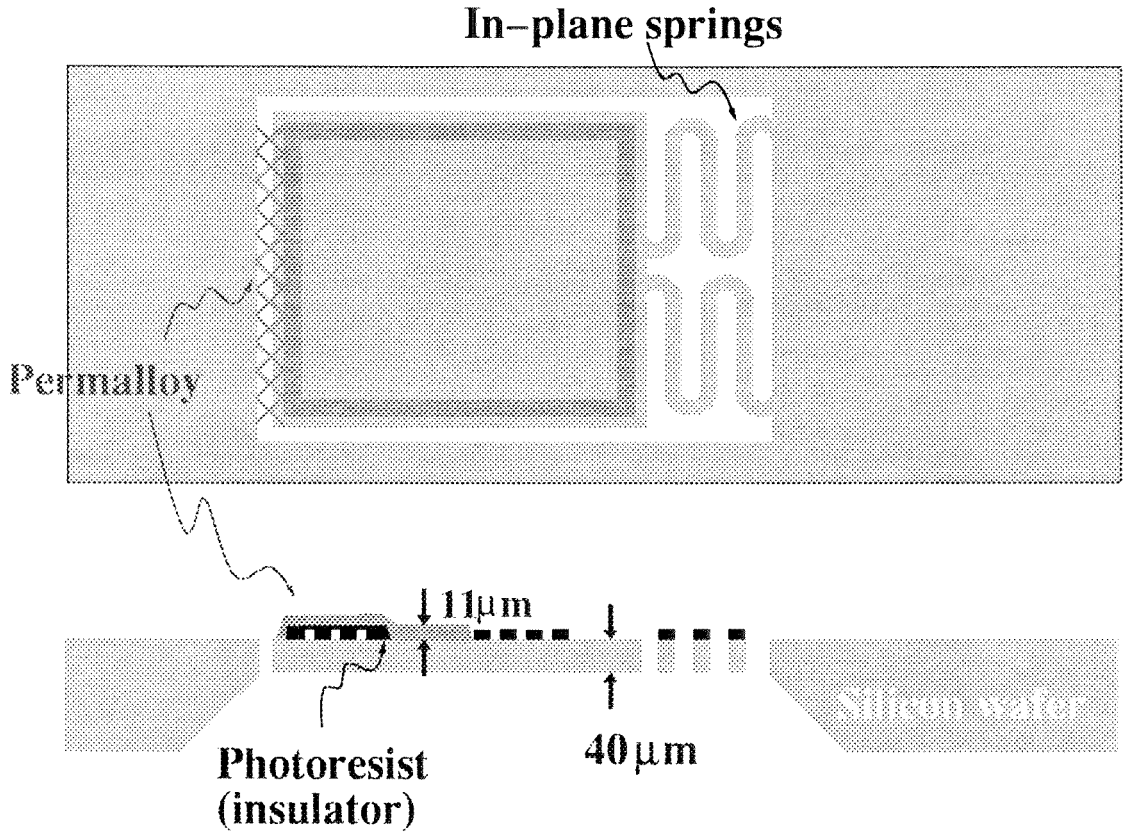


Figure 7.22: Design of the active microflap.

of the two could be decreased by better external solenoids, but the power dissipation of the passive flap can also improve. This is an important point, since the maximum usable deflection angle seems to be determined by the amount of heat–induced flap distortion. As the power dissipation of the flap improves, larger currents can be used and the angle range of the device expands.

Experiments

We stored holograms in analogous storage setups to those described for the passive micromirrors. One setup contained a lens system to image the micromirror deflection

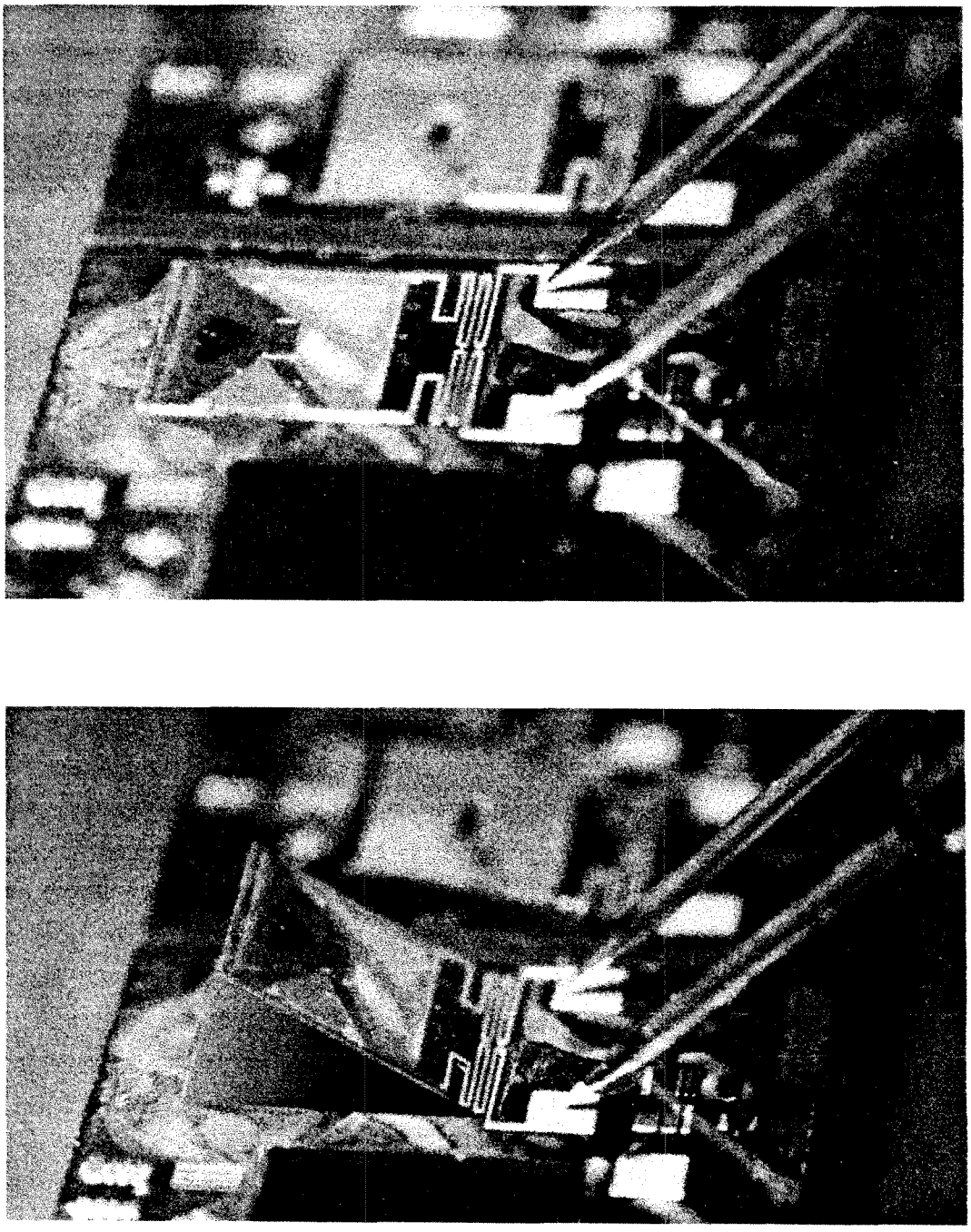


Figure 7.23: Active micromirror, at rest and deflected by a combination of magnetic field and coil current.

onto the crystal—in the second, we placed the micromirror as close as possible to the crystal. In both cases, the "raw" laser beam was reflected off the uncoated back surface (epitaxial silicon) of the active microflap. We found that the back surfaces of active micromirrors were relatively smooth except for a sharp divot in the center, corresponding to the point at which the electroplated permalloy on the front surface filled the center of the copper coil. This divot filled approximately 25% of the vertical size of the flap, and approximately 15–20% horizontally In both storage experiments, we used a cylindrical lens to focus the input beam vertically onto a region of the back of flap outside of this divot region. The reason we did not focus the beam horizontally was to avoid affecting the horizontal spatial frequency spread of the reflected beam at the crystal, which would have affected the crosstalk performance of the device. The reflected beams were not plane waves to begin with, and an additional broadening of the spatial frequency spectrum would have further reduced the number of holograms which we could store.

The flap was biased to approximately 45° by a permanent magnet mounted near the flap substrate. In both experiments, the separation between the biased flap and the surface of the permanent magnet was between 1 and 2mm. For this configuration, the deflection angle is a non–symmetric function of applied current (Figure 7.25). The limit on the amount of current which can be applied is the heat dissipation of the device—at high currents, the curvature of the device changes radically. Above currents of 80mA or so, the devices tended to be destroyed. To supply the driving current to the active micromirror for our storage experiments, we used an accurate current source (ILX Lightwave Model LDX–3207B) with a GPIB input.

On the left side of Figure 7.24, we show originals and reconstructions from the storage of nearly 100 holograms with a 4–F system in the reference beam. For this experiment, some of the holograms were placed too close together in angle and were not distinguishable—approximately 70–80 holograms could be individually distinguished out of the 100 stored. On the right half of the figure, reconstructions are shown from 100 holograms stored with the microflap placed right at the crystal surface. In this second experiment, all of the holograms were easily distinguishable. As with

the passive micromirror experiment without reference beam lenses, the crystal was placed in the Fourier transform plane of the object beam without a random phase plate. Because we had no way to shutter the reference beam without causing thermal resettling of the micromirror, the holograms were stored in rapid consecutive order with minimal horizontal movement of the object beam between exposures. With improvements in the thermal performance of the microflaps, we will be able to take more time between exposures to move the focussing object beam within the crystal, and avoid the nonlinear recording and image distortion that begins to occur in this 100 hologram experiment. Alternately, if the deflection of the object beam within the crystal can be done with stages which do not require significant move and settle delays, then the difficulty is also alleviated.

To measure the deflection speed of the device, we used a sinusoidal current input of moderate amplitude. We measured the angular deflection of the device by directing the deflected beam onto a CCD camera. The measured angular deflection as a function of driving frequency is shown in Figure 7.26. There is a strong mechanical resonance just above 65 Hz, which implies that the current generation device could be used for rapid deflections down to access times of approximately 16ms. The location of the resonance was empirically found to be a function of the external magnetic field, and could range from 50–80 Hz.

7.7 Rapid Bragg–matching

When we read out our angle–multiplexed holograms, we cannot simply assume that the angle which we used to store the hologram will be the best for readout. Small changes in crystal size (through temperature–induced expansion) or in bulk index of refraction may change the angle required for Bragg match [246, 390]. On top of this, there may be changes in the wavefront controller. For the best results, we should plan on finding the optimum Bragg match condition for each hologram before we detect it with the array detector.

Let's say that we are planning on reading 1000 holograms per second. We will need

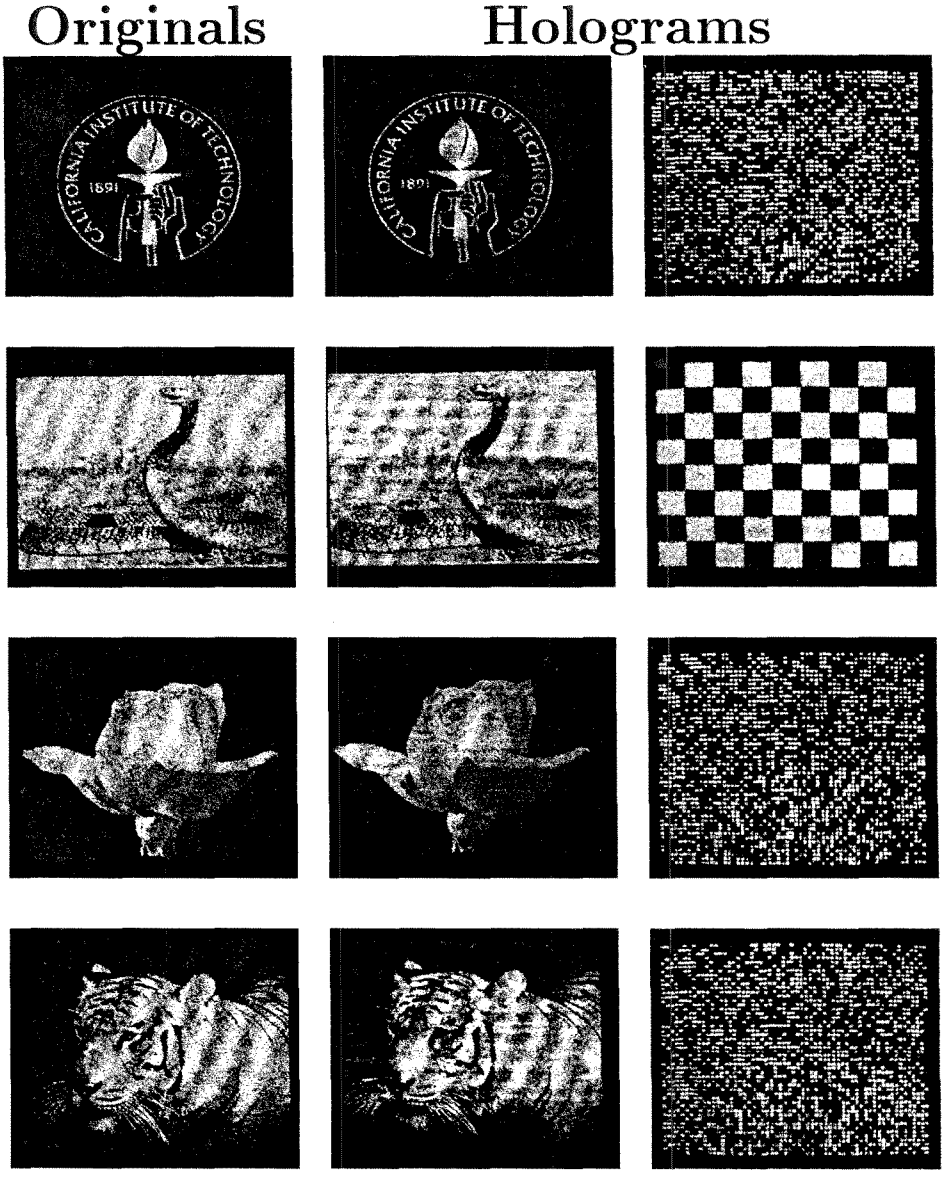


Figure 7.24: Originals and reconstructions from holograms stored with active micromirror

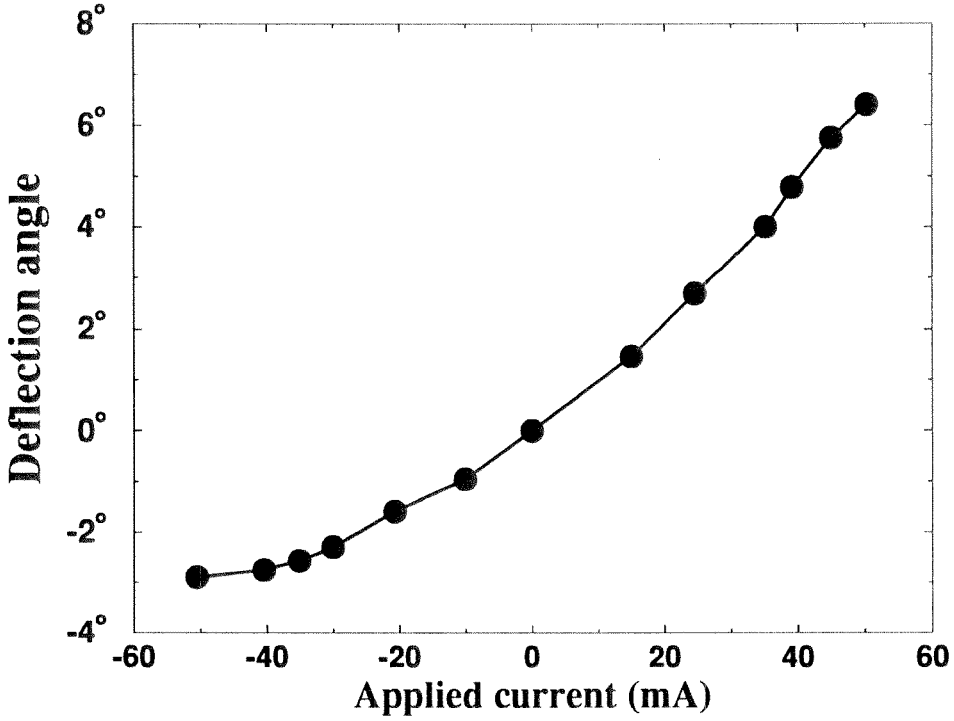


Figure 7.25: Deflection of the active microflap versus applied current.

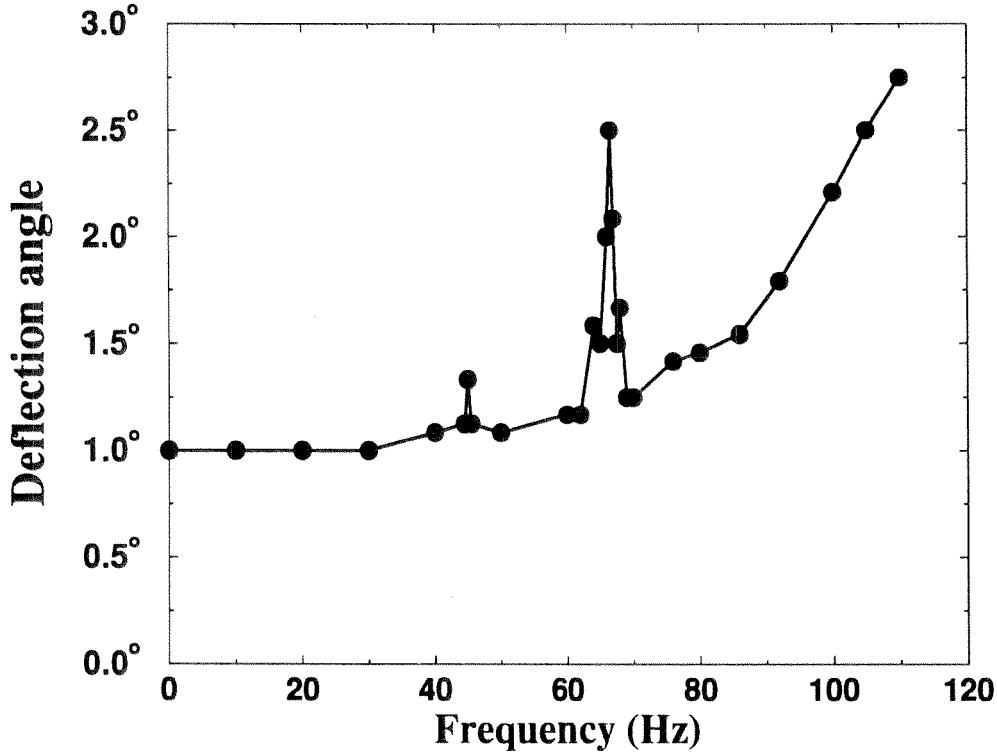


Figure 7.26: Frequency response of the active microflap.

to spend most of the 1 millisecond integrating the holographic reconstruction with the megapixel array detector. We can take at most 100μsec to find the Bragg–match condition. To accomplish this, we need two things:

- an angle deflector capable of high–resolution deflections ($< 0.1 \times$ the angular selectivity) in 10μsec. This gives us 10 move–and–measure cycles to find the Bragg–match condition.

- a way to detect the power in the entire reconstructed wavefront. We can't just use some of the pixels in the detector array, because each one gets very few photoelectrons during the short integration period. In addition, to get a reasonable measure, we would need to construct additional electronics to combine the results from several pixels. On the other hand, we can get enough detected photoelectrons by deflecting a small fraction of the reconstructed wavefront and focussing it onto a single detector. This allows us to keep most of the power in the reconstruction on the array detector, for the 900μsec of integration once we have found the Bragg–match condition.

We can perform rapid high–resolution angle tuning by translating the point source in the reference arm with a piezoelectric crystal. These devices require driving voltages on the order of 100–1000V. A piezoelectric crystal can have a frequency resonance around 10kHz–100kHz and can translate up to 100μm depending on the length of the device [389]. In comparison, the angular selectivity of a 1cm wide reference beam in the 90° geometry requires only a 5μm movement.[13] This can be further reduced by reducing the focal length of this reference beam lens.

For our second task, we need both power and array detection of the reconstructed output. We can use a permanent beamsplitter, perhaps diverting 1% of the power to the power detector and the rest to the array detector. If the reconstruction contains enough power for the array detector, then the power detector should be able to work with 1% of the output power. The power detector receives 1% $\times$ 500,000 $\times$ the power delivered to a single ON pixel of the array. However, we have only shrunk

[13] in the back focal plane of a lens of focal length 10cm.

the integration time by a factor of 100. If we had 1000 photoelectrons in each pixel of the array after 1ms, we'll have 50,000 in the power detector after 10μsec. This should give us the few bits of dynamic range that we'll need in order to determine the Bragg–match condition.

7.8 The Object beam

In Chapter 5, we decided that spatial multiplexing was needed because the storage capacity at one location is not sufficient to justify holographic memories. An additional consideration is the cost of an SLM and detector array for each storage location. To keep the cost per bit of capacity from becoming too high, we need to be able to store at multiple locations for each SLM and detector array pair. In the remainder of the chapter, we consider various ways of routing the object and reference beams to the crystal, and of routing the reconstructed holograms to the detector array. We begin with the object beam, and take up the reference beam in the next section.

7.8.1 Using just one SLM

One option for the object beam is to have just one SLM for the entire system, and then pipe the information–carrying beam to the appropriate location for storage. We can use 4–F systems as image "repeaters." This branched pipeline approach may be attractive if the SLM is extremely costly. We can use the same techniques to deflect the reconstructed images to a central array detector, with commensurately higher expectations on the switching speed.

Switched polarizing beamsplitters In Figure 7.27, we show two ways that we could perform image routing. The first (shown on the left side of the figure) is to use polarizing beamsplitter cubes to split the object beam into one of two paths. The path is chosen by the state of two liquid–crystal cells, which can either rotate the polarization by 90° or pass the beam unhindered. We always want the polarization within the 4–F systems to be vertical, so the second liquid–crystal cell rotates hor-

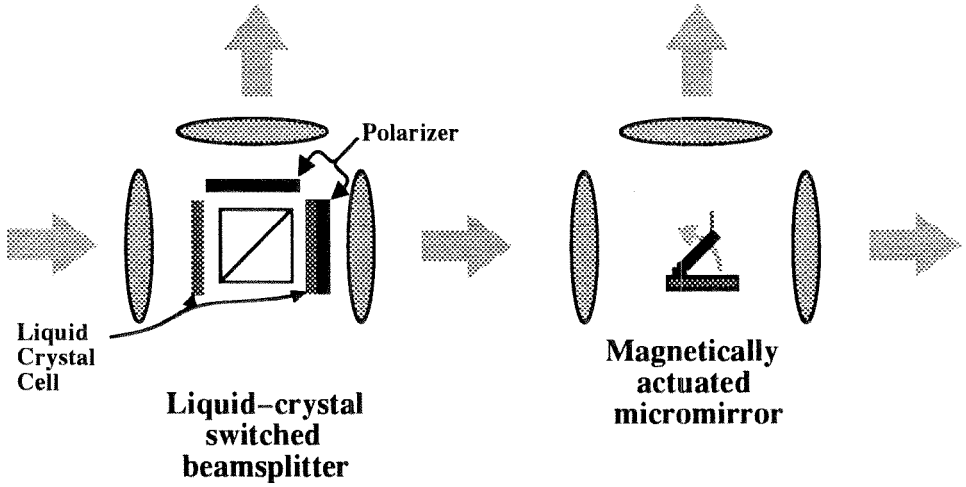


Figure 7.27: Methods of routing high–resolution images.

izontal polarization (which passes straight through the beamsplitter cube) back to vertical polarization. Polarizers are required only if the contrast ratio between the switched and unswitched state is insufficient.

We tested a liquid–crystal switch using two twisted nematic liquid crystal cells built by Jean–Jacques Drolet. Figure 7.28 shows the switching of a high–resolution image from one path to another. The figure shows two freeze frames from the video outputs from two separate CCD cameras. A video mixer was used to combine the two images; a full–screen image in the background, and an inset image in the lower left corner. The images were originally displayed on an SLM, and imaged to the center of a polarizing beamsplitter cube (the setup was very similar to the left half of Figure 7.27). Two additional 4–F systems sat at the exit faces of the polarizing beamsplitter cube, and imaged from the center of the beamsplitter to the two CCD cameras. The output of a square–wave generator was switched into the two liquid–crystal cells to deflect the SLM image down one path or the other. The free–running videotape of the switching demonstration showed only one dead frame between the two states, indicating a 40-50 msec switching time. Cells built in our laboratory have demonstrated switching times as low as 25msec [316]. The primary consideration is the gap between the two glass plates. The cells that we used were designed for 488nm with 10μm thickness.

Because we used the cells at 532nm, we placed polarizers after the two exit faces of

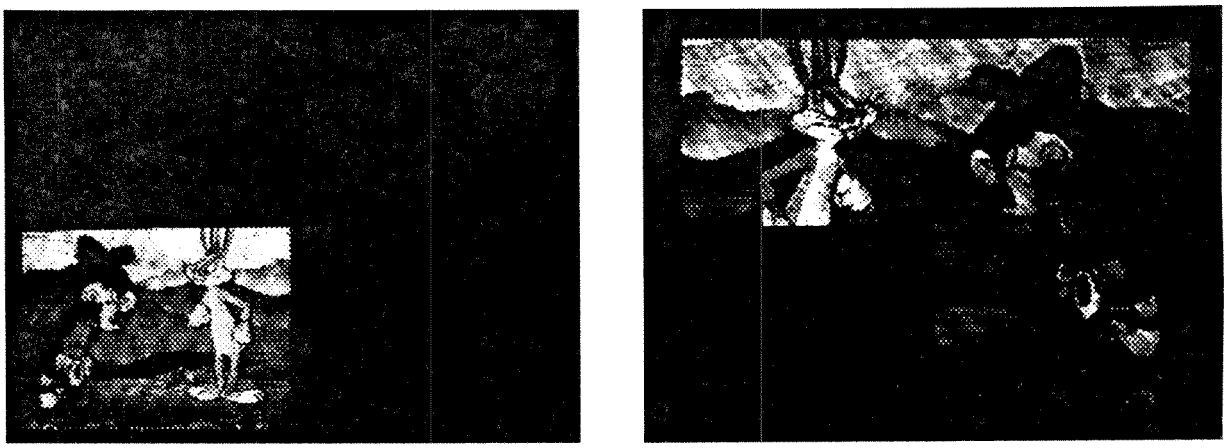


Figure 7.28: Deflection of high–resolution images using a liquid–crystal switched beamsplitter

the beamsplitter cube to increase contrast. We measured a contrast in the straight–through path of 13,500:1, and a contrast in the 90° deflection path of 650:1. In comparison, the contrast without the polarizers was 350:1 and 170:1, respectively. The efficiency without the polarizers was 76% in the straight path and 78% in the deflected path. These values could be conceivably increased by antireflection coating of the glass substrates before assembly of the cell.

Video–rate switching times are probably acceptable if we are just deflecting images towards their storage location. If, however, we are routing reconstructed holograms towards the array detector, we need faster switching. We can change from nematic liquid crystals to ferroelectric liquid crystals for this task.

Micromirrors A second method for deflecting high–resolution images is shown in the right half of Figure 7.27, using the micromirrors which we described in Section 7.6. Instead of positioning the image plane of the SLM (which may be as large as 2cm) at the micromirror surface, we can place the micromirror in the Fourier transform plane of the SLM. This allows us to make the micromirror only 4mm across or so, assuming lenses of focal length <100mm. We can get around any uncertainty in deflection angle of the micromirror by using a deflection stop. We just drive the micromirror with a sufficiently large field so that it opens until it contacts the stop. This should allow reasonable repeatability in the deflection angle.

However, we must ensure that the micromirrors are flat, or if curved, are of a known curvature. Any curvature of the micromirror will tend to distort our high–resolution image and make it unrecognizable. If the curvature of the mirror is always a known quantity, we can treat the surface as an additional lens and redesign the rest of the optics so that the image is properly focused at the detector array. Changes in curvature, either from mirror to mirror, or in one mirror due to temperature changes will be most difficult to deal with.

Conclusion It would seem that routing high–resolution images through multiple 4–F systems is not the best solution. If we are using switched beamsplitters, we quickly incur costs much larger than the cost of an individual SLM. The culprit is not the liquid–crystal cells, but the high quality polarizing beamsplitters and low f/# optics. For example, the Kopin 640×480 SLM costs ~$2000 with driver electronics—in comparison, a 1 inch polarizing beamsplitter cube costs ~$500. From a cost perspective, it is more economical to buy many SLMs. The cost of the detector array is still unknown, considering all of the features (see Chapter 1.3.1) that we require of it. For instance, a 750 × 480 pixel cooled CCD array with 100% fill factor, video read–out rate, but no error–correction circuitry, currently costs $12,000. However, we should also consider that the liquid crystal cells required in the readout arm must be high–quality ferroelectric cells of >2cm aperture, which will be of considerable expense themselves. There are tradeoffs between contrast ratio and efficiency, reducing either the power of the reconstructed signal beam reaching the detector array or the SNR (through multiple ghost images from other paths).

In comparison, the micromirror solution seems quite elegant. The micromirror itself can be made cheaply using batch techniques, leaving the bulk of the cost to the lenses in the 4–F systems. The efficiency of the system can be quite high if the mirrors are coated, since the beam strikes only the mirrors which are needed to deflect the beam. However, the micromirrors require additional development in several areas. We mentioned the need for a flat deflecting surface. In addition, the deflection time must fall to much less than 1ms (we want to spend *most* of the 1ms integrating the

reconstructed signal beam). And the springs must be incredibly reliable—with one thousand possible 0° to 45° swings each second, one micromirror might perform 31 billion open–close cycles in a year.

One final consideration for both of these methods is assembly of the storage system. Since the SLM data can take many paths to the detector array, each must be focused, aligned, and registered exactly. Complicating the task is the fact that each beamsplitter (or micromirror) participates in multiple paths. Care must be taken to align a single path from SLM to detector array, after which each path is added by aligning only one additional element. This difficult alignment process adds an additional cost to the total system cost.

7.8.2 Using multiple SLMs

It would seem that we are back to a system which resembles the one we designed in Chapter 5: an SLM, a large lens, a stack of crystals, a second large lens, and a detector array. Of course, we can put many such modules in the box we call our memory, but we are still essentially using one SLM for each set of storage crystals. This is not such a bad thing after all: our total system input rate, output rate, and capacity increases for each additional module that we add. In contrast, if we were to deflect the object beam to and from multiple lens systems, we would increase only the capacity (and possibly decrease our efficiency) for each additional storage location.

Before we move on to how what we might add to (or subtract from) the object arm we showed in Chapter 5, we should ask "Why do we even need a system which contains more than one module?" The first reason is that we may decide that we don't need one high–power laser for each SLM—we can cut down on the system size and power usage if we share one *recording* laser between multiple modules. The second reason is that we will be hard–pressed to reach our goal of 1 terabit of storage without using multiple modules. And the third reason is that we will need multiple modules to implement our solution to the volatility problem (See Chapter 8.1).

The scanning point source

One advantage to using a dedicated set of lenses for each SLM/CCD pair is that we can use deflection to reach more storage locations, just as we did with the vertically stacked spatial locations in our 160,000 hologram system. We need to use one of the deflection technologies which we reviewed earlier in this chapter to perform this deflection.

When we were describing the mirror array, we spoke of deflecting a focused spot across the surface of the mirror array. We said that this focused spot moved in the two transverse directions to implement angle scanning and was deflected or tilted to perform spatial movement of the spot. To create this movement of the focused spot in the *front* plane of a lens, we deflected a collimated laser beam at the back focal plane of the same lens. However, we originally created this collimated beam by spatially filtering a focused beam. In the remainder of this thesis, we consider designs which combine these two steps: we remove all of the collimation optics between the spatial filter and the focused spot plane (where the mirror array would be in the reference arm). We can then perform all of the angle deflection we need by translating this point source.

One point source that we can use is the tip of a single–mode fiber. The flexibility of the fiber allows us to place one end on a translation stage while the other is fixed elsewhere, near the output of the high–power laser. The same laser output can be coupled to fibers which go to the corresponding object and reference beams. The length of the fiber can be less than several meters. The problem of directing collimated beams to the correct module is reduced to a problem of directing raw laser beams to the correct fiber with high coupling efficiency. However, these are problems which have long been considered for switching speeds far higher than the ones we will need. The differences are in wavelength and perhaps in power level.

In the object beam, we can mount our fiber tip on a relatively slow pair of linear translation stages, in the back focal plane of our single collimation lens. We were going to have to use a spatial filter and a collimating lens anyway, and we have performed

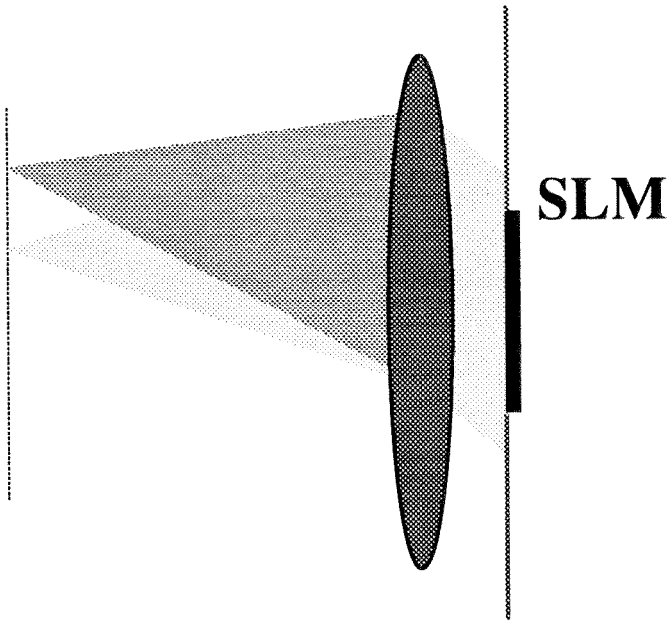


Figure 7.29: Control over the illumination angle of the SLM by position and tilt control of a point source.

the deflection in the same package! The illumination of the SLM at various angles is shown in Figure 7.29—note that we tilt the diverging cone towards the optical axis. This allows us to place the SLM close to the collimating lens as shown in the figure. The remainder of the object arm is identical to the systems pictured in Chapter 5: SLM, Fourier transform lens, stack of crystals, second Fourier transform lens, and the detector array. In Chapter 8, we will add phase–conjugate readout to the object beam we have pictured here. Our motivation there will be to avoid the effects of lens aberrations on the reconstructed signals received by the array detector. Since with phase–conjugate readout we do not need to worry about finding an image plane on the far side of a 4–F system, we can also place the the Fourier transform lens close to the SLM. Figure 7.30(a) shows the illumination of an SLM in a typical object arm. In comparison, Figure 7.30(b) shows the amount of space we save by putting the SLM closer to the lens f_{object}. In addition, this scheme decouples the range of usable illumination angles from the $f/\#$ of the Fourier transform lens.

One worry about fiber might be the stability of the interference pattern during storage. If the unclamped parts of the fibers move after the linear translators stop

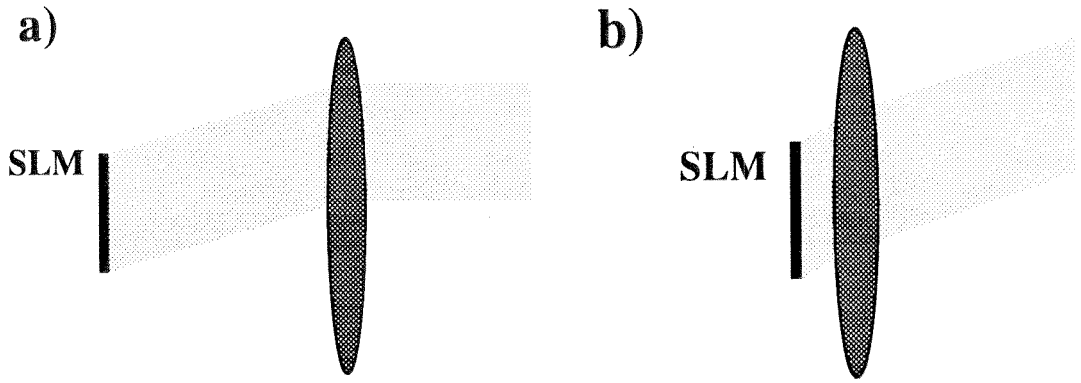


Figure 7.30: Illumination of the SLM in (a) a typical object arm; (b) an object arm for phase–conjugate readout.

moving, the optical path length may change. This will make our interference pattern shift in phase within the crystal. It may be necessary to mechanically clamp the fiber (with some slack) at some distance from the translation stages, and in addition, to impose some delay between movement of the fiber tip and exposure for recording. This will decrease our input data rate, though not dramatically.

Spatial light modulator/detector arrays

We mentioned in Chapter 1.6.5 that reflective SLMs with built–in detector arrays are being proposed and built [327, 328]. The advantage of these devices is that the SLM and detector array are guaranteed to be registered with each other and the object arm is self–aligning. Phase–conjugate readout can be used to redirect the holographic reconstructions back toward the original SLM, removing any aberrations as the light passes back through the lens. The recording/readout process is shown in Figure 7.31. Figure 7.31(a) shows the recording of holograms and Figure 7.31(b) shows readout onto the co–located detector array. In fact, the entire system could be constructed without a lens, with the SLM/detector array placed against one of the faces of the storage crystal [327, 328]. Our experience, however, has been that the detector array needs to be carefully shielded from any scattering of the strong reference beam, especially if the holograms are weak. An extremely effective way to shield the detector array is to limit the field of view with a lens and a small aperture.

In this case, however, the lens does not have to be high–performance, since we're going to use phase–conjugate readout anyway.

This combined SLM/detector array device is a difficult research goal. Our requirements for a stand–alone detector array—on efficiency, fill factor, pixilation, readout rate, additional circuitry for error correction—may still prove to exceed what it is possible to build on a reasonable budget.[14] Now one must add liquid crystal (or other) modulator pads and their circuitry without sacrificing on the performance of either SLM or detector array. The elegance and compactness of the resulting system may well be worth a certain amount of sacrifice in performance, though, especially if this sacrifice is solely in the number of pixels. The resulting loss in capacity and readout rate could always be made up by building multiple modules, especially if each module has become quite small.

7.9 Reference beams

We can use the idea of a scanning point source to also provide the deflection in the reference arm. Figure 7.32 shows how the different types of multiplexing can be implemented—these directions are analogous to the operation of the mirror array. Horizontal movement of the focused spot is angle multiplexing, while vertical movement implements fractal multiplexing. Spatial multiplexing is implemented by changing the direction of the central ray of the expanding cone of light. This can be done vertically, as we did in the 160,000 hologram system, as well as horizontally.

It is likely that we will not want the diverging cone of light to be symmetric. For instance, in our experimental demonstration of 10,000 holograms, the ratio of the width of the reference beam to its height was move than 6:1. We can implement some of this asymmetry in divergence angle by manipulating the size of the fiber tip. In addition, we can use a pair of lightweight diffractive optical elements to shape the diverging cone of light.

[14]See Chapter 1.3.1 for mroe details on the performance requirements of the detector array.

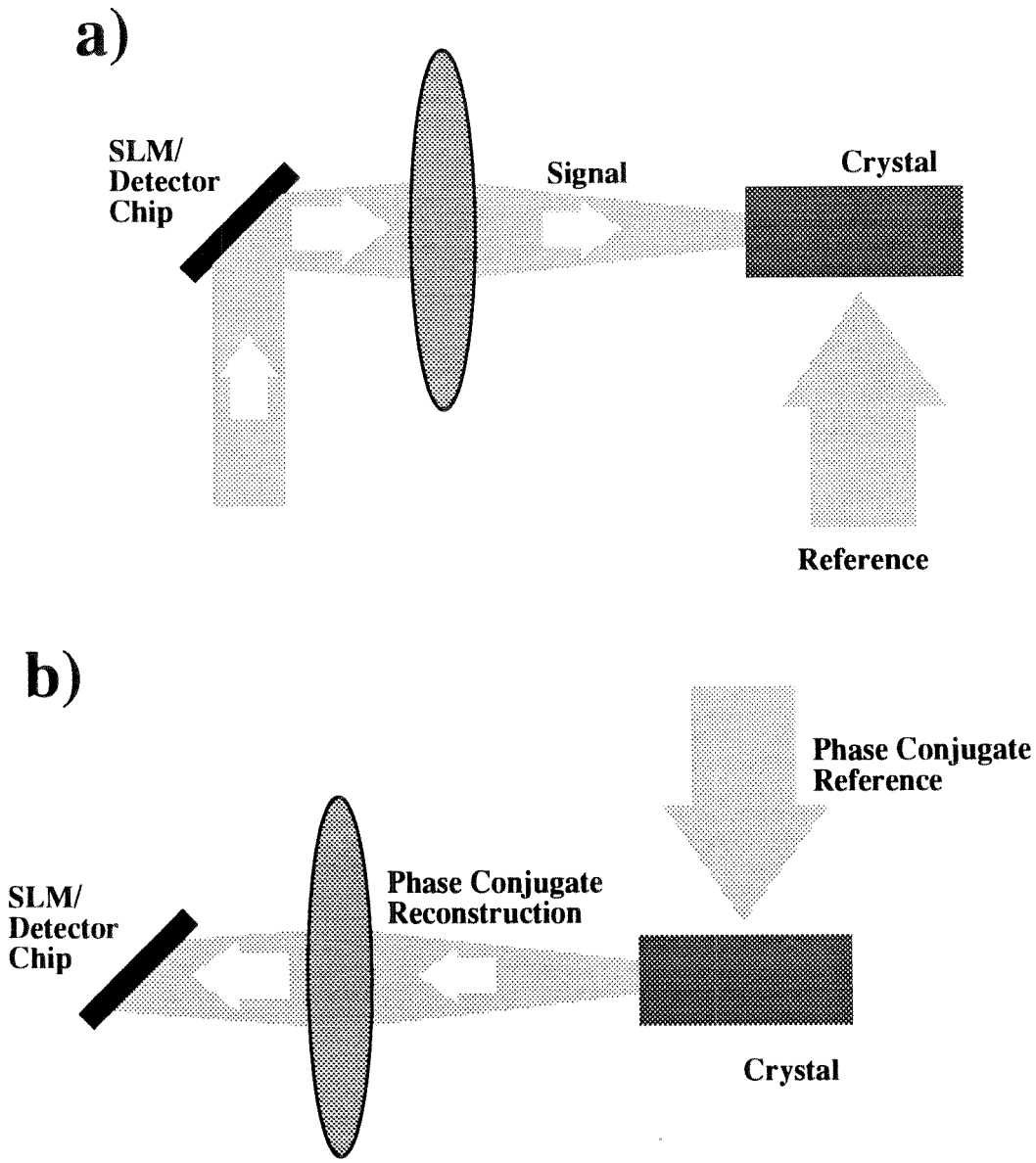


Figure 7.31: Operation of an integrated SLM/Detector array using phase–conjugate readout. (a) recording of a hologram using the on–chip modulator pads; (b) readout of the hologram onto the onchip detectors using the conjugate reference beam.

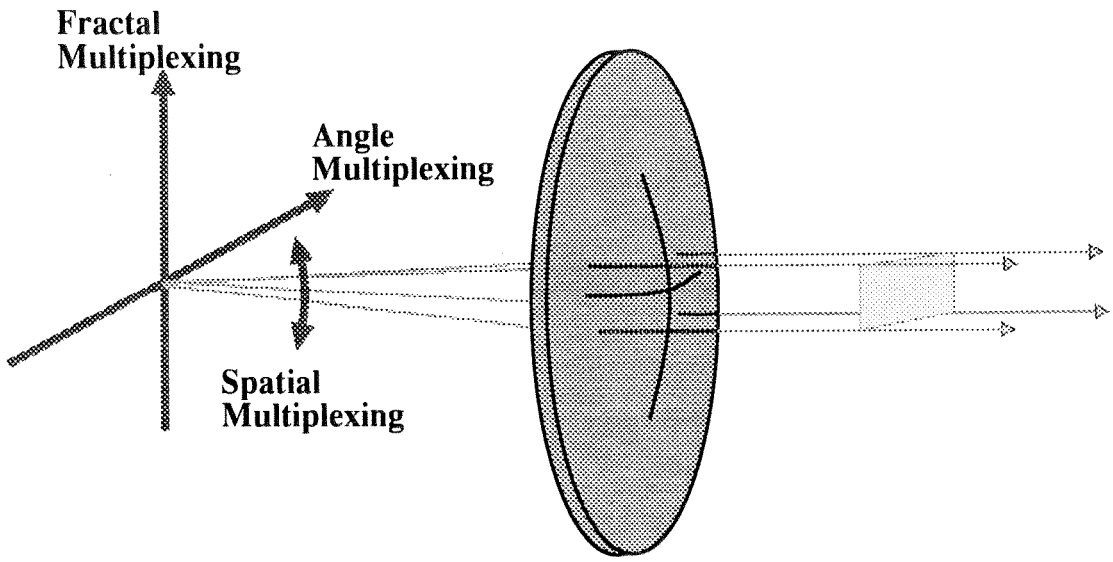


Figure 7.32: Implementation of angle, fractal, and spatial multiplexing by manipulation of a point source in the back focal plane of the reference arm's collimating lens. This plane is equivalent to the surface of the mirror array from Chapter 5.

7.9.1 Laser diodes

We mentioned that, for recording, the light from a single high–power laser could be coupled into fibers and directed to both the object and reference beam. For readout, there is no reason why we cannot use a different laser source, as long as the wavelength is similar. We could use a single laser diode, or a laser diode array. Of course, we are still talking about wavelengths in the green or blue—and at the time of this writing, blue laser diodes currently in development are limited to lifetimes of several hours [391]. However, there is an enormous research effort dedicated to bringing blue and green laser diodes to fruition, because of the hope for higher density in bit–oriented optical disk storage. It is probably safe to assume that green laser diodes of sufficient operating lifetime will become available sometime soon. Even if this does not happen, or if the output power is insufficient, there are many compact laser devices which are currently available in the green which use second harmonic frequency doubling. Since our requirements for a readout source are much more relaxed than for a recording source, an inexpensive doubled diode laser may be a solution for us.

If laser diode arrays (or arrays of doubled laser diodes) at our target wavelength become available, we can use the multiple laser diodes to perform some or all of the angle multiplexing task. In order to match all of the spatial frequencies in the image, we may even want to be able to perform slow temperature–tuning of the wavelength of the laser diode or diodes [349].

At this point, one might ask why we are implicitly assuming that we should use the older photorefractive materials/dopants which are sensitive in the blue–green wavelength region. Instead, we might use photorefractive materials which are sensitive in the red or infrared region, and then use well–developed laser diode technology. Our answer is that there is a whole lot more money and effort going into the move of laser diodes to *lower* wavelengths than there is going into the move of photorefractives to *higher* wavelengths. Even though each photorefractive material has different disadvantages, our overriding concern remains material dynamic range. So we should work with the photorefractive material which comes closest to the ideal behavior for us, and select the rest of our technology accordingly. In this thesis, we want to use crystals that are larger than $5mm^3$ in size, so $LiNbO_3$:Fe is our choice. If a new material is developed which works well in the red, we should of course switch to it; not because it works in the red, but because it works *better*. Given the rapid advances in compact laser sources and the slow advance in new photorefractive materials for storage, we should not give away material performance in exchange for a particular wavelength range.

7.9.2 Micromirrors

Since we have decided to work with a scanning point source and only one lens, it is difficult to perform angle multiplexing with the angle scanning of one micromirror. One way to use the micromirrors would be to set many of the devices in a line, and direct the beam across the wafer surface a millimeter or so above the surface. The beam would pass over micromirrors which are in the closed state and be deflected by the first open micromirror, through the Fourier transform lens to the crystal.

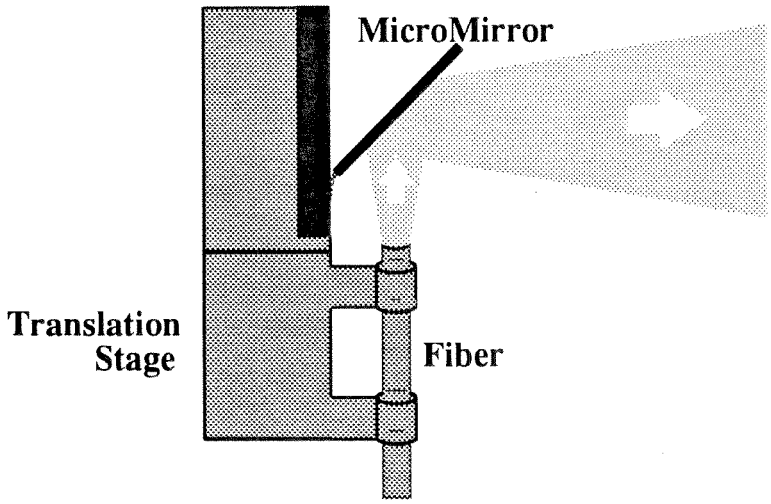


Figure 7.33: Deflection of the diverging output of the fiber tip by a micromirror. The deflection of the micromirror implements spatial multiplexing.

The problem with this is that we would have only coarse control over the eventual incidence angle, because we are limited in how closely we can space the micromirrors. We would have fine control over the position of the reference beam, through the deflection angle of the micromirror.

Instead of angle multiplexing, though, we could use the micromirror to perform spatial multiplexing. By deflecting the diverging beam soon after it leaves the point source, the micromirror plays the part that the tilt of the mirror strips played in the mirror array architecture. This is shown for a fiber tip in Figure 7.33. Linear translation of the point source with the micromirror becomes complicated. Since the micromirror contains a spring and is not being mechanically driven against a stop, we probably don't want to be moving the micromirror or its support. We can still perform the Bragg tuning with a piezoelectric crystal, however, by moving the point source independently of the micromirror. Since the micromirror is 4 mm wide, 100μm of movement should not cause too much of a problem.

If we do choose to move the micromirror, then we need to wait for the vibration of the micromirror to die down. Active damping of the micromirror is possible through adaptive feedback—the electromagnet is driven with a signal which keeps the micromirror pointed in the correct direction. Also, passive damping is a possibility in

future micromirror designs, either in more complicated mechanical spring designs or in the patterning of the permalloy structure. The latter idea is that the eddy currents induced by the movement of the flap through the magnetic field might possibly be used to damp the vibration.

7.9.3 Linear actuators

The simplest option for translating our scanning point source is a linear actuator. One might ask why this is an option now, when it was too slow in Chapters 5 and 6? The main difference is the amount of mass to be moved. In our experimental setup, the vertical stage moved a lens and mirror and associated mounting hardware, while the horizontal stage moved the entire vertical stage plus an additional mirror. Now we are talking about stages that move 2cm of single–mode fiber, a piezoelectric crystal, and some glue. Even with this radical change in the mass involved, we still need to be concerned about reducing the access time well below 1 millisecond.

If we are planning on having two orthogonal stages, one will have to carry the weight of the other. However, since we are not planning on doing many fractal rows, we can have one fiber tip for each fractal row and have just one stage for horizontal deflection, which moves the whole matrix of fiber tips. Piezoelectric crystals provide both horizontal and vertical fine tuning. Most likely, the speed and acceleration of the stage will limit the size of the horizontal deflection which can be performed within the required access time. For instance, in order to perform a 100μm movement within 1 millisecond, the stage must accelerate and decelerate at 40$\times$ gravity, reaching a top speed of 200 mm/sec. Fortunately, the speed of the stage is mostly independent of the load mass (until the point at which the load mass is comparable to the mass of the stage itself). As a result, we can use multiple fiber tips in the horizontal dimension as well, spaced apart by the maximum rapid deflection. At this point, we have replaced the problem of moving a fiber tip on a stage with a limited maximum velocity with the problem of efficiently and rapidly coupling light into the correct fiber. In addition, we may have to use a separate piezoelectric crystal (or pair, for

horizontal and vertical tuning) for each fiber tip. It will depend on the maximum load mass that the piezoelectric device can drive.

Note that we could replace the phrase "fiber tip" with "laser diode/fiber tip pair." We will always need the fiber tip, because the recording process requires light from the same source as the object beam. We can reduce the weight on our stage, however, by placing the laser diode off the translation stage, coupling the light into the same fiber used for carrying the light for recording. This might be especially attractive if the linear stage must also carry the weight of some DOEs for beam shaping.

Another solution for spatial multiplexing of the beam is to mount the fiber tip on a goniometer. This allows the rotation to occur about the fiber tip itself. The diverging cone of light will appear to be deflected at the point source itself, just as it did when light was reflected off the different mirror array facets. If the goniometer is not practical, then the center of rotation can be somewhere else along the fiber. Since rotation of the fiber now translates the fiber tip vertically at the same time, the fiber tip can be translated vertically to compensate and make the center of rotation appear to be at the fiber tip.

7.9.4 Liquid crystal beam steerers

Although the linear actuators are a feasible if clunky solution, liquid crystal beam-steerers appear to be the best solution to our deflection problems. We can place the beamsteerer against the surface of the crystal, deflecting the beam just before it enters the crystal. Why do we even need the Fourier transform lens and the scanning point source then? Well, first of all, we need to collimate the beam at some point before illuminating the beam steerer, so the space needed for the collimating lens cannot be removed.

Secondly, by using a piezoelectric crystal for our fine tuning, we substantially reduce the demands on our beam steering element. For instance, let's say we're going to use 3 fractal rows of 1000 holograms each, spaced by the 5^{th} null.[15] This is 9μm

[15]With a 2cm interaction length, for instance.

of movement in the back focal plane of a 70mm focal length lens. A piezoelectric tuner with an 100μm tuning range can perform fine tuning over 10 holograms. The maximum angle deflection (from the normal) that we require from our liquid crystal beam steerer is 2500 nulls, or 3.5°. This implies that each 2π cycle on our device occupies $8\mu m$. The presence of our piezoelectric device means that the resolution we need at this angle is only .08°. Our two equations are that

$$\frac{\lambda}{qb} = 3.5° \, (= .061\text{radians}) \tag{7.12}$$

and that

$$\frac{\lambda}{b}\left(\frac{1}{q-1} - \frac{1}{q}\right) < .08° \, (= .0014\text{radians}), \tag{7.13}$$

and our two unknowns are b—the periodicity of the electrodes—and q—the number of electrodes per period used for our maximum desired angle deflection.[16] The solution requires that more than 44 electrodes be available within each 8μm period, or lithography of better than 0.1μm. To avoid this problem, we can use 5 or so fiber tips, which would decrease the required number of electrodes to 9 (or a 0.5μm linewidth in the lithography). We will be able to decrease the amount of power in the higher (unwanted) diffracted orders by decreasing the electrode spacing to the smallest practical value (allowing us to increase q). The tradeoff here is between crosstalk to higher diffracted orders, and the complexity of the lithography and the number of I/O connections to the beamsteering device. Another option to reduce the demand on lithography is to cascade two or more beamsteerers of similar spacing.

Since we are only going to use three fractal rows, the limited amount of vertical deflection probably is not worth an entire liquid crystal steerer. An additional consideration is that the large deflection angles we would need are difficult to obtain with the liquid crystal beamsteerer. Instead, as we mentioned above, we can use several scanning fiber tips, either with common piezoelectric fine tuners or with individual

[16] We assume $\lambda = 488$nm.

fine tuning.

Since we are aiming at submillisecond access times, we are implicitly assuming that the liquid crystal devices will conform to this specification. This implies devices that are fabricated with ferroelectric liquid crystals. At this time, most research in liquid crystal beam steerers has been with nematic liquid crystal devices. [392,393] However, most devices have been designed to do the entire angle deflection task themselves, including fine tuning. Since we are implementing this task with a liquid crystal device combined with a piezoelectric fine tuner, our demands on the beamsteerer are more moderate. Although this should help, there is no guarantee that reliable and inexpensive transmissive ferroelectric beamsteerers are feasible. If so, then nematic devices can be used and we live with the best access times that they can give us.

The final consideration is spatial multiplexing. So far, we have considered solutions with a micromirror or goniometer. We might like to use a liquid crystal beamsteerer, though, to perform the spatial multiplexing. Although we cannot place one right at the point source, we could place a beamsteering element at some small distance in front. It should not be too close, or the diverging beam will only cover a few electrodes on the device. On the other hand, it should not be too far away, or the device may begin to noticeably affect the vertical incidence angle at the crystal. If we assume that we can use 0.2μm linewidths and that we don't want to use fewer than 7 electrodes per 2π period of the phase ramp, then the largest deflection angle from one device is 10°. This might certainly be a place where several cascaded devices could be used.

Chapter 8 Next–generation system architectures

Contents

In this thesis, we have demonstrated

- $M/\#$ improvement through oxidation/reduction treatment (Chapters 2–4)
- Storage of 10,000 holograms in a single storage location (Chapter 6), and
- Storage at multiple locations in a system which does not sacrifice rapid access (Chapters 5–6).

What do we need to add to this before we have a useful system? Well, we would like to

1. increase the capacity to the terabit level,
2. demonstrate high readout rate performance, and
3. incorporate the systems solutions that we need to ensure reliable long term storage despite the limitations of our material.

Our approach is to address these issues as we design a next–generation, large–scale, high–speed, angle–multiplexed holographic memory system. We build on the

techniques that we experimentally demonstrated in the first 6 chapters of the thesis. Having surveyed the available actuation technologies and system techniques in Chapter 7, we choose among them in this chapter to reach our goals of large capacity, rapid readout, and reliable storage.

We begin by designing our solution to the volatility problem, drawing from our discussions in Chapter 7.4. Then we consider the possibility of additional improvement in the dynamic range performance, through material and system improvements. Capacity and readout rate are both affected by the number of bits per image, so we will consider how to guarantee the delivery of high–bandwidth reconstructions to the detector array. Finally, incorporating our beam scanning discussions from Chapter 7.5, we propose several next–generation system architectures.

8.1 The volatility problem

The solution we propose for non–volatile readout is a combination of fixing and electrical copying. Our memory consists of several modules, shown as boxes in the top row of Figure 8.1. Each module contains an input SLM, some volume of storage material, a CCD camera for readout, and appropriate routing optics. The usable capacity of the memory uses about 90% of the modules (the shaded box), in which data has been recorded and thermally fixed. The remaining 10% of the modules are intentionally left empty, for purposes we will shortly describe. In later sections, we will assume that we will lose some fraction of the diffraction efficiency during the fixing process, in exchange for a significant increase in erasure time during illumination [123].

One of the knocks against thermal fixing is that all of the data within the crystal suffers the same fate—it is inconvenient to erase or modify one of the stored holograms in the crystal without destroying the other fixed holograms. However, we have set up our memory so that there is random access to any of the locations. Therefore, if we want to change one of the stored pages, we can simply write the new data in a different location and reposition the address "pointer" so that the new location is accessed. This process is shown in the second row of Figure 8.1. Since the access

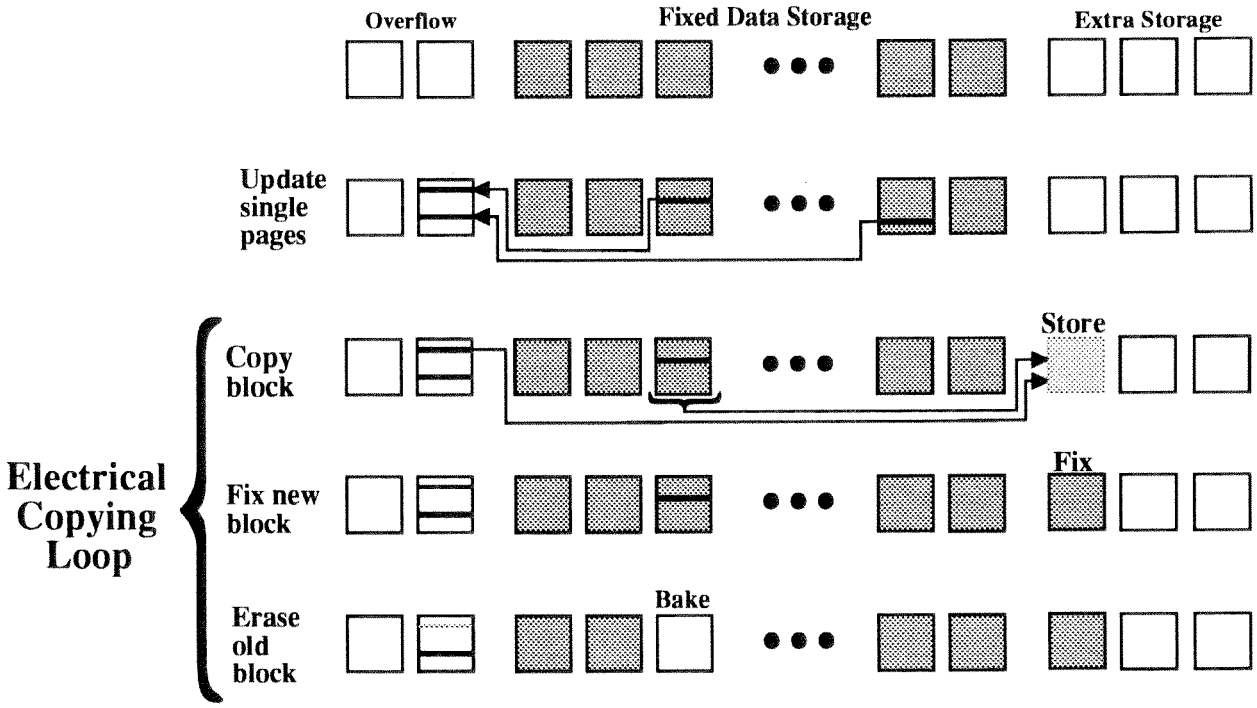


Figure 8.1: Proposed system solution to the volatility problem, using both fixing and periodic electrical copying. Shown are the operations of adding additional holograms, and of recopying a block of data to a fresh location.

time to any hologram is the same, the change in the physical location of the data is invisible to the user. A nice side effect is that the data can be immediately "unerased" by rewriting the address pointer.

If too many of the stored holograms become replaced like this, we can recopy the data into an empty module. This might also be something that we would want to do periodically anyway, either because gradual photovoltaic buildup degrades the SNR or because the lifetime of the fixed gratings is not infinite. The three parts of the copying process are shown in the bottom of Figure 8.1. First, the data is copied to the new location by electronic copying: each hologram is reconstructed and detected, the data displayed on the appropriate SLM and then recorded into the new module. Since the readout process is much faster than the recording process, readout cycles for the user can continue mostly unabated. After all of the data has been recollected and recopied into the new location and verified against the old data, the new module is then thermally fixed. This takes 5-20 minutes of heating, followed by cooldown

and then illumination to erase the electronic grating. A long period of illumination (10–30 minutes or more) may be required—however, the data is still available to the user in the old location. After the electronic grating is completely erased and the ionic grating revealed in the new module, the data pointers are redirected to the new module. At this point, the old module can be thermally erased to remove its ionic gratings, redistribute the filled traps, and prepare the crystal for new data.

The time required between copying cycles depends on the original diffraction efficiency ($M/\#$), the efficiency of the fixing operation, how fast the fixed holograms decay and/or degrade, and the threshold at which data begins to be lost through incorrect detection. Obviously, we are best off with good $M/\#$, efficient fixing, long term fixing, and low noise. However, we are only in trouble if the time between copying cycles is comparable with the length of the recopying procedure. We can anticipate that, with $\sim$100mW/cm^2 of incident power during recording, the erasure time constant will be 200–300 seconds. The entire duration of the recording schedule will be 2–3 $\times$ this time constant [1] plus settling and SLM load time, say 15–20 minutes. We need to heat the crystal for 5–20 minutes to fix the holograms (depending on temperature), and then illuminate for >10 time constants to erase the electronic gratings. So the total duration of the copying, from the reconstruction of the old holograms until the final movement of the data pointers to the new location, can be as little as 90 minutes. And this can be occurring simultaneously at several places in our memory. If we have 90 filled locations and 10 empty locations, we can recycle all of the data in <14 hours. So we really only require that the lifetime of the fixed gratings be longer than this. Of course, we would much prefer to recycle the data once a week or even once a month. This allows us to shut down the SLMs and recording lasers most of the time to save power.

A few minor notes before we move on. The overflow locations, in which we write single holograms to implement page update, must be built in a way that we can read the hologram immediately after recording. This may mean that we make double

[1] This total duration depends on the factor f from Chapter 2.2.1 and is independent of the number of holograms stored. See Chapter 2 for more details.

copies of these holograms in separate blocks, and then fix one of these blocks before it is completely filled. Our second consideration is the effect of elevated temperature on the liquid crystal beamsteerers. Assuming that the effects of different thermal expansion coefficients are dealt with, the liquid crystal material will undergo a phase change as the temperature is elevated. In order to return to the correct alignment, the material must be cooled in a careful schedule through the various phases [314]. Another possible solution is to place the beamsteerer farther from the crystal and surround both with thermally insulating material. In addition, the beamsteerer can be cooled with Peltier cooling while the crystal is heated with contact electrodes.[2]

8.1.1 Photovoltaic field buildup

The photovoltaic effect in $LiNbo_3$ not only assists the formation of the charge gratings, but also separates charges to build up a macroscopic field across the illuminated region. The drift contribution of this field opposes the charge transport of the photovoltaic effect.[3] This has the tendency to reduce the recording slope as we continue to expose the crystal by recording holograms. This is not a serious problem—we can always plan accordingly by increasing the scheduled exposure times gradually.

The real problem is that the presence of absorption causes this macroscopic field to have its own spatial profile. Since the local intensity is spatially dependent, the amount of macroscopic charge transport and the ensuing bulk field are also spatially varying. This creates a gradient index change through the electro–optic effect, a gradient which is increasing in time as the crystal continues to be exposed. What happens when we try to read out one of the first holograms we recorded? Well, the bulk index was uniform when the grating was first recorded, implying that the recording reference beam refracted upon entry to the crystal, but was otherwise unperturbed. Now, however, the reference beam refracts continually as it travels through the crystal. This means that we can only Bragg–match the grating during part of the travel of

[2] A frightful waste of power, but possibly necessary.

[3] If this field did not oppose the photovoltaic effect, we would probably be referring to these crystals as light–induced explosives.

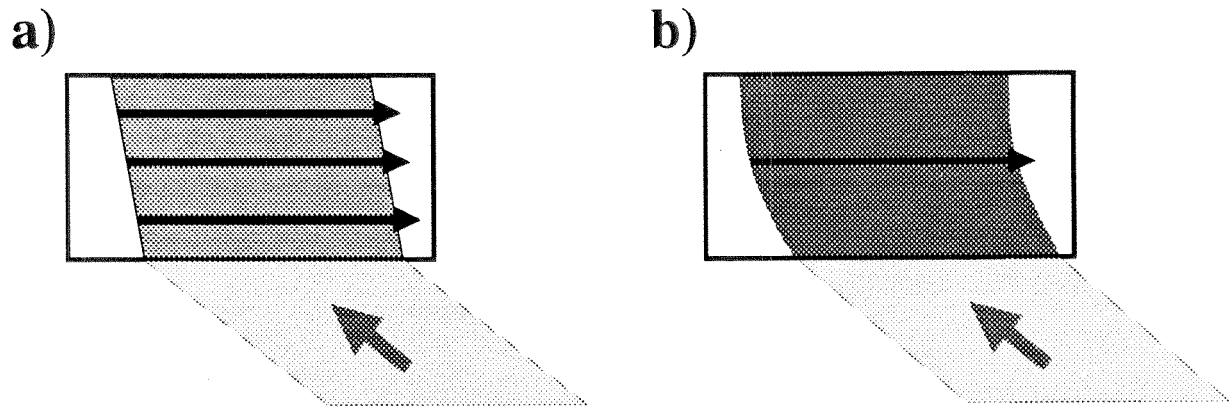


Figure 8.2: Effect of gradient index profile induced by the bulk photovoltaic in the presence of absorption. a) Reference beams used to store holograms early in the schedule refract only at the crystal surface. b) After extended illumination, reference beams refract continuously as they pass through the crystal, allowing perfect Bragg–match only in vertical slices.

the reference beam through the crystal, as shown in Figure 8.2. In the 90° geometry, the portions which get Bragg–matched consist of slices which are parallel to the original signal beam. If the hologram was recorded in the image plane, only a vertical slice of the hologram can be Bragg–matched with any one reference beam. As we tune the angle of the reference beam back–and–forth, the horizontal position of this slice moves back and forth. We observed this effect when we recorded 10,000 image plane holograms (Chapter 6.3.1). In Fourier plane storage, this selective Bragg–matching effect becomes a sliding spatial bandpass filter and is not so unambiguously obvious. We *have* observed that the angular selectivity is often broadened when thousands of weak holograms are recorded over a long total exposure time—however, we cannot attribute all of this to gradient index refraction of the reference beam induced by the photovoltaic effect.

This buildup of macroscopic field and the corresponding index gradient continues to occur while we read out holograms after recording. We have several options, to minimize this problem.

- *Fixing*—Since fixing is a compensatory process, the sign of the macroscopic field is reversed in the ionic copy of the total charge distribution. However, our

usual assumptions about effective compensation by the mobile ions, and about complete dispersal of the electronic charges upon exposure may or may not be valid for the macroscopic charge distribution.

- *Short the crystal faces*—After the holograms are recorded, we can electrically connect the crystal faces in an attempt to drive some of the photovoltaically driven charges to the other side of the crystal. The effectiveness may depend on how good an ohmic contact we can make. We can apply voltage across the crystal as well.

- *Pre–illuminate/Pre–fix the crystal*—We might consider setting up an initial macroscopic field in the crystal by pre–illumination. The charges will be trapped at the edge of the illuminated region. Since this occurs before recording, we have a fair amount of flexibility in the size and shape of the illuminated region. We can pick the direction of the gradient in the field by choosing which face to illuminate. If we choose, we can change the direction of the field after illumination by fixing and then reilluminating with a wider beam to disperse the compensating electronic charges.

8.2 More dynamic range

Our dynamic range target is the $\sim 3 \times 10^{-6}$ diffraction efficiency that we said we needed in Chapter 7.1.3. This is what we need to get our 1000 photons per ON detector pixel. We have already considered the optimization of our $LiNbO_3$ crystals in terms of oxidation state (or absorption coefficient). However, there remain a few more things that we can think about in terms of getting to our target diffraction efficiency.

The first thing we can do is back off on the number of holograms per storage location. Of course, this is not a gain in dynamic range so much as a reassignment to meet our photon budget. We saw in Chapter 4.2.4 that we could maximize the number of holograms per unit volume by using lots of storage locations with a smaller

number of holograms per location. A reasonable target number of holograms per location might be 2000–3000.

In addition to oxidation state, we can think about several other parameters of the $LiNbO_3$ crystal.[4] These might include iron doping level, the photovoltaic constant as set by the poling procedure, the stoichiometry of the crystal, and the concentration of the compensating H^+ ions which will eventually store the fixed holograms.

The tradeoffs are between better dynamic range and the various side effects of overdoing the particular material parameter. With doping level, there is the worry that the dark conductivity will start to become sizable at high doping levels. The relationship between the poling of the crystal and the photovoltaic effect is a hypothetical assumption, based on the observation that the PV effect depends on the asymmetry of the crystal lattice which is actually introduced by the poling. It stands to reason that the efficiency with which the crystal is poled can effect the photovoltaic constant, and certainly explains the wide variance of reported values in the literature (which we noted in Chapter 4.1.7). We might not be happy with photovoltaic constants which are overly large, because the distortions we mentioned in the previous section (gradients in bulk index changing in time as we continue to reconstruct stored holograms) would get more noticeable.

The effect of the stoichiometry of the $LiNbO_3$ crystal has not been extensively studied with holograms. Early reports are that the steady–state photogenerated change in index of refraction rises as the crystal stoichiometry moves from the congruent composition towards 50:50 composition [172]. However, a careful study of the effects of stoichiometry has to be separated from any simultaneous changes in iron doping (because of changing inclusion rates during growth), or photovoltaic constant (because of changes in poling effectiveness or local lattice behavior). If, after these effects have been separated, there is an improvement in dynamic range which is roughly monotonic with change in stoichiometry away from the congruent composition, then the tradeoff is between $M/\#$ and the optical quality/ease of manufacture of the crys-

[4]Again, the plan is to stick with $LiNbO_3$ until something obviously better comes along. A good rule of thumb is to look at which material is being used by system demonstrators to show the "state–of–the–art" in holographic storage.

tals. Noncongruent $LiNbO_3$ is difficult to grow in large size because the melt changes as the boule is pulled from the melt. The resulting strain along the pulling direction tends to lead to cracking and dislocations, and would be expected to get worse as the stoichiometry is moved away from congruency. There are double crucible methods which are designed to re–balance the ratio of Li_2O and Nb_2O_5 in the melt as it is depleted, but these will tend to add to the complexity/cost of the crystal fabrication.

The final consideration is that we are not really planning on using these holograms much until they are thermally fixed. So our $M/\#$ should be measured from the fixed holograms, including fixing efficiency. The optimization of the fixing procedure might include time, temperature, and ion concentration as control variables, and diffraction efficiency and fixed lifetime as variables for maximization. Since we are planning on periodic electronic copying anyway, a fixed lifetime of several months is probably enough for our purposes. Then we can bend our efforts in the fixing process towards increased dynamic range. There is also the possibility that the efficiency of the fixing process will be affected by doping level and absorption coefficient. For instance, a grating with an external diffraction efficiency of, say, 0.1% has much larger internal electric fields in a crystal of $2cm^{-1}$ absorption than it would in a crystal of $0.2cm^{-1}$. This will affect how well the ionic compensation process works.

8.3 Delivering the object beam

We saw in our experiments in Chapter 6.4 that we had to go to a lot of trouble to find the location of the pixel regions, in our analysis of the error performance of our holograms. This occurred because of aberrations encountered by signal beams going to the outer locations, but there was also a small amount of translation of images on the detector even between holograms stored at the same location.

In a practical system, we do not have the luxury of letting the detector pixels move around to find the appropriate SLM image pixels. We would really like to be operating in a regime where the spacing between the detector pixels are the same as (or an integral multiple of) that between the image of the SLM pixels. We can

deal with some mismatch of the magnification of SLM onto detector as long as the mismatch never changes. However, we will have detector pixels which are permanently redundant and extra processing steps required to decode the data.

In this sense, we need a way to return the object beam to the state that it had when it left the SLM. We can do this by reading out our stored holograms with a phase–conjugate reference beam. The signal beam which is readout is S^* instead of S, and it retraces the original path that it used to get to the crystal. Any phase distortions that the beam passes through are effectively undone, including lens aberrations.

Sounds great. Why isn't everyone already using this? Well, for one, the system is now very difficult to align—we can't even see whether the image is in focus until we make a hologram. We might think about placing a mirror in the exact Fourier transform plane so the image is retroreflected for alignment. Another possibility is to use a real–time four–wave phase–conjugate setup to align the array detector, and then place the storage crystal.

Now that the detector is aligned in the right place, we have to provide a phase–conjugate readout beam. This can be done in two ways: we can bring a second reference beam into what was the back of the crystal and then place the detector back where the SLM was. Alternately, we can rotate the crystal by 180° and then use the same reference arm to provide the reference beam. In this latter system, the vertical and horizontal incidence angle of the reference beam must flip over the optical axis of the reference arm, and the detector placed after a second lens, so that the system looks much like the ones we have built so far.

But wait. If the rotation axis of the crystal is not exactly at the Fourier transform plane, then the two lenses in the object beam will not be exactly 4F apart. While this is fine, how are we going to align them? One way is to place the rotation stage of the crystal where you will want it and then store two holograms: one, an image of all ON pixels, and the second, an alignment target or crosshair pattern. When you rotate the crystal and tweak the reference beam to read out the first hologram (not an simple task in itself, but we have tried this at Caltech and it can be done), you can then place this second lens to make the reconstructed wavefront into a plane

wave. The second hologram allows one to align and focus the array detector. The advantage of the rotating crystal method is that we only need to build one reference arm—the disadvantage is that we cannot quickly check that holograms are getting recorded correctly until the entire sequence is recorded.[5]

With real–time phase–conjugation, phase distortions can be totally removed because they are passed twice in rapid succession. With holographic storage, the second pass through a phase distortion occurs upon readout, which may be long after the original storage of the hologram. If the phase distortion has changed in the meantime, the image fidelity of the reconstruction at the original SLM plane (where the detector array is waiting) may suffer. In particular, we should worry about the region of the crystal between the face through which the signal enters and where it is stored in the crystal. As the crystal continues to be used to store and reconstruct holograms, the index profile of the crystal in this region is changing. This may have a profound effect on the ability of the system to reconstruct phase–conjugate replicas of the original signal beam, especially for holograms stored early on in the recording process.

8.4 Proposed architectures

The system we propose here is a combination of some of the ideas which we presented in the previous chapter, and in the previous sections of this chapter. The first version, using a conventional transmission mode SLM, is shown in Figure 8.3. The object arm contains a fiber tip on a translation stage, which allows us to illuminate the SLM at various angles. A polarizer is placed very close to the surface of the SLM, and the surface above each SLM is etched to either 0 or π phase depth to form a random phase plate in contact with the SLM. The modulated signal beam passes through an object beam lens of between 10–20cm focal length. In the Fourier transform plane is placed a slit which passes only the desired SLM order into the desired crystal location. All

[5] I should note that readout during the operation of the recording schedule will give a lot more reconstructed signal power than is expected after all holograms are stored. This is especially true for holograms written early in the sequence. In order for the camera to be able to detect these holograms without saturating, it may need more dynamic range or the ability to shorten exposure times.

the other slits are closed during storage at this location. If these only need to be shut during recording, this can be done with mechanical blade–type shutters. If we would like to open and close them during readout from various locations, we need to use liquid crystal shutters in combination with polarizers.

The reference beam is also introduced by a fiber tip, on a piezoelectric crystal which allows a small amount of translation in this plane for tuning the Bragg angle. A micromirror or second piezoelectric crystal rotates the tip of the fiber so that the exit angle can be controlled. After passing through a collimating lens, the plane wave illuminates a liquid crystal beam steerer at the crystal which controls the coarse reference beam angle. Alternatively, some of the other options mentioned in Chapter 7.9 could be used. Note that a set of reference beams is available on both sides of the crystal, while a second object arm is partially shown introducing light into the right-hand crystal (marked B_2). The beams are shown recording holograms in crystals B_1 and B_2. For readout, the reference beams swap crystals, enabling phase conjugate readout.

Upon readout, the phase–conjugate reconstruction continues back along the path it took on the way in, returning it to the SLM. We deflect approximately 1% of the beam to use for diffraction efficiency measurements when tuning the Bragg angle with the piezoelectric crystal. The remainder of the light passes back through the object beam lens and illuminates the detector. In Figure 8.3, we have drawn a system where a mechanical stage must translate the SLM and detector array to switch between recording and readout. However, alternative schemes are possible, as shown in Figure 8.4. Here we show the use of a beamsplitter to direct light either from the SLM to the crystal, or from the crystal to the array detector. The advantage here is that the system can be aligned once and there is no need for any translation of the array detector once the system is made. An alternate possibility is the integrated SLM/detector array that we discussed in Chapter 7.8.2 and showed in Figure 7.31. [327]

To increase the storage capacity of the system, we have added crystals within the scope of the same object beam. This increases the number of storage locations

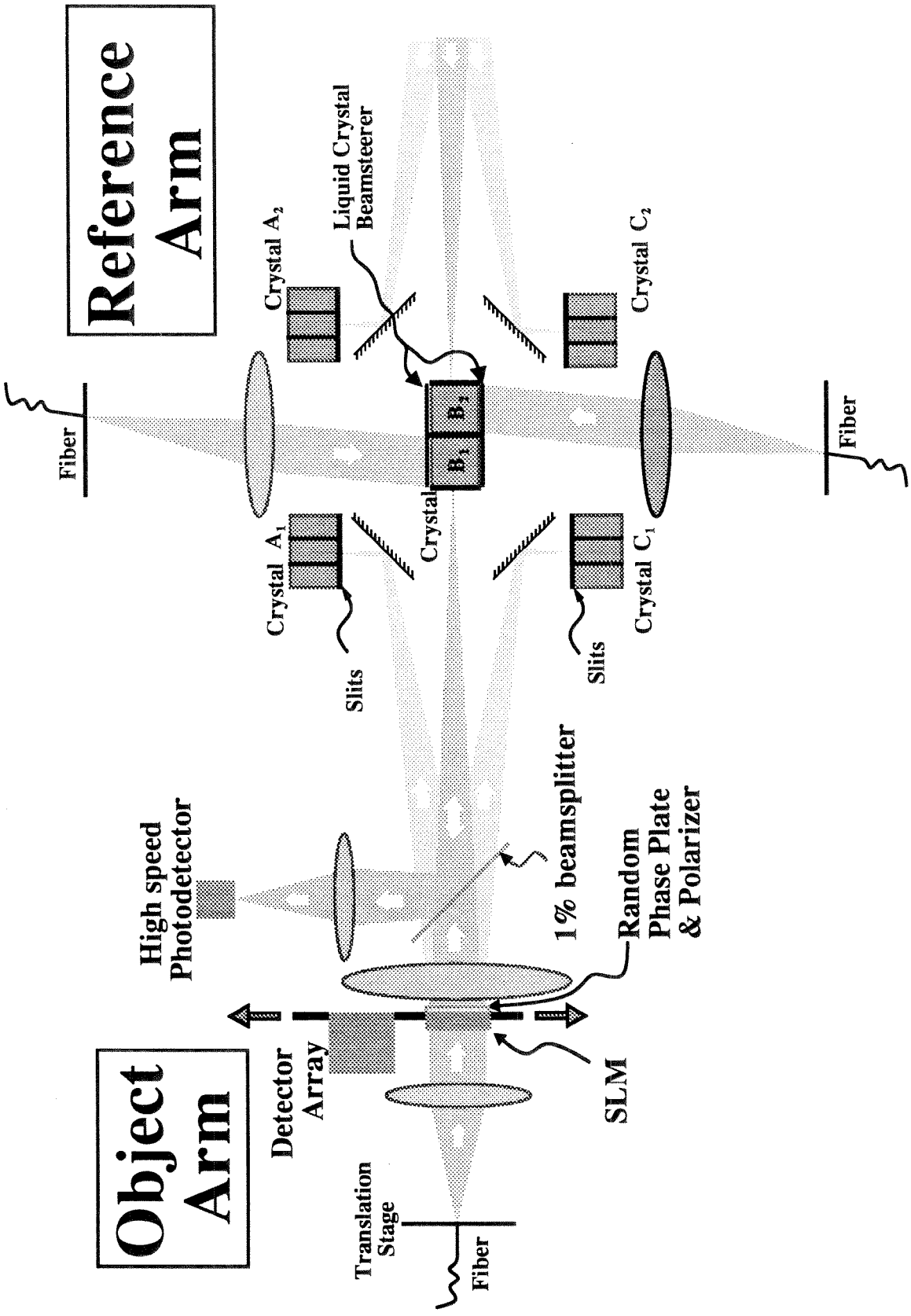


Figure 8.3: Prototype system using separate SLM and detector array.

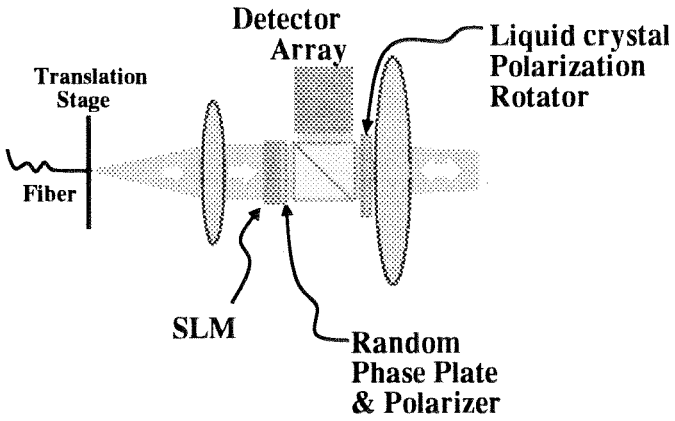


Figure 8.4: Alternate method for using separate SLM and detector array.

accessible to the same SLM and detector array. However, for each stack of crystals, we need to add another set of two reference beams: one for recording and one for readout. Although there is no room left in the horizontal plane, there is nothing that says that the plane of intersection of the beams has to lie in this plane. We can add reference beams which strike these crystals normal to the page. In Figure 8.5, we show what you would see if you looked at the set of storage locations from where the 1% beamsplitter is sitting. The stack of crystals labeled B corresponds to the center stack, and can be very tall since the phase–conjugate readout will return the reconstructions to the optical axis. The arrows denote the various reference beams striking from the side. Flanking this set of storage locations are two mirrors, in which the other two crystal stacks appear. From the point of view of the object beam lens and the deflection of the SLM, the signal beam needs to be deflected to the surface of this mirror in order to reach these outside storage locations. By placing the mirrors closer to the object beam lens, we can ensure that the Fourier transform plane corresponds to the front surfaces of the crystals where the slits are located. Note that the arrows striking the A and C locations arrive vertically.

Let's get an idea of the capabilities of this system, and also the volume required. If the SLM has 20μm pixels and the object beam focal length is $\sim$15cm in length, then each storage location needs to be a little bigger than 3mm or so. From our treatment in Chapter 4.2.4, we can expect to conservatively store 100,000 holograms

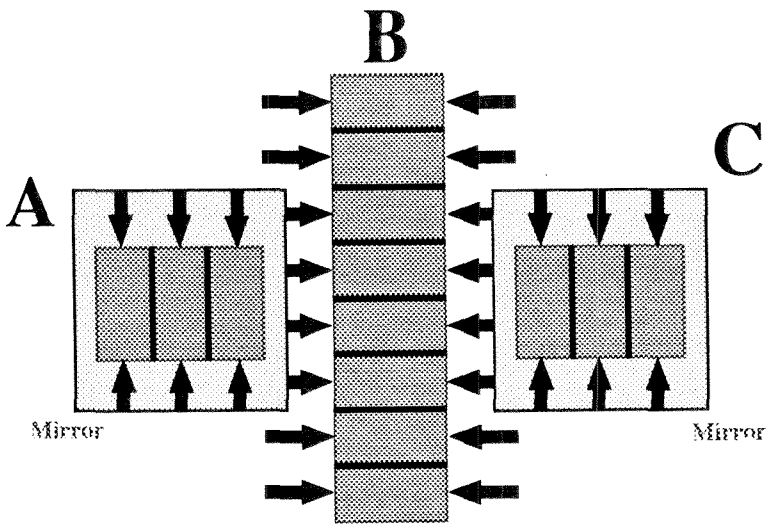


Figure 8.5: View of the crystal locations as seen from the Fourier transform lens. Arrows show the incidence of the reference beams.

per 4cm height of crystal and still meet our photon budget from Chapter 7.1.3. Let's assume that the B stack is 6cm tall, and A and C both 3cm, so that we can have 300,000 holograms stored per SLM. A conservative value for the number of pixels per hologram might be 300,000, and let's assume that we can get the readout rate up to 1000 frames per second.

Now we need the volume of each of these storage "units." Let's say that the object beam needs a 10×10 cm^2 cross section and is 25cm in length, from fiber tip to crystal. For the reference beam volume, we need a cross section of 8cm $\times$ 8cm and a length (beyond the object beam tube) of 12cm. There are four of these vertically, and 1 horizontally (the second is shared with the neighboring "unit," as we discussed above). In addition, we should include the volume where crystal stacks A and C are, about 7cm $\times$ 7cm $\times$ 10cm each. The total volume of the storage "unit," containing 1.5×10^{11} bits, is 7500cm^3. This unit is shaped like a 3–dimensional plus sign, which might make stacking them up difficult. However, we might envision different ways of getting crystal stacks A and C and their reference beams into the horizontal plane by using multiple mirrors to direct the object beam away from crystal stack B.

Let's say that we have 100 of these units together, of which 90 contain data and 10 are empty so that we can copy data into them as the fixed holograms degrade. We

can round the system volume up so as to include control electronics and the volume of the various lasers and switches needed at the other end of all those fibers. We estimate the following performance capabilities:

Total capacity:	13.5 Terabits
Total readout rate:	42.5 Gigabits/second
System volume	1 m^3

On the plus side, these numbers were made with conservative estimates on the number of pixels and the number of holograms per location. Fortunately, our design allows us to use relatively inexpensive optics and reduces the number of lasers for recording by using switched fibers. On the minus side, though, we have purchased 100 SLMs and 100 detector arrays, so the total cost is going to depend a lot on how much these cost. And there is the question of how many hydroelectric power plants are required for each of these systems...

8.5 Conclusion

What needs to be researched before a system with the desired capabilities can be seriously considered? I think the areas requiring focus parallel the sections in this chapter:

- *Fidelity of fixed holograms*—Research on fixing has concentrated on the lifetime of fixed holograms. For thermal fixing in $LiNbO_3$, this lifetime is already larger than what would be needed for a system like the one described here. Instead, the focus should turn to the effects of further illumination on the fidelity and error performance of the holograms.

- *More dynamic range*—In this thesis, we discussed increasing dynamic range by optimizing the oxidation state of $LiNbO_3$. In addition to this, there is the Fe doping level, the poling procedure, the stoichiometry of the crystal, and the efficiency of the fixing procedure as items of future research. There may be a tradeoff on the strength of the photovoltaic effect between dynamic range and

long–term fidelity after recording. As always, one should keep their eyes open for materials (or different dopants for $LiNbO_3$) which offer more (in the way of volatility or dynamic range) without sacrificing what $LiNbO_3$ has going for it (optical quality, ease of fabrication even for large samples, robustness).

- *Phase–conjugate readout*—The holographic storage field is just beginning to realize the difficulty of exact pixel matching. The difficulties of passing a high–resolution image through a birefringent sample and maintaining pixel match over a million pixels requires expensive optics, careful alignment, perfect polishing and high optical quality crystals. And after all that, holograms can only be stored at one storage location per SLM/detector pair. Unless the storage media can be mechanically translated very rapidly, high readout rate and high capacity will remain mutually exclusive goals. A possible solution is phase–conjugate readout, on which study has begun at Caltech [327, 328]. As with the fixing, the worry remains long–term fidelity of holograms as the crystal is illuminated.

- *System design*—The final step is designing systems which meet all of the requirements, and building them with components that can reach the performance goals.

I think that holographic storage can find its place in the world, either in existing applications or in ones which are waiting in the wings. I don't know if any of the ideas contained in this thesis will be involved in a practical realization of holographic storage—but at least I've had fun (and learned a lot) thinking them up!

Bibliography

[1] Rome Air Development Center. Memory capability study.

[2] W. L. Bragg. A new type of 'x–ray microscope'. *Nature*, 143:678, 1939.

[3] E. Hecht. *Optics, 2nd ed.* Addison–Wesley, 1988.

[4] D. Gabor. A new microscopic principle. *Nature*, 161:777, 1948.

[5] T. H. Maiman. Stimulated optical radiation in ruby. *Nature*, 187:493, 1960.

[6] E. N. Leith and J. Upatnieks. Reconstructed wavefronts and communication theory. *Journal of the Optical Society of America*, 52:1123, 1962.

[7] E. N. Leith and J. Upatnieks. Wavefront reconstruction with continuous–tone objects. *Journal of the Optical Society of America*, 53:1377, 1963.

[8] E. N. Leith and J. Upatnieks. Wavefront reconstruction with diffused illumination and three–dimensional objects. *Journal of the Optical Society of America*, 54:1295, 1964.

[9] J. W. Goodman. *Introduction to Fourier Optics.* McGraw–Hill, 1968.

[10] M. Born and E. Wolf. *Principles of Optics, 6th ed.* Pergamon Press, 1980.

[11] Y. N. Denisyuk. Photographic reconstruction of the optical properties of an object in its own scattered radiation field. *Sov. Phys. Dokl.*, 7:543, 1962.

[12] P. J. van Heerden. Theory of optical information storage in solids. *Applied Optics*, 2(4):393–401, 1963.

[13] R. W. Lucky. *Silicon dreams: Information, man, and machine.* St. Martin's Press, 1989.

[14] S. Jackson. *Classical Electrodynamics.* John Wiley & Sons, 1967.

[15] H. Kogelnik. Coupled wave theory for thick hologram gratings. *The Bell System Technical Journal*, 48(9):2909–2947, 1969.

[16] D. L. Staebler and J. J. Amodei. Coupled–wave analysis of holographic storage in $LiNbO_3$. *Journal of Applied Physics*, 43(3):1042–1049, 1972.

[17] F. G. Kaspar. Diffraction by thick, periodically stratified gratings with complex dielectric constant. *Journal of the Optical Society of America* **A**, 63:37–45, 1973.

[18] R. Alferness. Analysis of optical propagation in thick holographic gratings. *Applied Physics*, 7:29–33, 1975.

[19] L. Solymar. A general two–dimensional theory for volume holograms. *Applied Physics Letters*, 31(12):820–822, 1977.

[20] R. Magnusson and T. K. Gaylord. Solutions of the thin phase grating diffraction gratings. *Optics Communication*, 25(2):129–132, 1978.

[21] R. Magnusson and T. K. Gaylord. Diffraction efficiencies of thin phase gratings with arbitrary grating shape. *Journal of the Optical Society of America*, 68(6):806–809, 1978.

[22] D. J. Cooke, L. Solymar, and C. J. R. Sheppard. A three–dimensional vectorial theory for volume holograms. *International Journal of Electronics*, 46(4):337–356, 1979.

[23] R. Magnusson and T. K. Gaylord. Diffraction efficiencies of thin absorption and transmittance gratings. *Optics Communication*, 28(1):1–3, 1979.

[24] L. Solymar and C. J. R. Sheppard. A two–dimensional theory of volume holograms with electric polarization in the plane of the grating. *Journal of the Optical Society of America*, 69(4):491–495, 1979.

[25] P. St. J. Russell and L. Solymar. The properties of holographic overlap gratings. *Optica Acta*, 26(3):329–347, 1979.

[26] B. Benlarbi, D. J. Cooke, and L. Solymar. Higher order modes in thick phase gratings. *Optica Acta*, 27(7):885–895, 1980.

[27] B. Benlarbi and L. Solymar. Higher order modes in non–sinusoidal volume phase holograms. *International Journal of Electronics*, 48(4):351–359, 1980.

[28] M. G. Moharam, T. K. Gaylord, and R. Magnusson. Criteria for Bragg regime diffraction by phase gratings. *Optics Communication*, 32(1):14–18, 1980.

[29] M. G. Moharam, T. K. Gaylord, and R. Magnusson. Diffraction characteristics of three–dimensional crossed–beam volume gratings. *Journal of the Optical Society of America*, 70(4):437–442, 1980.

[30] M. G. Moharam, T. K. Gaylord, and R. Magnusson. Bragg diffraction of finite beams by thick gratings. *Journal of the Optical Society of America*, 70(3):300–304, 1980.

[31] L. Solymar and D. J. Cooke. *Volume holography and volume gratings.* Academic Press, 1981.

[32] R. J. Collier, C. B. Burckhardt, and L. H. Lin. *Optical Holography.* Academic Press, Inc., 1971.

[33] P. St. J. Russell. Optical volume holography. *Physics Reports*, 71(4):209–312, 1981.

[34] A. Ashkin, G. D. Boyd, J. M. Dziedzic, R. G. Smith, A. A. Ballman, J. J. Levinstein, and K. Nassau. Optically–induced refractive index homogeneities in $LiNbO_3$ and $LiTaO_3$. *Applied Physics Letters*, 9(1):72–74, 1966.

[35] J. J. Amodei, W. Phillips, and D. L. Staebler. Improved electrooptic materials for holographic storage applications. *IEEE Journal of Quantum Electronics*, 7(6):321, 1971.

[36] A. Ishida, O. Mikami, S. Miyazawa, and M. Sumi. Rh–doped $LiNbO_3$ as an improved new material for reversible holographic storage. *Applied Physics Letters*, 21(5):192–193, 1971.

[37] O. Mikami. Cu–diffused layers in $LiNbO_3$ for reversible holographic storage. *Optics Communication*, 11(1):30–32, 1974.

[38] E. Okamoto, H. Ikeo, and K. Muto. Holographic storage in U–doped $LiNbO_3$. *Applied Optics*, 14(10):2453–2455, 1975.

[39] G. T. Avanesyan, E. S. Vartanyan, R. S. Mikaelyan, R. K. Hovsepyan, and A. R. Pogosyan. Mechanisms of photochromic and photorefractive effects in doubly doped lithium niobate crystals. *Physica Status Solidi*, 126:245–252, 1991.

[40] K. Sayano, G. A. Rakuljic, A. Agranat, A. Yariv, and R. R. Neurgaonkar. Photorefractive dark conductivity in Cr–doped strontium barium niobate. *Optics Letters*, 14(9):459–461, 1989.

[41] I. R. Dorosh, Y. S. Kuz'minov, V. V. Osiko, and N. V. Tkachenko. Influence of Ce concentration on holographic sensitivity of barium–strontium niobate $(Sr_xBa_{1-x})_{1-y}(Nb_2O_6)_y$ crystals. *Soviet Physics–Solid State*, 23(2):345–346, 1981.

[42] R. R. Neurgaonkar, W. K. Cory, J. R. Oliver, M. D. Ewbank, and W. F. Hall. Development and modification of photorefractive properties in the tungsten bronze family crystals. *Optical Engineering*, 26(5):392–404, 1987.

[43] V. V. Voronov, E. K. Gulanyan, I. R. Dorosh, Y. S. Kuz'minov, A. L. Mikaelyan, V. V. Osiko, N. M. Polozkov, and A. M. Prokhorov. Photoelectric and photorefractive properties of cerium–doped barium strontium niobate crystals. *Soviet Journal of Quantum Electronics*, 9(9):1172–1175, 1979.

[44] G. Chanussot. Inhomogeneous instabilities in an illuminated crystal of doped $BaTiO_3$. *Ferroelectrics*, 13:313–315, 1976.

[45] D. Rytz, B. A. Weschler, M. H. Garret, C. C. Nelson, and R. N. Schwartz. Photorefractive properties of $BaTiO_3$:Co. *Journal of the Optical Society of America* **B**, 7(12):2245–2254, 1990.

[46] K. Buse, L. Holtmann, and E. Kratzig. Activation of $BaTiO_3$ for infrared holographic recording. *Optics Communication*, 85:183–186, 1991.

[47] B. A. Wechsler, M. B. Klein, C. C. Nelson, and R. N. Schwartz. Spectroscopic and photorefractive properties of infra–red sensitive rhodium–doped barium titanate. *Optics Letters*, 19(8):536–538, 1994.

[48] K. Sutter and P. Gunter. Photorefractive gratings in the organic crystal 2–cycloctylamino–5–nitropyridine doped with 7,7,8,8–tetracyanoquinodimethane. *Journal of the Optical Society of America* **B**, 7(12):2274–2278, 1990.

[49] S. Ducharme, J. C. Scott, R. J. Twieg, and W. E. Moerner. Observation of the photorefractive effect in a polymer. *Physics Review Letters*, 66(14):1846–1849, 1991.

[50] Y. P. Cui, Y. Zhang, P. N. Prasad, J. S. Schildkraut, and D. J. Williams. Photorefractive effect in a new organic–system of doped nonlinear polymer. *Applied Physics Letters*, 61(18):2132–2134, 1992.

[51] S. Ducharme, B. Jones, J. M. Takacs, and Z. Lei. Electric–field stabilization and competition of gratings in a photorefractive polymer. *Optics Letters*, 18(2):152–154, 1992.

[52] S. M. Silence, C. A. Walsh, J. C. Scott, and W. E. Moerner. C–60 sensitization of a photorefractive polymer. *Applied Physics Letters*, 61(25):2967–2969, 1992.

[53] S. M. Silence, C. A. Walsh, J. C. Scott, T. J. Matray, and R. J. Tweig. Subsecond grating growth in a photorefractive polymer. *Optics Letters*, 17(16):1107–1109, 1992.

[54] K. Tamura, A. B. Padias, H. K. Hall, and N. Peyghambarian. New polymeric material containing the tricyanovinylcarbazole group for photorefractive applications. *Applied Physics Letters*, 60(15):1803–1805, 1992.

[55] K. A. Verkhovskaya, V. M. Fridkin, A. V. Bune, and J. F. Legrand. Photoconducting ferroelectric polymers. *Ferroelectrics*, 134(1–4):7–15, 1992.

[56] L. P. Yu, W. K. Chan, Z. N. Bao, and S. X. F. Cao. Synthesis and physical measurements of a photorefractive polymer. *Journal of the Chemical Society—Chemical Communications*, 23:1735–1737, 1992.

[57] Y. Zhang, Y. P. Cui, and P. N. Prasad. Observation of photorefractivity in a fullerene–doped polymer composite. *Physical Review B—Condensed Matter*, 46(15):9900–9902, 1992.

[58] C. A. Walsh and W. E. Moerner. Two–beam coupling measurements of grating phase in a photorefractive polymer. *Journal of the Optical Society of America* **B**, 9(9):1642–1647, 1992.

[59] M. C. J. M. Donkers, S. M. Silence, C. A. Walsh, F. Hache, D. M. Burland, W. E. Moerner, and R. J. Tweig. Net two–beam–coupling gain in a polymeric photorefractive material. *Optics Letters*, 18:1044–1046, 1993.

[60] S. Ducharme, B. Jones, J. M. Takacs, and L. Zhang. Electric–field stabilization and competition of gratings in a photorefractive polymer. *Optics Letters*, 18(2):152–154, 1993.

[61] B. Kippelen, K. Tamura, N. Peyghambarian, A. B. Padias, and H. K. Hall. Photorefractive effect in a poled polymer containing the tricyanovinylcarbazole group. *Journal of Applied Physics*, 74(5):3617–3619, 1993.

[62] B. Kippelen, Sandalphon, N. Peyghambarian, S. R. Lyon, A. B. Padias, and H. K. Hall, Jr. New highly efficient photorefractive polymer composite for optical–storage and image–processing applications. *Electronics Letters*, 29:1873–1874, 1993.

[63] M. J. Sansone, C. C. Teng, A. J. East, and M. S. Kwiatek. Observation of the photorefractive effect in a dialkylaminonitrostilbene copolymer. *Optics Letters*, 18(17):1400–1402, 1993.

[64] M. C. J. M. Donckers, S. M. Silence, C. A. Walsh, F. Hache, and D. M. Burland. Net 2–beam–coupling gain in a polymeric photorefractive material. *Optics Letters*, 18(13):1044–1046, 1993.

[65] C. A. Walsh and W. E. Moerner. 2–beam coupling measurements of grating phase in a photorefractive polymer. *Journal of the Optical Society of America* **B**, 10(4):753, 1993.

[66] M. Liphard, A. Goonesekera, B. E. Jones, S. Ducharme, J. M. Takacs, and L. Zhang. High–performance photorefractive polymers. *Science*, 263:367–369, 1994.

[67] K. Meerholz, B. L. Volodin, Sandalphon, B. Kippelen, and N. Peyghambarian. A photorefractive polymer with high optical gain and diffraction efficiency near 100%. *Nature*, 371:497–500, 1994.

[68] W. E. Moerner, S. M. Silence, F. Hache, and G. C. Bjorklund. Orientationally enhanced photorefractive effect in polymers. *Journal of the Optical Society of America* **B**, 11:2132–2134, 1994.

[69] S. M. Silence, G. C. Bjorklund, and W. E. Moerner. Optical trap activation in a photorefractive polymer. *Optics Letters*, 19(22):1822–1824, 1994.

[70] C. Halvorson, B. Kraabel, A. J. Heeger, B. L. Volodin, K. Meerholz, Sandalphon, and N. Peyghambarian. Optical computing by use of photorefractive polymers. *Optics Letters*, 20:76–79, 1995.

[71] B. Kippelen, K. Meerholz, Sandalphon, B. Volodin, and N. Peyghambarian. Nonlinear photorefractive polymers. *Optical materials*, 4:354–357, 1995.

[72] B. Volodin, Sandalphon, K. Meerholz, B. Kippelen, N. V. Kukhtarev, and N. Peyghambarian. Highly efficient photorefractive polymers for dynamic holography. *Optical Engineering*, 34(8):2213–2223, 1995.

[73] F. S. Chen, J. T. LaMacchia, and D. B. Fraser. Holographic storage in lithium niobate. *Applied Physics Letters*, 13(7):223–225, 1968.

[74] F. S. Chen. Optically induced change of refractive indices in $LiNbO_3$ and $LiTaO_3$. *Journal of Applied Physics*, 40(8):3389–3396, 1969.

[75] P. Gunter, U. Fluckiger, J. P. Huignard, and F. Micheron. Optically induced refractive index changes in $KNbO_3$:Fe. *Ferroelectics*, 13:297–299, 1976.

[76] J. B. Thaxter. Electrical control of holographic storage in strontium–barium niobate. *Applied Physics Letters*, 15(7):210–212, 1969.

[77] R. L. Townsend and J. T. LaMacchia. Optically induced refractive index changes in $BaTiO_3$. *Journal of Applied Physics*, 41(13):5188–5192, 1970.

[78] J. P. Huignard and F. Micheron. High sensitivity read–write volume holographic storage in BSO and BGO crystals. *Applied Physics Letters*, 29:591, 1976.

[79] M. B. Klein. Beam coupling in undoped GaAs at 1.06 μm using the photorefractive effect. *Optics Letters*, 9:350–352, 1984.

[80] A. M. Glass, A. M. Johnson, D. H. Olson, W. Simpson, and A. A. Ballman. Four–wave mixing in semi–insulating InP and GaAs using the photorefractive effect. *Applied Physics Letters*, 44:948, 1984.

[81] J. J. Amodei. Analysis of transport processes during holographic recording in insulators. *RCA Review*, 32:185–197, 1972.

[82] J. J. Amodei. Electron diffusion effects during hologram recording in crystals. *Applied Physics Letters*, 18(1):22–24, 1971.

[83] L. Young, W. K. Y. Wong, M. L. W. Thewalt, and W. D. Cornish. Theory of formation of phase holograms in lithium niobate. *Applied Physics Letters*, 24(6):264–265, 1974.

[84] G. A. Alphonse, R. C. Alig, D. L. Staebler, and W. Phillips. Time–dependent characteristics of photo–induced space–charge field and phase holograms in Lithium Niobate and other photorefractive media. *RCA Review*, 36:213–229, 1975.

[85] V. L. Vinetskii and N. V. Kukhtarev. Theory of the conductivity induced by recording holographic gratings in nonmetallic crystals. *Soviet Physics—Solid State*, 16(12):2414–2415, 1975.

[86] D. W. Vahey. A nonlinear coupled–wave theory of holographic storage in ferroelectric materials. *Journal of Applied Physics*, 46(8):3510–3515, 1975.

[87] S. F. Su and T. K. Gaylord. Unified approach to the formation of phase holograms in ferroelectric crystals. *Journal of Applied Physics*, 46(12):5208–5213, 1975.

[88] N. V. Kukhtarev. Kinetics of hologram recording and erasure in electrooptic crystals. *Soviet Technical Physics Letters*, 2(12):438–440, 1976.

[89] M. G. Moharam and L. Young. Hologram writing by the photorefractive effect with gaussian beams at constant applied voltage. *Journal of Applied Physics*, 47(9):4048–4051, 1976.

[90] R. Magnusson and T. K. Gaylord. Use of dynamic theory of describe experimental results from volume holography. *Journal of Applied Physics*, 47(1):190–199, 1976.

[91] D. M. Kim, R. R. Shah, T. A. Rabson, and F. K. Tittel. Nonlinear dynamic theory for photorefractive phase hologram formation. *Applied Physics Letters*, 28(6):338–340, 1976.

[92] K. Blotekjaer. Theory of hologram formation in photorefractive media. *Journal of Applied Physics*, 48(6):2495–2501, 1977.

[93] M. G. Moharam and L. Young. Hologram writing by the photorefractive effect. *Journal of Applied Physics*, 48(8):3230–3236, 1977.

[94] M. G. Moharam and L. Young. Reading and optical erasure of holograms stored by the photorefractive effect in lithium niobate. *Applied Optics*, 17(17):2773–2778, 1978.

[95] N. V. Kukhtarev, V. B. Markov, S. G. Odulov, M. S. Soskin, and V. L. Vinetskii. Holographic storage in electrooptic crystals I. steady state. *Ferroelectrics*, 22:949–960, 1979.

[96] R. Jaura, T. J. Hall, and P. D. Foote. Simplified band transport model of the photorefractive effect. *Optical Engineering*, 25(9):1068–1074, 1986.

[97] G. C. Valley. Erase rates in photorefractive materials with two photoactive species. *Applied Optics*, 22(20):3160–3164, 1983.

[98] F. P. Strohkendl, J. M. C. Jonathan, and R. W. Hellwarth. Hole–electron competition in photorefractive gratings. *Optics Letters*, 11(5):312–314, 1986.

[99] G. C. Valley. Simultaneous electron/hole transport in photorefractive materials. *Journal of Applied Physics*, 59(10):3363–3366, 1986.

[100] M. C. Bashaw, R.-P. Ma, R. C. Barker, S. Mroczkowski, and R. R. Dube. Thoery of complementary holograms arising from electron–hole transport in photorefractive media. *Journal of the Optical Society of America* **B**, 7(12):2329–2338, 1990.

[101] F. Jariego and F. Agullo-Lopez. Holographic writing and erasure in unipolar photorefractive materials with multiple active centers: theoretical analysis. *Applied Optics*, 30(32):4615–4621, 1991.

[102] M. C. Bashaw, T.-P. Ma, and R. C. Barker. Comparison of single– and two–species models of electron–hole transport in photorefractive media. *Journal of the Optical Society of America* **B**, 9(9):1666–1672, 1992.

[103] P. Tayebati. Effect of shallow traps on electron–hole competition in semi–insulating photorefractive materials. *Journal of the Optical Society of America* **B**, 9(3):415–419, 1992.

[104] M. C. Bashaw, J. F. Heanue, A. Aharoni, J. F. Walkup, and L. Hesselink. Cross–talk considerations for angular and phase–encoded multiplexing in volume holography. *Journal of the Optical Society of America* **B**, 11(9):1820–1836, 1994.

[105] P. Yeh, A. E. Chiou, J. Hong, P. Beckwith, T. Chang, and M. Khoshnevisan. Photorefractive nonlinear optics and optical computing. *Optical Engineering*, 28(4):328–343, 1989.

[106] P. Gunter. Holography, coherent light amplification and optical phase conjugation with photorefractive materials. *Physics Reports*, 4:199–299, 1982.

[107] A. Yariv and D. M. Pepper. Amplified reflection, phase conjugation, and oscillation in degenerate four–wave mixing. *Optics Letters*, 1:16–18, 1977.

[108] A. Yariv. Four wave nonlinear optical mixing as real time holography. *Optics Communication*, 25:23–25, 1978.

[109] A. Yariv. *Quantum Electronics, 3^d ed.* John Wiley & Sons, 1989.

[110] S. Mallick, B. Imbert, H. Ducollet, J. P. Herriau, and J. P. Huignard. Generation of spatial subharmonics by two–wave mixing in a nonlinear photorefractive medium. *Journal of Applied Physics*, 63:5660–5663, 1988.

[111] D. C. Jones and L. Solymar. Competition between subharmonic and resonating beams for photorefractive gain in bismuth silicon oxide. *Optics Letters*, 14:743–744, 1989.

[112] M. Segev, B. Crosignani, A. Yariv, and B. Fischer. Spatial solitons in photorefractive media. *Physics Review Letters*, 68(7):923–926, 1992.

[113] R. L. Byer, J. F. Young, and R. S. Feigelson. Growth of high-quality $LiNbO_3$ crystals from the congruent melt. *Journal of Applied Physics*, 41(6):2320–2325, 1970.

[114] Various Authors. *Properties of lithium niobate.* EMIS Datareviews Series No. 5. INSPEC, 1989.

[115] A. M. Glass, D. von der Linde, and T. J. Negram. High–voltage bulk photovoltaic effect and the photorefractive process in $LiNbO_3$. *Applied Physics Letters*, 25(4):233–235, 1974.

[116] G. E. Peterson, A. M. Glass, and T. J. Negran. Control of the susceptibility of lithium niobate to laser–induced refractive index changes. *Applied Physics Letters*, 19(5):130–132, 1971.

[117] W. Phillips, J. J. Amodei, and D. L. Staebler. Optical and holographic storage properties of transition metal doped lithium niobate. *RCA Review*, 33:94–109, 1972.

[118] D. L. Staebler, J. J. Amodei, and W. Phillips. Multiple storage of thick phase holograms in $LiNbO_3$. In *VII International Quantum Electronics Conference.* Joint Council on Quantum Electronics, May 1972. QE–8, pg 611.

[119] W. Phillips and D. L. Staebler. Control of the Fe^{2+} concentration in iron–doped lithium niobate. *Journal of Electronic Materials*, 3(2):601–617, 1974.

[120] D. L. Staebler, W. J. Burke, W. Phillips, and J. J. Amodei. Multiple storage and erasure of fixed holograms in Fe–doped $LiNbO_3$. *Applied Physics Letters*, 26(4):182–184, 1975.

[121] J. J. Amodei and D. L. Staebler. Holographic pattern fixing in electro–optic crystals. *Applied Physics Letters*, 18(12):540–542, 1971.

[122] J. J. Amodei, W. Phillips, and D. L. Staebler. Improved electrooptic materials and fixing techniques for holographic recording. *Applied Optics*, 11(2):390–396, 1972.

[123] D. L. Staebler and J. J. Amodei. Thermally fixed holograms in $LiNbO_3$. *Ferroelectrics*, 3:107–113, 1972.

[124] D. L. Staebler. *Ferroelectric crystals*, volume 20 of *Topics in Applied Physics*, pages 101–132. Springer–Verlag, 1977.

[125] F. Micheron and G. Bismuth. Electrical control of fixation and erasure of holographic patterns in ferroelectric materials. *Applied Physics Letters*, 20(2):79–81, 1972.

[126] F. Micheron and G. Bismuth. Field and time thresholds for the electrical fixation of holograms recorded in $(Sr_{0.75}Ba_{0.25})Nb_2O_6$ crystal. *Applied Physics Letters*, 23(2):71–72, 1972.

[127] G. A. Alphonse and W. Phillips. Iron–doped Lithium Niobate as a read–write holographic storage medium. *RCA Review*, 37:184–205, 1976.

[128] J. P. Huignard, J. P. Herriau, and F. Micheron. Coherent selective erasure of superimposed volume holograms in $LiNbO_3$. *Applied Physics Letters*, 26(5):256–258, 1975.

[129] L. D'Auria, J. P. Huignard, and E. Spitz. Holographic read–write memory and capacity enhancement by 3–D storage. *IEEE Transactions on Magnetics*, MAG–9(2):83–94, 1973.

[130] L. d'Auria, J. P. Huignard, C. Slezak, and E. Spitz. Experimental holographic read–write memory using 3–D storage. *Applied Optics*, 13(4):808–818, 1974.

[131] D. Psaltis. Parallel optical memories. *BYTE*, 17(9):179, 1992.

[132] D. Psaltis and F. Mok. Holographic memories. *Scientific American*, 273(5):70–76, Nov 1995.

[133] E. G. Ramberg. Holographic information storage. *RCA Review*, 33:5–53, 1972.

[134] W. J. Burke and P. Sheng. Crosstalk noise from multiple thick–phase holograms. *Journal of Applied Physics*, 48(2):681–685, 1977.

[135] H. Lee. Cross–talk effects in multiplexed volume holograms. *Optics Letters*, 13(10):874–876, 1988.

[136] H. Lee, X.-G. Gu, and D. Psaltis. Volume holographic interconnections with maximal capacity and minimal crosstalk. *Journal of Applied Physics*, 65(6):2191–2194, 1989.

[137] H. Lee. Volume holographic global and local interconnecting patterns with maximal capacity and minimal first–order crosstalk. *Applied Optics*, 28(24):5312–5316, 1989.

[138] C. Gu, J. Hong, I. McMichael, R. Saxena, and F. Mok. Cross–talk–limited storage capacity of volume holographic memory. *Journal of the Optical Society of America* **A**, 9(11):1–6, 1993.

[139] C. Gu, A. Chiou, and J. Hong. Cross–talk noise in photorefractive interconnection. *Applied Optics*, 32(8):1437–1440, 1993.

[140] K. Curtis, D. Psaltis, and C. Gu. Crosstalk in wavelength multiplexed holographic memories. *Optics Letters*, 18(12):1001–1003, 1993.

[141] K. Curtis and D. Psaltis. Cross talk in phase–coded holographic memories. *Journal of the Optical Society of America* **A**, 10(12):2547–2550, 1993.

[142] G. P. Nordin and P. Asthana. Effects of cross talk on fidelity in page–oriented volume holographic optical data storage. *Optics Letters*, 18(18):1553–1555, 1993.

[143] A. Yariv. Interpage and interpixel cross talk in orthogonal (wavelength–multiplexed) holograms. *Optics Letters*, 18(8):652–654, 1993.

[144] M. C. Bashaw, M. Jeganathan, and L. Hesselink. Theory of two–center transport in photorefractive media for low–intensity, continuous–wave illumination in the quasi–steady–state limit. *Journal of the Optical Society of America* **B**, 11(9):1743–1757, 1994.

[145] K. Curtis and D. Psaltis. Cross talk for angle– and wavelength–multiplexed image plane holograms. *Optics Letters*, 19(21):1774–1776, 1994.

[146] X. Yi, P. Yeh, and C. Gu. Statistical analysis of cross–talk noise and storage capacity in volume holographic memory. *Optics Letters*, 19(19):1580–1582, 1994.

[147] C. Gu and P. Yeh. Form birefringence of layered media and volume gratings. *Journal of the Optical Society of America* **B**, 12(6):1094–1099, 1995.

[148] X. Yi, P. Yeh, and C. Gu. Cross–talk noise in volume holographic memory with spherical reference beams. *Optics Letters*, 20(17):1812–1814, 1995.

[149] X. Yi, S. Campbell, P. Yeh, and C. Gu. Statistical analysis of cross–talk noise and storage capacity in volume holographic memory: image plane holograms. *Optics Letters*, 20(7):779–781, 1995.

[150] D. Psaltis, X.-G. Gu, and D. Brady. Fractal sampling grids for holographic interconnections. In *Proceedings of the SPIE—963*. SPIE, 1988.

[151] D. Psaltis, D. Brady, X.-G. Gu, and S. Lin. Holography in artificial neural networks. *Nature*, 343(6256), 1990.

[152] C. X.-G. Gu. *Optical neural networks using volume holograms.* PhD thesis, California Institute of Technology, 1990.

[153] F. H. Mok. Angle–multiplexed storage of 5000 holograms in lithium niobate. *Optics Letters*, 18(11):915–917, 1993.

[154] K. Curtis, A. Pu, and D. Psaltis. Method for holographic storage using peristrophic multiplexing. *Optics Letters*, 19(13):993–994, 1994.

[155] A. Pu, K. Curtis, and D. Psaltis. A new method for holographic data storage in photopolymer films. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, pages 433–435. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[156] D. Psaltis, M. Levene, A. Pu, G. Barbastathis, and K. Curtis. Holographic storage using shift multiplexing. *Optics Letters*, 20(7):782–784, 1995.

[157] S. Campbell, X. Yi, and P. Yeh. Hybrid sparse–wavelength angle–multiplexed optical data storage system. *Optics Letters*, 19(24):2161–2163, 1994.

[158] D. Psaltis, D. Brady, and K. Wagner. Adaptive optical networks using photorefractive crystals. *Applied Optics*, 27(9):1752–1759, 1988.

[159] A. C. Strasser, E. S. Maniloff, K. M. Johnson, and S. D. D. Goggin. Procedure for recording multiple–exposure holograms with equal diffraction efficiency in photorefractive media. *Optics Letters*, 14(1):6–8, 1989.

[160] F. H. Mok, G. W. Burr, and D. Psaltis. A system metric for holographic memory systems. *Optics Letters, 1996.*

[161] G. W. Burr and D. Psaltis. Effect of the oxidation state of $LiNbO_3$:Fe on the diffraction efficiency of large scale holographic memories. In Photorefractive Materials, Effects, and Devices Conference, *1995. poster MPB3.*

[162] G. W. Burr and D. Psaltis. Improving diffraction efficiency of holographic memories via $LiNbO_3$:Fe oxidation state. In 1995 OSA Annual Meeting, *1995. paper MSS3.*

[163] G. W. Burr and D. Psaltis. Optimization of the oxidation state of $LiNbO_3$ for large scale holographic storage. Optics Letters, 1996.

[164] Y. Qiao, S. Orlov, and D. Psaltis. Electrical fixing of photorefractive holograms in $(Sr_{0.75}Ba_{0.25})Nb_2O_6$. *Optics Letters*, 18(12):1004–1006, 1993.

[165] A. Horowitz, A. Bekker, and B. Fischer. Image and hologram fixing method with $(Sr_xBa_{1-x})Nb_2O_6$ crystals. *Optics Letters*, 18(22):1964–1966, 1993.

[166] A. Kewitsch, M. Segev, A. Yariv, and R. Neurgaonkar. Selective page–addressable fixing of volume holograms in $Sr_{0.75}Ba_{0.25}Nb_2O_6$ crystals. *Optics Letters*, 18(15):1262–1264, 1993.

[167] S. Orlov, D. Psaltis, and R. R. Neurgaonkar. Spatial and temporal characteristics of electrically fixed holograms in photorefractive strontium–barium niobate. *Applied Physics Letters*, 64(7):824–826, 1994.

[168] P. Gunter, R. S. Cudney, J. Fousek, and M. Zgonik. Photoassisted ferroelectric domain fixing for volume hologram storage. In *OSA Annual Meeting—paper MB5*, 1995.

[169] A. Yariv, S. Orlov, G. Rakuljic, and V. Leyva. Holographic fixing, readout, and storage dynamics in photorefractive materials. *Optics Letters*, 20(11):1334–1336, 1995.

[170] G. Montemezzani and P. Gunter. Thermal fixing impure and doped $KNbO_3$ crystals. *Journal of the Optical Society of America* **B**, 7(12):2323–2328, 1990.

[171] V. Leyva, A. Agranat, and A. Yariv. Fixing of a photorefractive grating in $KTa_{1-x}Nb_xO_3$ by cooling through the ferroelectric phase transition. *Optics Letters*, 16(8):554–556, 1991.

[172] V. Leyva, D. Engin, X. Tong, M. Tong, A. Yariv, and A. Agranat. Fixing of photorefractive volume holograms in $K_{1-y}Li_yTa_{1-x}O_3$. *Optics Letters*, 20(11):1319–1321, 1995.

[173] Y. Furukawa, M. Sato, K. Kitamura, and F. Nitanda. Growth and characterization of off–congruent $LiNbO_3$ single crystals grown by the double crucible method. *Journal of Crystal Growth*, 128:909–914, 1993.

[174] G. I. Malovichko, V. G. Grachev, E. P. Kokanyan, O. F. Schirmer, K. Betzler, B. Gather, F. Jermann, S. Klauer, U. Schlarb, and M. Wohlecke. Characterization of stoichiometric $LiNbO_3$ grown from melts containing K_2O. *Applied Physics A*, 56:103–108, 1993.

[175] G. I. Malovichko, V. G. Grachev, and O. F. Schirmer. The effect of iron ions on the defect structure of lithium niobate crystals grown from K_2O containing melts. *Solid State Communications*, 89(3):195–198, 1994.

[176] P. Shah, F. K. Tittel T. A. Rabson, and T. K. Gaylord. Volume holographic recording and storage in Fe–doped $LiNbO_3$ using optical pulses. *Applied Physics Letters*, 24(3):130–131, 1974.

[177] D. von der Linde, A. M. Glass, and K. F. Rodgers. Multiphoton photorefractive processes for optical storage in $LiNbO_3$. *Applied Physics Letters*, 25(3):155–157, 1974.

[178] O. F. Schirmer and D. Von der Linde. Two–photon and x–ray–induced Nb^{4+} and O^- small polarons in $LiNbO_3$. *Applied Physics Letters*, 33(1):35–38, 1978.

[179] C.-T. Chen, D. M. Kim, and D. von der Linde. Efficient hologram recording in $LiNbO_3$:Fe using optical pulses. *Applied Physics Letters*, 34(5):321–324, 1979.

[180] J. Feinberg, D. Heiman, A. R. Tanguay Jr., and R. W. Hellwarth. Photorefractive effects and light–induced charge migration in barium titanate. *Journal of Applied Physics*, 51(3):1297–1305, 1979.

[181] V. V. Kulikov and S. I. Stepanov. Mechanisms of holographic recording and thermal fixing in photorefractive $LiNbO_3$:Fe. *Soviet Physics—Solid State*, 21(11):1849–1851, 1979.

[182] C.-T. Chen, D. M. Kim, and D. Von der Linde. Efficient pulsed photorefractive process in $LiNbO_3$:Fe for optical storage and deflection. *IEEE Journal of Quantum Electronics*, 16(2):126–129, 1980.

[183] H. Vormann and E. Kratzig. Two step excitation in $LiTaO_3$:Fe for optical data storage. *Solid State Communications*, 49(9):843–847, 1984.

[184] Y. Ming, E. Kratzig, and R. Orlowski. Photorefractive effects in $LiNbO_3$:Cr induced by two–step excitation. *Physica Status Solidi A*, 92:221–229, 1985.

[185] F. Jermann and E. Kratzig. Charge transport processes in $LiNbO_3$:Fe at high intensity laser pulses. *Applied Physics A*, 55:114–118, 1992.

[186] K. Buse and K. H. Ringhofer. Pyroelectric drive for light–induced charge transport in the photorefractive process. *Applied Physics A*, 57:161–165, 1993.

[187] F. Jermann and J. Otten. Light–induced charge transport in $LiNbO_3$:Fe at high light intensities. *Journal of the Optical Society of America* **B**, 10(11):2085–2092, 1993.

[188] F. Jermann and K. Buse. Light–induced thermal gratings in $LiNbO_3$:Fe. *Applied Physics B*, 59:437–443, 1994.

[189] H. Mabuchi, E. S. Polzik, and H. J. Kimble. Blue–light–induced infrared absorption in $KNbO_3$. *Journal of the Optical Society of America* **B**, 11(10):2023–2029, 1994.

[190] K. Buse, F. Jermann, and E. Kratzig. Infrared holographic recording in $LiNbO_3$:Fe and $LiNbO_3$:Cu. *Optical Materials*, 4:237–240, 1995.

[191] F. Jermann, K. Buse, M. Simon, and E. Kratzig. Laser–induced transient gratings in $LiNbO_3$:Fe. *Optical Materials*, 4:318–321, 1995.

[192] M. P. Petrov, S. I. Stepanov, and A. A. Kamshilin. Light diffraction in photorefractive ferroelectrics. *Ferroelectrics*, 21:631–633, 1978.

[193] E. Kh. Gulanyan, I. R. Dorosh, V. D. Iskin, A. L. Mikaelyan, and M. A. Malorchuk. Nondestructive readout of holograms in iron–doped lithium niobate crystals. *Soviet Journal of Quantum Electronics*, 9(5):647–649, 1979.

[194] M. P. Petrov, S. I. Stepanov, and A. A. Kamshilin. Light diffraction from the volume holograms in electrooptic birefringent crystals. *Optics Communication*, 29(1):44–48, 1979.

[195] H. C. Kulich. A new approach to read volume holograms at different wavelengths. *Optics Communication*, 64(5):407–411, 1987.

[196] R. A. Rupp, H. C. Kulich, U. Schurk, and E. Kratzig. Diffraction by difference holograms in electrooptic crystals. *Ferroelectrics Letters*, 8:25–30, 1987.

[197] S. Fries, J. Otten, K. H. Ringhofer, and R. A. Rupp. Frequency–difference holograms in lithium niobate. *Optics Communication*, 72(3,4):169–174, 1989.

[198] S. Fries, S. Bauschulte, E. Kratzig, K. Ringhofer, and Y. Yacoby. Spatial frequency mixing in lithium niobate. *Optics Communication*, 84(5,6):251–257, 1991.

[199] H.-C. Kulich. Reconstructing volume holograms without image field losses. *Applied Optics*, 30(20):2850–2857, 1991.

[200] S. Fries. Spatial frequency mixing in electrooptic crystals: Application to non-destructive read–out of optically erasable volume holograms. *Applied Physics A*, 55:104–113, 1992.

[201] H. C. Kulich. Transfer function for image formation of objects reconstructed from volume holograms with different wavelengths. *Applied Optics*, 31(14):2461–2477, 1992.

[202] D. Psaltis, F. Mok, and H.-Y. S. Li. Nonvolatile storage in photorefractive crystals. *Optics Letters*, 19(3):210–212, 1994.

[203] A. A. Gorokhovskii, R. K. Kaarli, and L. A. Rebane. Hole burning in the contour of a pure electronic line in an Shpolskii system. *JETP Letters*, 20:216–219, 1974.

[204] R. I. Personov and L. A. Bykovskaya. Stable gap in absorption spectra of solid solutions of organic molecules by laser irradiation. *Optics Communication*, 12:191–194, 1974.

[205] A. Renn, A. J. Meizner, U. P. WIld, and F. A. Burkhalter. Holographic detection of photochemical holes. *Chemical Physics*, 93:157–162, 1985.

[206] A. Renn and U. P. Wild. Spectral hole burning and hologram storage. *Applied Optics*, 26:4040–4042, 1987.

[207] A. J. Meixner, A. Renn, and U. P. Wild. Spectral hole burning and holography. I. Transmission and holographic detection of spectral holes. *Journal of Chemical Physics*, 91:6728–6736, 1989.

[208] A. Renn, A. J. Meixner, and U. P. Wild. Spectral hole burning and holography. II. Diffraction properties of two spectrally adjacent holograms. *Journal of Chemical Physics*, 92:2748–2755, 1990.

[209] T. W. Mossberg. Time–domain frequency–selective optical data storage. *Optics Letters*, 7:77–79, 1982.

[210] N. W. Carlson, L. J. Rothberg, A. G. Yodh, W. R. Babbitt, and T. W. Mossberg. Storage and time reversal of light pulses using photon echoes. *Optics Letters*, 8:483–485, 1983.

[211] H. Lin, T. Wang, and T. W. Mossberg. Demonstration of 8–Gbit/in^2 areal storage density based on swept–carrier frequency–selective optical memory. *Optics Letters*, 20(15):1658–1660, 1995.

[212] H. Schwoerer, D. Erni, and A. Rebane. Holography in frequency–selective media. III. Spectral synthesis of arbitrary time–domain pulse shapes. *Journal of the Optical Society of America* **B**, 12(6):1083–1093, 1995.

[213] E. S. Maniloff, S. B. Altner, S. Bernet, F. R. Graf, and A. Renn. Recording of 6000 holograms by use of spectral hole–burning. *Applied Optics*, 34(20):4140–4148, 1995.

[214] R. A. Linke, T. Thio, J. D. Chadi, and G. E. Devlin. Diffraction from optically written persistent plasma gratings in doped compound semiconductors. *Applied Physics Letters*, 65(1):16–18, 1994.

[215] R. L. MacDonald, R. A. Linke, J. D. Chadi, G. E. Devlin, and P. Becla. Thick plasma gratings using a local photorefractive effect in CdZnTe:In. *Optics Letters*, 19(24):2131–2133, 1994.

[216] R. L. MacDonald, R. A. Linke, G. E. Devlin, and M. Mizuta. Confirmation of the local nature of the plasma grating photorefractive effect. *Optics Letters*, 20(11):1322–1324, 1995.

[217] R. L. MacDonald and R. A. Linke. Optical phase conjugation using DX centers. *Optics Letters*, ?(?):?, 1996.

[218] D. P. Hand and P. St. J. Russell. Photoinduced refractive–index changes in germanosilicate fibers. *Optics Letters*, 15(2):102–104, 1990.

[219] T. Erdogan, A. Partovi, V. Mizrahi, P. J. Lemaire, W. L. Wilson, T. A. Strasser, and A. M. Glass. Volume gratings for holographic storage applications written in high–quality germanosilicate glass. *Applied Optics*, 34(29):6738–6743, 1995.

[220] J. A. Jenny. Holographic recording with photopolymers. *Journal of the Optical Society of America*, 60(9):1155–1161, 1970.

[221] F. P. Laming. Holographic grating formation in photopolymers—polymethylmethacrylate. *Polymer engineering and science*, 11(5):421–425, 1971.

[222] W. S. Colburn and K. A. Haines. Volume hologram formation in photopolymer materials. *Applied Optics*, 10:1636–1641, 1971.

[223] R. A. Bartolini, A. Bloom, and H. A. Weakliem. Volume holographic recording characteristics of an organic medium. *Applied Optics*, 15(5):1261–1265, 1976.

[224] R. A. Bartolini, A. Bloom, and J. S. Escher. Multiple storage of holograms in an organic medium. *Applied Physics Letters*, 28(9):506–507, 1976.

[225] A. Bloom, R. A. Bartolini, and H. A. Weakliem. Organic materials for optical devices. *Optical Engineering*, 17(5):446–452, 1976.

[226] D. Blitz and B. J. Pernick. Polarization properties of photopolymers for use in holographic and coherent optical systems. *Applied Optics*, 32(32):6501–6502, 1993.

[227] A. Fimia, N. Lopez, F. Mateos, R. Sastre, J. Pineda, and F. Amat-Guerri. New photopolymer used as a holographic recording material. *Applied Optics*, 32(20):3706–3707, 1993.

[228] U.-S. Rhee, H. J. Caulfield, J. Shamir, C. S. Vikram, and M. M. Mirsalehi. Characteristics of the Du Pont photopolymer for angularly multiplexed page–oriented holographic memories. *Optical Engineering*, 32(8):1839–1847, 1993.

[229] K. Curtis and D. Psaltis. Characterization of the DuPont photopolymer for three–dimensional holographic storage. *Applied Optics*, 33(23):5396–5399, 1994.

[230] F. Ghailane, G. Manivannan, and R. A. Lessard. Spiropyran–doped poly(vinyl carbazole): a new photopolymer recording medium for erasable holography. *Optical Engineering*, 34(2):480–485, 1995.

[231] U.-S. Rhee, H. J. Caulfield, C. S. Vikram, and J. Shamir. Dynamics of hologram recording in DuPont photopolymer. *Applied Optics*, 34(5):846–853, 1995.

[232] G. Zhao and P. Mouroulis. Second order grating formation in dry holographic photopolymers. *Optics Communication*, 115:528–532, 1995.

[233] A. Pu. Private communication.

[234] D. Psaltis and A. Pu. Holographic 3–D disk. *Optoelectronics—devices and technologies*, 10(3):333–342, 1995.

[235] A. Pu and D. Psaltis. High density recording in photopolymer–based holographic 3–d disks. *Submitted to Applied Optics*, May 1996.

[236] G. Barbastathis, M. Levene, and D. Psaltis. Shift multiplexing with spherical reference waves. *Submitted to Applied Optics*, May(?):???, 1996.

[237] G. Barbastathis, A. Pu, M. Levene, and D. Psaltis. Holographic 3–D disks using shift multiplexing. In *1995 Optical Data Storage meeting*, 1995. SPIE Vol 2514–38.

[238] R. V. Johnson and A. R. Tanguay Jr. Stratified volume holographic optical elements. *Optics Letters*, 13(3):189–191, 1988.

[239] G. P Nordin and A. R. Tanguay Jr. Photopolymer–based stratified volume holographic optical elements. *Optics Letters*, 17(23):1709–1711, 1992.

[240] G. P. Nordin, R. V. Johnson, and A. R. Tanguay Jr. Diffraction properties of stratified volume holographic optical elements. *Journal of the Optical Society of America* **A**, 9(12):2206–2217, 1992.

[241] A. Granger, L. Song, and R. A. Lessard. Multiple beam generation using a stratified volume holographic grating. *Applied Optics*, 32(14):2534–2537, 1993.

[242] R. De Vre and L. Hesselink. Analysis of photorefractive stratified volume holographic optical elements. *Journal of the Optical Society of America* **B**, 11(9):1800–1808, 1994.

[243] R. De Vre, M. Jeganathan, J. P. Wilde, and L. Hesselink. Effect of applied fields on the bragg condition and the diffraction efficiency in photorefractive crystals. *Optics Letters*, 19(12):910–912, 1994.

[244] J. J. Stankus, S. M. Silence, W. E. Moerner, and G. C. Bjorkland. Electric–field–switchable stratified volume holograms in photorefractive polymers. *Optics Letters*, 19(18):1480–1482, 1994.

[245] Holoplex. HM–100 holographic memory product.

[246] S. Campbell, S.-H. Lin, X. Yi, and P. Yeh. Absorption effects in photorefractive volume holographic memory. In *Topical meeting on Photorefractive Materials, Devices, and Effects*, 1995.

[247] Epson. VGA projector.

[248] Kopin. KVGA & LVGA projectors.

[249] Kopin. under development.

[250] H.-Y. S. Li. *Photorefractive 3–D disks for optical data storage and artificial neural networks.* PhD thesis, California Institute of Technology, 1994.

[251] A. Pu, K. Curtis, and D. Psaltis. Exposure schedule for multiplexing holograms in photopolymer films. *Submitted to Optical Engineering*, 1996.

[252] X. An and D. Psaltis. Experimental characterization of an angle–multiplexed holographic memory. *Optics Letters*, 20(18):1913–1915, 1995.

[253] A.A. Blok. Effect of data coding methods in holographic memory on characteristics of reconstructed images of data pages. *Optoelectronic, Imstrumentation and Data Processing*, 5:45–51, 1989.

[254] V.A. Dombrovskii, S.A. Dombrovoskii, and E.F. Pen. Noise immunity of holographic memory with paraphase data coding. *Optoelectronics, Instrumentation and Data Processing*, 2:63–69, 1988.

[255] S.A. Dombrovoskii. Effectiveness of using error-correcting codes in holographic storage. *Optoelectronics, Instrumentation and Data Processing*, 2:58–62, 1989.

[256] J. D. Hayes and M. A. Neifeld. Error correction schemes for highly parallel access memories. In *OSA Annual meeting—1994*, 1994.

[257] M. A. Neifeld and J. D. Hayes. Parallel error correction for optical memories. *Optical Memory and Neural Networks*, 3(2):87–98, 1994.

[258] M. A. Neifeld and M. McDonald. Error correction for increasing the usable capacity of photorefractive memories. *Optics Letters*, 19:1483–1485, 1994.

[259] B. H. Olson and S. C. Esener. Partial response precoding for parallel–readout optical memories. *Optics Letters*, 19(9):661–663, 1994.

[260] B. J. Goertzen and P. A. Mitkas. An eror–correcting code for volume holographic storage of a relational database. Submitted to Optical Computing 1995, 1995.

[261] W.-F. Hsu and A. A. Sawchuk. Volume optical memory smart pixel interfaces with 0.1—1.0 terabit per second performance. In *1995 OSA annual meeting*, 1995.

[262] W. Kawakami and K.-I. Kitayama. Optical error–correction coding encoder and decoder: design considerations. *Applied Optics*, 34(23):5064–5073, 1995.

[263] B. Olson and S. Esener. Coding and processing for 3–D parallel access optical storage. In *1995 OSA annual meeting*, 1995.

[264] S. K. Sridharan and M. A. Neifeld. Parallel error correction for page access optical memories. In *1995 OSA annual meeting*, 1995.

[265] J. F. Heanue, M. C. Bashaw, and L. Hesselink. Volume holographic storage and retrieval of digital data. *Science*, 265:749–752, August 5 1994.

[266] N. Negroponte. *Being digital.* Random House, 1995.

[267] P. J. van Heerden. A new optical method of storing and retrieving information. *Applied Optics*, 2(4):387–392, 1963.

[268] D. Psaltis. Class notes: Optical computing, winter 1994.

[269] F. M. Smits and L. E. Gallaher. Design considerations for a semipermanent optical memory. *The Bell System Technical Journal*, pages 1267–1278, July–August 1967.

[270] A. L. Mikaelian, V. I. Bobrinev, S. M. Naumov, and L. Z. Sokolova. Design principles of holographic memory devices. *IEEE Journal of Quantum Electronics*, 6(4):193–198, 1970.

[271] P. Graf and M. Lang. Geometrical aspects of consistent holographic memory design. *Applied Optics*, 11(6):1382–1388, 1972.

[272] B. Hill. Some aspects of a large capacity holographic memory. *Applied Optics*, 11(1):182–191, 1972.

[273] A. Vander Lugt. Design relationships for holographic memories. *Applied Optics*, 12(7):1675–1685, 1973.

[274] A. Vanderlugt. Packing density in holographic systems. *Applied Optics*, 14(5):1081–1087, 1975.

[275] S. Tao, D. R. Selviah, and J. E. Midwinter. Spatioangular multiplexed storage of 750 holograms in an Fe:$LiNbO_3$ crystal. *Optics Letters*, 18(11):912–914, 1993.

[276] J. H. Hong, I. McMichael, T. Y. Chang, W. Christian, and E. G. Paek. Volume holographic memory systems: techniques and architectures. *Optical Engineering*, 34(8):2193–2203, 1995.

[277] S. Tao, Z. H. Song, D. R. Selviah, and J. E. Midwinter. Spatioangular-multiplexing scheme for dense holographic storage. *Applied Optics*, 34(29):6729–6737, 1995.

[278] J. P. Huignard and F. Micheron. Control of sensitivity to hologram storage in $LiNbO_3$ using an accessory photovoltaic crystal. *Optics Communication*, 16(1):80–82, 1976.

[279] F. H. Mok, D. Psaltis, and G. W. Burr. Spatially– and angle–multiplexed holographic random access memory (invited paper). In *Proceedings of the SPIE, Vol 1773c*, 1992.

[280] G. W. Burr, K. Curtis, and D. Psaltis. Comparison of wavelength– and angle–multiplexed holographic memories. In *OSA 1993 Topical Meeting on Optical Computing*, 1993.

[281] F. H. Mok, D. Psaltis, and G. W. Burr. Large–scale holographic memories (invited paper). In *CLEO 1994*, page 7, 1994. paper CMB3.

[282] G. W. Burr, F. H. Mok, and D. Psaltis. Angle and space multiplexed holographic storage using the 90° geometry. *Optics Communications*, 117:49–55, 1995.

[283] G. W. Burr, F. H. Mok, and D. Psaltis. Large–scale holographic memory: Experiment results. In *1993 ISOM/ODS Meeting*, 1993.

[284] G. W. Burr, F. H. Mok, and D. Psaltis. Storage of 10,000 holograms in $LiNbO_3$:Fe. In *CLEO 1994*, page 9, 1994. paper CMB7.

[285] F. T. S. Yu, S. Wu, A. W. Mayers, and S. Rajan. Wavelength multiplexed reflection matched spatial filters using $LiNbO_3$. *Optics Communication*, 81(6):343–347, 1991.

[286] G. A. Rakuljic, V. Leyva, and A. Yariv. Optical data storage by using orthogonal wavelength–multiplexed volume holograms. *Optics Letters*, 17(20):1471–1473, 1992.

[287] H. Yamamoto, K. Maeda, S. Ishizuka, and T. Kubota. Real–time measurement of wavelength selectivity of reflection holograms. *Applied Optics*, 31(35):7397–7399, 1992.

[288] G. A. Rakuljic and V. Leyva. Volume holographic narrow–band optical filter. *Optics Letters*, 18(6):459–461, 1993.

[289] J. Rosen, M. Segev, and A. Yariv. Wavelength–multiplexed computer–generated volume holography. *Optics Letters*, 18(9):744–746, 1993.

[290] S. Yin, H. Zhou, F. Zhao, M. Wen, Z. Yang, J. Zhang, and F. T. S. Yu. Wavelength multiplexed holographic storage in a sensitive photorefractive crystal using a visible–light tunable diode laser. *Optics Communication*, 101:317–321, 1993.

[291] F. Zhao, H. Zhou, S. Yin, and F. T. S. Yu. Wavelength–multiplexed holographic storage by using the minimum wavelength channel separation in a photorefractive crystal fiber. *Optics Communication*, 103:59–62, 1993.

[292] J. Ma, B. Catanzaro, J. E. Ford, Y. Fainman, and S. H. Lee. Photorefractive holographic lenses and applications for dynamic focusing and dynamic image shifting. *Journal of the Optical Society of America* **A**, 11(9):2471–2480, 1994.

[293] F. T. S. Yu, S. Yin, and Z. H. Yang. Thick volume photorefractive crystal wavelength–multiplexed reflection–type matched filter. *Optical Memory and Neural Networks*, 3(2):207–215, 1994.

[294] H. Zhou, F. Zhao, and F. T. S. Yu. Diffraction properties of a reflection photorefractive hologram. *Applied Optics*, 33(20):4345–4352, 1994.

[295] K. Curtis. *3–D photopolymer disks for correlation and data storage, and cross–talk in volume holographic memories.* PhD thesis, California Institute of Technology, 1994.

[296] C. J. Chang–Hasnain, J. P. Harbison, C. E. Zah, L. T. Florez, and N. C. Andreadakis. Continuous wavelength tuning of two–electrode vertical cavity surface–emitting lasers. *Electronics Letters*, 27(11):1002–1003, 1992.

[297] M. C. Bashaw, R. C. Singer, J. F. Heanue, and L. Hesselink. Coded and mosaic wavelength multiplex volume holography. In *Photorefractive Materials, Effects, & Devices—paper TC3*, 1995.

[298] P. Yeh. Fundamental limit of the speed of photorefractive effect and its impact on device applications and material research. *Applied Optics*, 26(4):602–604, 1987.

[299] J. T. LaMacchia and D. L. White. Coded multiple exposure holograms. *Applied Optics*, 7(1):91–94, 1968.

[300] T. F. Krile, R. J. Marks II, J. F. Walkup, and M. O. Hagler. Holographic representations of space–variant systems using phase–coded reference beams. *Applied Optics*, 16(12):3131–3135, 1977.

[301] V. N. Morozov. Theory of holograms formed using a coded reference beam. *Soviet Journal of Quantum Electronics*, 7(8):961–964, 1977.

[302] T. F. Krile, M. O. Hagler, W. D. Redus, and J. F. Walkup. Multiplex holography with chrip–modulated binary phase–coded reference–beam masks. *Applied Optics*, 18(1):52–56, 1979.

[303] M. I. Jones, J. F. Walkup, and M. O. Hagler. Multiplex hologram representations of space–variant optical systems using ground–glass encoded reference beams. *Applied Optics*, 21(7):1291–1297, 1982.

[304] J. E. Ford, Y. Fainman, and S. H. Lee. Array interconnection by phase–coded optical correlation. *Optics Letters*, 15(19):1088–1090, 1990.

[305] C. Denz, G. Pauliat, G. Roosen, and T. Tschudi. Volume hologram multiplexing using a deterministic phase encoding method. *Optics Communication*, 85:171–176, 1991.

[306] C. Denz, G. Pauliat, G. Roosen, and T. Tschudi. Potentialities and limitations of hologram multiplexing by using the phase–encoding technique. *Applied Optics*, 31(26):5700–5705, 1992.

[307] C. Alves, G. Pauliat, and G. Roosen. Dynamic phase–encoding storage of 64 images in a $BaTiO_3$ photorefractive crystal. *Optics Letters*, 19(22):1894–1896, 1994.

[308] J. F. Heanue, M. C. Bashaw, and L. Hesselink. Recall of linear combinations of stored data pages based on phase–code multiplexing in volume holography. *Optics Letters*, 19(14):1079–1081, 1994.

[309] J. F. Heanue, M. C. Bashaw, and L. Hesselink. Encrypted holographic data storage based on orthogonal–phase–code multiplexing. *Applied Optics*, 34(26):6012–6015, 1995.

[310] A. N. Jain. *Fundamentals of digital image processing.* Prentice–Hall, 1989.

[311] A. Pu, R. Denkewalter, and D. Psaltis. Robot navigation using a peristrophic holographic memory. In *OSA Spring Topical Meetings—Optical Computing*, Salt Lake City, March 1995. Optical Society of America.

[312] H.-Y. S. Li and D. Psaltis. Three–dimensional holographic optical disks. *Applied Optics*, 33(17):3764–3774, 1994.

[313] P. G. de Gennes. *The Physics of Liquid Crystals.* Oxford University Press, 1974.

[314] I.-C. Khoo. *Liquid Crystals: Physical properties and nonlinear optical phenomena.* John Wiley & Sons, 1995.

[315] C. H. Gooch and H. A. Tarry. Optical characteristics of twisted nematic liquid-crystal films. *Electronics Letters*, 10(1):1–4, January 1974.

[316] Jean-Jacques P. Drolet, Jay S. Patel, Konstantinos G Haritos, Weihua Xu, Axel Scherer, and Demetri Psaltis. Hybrid-aligned nematic liquid-crystal modulators fabricated on VLSI circuits. *Optics Letters*, 20(21):2222–2224, November 1995.

[317] Paul M. Alt and Peter Pleshko. Scanning limitations of liquid-crystal displays. *IEEE Transactions on Electron Devices*, ED-21(2):146–155, February 1974.

[318] Jürgen Nehring and Allan R. Kmetz. Ultimate limits for matrix addressing of RMS-responding liquid-c rystal displays. *IEEE Transactions on Electron Devices*, ED-26(5):795–801, May 1979.

[319] Noel A. Clark and Sven T. Lagerwall. Submicrosecond bistable electro-optic switching in liquid crystals. *Applied Physics Letters*, 36(11):899–901, June 1980.

[320] D. C. O'Brien, R. J. Mears, T. D. Wilkinson, and W. A. Crossland. Dynamic holographic interconnects that use ferroelectric liquid–crystal spatial light modulators. *Applied Optics*, 33(14):2795–2803, 1994.

[321] D. J. McKnight, K. M. Johnson, and R. A. Serati. Electrically addressed 256 by 256 liquid–crystal–on–silicon spatial light modulator. *Optics Letters*, 18(24):2159–2161, 1995.

[322] Kristina M. Johnson, Douglas J. McKnight, and Ian Underwood. Smart spatial light modulators using liquid crystals on silicon. *IEEE Journal of Quantum Electronics*, 29(2):699–714, February 1993.

[323] Douglas J. McKnight, Kristina M. Johnson, and Roylnn A. Serati. 256 × 256 liquid-crystal-on-silicon spatial light modulator. *Applied Optics*, 33(14):2775–2784, May 1994.

[324] Ian Underwood, David G. Vaas, Antony O'Hara, Dwayne C. Bu rns, Peter W. McOwan, and James Gourlay. Improving the performance of liquid-crystal-over-silicon spatial l ight modulators: Issues and achievements. *Applied Optics*, 33(14):2768–2774, May 1994.

[325] Displaytech. Head–mounted display product.

[326] Douglas J. McKnight. Continuous view of dc-balanced images on a ferroelectric liquid-cr ystal spatial light modulator. *Optics Letters*, 19(18):1471–1473, September 1994.

[327] Jean-Jacques P. Drolet, George Barbastathis, Jay S. Patel, and Demetri Psaltis. Liquid crystal devices for volume holographic memories. In *OSA Annual Meeting*, September 1995.

[328] Jean-Jacques P. Drolet, Demetri Psaltis, and Jay S. Patel. Nematic liquid crystal analog modulators fabricated on VLSI circuitry. In *OSA Spring Topical Meetings*, Salt Lake City, March 1995. Optical Society of America. post-deadline paper.

[329] G. W. Burr, F. H. Mok, and D. Psaltis. Spatial and angle multiplexing using the long interaction length architecture. In *OSA 1993 Annual Meeting*, 1993. Paper TuH6.

[330] E. S. Maniloff and K. M. Johnson. Maximized photorefractive holographic storage. *Journal of Applied Physics*, 70(9):4702–4707, 1991.

[331] E. S. Maniloff and K. M. Johnson. Effects of scattering on the dynamics of holographic recording and erasure in photorefractive lithium niobate. *Journal of Applied Physics*, 73(2):541–547, 1993.

[332] Y. Taketomi, J. E. Ford, H. Sasaki, J. Ma, Y. Fainman, and S. H. Lee. Incremental recording for photorefractive hologram multiplexing. *Optics Letters*, 16(22):1774–1776, 1991.

[333] P. Gunter and J.-P. Huignard. Photorefractive effects and materials. In P. Gunter and J.-P. Huignard, editors, *Topics in Applied Physics—Vol. 61: Photorefractive Materials and Their Applications I—Fundamental Phenomena*, pages 7–73. Springer–Verlag, 1988.

[334] H. Kurz. Lithium niobate as a material for holographic information storage. *Philips Technical Review*, 37(5/6):109–120, 1977.

[335] H. Kurz, E. Kratzig, W. Keune, H. Engelmann, U. Gonser, B. Dischler, and A. Rauber. Photorefractive centers in $LiNbO_3$—studied by optical, Mossbauer, and EPR—methods. *Applied Physics*, 12:355–368, 1977.

[336] J. R. Carruthers, G. E. Peterson, M. Grasso, and P. M. Bridenbaugh. Nonstoichiometry and crystal growth of lithium niobate. *Journal of Applied Physics*, 42(5):1846–1851, 1971.

[337] P. Bridenbaugh. Private communication.

[338] L. O. Svaasand, M. Eriksrud, G. Nakken, and A. P. Grande. Solid–solution range of $LiNbO_3$. *Journal of Crystal Growth*, 22:230–232, 1974.

[339] M. P. Petrov, S. I. Stepanov, and A. V. Knomenko. *Photorefractive crystals in coherent optical systems*, volume 59 of *Springer Series in Optical Sciences*. Springer–Verlag, 1991.

[340] C. Gu, J. Hong, H.-Y. Li, D. Psaltis, and P. Yeh. Dynamics of grating formation in photovoltaic media. *Journal of Applied Physics*, 69(3):1167–1172, 1991.

[341] R. von Baltz and Ch. Lingenfelder. Nonlocal photovoltaic response function for the interpretation of hologram writing in ferroelectric crystals. *Applied Physics A*, 32:13–18, 1983.

[342] R. A. Rupp, R. Sommerfeldt, K. H. Ringhofer, and E. Kratzig. Space charge field limitations in photorefractive $LiNbO_3$:Fe crystals. *Applied Physics B*, 51:364–370, 1990.

[343] R. S. Weis and T. K. Gaylord. Lithium niobate: summary of physical properties and crystal structure. *Applied Physics A*, 37:191–203, 1985.

[344] D. W. Oxtoby and N. H. Nachtrieb. *Principles of modern chemistry*. Saunders college publishing, 1986.

[345] E. Kratzig. Photorefractive effects and photoconductivity in $LiNbO_3$:Fe. *Ferroelectrics*, 21:635–636, 1978.

[346] E. Kratzig and H. Kurz. Spectroscopic investigation of photovoltaic effects in doped $LiNbO_3$. *Journal of the Electrochemical Society*, 124(1):131–134, 1977.

[347] L. Young, M. G. Moharam, F. El Guibaly, and E. Lun. Hologram writing in lithium niobate: beam coupling and the transport length in the bulk photovoltaic effect. *Journal of Applied Physics*, 50(6):4201–4207, 1979.

[348] R. J. Woods and L. Young. Spatial phase shift during hologram writing in lithium niobate. *Ferroelectrics*, 46:275–280, 1983.

[349] A. Yariv. *Optical Electronics, 4^{th} ed.* Saunders College Publishing, 1991.

[350] H.-Y. Li and D. Psaltis. Double grating formation in anisotropic photorefractive crystals. *Journal of Applied Physics*, 71(3):1394–1400, 1992.

[351] M. Carrascosa, J. M. Cabrera, and F. Agullo–Lopez. Steady holographic gratings formed in photorefractive materials: influence of material parameters. *IEEE Journal of Quantum Electronics*, 27(3):509–515, 1991.

[352] G. W. Burr, F. H. Mok, and D. Psaltis. Large scale volume holographic storage in the long interaction length architecture. In *Proceedings of the SPIE, Vol 2297-47*, 1994.

[353] G. W. Burr, X. An, F. H. Mok, and D. Psaltis. Large–scale rapid access holographic memory. In *1995 Optical Data Storage meeting*, 1995. SPIE Vol 2514–39.

[354] J. P. Huignard, F. Micheron, and E. Spitz. Optical systems and photosensitive materials for information storage. In *Optical properties of Solids*, chapter 16, pages 847–925. North Holland, Amsterdam, 1976.

[355] X. An, G. W. Burr, and D. Psaltis. Angle–multiplexed holographic storage using acousto–optic deflectors. In *1995 OSA Annual Meeting*, 1995. paper MP6.

[356] J. L. Pezzanti and R. A. Chipman. Cascaded polarizing beamsplitter cubes in imaging–systems. *Optical Engineering*, 33(5):1543–1549, 1995.

[357] J. L. Pezzanti and R. A. Chipman. Linear polarization uniformity measurements taken with an imaging polarimeter. *Optical Engineering*, 34(6):1569–1573, 1995.

[358] D. B. Neumann and H. W. Rose. Improvement of recorded holographic fringes by feedback control. *Applied Optics*, 6(6):1097–1104, 1967.

[359] D. R. MacQuigg. Hologram firnge stabilization method. *Applied Optics*, 16(2):291–292, 1977.

[360] M. G. Moharam. Cross–talk and cross–coupling in multiplexed holographic gratings. In *Practical Holography III—SPIE Vol 1051*, pages 143–147. SPIE, 1989.

[361] M. C. Bashaw, J. F. Heanue, and L. Hesselink. Crosstalk control for multiplex holography. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, pages 224–226. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[362] X. Yi, P. Yeh, and C. Gu. Cross–talk noise and storage density in holographic memory. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, pages 436–438. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[363] J. V. Alvarez–Bravo, N. Bolognini, and L. Arizmendi. Cross–talk in multiplexed holograms using angular selectivity in $LiNbO_3$. *Optical Materials*, 4:414–418, 1995.

[364] C. Gu, H. Fu, and J.-R. Lien. Correlation patterns and cross–talk noise in volume holographic optical correlators. *Journal of the Optical Society of America* **A**, 12(5):861–868, 1995.

[365] J. F. Heanue, M. C. Bashaw, and L. Hesselink. Sparse selection of reference beams for wavelength– and angular–multiplexed volume holography. *Journal of the Optical Society of America* **A**, 12(8):1671–1676, 1995.

[366] J. Lembcke, C. Denz, and T. Tschudi. General formalism for angular and phase–encoding multiplexing in holographic image storage. *Optical Materials*, 4:428–432, 1995.

[367] H. Vormann, G. Weber, S. Kapphan, and E. Kratzig. Hydrogen as origin of thermal fixing in $LiNbO_3$:Fe. *Solid State Communications*, 40:543–545, 1981.

[368] F. El Guibaly. Information storage in ferroelectrics: real–time monitoring of refractive indices. *Canadian Journal of Physics*, 66:659–661, 1988.

[369] M. Carrascosa and F. Agullo-Lopez. Theoretical modeling of the fixing and developing of holographic gratings in $LiNbO_3$. *Journal of the Optical Society of America* **B**, 7(12):2317–2322, 1990.

[370] L. Arizmendi, P. D. Townsend, M. Carrascosa, J. Baquedano, and J. M. Cabrera. Photorefractive fixing and related thermal effects in $LiNbO_3$. *Journal of Physics: Condensed Matter*, 3:5399–5406, 1991.

[371] R. Muller, L. Arizmendi, M. Carrascosa, and J. M. Cabrera. Determination of H concentration in $LibO_3$ by photorefractive fixing. *Applied Physics Letters*, 60(26):3212–3214, 1992.

[372] M. Carrascosa and L. Arizmendi. High–temperature photorefractive effects in $LiNbO_3$:Fe. *Journal of Applied Physics*, 73(6):2709–2713, 1993.

[373] A. Yariv, V. Leyva, and G. Rakuljic. Relaxation and lifetime of "fixed" charge holograms. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, page postdeadline. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[374] R. Muller, L. Arizmendi, M. Carrascosa, and J. M. Cabrera. Time evolution of photorefractive fixing processes in $LiNbO_3$. *Optical Materials*, 4:290–293, 1995.

[375] D. Brady, K. Hsu, and D. Psaltis. Periodically refreshed multiply exposed photorefractive holograms. *Optics Letters*, 15(14):817–819, 1990.

[376] E. G. Paek, E. C. Jung, Y. Silberberg, and T. S. Ravi. Enhanced nonlinear recording using a thermoplastic plate. *Applied Physics Letters*, 59(9):1019–1021, 1991.

[377] Y. Qiao, D. Psaltis, C. Gu, J. Hong, P. Yeh, and R. R. Neurgaonkar. Phase-locked sustainment of photorefractive holograms using phase conjugation. *Journal of Applied Physics*, 70(8):4646–4648, 1991.

[378] H. Sasaki, Y. Fainman, J. E. Ford, Y. Taketomi, and S. H. Lee. Dynamic photorefractive optical memory. *Optics Letters*, 16(23):1874–1876, 1991.

[379] S. Boj, G. Pauliat, and G. Roosen. Dynamic holographic memory showing readout, refreshing, and updating capabilities. *Optics Letters*, 17(6):438–440, 1992.

[380] S. Piazzolla, B. K. Jenkins, and Jr. A. R. Tanguay. Single–step copying process for multiplexed volume holograms. *Optics Letters*, 17(9):676–678, 1992.

[381] Y. Qiao and D. Psaltis. Sampled dynamic holographic memory. *Optics Letters*, 17(19):1376–1378, 1992.

[382] H. Rajbenbach, S. Bann, and J. P. Huignard. Long–term readout of photorefractive memories by using a storage/amplification two–crystal configuration. *Optics Letters*, 17(23):1712–1714, 1992.

[383] Y. Qiao and D. Psaltis. Photorefractive holographic memories with dynamic copying. *International Journal of Optical Computing*, 2:185–199, 1993.

[384] S. Campbell, P. Yeh, C. Gu, S. H. Lin, C.-J. Chen, and K. Y. Hsu. Optical self–enhancement of photorefractive holograms. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, pages 430–432. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[385] P. Yeh, C. Gu, C.-J. Cheng, and K. Y. Hsu. Optical restoration of photorefractive holograms. *Applied Physics B*, 61(5):511–514, 1995.

[386] P. Yeh, C. Gu, C.-J. Cheng, and K. Y. Hsu. Hologram restoration and enhancement in photorefractive media. In *Nonlinear Optics: Materials, Fundamentals, and Applications*, pages 421–423. IEEE/Lasers and Electro–Optics Society and OSA, July 1994.

[387] T. Dellwig, C. Denz, T. Rauch, and T. Tschudi. Coherent refreshment and updating for dynamic photorefractive optical memories using phase conjugation. *Optics Communication*, 119:333–340, 1995.

[388] A. Yariv and P. Yeh. *Optical Waves in Crystals.* John Wiley & Sons, 1984.

[389] Physik Intrumente. Low–voltage piezoelectric translator data sheet.

[390] R. De Vre, J. F. Heanue, K. Gurkan, and L. Hesselink. Bragg detuning effects in photorefractive crystals. In *Topical meeting on Photorefractives Materials, Effects, and Devices*, 1995.

[391] N. Nakayama, S. Itoh, T. Ohata, K. Nakano, and H. Okyuama. Room temperature continuous operation of blue–green laser diodes. *Electronics Letters*, 29(16):1488–1489, 1993.

[392] Paul F. McManamon, Edward A. Watson, Terry A. Dorschner, and Lawrence J. Barnes. Applications look at the use of liquid crystal writable gratings f or steering passive radiation. *Optical Engineering*, 32(11):2657–2664, November 1993.

[393] Terry A. Dorschner and Daniel P. Resler. Optical beam steerer having subaperture addressing, March 1992. U.S. Patent 5,093,740.

[394] G. W. Burr, F. H. Mok, and D. Psaltis. Large–scale holographic memory: Experimental results (invited paper). In *Proceedings of the SPIE, Vol 2026–62*, 1993.

[395] G. W. Burr. Holographic random–access memory. *SPIE Optical Processing & Computing newsletter*, November 1993.