Long–term Large–scale Holographic Storage in $LiNbO_3$:Fe

Thesis by

Xin An

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1998

(Submitted April 27, 1998)

To my parents

Acknowledgements

First, I owe a whole-hearted thank you to Professor Demetri Psaltis. He is a wonderful advisor who has created a first-class research group and a wonderful culture. I am grateful to him for his patience and trust when I was struggling to find the right direction in the first two years of my study. I appreciate his finding the experimentalist inside of me with ever-increasing self-confidence, and his guidance in making me a true PhD candidate with solid theoretical understanding built upon invaluable lab experiences. I thank him for his confidence in me in giving me the latitude and the largest lab as well as the biggest collection of lab equipments to do things in my own ways many times (though sometimes not necessarily the right ones). Professor Psaltis is also a dedicated and understanding teacher. I would never forget rehearsing my first conference talk sentence by sentence with him. I could not have accomplished this without his help and guidance.

It is always true that the life here at Caltech is extremely tough, but it is also always fun to be in a professional family like Psaltis' group. I would like to thank Dr. Sidney Li for teaching me the basics of volume holography and lab skills, and for his example as an all-around researcher. I am greatly indebted to Dr. Geoffrey Burr for being a wonderful partner during one of the toughest and most fruitful year of my research. Besides the experiences we shared and the knowledge I learned from him, his inspiration and constant "push" are also very rewarding. A special thank you goes to Dr. Allen Pu for being a good friend and for the movie nights, the beers, the paintball games we had together, and for introducing me to the world of "Married with Children" (though I still got married after watching so many episodes). I am also grateful for the many insightful discussions with Michael Levene, and for my first record of staying up for forty-eight hours straight with him. Thanks to Dr. George Barbastathis for the discussions on many research projects, and for his fondness of every Chinese restaurant I picked. I would also like to thank Ernest Chuang for

being a friend and two–year roommate, and for the many discussions in our kitchen. Special thanks to Dr. Fai Mok. Without his milestone experiment and great idea which started the HRAM project, I probably would not be in this position to say this many thank you's. I thank Dr. Annette Grot for helping us in making the random–phase plate we used in the big experiment. Many thanks to Wenhai Liu and George Panotopoulos for their collaboration on my latest research project. I am very grateful to Yayun Liu for her help and technical support in all my work, and to Lucinda Acosta for the birthday cakes, and for making the group a warm family.

I would like to thank all the other friends and colleagues in the group, past and present, for their help: Dr. Kelvin Wagner, Dr. Yong Qiao, Dr. Kevin Curtis, Dr. Chuan Xie, Dr. Jiafu Luo, Dr. David Marks, Dr. Jean–Jacques Drolet, Robert Denkewalter, Greg Billock, Shaw Wang, Ali Adibi, Greg Steckman, Christophe Moser, Yunping Yang, Jose Mumbru, Dimitris Sakellariou. I also thank Helen Carrier, Lyn Hein and David Seiving.

I could never say thank you enough to my parents, for their love, support, and encouragement. They are always a big part of my life, and everything that I have accomplished. I do not know how to write about their impact on me because no matter what I say, it would always be an understatement. I am also greatly indebted to my wonderful younger sister, who fills in my position and takes care of things which I am supposed to take care of as the elder son. I thank my grandparents, especially my late grandfather who would have been the happiest person to know that I would get a PhD degree from Caltech. Many thanks to my uncles and aunts who have always treated me as if I were their own son.

Finally, I would like to thank my loving wife, Xiaolei, for being a wonderful and understanding spouse, friend, and companion. I would not say that I did all this for her, since she would have probably preferred that I spent more time with her, but it always feels great to know that I can go back home to her after a day's work.

Abstract

The organization of this thesis very much coincides with the progress and "evolution" of my work here at Caltech. I came in at a time when our research activities entered a phase of full-scale realization of holographic storage after theoretical model building and proof-of-principle investigations were successfully carried out. More and more efforts have been dedicated to various experimental examination and demonstrations. Therefore, the major portion of this thesis details results from these efforts. Theoretical treatments are developed when necessary and only serve as the starting point and justification of further, serious experimentation.

The centerpiece of my work described in this thesis is a large-scale fast random-access holographic memory using $LiNbO_3$:Fe. It is not only our first attempt to incorporate many important disciplines and understandings of holographic memory design, engineering and experimentation to construct a real working prototype, but also a tool and test-bed for later characterization and investigation of other aspects in the application of holographic storage technology. This system is described in detail in Chapter 2 and 3 and has been used repeatedly and extensively through all of our works. Another two threads in the work presented here are the use of the M/# as a system metric to evaluate the dynamic range limitations, and of the Signal-to-Noise Ratio (SNR) and Bit-Error Rate (BER) to characterize the system error performance.

In Chapter 2, the design of a large-scale random-access holographic memory using $LiNbO_3$:Fe is discussed in detail. High, system dynamic-range-limited storage capacity is demonstrated by using angle, fractal and spatial multiplexing with a key custom-designed component—the segmented mirror array. The SNR and BER obtained from the reconstructed information are comparable to those of conventional CD-ROMs.

Fast random access to the memory contents is materialized in a separate system using an acousto-optic deflector (AOD) as the addressing device and an electro-optic

modulator (EOM) to compensate for the Doppler shift. Chapter 3 discusses the design issues and presents experimental demonstration of holographic storage using the system. The design and application of an optical phase–lock loop using the AOD and EOM for phase stabilization are also described at the end of this chapter.

Chapter 4 and 5 address two methods of thermal fixing to solve the volatility problem in holographic memories using photorefractive materials. First, "Low–High–Low" fixing is described in Chapter 4, along with the characterization of system error performance of non–volatile holographic storage using thermal fixing. A novel "incremental fixing schedule" is introduced to improve the system fixing efficiency. Experimental demonstration of a large–scale non–volatile memory with good error performance is also presented.

Chapter 5 shows theoretical treatment and experimental demonstration of high–temperature recording in $LiNbO_3$:Fe. Different charge transport mechanisms and their influence on the dynamics of holographic recording as well as the system dynamic range are discussed in detail. The two thermal fixing methods are examined and compared in terms of the M/#.

In Chapter 6, a very important holographic noise source, the inter–pixel grating noise, is evaluated theoretically based on a linear (small–signal) model, followed by experimental investigation of its influence on the system error performance of a large–scale memory. Random–phase modulation in the signal beam is discussed and demonstrated as an effective way to suppress this holographic noise.

Advisor: Professor Demetri Psaltis

Contents

Acknowledgements iv

Abstract vi

1 Introduction 1

1.1 Holography 1

1.2 Volume holographic storage: early works 2

1.3 Volume holographic storage: recent interest 3

1.4 Outline of thesis 5

2 Large–scale holographic memory: Design and Experiment 7

2.1 Single–location storage 8

2.2 Multiple–location storage 17

2.3 Segmented mirror array 18

2.4 Image–plane vs. Fourier–transform–plane storage 21

2.5 Memory system design 22

2.6 Storage of 10,000 holograms at a single location 24

2.7 Demonstration of the 160,000 hologram system 31

2.8 Appendix: Complications in the M/# 42

3 Fast random–access holographic memory 50

3.1 Readout using AOD 51

3.2 Recording using AOD and EOM 57

3.3 Single–AOD setup 67

3.4 Optical Phase–Lock Loop (OPLL) 70

4 Thermal fixing of 10,000 holograms 78

4.1 Theoretical model 80

4.2 Application of thermal fixing in large-scale memory 92
4.3 Incremental fixing schedule 96
4.4 Miscellaneous topics 107
4.5 Appendix: Derivation of the theoretical model 112

5 High-temperature recording in $LiNbO_3$:Fe **117**
5.1 Theoretical model 119
5.2 Experiment results 141
5.3 Miscellaneous topics 151
5.4 Appendix A: Derivation of the extended Kukhtarev equations 157
5.5 Appendix B: Derivation of the simplified model 160

6 Inter-pixel grating noise **162**
6.1 What is inter-pixel grating noise? 164
6.2 Why is inter-pixel grating noise important? 165
6.3 Temporal development of inter-pixel gratings 171
6.4 Formation of inter-pixel grating noise 177
6.5 Random-phase modulation in the signal 191
6.6 Experiment results 194

Bibliography **204**

List of Tables

4.1 Material parameters of $LiNbO_3$. 86
4.2 Other parameters used in numerical evaluation. 87

5.1 The writing slope, erasure time constant, and the resulting M/# of holographic storage at different temperatures. 148
5.2 The coefficients for **c**–axis expansion in congruent $LiNbO_3$. 153
5.3 The coefficients for **a**–axis expansion in congruent $LiNbO_3$. 153

List of Figures

2.1 Angle definition and K–space description. 10
2.2 Holographic storage in the 90° geometry. 11
2.3 Measured angle selectivity in 90° geometry. 11
2.4 K–sphere description of angle and fractal multiplexing: (a) angle multiplexing; (b) fractal multiplexing. 13
2.5 Space–Bandwidth Product (SBP) loading. 14
2.6 Exposure schedule. 15
2.7 One–dimensional illustration of the 4F system for angle deflection. . . 18
2.8 Schematic of the segmented mirror array. 19
2.9 Operation of the segmented mirror array. 20
2.10 Experimental setup of the large–scale memory. 23
2.11 Sample reconstructions from the 10,000 Fresnel–plane holograms. . . 27
2.12 Normalization by the background illumination profile. 28
2.13 Evolution of the system SNR as a function of the number of holograms. 29
2.14 Accessibility of the fractal rows by the reference beam. 33
2.15 Images showing lines from random phase plate: (a) image through crystal; (b) one of 10,000 holograms at top location. 34
2.16 Set of 60 × 50 pixel regions used to avoid random phase plate lines. . 34
2.17 Horizontal scanning of the signal beam. 35
2.18 Effects of the horizontal scanning. 35
2.19 Sample reconstructions from 10,000 Fourier–transform–plane holograms at the center location (#9). 38
2.20 Sample reconstructions from 10,000 Fourier–transform–plane holograms at the bottom location (#1). 39
2.21 Sample reconstructions from 10,000 Fourier–transform–plane holograms at the top location (#16). 40

2.22 Analysis of the sample reconstructions from Figure 2.19—2.21. 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. 41
2.23 Sample reconstructions from the 1,000 holograms at each of the 16 storage locations 43
2.24 SNR as a function of hologram location. 44
2.25 Time evolution of the space–charge field under different conditions. . 48

3.1 K–space description of the interaction between light and sound. . . . 52
3.2 4F system imaging the reference beam onto the crystal. 54
3.3 Angle selectivity curve of a 90°–geometry hologram readout by AOD. 54
3.4 Diffraction efficiency of the AOD as a function of the RF frequency. . 57
3.5 Operation of an EOM. 59
3.6 Experimental setup of a holographic storage system using AOD and EOM. 59
3.7 Optimizing the performance of the EOM. 62
3.8 Diffraction efficiency of the sampled holograms. 64
3.9 Sample reconstructions from the 1,000 holograms recorded by AOD and EOM. 65
3.10 Histograms from one sample reconstruction. 66
3.11 Large–scale random–access holographic memory desgin using AOD and EOM. 66
3.12 Single–AOD setup. 68
3.13 Angle selectivity in the single–AOD setup. 68
3.14 Diffraction efficiency of 300 holograms stored in the Single–AOD setup. 69
3.15 Phase modulation by the AOD. 70
3.16 Phase stabilization system using AOD and EOM: PD: photodector; PS: RF phase shifter. 72
3.17 Processing in the computer–based PLL. 73

3.18 Experiment with optical PLL: (a) with phase stabilization; (b) without phase stabilization. 74

3.19 Coherent addition and subtraction. 76

3.20 Improved PLL system design. 77

4.1 Schematic of thermal fixing process. 81

4.2 Degree of compensation as a function of proton density. 87

4.3 The M/# as a function of the absorption coefficient of the material: (a) the M/# of the original system; (b) the M/# of thermally fixed system; (c) thermal fixing efficiency. 91

4.4 Experimental setup for thermal fixing. 92

4.5 SNR degradation with thermal fixing. 94

4.6 Means and variances of the ON and OFF signals as a function of the number of holograms. 95

4.7 Holographic recording and erasure behavior under different conditions. 99

4.8 An example of the incremental fixing schedule. 102

4.9 System fixing efficiency as a function of the number of heating treatments. 103

4.10 Sample reconstructions from the 10,000 fixed holograms using incremental fixing schedule. 105

4.11 Characteristics of the sample reconstructions from the 10,000 fixed holograms. 106

4.12 η_{fixing} and $|1-\Delta|^2$ as a function of the oxidation/reduction state of the material. 109

5.1 The absorption as a function of the temperature. 125

5.2 The dynamics of holographic recording and erasure at high temperature ($150°C$): (a) electronic grating N_{D1}^{+}; (b) protonic grating n_{p1}; (c) space-charge field E_1. 127

5.3 Writing/erasure asymmetry of a transmission-geometry hologram ($\theta = 30°$) written at room temperature. 130

5.4 Behavior of the electronic and protonic grating during erasure at elevated temperature (150°C): (a) electronic grating N_{D1}^{+}; (b) protonic grating n_{p1}; (c) space–charge field E_1. 132

5.5 Comparison between the simplified model and the complete treatment (at 150°C): (a) transmission geometry ($\theta = 30°$); (b) 90° geometry. . 134

5.6 Comparison between (a) 90° and (b) transmission–geometry recording. 136

5.7 Normalized M/# as a function of the grating vector in high–temperature recording (150°C). 137

5.8 The ratio of the M/# in high–temperature (150°C) and room–temperature recording as a function of the grating vector. 139

5.9 Dynamics of the electronic charge grating N_{D1}^{+} at (a) room temperature, (b) 120°C, and (c) 150°C (transmission geometry $\theta = 30°$). . . . 140

5.10 The ratio of the M/# in high–temperature and room–temperature recording as a function of the temperature in transmission–geometry ($\theta = 30°$). 140

5.11 Experimental setup for the M/# measurement at elevated temperatures. 142

5.12 Angle selectivity curve of an unfixed and a thermally fixed hologram. 143

5.13 Writing/erasure dynamics at high temperature. 144

5.14 Another example of writing/erasure dynamics at high temperature. . 145

5.15 A typical writing curve at 150°C. 147

5.16 A typical erasure curve at 150°C. 147

5.17 Photo of the experimental setup for 600–hologram storage. 150

5.18 Diffraction efficiency of the 600 fixed holograms. 151

5.19 Bragg shift in the selectivity curves due to **c**–axis contraction. 155

5.20 K–space description of phase–mismatching due to thermal contraction along the **a**-axis. 155

5.21 Reconstruction of one hologram written at 150°C. 157

6.1 K–space description of inter–pixel grating noise. 165

6.2 SNR degradation as a function of the number of holograms. 167

6.3 Mean of ON/OFF pixels as a function of reference beam scanning: (a) a "blank" hologram and its neighbor from 10,000 stored holograms; (b) a single hologram. 168
6.4 Strong two–beam coupling effect. 170
6.5 Recording of a 90°–geometry hologram and inter–pixel gratings (a) 90°–geometry, (b) transmission–geometry $\theta_{out} = 0.1°$, (c) transmission–geometry $\theta_{out} = 1°$. 172
6.6 Different recording and erasure behaviors of (a) an information grating and (b)(c) inter–pixel gratings during multiple hologram recording. . 173
6.7 Ratio of the diffraction efficiencies of the inter–pixel grating $\vec{K}_{ij}$ and the information grating $\vec{K}_{ri}$ as a function of the number of holograms recorded with the same schedule. 176
6.8 Signal cone in Fourier–transform–plane recording. 178
6.9 Signal plane in Fourier–transform–plane recording. 178
6.10 Links inducing crosstalk between a pair of pixels separated by horizontal and vertical distances d_x, d_y. 179
6.11 The number of links of every pixel in a 100×100 pixel image. 181
6.12 The number of "links" of the center pixel as a function of the number of pixels N in one dimension on the image. 182
6.13 The grating vector of the inter–pixel grating. 183
6.14 The "contribution" the pixel at (x, y) makes to the noise buildup at the center pixel of a 100×100 image: (a) worst case where all the inter–pixel gratings are of the same strength; (b) variations in the grating strength are considered. 187
6.15 The normalized number of "links" of the center pixel as a function of the number of pixels N in one dimension on the image. 188
6.16 "Connections" arising from wide Bragg selectivity. 189
6.17 The rediffraction by pixel i via noise grating $\vec{K}_g$. 190
6.18 Random–phase modulation to reduce the strong two–beam coupling effect. 193

6.19 Experimental setup for the plane–wave signal experiment. 194

6.20 Liquid Crystal Phase Retarder operating in intensity–modulation mode: (a) in between crossed polarizers; (b) in between parallel polarizers. . 196

6.21 Ratio of the diffraction efficiencies of the inter–pixel grating and the hologram as a function of the number of holograms (a) without phase modulation; (b) random–phase modulation all the time; (c) π phase shift only when the two beams were ON at the same time. 197

6.22 Experimental setup for random–bit patterns. 198

6.23 SNR as a function of the number of holograms: (a) without random–phase modulation; (b) with random–phase modulation. 202

6.24 OFF signal as a function of the number of holograms: (a) without random–phase modulation; (b) with random–phase modulation. . . . 202

Chapter 1 Introduction

Contents

1.1 **Holography** 1

1.2 **Volume holographic storage: early works** 2

1.3 **Volume holographic storage: recent interest** 3

1.4 **Outline of thesis** 5

1.1 Holography

Holography dates back to 1948, when Dennis Gabor developed the theory of wavefront reconstruction in an attempt to improve electron microscopy [1, 2]. When a coherent reference beam interferes with a signal beam, the intensity of the interference pattern contains both the amplitude and phase information of the beams. Since photographic materials such as film record the intensity of the exposing light, this method allows the recording of the whole information of the signal beam. Gabor coined the term hologram from the Greek words holos (meaning whole) and gramma (meaning message). The 1971 Nobel Prize in Physics was awarded to him for this invention.

Further progress in the field was stymied during the next decade because light sources available at the time were not truly coherent. This barrier was overcome in 1960 with the invention of the laser. In 1962 Leith and Upatnieks recognized from their work in side–looking radar that holography could be used as a 3–D visual medium. They improved upon Gabor's original idea by using a laser and an off–axis technique [3]. The result was the first laser transmission hologram of 3–D objects with good reconstruction quality. The basic off–axis technique of Leith and Upatnieks is

still the staple of holographic methodology. With this breakthrough in recording technique came new applications besides the original application of imaging. For example, 3–D image reconstruction with laser [4] or white light [5], information storage [6–8], filtering [9], aberration correction in imaging systems [10, 11], and information processing in synthetic–aperture radar [12] were all proposed and demonstrated during the 1960's and early 1970's.

The material used in most early experimental work was photographic film. The photographic film is typically a thin, two–dimensional (2–D) layer of recording material. It is sensitive and has good spatial frequency response. Thin holograms lack sensitivity to changes in the angle of the reconstruction beam. This enables the construction of shift–invariant filters for correlation and information processing. However, thin film is not a good storage material because it requires developing between recording, and readout and restricts the capacity to approximately one hologram per spatial location on the film.

1.2 Volume holographic storage: early works

The limited capacity of 2–D holographic storage can be overcome by using volume holography. In volume holography, the recording medium is a three–dimensional (3–D) solid. Because the interference pattern is recorded throughout the storage material, thick holograms are sensitive to changes in the propagation properties of the readout beam. This effect is similar to the phenomenon first observed in X–ray diffraction from crystals by Bragg in 1942 [13]. Kogelnik [14] analyzed the diffraction of light from a hologram written in a volume by two interfering plane waves. His result showed how Bragg selectivity applied to thick absorption or phase holograms. Later analysis [15–18] discussed Bragg selectivity and diffraction from materials whose thickness is between the thick and thin regimes. Using the Bragg selectivity inherent to volume holography, multiple holograms can be stored and retrieved independently in the same volume by changing the propagation properties of the readout beam. As a result, the information storage capacity is greatly increased by volume holography.

A very important type of material for volume holography is photorefractive crystal. Photorefractive effect was discovered in ferroelectric inorganic crystals, by researchers attempting to perform second–harmonic generation. They observed that the index of refraction would change as lithium niobate crystal absorbed light [19]. This effect was first called "optical damage" since the index change distorted the optical beams. Photorefractive effect generally occurs in three steps:

- Upon illumination, the electrons (holes) are excited from the donors into the conduction (valence) band and become free carriers.
- The free carriers migrate under different charge transport mechanisms such as diffusion, drift, and/or photovoltaic effect, followed by re–trapping by the acceptors (traps). This leads to charge separation which generates an electric field— the space–charge field.
- The space–charge field modulates the refraction index of the material through the linear (Pockel's) electro–optic effect, to form phase holograms.

Many theoretical models for the behavior of the photorefractive effect were introduced in 1970's [20–34], but the Kukhtarev equations have been universally accepted as the general model for the operation of the photorefractive effect [27,34].

Chen et al. [35] first realized the potential of this effect in translating the information embedded in the hologram to the index modulation and proposed volume holographic storage using photorefractive materials such as $LiNbO_3$. Volume holographic storage in photorefractive materials was demonstrated in the late 1960's and early 1970's [35–38]. However, because of the lack of supporting optical devices (lasers, spatial light modulators (SLM), 2–D detector arrays, beam deflectors), these works did not lead to a commercial product at the time.

1.3 Volume holographic storage: recent interest

Holographic storage has undergone resurgence in recent years mainly because the dramatic advances in electronic computers lead to ever–increasing demand for in-

expensive mass storage devices. Processor speeds for supercomputers are moving beyond the billion instructions per second range. With this high performance, the data access time and data transfer rate between CPU and primary memory, and between primary and secondary memories, need to improve significantly. In addition, the storage capacity of memory devices has to be greatly increased since data accumulate quickly with ultra–high–speed computation and new applications such as multimedia have emerged. However, there exists an imparity between the improvement of computation and memory devices. For instance, over the past decade, disk drive performance has improved by a factor of less than 10. In comparison, CPU performance has increased by a factor of 1,000. Therefore, a 3–D memory technology which offers very high storage capacity and 2–D parallel access at the same time, such as volume holographic storage, is very much in demand.

In addition to the increased demand and regained interest, the modern opto–electronic technology has provided us with a rich variety of compact, inexpensive, and high performance semiconductor laser diodes, liquid crystal SLMs, CCD detector arrays, as well as computer–controlled apparatus such as precision beam scanners. As a result, there is no technical impediment to making holographic memory systems.

All of these, however, have not materialized volume holographic memories as practical mass storage systems. Lack of "perfect" storage material and pessimistic predictions on the noise–limited storage capacity had once cast a shadow on the future of this new technology until Mok et al. successfully demonstrated the storage and high–fidelity recall of 5,000 holograms in $LiNbO_3$ [39]. Since then, a methodology has been established to characterize the performance of holographic memory using photorefractive materials [40], hologram multiplexing has been extensively studied, better understood and utilized, new multiplexing techniques have been invented, and new storage materials have been found.

Whether or not holographic storage will become commercialized in the near future is still a question. The most important factors are the expense and novelty of the technology, as well as market niche when it is faced with competition from mature storage technologies such as magnetic, semiconductor, and conventional optical

memories. This is the reason that the main research focus has to be shifted to the system–level investigation and the experimental approach has to be more application–oriented than purely theoretical. And I believe that this is exactly what our group has been doing and has accomplished in the past several years.

1.4 Outline of thesis

The organization of this thesis very much coincides with the progress and "evolution" of my work here at Caltech. Since I came in at a time when we were undergoing a transition from theoretical modeling to full–scale realization of holographic storage, the major portion of the thesis details results of various experimental investigations and demonstrations. Theoretical treatments are developed when necessary and only serve as the starting point and justification of further, serious experimentation.

The centerpiece of my work described in this thesis is a large–scale fast random–access holographic memory using $LiNbO_3$:Fe. It is not only our first attempt to incorporate many important disciplines of holographic memory design, engineering and experimentation to construct a real working system, but also a tool and test–bed for later investigation of other aspects in the application of holographic storage technology. This system is described in detail in Chapter 2 and 3 and has been used repeatedly and extensively through all of our works. In Chapter 2, high, system dynamic–range–limited storage capacity is demonstrated by using angle, fractal and spatial multiplexing with a key custom–designed component—the segmented mirror array. Chapter 3 discusses the design and realization of a fast random–access memory system using an acousto–optic deflector (AOD) as the addressing device and an electro–optic modulator (EOM) to compensate for the Doppler shift.

Chapter 4 and 5 address two methods of thermal fixing to solve the volatility problem in holographic memories using photorefractive materials. First, "Low–High–Low" fixing is described in Chapter 4, along with the characterization of system error performance of non–volatile holographic storage using thermal fixing. A novel "incremental fixing schedule" is introduced to improve the system fixing efficiency.

Experimental demonstration of a large-scale non-volatile memory with good error performance is also presented. Chapter 5 shows theoretical treatment and experimental demonstration of high-temperature recording in $LiNbO_3$:Fe. Different charge transport mechanisms and their influence on the dynamics of holographic recording as well as the system dynamic range are discussed in detail. The two fixing methods are examined and compared in terms of the M/#.

In Chapter 6, a very important holographic noise source, the inter-pixel grating noise, is evaluated theoretically based on a linear (small-signal) model, followed by experimental investigation of its influence on the system error performance of a large-scale memory. Random-phase modulation in the signal beam is discussed and demonstrated as an effective way to suppress this holographic noise.

Chapter 2 Large–scale holographic memory: Design and Experiment

Contents

2.1 Single–location storage 8

2.1.1 Multiplexing techniques and geometric limitations 8

2.1.2 Dynamic–range limitations and M/# 13

2.2 Multiple–location storage 17

2.3 Segmented mirror array 19

2.4 Image–plane vs. Fourier–transform–plane storage 21

2.5 Memory system design 22

2.6 Storage of 10,000 holograms at a single location 24

2.6.1 10,000 Fresnel–plane holograms 24

2.6.2 Characterization of system error performance 26

2.7 Demonstration of the 160,000 hologram system 31

2.7.1 The 160,000 hologram system 31

2.7.2 30,000 Fourier–transform–plane holograms at 3 locations . . 36

2.7.3 160,000 Fourier–transform–plane holograms at each of the 16 locations 42

2.8 Appendix: Complications in the M/# 42

Full–scale realization of a holographic storage system has always been a great challenge since the concept of volume holographic memory was first conceived 30 years ago. Lack of appropriate recording material and pessimistic predictions on the storage capacity and error performance had once cast a shadow on the development of this new technology until Mok et al. successfully demonstrated the storage of 5,000

holograms in a $LiNbO_3$:Fe crystal [39]. Although some of the problems still exist (for example, the "perfect" material has not been found yet), it is essential to explore the potentials of the technology on the system level, to establish systematic evaluation methods, and to set up a test–bed to characterize the system performance. It is for this reason that we set out to construct a large–scale holographic storage system and conducted series of experiments on it. After all, our philosophy is **"If you don't do it, you don't really know what the problems are, and you don't really know how to solve them!"**

In this chapter, I describe a large–scale random–access holographic memory using $LiNbO_3$:Fe crystals. I address most of the issues associated with the implementation of holographic memories to achieve very high storage capacities. In Section 2.1—2.5, I discuss the issues related to the system design in detail. In the remainder of this chapter, I present the experimental results of the memory system for large–scale storage.

2.1 Single–location storage

Holograms are the recording of the interference pattern created by mixing two coherent laser beams—the signal beam which carries the information to be stored and the reference beam. The information is embedded in the interference pattern and distributed in a three–dimensional volume. To retrieve the information, the same reference beam is used. The reconstruction is a result of the summation (space integral) of the diffraction off the entire volume (i.e., the interaction region). A change in the propagation properties of the reference beam results in so–called Bragg match or mismatch—constructive or destructive interference similar to what happens to the reflection and transmission of the light on multi–layer–dielectric–film interface. When Bragg mismatch takes place, a new hologram could be stored in the same volume, with no (or minimal) crosstalk from the others. This is the basic concept of single–location storage.

There are two important questions regarding single–location storage. The first

is how multiple holograms can be superimposed in the same volume. This will be addressed in Section 2.1.1, in the context of multiplexing techniques. The second is how many holograms we can superimpose in the same volume. This is the topic of Section 2.1.2, where the dynamic–range limitations of a holographic memory system are discussed.

2.1.1 Multiplexing techniques and geometric limitations

In this section, we discuss some of the available methods for hologram multiplexing—creation of reference beam wave–front for storage and independent retrieval of information. These methods can be divided into two general classes [41],

- Bragg–matched readout—crosstalk suppression by Bragg–mismatch. The techniques include angle [42], wavelength [43–47], phase code [48–54], and shift [55] multiplexing.

- Displaced readout—more than one hologram can be accessed by the same reference beam. The reconstructions do not overlap in the spatial frequency domian, so the undesired ones can be eliminated by spatial filtering. The techniques include fractal [56, 57] and peristrophic [58] multiplexing.

Not all the multiplexing techniques can be employed at the same time. Some require complicated hardware. In the discussion below, we only describe the ones we used in our large–scale holographic memory.

Angle multiplexing and 90° geometry The most common technique to store multiple holograms in the same volume within a thick recording medium is angle multiplexing. In angle multiplexing, a discrete set of reference plane waves which vary in illumination angle is used. The wave–vectors for all of these beams lie in a single plane (the interaction plane) which includes the central wave vector of the signal beam. 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

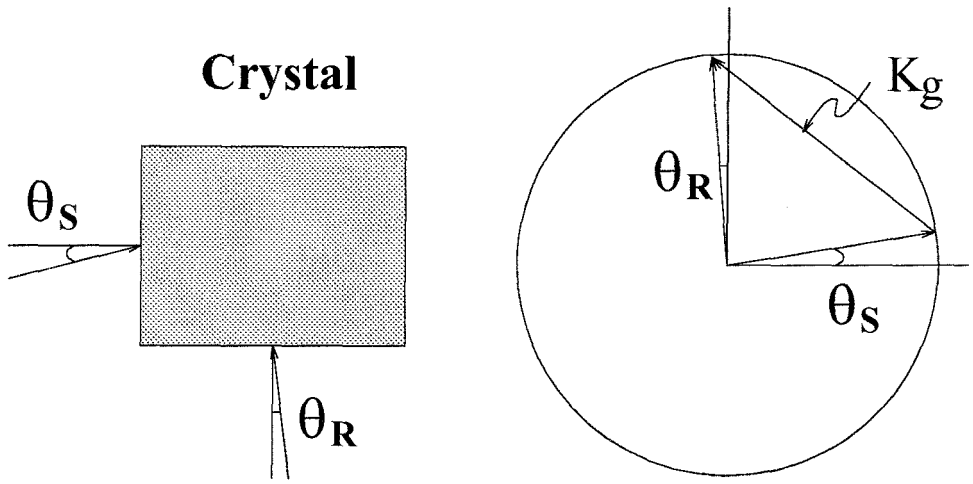

Figure 2.1: Angle definition and K–space description.

to this as the selectivity function. When the readout beam is identical to the wave–front used for recording, the diffracted signal is at a central peak in the selectivity function and is "Bragg–matched." 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 [59]

$$\Delta\theta = \frac{\lambda}{L}\frac{\cos\theta_S}{\cos(\theta_R - \theta_S)}, \tag{2.1}$$

where θ_S and θ_R are the incidence angles of the signal and reference beams, and L is the thickness of the crystal (the interaction length) as shown in Figure 2.1.

For holograms formed by plane–wave signal (no information) and reference beams, a change of $\Delta\theta$ in the reference beam yields a "fresh" angle where a new hologram can be stored without crosstalk from the others. However, for information–bearing signals which contain multiple spatial frequencies (different θ_S), the Bragg–mismatch condition ($\Delta\theta$) for each spatial frequency is different. As a result, all the spatial frequencies in the signal can not be suppressed at the same time, leading to crosstalk between holograms [60–67].

For maximal storage capacity with minimal crosstalk within a given angular range, small angular selectivity is very much preferred. The 90° geometry, shown in Fig-

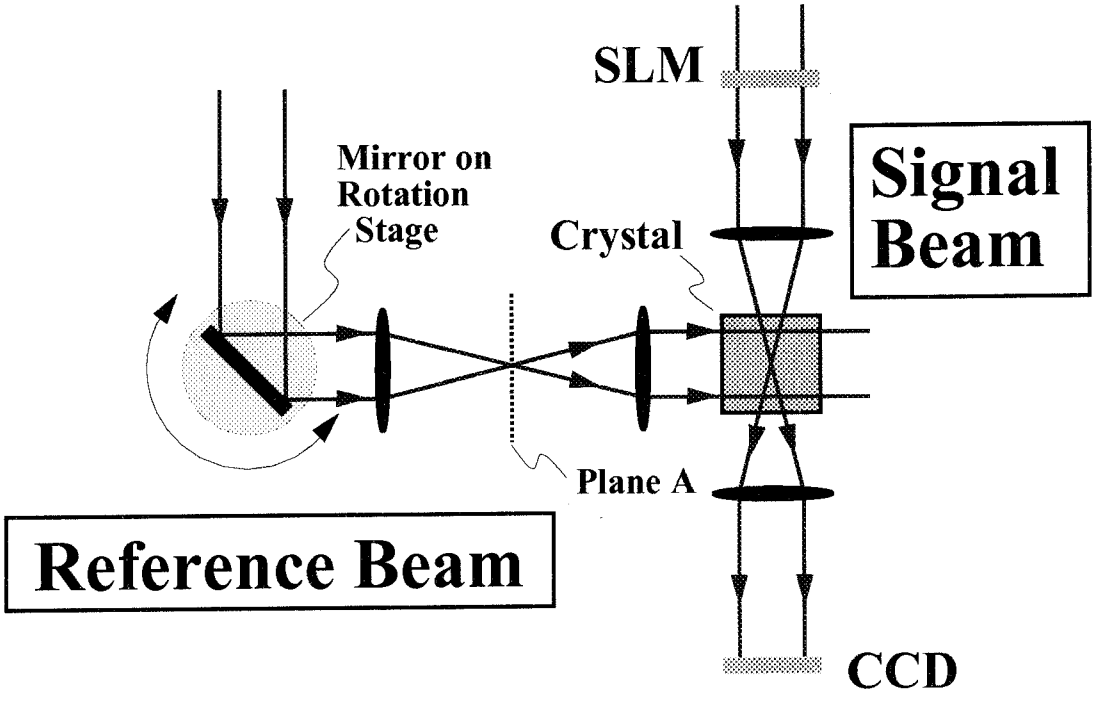

Figure 2.2: Holographic storage in the 90° geometry.

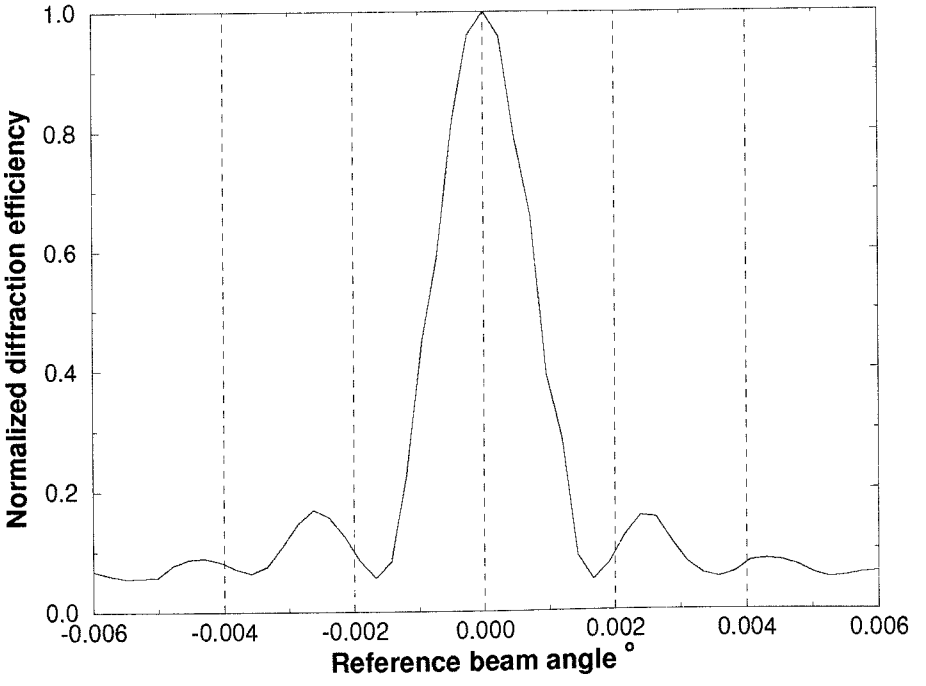

Figure 2.3: Measured angle selectivity in 90° geometry.

ure 2.2, where the signal and reference beam enter the photorefractive crystal through orthogonal faces, has the strongest angular selectivity (λ/L), thus allowing optimal use of the angular scanning range of the reference beam. Figure 2.3 shows a measured angular selectivity curve in the 90° geometry for a crystal of thickness $L = 20$ mm. The measured Bragg selectivity is 0.0016°, which is slightly larger than the theoretical value $\lambda/L = 0.0014°$ due to absorption. All angles listed here are external angles. In the 90° geometry, the interaction length can be increased simply by extending the thickness of the crystal, increasing Bragg selectivity as well as diffraction efficiency [68,69]. Other advantages include reduced Fresnel loss as a result of orthogonal entry of the light. Noise collected by the detector due to bulk scattering, which tend to cluster along the direction of the reference beam, is also strongly suppressed.

Geometric limitations and fractal multiplexing The number of holograms that can be stored at a single location by angle multiplexing is determined by the total number of distinct reference beam angles the optical system can deliver. It is often limited by the Space–Bandwidth Product (SBP) of the angle scanner, especially in the case of non–mechanical devices such as acousto–optic or liquid crystal deflectors. For example, current AOD technology provides SBP on the order of 1,000 or so. Therefore, storage of more than 1,000 angle–multiplexed holograms at the same location requires more than one SBP–limited scanner. If the angle deflector is capable of an arbitrarily large number of resolvable angles, then the geometric limitations is imposed by the numeric aperture of the lens system in the reference path. The entire angular range required for multiplexing is not simply a product of the number of holograms and a constant $\Delta\theta$. $\Delta\theta$ increases as the reference beam deviates from the optical axis as a result of less strong Bragg selectivity and worse plane–wave quality due to the aberrations in the optics. In addition, to reduce the crosstalk from neighboring holograms, we always prefer, as a rule of thumb, to set the inter–hologram spacing about five times the Bragg selectivity. Therefore, a large numeric aperture is required for the large range of angle deflection, which is very difficult, even impossible in many cases.

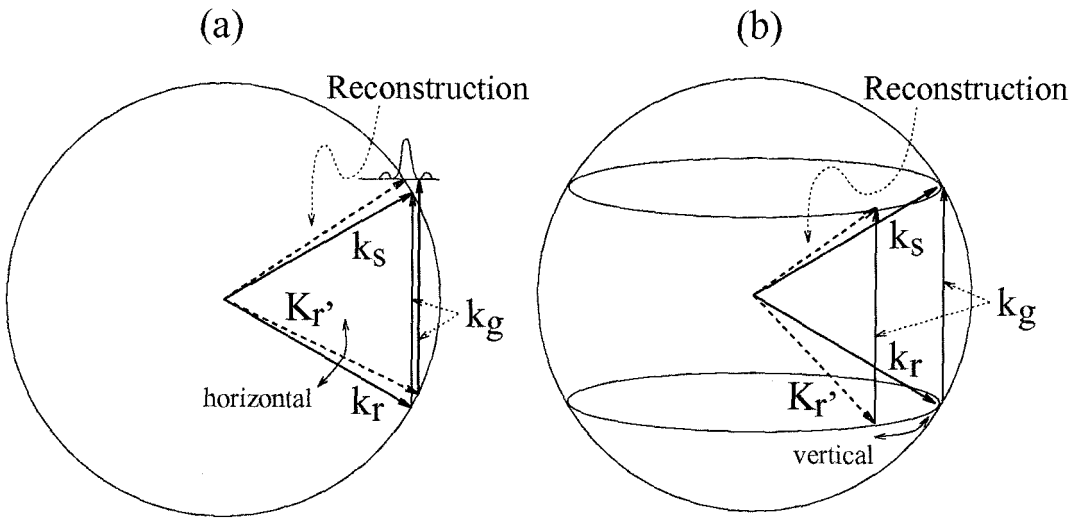

Figure 2.4: K–sphere description of angle and fractal multiplexing: (a) angle multiplexing; (b) fractal multiplexing.

It is important to notice that the Space–Bandwidth Product requirement for angle multiplexing is one dimensional (horizontal, in the interaction plane). In order to alleviate the SBP "overloading" in this dimension, we can take advantage of the 2–D nature of optics, and spread out the SBP loading in the vertical dimension by using another multiplexing technique: Fractal multiplexing.

In angle multiplexing, reference beam angles used for separate holograms must be distinct. This distinction is usually created by angle changes in the plane of interaction of the reference and signal beams (i.e., horizontally), due to Bragg selectivity as shown in Figure 2.4(a). Angle changes orthogonal to this plane of interaction (i.e., vertically) have little effect on the phase–buildup of the reconstructed hologram. This is shown in Figure 2.4(b). The reference beams which are deflected vertically can still readout the hologram [57]. However, the reconstructed holograms are tilted slightly from the original signal axis, by the same vertical angle deflection in the reference beam. In practice, the vertical spatial bandwidth of the signal is limited. Therefore, when the vertical deflection of the reconstruction becomes larger than the vertical bandwidth of the signal, no signal will be detected by the detector array centered on the signal axis. As a result, another new hologram can be recorded at this vertical angle with no crosstalk from the others. This is called fractal multiplexing, or out–of–plane

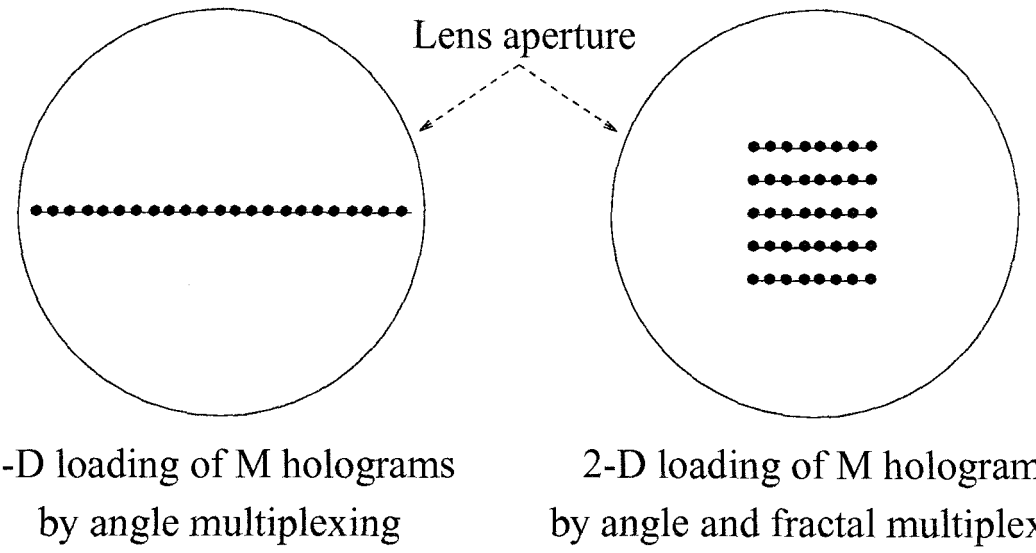

Figure 2.5: Space–Bandwidth Product (SBP) loading.

multiplexing.

With fractal multiplexing, multiple holograms can be divided into several "fractal rows" (vertical angles addressed by a "vertical" angle scanner), with each row containing a number of angularly multiplexed holograms (addressed by a "horizontal" angle scanner). This arrangement is described in Figure 2.5. It eases the SBP requirement on the horizontal angle scanner. Therefore, the storage and access of a large number of holograms can be achieved with commercial angle scanners and off–the–shelf optical components.

2.1.2 Dynamic–range limitations and M/#

Photorefractive crystals are widely used for large–scale holographic storage. During recording of a hologram, the crystal is exposed to a spatially–varying interference pattern created by the signal and reference beam. Charge carriers are excited into the conduction band by the illumination. They migrate by diffusion, drift, and photovoltaic effect, and then become trapped at new sites, leading to charge separation inside the material. The electric field set up by the charge grating modulates the index of refraction of the crystal via the electro–optic effect, giving rise to a phase

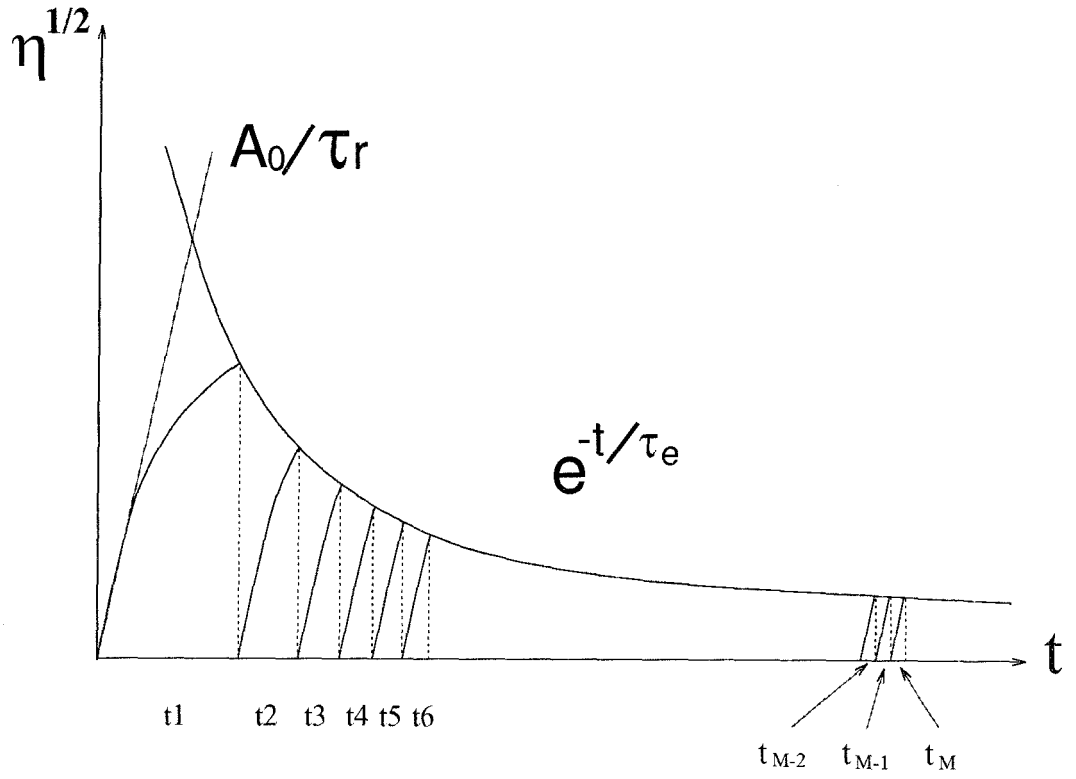

Figure 2.6: Exposure schedule.

hologram which is a copy of the exposing interference pattern. Upon readout, the same reference beam is used to reconstruct the signal, which is then collected by a 2D detector array to retrieve the information. The recording can be repeated for many different signals to store multiple holograms. However, the redistribution of space charges caused by repeated exposures erases the existing gratings, leading to the loss of previously stored information. This means that the storage of multiple holograms involves both recording and erasure behavior. Therefore, the dynamic range of a photorefractive holographic storage system is not the same as the material dynamic range of the photorefractive crystal since subsequent recording erases the previously stored holograms and in effect releases part of the dynamic range consumed. Limitations on diffraction efficiency of the holograms does not come from material saturation alone, and the measurement of saturation diffraction efficiency is not very reliable because of the non–linear effects such as two–beam coupling as well as the photovoltaic effect.

Because of the erasure during recording, an exposure schedule is required for

storing multiple holograms of equal diffraction efficiency [56]. The exposure schedule is a carefully chosen set of decreasing exposure times. Initial holograms start with a large diffraction efficiency, and are erased by the exact amount needed to match the diffraction efficiency of the final short exposure. This is shown in Figure 2.6. To calculate the proper set of exposure times, each hologram is assumed to evolve during recording as

$$h(t) = A_0(1 - e^{-t/\tau_r}) \approx \left(\frac{A_0}{\tau_r}\right) t, \tag{2.2}$$

and decay during erasure as

$$h(t) = h_0 e^{-t/\tau_e}, \tag{2.3}$$

where A_0 is the saturation strength of the hologram. τ_r is the characteristic recording time constant, and τ_e the characteristic erasure time. h_0 is the initial hologram strength at the beginning of erasure. Using a simple "backwards" recursion algorithm, the set of exposure times t_1, t_2, ..., t_{M-1} can be calculated from the last exposure time t_M, where M is the total number of holograms. For a large number of holograms M, under the assumption that the first hologram is exposed long enough to reach saturation, the final equalized diffraction efficiency is

$$\eta = \left[\left(\frac{A_0}{\tau_r}\right) t_M\right]^2 = \left[\left(\frac{A_0}{\tau_r}\right) \frac{\tau_e}{M}\right]^2. \tag{2.4}$$

We define M/# as [40,41,56]

$$M/\# = \left(\frac{A_0}{\tau_r}\right) \tau_e, \tag{2.5}$$

so that

$$\eta = \left[\frac{M/\#}{M}\right]^2. \tag{2.6}$$

The M/# is the constant of proportionality between the diffraction efficiency and the number of holograms squared. It is a system metric for a holographic memory, determined by both the material and system parameters. It predicts the highest possible diffraction efficiency that can be achieved for a given number of holograms. On the other hand, by taking into account the background scattering level and holographic noise, it gives a very accurate estimate of the storage capacity of a holographic storage system.

The M/# can be measured relatively easily from the recording and erasure behavior of a single hologram. The crystals we used in our memory system have M/#'s ranging from 1 to 2. If 10,000 holograms are to be stored in the system, the expected diffraction efficiency is on the order of 10^{-8}. However, the M/# reflects an ideal recording schedule, in which the first exposure time is infinitely long. In our case, in order to avoid the buildup of holographic noise due to long exposures, the first exposure time is always empirically set to about 10 seconds. It means, in practice, we can only achieve effectively a fraction of the M/#. In addition, the M/# drops with prolonged exposure due to the buildup of the DC screening field. As a result, the final diffraction efficiency of 10,000 holograms is expected to be on the order of 10^{-9}, which is several times (≈ 10) higher than the noise background in a fresh crystal in our experimental environment. This suggests that 10,000 holograms is a reasonable objective in our memory system due to the dynamic–range limitation.

2.2 Multiple–location storage

Angle and fractal multiplexing are the techniques to superimpose multiple holograms at the same spatial location—same volume in the crystal. The total number of holograms we can record by angle and fractal multiplexing is mainly determined by the system dynamic range, namely, the M/#. In order to further increase the storage capacity, spatial multiplexing can be used to store holograms at different locations on the crystal. Spatial multiplexing is not limited by the dynamic range but by the size of the signal transverse dimension, the recording medium and the numeric aperture

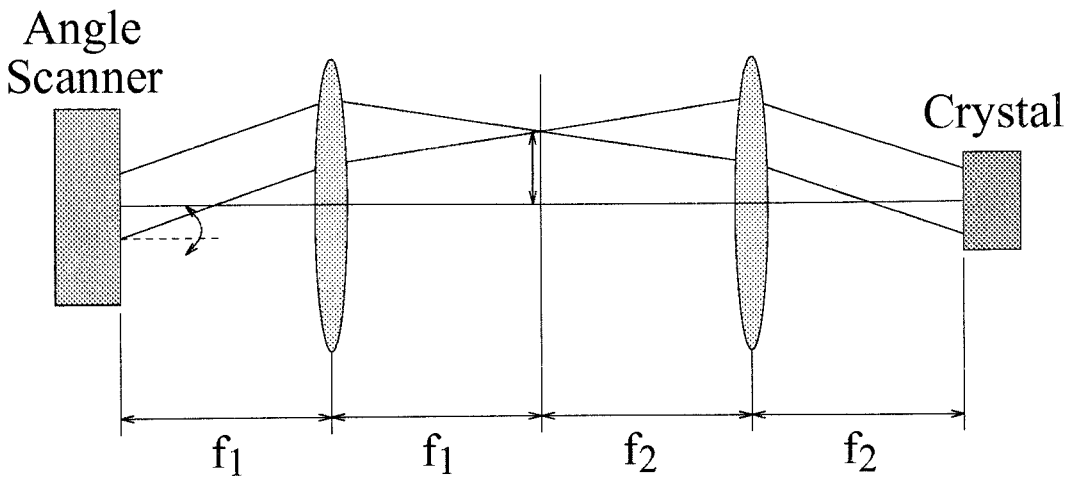

Figure 2.7: One–dimensional illustration of the 4F system for angle deflection.

of the optical system, especially of the signal arm.

2.3 Segmented mirror array

The segmented mirror array is the key element of our holographic storage system. It enables us to do angle, fractal, and spatial multiplexing with the same reference beam optics [68–70]. For angle and fractal multiplexing, an easy and widely used implementation is the 4F system shown in Figure 2.7. The 2–D angle scanner is placed in the front focal plane of the first lens and controls the 2–D movement of the focused reference beam in the center plane of the 4F system. In turn, this movement is converted to a 2–D angle deflection at the crystal for angle and fractal multiplexing. Note that an angle change in this center plane becomes a position change of the reference beam on the crystal for spatial multiplexing. Therefore, the 4F system can be "folded" by inserting a motorized mirror at the center plane. By controlling the vertical tilt of this mirror, we will be able to do angle and fractal multiplexing at different spatial locations in the crystal. Instead of the motorized mirror, we use a custom segmented mirror array at this center plane. The schematic of the mirror array is shown in Figure 2.8. It consists of 16 identical mirror strips. Each mirror strip is 75mm long, and each contains a set of 16 mirror facets with different vertical tilts

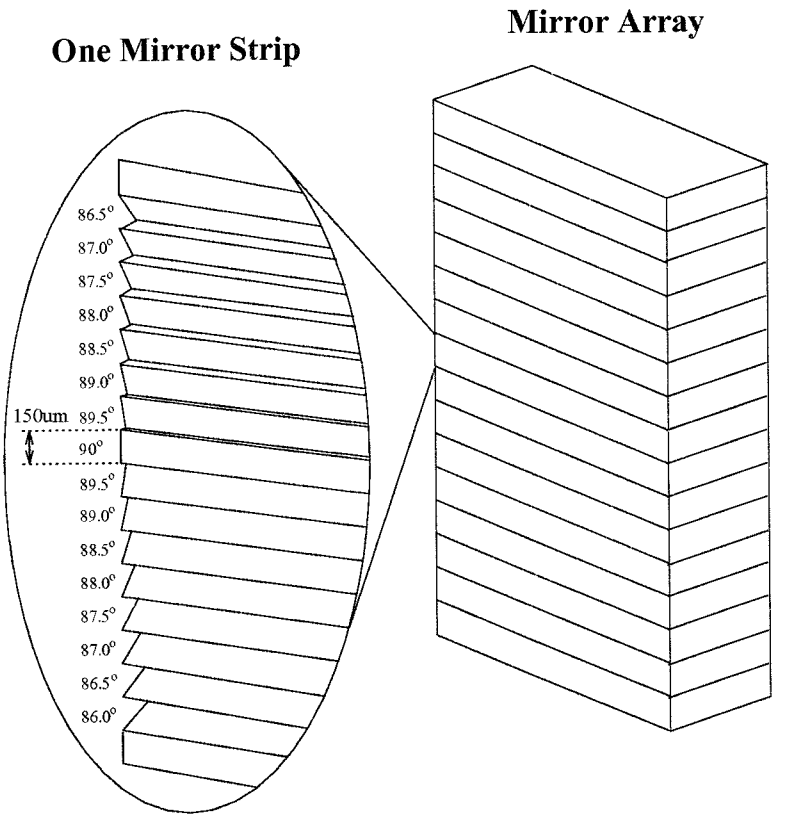

Figure 2.8: Schematic of the segmented mirror array.

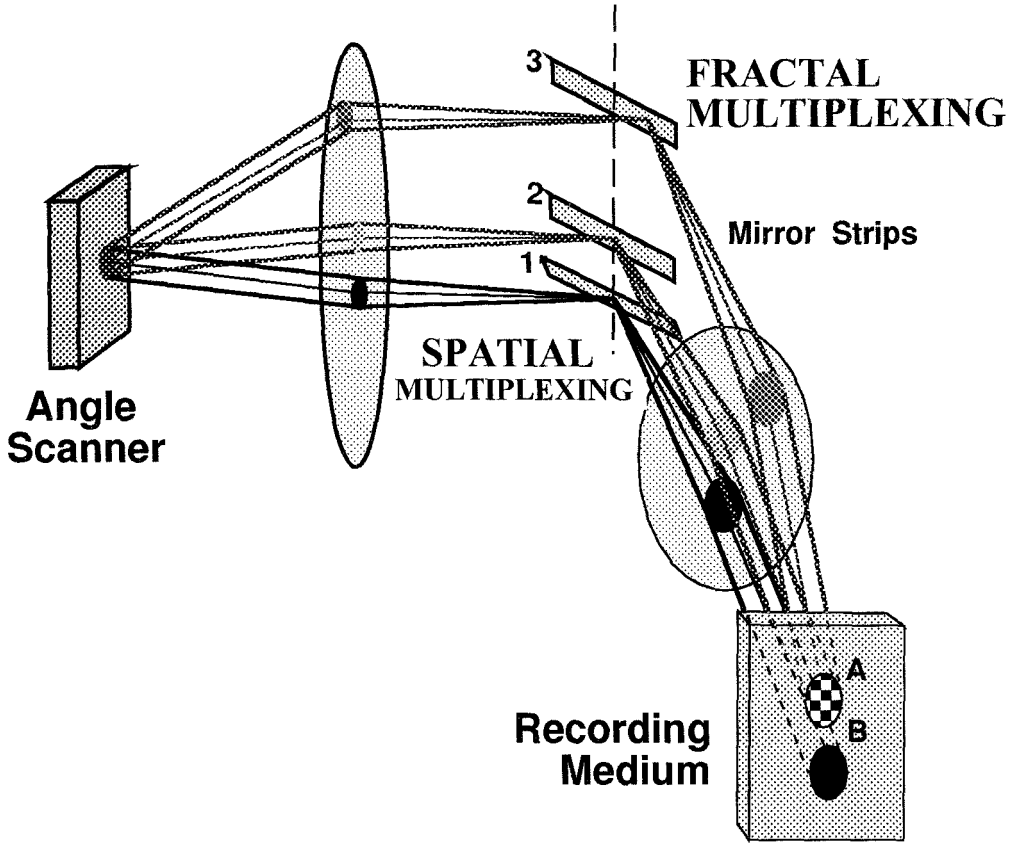

Figure 2.9: Operation of the segmented mirror array.

for spatial multiplexing at 16 locations. The angle change between two neighboring facets is 0.5°. Each mirror facet is 150μm wide. The width of an entire strip of 16 facets is 2.4mm. To scan the reference beam horizontally for angle multiplexing at a certain location, the horizontal angle scanner moves the focused reference beam along a mirror facet with the corresponding tilt angle. For fractal multiplexing, in order to have the reference beam deflected vertically while remaining at the same location on the crystal, the focal spot is moved to the mirror facet with the same vertical tilt on another strip. The vertical movement between these two strips creates the vertical angle change needed for fractal multiplexing. Therefore, the vertical angle scanning is not continuous but consists of a set of 16 discrete angles, which we refer to as the 16 fractal rows.

Figure 2.9 shows the typical operation of the mirror array. Mirror facet 1 and 2

perform spatial multiplexing at location A and B. Mirror facet 3 is parallel to facet 2, but displaced vertically (because it is on another mirror strip). Beams directed to facet 3 also arrive at storage location A, but with a different vertical incidence angle than those arriving via facet 2. If the vertical separation between facet 2 and 3 is larger than the vertical bandwidth of the images being stored, then each can be used to store and retrieve holograms. With the mirror array, the system is capable of storing holograms at 16 different spatial locations, with as many as 16 fractal rows at each location. Given that the number of holograms that can be stored at a single location is 10,000, the total storage capacity of the system is 160,000 holograms. Fast, random access to the stored holograms can be realized if non–mechanical angle scanners, such as AODs, are used.

2.4 Image–plane vs. Fourier–transform–plane storage

Since holography allows one to copy and recreate an arbitrary wavefront, it is possible to store either image–plane holograms or Fourier–transform–plane holograms [71]. Recording multiple holograms in the image plane is easier, because intensity tends to be more uniform. However, image–plane holograms are very susceptible to material imperfections since the information is localized. To the contrary, in Fourier–transform–plane recording, the information of each pixel is distributed throughout the recording volume. Therefore, localized burst errors (such as surface defects, dust particles on the crystal, etc.) do not affect the quality of the hologram much. In addition, Fourier–transform–plane holograms are less sensitive to translation alignment errors—higher tolerance for misalignment of removable medium. In addition, since the light intensity is higher with Fourier–transform–plane recording, holographic recording is faster. Most important of all, because of the small transverse size of the signal on the crystal, Fourier–transform–plane recording offers the highest storage capacity under our experiment conditions by allowing more storage locations in a

given volume. However, it is difficult to record multiple Fourier-transform-plane holograms. In the Fourier spectrum, the higher spatial frequencies (where most of the information is) are usually much weaker than the lower frequencies. This makes the spatial profile of the modulation depth very uneven within the crystal and causes distortions due to spatial filtering effect, such as edge enhancement in the stored holograms. To solve this problem, a random-phase plate can be used to diffuse the peak energy in the FT pattern and make the signal intensity well distributed.

As a trade-off between the two, we can also record Fresnel-plane holograms by placing the crystal beyond the focal plane in the signal arm. With Fresnel-plane recording, the signal beam energy is diffused in a larger cross-sectional area, reducing image distortion and improving the M/# [41].

2.5 Memory system design

In all of our experiments described in this chapter, we used mechanical angle scanners to control the angle and position of the reference and signal beam. Demonstration of fast random-access will be discussed in the next chapter. The experimental setup is shown in Figure 2.10. An argon laser with wavelength 488nm is used as the light source. The reference arm contains an XY mechanical scanner (Newport Linear Stage 850B) which moves the focused reference beam horizontally for angle multiplexing, and vertically for fractal and spatial multiplexing. The mirror array is placed at the center plane of the 4F system consists of lenses f_{in} and f_{out}. A large polarizing beamsplitter cube is used 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 facet, yet still avoid crosstalk to other storage locations. 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 90°-geometry storage. Not shown here in the drawing are four cylindrical lenses which magnify the horizontal dimension of the reference beam to make full use of the crystal thickness while keeping the vertical extent small,

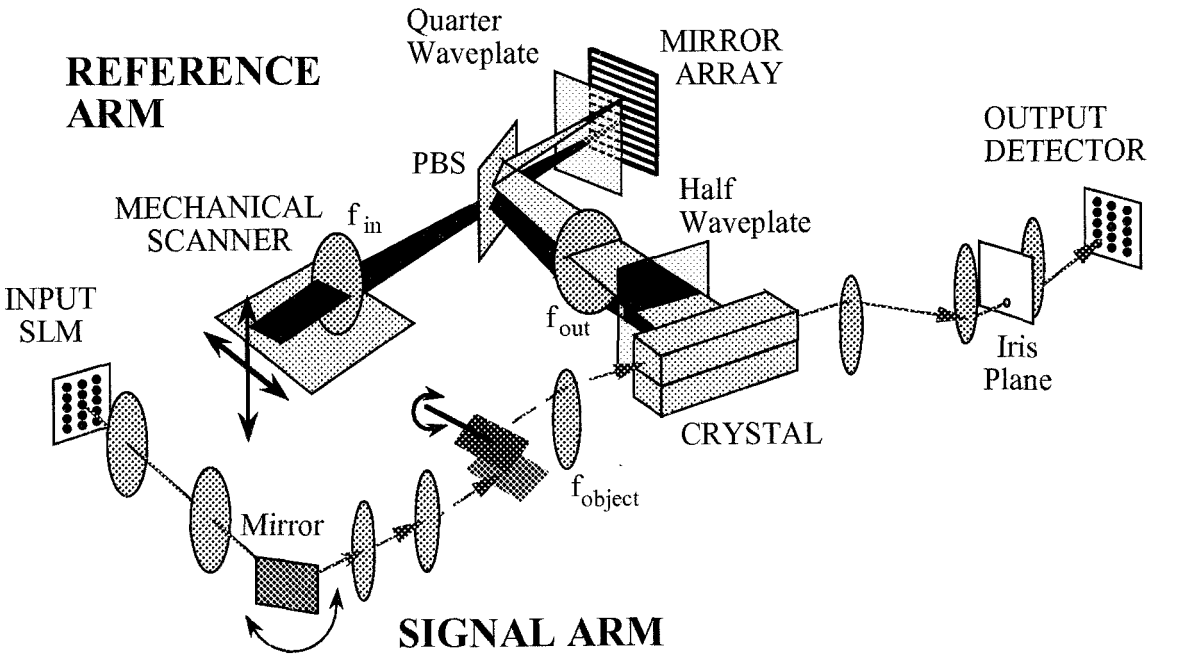

Figure 2.10: Experimental setup of the large–scale memory.

and two mirrors in a periscope arrangement to convey this beam onto the mechanical scanner.

In the signal arm, the information presented on the liquid–crystal Spatial Light Modulator (SLM) is imaged to a mirror mounted on a rotation stage. This horizontal deviation is not required in the theoretical design of our system, but it has an important role in its practical realization. One of the difficulties of large–scale holographic storage is that as the amount of holographic noise and fixed pattern energy (from dust, fixed patterns on the SLM display, or common features among the presented images) 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. This increases the diffraction efficiency of the holograms as well as the effective M/# of the storage system. We use the periscope in the signal arm 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 with horizontal rotation axis, 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 signal arm, and deflection originates from the optical axis of the lens f_{object}. At this point, the object beam has deflected by an arbitrary 2D angle. On the far side of lens f_{object}, this angle determines where the Fourier transform of the displayed information arrives. The distance between the crystal and lens f_{object} can be adjusted to allow either Fourier–plane or Fresnel–plane recording. An identical lens after the crystal completes the 4F system with lens f_{object} to form the output image. As the last stage, another 4F system is used to scale the images onto a CCD detector array and filter out scattered light. A PC computer is used to read in the reconstructed images for further data analysis.

2.6 Storage of 10,000 holograms at a single location

In this section, I demonstrate the storage of 10,000 Fresnel–plane holograms at a single location in the crystal. I present the characterization of the system error performance to gain more insight into the buildup and influence of various noise sources. The procedures and the set of parameters we adopted here became a routine system evaluation process in all of our later experiments. I also propose and demonstrate a "post–recording" profile compensation technique to eliminate most of the deterministic noise in the system.

2.6.1 10,000 Fresnel–plane holograms

Using Fresnel–plane storage, 10,000 holograms were recorded at one storage location of a $LiNbO_3$ crystal. The crystal, a 0.01% Fe–doped 90°–geometry $LiNbO_3$ of dimensions 22mm×22mm ×12mm, was displaced beyond the Fourier–transform plane by

80mm (f_{object} = 300mm). At this point, the DC portion of the expanding image was approximately 4.8mm high × 3.6mm wide. The images were displayed on a 640×480 pixel VGA monitor, and sampled for the 480×440 pixel SLM. Both random bit patterns and a standard chessboard pattern were stored. The CCD detector array was the Photometrics Imagepoint camera with 8-bit dynamic range and 752×480 pixels. The images were over-sampled. Five fractal rows were used for storage with 2,000 holograms stored on each. The vertical spacing between fractal rows was 5mm (2.4°); the horizontal spacing between holograms was 15μm (0.007°, approximately 4 times wider than the Bragg selectivity). The angles listed are the external angles at the crystal face, the distances are movements of the focused reference beam on the mirror array. The vertical angular bandwidth of the images was about 2°, so the fractal spacing completely displaced unwanted reconstructions off the detector array. In the center of the 4F system in the reference arm, the segmented mirror array was replaced by a mirror since spatial multiplexing was not used.

In order to avoid the noise buildup and make full use of horizontal extent of the crystal, the signal beam was displaced sideways during the recording. At the beginning of the recording on each fractal row, the signal was moved towards the near side (with respect to the reference beam) of the crystal by a small amount. During the recording of the first row of 2,000 holograms, the signal was displaced by a smaller amount for every 200 holograms to further reduce the chance of noise buildup due to the long exposures. Because of the bulk absorption, the reference beam is attenuated along its path of propagation through the crystal. Therefore, the erasure time constant was not the same for all the 10,000 holograms because of the displacement of the signal beam on the crystal. Experimentally, we determined the time constant to range from 3,100 seconds to 3,800 seconds depending on the signal beam displacement. The last exposure time was chosen as 0.26 second, the initial exposure lasted 7 seconds, and the total exposure time was 134.6 minutes. The average exposure time per hologram was 0.81 seconds. Several reconstructions are shown in Figure 2.11. These images were taken with camera integration times between 0.5 and 1 second. The average diffraction efficiency was approximately 5×10^{-9}. The average power in

the reconstructions (which were already half dark) was approximately 7 times the background scatter (measured before storage).

2.6.2 Characterization of system error performance

To characterize the noise performance in this angle multiplexed memory, we normalized the reconstruction by an overall illumination profile [72]. 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.

In the noise analysis, 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 an estimate of the Probability Density Function (PDF). We can obtain several parameters from these two PDFs to describe the error performance of the holographic memory:

- the measured Bit-Error Rate (BER) 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

$$SNR = \frac{\mu_1 - \mu_0}{\sqrt{\sigma_1^2 + \sigma_0^2}}, \tag{2.7}$$

  where μ_0 and μ_1 are the means of the OFF and ON pixels, respectively. σ_0 and σ_1 are their variances. The SNR is a useful indicator of the error performance: a large SNR indicates a low BER.

- the estimated Bit-Error Rate (BER). We estimate the shape of the PDF's by assuming that they follow a particular distribution, such as the Gaussian or χ^2

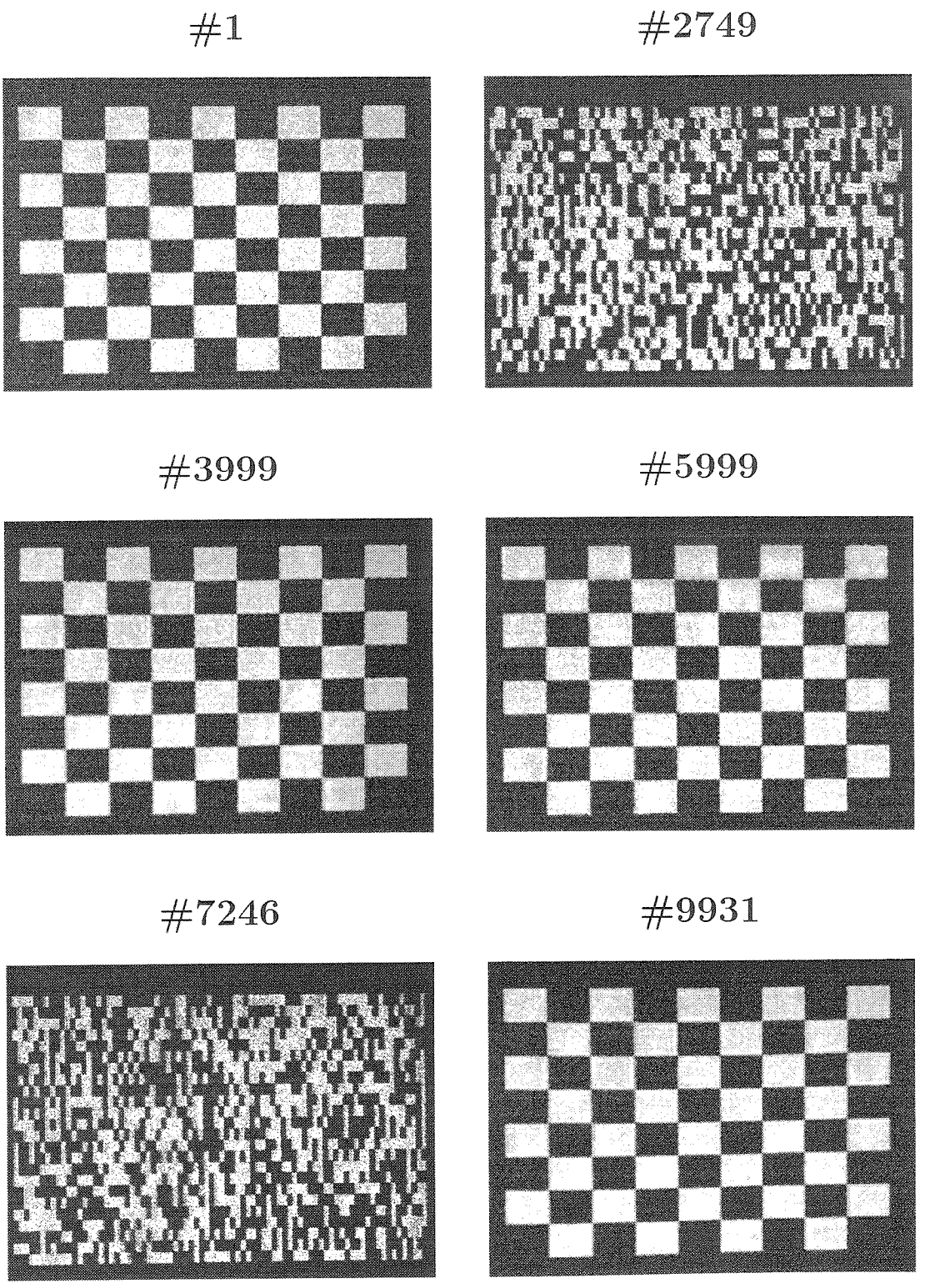

Figure 2.11: Sample reconstructions from the 10,000 Fresnel–plane holograms.

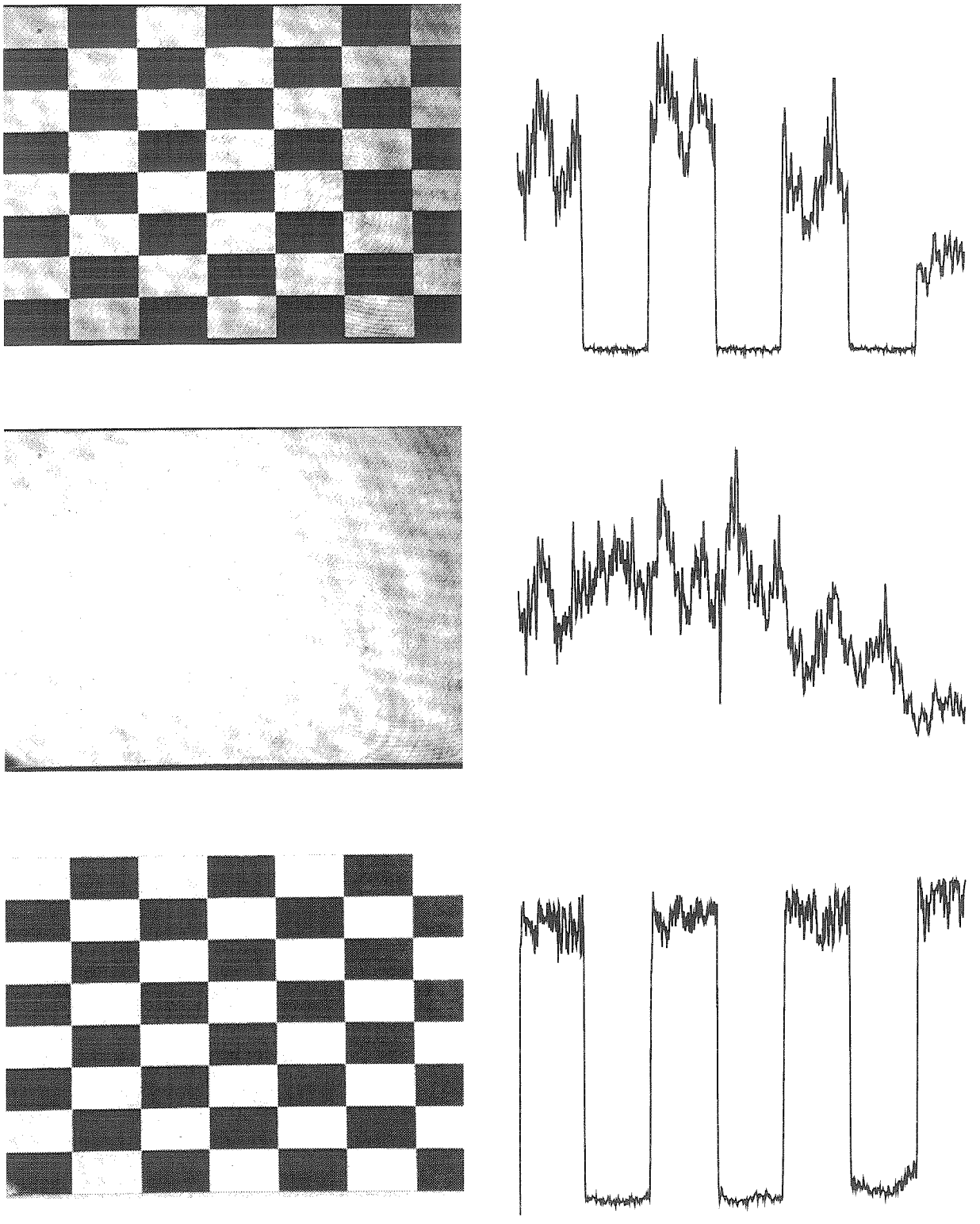

Figure 2.12: Normalization by the background illumination profile.

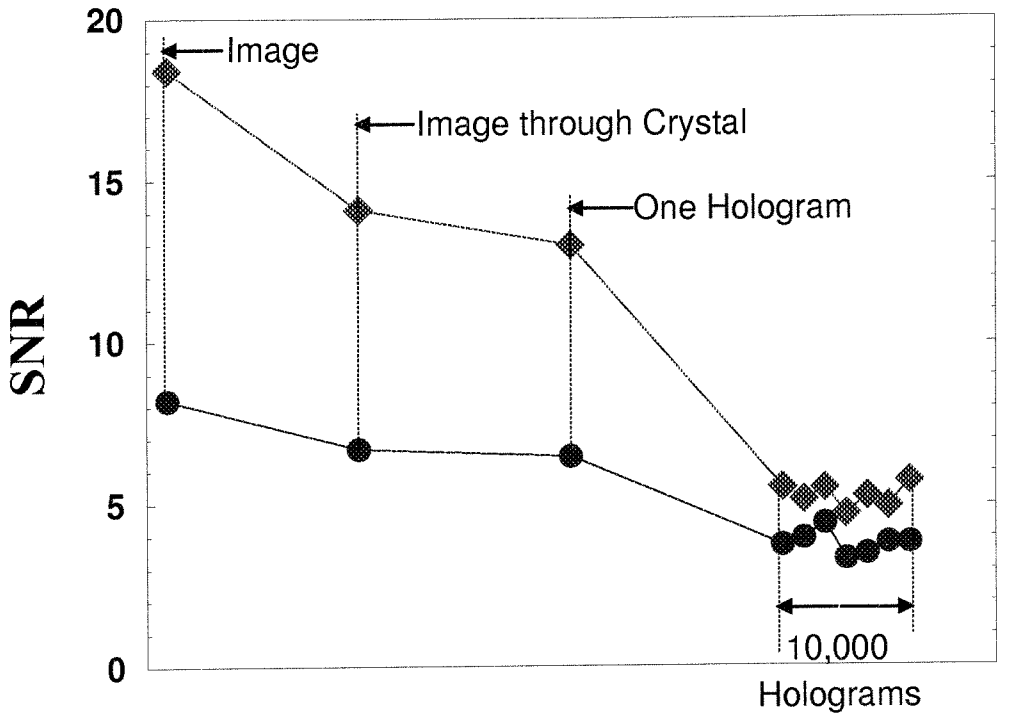

Figure 2.13: Evolution of the system SNR as a function of the number of holograms.

distribution. This is often necessary if the error rate is smaller than the inverse of the number of samples.

The normalization procedure is described in Figure 2.12. 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 2.12.

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 for the storage of 10,000 holograms.

Each chessboard image was also normalized by a blank page image. The SNR for these various conditions is summarized in Figure 2.13. 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 is because the normalization procedure can not effectively correct the errors introduced by the crystal surface imperfections, bulk scattering, and the scattered noise from holographic recording because they are not on an image plane in the signal path.

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 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, the spatially varying modulation depth, and the bulk scattering. In this experiment, 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 5^{th} null of the angular selectivity curve), development of inter–pixel and other noise gratings over the long exposure sequence, and nonuniform erasure of the recorded holograms due to absorption in the crystal and the displacement of the signal beam.

2.7 Demonstration of the 160,000 hologram system

In this section, we demonstrate the full storage capacity of our memory system with the segmented mirror array for angle, fractal, and spatial multiplexing. In all our demonstrations, we used mechanical angle scanners. Fast–random access to the memory using non–mechanical scanners is discussed in the next chapter.

2.7.1 The 160,000 hologram system

Preparation of the reference arm: In the reference arm, the segmented mirror array was used to provide spatial multiplexing at 16 locations on the crystal. Basically, because there are 16 mirror strips in the stack, the mirror array provides 16 fractal rows (vertical angles) for fractal multiplexing at each of the 16 locations. However, the limited numeric aperture of the lens f_{out} affects the number of fractal rows usable for multiplexing, as well as the range of multiplexing angles available on each fractal row, especially for the locations at the two ends. As shown in Figure 2.14, only the reference beams for the locations in the middle can access all the 16 rows with good quality. For another location corresponding to a certain vertical tilt of the mirror facet, not all the 16 vertically tilted reference beams are able to clear the aperture of lens f_{out}. In addition, for those beams which pass through the edge of the lens, distortion occurs as a result of the aberration. In the original system design, a plano–convex lens of 120mm focal length and 100mm diameter ($f/\# = 1.2$) was used as the output lens f_{out}. At the top and bottom locations, there were only 3 fractal rows available. In order to reach our target of 10,000 holograms at each location, 3,333 holograms must be stored on each 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. To solve this problem, we used a Fresnel lens of 70mm focal length and 70mm square aperture. The result is a dramatic increase of both the horizontal angle multiplexing and the number of usable fractal rows for

every location. Even though the reference beam generated by this plastic lens no longer resembled a well–collimated plane wave, no significant difference in the angular selectivity was observed, indicating the result of good auto–correlation carried out in the reconstruction as long as the same reference beam was used.

With the Fresnel lens, 4 fractal rows can be employed to store 2,500 angle multiplexed holograms each, yielding a total 10,000 holograms at one storage location. The vertical translation between fractal rows was 4.8mm. The inter–hologram spacing on the same row was 10μm, which was about 5 times wider than the angle selectivity.

Preparation of the signal arm: The storage medium we used in the experiments was a 90°–geometry $LiNbO_3$:Fe crystal of dimensions 15mm × 20mm × 40mm. It was 0.015% Fe–doped, with absorption coefficient approximately equal to 0.55cm^{-1}. In order to have 16 storage locations in the 4cm–high crystal, we used Fourier–transform–plane recording. The output lens in the signal was an f_{object}=200mm achromat lens of 70mm aperture. The magnification of the imaging optics was 0.67. Therefore, the vertical extent of the signal on the crystal was less than 0.24cm. The total vertical extent of the 16 locations was then 3.9cm.

A random–phase plate was used to spread out the DC energy in the Fourier–transform pattern to avoid distortion and noise caused by over–exposure and erasure at the focused DC spot of the image. The random–phase plate was binary, with phase excursion of 0 and π. The particular phase plate we used in the experiment contained one phase region for every 8×8 SLM pixels. We chose the 8×8 plate simply because it came out with the best quality. It has the smallest transition width between the regions of 0 and π, but there were still dark lines showing up in the detected images Figure 2.15. Therefore, the images to be stored contained 60×50 "superpixels" with 8×8 SLM pixels on each to match the phase plate. SLM pixels corresponding to superpixel boundaries were explicitly turned off in order to hide these dark lines. This is shown in Figure 2.16.

In order to use all of the horizontal extent of the crystal within each vertical location, we deflected the signal beam horizontally to spread the 10,000 holograms

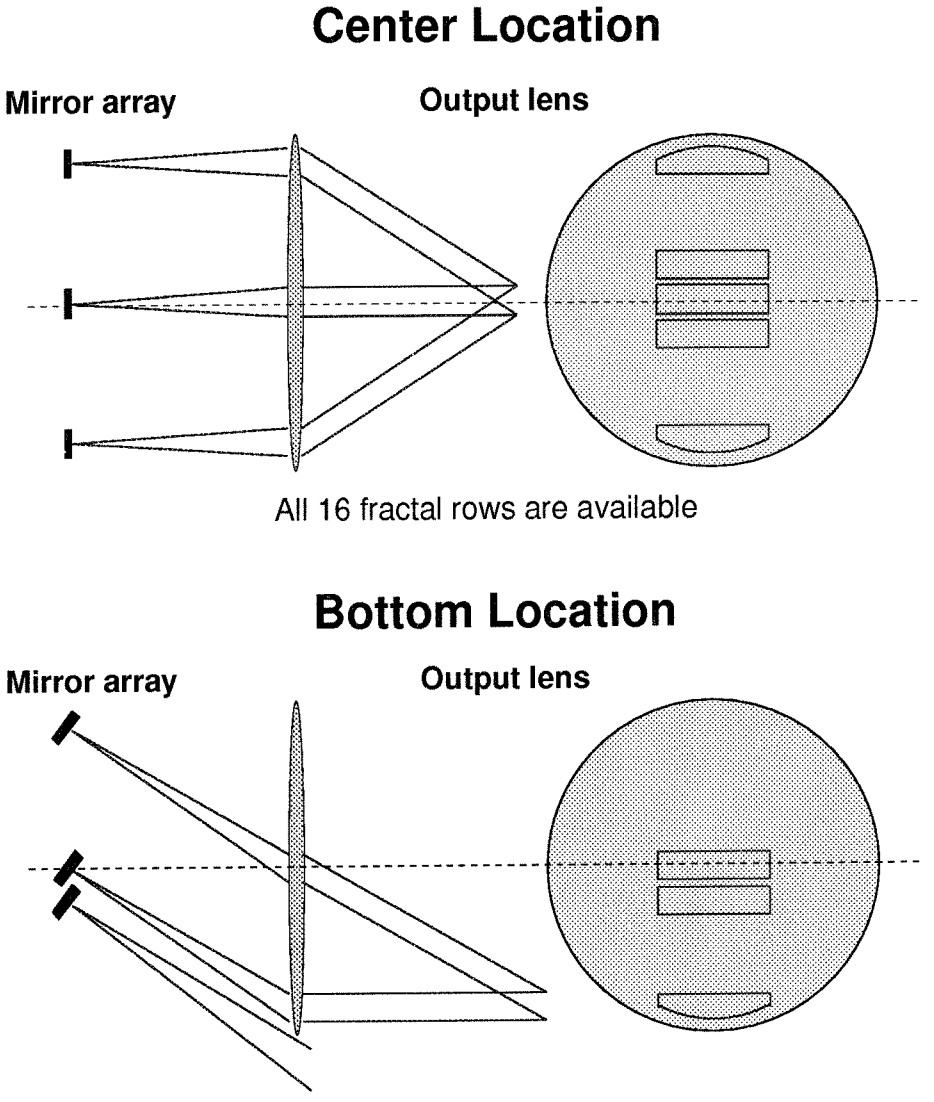

Figure 2.14: Accessibility of the fractal rows by the reference beam.

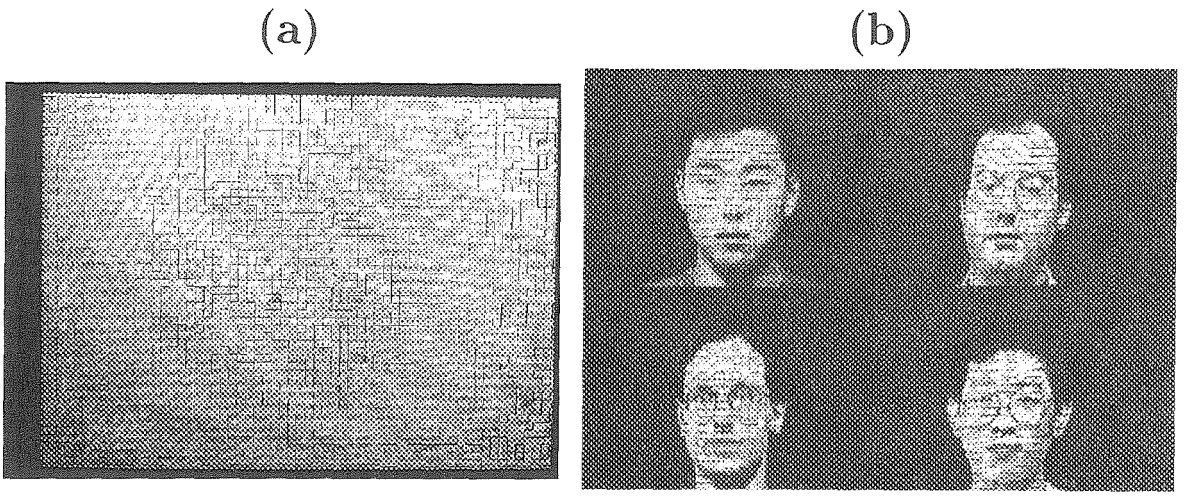

Figure 2.15: Images showing lines from random phase plate: (a) image through crystal; (b) one of 10,000 holograms at top location.

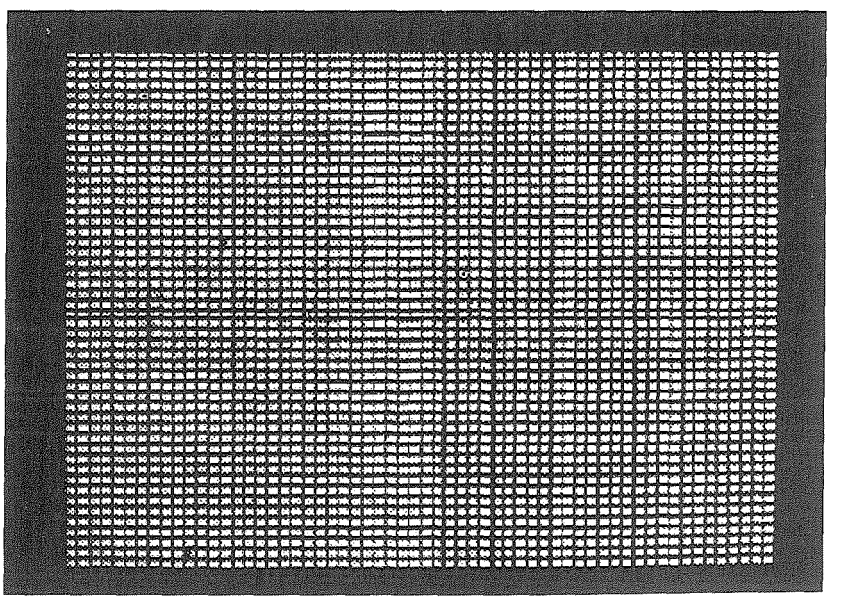

Figure 2.16: Set of 60 $\times$ 50 pixel regions used to avoid random phase plate lines.

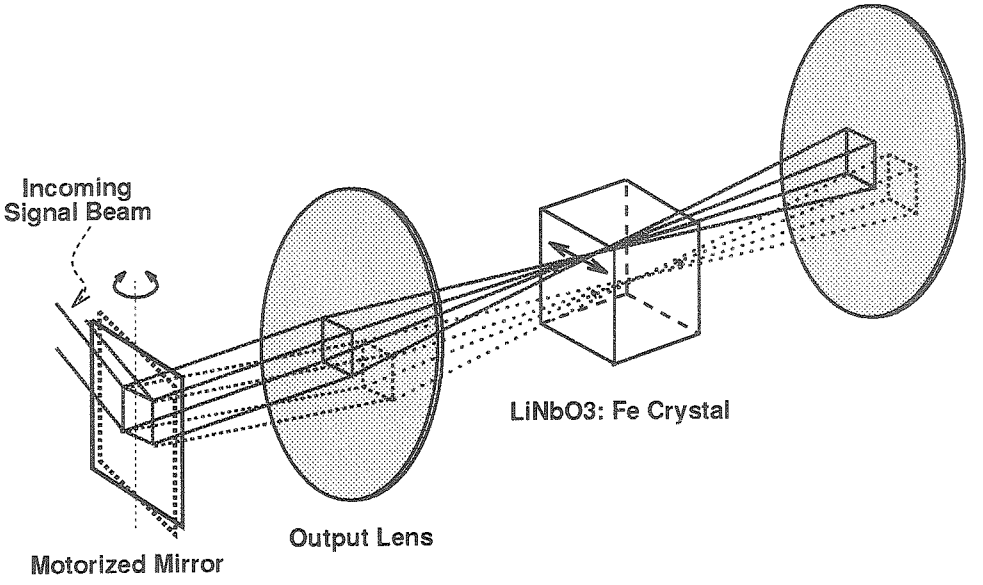

Figure 2.17: Horizontal scanning of the signal beam.

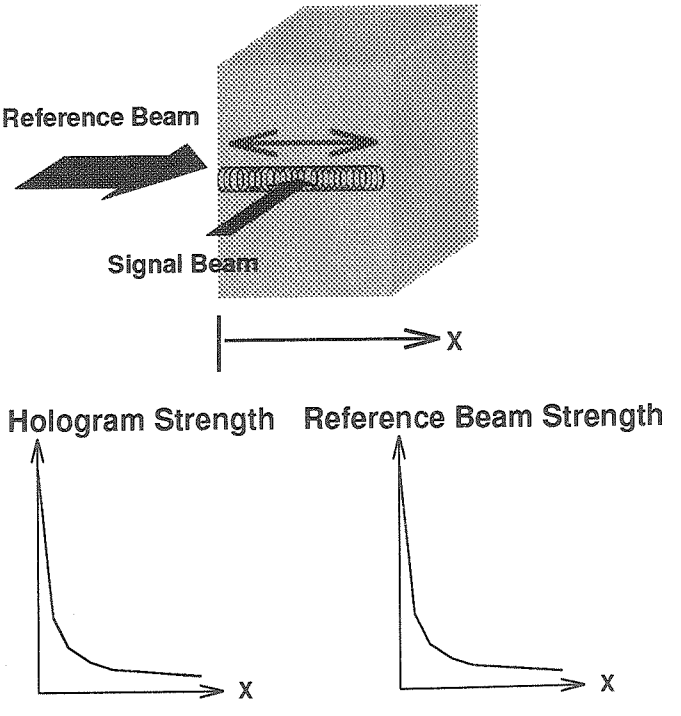

Figure 2.18: Effects of the horizontal scanning.

throughout the crystal. This is shown in Figure 2.17. However, unlike the previous case in which we deflected the signal beam once every fractal row, here we displaced it for every hologram. For each location, we deflected the signal horizontally across the width of the crystal, completing 50 complete cycles during the storage of 10,000 holograms. In addition, since the reference beam was slightly larger than the size of the focused signal, we deflected the signal beam vertically within the storage location by a small random amount along with each horizontal movement. This vertical deflection of the signal beam was constrained to be within reasonably tight limits, so that the signal and reference achieved good overlap for all holograms. We observed no strong trend in diffraction efficiency between holograms written with signal 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 (Figure 2.18). 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 could be compensated by minor adjustment of the recording times for these last few holograms.

2.7.2 30,000 Fourier–transform–plane holograms at 3 locations

30,000 holograms were recorded at locations #1, #9, and #16, corresponding to the bottom, center, and top locations of our 160,000 hologram system, with 10,000 holograms at each location. Each set of 10,000 holograms at a location was stored with 4 fractal rows of 2,500 holograms each. The vertical angle difference between fractal rows is 4.8mm (3.9°). Within each fractal row, all holograms were spaced by 10μm (0.008°). For each set of 10,000 holograms, the last exposure time was 0.52 second and erasure time constant was 6,000 seconds. The recording and erasure were slower than those in the recording of 10,000 Fresnel–plane holograms because the

mirror array caused a lot of loss in the light intensity. The first exposure time was 3.89 seconds, with a total recording time as 201 minutes, and the average exposure time per hologram 1.2 seconds. The average diffraction efficiency was approximately 4×10^{-9}.

We captured reconstructions from the center, top, and bottom locations, using a 1 second exposure with the Photometric 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. Some of the sample reconstructions are shown in Figure 2.19, 2.20, 2.21.

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 superpixel region to create sample points for analysis. This gives us, on average, 1,500 ON pixel regions and 1,500 OFF regions. The effects of SLM nonuniformity across a superpixel were averaged out in this process.

All the reconstructions were retrieved without any measured errors. The characteristics of these reconstructions are plotted in Figure 2.22, 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. Note that one of the reconstructions in Figure 2.22 (marked with an arrow) appears to be much weaker than the others. This particular reconstruction was captured with a 0.2 second exposure. Note that the SNR and probability of error are unaffected, though, 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

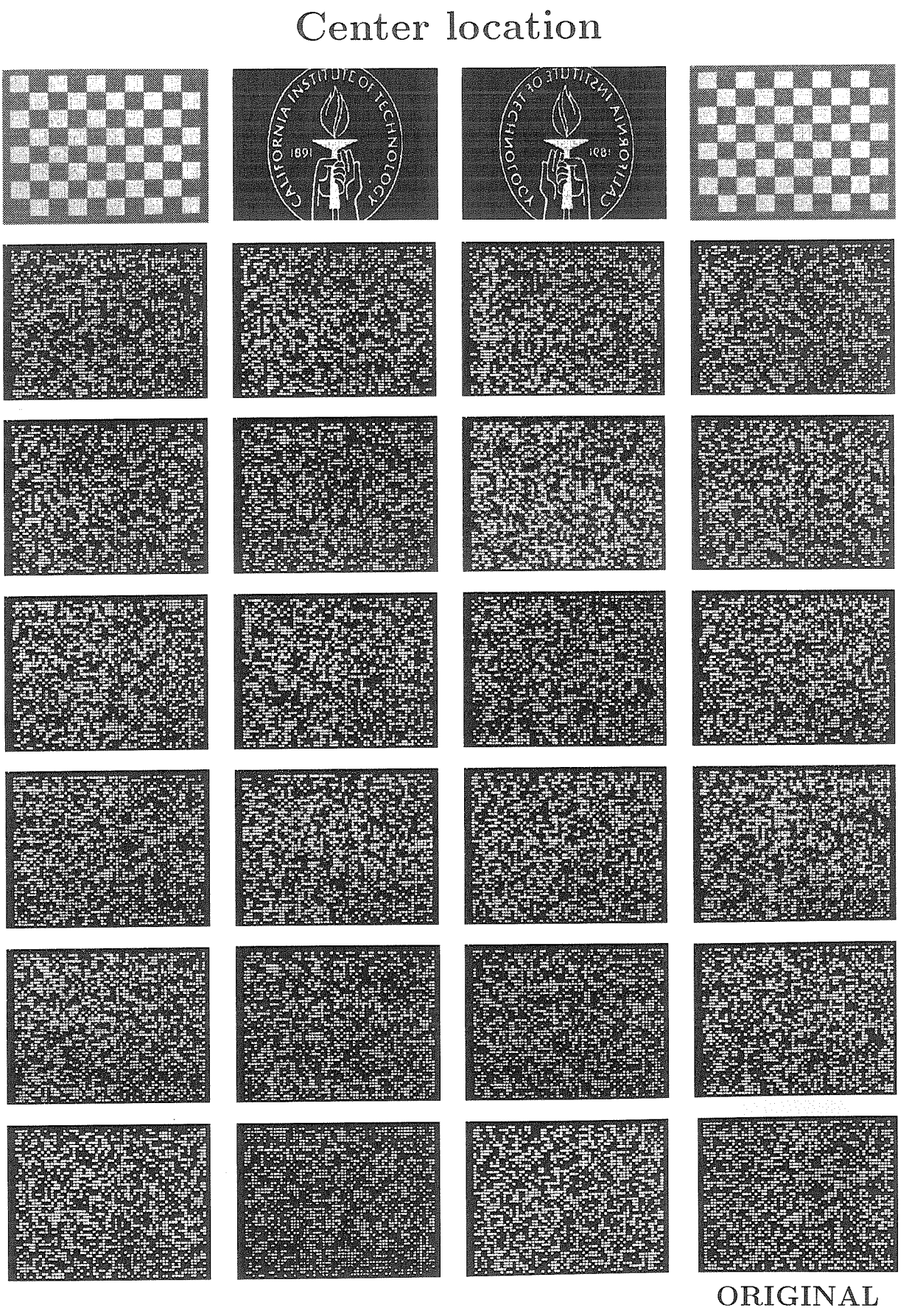

Figure 2.19: Sample reconstructions from 10,000 Fourier–transform–plane holograms at the center location (#9)

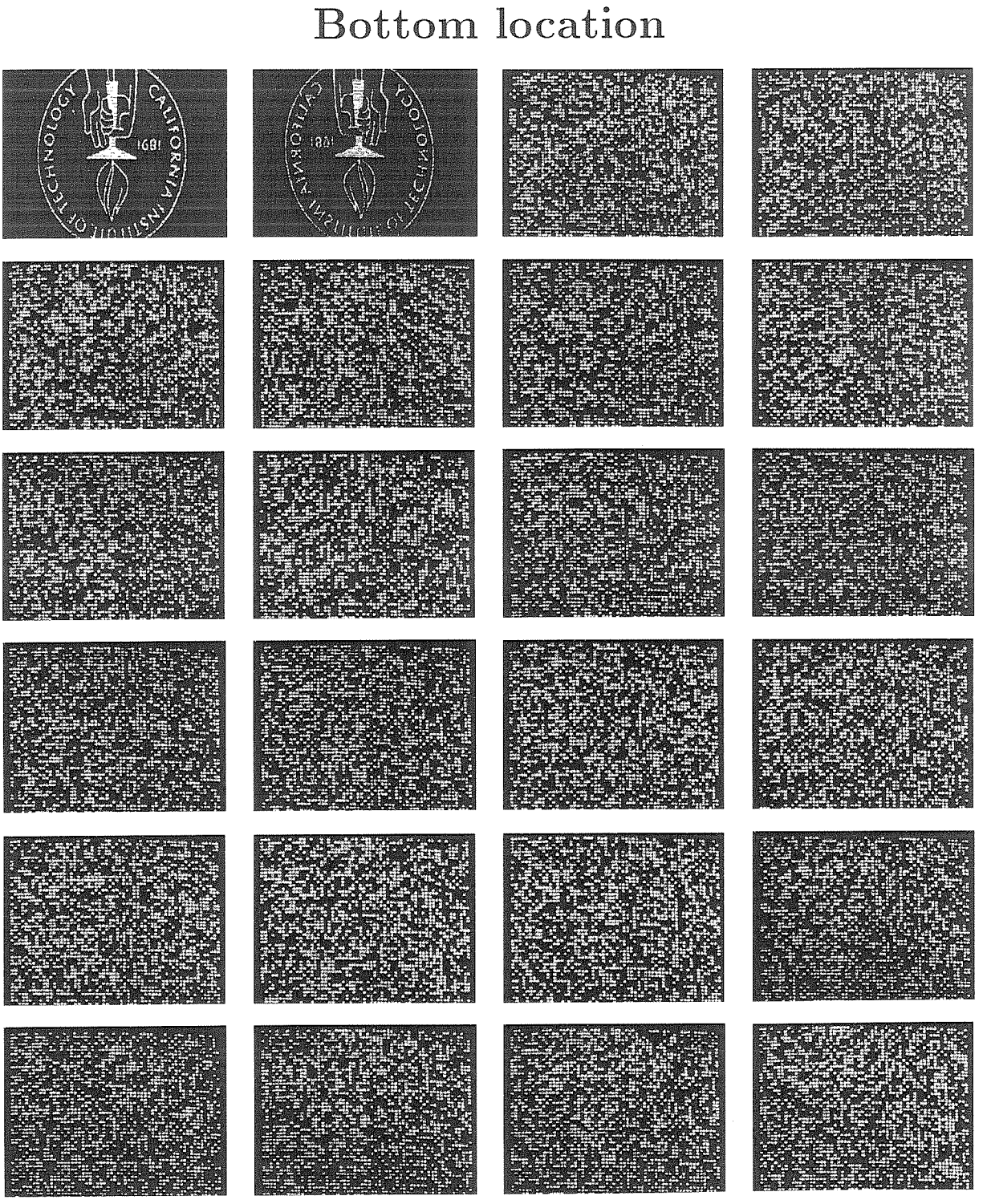

Figure 2.20: Sample reconstructions from 10,000 Fourier–transform–plane holograms at the bottom location (#1)

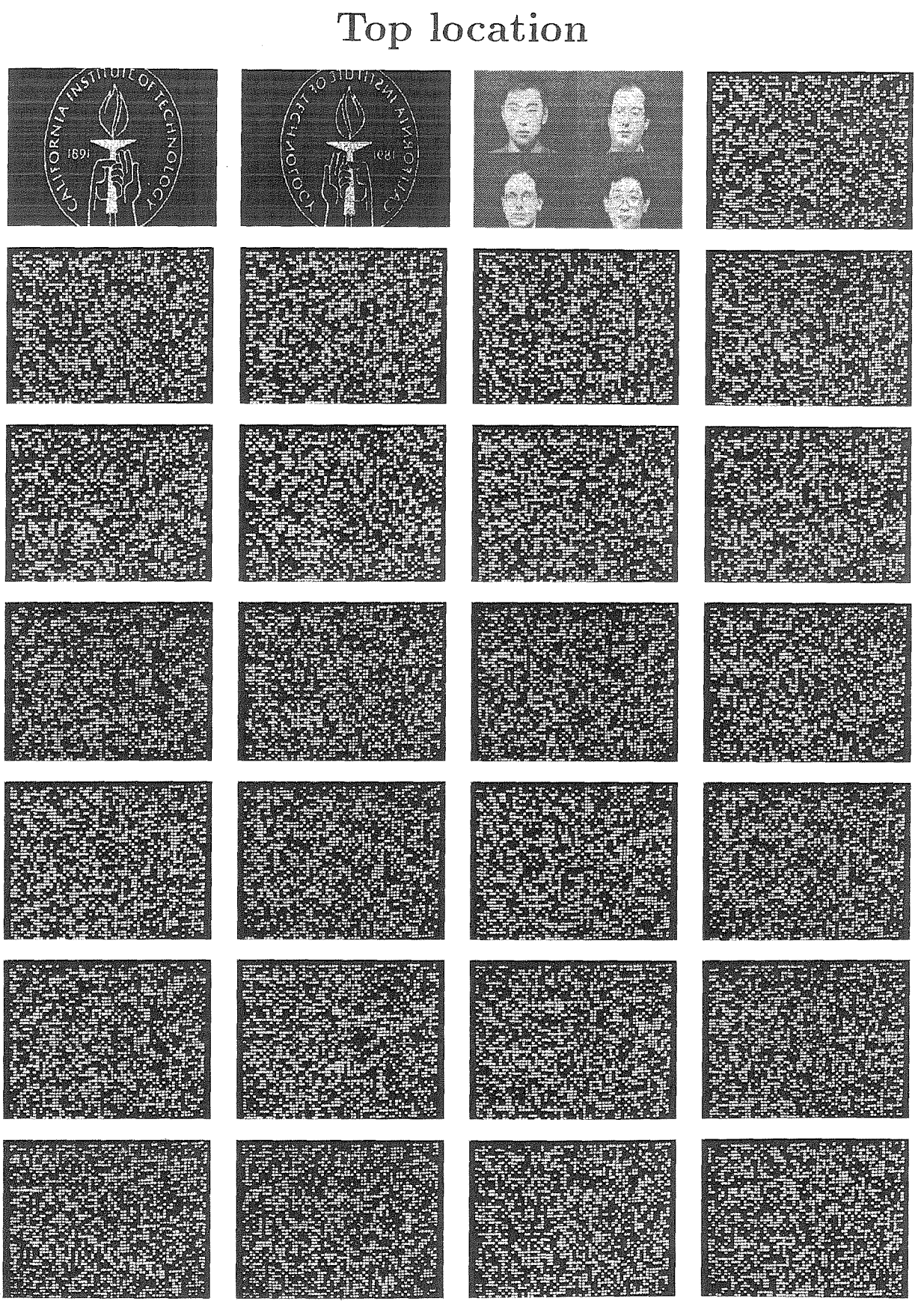

Figure 2.21: Sample reconstructions from 10,000 Fourier–transform–plane holograms at the top location (#16).

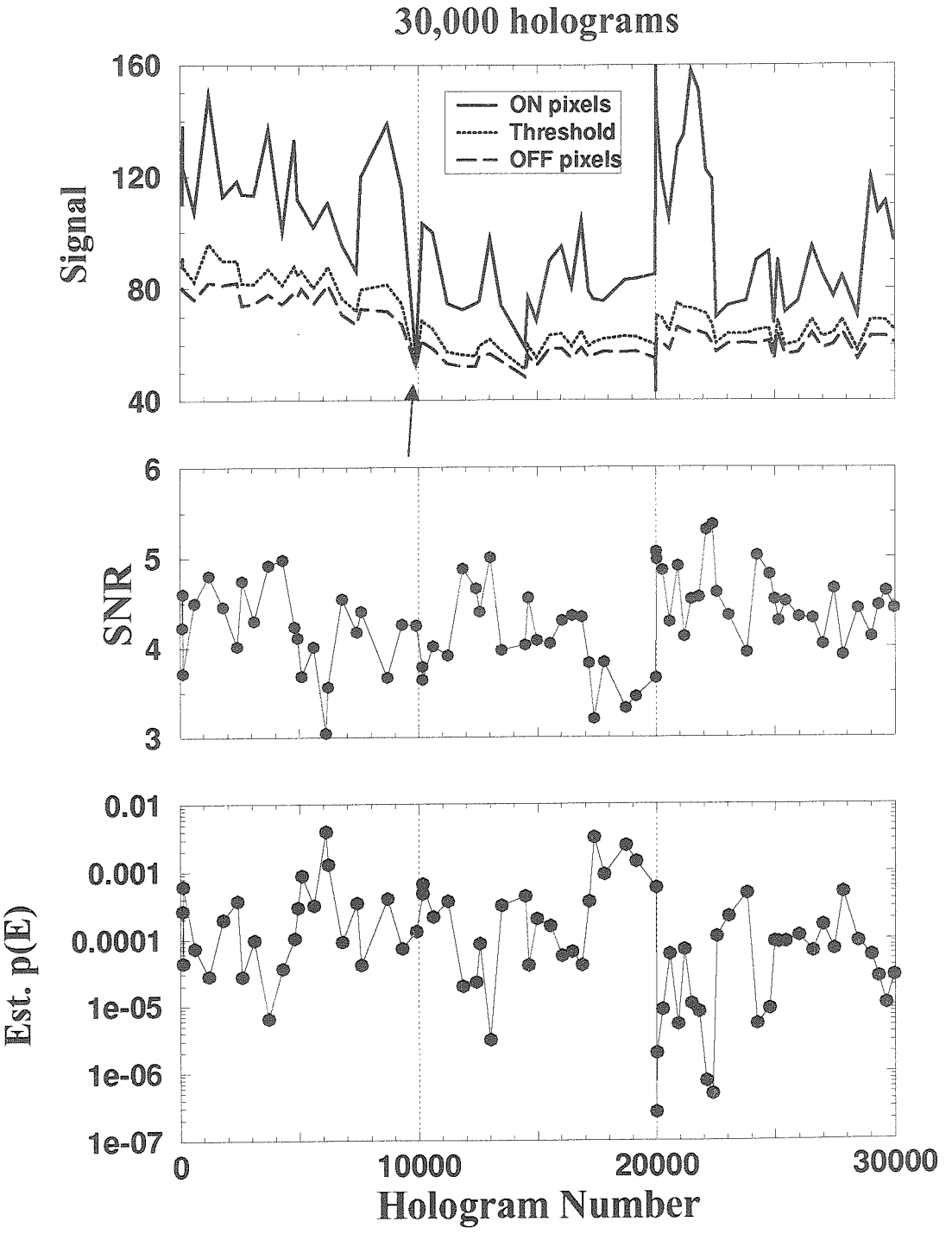

Figure 2.22: Analysis of the sample reconstructions from Figure 2.19—2.21. Plotted as a function of hologram number are the means of the ON and OFF pixel regions, value of the optimal threshold, SNR, and both measured and estimated probability of error.

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.

2.7.3 160,000 Fourier–transform–plane holograms at each of the 16 locations

In this experiment, 16,000 Fourier–transform–plane holograms were recorded at the 16 storage locations, with 1,000 hologram at each location. Some sample reconstructions are shown in Figure 2.23. There were no measured errors in all the sample reconstructions indicating that there was no crosstalk from other locations. In Figure 2.24, we show SNR and estimated Bit–Error Rate (Pe) as a function of spatial location for a single hologram and 1,000 holograms in each location.

2.8 Appendix: Complications in the M/#

In this section, we derive the expressions for the zero and first–order spatial harmonics of the space–charge field during its initial evolutions [1]. Many theoretical models for the behavior of the photorefractive effect were introduced in the 1970's [20–34], but the Kukhtarev equations have been universally accepted as the general model for the operation of the photorefractive effect [27, 34]. Many papers have been published which expand on the basic set of equations. In particular, we refer to three almost simultaneously published papers [73–75] in which the photovoltaic current is proportional to the local absorber concentration $(N_D - N_D^+)$, instead of to the bulk absorption coefficient α [76, 77].

The reason we repeat the derivations here is to show the complications in the M/# of a large–scale holographic storage system. We start with the one–dimensional

[1] In other words, the regime where the space–charge field is too small to affect holographic recording.

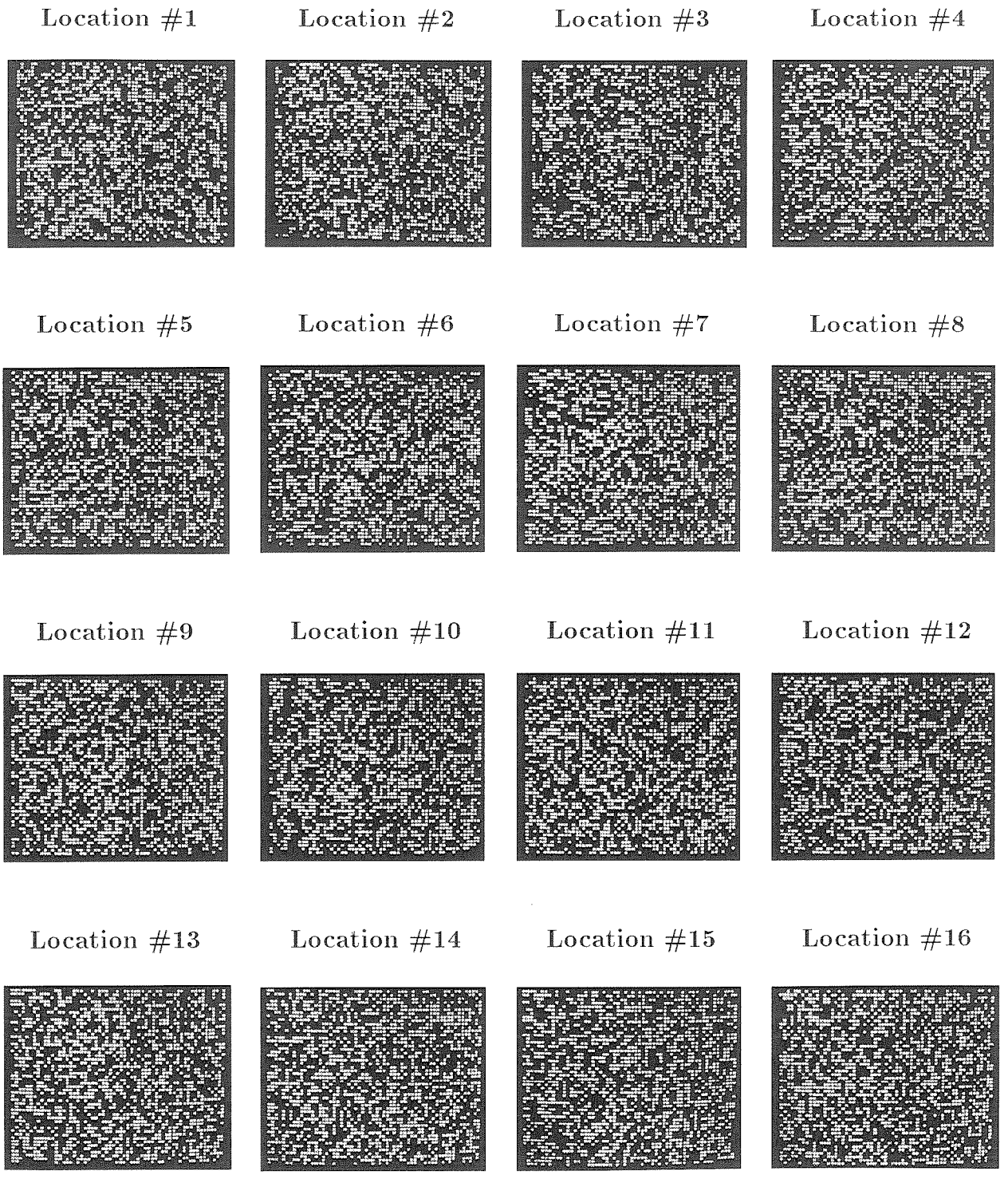

Figure 2.23: Sample reconstructions from 1,000 holograms at each of the 16 storage locations.

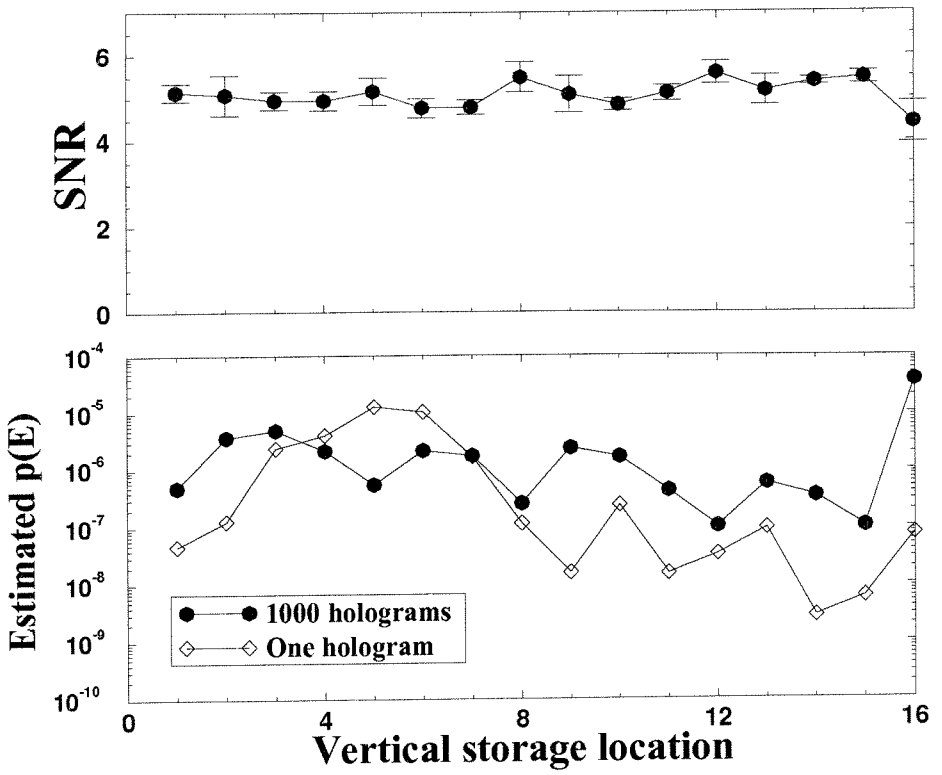

Figure 2.24: SNR as a function of hologram location.

Kukhtarev equations listed below,

$$\frac{\partial N_D^+}{\partial t} = (SI+\beta)(N_D - N_D^+) - \gamma_R N_D^+ n_e, \tag{2.8}$$

$$\frac{\partial n_e}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{q}\frac{\partial J_e}{\partial x}, \tag{2.9}$$

$$J_e = q\mu_e n_e E + K_B T \mu_e \frac{\partial n_e}{\partial x} + P(N_D - N_D^+)I, \tag{2.10}$$

$$\epsilon\frac{\partial E}{\partial x} = q(N_D^+ - N_A - n_e), \tag{2.11}$$

where N_D, N_D^+, N_A, n_e are the density of the dopants, ionized donors, acceptors, and free electrons, respectively. I is the light intensity pattern. S, β, γ_R, and P are all material parameters—the photorefractive cross-section, thermal generation rate, recombination rate, and photovoltaic constant, respectively. K_B is Boltzmann's constant. T is the absolute temperature. J_e is the electron current density. μ_e is the mobility of the electrons.

In order to reach closed-form solutions to these equations, three assumptions have to be made: First, $m = \frac{I_1}{I_0} \ll 1$ [2]. This allows us to linearize the equations, by writing the variables in terms of a DC component and a sinusoidal component (spatial perturbation) of the same spatial frequency, K, as that of the intensity interference patterns. As a result, when the intensity pattern illuminating the crystal is

$$I = I_0 + I_1 e^{jKx}, \tag{2.12}$$

then the following variables are expected to be

$$N_D^+ = N_{D0}^+ + N_{D1}^+ e^{jKx}, \tag{2.13}$$

$$n_e = n_{e0} + n_{e1} e^{jKx}, \tag{2.14}$$

$$J_e = J_{e0} + J_{e1} e^{jKx}, \tag{2.15}$$

$$E = E_0 + E_1 e^{jKx}. \tag{2.16}$$

Note that E_0 can either be an externally applied electric field, or an internal field which builds up over the illuminated region during exposure. When we substitute these variables into Equation 2.8—2.11, they can be separated into two set of equations: DC terms and first-order terms.

We first notice that the change of the DC ionized donor (ΔN_{D0}^+) is identical to that of the DC density of conduction (free) electrons (n_{e0}) [3]. This is because the DC current density of the electrons (J_0) has no spatial dependence. The change of n_{e0} can be derived as

$$n_{e0}(t) = \frac{SI_0(N_D - N_{D0}^+)}{\gamma_R N_{D0}^+} \left(1 - e^{-t(\gamma_R N_{D0}^+)}\right), \tag{2.17}$$

where the steady-state value n_{e0} is

$$n_{e0} = \frac{SI_0(N_D - N_{D0}^+)}{\gamma_R N_{D0}^+}, \tag{2.18}$$

[2] where I_0, I_1 are the DC and first-order spatial perturbation in the exposing light pattern.

[3] We assume $\beta = 0$. There are no free electrons by thermal excitation.

which is much less than the density of the dopants and the acceptors. This allows us to make the second assumption: $N_D, N_A \gg n_{e0}$, which much simplifies the following theoretical treatment. As a result, the DC ionized donor N_{D0}^{+} can be derived from the DC Poisson's relationship.

$$N_{D0}^{+} = N_A. \tag{2.19}$$

So the rate of change in $n_{e0}(t)$ is determined by $\gamma_R N_A$. This term is much larger than the excitation rate $SI_0(N_D - N_A)$. Therefore, n_{e0} reaches a very small value relative to N_{D0}^{+} much faster than the photorefractive response time. This is the reason that we can make yet another assumption: $\partial n_{e0}/\partial t = 0$, to further simplify the solution to the Kukhtarev equations.

The buildup of the DC electric field E_0 is similar to charging up a capacitor by a virtual field—the photovoltaic field. It can be described as

$$J_{e0} = q\mu_e n_{e0} E_0 + P(N_D - N_A) I_0, \tag{2.20}$$

$$\frac{\partial \rho}{\partial t} = -J_{e0}, \tag{2.21}$$

$$E_0 = \frac{\rho}{\epsilon}, \tag{2.22}$$

where ρ is the surface charge density at the boundary of the illuminated region. The solution to this set of equations is

$$E_0 = -E_{0ph}(1 - e^{-t/\tau_{di}}), \tag{2.23}$$

where E_{0ph} is the characteristic photovoltaic field,

$$E_{0ph} = \frac{P\gamma_R N_A}{q\mu_e S}, \tag{2.24}$$

τ_{di} is the dielectric (relaxation) time of the electrons,

$$\tau_{di} = \frac{\epsilon}{q\mu_e n_{e0}}. \tag{2.25}$$

This shows that, during holographic recording, the DC space-charge field increases to "screen" the photovoltaic field.

The first-order Kukhtarev equations are

$$\frac{\partial N_{D1}^{+}}{\partial t} = SI_1(N_D - N_A) - SI_0 N_{D1}^{+} - \gamma_R N_{D1}^{+} n_{e0} - \gamma_R N_A n_{e1}, \tag{2.26}$$

$$\frac{\partial n_{e1}}{\partial t} = \frac{\partial N_{D1}^{+}}{\partial t} + j\frac{K}{q} J_{e1}, \tag{2.27}$$

$$J_{e1} = q\mu_e(n_{e0}E_1 + n_{e1}E_0) + jK_BTK\mu_e n_{e1} + P[(N_D - N_A)I_1 - N_{D1}^{+}I_0], \tag{2.28}$$

$$E_1 = -j\frac{q}{K\epsilon}(N_{D1}^{+} - n_{e1}). \tag{2.29}$$

Under the quasi-steady state assumption ($\partial n_e/\partial t = 0$), the above equations become a set of first-order differential equations. The development of the space-charge field is then found to be

$$\tau\frac{\partial E_1}{\partial t} = -E_1 + E_{1\infty}, \tag{2.30}$$

where the characteristic time constant τ and the saturation value $E_{1\infty}$ are

$$\tau = \tau_{di}\frac{E_q}{E_{ue}}\frac{E_0 + j(E_{ue} + E_d)}{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + j(E_q + E_d)}, \tag{2.31}$$

$$E_{1\infty} = -jmE_q\frac{(E_0 + E_{0ph}) + jE_d}{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + j(E_q + E_d)}. \tag{2.32}$$

E_q, E_{ue}, E_d are the space-charge limited field, drift and diffusion field, respectively.

$$E_q = \frac{qN_A(N_D - N_A)}{K\epsilon N_D}, \tag{2.33}$$

$$E_{ue} = \frac{\gamma_R N_A}{K\mu_e}, \tag{2.34}$$

$$E_d = \frac{K_BTK}{q}. \tag{2.35}$$

Since the DC field E_0 changes with time during exposure, the evolution of E_1 is not exactly exponential, and there is no closed-form solution to Equation 2.30.

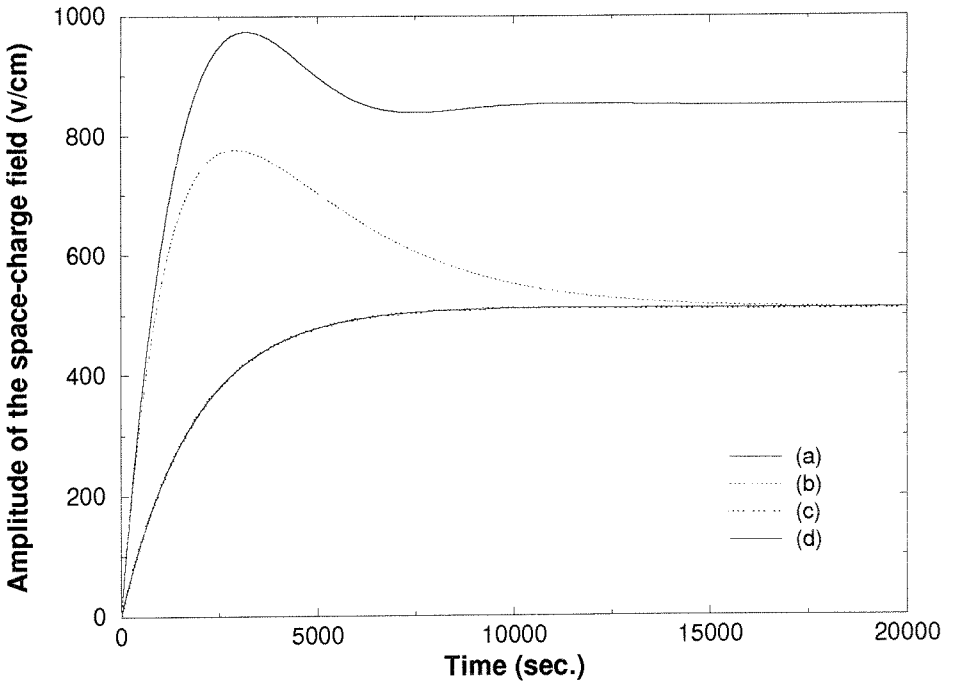

Figure 2.25: Time evolution of the space–charge field under different conditions.

We numerically evaluated the development of E_1 under different conditions using the same method found in Reference [73]:

- Short–circuit: $E_0 = 0$ all the time.
- Open–circuit: E_0 builds up following Equation 2.23.
- Open–circuit: after prolonged exposure (or pre–exposure) so that $E_0 = -E_{0ph}$.
- With no photovoltaic effect: $E_{0ph} = 0$.

The corresponding time evolutions of E_1 during recording with the same modulation depth m are plotted in Figure 2.25. Curve (a) is for the short–circuit condition. Curve (b) is for the open–circuit which is the most common condition in 90°–geometry recording. Curve (c) is obtained after very long exposure so that the DC field fully "screens" the photovoltaic field. Because of the screening, it is similar to Curve (d) which corresponds to the case without the photovoltaic effect at all. The writing rate in Curve (a) is much higher than that in Curve (c). In practice, the writing slope of Curve (a) represents the recording rate at the beginning of holographic recording

with a "fresh" crystal. As the crystal is further exposed for multiple hologram storage, the recording rate becomes smaller and smaller, approaching that of Curve (c). As a result, the M/# of the memory system decreases. The M/# that is predicted in Reference [40, 41] under short–circuit condition, or experimentally determined from single–hologram recording and erasure behavior [4], is only the best–scenario value.

[4]The exposure is usually much less than what it takes for the DC field to fully build up.

Chapter 3 Fast random–access holographic memory

Contents

3.1 Readout using AOD **51**

3.1.1 Operation of an AOD 51

3.1.2 Angular selectivity using AOD 53

3.1.3 Storage capacity and data transfer rate with AOD 55

3.2 Recording using AOD and EOM **57**

3.2.1 System design 57

3.2.2 Optimizing EOM 60

3.2.3 Storage of 1,000 holograms 63

3.2.4 Large–scale memory using AOD and EOM 65

3.3 Single–AOD setup **67**

3.4 Optical Phase–Lock Loop (OPLL) **69**

3.4.1 Optical phase stabilization using AOD and EOM 71

3.4.2 Phase stabilization by OPLL 72

3.4.3 Coherent addition and erasure of holograms 74

3.4.4 Improved OPLL design for holographic memory 75

One of the advantages that distinguishes volume holographic memory from other storage technologies is the parallel readout. When accessing memory content of a holographic storage system, a page of information, instead of a bit of data such as in a CD–ROM, is retrieved at a time. Therefore, the data transfer rate of a holographic memory is the product of the accessing (scanning) speed of the addressing device and the number of pixels (bits) on each hologram. As a result, volume holographic storage

has the potential to provide higher data transfer rate. In the previous demonstrations of a large–scale holographic random–access memory, we used mechanical scanners as the addressing device. They are slow, and prone to errors in re–locating the stored holograms due to their low accuracy and the backlash problem. Non–mechanical angle scanners, on the other hand, provide much faster scanning rate, accurate addressing without backlash or instabilities due to moving parts. For example, an Acousto–Optic Deflector (AOD) is capable of scanning and changing the angle deflection of the output beam in less than 100μsec. Given the number of pixels on the SLM as 1,000×1,000, the data transfer rate can then be as high as 10Gbits/sec. Volume holographic storage using AOD's has been demonstrated [72, 78, 79]. Recording and retrieval of a temporal bit sequence using the AOD in an optical buffer memory has also been demonstrated [80].

In this chapter, we discuss a holographic storage system using AOD as the addressing device. We address the issues related to memory–content addressing by an AOD in Section 3.1. In Section 3.2, we show the design and experiment result of a memory setup using an AOD for angle multiplexing and an Electro–Optic Modulator (EOM) to compensate for the Doppler–shift introduced by the AOD. In Section 3.3, we demonstrate an alternative configuration using a single AOD for both recording and readout. In the remainder of this chapter, we present an optical–phase stabilization system using an AOD and EOM for active compensation of the phase drifts in the optical paths.

3.1 Readout using AOD

3.1.1 Operation of an AOD

Acousto–optic devices utilize the interaction between light and sound to control the frequency, intensity, and direction of the optical beam. This type of control enables a large number of applications involving the transmission, display, and processing of information. In our experiments, we used a special type—the Acousto–Optic Deflec-

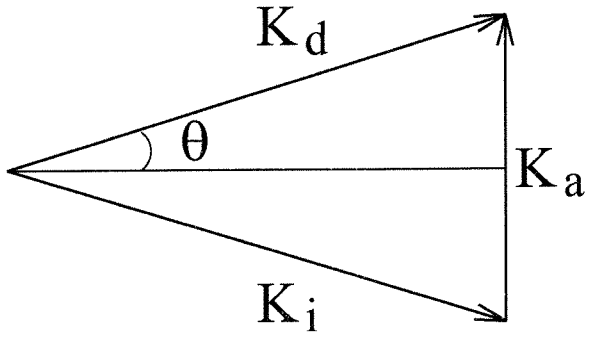

Figure 3.1: K–space description of the interaction between light and sound.

tors (AOD) that induce Bragg diffraction of light by acoustic wave propagating inside the device.

Bragg diffraction of light by sound can be described as two sets of conservation conditions [81]. The first one is the conservation of momentum which can be written as

$$\vec{K}_d = \vec{K}_i + \vec{K}_a, \tag{3.1}$$

where $\vec{K}_i, \vec{K}_d$ and $\vec{K}_a$ are the wave vectors of the incident, diffracted, and the sound (acoustic wave), respectively. This condition is equivalent to the Bragg (phase–matching) condition, as shown in Figure 3.1.

The second is the conservation of energy

$$\omega_d = \omega_i \pm \omega_a, \tag{3.2}$$

where ω_i, ω_d, and ω_a are the angular frequencies of the incident, diffracted, and the sound, respectively. This relationship indicates that the diffracted beam is shifted in frequency by an amount equal to the sound (acoustic) frequency. This is called Doppler shift.

Given that $\omega_d, \omega_i \gg \omega_a$, we have

$$|K_d| \approx |K_i|.$$

An expression can then be found relating the angle deflection in the diffracted beam, $\Delta\theta$ (inside the material) to the change of the acoustic frequency Δf_a. It is

$$\Delta\theta = \frac{\lambda}{n\nu_s}\Delta f_a, \tag{3.3}$$

where λ is the vacuum wavelength of the incident beam. n is the index of refraction of the material. ν_a is the acoustic velocity inside the material. This is the basic relationship that governs the operation of an AOD as an angle deflector and scanner.

3.1.2 Angular selectivity using AOD

As shown in the previous chapter, 90°–geometry storage is the optimal choice for angle multiplexing in a volume holographic memory. In 90° geometry, the angle selectivity is determined by the aperture of the reference beam on the crystal—the interaction length L. For the best angle selectivity in the 90° geometry (both signal and reference beam enters the crystal orthogonally), the first–null of the angle selectivity curve is at

$$\delta\theta_x = \lambda/L, \tag{3.4}$$

where $\delta\theta_x$ is the angle change at the crystal.

Normally, the reference beam is imaged onto the crystal by a 4F system shown in Figure 3.2. To generate the first–null angle deflection $\delta\theta_x$ at the crystal, a change in the deflection angle of the reference beam $\delta\theta_a$ at the AOD should be

$$\delta\theta_a = \delta\theta_x \frac{f_1}{f_2}, \tag{3.5}$$

where f_1 and f_2 are the focal lengths of the two lenses in the 4F imaging system. Since $L = Af_2/f_1$, with A as the aperture of the reference beam on the AOD, Equation 3.4

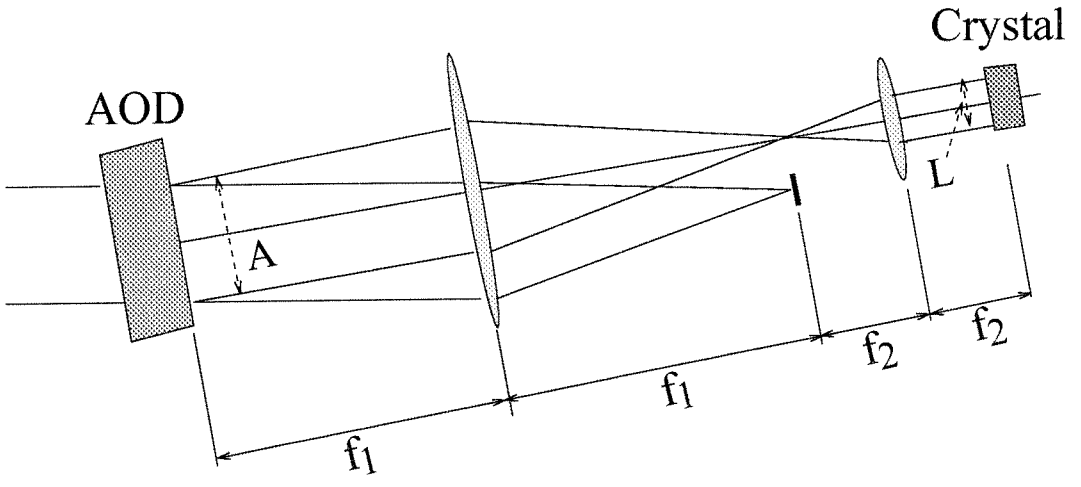

Figure 3.2: 4F system imaging the reference beam onto the crystal.

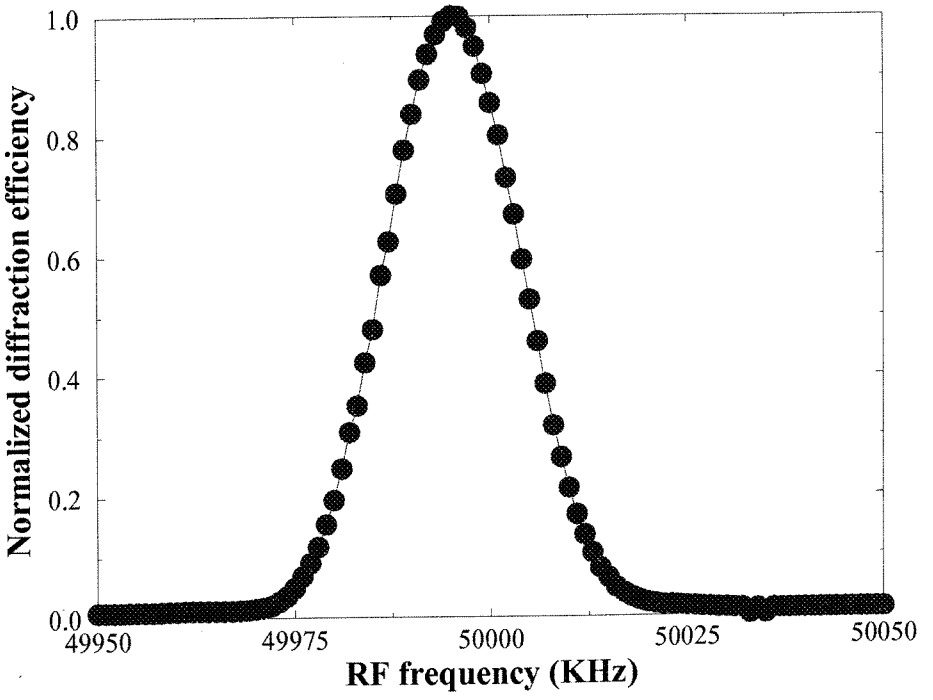

Figure 3.3: Angle selectivity curve of a 90°–geometry hologram readout by AOD.

and 3.5 yield

$$\delta\theta_a = \frac{\lambda}{A}, \tag{3.6}$$

which indicates that the Bragg selectivity of the storage system is equivalent to the diffraction limit of the AOD. Figure 3.3 presents an example where the selectivity curve of a recorded hologram is obtained by scanning the AOD in the frequency range from 40MHz to 60MHz. The half-width of this angle selectivity curve is approximately 33KHz. Given the magnifying ratio of the 4F system (f_2/f_1) used in this measurement as 0.2, this frequency can be converted into a change of the deflection angle of the reference beam by Equation 3.3. The resulting angle change is 0.0007°, which is the same as the diffraction limit of the AOD with a 40mm-aperture reference beam on the AOD.

3.1.3 Storage capacity and data transfer rate with AOD

For an Acousto-Optic Deflector (AOD), its center frequency and aperture together determine a very important parameter—the Time-Bandwidth Product (TBP). Namely, it is the product of the bandwidth and the switching time of the AOD.

$$TBP \equiv BT_s = B\frac{A}{\nu_a}, \tag{3.7}$$

where B is the bandwidth, T_s is the time it takes for the acoustic wave to travel through the aperture of the illumination.

In a holographic memory using angle multiplexing, the storage capacity is determined by the number of distinctive reference beam angles the addressing device can provide. This can be translated into the Space-Bandwidth Product (SBP) requirement of the addressing device. The SBP of an AOD is determined by its angular range and resolution, or equivalently, by its bandwidth B and the frequency change

$\Delta f_a|_{d.l.}$ which corresponds to an angle change at the diffraction limit.

$$SBP \equiv \frac{B}{\Delta f_a|_{d.l.}}. \tag{3.8}$$

From Equation 3.3, this frequency change can be derived as

$$\Delta f_a|_{d.l.} = \frac{\nu_a}{A}. \tag{3.9}$$

Therefore, the SBP of an AOD can be written as

$$SBP = B\frac{A}{\nu_a} = TBP. \tag{3.10}$$

The SBP of an AOD is the same as its TBP. As a result, the maximal number of holograms one can store in the holographic memory using AOD as the addressing device is determined by its TBP.

The switching time T_s of an AOD determines the speed of an AOD in changing the deflection angle of the output beam, and thus the data access and transfer rate of a holographic memory using AOD. It is important to note here that one can always increase the data transfer rate of the memory by reducing the aperture of the reference beam on the AOD. On the other hand, since the Time–Bandwidth Product of the AOD becomes less with smaller aperture, the total capacity of the memory decreases.

Another interesting point to notice here is that the diffraction efficiency (the ratio of the intensities of the diffracted (deflected) light and the incident light) can be controlled by the intensity of the acoustic wave [82]. In our experiment, we achieved a peak diffraction efficiency of 86% when the AOD was tuned at its center frequency 50MHz. However, in order to have uniform diffraction efficiency across the entire frequency range, we detuned the AOD to obtain a fairly flat response of 22%. This is shown in Figure 3.4. The main purpose of doing this was to simplify the exposure schedule.

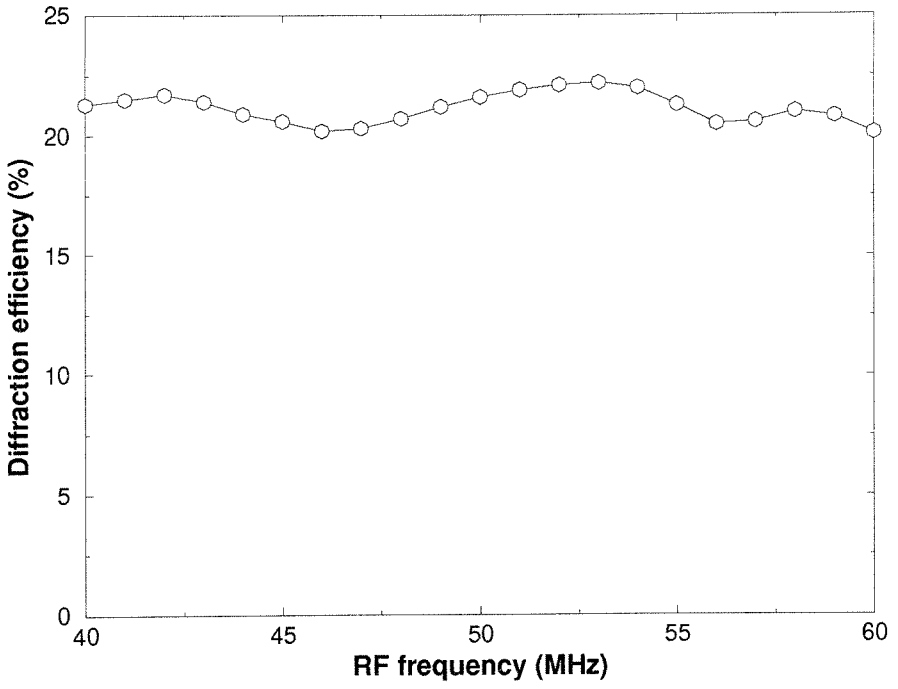

Figure 3.4: Diffraction efficiency of the AOD as a function of the RF frequency.

3.2 Recording using AOD and EOM

3.2.1 System design

In previous discussions, we showed that an AOD can be used in the reference beam to deliver angle deflection needed to address an angularly multiplexed memory. Here, we address the issues associated with holographic storage using the AOD.

We learn from Equation 3.2 that the frequency of the deflected reference beam from the AOD as a result of Bragg diffraction is shifted upward or downward by the frequency of the acoustic wave traveling inside the device. This effect is called Doppler shift. The amount of the frequency shift is the same as that of the acoustic wave (the driving RF signal) inside the device. This results in a rapid fringe movement when this diffracted beam interferes with the signal beam. Normally, the frequency shift is on the order of MHz, which is much faster than the frequency response of holographic materials. In order to create stationary interference patterns for effective holographic recording, a frequency–compensating device has to be used.

We can use another AOD (or an AO modulator (AOM) [78]) in the signal beam to

compensate for the Doppler shift. The drawback is that it introduces unwanted shift of the signal on the storage material. Alternatively, we can also use a single AOD before the laser beam is split into two—the signal and reference beams to create stationary interference pattern (this will be discussed later in this chapter). Another choice is to use an Electro–Optic Modulator (EOM).

An EOM utilizes the electro–optic effect of some materials to modulate their refraction indices, hence the phase of the optical beam traveling through, by an applied electric field. It normally operates on the "raw" laser beam and output an undeflected phase–modulated beam. When an EOM is driven by an AC signal at frequency ω_a, the output beam contains many spectral orders [83],

$$e^{j\omega t}e^{j\delta\cos\omega_a t} = e^{j\omega t}\left(J_0(\delta) + J_1(\delta)e^{\pm j\omega_a t} + J_2(\delta)e^{\pm 2j\omega_a t} + \cdots\right), \tag{3.11}$$

where ω is the light frequency. δ is the phase modulation index which can be written as

$$\delta = \frac{\pi n^3 \Gamma d}{\lambda} E, \tag{3.12}$$

where n is the refraction index of the material. Γ is the effective electro–optic coefficient. d is the thickness of the material and E is the amplitude of the AC signal. A graphic description of the operation of an EOM is shown in Figure 3.5. The first–order sidebands in the phase–modulated output are shifted by the frequency of the driving electric field (the AC signal). Therefore, if we use the same RF signal that controls the AOD to drive the EOM as well ($\omega_a = 2\pi f_a$), one of the first–order sidebands generated by the EOM will interfere with the Doppler–shifted reference beam to form a stationary interference pattern for holographic recording.

We constructed an experimental setup using an AOD and an EOM for holographic storage. It is shown in Figure 3.6. In the reference path, a CrystalTech Model 4050–2 Te_2O_3 slow–shear AOD was used to deflect the reference beam for angle multiplexing. The center frequency of the AOD is 50MHz. The acoustic velocity inside the Te_2O_3

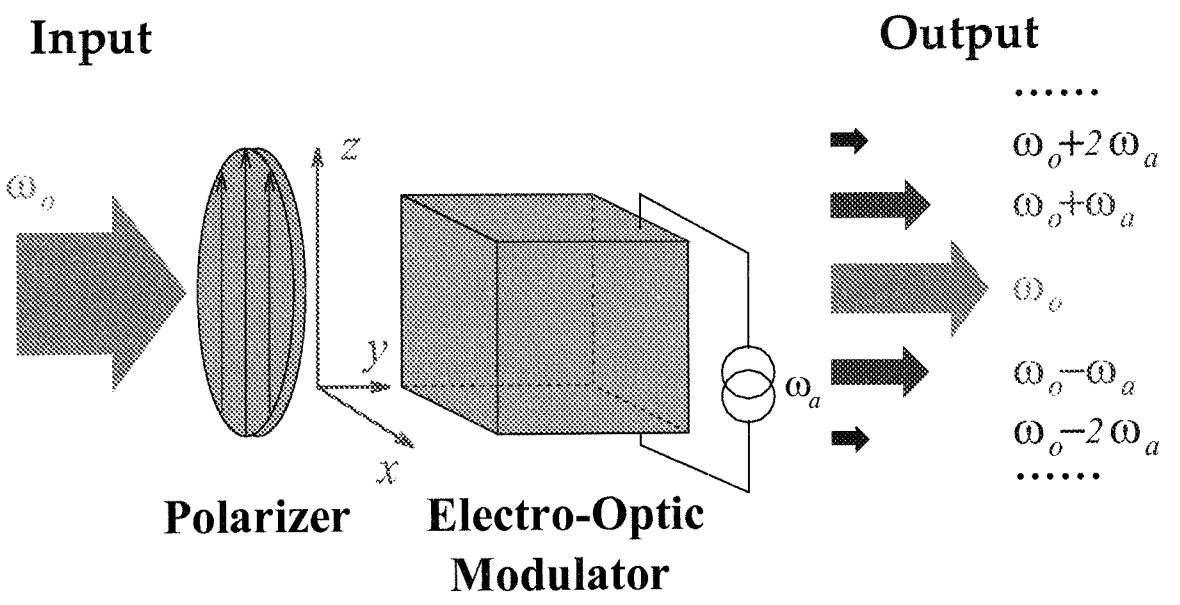

Figure 3.5: Operation of an EOM.

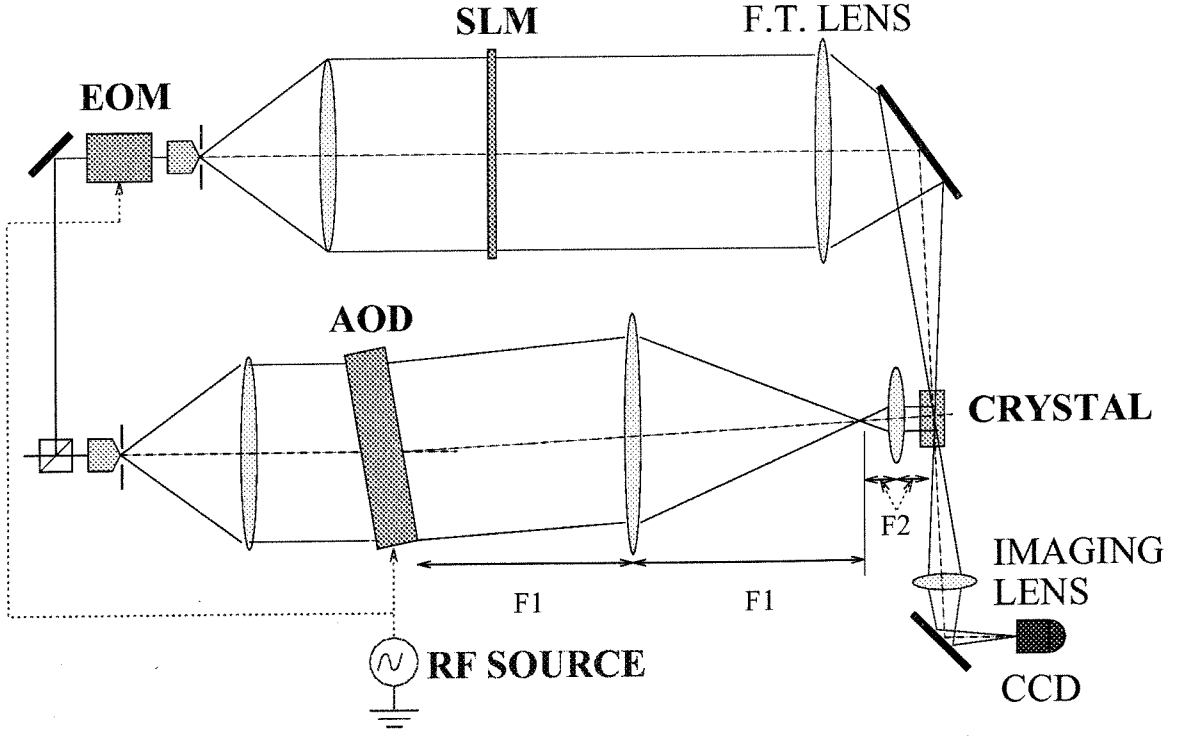

Figure 3.6: Experimental setup of a holographic storage system using AOD and EOM.

material is ν_a = 617 m/sec. The Time–Bandwidth Product of this AOD is 1,000. For the full frequency range from 33MHz to 66MHz, we obtained a deflection of 1.5°. The AOD was placed at the center of a 4F system consisting of two cylindrical lenses. This configuration guaranteed the usage of the full 40mm aperture on the AOD for optimal multiplexing capacity. Another 5:1 4F system imaged the deflected beam onto the crystal, demagnifying the illuminated aperture of the AOD to 8mm. In the meantime, the angle deflection at the crystal was expanded to 7.5°. The frequency corresponding to the first–null angle of the Bragg selectivity was 33KHz. Given the total frequency range of the AOD as 33MHz, the number of holograms that can be stored by angle multiplexing at the first–null is 1,000, the same as the TBP.

In the signal path, we used a New Focus Model 4002 broadband Electro–Optic Modulator (EOM) using Mg–doped $LiNbO_3$ to operate on the "raw" signal beam. The phase–modulated signal beam was then expanded, illuminated the SLM, and directed to the crystal. The crystal was placed in front of the Fourier–transform plane for Fresnel–plane recording. In this experiment, the SLM consisted of a slide which was rotated several degrees between each exposure. Imaging optics was used following the crystal and delivered the reconstruction onto the CCD detector—a Photometrics Imagepoint CCD array.

3.2.2 Optimizing the EOM

As shown in Equation 3.11, only one of the first–order sidebands among the many spectral orders in the output from the EOM is capable of compensating the Doppler–shifted reference beam. This raises a serious problem in the overall light efficiency in the signal beam since most of the signal intensity is wasted. In addition, to maintain proper modulation depth for holographic recording, more power has to be allocated to the signal beam. As a result, the wasted yet strong signal gives rise to noise due to signal beam fanning and the buildup of noise gratings (for detailed analysis of this type of noise source, refer to Chapter 6).

To achieve optimal light efficiency and noise performance in holographic recording,

it is essential to improve the efficiency of the phase modulation, which is the portion of the light intensity obtained in the first–order sideband. The light efficiency can be improved by optimizing the phase modulation index δ (Equation 3.12), which is directly controlled by the amplitude of the AC signal used to drive the device. The EOM we used in the experiment was a broadband device with $V_\pi = 97\text{V}$ [1]. The maximum RF power we can use to drive the modulator is 10W. Therefore, with a 50ohm load in parallel with the 30pF modulator crystal, the peak voltage is 32V which only induces about 1 radian phase shift. In other words, δ is only 1. This yields a value for $J_1(\delta)$ of 0.44. Only about 19% of the input optical power is converted into each first–order sideband, and almost all the rest of the power ($\approx 59\%$) remains in the baseband. As a result, 80% of the signal intensity is not only wasted, but also causes a lower modulation depth for recording. This results in a much smaller $M/\#$ (< 0.1) [2] than $\approx$0.8 which was obtained under normal recording conditions. In addition, the holographic noise increases significantly as we tried to balance the writing beams to improve the modulation depth.

To improve the performance of the EOM, we can raise the phase modulation index δ by applying a DC bias on top of the RF signal. However, because of the relatively high V_π and the parallel resistance load, this requires a high–power DC voltage supply which was not available at the time. Instead, we partly solved the problem by optimizing the polarization state of the signal beam.

It should be noted first that most of the wasted light remains in the baseband and the unused first–order sideband. Because the unused sideband always coexists with the useful one, we can only manage to reduce the residual power in the baseband to improve the efficiency.

An EOM controls the phase of an optical beam by modulating the refraction index of the material via the electro–optic effect. For pure phase modulation, the input optical beam should be polarized along the eigen–axis whose refraction index is modulated by an applied electric field. Otherwise, amplitude modulation occurs

[1] the amplitude of the driving electric field to generate a π phase shift ($\delta = \pi$).

[2] Another reason for the low $M/\#$ is the low damage threshold of the EOM. We could not adjust the beam ratio freely to optimize the modulation depth.

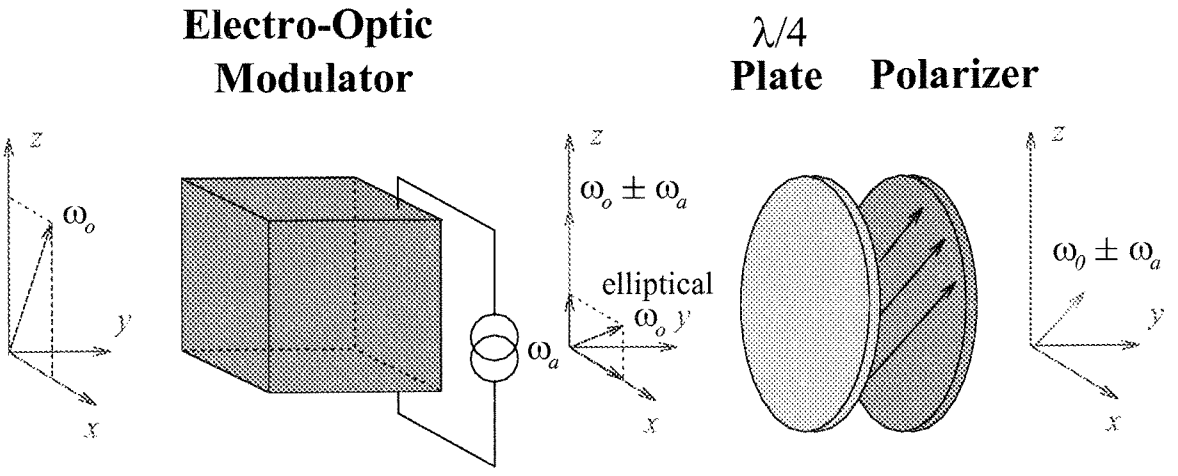

Figure 3.7: Optimizing the performance of the EOM.

because of the interaction between the phase–modulated and the unmodulated eigen–modes. We can take advantage of the amplitude modulation to eliminate (or reduce) the baseband (unmodulated) signal. When the input polarization is not along the designated axis for pure phase modulation, the baseband signal and the unmodulated light along the other axis [3] generally form an elliptically polarized light at the exit. The axes ratio and the direction of rotation of this elliptically polarized output depend on the input polarization, the phase modulation index δ, and the phase delay between the two eigen–modes at the exit. A quarter–wave plate can be use to convert this elliptically polarized light back to linear polarization. Normally, this linearly polarized light (unmodulated) can not be completely eliminated because it is unlikely to be exactly orthogonal to the polarization of the modulated sideband signal. However, by adjusting the input polarization and output polarization optics, we could achieve less than 10% baseband in the output signal —a reduction by a factor of 5, at the expense of some loss ($\approx$ 35%) in the signal intensity. The two first–order sidebands retained 45% of the total intensity. The optimization of the EOM is graphically summarized in Figure 3.7.

Another optimization was achieved later by using a Liquid Crystal Phase Retarder

[3] In Mg:$LiNbO_3$ which is used in our EOM, only one eigen–mode (z or c–axis) experiences significant modulation by the applied field (along z) on the refraction index (> 100 times stronger than others).

(LCPR) (Meadowlark D1040). Instead of a quarter-wave plate, we used a half-wave plate to align the output polarization with the Liquid Crystal Phase retarder (whose fast and slow axis are at 45° to the lab axes). By controlling the phase difference along the two eigen-modes of the LCPR, we achieved nearly complete elimination of the baseband signal at the expense of 55% loss in the signal intensity. This configuration also verified the phase delay between the two eigen-modes in the EOM we measured using the method found in [84].

It should be noted that our optimization goal is not obvious, as we need to increase the amount of light modulated by the correct order, as well as minimizing the light remaining in the unmodulated baseband. The exact determination of the configuration depends on the relative importance of these two parameters. The loss of the signal intensity has to be considered as well because the maximal power we can allocate in the signal is limited by the damage threshold of the EOM, which is only 200mW. As a result, the optimization we achieved was evaluated by the degree of success in our later demonstration of holographic storage using the AOD and EOM.

With the optimization, the $M/\#$ of the system was raised to 0.34, corresponding to an increase in the diffraction efficiency by a factor of 10. The holographic noise was also significantly suppressed.

3.2.3 Storage of 1,000 holograms

Using the AOD and EOM1,000 holograms were recorded in the memory system. The exposure schedule was generated with a time constant of 510 seconds, starting from the last exposure time as 0.48 seconds. The first exposure was 11.3 seconds. The average diffraction efficiency of the holograms was 6.5×10^{-8}. A quick scanning by the AOD at the end of recording was performed to retrieve some of the recorded holograms. The result is shown in Figure 3.8, where hologram #0—50, 200—250, 400—450, 600—650, 800—850, and 950—1,000 were sampled at the corresponding RF frequencies used for recording.

Some of the sample reconstructions are shown in Figure 3.9. The detected pixel

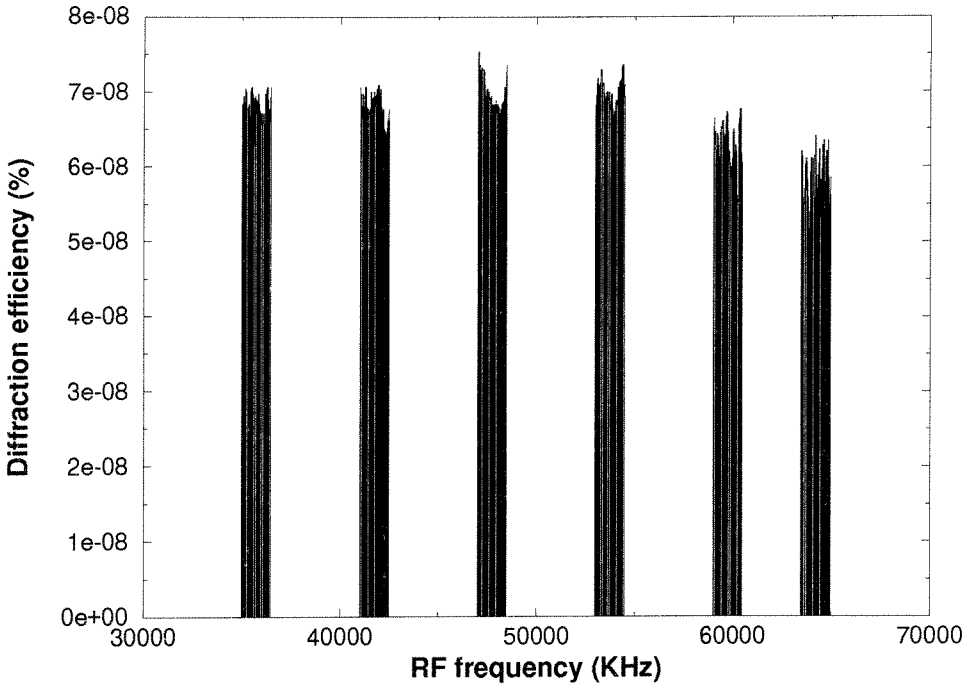

Figure 3.8: Diffraction efficiency of the sampled holograms.

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 in this system. The PDFs of OFF and ON pixels from a sample reconstruction are shown in Figure 3.10. Totally, approximately 200,000 pixels from the image are represented. An optimal threshold was empirically determined, giving a measured raw probability of error of 10^{-5}. The reason for the measured errors is that we used the first null of the selectivity curve as the inter-hologram spacing. Therefore, crosstalk noise was much higher than in the previous experiments and it became even worse when the reference beam deviated from its orthogonal direction.

3.2.4 Large-scale memory using AOD and EOM

Here, as a combination of the system architecture described in Chapter 2 and the above demonstration, we present a page-formatted fast random-access holographic

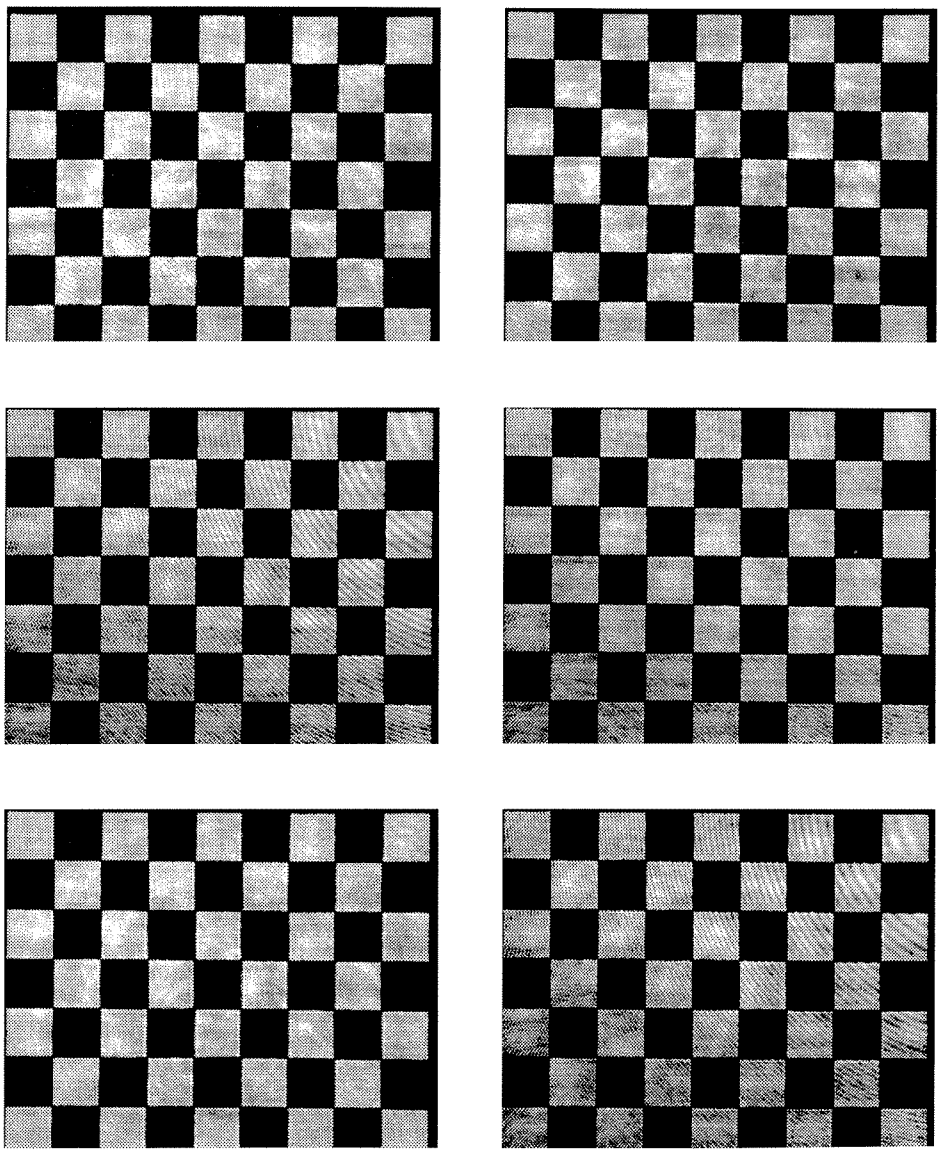

Figure 3.9: Sample reconstructions from the 1,000 holograms recorded using AOD and EOM.

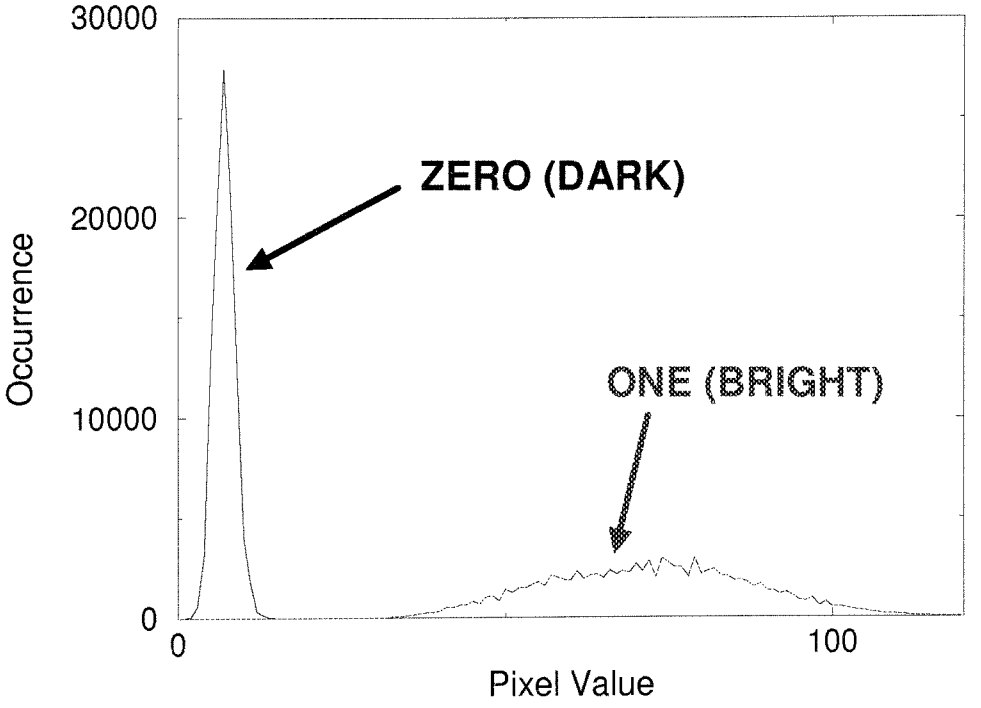

Figure 3.10: Histograms from one sample reconstruction.

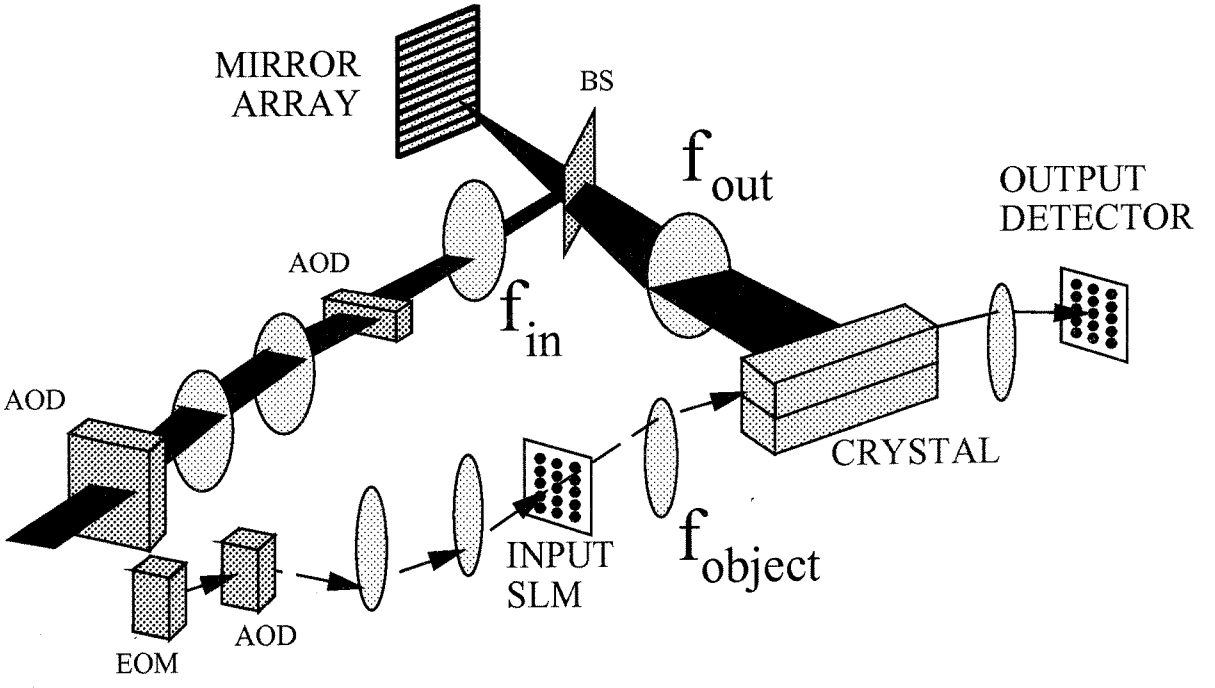

Figure 3.11: Large-scale random-access holographic memory desgin using AOD and EOM.

memory to store 160,000 holograms. The segmented mirror array (Section 2.1.3) allows rapid access to any of the stored holograms with the AOD as the addressing devices. The memory consists of 16 vertically spaced locations, each containing 10,000 holograms. Each location is organized as 10 fractal rows of 1,000 angularly multiplexed holograms. This assignment requires a horizontal SBP of 1,000 and a vertical SBP of 160, comfortably within the capabilities of commercial AODs. The total storage capacity is 160Gbits and a random access time of less than 100μs.

The system design is shown in Figure 3.11 [41,85]. A segmented mirror array and two crossed AODs allow the reference arm of the system to control both the position and angle (horizontal and vertical) of incidence of the reference beam. Another AOD is used in the signal arm to direct the signal beam onto each of the 16 storage locations. The Doppler shift introduced by the AODs is compensated by an EOM.

3.3 Single–AOD setup

The system diagram of the experimental apparatus is shown in Figure 3.12. We used the same slow–shear Te_2O_3 AOD with a 42mm×8mm aperture. In the signal arm, we used a 5:1 imaging system to illuminate the SLM (Epson liquid crystal television device, 480×440 pixels in an area 27mm×20mm). A second imaging system transferred the page displayed on the SLM onto a cooled CCD (752×480 pixels) whose video output was digitized and transferred to the computer memory for evaluation. The $LiNbO_3$ crystal ($2\times1\times1cm^3$, 45°–cut) is placed just beyond the Fourier–transform plane at the center of this second imaging system. In the reference arm, a 1:1 imaging system was used to image the AOD aperture onto the crystal so that the two beams interact at right angles.

The angular selectivity of this system is

$$\delta\theta = \lambda/A \approx 0.0035°,$$

where λ=488nm, and A=8mm is the diameter of the illuminated region. From Equa-

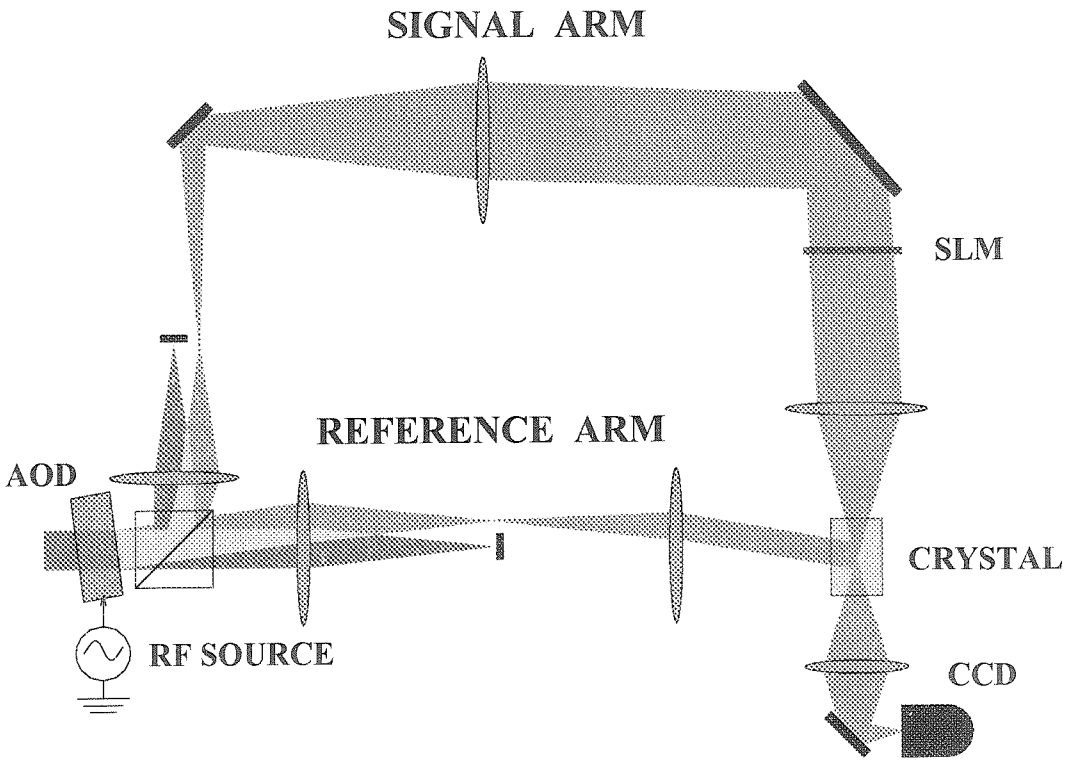

Figure 3.12: Single–AOD setup.

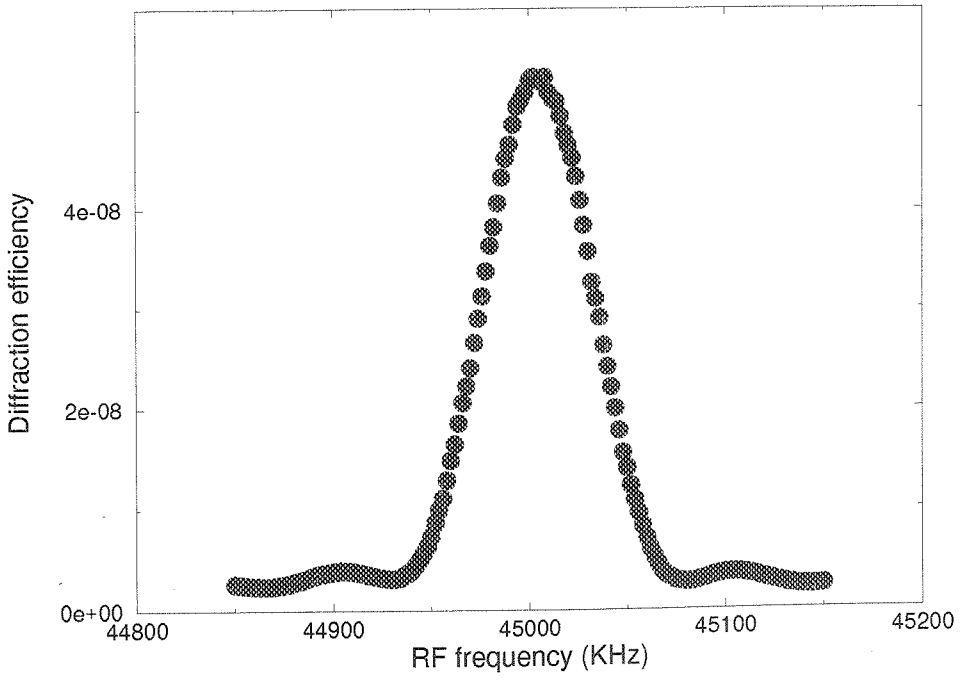

Figure 3.13: Angle selectivity in the single–AOD setup.

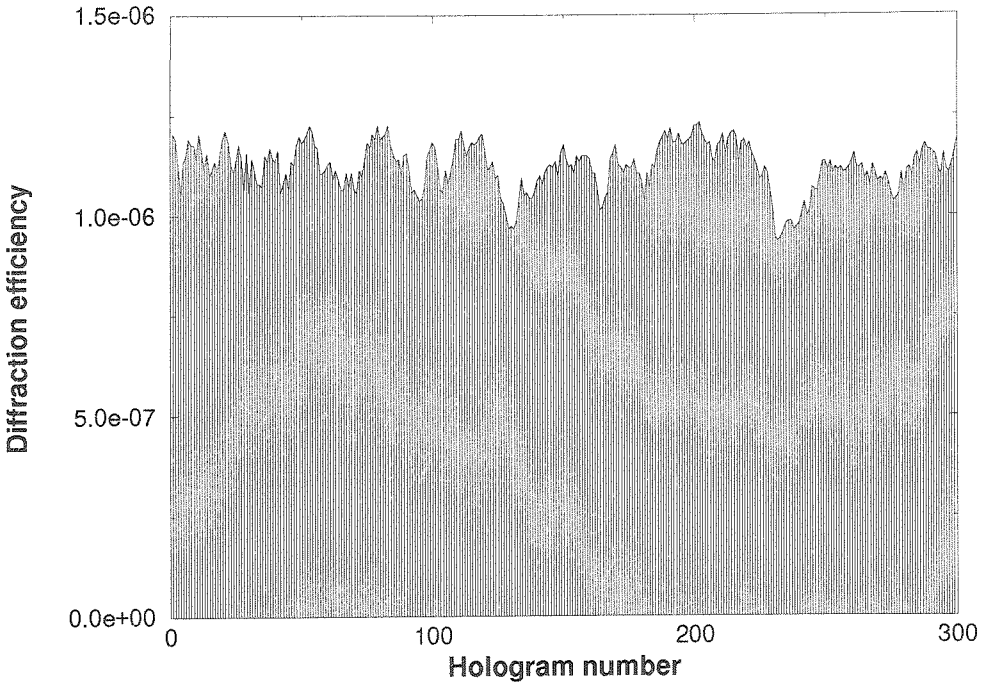

Figure 3.14: Diffraction efficiency of 300 holograms stored in the Single–AOD setup.

tion 3.3, the minimum frequency distance between two neighboring holograms is approximately 75 KHz. Given the bandwidth of the AOD as 33 MHz, the maximum number of holograms we can angularly multiplex in this system is about 400.

The measured angular selectivity of a recorded chessboard pattern is around 80 KHz as shown in Figure 3.13. Using 80 KHz as the frequency spacing, we stored a total of 300 holograms. Of these holograms 11 were chessboard patterns used for measuring the performance of the reconstructions (evenly distributed in the recording sequence), another 11 were all–ON patterns used to normalize the rest of the holograms, and the rest were random–bit patterns. Figure 3.14 shows the diffraction efficiency of the 300 holograms. The average diffraction efficiency is 1.12×10^{-6}. The maximum deviation of the diffraction efficiency from the mean value is less than 15%. The average SNR is 3.9, and the probability of error is 1.2×10^{-5}.

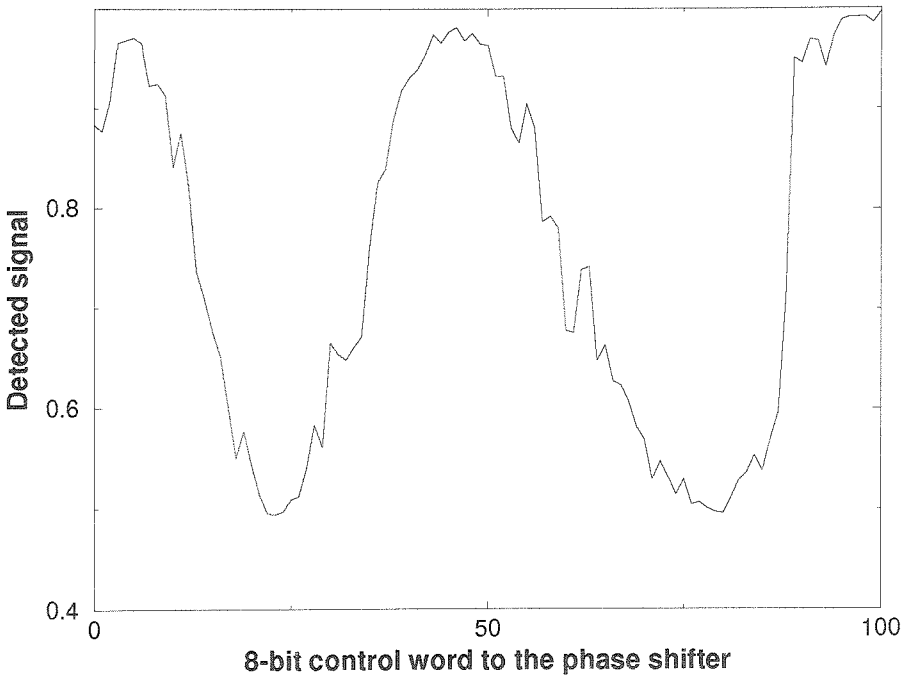

Figure 3.15: Phase modulation by the AOD.

3.4 Optical Phase–Lock Loop (OPLL)

Phase stabilization techniques are important for interferometric systems including holographic memory that are very sensitive to the environmental phase perturbations. Passive phase–compensation was proposed to stabilize an interferometric system by using circularly polarized beams whose intensity is not sensitive to the phase drift [86]. However, for the application of holographic memories and characterization of holographic materials, active phase–compensation techniques have to be used because of the direct influence of phase perturbations on holographic recording. In most of the phase–stabilization system reported, piezo–mirrors were used in one of the optical paths to generate a "beat" signal for heterodyne detection of the phase drift and a displacement for phase compensation [87–90]. The use of an Electro–Optic Modulator (EOM) to actively compensate the phase perturbations was also reported [91, 92].

The holographic memory system we constructed using AOD and EOM came with an extra bonus: it offers a convenient way to construct a phase-compensation system to stabilize holographic recording. The phase difference between the two writing

beams can be retrieved by interfering different temporal–frequency components in the beams for heterodyne detection. Phase compensation can be implemented by phase–modulating the AOD or the EOM. To demonstrate phase modulation by the AOD, we constructed an interferometer using the signal and reference arm of the memory system. The AOD and EOM were both operating at 50MHz. The RF signal that drove the AOD first passed through a narrow–band delay–line phase shifter which was controlled by a computer. A photodetector was placed at the interference pattern, behind an iris. The fringe movement caused by the change in the phase of the optical arm containing the AOD resulted in an alternating signal, as shown in Figure 3.15. Ideally, this alternating signal is sinusoidal as the result of phase "rotation" by the phase shifter. However, non–ideal behavior and fluctuations were observed due to the non–linear response of the phase shifter and the phase drift in the environment.

3.4.1 Optical phase stabilization using AOD and EOM

We designed and demonstrated an Optical Phase–lock Loop (OPLL) using AOD and EOM. The optical system of the phase–lock loop is the same as the normal holographic recording system shown in Figure 3.6. As a result, holographic recording and active phase stabilization can be carried out at the same time.

Because effective heterodyne detection requires the removal of spatial variations in the interference at the detector, holograms are needed to make the interfering beams co–linear. There are many "beat" signals available during holographic recording. They are generated by the interference between

- diffracted reference and the unmodulated baseband signal with "beat" frequency f_a;
- diffracted reference and the unused first–order sideband signal with "beat" frequency $2f_a$;
- diffracted reference and the other higher–order sideband signal with "beat"

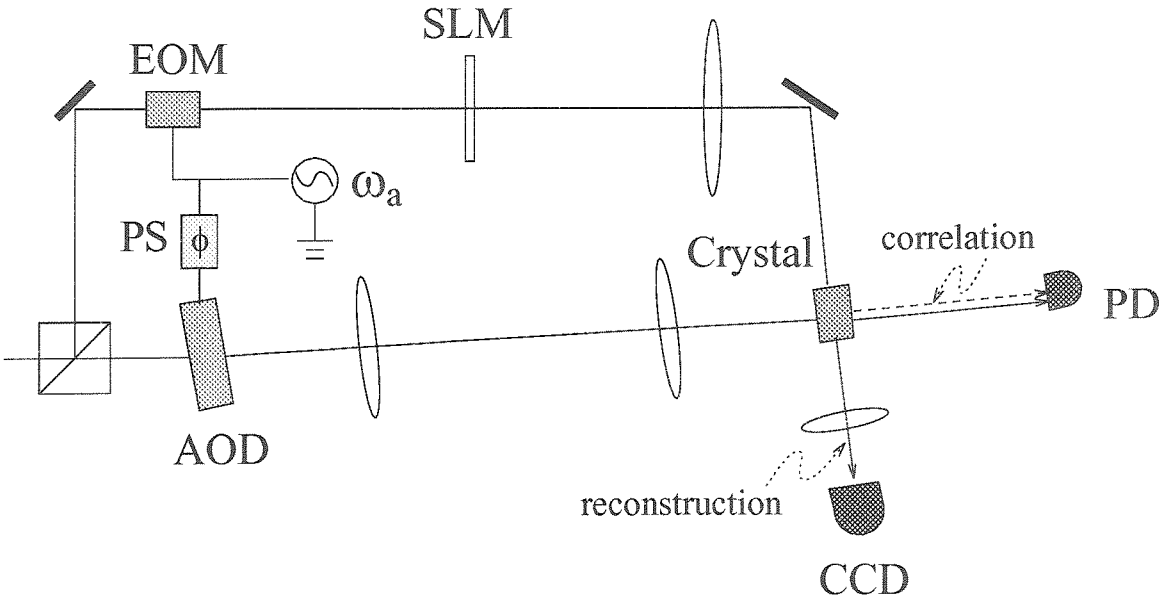

Figure 3.16: Phase stabilization system using AOD and EOM: PD: photodector; PS: RF phase shifter.

frequency as multiples of f_a;

- undiffracted reference and the sideband signal with "beat" frequency $\pm f_a, \pm 2f_a, \cdots$.

After the optimization of the EOM as described in the previous section, the baseband and higher–order sidebands in the signal are greatly suppressed. Therefore, the heterodyne signal generated by them is too weak to be detected by the high–speed photodetector we used. The interference between the undiffracted reference beam and the first–order sidebands in the signal yields much stronger signal. However, because the undiffracted reference interferes with both the plus and minus first–order sidebands, the resulting signal is an AM signal with a carrier frequency the same as the acoustic frequency $f_a \approx 50$MHz. This makes it very difficult to retrieve the phase information embedded in the amplitudes. A possible solution is to mix the AM signal with a local oscillator with the same frequency. But the Low–Pass Filter (LPF) output is noisy because it contains a DC term much higher than the perturbation which carries the phase information. In addition, the undiffracted reference and the baseband can write the same hologram continuously, consuming a large portion of

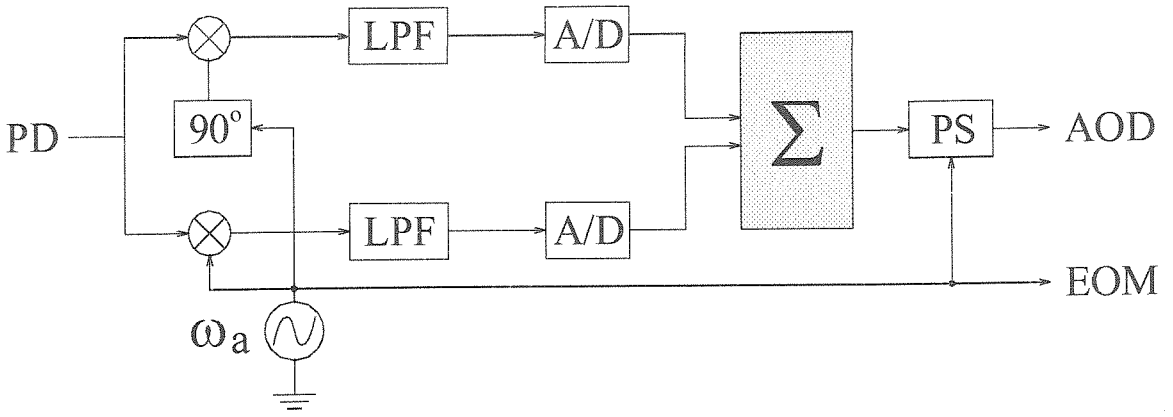

Figure 3.17: Processing in the computer–based PLL.

the dynamic range.

To simplify matters, we did not monitor the phase all the time in our experiment. Instead, we only monitored and compensated the phase difference before the recording of a hologram. The heterodyne signal was received along the "correlation" direction—between the transmitted first–order reference beam and the diffracted baseband signal (EOM was turned off). This configuration (as shown in Figure 3.16) yielded more intensity at the detector and generated an FM signal which is much easier to process. The drawback is that we can not track and compensate the phase drift during recording.

3.4.2 Phase stabilization by OPLL

A hologram was written in a DuPont photopolymer to a diffraction efficiency of 5%. The EOM was then turned off—the output signal was unmodulated. The transmitted reference beam interfered with the diffracted signal, resulting in a heterodyne signal which was in the form of an FM signal

$$\cos(2\pi f_a t + \phi(t)),$$

where $\phi(t)$ is the phase difference between the signal and reference beam. This heterodyne signal was then detected by a high–speed photodetector (Photodyne 1500XP).

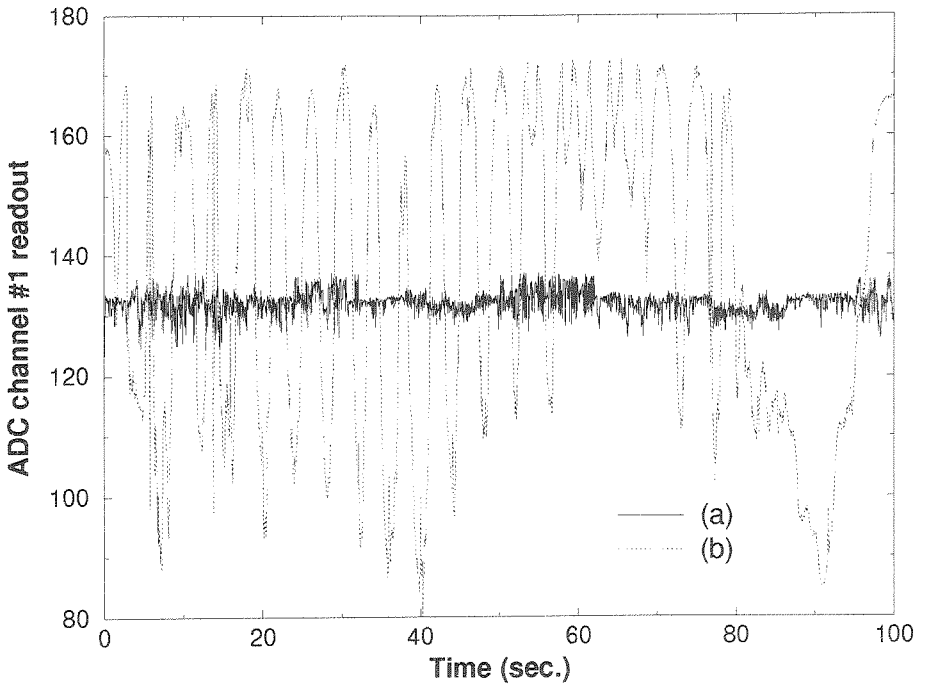

Figure 3.18: Experiment with optical PLL: (a) with phase stabilization; (b) without phase stabilization.

The detected signal was processed by a computer–based phase–compensation unit shown in Figure 3.17. A frequency reference was used as the local oscillator. After the low–pass filter, the sine and cosine of the phase information were sampled into the computer, which acted basically as an integrator and controlled a narrow–band RF phase shifter to generate a phase correction. The output of the phase shifter was directed to the AOD where the electronic phase in the acoustic signal was converted into the optical phase in the diffracted reference beam to compensate the phase drift.

The control algorithm was written in Turbo C^{++}. The sampling rate of the system was 20KHz, and the processing rate was 250Hz. For demonstration of phase stabilization by the system, we monitored one of the digitized signals to the computer as an indicator of the phase variation in the holographic memory setup. This signal is shown in Figure 3.18(a). For comparison, we show a signal obtained separately without phase stabilization in the same plot (Figure 3.18(b)). Both of the curves were obtained when the setup was subjected to mechanical vibration. Statistical analysis in the received signal over a period of 8 hours showed that, with phase stabilization,

the mean phase variation was zero with a standard deviation of 2.1° (worst case was 5.8°). Without the stabilization, the standard deviation was 23° (worst case was more than 180°).

3.4.3 Coherent addition and erasure of holograms

An optical phase stabilization system is not only capable of compensating for the phase drift in a holographic storage system; it can also be used to maintain a certain phase difference between the writing beams. An application of this is the coherent erasure of holograms [93–97].

Coherent addition and erasure of holograms was demonstrated in a similar setup using a real–time recording material—a $LiNbO_3$ crystal. In the experiment, one plane–wave hologram was first written to a certain level (10^{-5} diffraction efficiency) before the phase–lock loop was turned on [4]. Then this hologram was repeatedly written by the same exposure. Before each recording, the phase–lock loop was activated to stabilize the phase at either ϕ_0 or $\phi_0 + \pi$, depending whether it was for addition (positive hologram recording) or erasure (negative hologram recording)[5]. This process was carried out for multiple times and the diffraction efficiency of the hologram was observed at the reconstruction direction. The result is plotted in Figure 3.19. At the end of the experiment, a slight shift in the readout frequency from the recording frequency had to be used to obtain peak efficiency. We suspect that this is because of the hologram–bending effect as a result of the prolonged exposure [98–100].

3.4.4 Improved OPLL design for holographic memory

The above experiments demonstrate the use of an optical phase stabilization system to compensate and control the phase difference between the two writing beams in a holographic memory setup to stabilize hologram recording and add or subtract

[4] This is because of the noise level of the detector and the limited dynamic range of the ADC we used.

[5] ϕ_0 is a constant representing the mean of the phase difference between the two writing beams determined by their path difference.

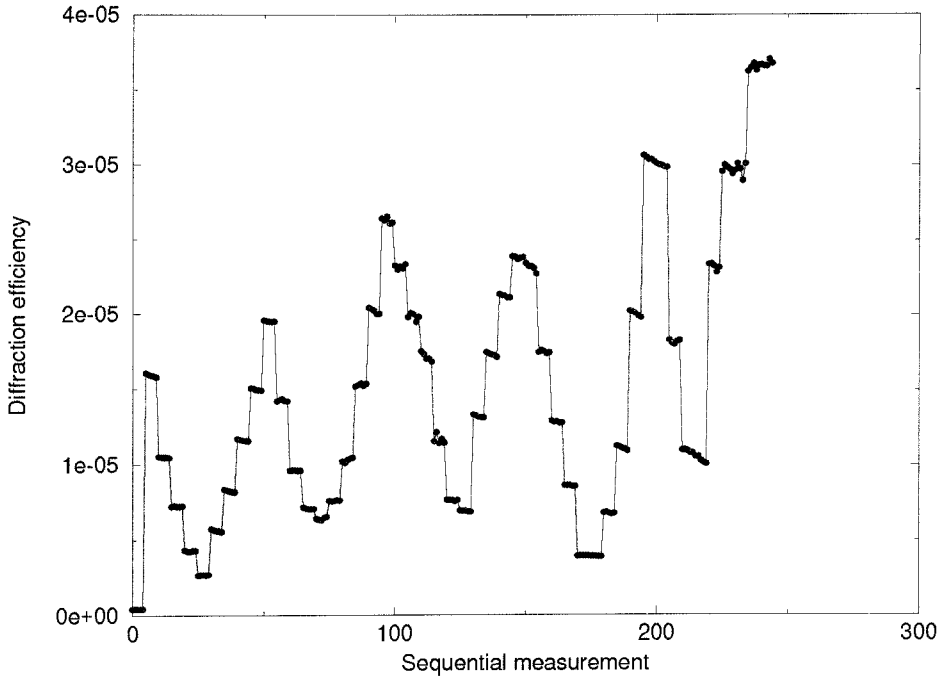

Figure 3.19: Coherent addition and subtraction.

holograms coherently. However, the system has several drawbacks:

- The RF circuitry is difficult to construct and noisy.
- The use of a computer and a narrow–band RF phase shifter to implement phase compensation limits the frequency response of phase stabilization and the range of multiplexing in multiple hologram recording.
- Phase compensation is not real–time during holographic recording. This would be good enough for normal recording condition in the lab where the phase drift is not significant over a short period of time. But it is problematic in some applications where holographic recording is conducted on site with more phase fluctuation.

Here we propose a new system design which improves the overall performance of the optical phase stabilization system. The setup is shown in Figure 3.20. The difference between this new design and the old system is the use of a "probe" beam in the reference arm. This probe beam is generated by another acoustic frequency which

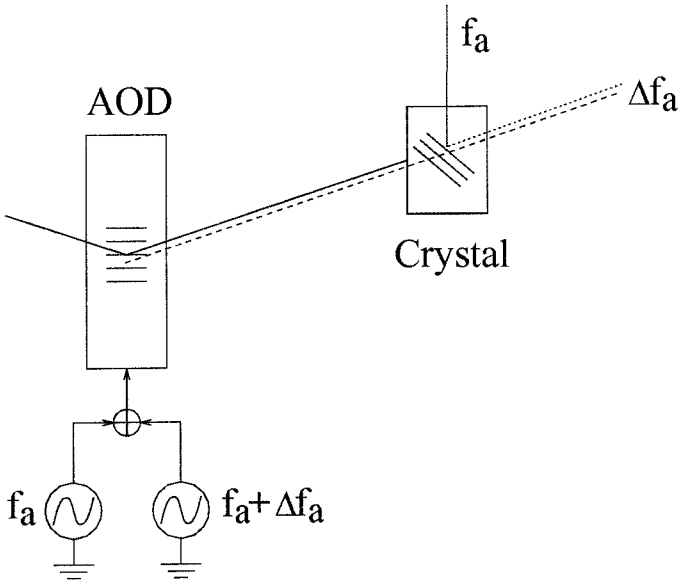

Figure 3.20: Improved PLL system design.

is shifted by Δf_a from what is used to deflect the reference beam for multiplexing. Δf_a is chosen to be faster than the frequency response of the holographic material to avoid recording unwanted holograms. In the meantime, Δf_a shall be well within the diffraction limit to make the probe beam and the reference beam nearly co–linear. This probe beam interferes with the diffraction off the hologram by the signal beam at the detector along the correlation direction to generate the heterodyne signal. An LPF filters out all higher–order harmonics except for the FM "beat" signal created by the probe and the diffracted first–order sideband signal which oscillates at Δf_a. For a $LiNbO_3$–based memory system, Δf_a can be chosen within the range from several KHz to tens of KHz. Therefore, a novel PLL circuit can be used to sense the change in the phase and compensate for it. The phase stabilization can also be implemented at any time.

This new design is inspired by and bears a lot of resemblance to the systems that measure the amplitude and phase of a holographic grating by direct heterodyne detection [101, 102]. Therefore, the same system can be used to monitor the buildup of a space–charge field in photorefractive materials during recording.

Chapter 4 Thermal fixing of 10,000 holograms

Contents

4.1 Theoretical model 80

4.1.1 Thermal fixing efficiency 82

4.1.2 Numerical evaluation 85

4.1.3 Dynamics in hologram revealing 87

4.1.4 The M/# of thermally fixed memory 90

4.2 Application of thermal fixing in large–scale memory . . 91

4.2.1 Experiment procedure 91

4.2.2 SNR degradation with thermal fixing 93

4.3 Incremental fixing schedule 96

4.3.1 Dynamics and exposure schedule 96

4.3.2 Development of incremental fixing schedule 99

4.3.3 The M/# with incremental fixing schedule 102

4.3.4 Thermal fixing of 10,000 holograms using incremental fixing schedule 103

4.4 Miscellaneous topics 107

4.4.1 Multiple recording and fixing treatments 107

4.4.2 Repeated recording and fixing treatment 109

4.4.3 Selective fixing procedures 111

4.5 Appendix: Derivation of the theoretical model 111

In this chapter, we discuss thermal fixing to solve the volatility problem of holographic storage systems using $LiNbO_3$:Fe. The volatility of the stored data is a serious impediment for the practical realization of photorefractive holographic memories.

The same charge excitation and transport mechanism responsible for the writing of holographic gratings also erases them upon further illumination, leading to the loss of stored information during readout. Amodei and Staebler [103] found that optically stored holograms in $LiNbO_3$:Fe could be stabilized against readout by heating the crystal during or after the writing process to above $100°C$ followed by uniform illumination at room temperature. They proposed that at high temperatures, where the ionic conductivity dominates the electronic conductivity, the light–induced space charge distribution is compensated by ionic space charge of opposite sign. At room temperature, the ionic conductivity is very small compared to the photoconductivity of the electrons. Therefore, the ionic replica of the holographic gratings remains nearly intact and the stored information can be read out without erasure. Bollmann et al. [104] suggested that OH^- ions which enter the crystals during growth were likely responsible for thermal fixing. The presence of these ions can be detected by their optical absorption at around 2.87μm. Vormann et al. [105] confirmed the importance of OH^- concentration in the fixing process, but identified the associated H^+ protons as the mobile ions. This was further corroborated by Klauer [106] in 1991.

Theoretical models have been developed to describe the formation of thermally fixed holograms in $LiNbO_3$:Fe [107–111], based on a set of equations similar to those first put forward by Kukhtarev [27, 34]. Thermal fixing of multiple holograms in $LiNbO_3$:Fe has also been demonstrated [112–114].

In this chapter, we start with the introduction to the theoretical treatment of thermal fixing. We present experimental characterization of the effects of the thermal fixing process on system error performance of a large–scale holographic memory. We show that when more than 1,000 holograms are stored, the loss in the Signal–to–Noise Ratio (SNR) due to the fixing process is a major obstacle to the implementation of non–volatile, large–scale holographic storage systems. We propose and experimentally demonstrate a novel "incremental fixing schedule" which improves the overall fixing efficiency. We demonstrate experimentally that for 10,000 fixed holograms, this new fixing method yields an SNR that is better than the one obtained when we simply record and evaluate 10,000 unfixed holograms.

4.1 Theoretical model

Although there are still some questions remaining in the interpretation of the thermal fixing process in $LiNbO_3$:Fe [115,116], it is generally believed that it arises from the interaction of two simultaneous effects. Free electrons can be excited by illumination at suitable wavelengths from available donor Fe^{2+} sites into the conduction band. They migrate and become trapped at acceptor Fe^{3+} sites, leading to space charge redistribution. At the same time, protons in the material can move under the effect of an electric field or a concentration gradient if the temperature is high enough. Thermal fixing can be carried out in one of two ways. The first, described graphically in Figure 4.1, consists of three steps—recording, heating, and revealing. During holographic recording, upon illumination by spatially–varying intensity patterns, photo–excited electrons migrate under various transport mechanisms and become trapped at different acceptor sites, resulting in spatially modulated charge concentration which is a replica of the exposing light pattern. This generates a space–charge field which, in turn, gives rise to the spatial modulation of the refractive index of the material via the electro–optic effect. After the recording is finished, the crystal is heated at elevated temperatures where the protonic conductivity dominates. The field–induced proton migration compensates the spatially–varying electronic charge distribution. In the revealing process, the crystal is cooled down to room temperature and illuminated by uniform light, either coherent or incoherent. The electronic gratings are partially erased, leaving a net space–charge field which is stable against further optical erasure.

In the second scheme, holograms are recorded in the crystal at elevated temperatures. Because of the continuous compensation by the mobile protons, the net space–charge field is kept small. Since it is the space–charge field which inhibits further holographic recording and causes most of the scattered noise, stronger gratings with less noise can be recorded in this way [112]. However, since the revealing process has to be carried out at lower temperatures, the fixed holograms suffer from distortion due to thermal contraction and change of refraction index of the material [113]. In this chapter, we present the theoretical treatment and experimental results of thermal

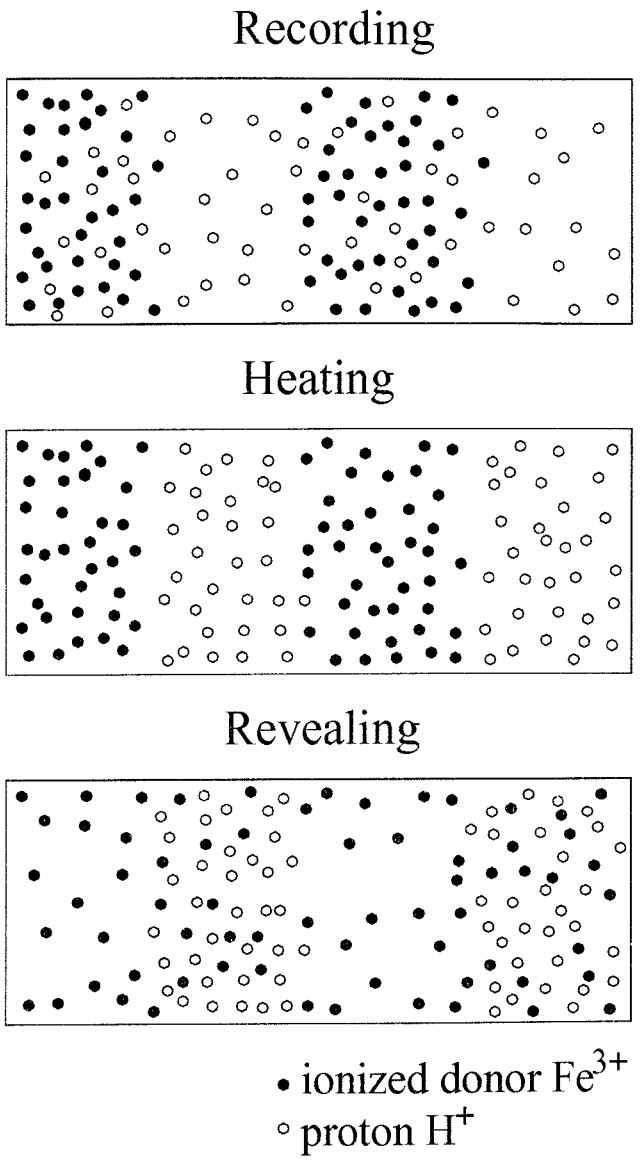

Figure 4.1: Schematic of thermal fixing process.

fixing using the first scheme, leaving the second one as the topic of the next chapter.

4.1.1 Thermal fixing efficiency

Both light-induced and thermally induced redistribution of charge in the photorefractive material can be described with a set of equations similar to the Kukhtarev equations. The following is the simplified one-dimensional version, where x corresponds to the **c**-axis of the $LiNbO_3$:Fe crystal.

$$\frac{\partial N_D^+}{\partial t} = (SI+\beta)(N_D - N_D^+) - \gamma_R N_D^+ n_e, \tag{4.1}$$

$$\frac{\partial n_e}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{q}\frac{\partial J_e}{\partial x}. \tag{4.2}$$

$$J_e = q\mu_e n_e E + K_B T \mu_e \frac{\partial n_e}{\partial x} + P(N_D - N_D^+)I, \tag{4.3}$$

$$\frac{\partial n_p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x}, \tag{4.4}$$

$$J_p = q\mu_p n_p E - K_B T \mu_p \frac{\partial n_p}{\partial x}, \tag{4.5}$$

$$\epsilon\frac{\partial E}{\partial x} = q(N_D^+ - N_A + n_p - n_{p0} - n_e), \tag{4.6}$$

where N_D, N_D^+, N_A, n_e are the density of the dopants, ionized donors, acceptors, and free electrons, respectively. n_p is the proton density and n_{p0} is its average or initial value. $I = I_0\ (1+me^{-jKx})$ is the light intensity pattern, where m is the modulation depth, and K is the grating vector. S, β, γ_R, and P are all material parameters—the photorefractive cross-section, thermal generation rate, recombination rate, and photovoltaic constant, respectively. K_B is Boltzmann's constant. T is the absolute temperature. J_e and J_p are the electron and proton current density, respectively. μ_e and μ_p are the mobility of the electrons and protons. Under three basic assumptions:

- low modulation-depth ($m \ll 1$);
- $N_D, N_A \gg n_e$ [117];
- quasi-steady state (n_e reaches equilibrium almost instantaneously—the spatial modulation of free electron concentration builds up at a rate characterized by

$\tau = 1/(\gamma_R N_A)$, much faster than the others. Therefore, $\partial n_e/\partial t$ could be treated as 0 [1]).

Equations 4.1—4.6 can be solved analytically [110]. To further simplify matters, we neglect the thermal excitation ($\beta = 0$). The analytic solutions to the steady–state space–charge field after recording at room temperature is

$$E_{original}\,|_{t\to\infty} = -jmE_q\,\frac{(E_0 + E_{0ph}) + jE_d}{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + j(E_q + E_d)}, \tag{4.7}$$

where $E_{original}\,|_{t\to\infty}$ is the amplitude of the saturation space–charge field. The characteristic field terms used in the equation are E_q, the limiting space–charge field of dopants (represents the upper limit on the strength of the space–charge field imposed by the density of either the available donors ($N_D - N_A$) or the deep acceptors N_A), the diffusion field E_d, and the photovoltaic field E_{0ph}. They are defined as

$$E_q = \frac{qN_A(N_D - N_A)}{K\epsilon N_D}, \tag{4.8}$$

$$E_d = \frac{K_B T K}{q}, \tag{4.9}$$

$$E_{0ph} = \frac{P\gamma_R N_A}{q\mu_e S}. \tag{4.10}$$

Because the electronic grating strength is translated to the protonic grating strength via proton compensation in the fixing process and determines the final diffraction efficiency of the fixed hologram, we rewrite the space–charge field as a function of the electronic grating strength.

$$E_{original} = -j\frac{q}{K\epsilon}N_{D1}^{+}, \tag{4.11}$$

where N_{D1}^{+} is the amplitude of the electronic grating after recording. This is also a general relationship without steady–state assumptions.

[1] One should note that this is only because $\partial n_e/\partial t$ is much larger than, for example, $\partial N_D^{+}/\partial t$ at the beginning so n_e reaches equilibrium and could be treated as a constant while others are still changing.

After holographic recording, the crystal is heated to elevated temperatures where the protons become mobile. If we neglect the thermal decay of the electronic grating during heating by limiting the heating temperature well below $180°C$ (roughly the temperature needed to thermally ionize the donors [118]), the steady-state proton grating amplitude n_{p1} becomes

$$n_{p1} = -\left(\frac{1}{1 + E_d/E_{qp}}\right) N_{D1}^{+}, \tag{4.12}$$

where $E_{qp} = qn_{p0}/K\epsilon$ is the limiting space-charge field of the protons. The minus sign in the equation represents the 180° phase shift between the electronic and the protonic gratings. This indicates that, because of proton diffusion, the electronic grating can not be completely compensated by the proton grating. However, given $E_{qp} \gg E_d$ which is valid in most cases, we have

$$n_{p1} \approx -N_{D1}^{+}.$$

The electronic grating is nearly fully compensated by the protonic one and the net electric field after heating is close to zero. This is the reason that usually no diffraction can be detected right after heating.

During revealing at room temperature, the electronic grating is partially erased by uniform illumination, resulting in a net space-charge field whose steady-state solution is [111]

$$E_{fixed} = -j\frac{q}{K\epsilon}\frac{\frac{N_D-N_A}{N_D}E_{0ph} + jE_d}{\frac{N_D-N_A}{N_D}E_{0ph} + j(E_d + E_q)}n_{p1}, \tag{4.13}$$

where E_{fixed} is the amplitude of the space-charge field of the fixed grating. Here we assume that the revealing process is long enough so that the induced DC field completely screens the photovoltaic field, $E_0 = -E_{0ph}$. E_{fixed} modulates the refraction index of the material via the electro-optic effect to form the thermally fixed phase hologram.

We define the thermal fixing efficiency as the ratio between the diffraction efficiencies of the thermally fixed and the original hologram before fixing. Under small–signal assumption (valid for the application of large–scale holographic memories), the diffraction efficiency of a hologram is proportional to the square of the space–charge field. Therefore, the thermal fixing efficiency can be written as

$$\eta_{fixing} \equiv \frac{\eta_{fixed}}{\eta_{original}} = \left[\frac{E_{fixed}}{E_{original}}\right]^2. \tag{4.14}$$

By neglecting both the thermal decay of the electronic grating and proton diffusion so that $n_{p1} = -N_{D1}^{+}$ after heating (the electronic charge grating is fully compensated by the protonic grating), the fixing efficiency can be rewritten as (combining Equation 4.11 and 4.13)

$$\eta_{fixing} = \frac{(\frac{N_D - N_A}{N_D} E_{0ph})^2 + E_d^2}{(\frac{N_D - N_A}{N_D} E_{0ph})^2 + (E_d + E_q)^2}. \tag{4.15}$$

After revealing, the net space–charge field E_{fixed} exists to prevent further erasure of the electronic charge grating by diffusion. The stronger the electron diffusion is, the stronger this field needs to be, and the larger the thermal fixing efficiency η_{fixing} becomes. This could be clearly seen in the above equation if we make the assumption that $N_A \approx N_D$ and reduce it to

$$\eta_{fixing} = \frac{1}{1 + E_q/E_d}. \tag{4.16}$$

As a result, larger thermal fixing efficiency is expected in the 90° geometry than in the transmission geometry because of stronger diffusion (larger K).

4.1.2 Numerical evaluation

Similar theoretical analysis and numerical evaluations have been shown in a 1990 paper by Carrascosa et al. [109]. However, although they did take the photovoltaic effect into account, the photovoltaic term is missing in their formula. Their sim-

Variable		Value	Reference
ϵ	Dielectric constant	2.6×10^{-12} F/cm	[119]
μ	Electron mobility	16 cm^2/v sec	[120]
n_o	Ordinary index of refraction	2.3489	[121]
Γ_{13}	Electro-optic coefficient	10×10^{-10} cm/v	[121]

Table 4.1: Material parameters of $LiNbO_3$.

ulation predicted fixing efficiencies as low as 0.003 in terms of the ratio of grating amplitudes. The reason is that the parameters they chose in the simulation was for a heavily-doped, strongly absorptive crystal. And also they only analyzed the influence of the doping level on the fixing process, neglecting that of the reduction/oxidation state which seems to us to be equally important in practice. Rather complete theory and many experiment results were reported in Reference [110, 111]. They showed a strong dependence of the fixing efficiency on the doping level as well as the reduction/oxidation state of the material. However, in most cases, their experiment conditions were different from what we have in a large-scale memory system. So, what do we expect from thermal fixing in practice?

In a practical large-scale holographic memory system, we typically prefer to use a medium-doped, deeply oxidized crystal to optimize the dynamic range of the system [40, 41]. For example, in our 160,000 hologram storage system (Chapter 2), the $LiNbO_3$:Fe crystal we used was 0.015%-doped. Its absorption at 488nm was 0.55cm^{-1}. In the numeric evaluation of our theoretical model, we used a set of parameters close to those found in Reference [40, 41] because of the good agreement between theory and experiment shown there. They are listed in Table 4.1 and 4.2.

We assigned these values in the equations for numerical evaluation of the thermal fixing efficiency in our experiment. The theoretical prediction (Equation 4.15) gives an expected fixing efficiency of 29%. In Equation 4.12, assuming that the proton concentration n_{p0} is 10^{19} cm^{-3} [122, 123], we have $|n_{p1}/N_{D1}^{+}| = 99.2\%$. This justifies the assumption $n_{p1} = -N_{D1}^{+}$ in our treatment. Another interesting point to notice here is that n_{p0} can be as low as $0.1N_A$ and the electronic grating is still more than 90%

Variable		Value
N_D	0.015 mol % Fe–doped	2.83×10^{18} cm^{-3}
N_A	$\alpha = 0.55\text{cm}^{-1}$	2.82×10^{18} cm^{-3}
K	90° geometry, $\lambda = 488$nm	4.27×10^{5} cm^{-1}
E_d	diffusion field	1.11×10^{4} V/cm
E_q	limiting space–charge field	$1.44\times10^{-13}\frac{N_A(N_D-N_A)}{N_D}$ V/cm
E_{0ph}	photovoltaic field	$1.09\times10^{-14}N_A$ V/cm

Table 4.2: Other parameters used in numerical evaluation.

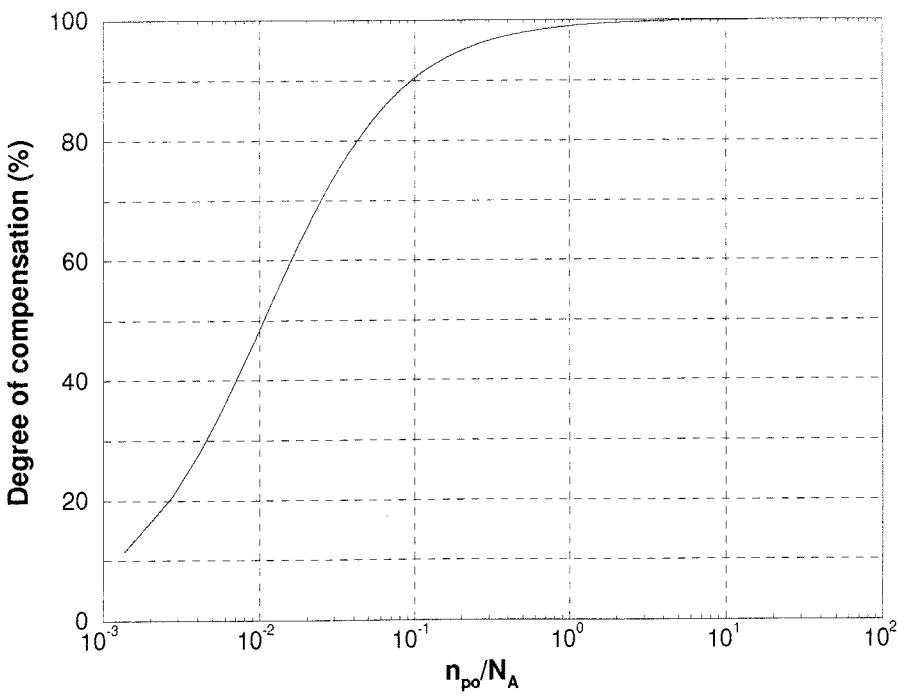

Figure 4.2: Degree of compensation as a function of proton density.

compensated—$|n_{p1}/N_{D1}^{+}| > 90\%$ (see Figure 4.2). Because the proton concentration n_{p0} affects the rate of dark decay of the protonic gratings, a trade–off can be reached for longer shelf–life of the fixed holograms by reducing the proton concentration while suffering only modest loss in the fixing efficiency.

4.1.3 Dynamics in hologram revealing

During revealing, a space–charge field emerges as the result of the partial erasure of the electronic grating by uniform illumination. This field modulates the index of refraction of the material via the electro–optic effect to form the thermally fixed phase hologram. The time evolution of this field can be described as

$$\tau \frac{\partial E_1}{\partial t} = -E_1 + E_{1\infty}, \tag{4.17}$$

where $E_{1\infty}$ is the saturation value, and τ is the time constant of hologram revealing.

$$E_{1\infty} = -j\frac{q}{K\epsilon}\frac{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + jE_d}{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + j(E_d + E_q)}n_{p1}, \tag{4.18}$$

$$\tau = \tau_{di}\frac{E_q}{E_{ue}}\frac{E_0 + j(E_{ue} + E_d)}{(E_0 + \frac{N_A}{N_D}E_{0ph}) + j(E_d + E_q)}, \tag{4.19}$$

where $E_{ue} = \gamma_R N_A / K\mu_e$ is the drift field of electrons. Evaluated after long exposure so that $E_0 = -E_{0ph}$, $E_{1\infty}$ becomes E_{fixed} as described by Equation 4.13.

The time evolution of the amplitude of the electronic grating N_{D1}^{+} can be derived using Equation 4.17 from the Poisson's relationship. Assuming that the electronic grating is fully compensated by the protonic one and the latter does not decay at room temperature, the dynamics of N_{D1}^{+} during revealing can be written as

$$N_{D1}^{+}(t) = -(1 - \Delta)n_{p1} - \Delta n_{p1}e^{-t/\tau}, \tag{4.20}$$

where Δ is

$$\Delta = \frac{(E_0 + \frac{N_A}{N_D} E_{0ph}) + jE_d}{(E_0 + \frac{N_A}{N_D} E_{0ph}) + j(E_d + E_q)}. \tag{4.21}$$

When Δ is evaluated at the end of the revealing where $E_0 = -E_{0ph}$, it can be related to the thermal fixing efficiency η_{fixing} by

$$\eta_{fixing} = |\Delta|^2. \tag{4.22}$$

We can rewrite Equation 4.20 as

$$N_{D1}^{+}(t) = N_{D1}^{+}(\infty) + [N_{D1}^{+}(0) - N_{D1}^{+}(\infty)]e^{-t/\tau},$$

or

$$N_{D1}^{+}(t) = N_{D1}^{+}(0)e^{-t/\tau} + N_{D1}^{+}(\infty)(1 - e^{-t/\tau}), \tag{4.23}$$

where

$$N_{D1}^{+}(0) = -n_{p1}$$

is the electronic grating strength at the beginning of revealing. $N_{D1}^{+}(\infty)$ is its saturation value, which can be written as

$$N_{D1}^{+}(\infty) = (1 - \Delta)\, n_{p1}.$$

We notice that the first part in Equation 4.23 is the same as the dynamics of the electronic grating during normal optical erasure without thermal fixing. An interesting conclusion can be drawn from this that although the time constant of hologram revealing is the same as that in normal hologram erasure, the rate of change (decay) in the amplitude of the electronic grating is less during revealing. This could be

explained by evaluating the time derivative of Equation 4.23 at (any) time t_0.

$$\frac{\partial N_{D1}^{+}}{\partial t}\Big|_{t=t_0} = -\frac{N_{D1}^{+}(0) - N_{D1}^{+}(\infty)}{\tau} e^{-t_0/\tau}. \tag{4.24}$$

This is clearly less than the rate of decay during normal erasure which is

$$\frac{\partial N_{D1}^{+}}{\partial t}\Big|_{t=t_0} = -\frac{N_{D1}^{+}(0)}{\tau} e^{-t_0/\tau}. \tag{4.25}$$

This is a very important conclusion that we would use in the development of a special exposure schedule to improve the overall efficiency of thermal fixing.

4.1.4 The M/# of thermally fixed memory

The M/# is a very important measure of the system dynamic range of a large–scale holographic memory. Complete theoretical treatment on the M/# and its dependence on the material as well as the system parameters could be found in Reference [40,41]. The oxidation/reduction state of the material is one of the important factors that influence the M/#. In general, if a $LiNbO_3$:Fe crystal is heavily reduced, there are more donor sites Fe^{2+} available. Therefore, the absorption is stronger, so is the sensitivity in holographic recording. However, reduction treatment reduces the density of the available acceptors Fe^{3+}, making it difficult to trap the electrons to create space–charge separation effectively. In addition, because of the strong absorption, the diffracted signal is attenuated much more inside the material, resulting in less observable diffraction efficiency. At the other end of the spectrum where the crystal is heavily oxidized, holographic recording is also ineffective because of the weak absorption. As a result, there is an optimal oxidation/reduction state (or absorption coefficient) for the highest possible M/# of a memory system. This is shown in Figure 4.3(a) [41].

With thermal fixing applied at the end of holographic recording, the final diffraction efficiency of the fixed holograms is the product of the original diffraction efficiency and the thermal fixing efficiency. As a result, the M/# of the thermally fixed memory

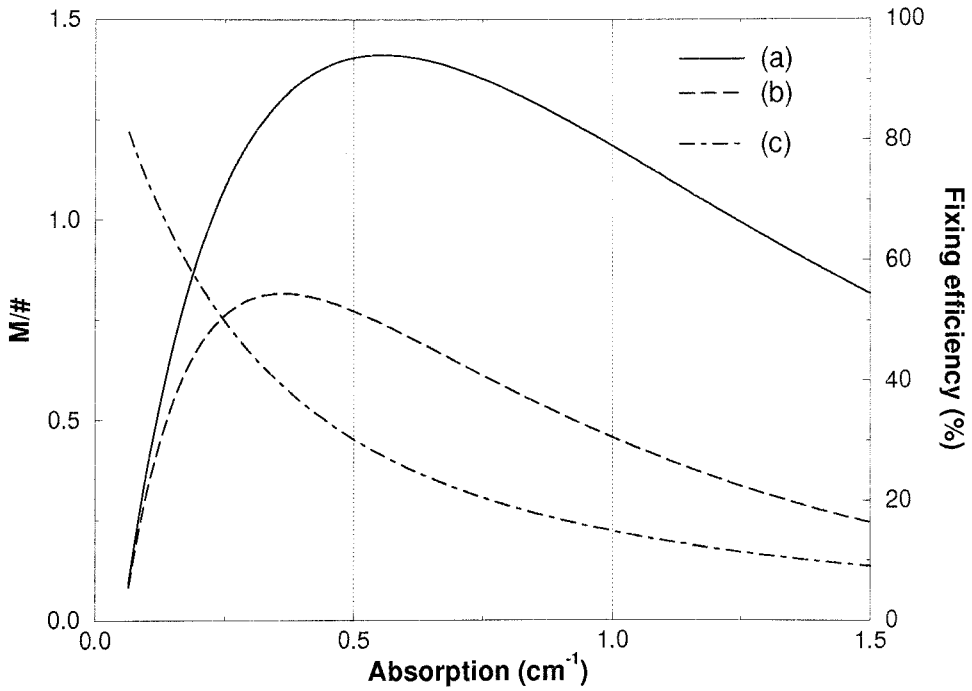

Figure 4.3: The M/# as a function of the absorption coefficient of the material: (a) the M/# of the original system; (b) the M/# of thermally fixed system; (c) thermal fixing efficiency.

system is the product of that of the original system without fixing and the square root of the fixing efficiency. We learn from Equation 4.15 that the fixing efficiency also depends on the reduction/oxidation state of the crystal. In other words, it is a function of the absorption coefficient as shown by curve (c) in Figure 4.3. It suggests that as the absorption becomes weaker, the fixing efficiency is higher. However, when the absorption is weaker, there are fewer donor sites Fe^{2+} available, so the M/# of the original system without fixing is smaller (curve (a)). As a result, an optimal absorption exits for the highest possible M/# after fixing, as shown in curve (b) in Figure 4.3. The optimal absorption is $0.37cm^{-1}$—not too far from what we have ($0.55cm^{-1}$) with the system. This means that there is little room for improvement in our experiment by changing the reduction/oxidation state of the crystal. The figure shows that the M/# of the system with thermal fixing applied at the end of the recording is about 0.7. Therefore, the expected diffraction efficiency shall be very close to the noise level when 10,000 holograms are recorded and thermally fixed.

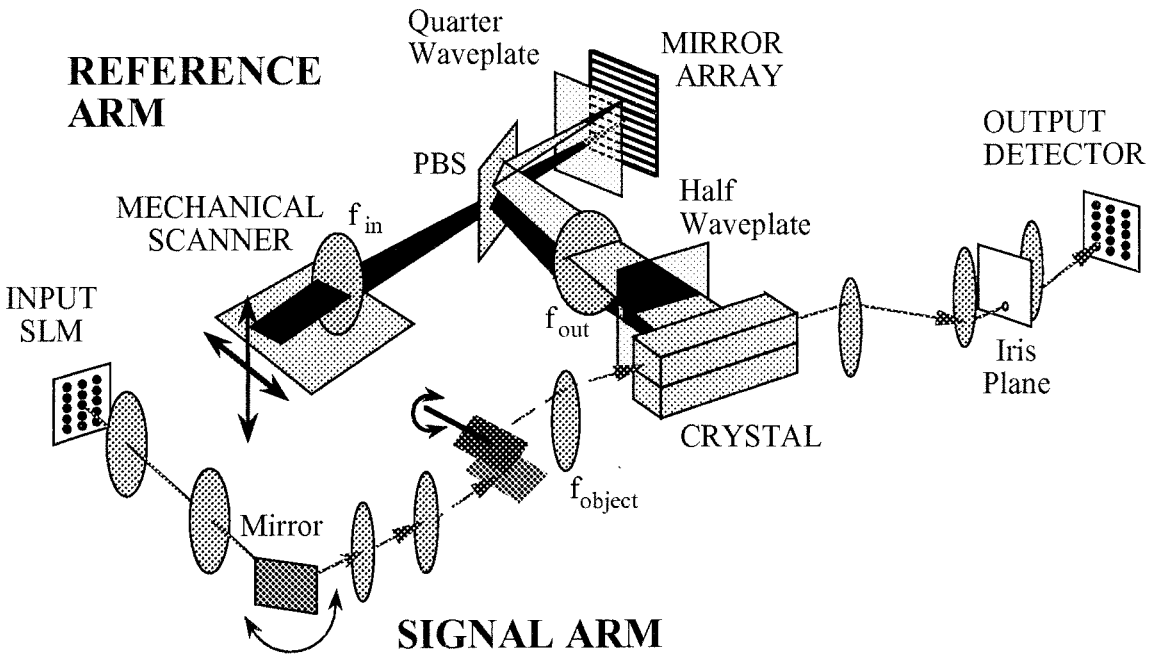

Figure 4.4: Experimental setup for thermal fixing.

4.2 Application of thermal fixing in large–scale memory

4.2.1 Experiment procedure

The fixing efficiency η_{fixing} is the portion in the strength of the original hologram attained after the revealing process. Numbers ranging from as low as $\eta_{fixing} \approx 10^{-5}$ [109] to as high as 70% [111] have been reported. Most of the reported data are from single hologram fixing experiments. We used our large scale holographic memory setup (Chapter 2) to carry out a systematic study of the thermal fixing efficiency and error performance that can be obtained when a large number of holograms are superimposed. The system diagram is shown in Figure 4.4. It is capable of storing 160,000 holograms using angle, fractal, and spatial multiplexing. In the fixing experiments, we stored up to 10,000 holograms at a single location.

The storage material was a 90°–geometry $LiNbO_3$ crystal with the **c**–axis at 45° to

the vertical faces. The crystal was 0.015% Fe–doped of dimension $2\times1.5\times4\text{cm}^3$. The crystal was placed 65mm beyond the focal plane of the Fourier–transform lens in the signal arm (focal length $f_{object} = 200\text{mm}$) for Fresnel–plane recording. The images to be stored were all binary random–bit patterns. Each pattern consisted of 25×80 bits that are represented by the 480×440 pixels on the input Spatial Light Modulator (SLM). The reconstructions were captured by a cooled scientific CCD camera and transferred to a computer for data analysis. The data analysis consists of registration of the 25×80 pixels and construction of the histograms of the OFF and ON pixels. Using these histograms as the estimates of the probability density functions of the binary signals, we used the same definition for the SNR as in Chapter 2.

4.2.2 SNR degradation with thermal fixing

We analyzed the SNR of the reconstructions for the storage of one hologram, and for the storage of 100, 300, 1,000, 5,000, and 10,000 original and thermally fixed holograms. In all cases the angular separation between neighboring holograms was about 5 times the angle to the first null of the selectivity curve of each hologram. Five fractal rows were used for the recording of 5,000 and 10,000 holograms, with 1,000 and 2,000 angular–multiplexed holograms on each row. An exposure schedule was followed to equalize the diffraction efficiency of the recorded holograms. The exposure schedules for storing different numbers of holograms were selected so that the equalized diffraction efficiency was the same in all cases ($\approx 7 \times 10^{-9}$). This is accomplished by recording, for example, the 100 holograms as if they were the last 100 holograms in the 10,000 hologram sequence. The results are summarized in Figure 4.5.

In Figure 4.5, the solid circles show the SNR of the equalized holograms before fixing. We attribute the reduction of SNR as the total number of holograms increases to the development of fanning, inter–pixel and other noise gratings over the long exposure sequence, and non–uniform erasure of the recorded holograms due to the bulk absorption in the crystal.

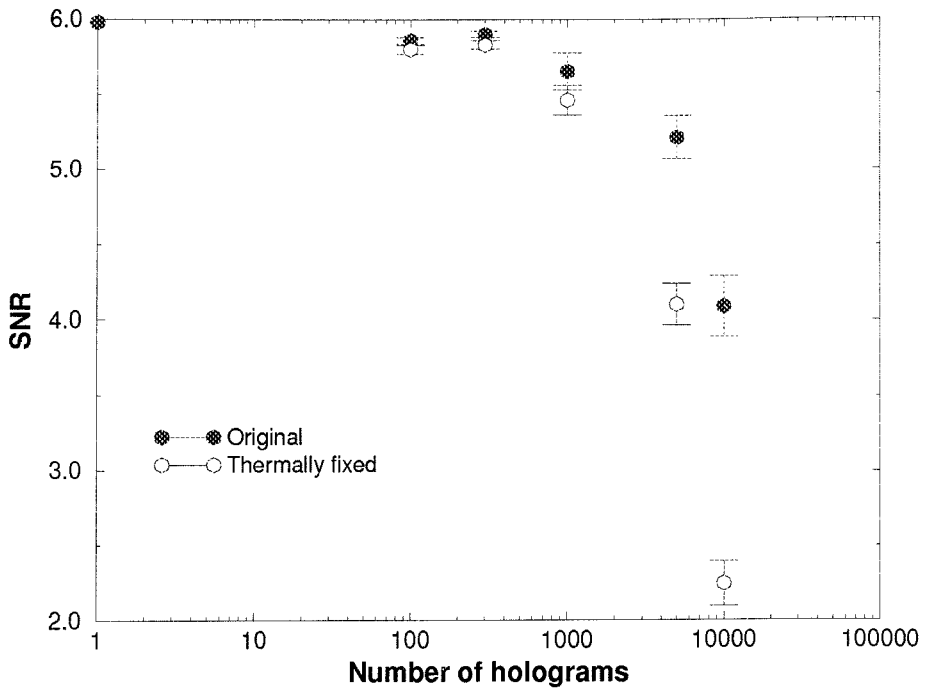

Figure 4.5: SNR degradation with thermal fixing.

Thermal fixing was carried out by heating the crystal after recording in an oven preheated to $120°C$. The crystal was allowed to cool off and then repositioned in the setup and illuminated with UV light to reveal the fixed holograms. We did not use the reference beam for the revealing because even with continuous scanning of the reference beam, a substantial increase in the background noise was observed. The average diffraction efficiency of the thermally fixed holograms measured from all the experiments was 1.8×10^{-9}, leading to an average thermal fixing efficiency of 26%. This is in good agreement with our theoretical prediction.

The SNR of the thermally fixed holograms is represented by the empty circles in Figure 4.5. It remains close to that of the original holograms at the beginning, indicating faithful copying by thermal fixing. However, as the total number of holograms becomes larger, the SNR of the fixed holograms is significantly lower. We can not attribute this drop to detector noise or scattering since the diffraction efficiency and the fixing efficiency were both independent of the number of holograms stored. Therefore, it must be attributed to a noise source that is "introduced" by the fixing process that becomes more prevalent when more (not weaker) holograms are stored. Detailed

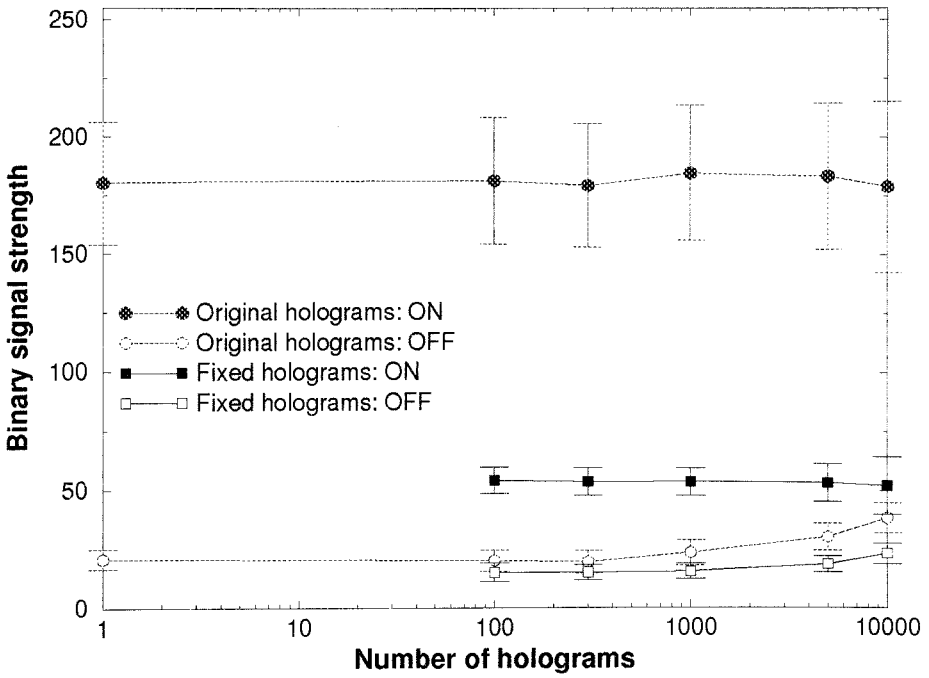

Figure 4.6: Means and variances of the ON and OFF signals as a function of the number of holograms.

examination of the experimental data (Figure 4.6) shows two effects: First, the mean and variance of the OFF pixels rises as the number of holograms increases whereas the mean of the ON pixels stays nearly constant. We attribute this to the holographic noise that arises from the buildup of noise gratings (e.g., inter–pixel gratings) during the prolonged exposures. Second, the drop in the mean and variance of the ON pixels due to fixing is higher than the drop of the OFF pixels. This seems to indicate that the unwanted noise gratings (mostly with large grating periods) experience higher fixing efficiency. This is surprising because theory predicts (Equation 4.15) that the fixing efficiency drops off at lower spatial frequencies. Independent measurements of the fixing efficiency of as a function of spatial frequency have confirmed this unexpected trend. We do not yet have a satisfactory explanation for this observation.

For 10,000 thermally fixed holograms the average SNR dropped from 5 to 3.5, and many errors were detected from the originally error–free reconstructions. Because of the dynamic–range limitations and other practical concerns, with 10,000 holograms written at the same storage location in the $LiNbO_3$ crystal, the average diffraction

efficiency was approximately 7×10^{-9}, about 10 times the background noise (bulk scattering, material surface defects, etc.) of a fresh crystal. Furthermore, the long exposure required to record multiple holograms increased the scattered noise level (fanning, inter–pixel gratings, etc.). At the end of the recording, the holograms were only 5—6 times stronger than the increased noise floor, corresponding to an SNR of approximately 5—the minimum for error–free reconstruction and low estimated bit–error rate (raw BER$\approx 10^{-6}$, comparable to that of a CD–ROM). With thermal fixing applied, the holograms suffered further loss in their diffraction efficiencies due to the low fixing efficiency. As shown in previous sections, the M/# is only about 0.7. The average diffraction efficiency practically achieved from the 10,000 fixed holograms was only 2×10^{-9}—too close to the noise level to resolve the binary patterns without error.

4.3 Incremental fixing schedule

With thermal fixing applied at the end of the recording to solve the volatility problem of holographic memories, the diffraction efficiency of the fixed holograms is smaller. Therefore, the limitation on the storage capacity imposed by the material dynamic range becomes more serious. Based on the discussion in Section 4.1.4, we believe that optimization of the oxidation/reduction state of the material will not lead to quantum improvement in the M/#. Therefore, the question becomes: Can we do something else to improve the fixing efficiency and hence the M/# of the system?

4.3.1 Dynamics and exposure schedule

In the context of a large–scale holographic memory, thermal fixing efficiency can be defined in two ways. The first is the fixing efficiency for an **individual** hologram as shown in Equation 4.15. The second one is the **system** thermal fixing efficiency. It is defined as the ratio of the final diffraction efficiency of multiple thermally–fixed holograms with t_1 as the recording time for the first hologram, to the diffraction efficiency of the same number of unfixed holograms starting with the same first exposure

time. Because of the dynamic nature of a memory system using $LiNbO_3$ crystals, the recording of multiple holograms involves both writing and erasure behaviors. In order to store all the holograms with equal strength, we need to use a carefully chosen set of decreasing exposure times as the exposure schedule. The difference between the fixing efficiencies for an individual hologram and a multiple–hologram system lies in the fact that the former is determined solely by the material and system parameters, while the latter also depends on the writing and erasure dynamics thus can be improved by optimizing the exposure schedule.

For detailed explanation, we first look at the evolution of the holographic grating under optical erasure in different cases. Although another charge grating, the protonic grating is involved when thermal fixing is applied, it always follows its electronic counterpart at the end of heating. Therefore, we are able to keep track of the evolution of the holograms by studying the dynamics of the electronic charge grating. For the erasure for an unfixed hologram, we have

$$h(t) = h_0 e^{-t/\tau_e}, \tag{4.26}$$

where $h(t)$ is the electronic grating strength as a function of time. This is simply an exponential decay with initial value h_0 and erasure time constant τ_e.

For the erasure of a hologram after heating where proton compensation takes place, although the migration of photo–electrons tends to smooth out the spatially–varying charge concentration as in the normal case, the electronic grating is still modulated by the existing protonic grating formed during heating. As a result, the electronic grating does not drop exponentially to zero. Instead, it experiences a slower decay with a non–zero steady–state value. Here, we rewrite Equation 4.19 in the form of grating strength, by replacing N_{D1}^{+} with h' and τ with τ_e [2]. In addition, we make the substitution of $\sqrt{\eta_{fixing}}$ for Δ [3]. The dynamics of the electronic grating strength is

[2] We make the assumption that the imaginary term in the complex time constant is negligible. Therefore, the "real" decay time constant is also τ_e. This is valid provided that the total erasure time is not very long compared to τ_e.

[3] This is made under the assumption that $N_A \approx N_D$. In other cases where this is not satisfied, there is a substantial (moving) phase shift between $N_{D1}^{+}(0)$ and $N_{D1}^{+}(\infty)$ and the dynamics becomes

then

$$h'(t) = (1 - \sqrt{\eta_{fixing}})h'_0 + \sqrt{\eta_{fixing}}h'_0 e^{-t/\tau_e}, \tag{4.27}$$

where $h'(t)$ is the electronic grating strength as a function of time. h'_0 is its initial value immediately after heating.

The recording and erasure behaviors in different cases are plotted in Figure 4.7. This figure also shows a simple example of how the exposure schedule is made. First, hologram #1 is recorded for t_1. Since the hologram would decay upon further exposure for hologram #2, the crystal could only be exposed for t_2 which is shorter than t_1 to make the strength of the two holograms equal. However, if the crystal is heated after the first hologram is written, the electronic grating strength decays much slower. As a result, we are able to record hologram #2 for t'_2 which is significantly longer, making the diffraction efficiencies of both holograms larger. If we keep doing this, recording one hologram and then heating the crystal, for multiple times, we can have stronger holograms at the end. However, because it takes a certain amount of time to heat and cool the crystal, this is not a practical procedure if we want to record and thermally fix many holograms. Instead, we break up the entire recording sequence of M holograms into M_s sets. Each set consists of equal number of holograms M_h=M/M_s. When the recording of a set of holograms is finished, the crystal is heated in an oven for a certain amount of time. After it cools down to room temperature, we record another set of M_h holograms. The procedure is repeated until all of the M holograms are recorded. Then as the final stage, we heat the crystal again and illuminate it with uniform light to reveal all the fixed holograms. We refer to this procedure as the "incremental fixing schedule." Intuitively, when M_s is smaller (the heating treatment is done less often), the normal optical erasure becomes more dominant, so the hologram strength at the end of the recording is weaker. On the other hand, with smaller M_s's, the hologram recording takes less time and effort. As a result, there is always a trade–off in practice.

more complicated.

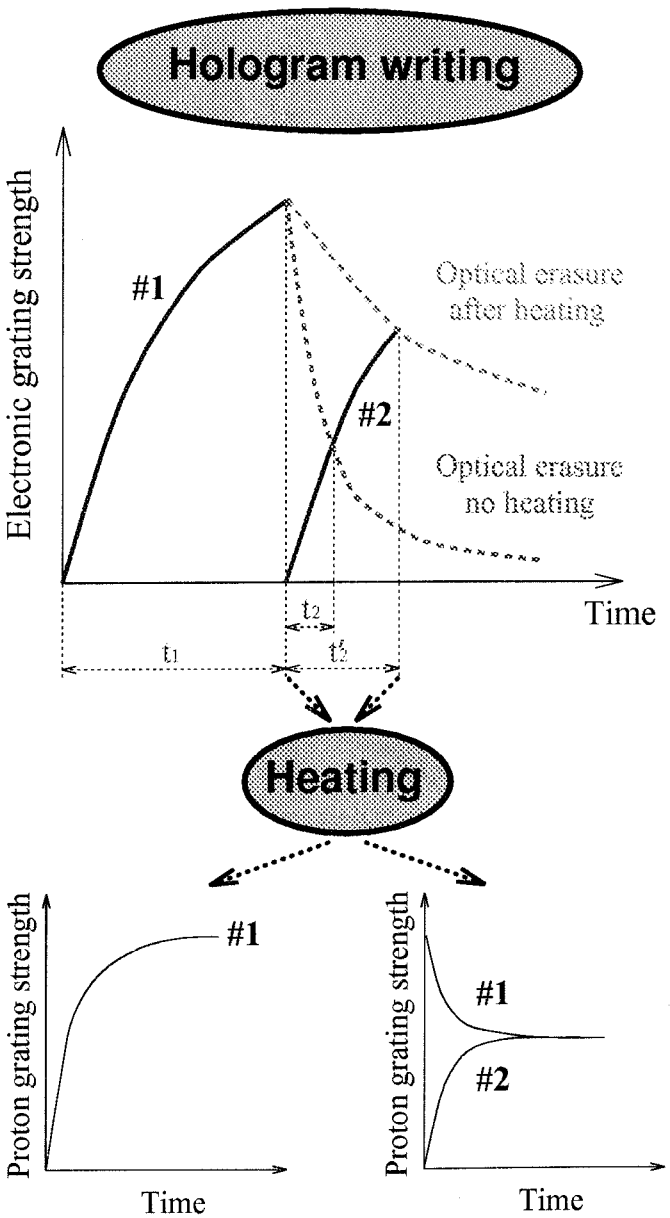

Figure 4.7: Holographic recording and erasure behavior under different conditions.

4.3.2 Development of incremental fixing schedule

In an exposure schedule for normal holographic recording, equal diffraction efficiency of all the holograms is ensured by making the diffraction efficiencies of the two neighboring holograms in the recording sequence, the m^{th} and the $(m+1)^{th}$, the same. In other words, the $(m+1)^{th}$ hologram is recorded for t_{m+1} so that it is equal to the m^{th} hologram which is recorded for t_m and erased for t_{m+1}.

$$h_0(1-e^{-t_m/\tau_e})e^{-t_{m+1}/\tau_e} = h_0(1-e^{-t_{m+1}/\tau_e}), \tag{4.28}$$

where t_m and t_{m+1} are the exposure times for the m^{th} and $(m+1)^{th}$ holograms, respectively. Typically, the exposure schedule is generated using a backward "recursive" algorithm, given the total number of holograms to be stored M, the erasure time constant τ_e, and the exposure time for the last hologram t_M. Alternatively, an exposure schedule can be calculated using a forward method starting with t_1 as the exposure time for the first hologram. The final diffraction efficiency of the system is then the efficiency of the decayed first hologram due to the exposures needed to record the rest (M-1) of the holograms with equal strength. In a purely theoretical treatment where t_1 can be infinitely long so that the first hologram is written to saturation, the diffraction efficiency approaches the upper limit imposed by the system dynamic range as described by the M/#. For the practical implementation of a large–scale memory which favors a much shorter t_1 to avoid the buildup of holographic noise, we choose $t_1 = \tau_e/R_1$ ($R_1 > 1$) and use it in Equation 4.28. The exposure time for the m^{th} hologram is then

$$t_m = \frac{\tau_e}{R_1 + m - 1}, \tag{4.29}$$

where $m = 1, 2, ..., M$.

To design the incremental fixing schedule, the only modification of Equation 4.28 is that instead of treating single holograms at a time, we should deal with sets of holograms, that is, to make the decreased diffraction efficiency of the n^{th} set of holograms

due to optical erasure after heating match that of the $(n+1)^{th}$ set of holograms. During the recording of the holograms in a set, the optical erasure is the same as in the normal case. Therefore, a normal exposure schedule is applied to the storage of each set of holograms. Given the number of holograms in a set as M_h and the erasure time constant τ_e, only the recording time for the first hologram in the set, t_1^n, is needed to specify the exposure schedule for the entire n^{th} set. The values of t_1^n $(n = 1, 2, ..., M_s)$ are calculated using the expression below:

$$\begin{aligned}(1-\sqrt{\eta_{\text{fixing}}})h_0^n + \sqrt{\eta_{\text{fixing}}}h_0^n exp\left[-\sum_{i=1}^{M_h} t_i^{n+1}/\tau_e\right] &= h_0^{n+1},\\ h_0^n &= h_\infty(1-e^{-t_{M_h}^n/\tau_e}),\\ h_0^{n+1} &= h_\infty(1-e^{-t_{M_h}^{n+1}/\tau_e}),\end{aligned} \tag{4.30}$$

where h_0^n and h_0^{n+1} are the strengths of the n^{th} and $(n+1)^{th}$ sets of holograms after recording, respectively. Correspondingly, $t_{M_h}^n$ and $t_{M_h}^{n+1}$ are the last exposure times for each set. h_∞ is the saturation grating strength. The left side of Equation 4.30 accounts for the decay of the n^{th} set of holograms during the recording of the $(n+1)^{th}$ set, which takes $\sum_{i=1}^{M_h} t_i^{n+1}$. Each t_i^n is simply derived from the normal exposure schedule starting with t_1^n (Equation 4.29).

Combining Equation 4.29 and 4.30, the first exposure time t_1^n for the n^{th} set is

$$t_1^n = \frac{\tau_e}{R_1 + \sqrt{\eta_{\text{fixing}}}(n-1)M_h}, \tag{4.31}$$

where $n = 1, 2, ..., M_s$. $R_1 = \tau_e/t_1$ corresponds to the choice for the first exposure time for the entire incremental fixing schedule. We note here that the exposure time of the first hologram in the n^{th} set, t_1^n, corresponds to the $[(n-1)M_h + 1]^{th}$ hologram in the entire exposure sequence. In the normal exposure schedule (Equation 4.29), the exposure time for this hologram would be

$$\frac{\tau_e}{R_1 + (n-1)M_h}.$$

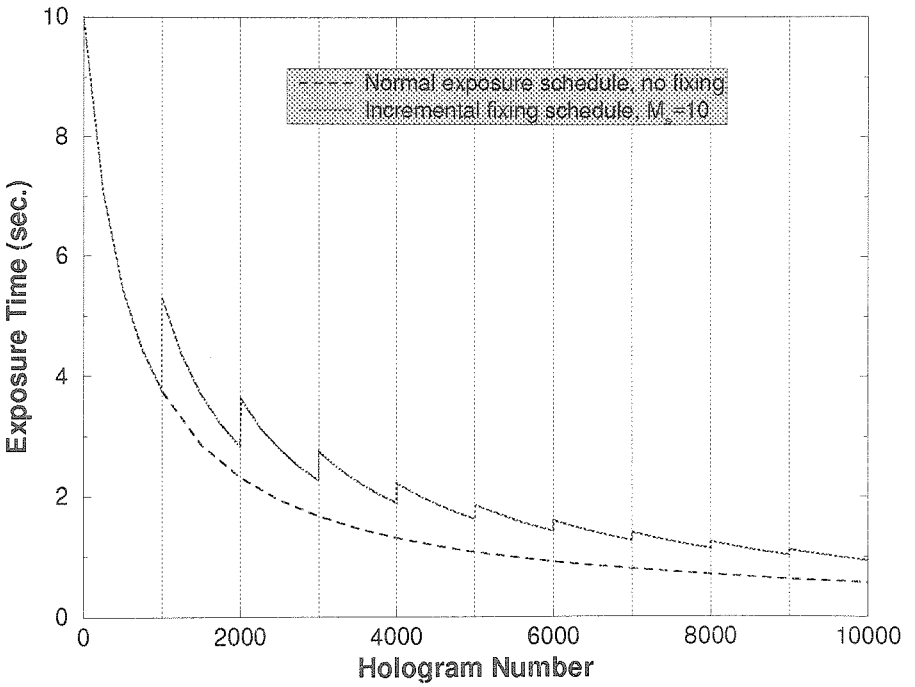

Figure 4.8: An example of the incremental fixing schedule.

Comparing the above expression with Equation 4.31, the exposure time for this hologram ($\#[(n-1)M_h+1]$) is increased with the incremental fixing schedule because $\eta_{\text{fixing}} < 1$. This is shown in Figure 4.8. The exposure schedules are adopted for the storage of 10,000 holograms ($M = 10,000$). The number of heating treatments M_s is 10, so the number of holograms in each set M_h is 1,000. The erasure time constant is taken as 6,000 seconds, and the first exposure time $t_1 = 10$ seconds ($R_1 = 600$). Because the crystal is heated after the recording of each set of 1,000 holograms, the exposure times at hologram #1,001, 2,001, ..., 9,001 are increased, resulting in overall stronger hologram strength at the end of the recording.

4.3.3 The M/# with incremental fixing schedule

After the entire recording sequence is finished, we need to heat the crystal once again and reveal the fixed holograms using uniform illumination. The diffraction efficiency of the fixed holograms is the product of the efficiency obtained at the end of the recording using the incremental fixing schedule, and the thermal fixing efficiency for

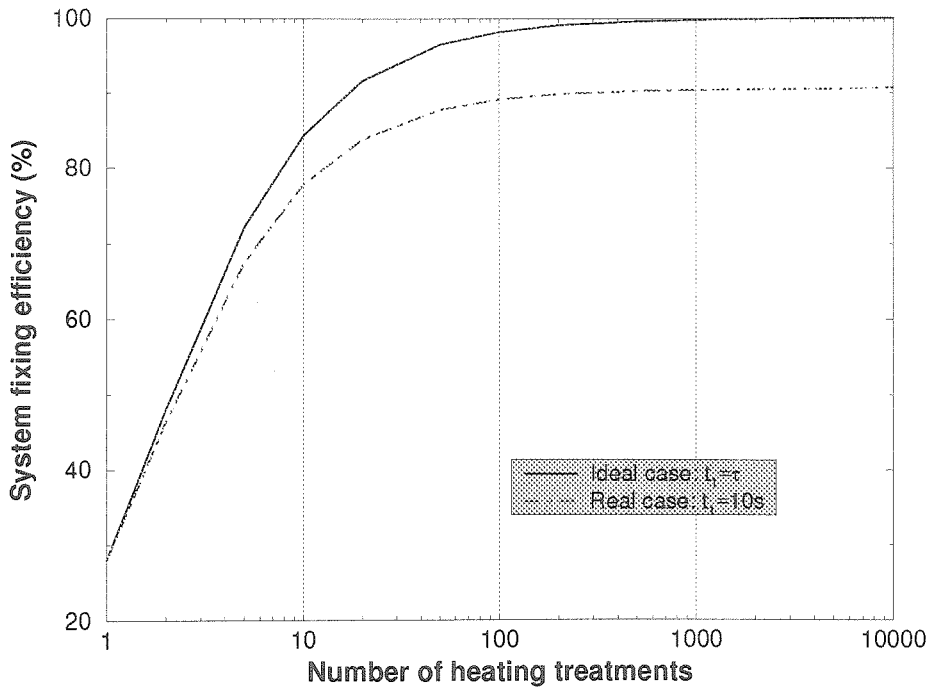

Figure 4.9: System fixing efficiency as a function of the number of heating treatments.

an individual hologram (Equation 4.15). The ratio of this diffraction efficiency and that of the M unfixed holograms is the **system** fixing efficiency. It is plotted as the function of the number of heating treatments (the number of sets, M_s) in Figure 4.9. The total number of holograms to be stored M is 10,000, and the erasure time constant τ_e is taken as 6,000 seconds. Curve (a) is for the practical case in which the recording is started with $t_1 = 10$ seconds. It shows that as M_s increases (the crystal is heated more often), the system fixing efficiency becomes higher. Curve (b) shows an ideal case in which t_1 can be infinitely long. It predicts that as M_s becomes closer to M, the system fixing efficiency approaches 100%. In other words, at the limit where $M_s = M$ (heating is applied after the recording of every hologram), the gain in the exposure schedule from incremental heating treatment compensates all the loss due to thermal fixing. The maximal M/# of the memory system with incremental fixing schedule is the same as that of the original, volatile one.

4.3.4 Thermal fixing of 10,000 holograms using incremental fixing schedule

In the same setup shown in Figure 4.5, 10,000 holograms were recorded and thermally fixed using an incremental fixing schedule. In the reference arm, the segmented mirror array was replaced by a mirror since no spatial multiplexing was required for the storage of 10,000 holograms. The incremental fixing procedure requires frequent removal and repositioning of the crystal. However, since no diffraction can be observed right after the crystal is heated, there is no reliable reference for accurate repositioning. To increase the tolerance to repositioning error, the cylindrical lenses in the original design were removed to improve the plane–wave quality of the reference beam. This change was proved to be helpful in guaranteeing reliable repositioning. Five fractal rows were used for storage with 2,000 holograms stored on each. The vertical spacing between fractal rows was 2.4°, and the horizontal spacing between neighboring holograms was 0.007°, approximately 4 times the measured Bragg selectivity. In order to avoid overlapping which might occur due to imperfect repositioning, the angle spacing between the two adjacent sets of holograms on the same fractal row was set to be 10 times wider than the inter–hologram spacing.

In the incremental fixing schedule, 10,000 holograms were divided into 10 sets ($M_s = 10$), two on each fractal row. The crystal was heated in an oven preheated to $120°C$ for 30 minutes after each set of $M_h = 1,000$ holograms were recorded. The erasure time constant τ_e was empirically determined to be 6,000 seconds. The first exposure time t_1 was chosen to be 10 seconds. The exposure time for the $10,000^{th}$ hologram ($t_{10,000}$) was calculated from Equation 4.31 and 4.29 to be 0.94 second. It was nearly twice as long as the final exposure time using the normal exposure schedule without thermal fixing. However, because of the loss in revealing ($1 - \eta_{fixing} \approx 70\%$), the final, **system** fixing efficiency is expected to be 78% (see Figure 4.9). This is more than twice the **individual** fixing efficiency, without the incremental schedule.

After 10,000 holograms were recorded and thermally fixed, 30 sample reconstructions were retrieved by a cooled scientific CCD camera. They were well distributed

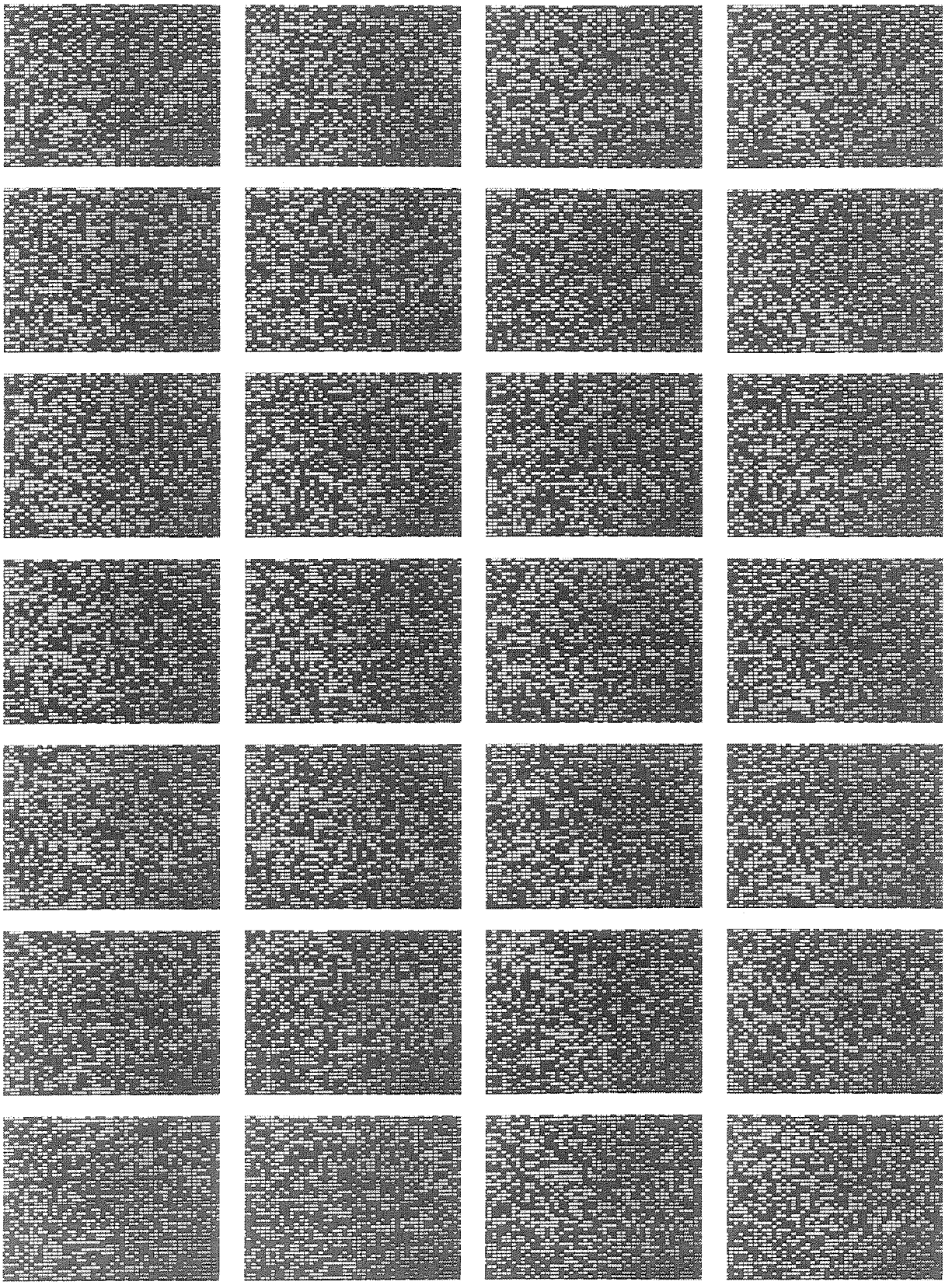

Figure 4.10: Sample reconstructions from the 10,000 fixed holograms using incremental fixing schedule.

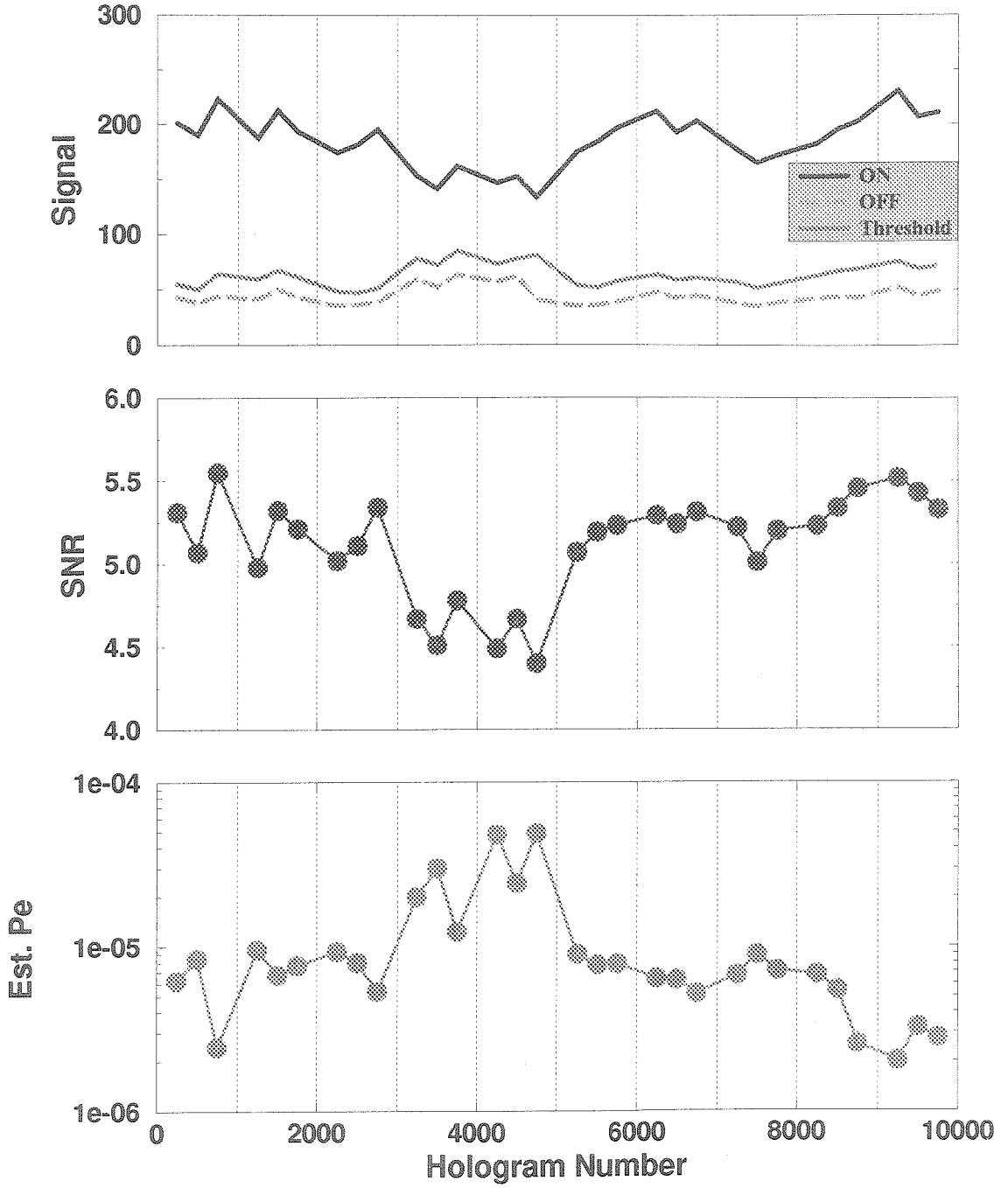

Figure 4.11: Characteristics of the sample reconstructions from the 10,000 fixed holograms.

among the 10,000 holograms. Some are shown in Figure 4.10. The same data analysis procedure as described in the previous sections was used to analyze the error performance. The average diffraction efficiency is approximately 4.6×10^{-9}. Compared to that of the 10,000 holograms without fixing ($\eta \approx 7 \times 10^{-9}$), the incremental fixing schedule yields a system fixing efficiency as roughly 66%, close to the estimated value from Figure 4.9. All the sample reconstructions were retrieved without any measured errors. The characteristics of these reconstructions are plotted in Figure 4.11, including the means of the ON and OFF pixel regions, the optimal threshold, the SNR, and estimated probability of error of the reconstructed holograms. It is very interesting to notice that, even with the lower diffraction efficiency, the error performance of the 10,000 fixed holograms is comparable to or even slightly better than that of the unfixed holograms (Figure 4.5). This is not surprising because the repeated heating reduces the electric field inside the crystal, resulting in less scattered noise. Similar effect has been observed by other researchers [112].

4.4 Miscellaneous topics

4.4.1 Multiple recording and fixing treatments

Thermal fixing solves the volatility problem of holographic memories by translating the stored information from its electronic copy to protonic copy. Thermally fixed holograms are basically Read–Only Memory (ROM). Here, we would like to address the question of whether holographic memories with thermal fixing applied are re–writable? In other words, can a new set of holograms be added in the memory with no or minimal interference with the old memory content? The key issue is how much of the original fixed holograms is retained after the recording and fixing of new holograms.

There is a slight difference between the procedures of rewriting holograms in a thermally fixed memory and the incremental fixing schedule we described previously. However, it causes a great difference in the final results. In the incremental fixing

schedule, we did not reveal the holograms before the recording of another set of holograms. Instead, we took advantage of the slower decay rate of the electronic gratings of the preciously stored set of holograms after heating, during the recording of the new set to achieve higher diffraction efficiency at the end.

In a memory storing thermally fixed holograms, the electronic and protonic copies of the holograms already reach their steady–state values after heating and revealing. The thermally fixed holograms are the result of the difference of these two copies. The more the electronic gratings are erased by uniform illumination in revealing, the larger this difference is, and in turn, the higher the diffraction efficiency of the fixed holograms becomes. When a new set of holograms are added to the memory content, the crystal must be heated again for thermal fixing. At high temperatures, there exist two processes: First, the protons migrate to form the copy of the newly recorded holograms. In the meantime, the protonic copy of the previously stored holograms decays to match their electronic counterparts that are partially erased in revealing. As a result, the previously stored holograms are erased "thermally" after the addition of new holograms.

The electronic grating remained after revealing, $N_{D1}^{+}(\infty)$ is

$$\begin{aligned} N_{D1}^{+}(\infty) &= (1-\Delta)N_{D1}^{+}(0) \\ &= \frac{jE_q}{\frac{N_D-N_A}{N_D}E_{0ph}+j(E_d+E_q)}N_{D1}^{+}(0), \end{aligned} \tag{4.32}$$

where $N_{D1}^{+}(0)$ is the original value after recording. After recording and thermally fixing a set of new holograms, the protonic copy of the old holograms becomes

$$n'_{p1} = -N_{D1}^{+}(\infty). \tag{4.33}$$

Therefore, the previously stored holograms is weakened by a factor of $|1-\Delta|^2$. When the memory is updated by x times, the first recorded and thermally fixed set of holograms are weakened by a factor of $|1-\Delta|^{2x}$. Using the same set of parameters to evaluate this decay, we obtained a value of 82% with a thermal fixing efficiency at

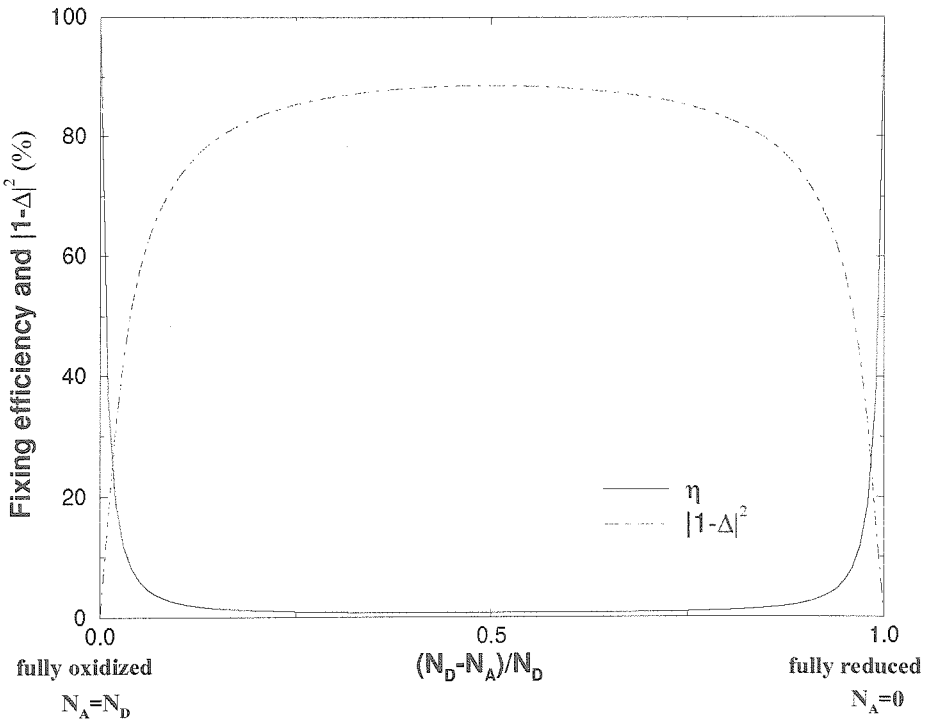

Figure 4.12: η_{fixing} and $|1-\Delta|^2$ as a function of the oxidation/reduction state of the material.

29%. After rewriting and fixing 10 times, only about 2% of the first set of holograms remains.

We notice that the thermal fixing efficiency η_{fixing} is Δ^2 which is evaluated at $E_0 = -E_{0ph}$ (after long revealing), while the diffraction efficiency of the fixed holograms remained after updating the memory is characterized by the term $|1-\Delta|^2$. It is clear that these two terms go against each other, as shown in Figure 4.12. In other words, when the fixing efficiency is increased, $|1-\Delta|^2$ is always decreased, and vice versa. The thermal fixing efficiency and $|1-\Delta|^2$ can not be optimized at the same time by adjusting the oxidation/reduction state of the material.

To equalize the strengths of the fixed holograms after multiple updating, a special recording and fixing schedule similar to the normal exposure schedule can be applied. However, the loss in the diffraction efficiency still exists and results in a much smaller effective M/# than that of either the original system or with the incremental fixing schedule.

In conclusion, a thermally fixed memory system **can not** be effectively updated by multiple recording and fixing treatments. In other words, non–volatile holographic storage systems with thermal fixing are Write–Once–Read–Many (WORM) memories, not re–writable ones.

4.4.2 Repeated recording and fixing treatment

For the storage of a few holograms where the final diffraction efficiency is the key issue, repeated recording and fixing treatment can be applied. During holographic recording in photorefractive materials at room temperature, an upper limit is imposed on the final hologram strength by the interaction between different charge transport mechanisms (different characteristic fields). Electric fields developed during the recording eventually inhibit further hologram growth before the available donors or acceptors are depleted. To the contrary, if the crystal is heated after recording, the net electric field is almost zero as a result of proton compensation, leaving room for further growth of the electronic charge grating when another round of recording is carried out. Therefore, a hologram can be recorded to saturation, heated, and then recorded once again to saturation, followed by another heating treatment. This procedure can be repeated by many times until the final limit—the donor or acceptor density ($(N_D - N_A)$ or N_A) is reached.

When a modulated exposing pattern is used instead of the uniform illumination for revealing, the resulting space–charge field is simply a linear summation of the results of revealing and normal holographic recording.

$$E_{1\infty} = -jE_q \frac{\left[(E_0 + \frac{N_A}{N_D}E_{0ph}) + jE_d\right] Rn_{p1} + m\left[(E_0 + E_{0ph}) + jE_d\right]}{(E_0 + \frac{N_A}{N_D}E_{0ph}) + j(E_d + E_q)}, \tag{4.34}$$

where n_{p1} represents the protonic grating formed during heating. m is the modulation depth of the new recording. $R = N_D/N_A(N_D - N_A)$, is a measure of the oxidation/reduction state of the crystal. There are two terms in the numerator: the first is the result of revealing by uniform illumination (refer to Equation 4.18), the sec-

ond is from normal holographic recording (refer to Equation 4.7). When the crystal is oxidized, $N_A \approx N_D$. Therefore, these two terms are almost in phase, and the electric field is simply increased by an amount corresponding to the result from revealing.

If we use Equation 4.32 and rewrite the above equation in terms of the electronic grating, it becomes

$$N_{D1}^{+\prime} = \frac{E_q}{E_d + E_q}\left[N_{D1}^{+} + \frac{N_A(N_D - N_A)}{N_D}m\right], \tag{4.35}$$

where $N_{D1}^{+}, N_{D1}^{+\prime}$ are the amplitudes of the electronic grating before and after heating and hologram rewriting. This expression for the saturation values is reached by assuming $N_A \approx N_D$. As this heating–plus–rewriting process is carried out many times, the electronic grating can be enhanced to the upper limit where depletion of either the donors or the acceptors occurs. Beyond this limit, the second term which is responsible for rewriting is zero, there is no further increase in the electronic grating strength.

4.4.3 Selective fixing procedures

A special application of the repeated recording and fixing treatment is selective revealing of the fixed holograms. As shown in Equation 4.35, when there is a 180° phase shift between the exposing light patterns, the electronic grating can be completely erased as the result of "negative" recording. Therefore, the protonic grating formed during heating can be fully revealed and 100% fixing efficiency can be obtained.

In addition, although the complex time constant is always the same during revealing, the rate of change or the characteristic time constant can be greatly improved by selective revealing. This is similar to the fast update of dynamic photorefractive optical memory reported in [96].

4.5 Appendix: Derivation of the theoretical model

In Section 2.8, we described the development of the space–charge field in room-temperature recording. Here we derive its evolution after thermal fixing is applied. The basic set of Kukhtarev equations have to be expanded to incorporate another mobile species, the protons, to explain the thermal fixing process. The continuity equation and current density of the protons are

$$\frac{\partial n_p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x}, \tag{4.36}$$

$$J_p = q\mu_p n_p E - K_B T \mu_p \frac{\partial n_p}{\partial x}. \tag{4.37}$$

Recording: In room–temperature recording, the space–charge field builds up in the same way as described in Section 2.8. At any moment in the evolution, it is related to the amplitude of the electronic charge grating N_{D1}^{+} by the Poisson's relationship

$$E_1 = -j\frac{q}{K\epsilon}N_{D1}^{+}. \tag{4.38}$$

Heating: During heating, if we neglect the thermal excitation of the electrons, the electronic grating remains the same—N_{D1}^{+} is constant. The only "free" mobile species are the protons. They migrate under the space–charge field of the electronic charge grating. As the protonic grating builds up to compensate its electronic counterpart, the net space–charge field decreases. An equilibrium is established when the diffusion of the protons equals their drift by the space–charge field. If we ignore the diffusion of the protons [4], then from the continuity equation, the net field at the end is zero. Therefore, the protonic grating fully compensate its electronic counterpart, as indicated in the Poisson's relationship.

If the proton diffusion is included, the development of the protonic grating can be

[4]The diffusivity or the mobility of the protons is indeed very small. It is the sheer number of "free" protons that makes the protonic conductivity much larger than the electronic conductivity (thermal or photo–conductivity).

derived from the first-order current density J_{p1}

$$J_{p1} = q\mu_p n_{p0} E_1 + q\mu_p n_{p1} E_0 - jK_B T K \mu_p n_{p1}, \tag{4.39}$$

where E_0, E_1 are the DC component and spatial perturbation of the space-charge field. The evolution of E_0 as a result of the change in the surface charge density is described by

$$\frac{\partial \rho}{\partial t} = -J_{p0} = -q\mu_p n_{p0} E_0. \tag{4.40}$$

The solution to this equation is

$$E_0 = -E_{0ph} e^{-t/\tau_{di}^p}, \tag{4.41}$$

where τ_{di}^p is the dielectric time of the protons,

$$\tau_{di}^p = \frac{\epsilon}{\sigma_p} = \frac{\epsilon}{q\mu_p n_{p0}}. \tag{4.42}$$

At equilibrium, both E_0 and J_{p1} are zero. From Equation 4.39, the amplitude of the protonic grating is related to the spatial perturbation in the space-charge field by

$$n_{p1} = -jn_{p0}\frac{E_1}{E_d}. \tag{4.43}$$

The first-order Poisson's relationship is

$$E_1 = -j\frac{q}{\epsilon K}(N_{D1}^+ + n_{p1}), \tag{4.44}$$

where N_{D1}^+ is the electronic copy of the information recorded at room temperature. Combining Equation 4.43 and 4.44, the steady-state value of the protonic grating is

$$n_{p1} = -\left(\frac{1}{1 + E_d/E_{qp}}\right) N_{D1}^+, \tag{4.45}$$

where $E_{qp} = qn_{p0}/K\epsilon$ is the space–charge limited field of the protons. The net space–charge field at the end of heating becomes

$$E_1 = -j\frac{E_d E_{qp}}{E_d + E_{qp}}\frac{N_{D1}^+}{n_{p0}}. \tag{4.46}$$

The complete solution describing the time evolution of the protonic grating during heating is

$$\tau_p \frac{\partial n_{p1}}{\partial t} = -n_{p1} + n_{p1}(\infty), \tag{4.47}$$

where

$$\tau_p = \frac{1}{1 + \frac{E_d + jE_0}{E_{qp}}}\tau_{di}^p, \tag{4.48}$$

$$n_{p1}(\infty) = -\frac{1}{1 + \frac{E_d + jE_0}{E_{qp}}}N_{D1}^+. \tag{4.49}$$

Evaluated after enough heating such that $E_0 = 0$, $n_{p1}(\infty)$ becomes the same as in Equation 4.45.

Revealing: After heating, the crystal is illuminated by uniform light to reveal the holograms. The conductivity of the protons at room temperature is much less than the photoconductivity of the electrons, so n_{p1} can be treated as constant. The electronic charge grating is partially erased during revealing, leaving a net space–charge field which modulates the refraction index of the material to form the thermally fixed holograms.

During revealing, the protonic grating formed at elevated temperatures remains intact, and the DC field increases again to screen the photovoltaic field. The first–

order equations that govern the erasure of the electronic grating become

$$\frac{\partial N_{D1}^{+}}{\partial t} = -SI_0 N_{D1}^{+} - \gamma_R n_{e0} N_{D1}^{+} - \gamma_R n_{e1} N_{D0}^{+}, \tag{4.50}$$

$$\frac{\partial n_{e1}}{\partial t} = \frac{\partial N_{D1}^{+}}{\partial t} + j\frac{K}{q} J_{e1}, \tag{4.51}$$

$$J_{e1} = q\mu_e n_{e1} E_0 + q\mu_e n_{e0} E_1 + jK_B T K \mu_e n_{e1} - PI_0 N_{D1}^{+}, \tag{4.52}$$

$$E_1 = -j\frac{q}{K\epsilon}(N_{D1}^{+} - n_{p1}). \tag{4.53}$$

Under the same assumptions we made in Section 2.8, we have

$$N_{D0}^{+} = N_A, \tag{4.54}$$

$$n_{e0} = \frac{SI_0(N_D - N_A)}{\gamma_R N_A}, \tag{4.55}$$

$$\frac{\partial n_{e1}}{\partial t} \approx 0. \tag{4.56}$$

Using Equation 4.50— 4.52 and 4.54— 4.56, the quasi–steady–state value of the first–order harmonic in the free electron concentration [5] is found to be

$$n_{e1} = -j\frac{K\epsilon}{q\gamma_R N_A}\frac{\partial E_1}{\partial t} - j\frac{K\epsilon SI_0 N_D}{q\gamma_R N_A^2} E_1 + \frac{SI_0 N_D}{\gamma_R N_A^2} n_{p1}. \tag{4.57}$$

From Equation 4.51 and 4.53, we also have

$$\frac{\partial E_1}{\partial t} = \frac{1}{\epsilon} J_{e1}. \tag{4.58}$$

Combining this equation with Equation 4.52, the dynamics of the space–charge field is found to be

$$\tau\frac{\partial E_1}{\partial t} = -E_1 + E_{1\infty}, \tag{4.59}$$

[5] We made yet another assumption $\gamma_R N_A \gg SI_0 \frac{N_D}{N_A}$, which is valid in most cases.

where

$$E_{1\infty} = -j\frac{q}{K\epsilon}\frac{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + jE_d}{\left(E_0 + \frac{N_A}{N_D}E_{0ph}\right) + j(E_d + E_q)}n_{p1}, \tag{4.60}$$

$$\tau = \tau_{di}\frac{E_q}{E_u}\frac{E_0 + j(E_u + E_d)}{(E_0 + \frac{N_A}{N_D}E_{0ph}) + j(E_d + E_q)}. \tag{4.61}$$

Since the DC field changes with time, the time constant and saturation value in this equation change with time too. So at any instant, the solution to Equation 4.58 is an exponential function; however, the overall time evolution takes a more complex form. If the revealing is long enough for the DC field to fully screen the photovoltaic field, the final strength of the fixed hologram is

$$E_{fixed} = -j\frac{q}{K\epsilon}\frac{\frac{N_D - N_A}{N_D}E_{0ph} + jE_d}{\frac{N_D - N_A}{N_D}E_{0ph} + j(E_d + E_q)}n_{p1}. \tag{4.62}$$

Chapter 5 High–temperature recording in $LiNbO_3$:Fe

Contents

5.1 Theoretical model 118

5.1.1 Extended Kukhtarev equations 118

5.1.2 Steady–state solutions 121

5.1.3 Numeric evaluation of the model—the dynamics 123

5.1.4 The M/#: writing/erasure asymmetry 127

5.1.5 Simplified model 132

5.1.6 Predictions of the M/# 134

5.2 Experiment results 140

5.2.1 Writing/erasure dynamics 143

5.2.2 Experimental procedure for the M/# measurement 145

5.2.3 Experiment results and comparison to theory 147

5.2.4 Multiple holograms storage 148

5.3 Miscellaneous topics 150

5.3.1 Thermal contraction and change of refraction index 150

5.3.2 Bragg shift and image distortion 152

5.4 Appendix A: Derivation of the extended Kukhtarev equations 156

5.5 Appendix B: Derivation of the simplified model 159

In this chapter, we discuss another method of thermally fixing stored holograms to solve the volatility problem of holographic memories using photorefractive materials. As we mentioned in previous discussions, thermal fixing can be carried out in

one of two ways. The first is the one we discussed in detail in Chapter 4. Because there are three steps involved in the recording and thermal fixing of holograms in a photorefractive crystal—room–temperature recording, high–temperature fixing, and room–temperature revealing and readout, the procedure is also called "Low–High–Low" fixing. The second one is high–temperature recording. Namely, holograms are directly recorded in the storage medium at elevated temperatures, followed by the same cooling and developing procedure to reveal them at room temperature. This is also called "High–Low" fixing. In high–temperature recording where the proton conductivity dominates, protonic gratings compensate ("screen") the electronic charge gratings rapidly and continuously, resulting in a small net space–charge field. As we recall from normal holographic recording at room temperature, it is the space–charge field associated with the electronic charge separation that finally stops further growth of the holograms. Therefore, conceivably, much stronger holograms could be recorded by high–temperature recording, limited only by either the density of available deep acceptors Fe^{3+} N_A (for reduced crystals) or donors Fe^{2+} $(N_D - N_A)$ (for oxidized crystals). The electric field only imposes an upper limit when the breakdown space–charge field of the material is reached after revealing [124]. In addition, since the electric field inside the material is very weak during holographic recording at elevated temperatures, some of the field–induced (holographic) noise sources (caused by two–beam coupling effect in which strong diffraction interacts with the writing beams such as self–diffraction, hologram bending, etc.) are greatly suppressed. As a result, it would be possible to achieve higher diffraction efficiencies with less noise in a large–scale holographic memory. This is the main reason that we set out to do the investigation on this topic.

Application of high–temperature recording to thermally fix stored holograms was first demonstrated by Staebler in 1975 [112] where more than 500 holograms were stored and fixed in a $LiNbO_3$:Fe crystal by holographic recording at $160°C$. Bragg shift due to thermal contraction (and change of refraction index), as well as the improvement in image quality of the reconstructed holograms were also observed [1].

[1] In their words, this improvement was attributed to the reduction in the "buildup of optical

Meyer et al. also conducted a series of experiments on high-temperature recording, with a conclusion that the "best results were obtained by writing and fixing the phase grating simultaneously at 180°C" [107]. High-temperature recording in $LiNbO_3$:Cu as well as in $KNbO_3$ were also carried out by Hertel, Sommerfeldt, and Montemezzani [108, 125–127]. Theoretical models have been established to explain high-temperature recording process [107–109, 127]. Among them, the one given by Montemezzani et al. is the most complete (with no photovoltaic effect though).

5.1 Theoretical model

In this section, we develop a theoretical model to describe holographic recording in $LiNbO_3$:Fe at elevated temperatures. We describe the writing/erasure asymmetry and its influence on the M/#. A simplified treatment is also presented which offers more insight into the influence of different charge transport mechanisms on the recording dynamics. We show theoretical predictions on the M/# in high-temperature recording based on the numerical evaluation of the model.

5.1.1 Extended Kukhtarev equations

High-temperature recording in photorefractive materials such as $LiNbO_3$:Fe is different from normal holographic recording and erasure because there are two mobile species—electrons and protons, simultaneously involved in the charge transport and redistribution. The formation of electronic charge gratings is the same as in room-temperature holographic recording, involving photo-excitation, diffusion and drift of free carriers—electrons to spatially separate the charges. In $LiNbO_3$:Fe crystals, the photovoltaic effect must be considered as well to fully describe the photorefractive process. At elevated temperatures where the protons become mobile, they drift under the space-charge field set up by the electronic charge separation. As the protonic gratings become stronger, the net space-charge field decreases. At the end, the diffusion of the protons themselves will counter-balance the drift to reach equilibrium.

scattering effects during storage."

Here, we only describe the one–dimensional case where the spatial dimension, x, is along the **c**–axis of the photorefractive crystal. To further simplify matters, we ignore the effects of thermal excitation (its influence will be discussed later).

To describe the formation of both the electronic and protonic charge gratings, the traditional Kukhtarev equations [27,34] have to be extended to accommodate for the extra carriers. The extended set of equations describing both the light–induced and thermal–induced charge transport are listed as below:

$$\frac{\partial N_D^+}{\partial t} = SI(N_D - N_D^+) - \gamma_R N_D^+ n_e, \tag{5.1}$$

$$\frac{\partial n_e}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{q}\frac{\partial J_e}{\partial x}, \tag{5.2}$$

$$J_e = q\mu_e n_e E + K_B T \mu_e \frac{\partial n_e}{\partial x} + P(N_D - N_D^+)I, \tag{5.3}$$

$$\frac{\partial n_p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x}, \tag{5.4}$$

$$J_p = q\mu_p n_p E - K_B T \mu_p \frac{\partial n_p}{\partial x}, \tag{5.5}$$

$$\epsilon\frac{\partial E}{\partial x} = q(N_D^+ - N_A + n_p - n_{p0} - n_e), \tag{5.6}$$

where N_D, N_D^+, N_A are the densities of the dopants, the ionized donors, and the deep acceptors, respectively. n_e is the density of free electrons. n_p is the density of protons, and n_{p0} is its average or DC value. S is the photorefractive cross–section. γ_R is the recombination rate. P is the photovoltaic constant. J_e, J_p are the electronic and protonic current density, respectively. E is the space–charge field. μ_e, μ_p are the mobility of the electrons and protons, respectively. K_B is the Boltzmann Constant and T is the absolute temperature, in Kelvin.

Under the same assumptions we made in Chapter 4—low–modulation depth, and quasi–steady state of the free electron concentration, these equations can be linearized and solved analytically. Among the zero–order solutions, there are some of particular interest. Shown below is the set of equations describing the buildup of the DC

(screening) field during high–temperature recording,

$$\tau_{dc}\frac{\partial E_0}{\partial t} = -E_0 + E_{0ss}, \tag{5.7}$$

$$\tau_{dc} = \tau_{di}\frac{1}{1+\sigma_p/\sigma_e}, \tag{5.8}$$

$$E_{0ss} = -E_{0ph}\frac{1}{1+\sigma_p/\sigma_e}, \tag{5.9}$$

where τ_{di} is the dielectric (relaxation) time constant of the electrons. E_{0ph} is the characteristic photovoltaic field. σ_e and σ_p are the conductivity of the electrons and protons, respectively. The equations suggest that, at elevated temperatures where the protonic conductivity dominates, the buildup of the DC (screening) field becomes much faster than during room–temperature recording because of the contribution of the extra mobile species. In addition, the DC field settles at a much lower value than that in room–temperature recording. As a result, it is no longer strong enough to completely "screen" the photovoltaic field. In normal holographic recording at room temperature, the existence of the photovoltaic field causes writing/erasure asymmetry, and in turn, improves the dynamic range of a memory system. However, as the DC field increases, it gradually screens the photovoltaic field to reduce the degree of asymmetry, resulting in smaller M/#. In high–temperature recording, the DC field is always very small as the above equations indicate. As a result, the writing/erasure asymmetry always exists and the M/# of the system does not decrease after long exposure. This is most desirable in the holographic storage systems.

The first–order solutions to the extended Kukhtarev equations could be expressed by two rate equations for the ionized donors and protons, coupled by Poisson's rela-

tionship.

$$\frac{\partial N_{D1}^{+}}{\partial t} = -j\mu_e K n_{e0} \times \frac{m\left[(E_0 + E_{0ph}) + jE_d\right] - R\left[(E_0 + \frac{N_A}{N_D}E_{0ph}) + jE_d\right] N_{D1}^{+} + E_1}{1 + \frac{E_d}{E_{ue}} - j\frac{E_0}{E_{ue}}} \quad (5.10)$$

$$\frac{\partial n_{p1}}{\partial t} = -j\mu_p K n_{p0}\left[(E_0 - jE_d)\frac{n_{p1}}{n_{p0}} + E_1\right] \quad (5.11)$$

$$E_1 = -j\frac{q}{\epsilon K}(N_{D1}^{+} + n_{p1}), \quad (5.12)$$

where E_{0ph}, E_d, and E_{ue} are the photovoltaic, diffusion, and electronic drift field, respectively. $R = N_D/N_A(N_D - N_A)$ is a constant determined by the oxidation/reduction state of the material. m is the modulation depth in the exposing light pattern.

5.1.2 Steady–state solutions

The closed–form steady–state solutions of the first–order perturbations terms are obtained by making the time derivatives in Equation 5.10 and 5.11 zero. The saturation space–charge field after high–temperature recording is

$$E_1 = -jmE_q \frac{(E_0 + E_{0ph}) + jE_d}{(E_0 + \frac{N_A}{N_D}E_{0ph}) + j(E_d + E_q) + \frac{(E_0 + \frac{N_A}{N_D}E_{0ph}) + jE_d}{E_d + jE_0} R n_{p0} E_q}. \quad (5.13)$$

Here in this equation, we identify an extra term in the denominator as a result of high–temperature recording. Without this term, Equation 5.13 could be reduced to the one corresponding to the result from normal holographic recording (refer to Equation 4.7). Normally, this extra term is very large (Rn_{p0} is large). Therefore, the final space–charge field E_1 is very small. This is the result of the continuous compensation by the mobile protons. This equation is in good agreement to the result Montemezzani obtained in his theory [127], except for the extra photovoltaic field. To gain more insight into the underlying mechanisms of high–temperature recording, we could further simplify this equation by assuming E_0 is zero and N_A/N_D is approximately 1 (for weakly absorptive crystals). The new expression for E_1 is

then

$$E_1 = -j\frac{m}{Rn_{p0}}E_d. \tag{5.14}$$

This indicates that the final space-charge field is determined by diffusion. It agrees with our expectation that it is the diffusion that finally inhibits further growth of the protonic grating. As a result, complete compensation is not possible because a net field is needed to prevent the protonic grating from decreasing by diffusion.

The steady-state solution for the protonic grating strength n_{p1} is

$$n_{p1} = \frac{-mE_q n_{p0}\left[(E_0 + E_{0ph}) + jE_d\right]}{(E_d + jE_0)\left[(E_0 + \frac{N_A}{N_D}E_{0ph}) + j(E_d + E_q)\right] + \left[(E_0 + \frac{N_A}{N_D}E_{0ph}) + jE_d\right] Rn_{p0}E_q}. \tag{5.15}$$

This expression could also be further reduced under the same assumptions, to a much simplified yet intuitive solution below,

$$n_{p1} = -m\frac{N_A(N_D - N_A)}{N_D}. \tag{5.16}$$

This clearly shows that, in high-temperature recording, the grating strength is only limited by either the density of the acceptors Fe^{3+}—N_A, or that of the donors Fe^{2+}—$(N_D - N_A)$. This also suggests that, although the net electric field is not zero, it is not capable of stopping the growth of holographic gratings like in the normal room-temperature recording because the diffusion of the protons is weak.

The complete expression for the electronic grating strength N_{D1}^{+} can be derived in a similar way. We neglect it here because it is in a very complex form and does not provide more information. It should not be surprising that we reach the same form as in Equation 5.16 under the same assumptions. The only difference would be in the signs, indicating a 180° phase shift between the two gratings.

5.1.3 Numeric evaluation of the model—the dynamics

The set of first-order equations (Equation 5.10—5.12) can be easily solved analytically by using the short-circuit ($E_0 = 0$) assumption. The results would be simply exponential functions. However, under open-circuit condition that is more common in practice, E_0 changes with time during exposure. In this case, a closed-form solution is not available. Therefore, we chose to use computer simulation to solve the dynamics of high-temperature recording numerically. The numeric evaluation we used was based on the 4^{th}-order Rouge-Kutta method [73].

For numeric evaluation, we assigned numerical values to the variables found in the theoretical model. Some of parameters are the same as in Chapter 4. They are listed below:

- Photorefractive cross-section: $S = 3.95 \times 10^{-19}\text{cm}^{-2}$ or equivalently for the photo-generation rate, $S = 0.97\text{cm}^2/\text{J}$
- Recombination rate: $\gamma_R = 2 \times 10^{-9}\text{cm}^3/\text{s}$
- Doping level: 0.015 mol % Fe-doped, $N_D = 2.83 \times 10^{18}\text{cm}^{-3}$
- Absorption coefficient $\alpha = 0.55\text{cm}^{-1}$ at 488nm at room temperature: $N_A = 2.82 \times 10^{18}\text{cm}^{-3}$

In general, the absorption of the material varies with temperature, and so are the photorefractive cross-section and recombination rate. The recombination rate, which is determined by the lifetime of the free electrons in the conduction band, should increase with higher temperatures. The photorefractive cross-section S contains the absorption, and is related to the oscillation strength of the Fe^{2+} transition f, as well. This factor f, which is essentially the fraction of photons absorbed which result in an electron in the conduction band, is also temperature-dependent. We measured the absorption of a transmission-geometry crystal as a function of temperature within the range from $27°C$ to $180°C$ and found no significant variation (see Figure 5.1). Therefore, the same absorption coefficient found at room temperature is used. In

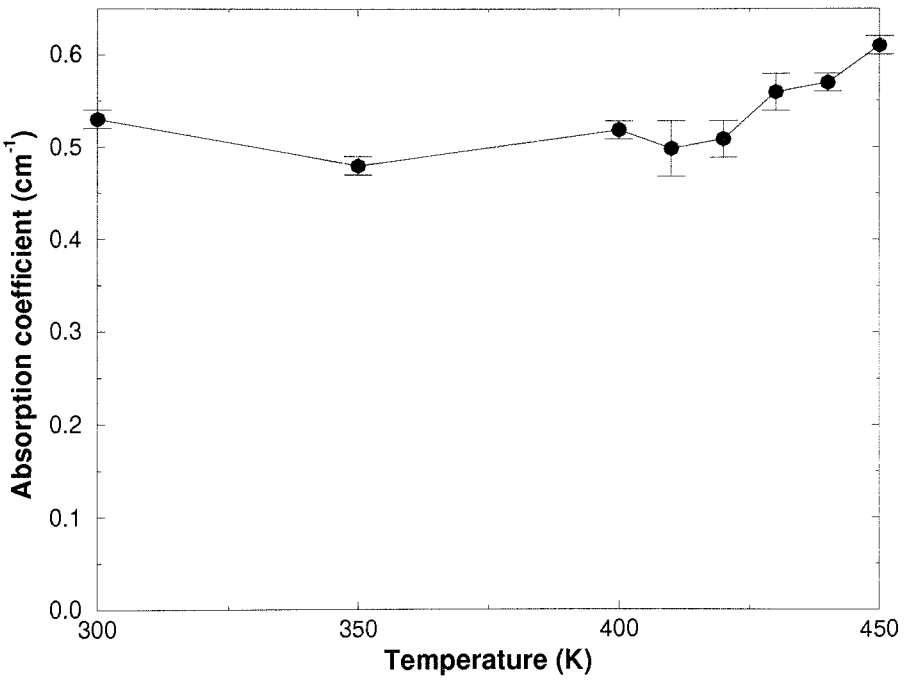

Figure 5.1: The absorption as a function of the temperature.

addition, because we could not find any relevant values for γ_R and S at elevated temperatures from the literature, we still use their room-temperature values. The same relationship between the absorption coefficient and the donor density $(N_D - N_A)$ from Chapter 4 is also used. The density of the protons is assumed to be 10^{19}cm^{-3}.

The conductivity of $LiNbO_3$ has been studied in a wide temperature range which can be divided into three parts corresponding to the different dominant conduction mechanisms. At low temperatures (< 400K), a hopping mechanism of electrons or small polarons is assumed. At temperatures between 400K to 1000K, the conduction is governed by mobile protons. Intrinsic conduction is believed to dominate at temperatures higher than 1000K. Within each of the temperature ranges, the conductivity follows an Arrhenius-type dependence on the absolute temperature.

The electrical conductivity of $LiNbO_3$ at room temperature is about 10^{-18}mho/cm, estimated from the thermal (dark) decay of a hologram written in the crystal [128]. Because of the very small value, direct current conductivity is very hard to measure. In addition, the displacement current caused by the pyroelectric effect from random temperature fluctuation is comparable or even higher than the conduction current.

As a result, the electrical conductivity at room temperature is usually measured by indirect methods, and the reported values cover a large range. The only value widely accepted is the activation energy—0.2ev [34, 129–134]. Most of the scattering of the reported values occurs in the pre-exponential coefficient in the Arrhenius function. Here, instead of adopting the reported temperature-dependence of the electron mobility, we chose the widely accepted value of the activation energy and the room-temperature mobility found by Alphonse and Shah [23,135] to determine the electron mobility at higher temperatures.

The temperature dependence of the protonic mobility is found in [115]. It is in the form of diffusion constant of the OH^- interstitial ions OH_i^-, which is believed to be the same as that of hydrogen diffusing into and inside $LiNbO_3$ crystals.

$$D_p = 0.081e^{-12540/T}, \tag{5.17}$$

where D_p is the diffusion constant of the protons. T is the absolute temperature. The corresponding activation energy is 1.08 ± 0.02ev, which is in close agreement with other reported values from 1.0ev [128], 1.1ev [133], to 1.17 ± 0.02ev [106, 136]. Similar pre-exponential coefficient $0.1\text{cm}^2/\text{s}$ was also reported by Kovacs [123]. The Einstein relationship was used to obtain the value of the proton mobility.

The photovoltaic effect is also temperature-dependent because the photovoltaic constant P contains the absorption coefficient. However, as stated earlier, we did not find significant variation of the absorption within the temperature range we used. The same P was used in our numerical evaluation. We used the same formula for the photovoltaic field in Chapter 4, with slight modification to accommodate the temperature-dependence of the electron mobility.

For a typical behavior of the dynamics in high-temperature recording, we chose the example of symmetric transmission-geometry recording in which the two ordinarily polarized writing beams are symmetric about the normal to the entrance face. The angle between the beams outside the material was 30°, corresponding to a grating vector $K = 6.74 \times 10^4\text{cm}^{-1}$. The temperature was set at $150°C$ and the density of

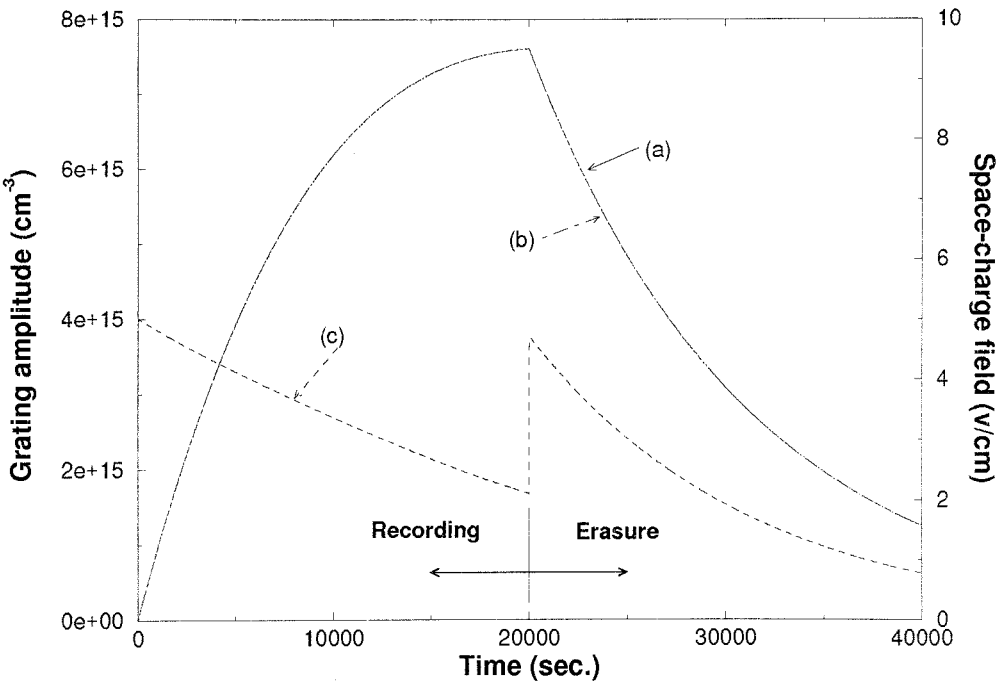

Figure 5.2: The dynamics of holographic recording and erasure at high temperature ($150°C$): (a) electronic grating N_{D1}^{+}; (b) protonic grating n_{p1}; (c) space–charge field E_1.

the protons n_{p0} was assumed to be 10^{19}cm^{-3}.

Figure 5.2 shows the dynamics of the recording and erasure of one hologram at $150°C$. Curve (a) and (b) are for the electronic grating N_{D1}^{+} and the protonic grating n_{p1}, respectively. It is very difficult to tell them apart because the protonic grating catches up with its electronic counterpart rapidly at the beginning and follows closely ever since. Curve (c) shows the dynamics of the space–charge field E_1. Because E_1 arises from the difference between the electronic and protonic gratings, generally it is much smaller than what it would be during normal holographic recording at room temperature. At the beginning, however, E_1 experiences a rapid increase since the protonic grating is lagging behind. As the protonic grating keeps growing to compensate the electronic one, the space–charge field reaches a maximum and begins to fall off as a result of the diminishing difference between the two gratings. The time it takes for E_1 to reach this turning point is an indicator of how fast the protons compensate the electronic grating. The saturation value of E_1 (described by Equa-

tion 5.13) during recording provides information about the degree of compensation. At 20,000 seconds, the recording is "turned off" by making the modulation depth m zero. The hologram is then erased by uniform illumination of the same intensity. It is interesting to note that E_1 reaches another peak before it starts to decrease. This is because that at the point when the recording is turned off, the protons continue to move to further compensate the electronic grating, which already begins to decay. As a result, Curve (a) and (b) would cross shortly after, leading to a zero electric field (this is not shown in Curve (c) since the resolution we chose for the simulation was not fine enough). From this point on, the space–charge field experiences a 180° phase shift and rises. It reaches the peak and starts to decrease after the protons lose their "momentum" and start to follow the electronic grating again.

5.1.4 The M/#: writing/erasure asymmetry

The M/# is a very important system metric. It is the product of the writing slope A_0/τ_r and the erasure time constant τ_e. In a photorefractive crystal without strong photovoltaic effect, the time constant of holographic recording is real. Therefore, $\tau_r = \tau_e$, the M/# is determined solely by A_0, the saturation grating strength. However, in photovoltaic materials such as $LiNbO_3$:Fe, the presence of a strong photovoltaic effect results in a complex time constant, introducing oscillation in the recording. This could be explained as below (neglecting the time evolution of the DC field),

$$h(t) = A_0(1 - e^{-t/\tau}), \tag{5.18}$$

$$1/\tau = 1/\tau_e + j\omega_g, \tag{5.19}$$

where $h(t)$ is the holographic grating strength (field term, unitless). A_0 is its saturation value. τ_e is the erasure time constant. ω_g is the "moving" (oscillation) frequency.

Combining the above two equations, $h(t)$ could be rewritten as [2]

$$h(t) = A_0(1 - e^{-t/\tau_e}e^{-j\omega_g t}). \tag{5.20}$$

Intuitively, the observed buildup of a holographic grating (the absolute value of $h(t)$) could be treated as the result of the competition between a stationary grating and a "moving" one. This competition leads to oscillation [41, 137] which, in turn, results in a larger rate of buildup in the observed diffraction. Taking the absolute value of this grating strength and the time derivative at $t = 0$, we have the expression for the observed writing slope

$$\frac{\partial |h(t)|}{\partial t}\bigg|_{t=0} = \frac{A_0}{\tau_e}\sqrt{1 + \tau_e^2\omega_g^2}. \tag{5.21}$$

During erasure which is described by $h(t) = h_0\, e^{-t/\tau}$, the phase oscillation term is gone because of intensity detection. Therefore, the erasure dynamics can be written as [3]

$$|h(t)| = |h_0|e^{-t/\tau_e}. \tag{5.22}$$

As indicated by Equation 5.21 and 5.22, the effective writing time constant τ_r is shorter than the erasure time constant τ_e by a factor of $\sqrt{1 + \tau_e^2\omega_g^2}$. As a result, the writing is faster than the erasure. This is shown in Figure 5.3. Taking the product of the writing slope (Equation 5.21) and the erasure time constant τ_e, the M/# becomes

$$M/\# = A_0\sqrt{1 + \tau_e^2\omega_g^2}. \tag{5.23}$$

It is increased by a factor of $\sqrt{1 + \tau_e^2\omega_g^2}$. Strictly speaking, because the DC (screening)

[2] A more complete form, including the spatial variations, can be found in reference [73]

$$\delta E = mE_{sc}\left[-cos(Kx + \phi) + exp(-t/\tau_e)cos(Kx + \omega_g t + \phi)\right].$$

[3] Even without taking the absolute value, the decay behavior is very close to that of an exponential decay with τ_e as the time constant, without oscillation.

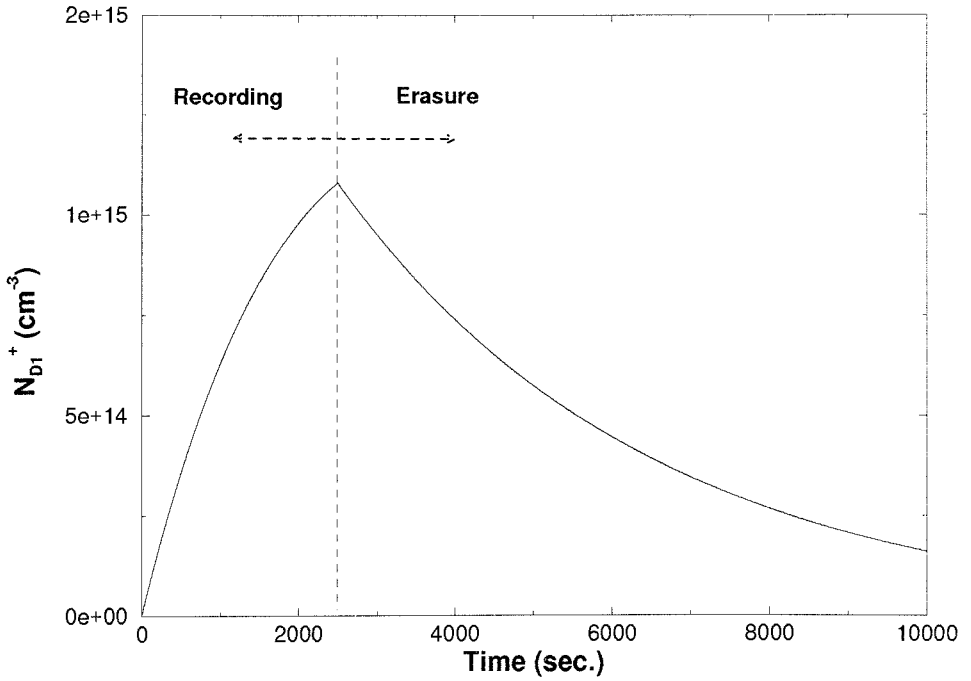

Figure 5.3: Writing/erasure asymmetry of a transmission-geometry hologram ($\theta = 30°$) written at room temperature.

field changes with time (its characteristic time constant is τ_{di}), the exact behavior of grating buildup in $LiNbO_3$:Fe is more complex than simply being in the form of an exponential function. However, the term $\sqrt{1+\tau_e^2\omega_g^2}$ serves as a realistic indicator as of the influence of the writing/erasure asymmetry on the improvement of the M/# in normal room-temperature recording. As seen from Section 5.1.1, during recording at elevated temperatures, the DC field is negligible because of the continuous compensation of the electronic grating by the mobile protons. Therefore, in high-temperature recording, the buildup of holographic gratings is indeed exponential. The degree of asymmetry in the writing/erasure behavior could be described accurately by this term.

One of the reasons that makes the M/# a good system metric is that it can be measured by monitoring the writing and erasure of a single hologram. At room temperature and, of course, with weak holograms [4], the observed diffraction efficiency is proportional to the square of the space-charge field. The space-charge field, in turn,

[4]The product of the photorefractive coupling constant and the thickness is much smaller than 1.

is directly related to the spatial perturbation of the ionized donor density N_{D1}^{+}. Therefore, what is observed is exactly what is recorded. The measurement from the time evolution of the diffraction efficiency during recording and erasure can be employed directly to determine the M/#. However, during holographic recording at elevated temperatures, the observed diffraction is only a measure of the magnitude of the space–charge field, which is directly related to neither of the two gratings, electronic and protonic, but to their difference. So to measure what is really being recorded, we have to find other means. Note here that it is the protonic copy of the stored information which remains nearly intact after the cooling and revealing process and determines the final diffraction efficiency of the fixed hologram. Therefore, n_{p1} comes naturally as a candidate for the calculation and measurement of the M/#. To monitor the growth of protonic grating, we could write a hologram at high temperature, then cool the crystal down to room temperature to reveal it. With fixing efficiency measured beforehand, the strength of the protonic grating could be determined [5]. By repeating this for different recording times, we could obtain the writing slope. In a similar way, the erasure curve could also be obtained. This is a theoretically plausible procedure, however, not practical. Because of the latency of the heating device (we were not able to cool the crystal down immediately after recording), the measured grating strength should be stronger than the actual one since the protons continue to migrate to further compensate for the electronic charge grating as long as the temperature is still high enough to keep them mobile. The extra increase of the protonic grating due to the latency is more significant at the beginning where the protonic grating does not follow its electronic counterpart well. This is problematic because the behavior at the beginning of recording is of most interest to the measurement of the M/#. As a result, this leaves us only one choice in the calculation and measurement of the M/#—the electronic grating N_{D1}^{+}.

Now let us see how we can calculate and measure the M/# from electronic grating buildup. Figure 5.4 explains the basic idea where a single hologram is recorded at

[5]The fixing efficiency is not needed when we want to compare the M/#'s of non–volatile systems using "High–Low" and "Low–High–Low" fixing procedures.

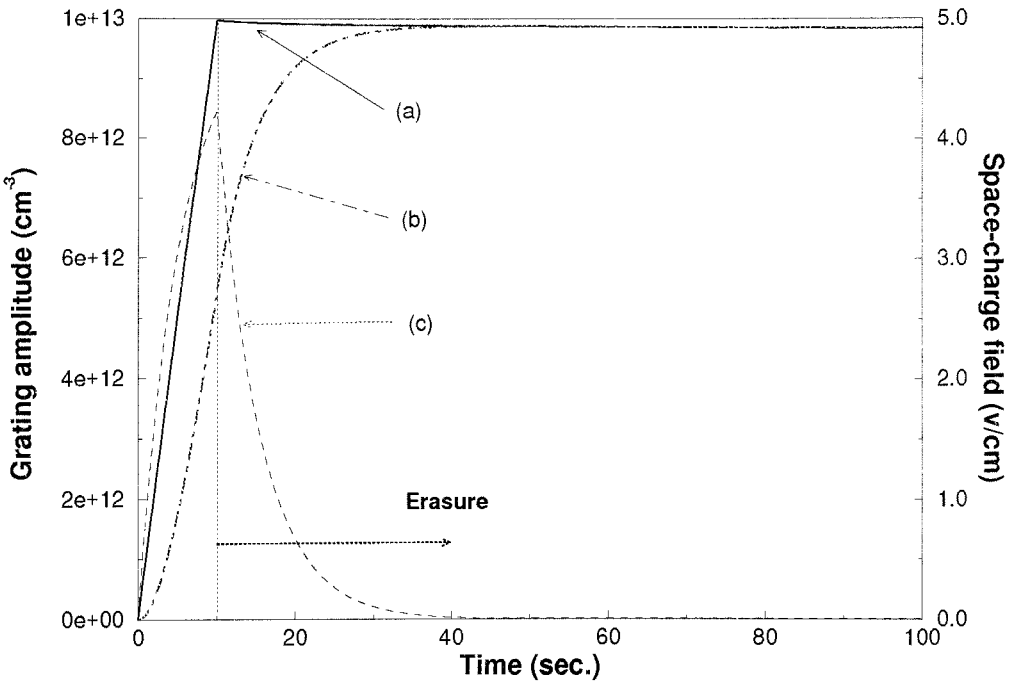

Figure 5.4: Behavior of the electronic and protonic grating during erasure at elevated temperature ($150°C$): (a) electronic grating N_{D1}^{+}; (b) protonic grating n_{p1}; (c) space–charge field E_1.

high temperature ($150°C$) for a short period of time and then erased by the same amount of exposing intensity. We notice that the protonic grating keeps building up even after the erasure begins. Finally, after a certain period of time, the protonic grating catches up with its electronic counterpart, and both of them decay at nearly the same rate. This indicates that after a certain amount of erasure, one can not tell the difference between the two gratings. In other words, if one starts to look at the behavior of the protonic grating later, beyond the point where it catches up, it looks as if it followed the same recording and erasure dynamics of the electronic grating. Furthermore, one should note that the M/# is a good measure for the system dynamic range only in the context of a large–scale holographic memory system where many holograms are superimposed in the same volume, with low diffraction efficiencies. In multiple–hologram recording, most of the holograms would experience fairly long erasure during the exposure for the rest of the holograms in the recording sequence. Therefore, for these holograms, it is sufficient for the protonic gratings to "fully" catch

up and follow their electronic counterparts. The observed behavior of the electronic and protonic gratings are the same. This justifies the use of the electronic grating N_{D1}^{+} as the target for calculation and measurement of the M/#.

5.1.5 Simplified model

In this section, we establish a way to theoretically predict the M/# of holographic storage at elevated temperatures. We learn from previous discussions that the DC field changes with time during recording, so a closed–form solution can not be obtained from the theoretical model. In addition, the existence of two mobile species makes it difficult to reach a "clean" and insightful expression for the M/#, similar to the one we have for normal holographic recording (Equation 5.23).

To assess the M/# of high–temperature recording and, especially, to understand the writing/erasure asymmetry and its influence on the M/#, as well as the interaction between the diffusion and photovoltaic effect in causing the asymmetry, we will make further assumptions to simplify the theoretical treatment. Since there are two mobile species involved in high–temperature recording, there must be two exponential terms in the dynamics. However, because the proton conductivity is much larger than the photoconductivity of the electrons, one of the exponential terms associated with the protons must be changing much faster than the other one, which is determined by the electron conductivity. Therefore, the double–exponential solution could be reduced to a single exponential one. We assume that the electric field inside the material is zero because of the fast, nearly complete compensation of the electronic grating by the protonic one. As a result, there is no interaction (coupling) between these two gratings, and the dynamics of the grating buildup can be reduced to a single–exponential term, determined only by diffusion and photovoltaic effect. Under these assumptions, the two first–order rate equations 5.10 and 5.11 are decoupled. The dynamics of the electronic charge grating is then simplified to

$$\tau \frac{\partial N_{D1}^{+}}{\partial t} = -N_{D1}^{+} + N_{D1}^{+}(\infty). \tag{5.24}$$

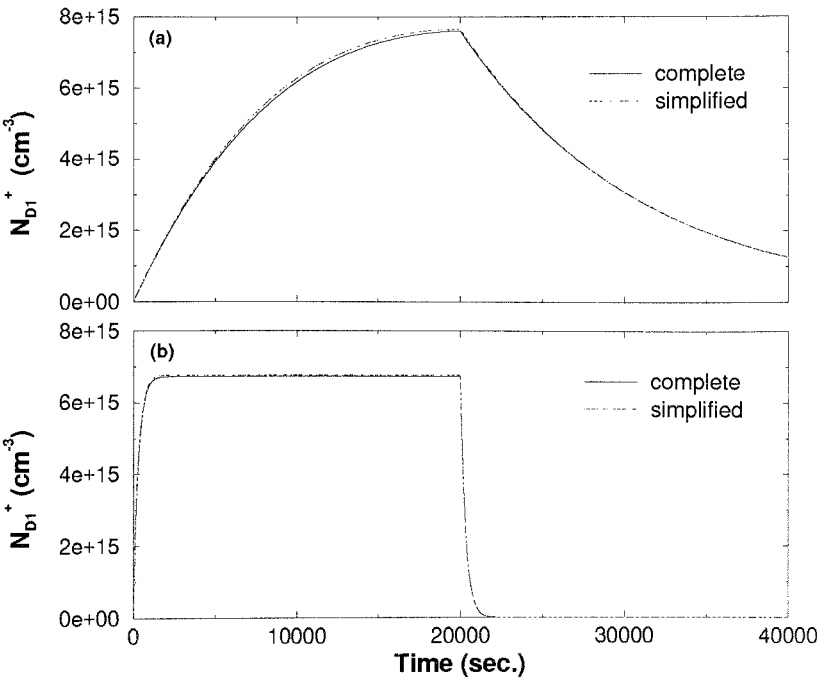

Figure 5.5: Comparison between the simplified model and the complete treatment (at 150°C): (a) transmission geometry ($\theta = 30°$); (b) 90° geometry.

The complex time constant τ is

$$\tau = \frac{\tau_{di}(1 + \frac{E_d}{E_u})}{\frac{E_d}{E_q} - j\frac{N_A}{N_D}\frac{E_{0ph}}{E_q}}, \tag{5.25}$$

where τ_{di} is the dielectric time constant. The saturation value of N_{D1}^{+} is

$$N_{D1}^{+}(\infty) = m\frac{N_A(N_D - N_A)}{N_D}, \tag{5.26}$$

which is in the same form as the one we derived from the complete treatment. Again this shows that, because of the zero-field assumption, N_{D1}^{+} is only limited by the density of the available donors Fe^{2+} or acceptors Fe^{3+} (in our case where the crystal is oxidized, it is mainly limited by $(N_D - N_A)$, the density of Fe^{2+}).

Figure 5.5 shows the computer-simulated result of the dynamics of the electronic grating N_{D1}^{+} in high-temperature recording and erasure. Plot (a) is for the transmission geometry where the outside angle between the two beams is 30°. Plot (b) is for the 90° geometry. The solid curves in both plots show the time evolution of

N_{D1}^{+} solved by the complete theoretical model (Equation 5.7—5.12). The dot–dashed curves are the result of the simplified treatment. As shown in this figure, the solutions evaluated by both the simplified model and the complete theoretical treatment are in close agreement to each other. From this point on, we shall use the closed–form solutions to the simplified equations to theoretically predict the performance of high–temperature recording, in terms of the improvement in the M/# with respect to that obtained in room–temperature holographic recording.

5.1.6 Predictions of the M/#

The M/# is a measure of the dynamic range of a memory system. It depends on the material parameters—the doping level, reduction/oxidation state, photo-sensitivity, diffusion, photovoltaic and electro–optic effect. In addition, the M/# also depends on the system parameters such as the exposing light intensity, modulation depth, aperture of the writing beams, bulk absorption and the interaction length [40,41]. Here in this section, we focus on the material–dependence of the M/# [6] and establish a way to compare the performance of high–temperature and room–temperature recording based on the improvement in the M/#, not its absolute value.

From Equation 5.25, we can extract the erasure time constant τ_e and the "moving" (oscillation) frequency ω_g. As mentioned earlier, the product of these two terms indicates the improvement in the M/# as a result of the writing/erasure asymmetry. It is now written as

$$\tau_e \omega_g = \frac{N_A}{N_D} \frac{E_{0ph}}{E_d}. \tag{5.27}$$

This expression clearly shows how the M/# is affected. N_A/N_D ($0 \leq N_A/N_D \leq 1$) is a measure of the oxidation/reduction state of the crystal. This expression suggests that, to obtain higher M/#, N_A/N_D should be as close to 1 as possible—the crystal is

[6] In other words, we analyze the "local" M/# [41] which is determined by the properties of the storage medium, instead of the "externally observed" M/# which also depends on the system parameters.

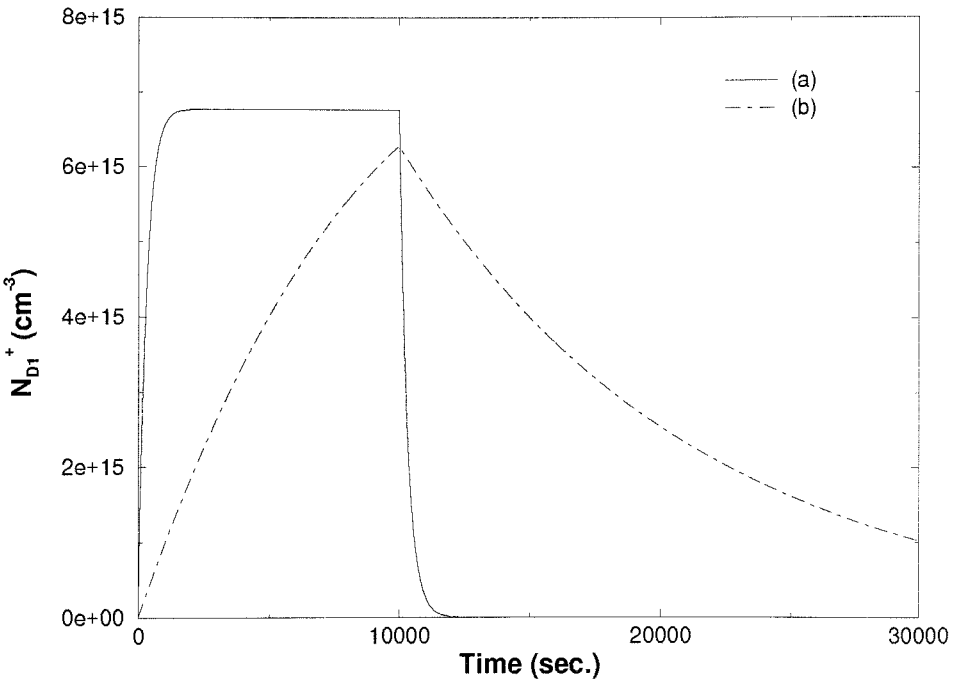

Figure 5.6: Comparison between (a) 90° and (b) transmission–geometry recording.

heavily oxidized (transparent) [7]. In addition, $\tau_e\omega_g$ is proportional to the characteristic photovoltaic field E_{0ph}, corroborating the role the photovoltaic effects plays in causing the oscillation and the writing/erasure asymmetry. On the other hand, it indicates that there is less improvement in the M/# if the diffusion is stronger. This could be explained by the fact that, with no field (thus no drift), the erasure is determined by diffusion alone. The decay due to diffusion has a time constant proportional to the grating period (inversely proportional to the grating vector). Therefore, for larger diffusion fields (smaller grating periods), the erasure is faster, reducing the degree of writing/erasure asymmetry, and in turn, making the M/# smaller [8].

The M/# of high–temperature recording vs. the grating vector: We evaluated the cases for both transmission–geometry and 90°–geometry recording at

[7]On the other hand, the material becomes less sensitive as the density of Fe^{2+}, $(N_D - N_A)$ decreases. A trade–off must be made to optimize the overall performance similar to the treatment found in Reference [40,41].

[8]This is for the M/# before revealing. E_d will come to the picture when the thermal fixing efficiency is included to obtain the final M/# of the system. Because the fixing efficiency depends on E_d, there should be a bump (optimum) in the curve relating the M/# and the grating vector.

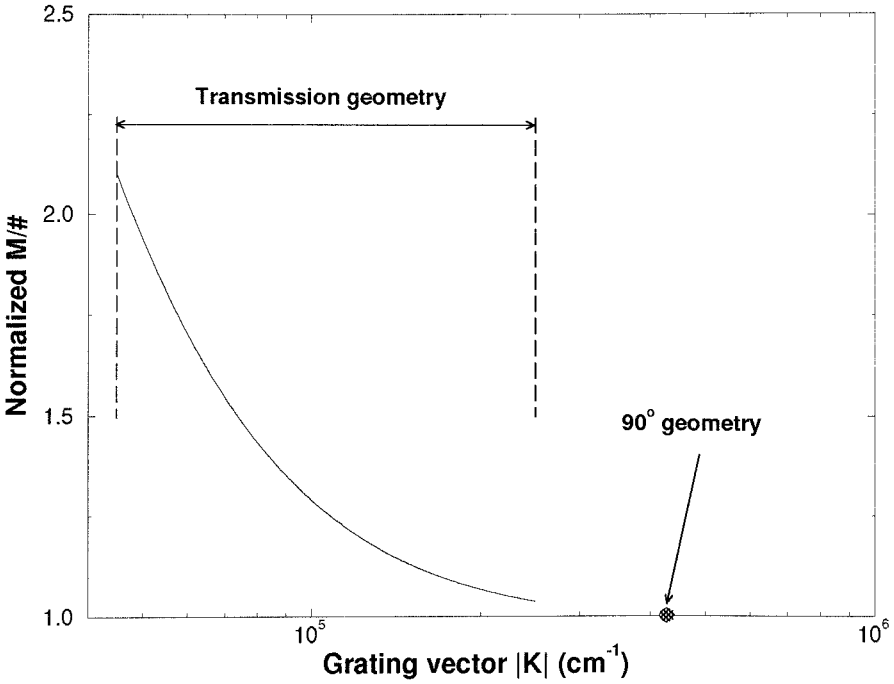

Figure 5.7: Normalized M/# as a function of the grating vector in high–temperature recording (150°C).

150°C. The results are shown in Figure 5.6 for comparison. In transmission–geometry recording shown by Curve (a), a profound asymmetry is observed between the writing and erasure behaviors. To the contrary, as Equation 5.27 predicts, the erasure is nearly as fast as the writing in 90°–geometry recording (Curve (b)). Because in both cases, the saturation diffraction efficiency is only limited by either the density of Fe^{2+} or Fe^{3+}, the M/# in 90°–geometry recording is smaller.

The M/# of high–temperature recording is obtained by combining Equation 5.23 and 5.27.

$$M/\# = A_0\sqrt{1+(\frac{N_A}{N_D}\frac{E_{0ph}}{E_d})^2}. \tag{5.28}$$

Its dependence on the grating vector $|\vec{K}|$ is plotted in Figure 5.7. In order to highlight the influence of material parameters, we did not include system parameters to obtain an absolute value of the M/#. Instead, we normalized the M/#'s by its value in 90°–geometry recording. Since A_0, the saturation grating strength, is always the same in

high–temperature recording, the only varying term is the diffusion field E_d.

The improvement in the M/# by high–temperature recording vs. the grating vector: To compare the performance of high–temperature and room–temperature recording in terms of the M/#, the saturation grating strength has to be considered since the electronic charge grating stops growing under the space–charge field before it can reach the upper limit imposed by the doping and reduction/oxidation state of the material in room–temperature recording. The material–dependent ("local") M/# at room temperature can be written in a form similar to the one in [41], except for the consideration of bulk asorption, beam aperture, etc.

$$\begin{aligned} M/\# &= A_0 \sqrt{\frac{E_{0ph}^2 + E_d^2}{\left(\frac{N_A}{N_D} E_{0ph}\right)^2 + (E_d + E_q)^2}} \sqrt{1 + (\tau_l \omega_l)^2}, \\ \tau_l &= \tau_{di} \frac{1 + E_d/E_u}{1 + E_d/E_q}, \\ \omega_l &= \frac{1}{\tau_{di}} \frac{N_A}{N_D} \frac{E_{0ph}}{E_q} \frac{1}{1 + E_d/E_u}, \end{aligned} \tag{5.29}$$

where A_0 represents space–charge–limited saturation value. It is the same as that in Equation 5.28 if the same modulation depth is used in both cases. To reach Equation 5.29, short–circuit condition ($E_0 = 0$) has to be applied. Otherwise, it is only the instantaneous M/# at the beginning of holographic recording. The M/# shall drop as the DC field increases during exposure. In other words, Equation 5.29 represents the best–scenario M/# of a holographic storage system at room temperature.

The ratio of the M/# at high and room temperature is obtained by dividing Equation 5.28 by 5.29. It is a function of the grating vector $|\vec{K}|$ as shown in Figure 5.8. It shows significant improvement in the M/# in the low spatial–frequency range in transmission–geometry recording. However, for 90° geometry, the M/# of high–temperature recording is only half of that obtained at room temperature. This could be attributed to two reasons: First, because of the strong photovoltaic field and higher spatial frequency in 90°–geometry recording ($E_d > E_q$), the saturation grating strength in room–temperature recording becomes quite close to the space–

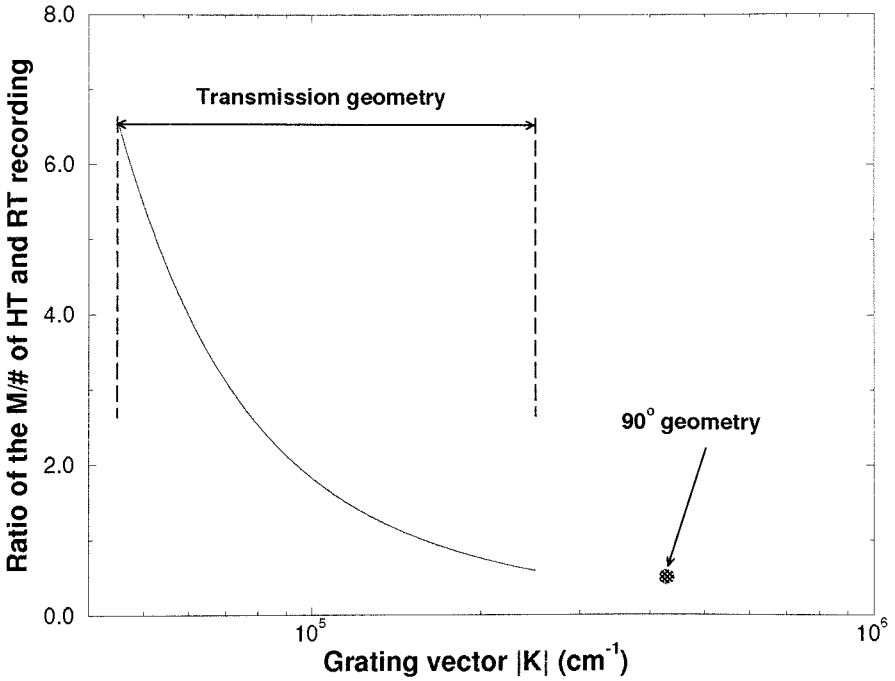

Figure 5.8: The ratio of the M/# in high-temperature ($150°C$) and room-temperature recording as a function of the grating vector.

charge limit [9]. Therefore, 90°-geometry recording does not gain as much at elevated temperatures. On the other hand, the degree of asymmetry in writing and erasure dynamics is always less with larger spatial frequencies. At higher temperatures, it becomes even worse since E_{0ph} decreases while E_d increases with larger electron mobilities. Therefore, in 90°-geometry recording, the loss due to decreased writing/erasure asymmetry overwhelms the gain in the saturation strength. As a result, the M/# of high-temperature recording is smaller than its room-temperature value.

The improvement in the M/# by high-temperature recording vs. temperature: The influence of the temperature on the M/# is also evaluated. Figure 5.9 shows the dynamics of a transmission-geometry grating at room temperature, $120°C$ and $150°C$, respectively. As the temperature rises, the recording and erasure both become faster. From numeric simulation of the same transmission-geometry holo-

[9]The term $\sqrt{\frac{E_{0ph}^2+E_d^2}{\left(\frac{N_A}{N_D}E_{0ph}\right)^2+(E_d+E_q)^2}}$ in Equation 5.29 shows how the different fields (transport mechanisms) interact and determine the final saturation value of recording at room temperature.

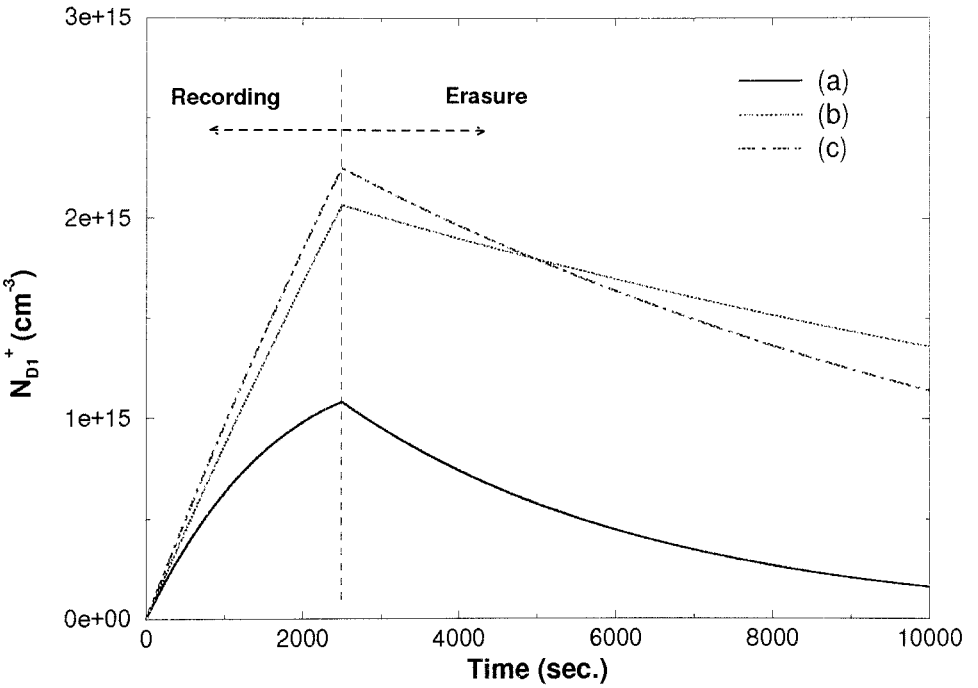

Figure 5.9: Dynamics of the electronic charge grating N_{D1}^{+} at (a) room temperature, (b) $120°C$, and (c) $150°C$ (transmission geometry $\theta = 30°$).

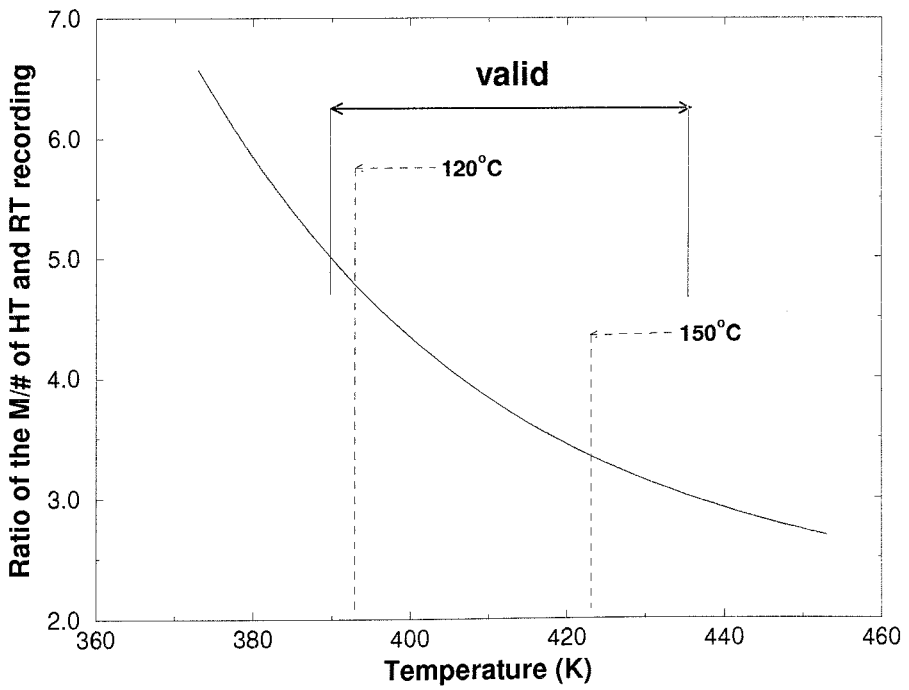

Figure 5.10: The ratio of the M/# in high-temperature and room-temperature recording as a function of the temperature in transmission-geometry ($\theta = 30°$).

gram, the ratio of the M/# of high and room temperature recording as a function of the temperature is also obtained and shown in Figure 5.10. It shows that the gain in the M/# becomes less as the temperature rises. This is the result of less asymmetric writing and erasure behaviors with higher temperatures which could be clearly seen from Equation 5.27. It should be noted that this curve is only valid at temperatures high enough for sufficient proton compensation to take place and validate the zero–field assumption we made earlier. In addition, the temperature should be well below 180° where thermal ionization takes place.

In conclusion, in high–temperature recording, holographic recording becomes faster with higher temperatures. In the meantime, however, the M/# sees less improvement by high–temperature recording since the erasure becomes much faster as well. High–temperature recording results in the highest diffraction efficiency in a single hologram, which is only limited by the reduction/oxidation state of the material. On the other hand, the M/# is worse than normal room–temperature recording in 90° geometry which is the optimal system configuration for large–scale holographic memory using angle multiplexing. As a result, high–temperature fixing is not a practical solution to the volatility problem in mass holographic storage.

5.2 Experiment results

The experimental setup for the measurement of the M/# at elevated temperatures is shown in Figure 5.11. We used the 488nm line from a Coherent Innova 300 argon laser as the light source. The laser beam was collimated to an aperture of 8mm diameter. After an Acousto–Optic Deflector (AOD), the beam was split into two writing beams, symmetric about the normal of the entrance face on the crystal. One of the beams was imaged onto the crystal by a 1:1 4F system and was also used to read out the hologram. This 4F system allowed the continuous scanning of the incident angle of this reading beam at the same location on the crystal. Each of the two writing beams contained approximately 35mW/cm^2. Preferably, for readout, the reading beam should be much weaker to avoid interfering with the existing gratings.

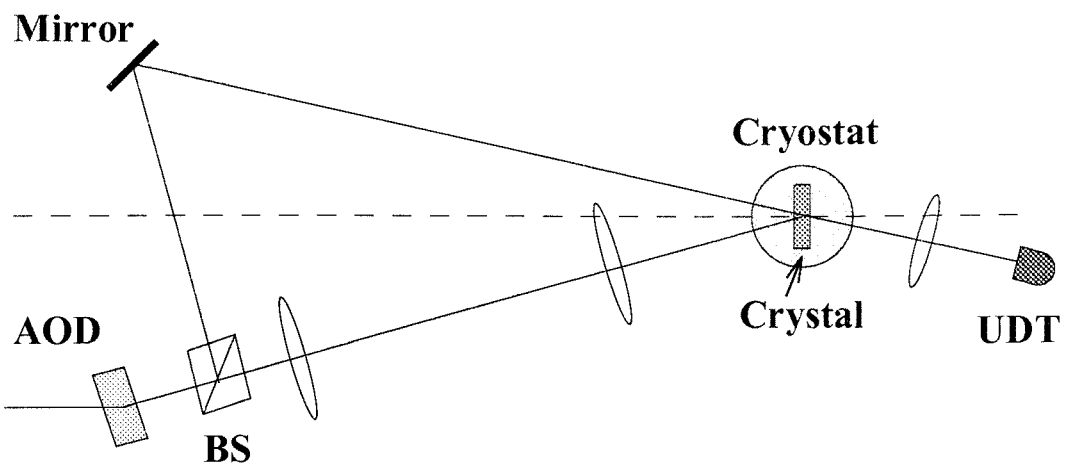

Figure 5.11: Experimental setup for the M/# measurement at elevated temperatures.

However, to observe the very weak diffraction during high–temperature recording, the same light intensity was used. In other cases where the hologram was measured after revealing, the readout intensity was deceased to 2.3mW/cm^2. A photo–detector, UDT Model 370, was used to sense the diffracted signal. In order to reduce the influence of the readout beam on the hologram, the faster analog output from the detector was used instead of the standard GPIB digital output. The analog signal from the UDT was sampled by a 8–bit A/D converter. The frequency response of this data acquisition system was tested to be 20Hz. Therefore, the settling time for each readout was set at 0.1 second. The AOD was tuned to center at 50MHz. At this RF frequency, the angle between the two writing beams outside the material was 30.3°, corresponding to a grating vector of magnitude as $6.73 \times 10^4 \text{cm}^{-1}$ inside the crystal. Both writing beams were vertically (ordinarily) polarized at the crystal.

The recording material was a transmission–geometry $LiNbO_3$:Fe crystal, of dimension 10mm$\times$ 10mm$\times$ 1mm. The absorption of the crystal is 0.55cm^{-1} at room temperature. The crystal was mounted on a copper holder inside a Cryostat (Oxford Instrument Optistat^{CF-V}). The Cryostat is capable of providing a wide range of temperatures, from 4K to 500K. In our experiment, we only used its heating function for thermal fixing. To avoid ambient behaviors caused by heating, the cryostat was pumped down to 70mTorr by a mechanical pump. At this vacuum level, we observed

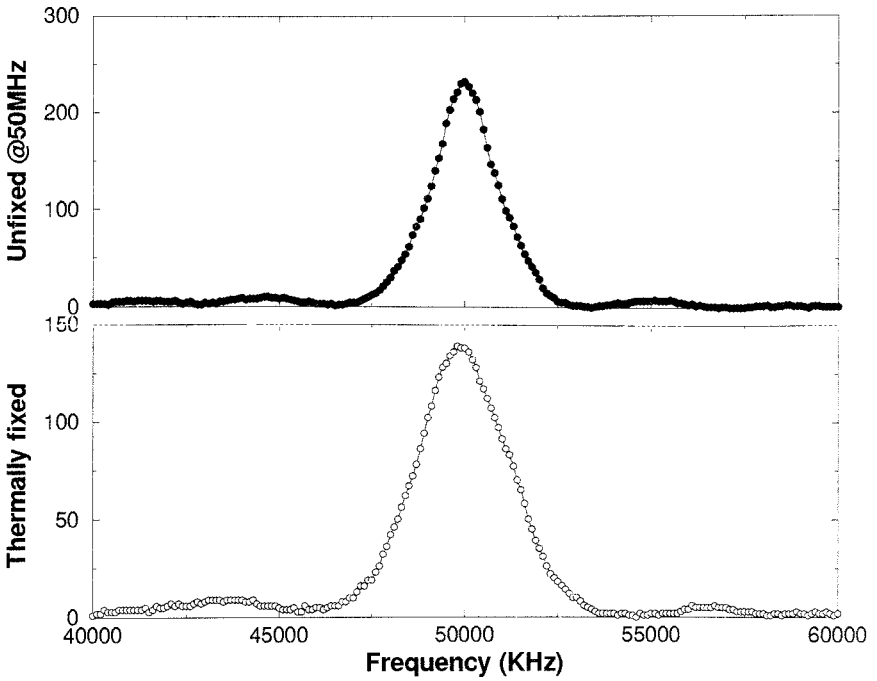

Figure 5.12: Angle selectivity curve of an unfixed and a thermally fixed hologram.

no significant influence of heating on holographic recording.

The reason for using the AOD is to deliver accurate angle deflection. This is very important because the selectivity curve of the hologram recorded at high temperatures is shifted as a result of the thermal contraction and change of refraction index of the material. Therefore, measurement of selectivity curve with fine resolution was required for reliable reading of the diffraction efficiency of the fixed holograms at room temperature. Figure 5.12 shows the angle selectivity curves of an unfixed and a thermally fixed hologram written at $150°C$. Both holograms were recorded using an RF frequency of 50MHz. After recording, the AOD scanned the reference beam to obtain the selectivity curve. The curve of the unfixed hologram peaks at exactly the same frequency used for recording. However, the curve of the thermally fixed hologram is shifted due to thermal contraction and change of refraction index.

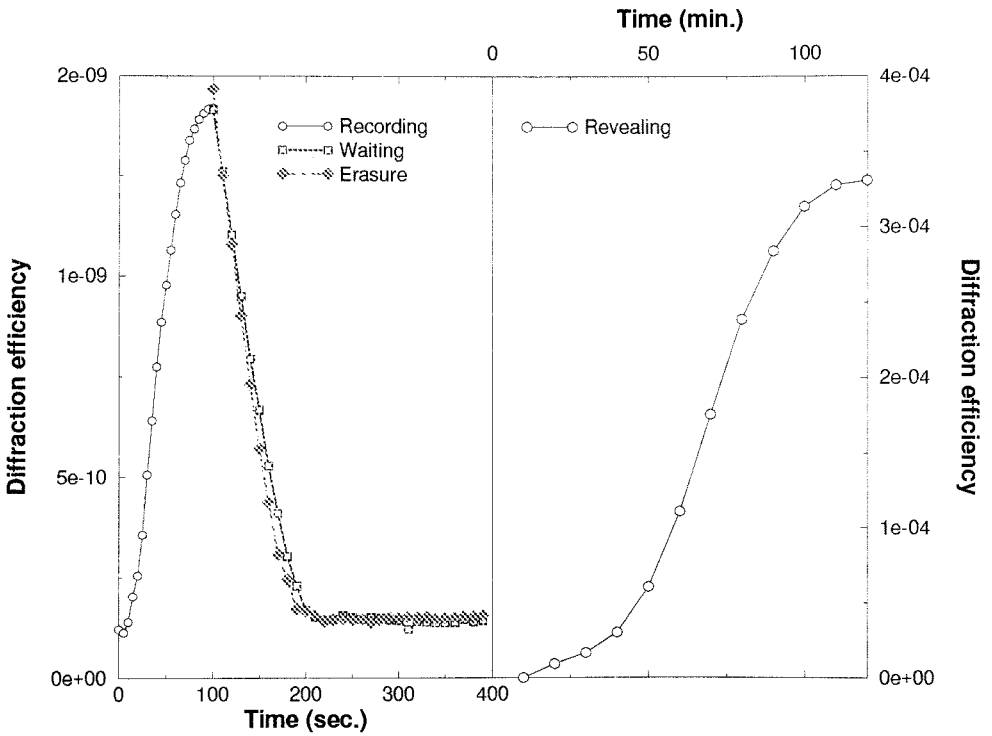

Figure 5.13: Writing/erasure dynamics at high temperature.

5.2.1 Writing/erasure dynamics

We started our experiment by monitoring the time evolution of the diffraction directly observed during holographic recording and erasure at $150°C$. The results of a typical measurement are shown in Figure 5.13. The first part of the curve corresponds to what was observed during recording. The diffracted signal, though much weaker than what it would be under normal recording conditions, clearly rises up, as a result of the initial increase in the difference between the two gratings during recording. After recording, two curves were taken. The first one is for hologram erasure with a Bragg–mismatched beam (the AOD was operating at 40MHz, beyond the 3^{rd}–null of the angle selectivity curve). This erasure beam was tuned to have approximately the same intensity as that of the total intensity of the two writing beams for recording ($\approx 68\text{mW/cm}^2$). To obtain the second one, we simply blocked all the writing beams to let the protonic grating fully compensate its electronic counterpart. The rate of decay in the first case is determined by photoconductivity of the electrons as well as the (thermal) conductivity of the protons; while in the second case, it is mainly

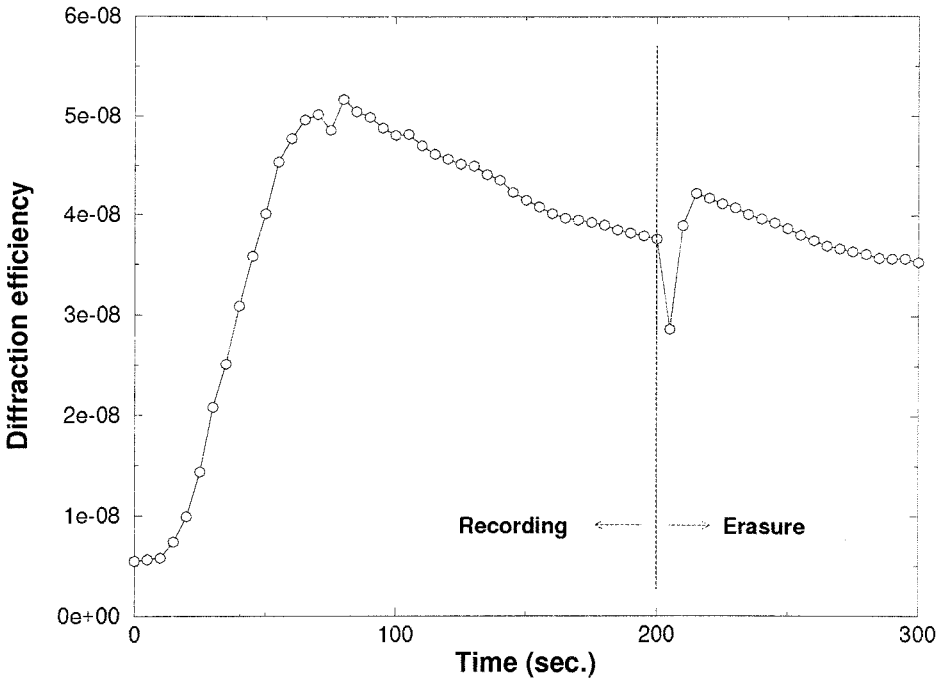

Figure 5.14: Another example of writing/erasure dynamics at high temperature.

determined by the proton conductivity alone. The fact that the two decaying rates are close to each other indicates that at this temperature, the proton conductivity is much higher than the photoconductivity of the electrons as observed by many researchers. Strictly speaking, the rates of change in both of the cases should also depend on the thermal conductivity of the electrons. However, the diffraction drops rapidly because of proton compensation; further decay caused by electron thermal conductivity could not be resolved from the fairly high noise level due to scattering and back-reflection from the windows on the Cryostat.

The revealing phase is shown in the second part of the plot. Instead of revealing the fixed hologram using either of the two laser beams, we used UV erasure to avoid the buildup of scattered noises. Because UV illumination heats up the material by 10–20K, we used the AOD to scan through the entire angle range to measure the selectivity curve and then sampled the peak value as the readings of the revealed hologram strength. The final diffraction efficiency of the fixed hologram is 3.3×10^{-6}.

Another experiment was carried out in an attempt to reproduce the result found in Reference [112], with stronger writing beams. We obtained similar dynamics which

is shown in Figure 5.14. However, the observed diffraction efficiency during recording was much less.

5.2.2 Experimental procedure for the M/# measurement

The measurement of the M/# consists of two steps—the measurement of the writing slope and the erasure time constant from single-hologram recording and erasure. Under normal recording conditions, this could be done in real-time, by monitoring the time evolution of the diffracted signal during holographic recording and erasure. However, as we have shown previously in this chapter, the observed diffraction at high temperatures could not be related directly to the hologram strength. Instead, we need to use the strength of the electronic grating as an indicator of the hologram strength for calculation and measurement of the M/# for high-temperature recording. The basic difference between the experimental procedures to measure the M/# under the two recording conditions is that the grating strength in high-temperature recording could only be determined after revealing.

To measure the writing slope, a hologram was recorded at elevated temperature for a certain duration of time. Instead of cooling the crystal immediately after recording, we turned off the writing beams and let the crystal remain at the high temperature for the protonic grating to "fully" compensate the electronic one. After 10 minutes, we switched off the heater. It takes more than 2 hours for the crystal to cool down to room temperature because of the vacuum condition inside the Cryostat. UV illumination was used to reveal the fixed hologram, for 150 minutes. After revealing, the selectivity curve was taken and the peak value was sampled to determine the final diffraction efficiency of the fixed hologram. To refresh the crystal for a new measurement, it was heated up to $200°C$ for 60 minutes. By repeating this procedure many times, with different recording time, the writing curve of the electronic grating at high temperatures could be obtained. A typical writing curve is shown in Figure 5.15. Linear regression on the square root of the diffraction efficiencies as a function of the recording time yields the writing slope.

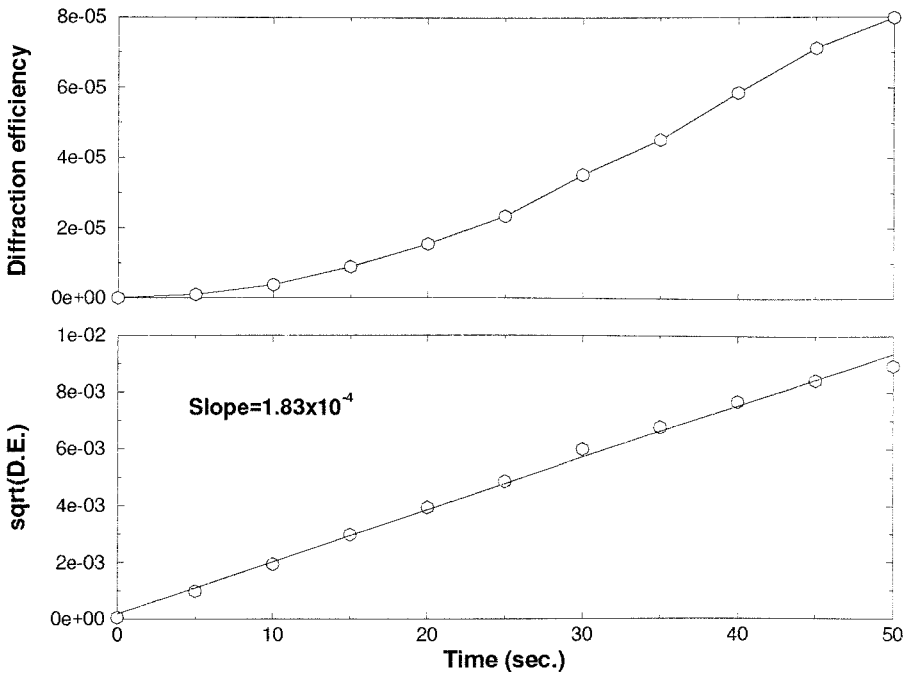

Figure 5.15: A typical writing curve at $150°C$.

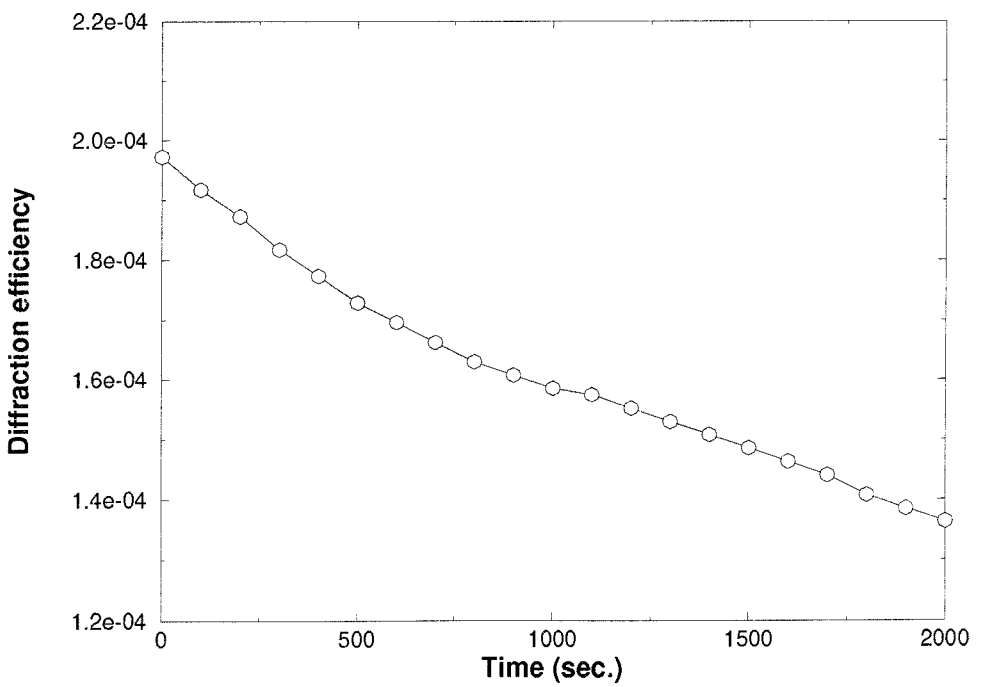

Figure 5.16: A typical erasure curve at $150°C$.

Temperature ($°C$)	Writing slope	Erasure time constant	M/#
Room temperature	1.49×10^{-4}	4100	0.61
120	1.72×10^{-4}	6100	1.04
150	1.84×10^{-4}	4800±200	0.88

Table 5.1: The writing slope, erasure time constant, and the resulting M/# of holographic storage at different temperatures.

To obtain the erasure curve, a single hologram was first recorded to the same level (with the same recording time). Then the crystal was illuminated for a certain amount of time, by one of the writing beams, Bragg–mismatched by resetting the frequency of the RF signal that drives the AOD to 40MHz. Again, after the erasure was done, we waited for 10 minutes before we turned off the heater. The same cooling, revealing procedure and selectivity curve measurement, were applied to yield the final diffraction efficiency of the fixed hologram. Multiple data points were obtained by repeating the procedure with increasing erasure time intervals. Between each measurement, the crystal was refreshed in a similar way. An erasure curve is shown in Figure 5.16.

5.2.3 Experiment results and comparison to theory

Two experiments were carried out at $120°C$ and $150°C$. The results are summarized in Table 5.1 along with the data we obtained from room–temperature recording.

From Table 5.1, the M/#'s we obtained in high–temperature recording are higher than that of normal storage at room–temperature. To compare the M/#'s of non–volatile storage systems using high–temperature recording and "Low–High–Low" fixing, the thermal fixing efficiency has to be included. Combining the fixing efficiency measured from previous experiment and the M/# of room–temperature recording listed in Table 5.1, the M/# of "Low–High–Low" fixing is 0.39. As a result, the M/# of high–temperature recording is nearly three times higher ($2.6^{\times}$ at $150°C$ and $3.1^{\times}$ at $120°C$), corresponding to an increase in the diffraction efficiency by a factor of 10.

However, the improvement in the measured M/#'s is not as much as what the

theory predicts. For example, for recording at 150°C, our theory suggests an M/# which is 3.5$^{\times}$ larger than that with "Low–High–Low" fixing (Figure 5.10). We suspect the reason to be the thermal conductivity of the electrons. Thermal excitation of electrons is uniform throughout the volume. Therefore, it acts like additional uniform illumination that speeds up the recording and erasure by providing more free carriers [10]. On the other hand, it reduces the effective modulation depth to

$$m_{eff} = \frac{m}{1+\beta/SI_0} = \frac{m}{1+\sigma_{th}/\sigma_{ph}}, \tag{5.30}$$

where σ_{th} and σ_{ph} are the thermal and photo–conductivity, respectively. This effectively reduces the recording rate. As a result, the M/# we obtained from experiment should be lower than the theoretical prediction without thermal conductivity. In addition, we suspect that the system parameters which we did not include in our predictions of the M/# might affect the final results as well.

5.2.4 Multiple holograms storage

In a separate setup shown in Figure 5.17, we demonstrated storage of multiple hologram in a 90°–geometry crystal at 150°. The crystal was a 0.015% Fe–doped $LiNbO_3$, of dimensions 5mm$\times$ 5mm$\times$ 8mm. Its absorption coefficient at 488nm was about 0.5cm^{-1}. The diameter of the reference beam at the crystal was about 4.5mm. The measured selectivity was 0.007°, corresponding to a linear horizontal translation of 21μm in the periscope (Chapter 2). In order to reduce crosstalk noise, we chose an inter–hologram spacing of 60μm, which is at about the third–null of the Bragg selectivity curve of the holograms. Because of the small size of the crystal and the limited aperture of the Cryostat, we did not use the entire horizontal angular range of the reference beam. Instead, we recorded 600 holograms on 3 fractal rows which were spaced every other mirror strip—4.8mm apart vertically. In the signal arm, a 2:1 4F system demagnified the image which was displayed on the Epson TVT6000 LCTV

[10]The measured time constants are much smaller than the ones predicted by the theory. However, in room–temperature recording, the measured time constants are always in close agreement (within 20%) with the theoretical predictions based on the same set of numerical values.

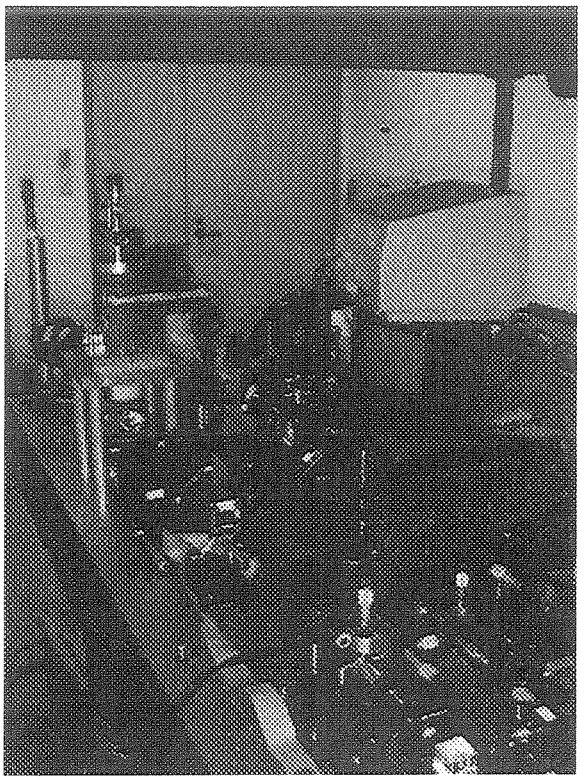

Figure 5.17: Photo of the experimental setup for 600–hologram storage.

(480×440 pixels), followed by a Fourier–transform lens of 300mm focal length. The focused signal beam was 4mm away from the exit face of the crystal.

An exposure schedule was used to equalize the strength of all the holograms. The erasure time constant was first chosen based on the measurements from the previous experiments, but determined largely by trial and error since an erasure measurement at high temperature is very time–consuming. During reconstruction, a noticeable Bragg shift was observed (for detailed characterization, refer to Section 5.3.2). We plot the diffraction efficiency of all the 600 reconstructed holograms as the result of continuous reference beam scanning in Figure 5.18. We notice that, because of the prolonged readout for all the stored holograms, the ones near the end become more noisy. This is the result of the buildup of holographic noise due to the reference beam exposure. It could be cured by illuminating the crystal with UV light once in a while [112]. The average diffraction efficiency was 2.2×10^{-7}, corresponding to an M/# of 0.27. The M/# we obtained in a separate measurement at room temperature

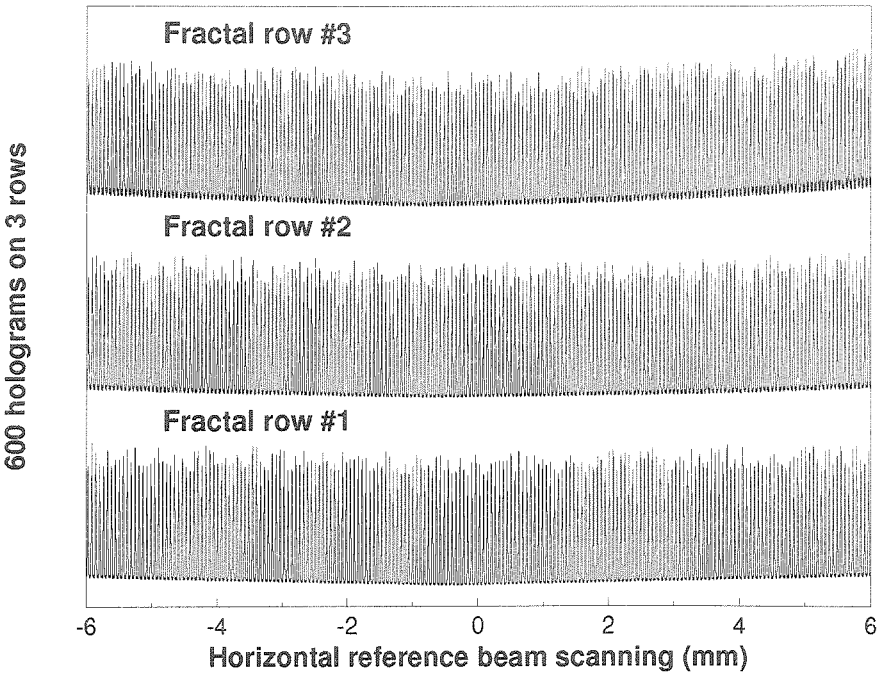

Figure 5.18: Diffraction efficiency of the 600 fixed holograms.

of the same crystal in the same setup was 0.82. The drop in the M/# is 50% more than the theoretical prediction. We can not attribute this drop to the short exposure schedule we usually adopted to reduce the buildup of holographic noise [11]. Instead, the low M/# could be the result of electron thermal conductivity, non-plane wave signal, scattering and Fresnel loss from the windows of the Cryostat.

5.3 Miscellaneous topics

5.3.1 Thermal contraction and change of refraction index

For thermal fixing using high-temperature recording, the recording and readout are carried out at different temperatures. Therefore, the temperature dependence of the refraction index and thermal expansion/contraction have to be considered.

$LiNbO_3$ is not completely defined by its chemical formula. Large deviations from this stoichiometric composition may exist, characterized by two independent vari-

[11]The fraction of the M/# we used—Mt_M/τ_e was more than 95%.

ables, namely the L_i/N_b ratio and ratio between the sum of the metals and oxygen. Many properties of the material, including thermal expansion/contraction, change of refraction index, are significantly affected by the stoichiometry.

For congruent $LiNbO_3$ crystals widely used in holographic memories, the temperature dependence of the ordinary and extraordinary refraction indices can be expressed in a similar way as the Sellmeier oscillator [138] which is used to describe their dispersion. The temperature-dependent Sellmeier equations are [139–141]

$$\begin{aligned} n_o^2 &= 4.9048 + \frac{1.178 \times 10^5 + 2.3416 \times 10^{-2} F}{\lambda_0^2 - (2.1802 \times 10^2 - 2.9671 \times 10^{-5} F)^2} \\ &\quad -2.7153 \times 10^{-8} \lambda_0^2 + 2.1429 \times 10^{-8} F, \end{aligned} \tag{5.31}$$

$$\begin{aligned} n_e^2 &= 4.5820 + \frac{0.9921 \times 10^5 + 5.2716 \times 10^{-2} F}{\lambda_0^2 - (2.1090 \times 10^2 - 4.9143 \times 10^{-5} F)^2} \\ &\quad -2.1940 \times 10^{-8} \lambda_0^2 + 2.2971 \times 10^{-7} F, \end{aligned} \tag{5.32}$$

$$F = (T - T_0)(T + T_0 + 546), \tag{5.33}$$

where λ_0 is the vacuum wavelength in nm. T is the temperature in Celsius. T_0 is a constant—24.5°C, the temperature of the original measurement in [140]. Evaluated at room temperature, these equations yield an n_0 of 2.3487, close to the value 2.3489 reported in [121]. In addition, the influence of impurity doping on the index of refraction has been studied [142], however, only under room temperature conditions.

Axial thermal expansion data have been reported for $LiNbO_3$ from X-ray diffraction [143–148], dilatometer [144,149–153], interferometer [154], and capacitance dilatometer [155,156] measurements. Gallagher et al. [149,150] observed that changes in stoichiometry only have a significant effect on the thermal expansion of the **c**–axis; the expansion along the other two axes (called **a**–axis) is virtually identical for stoichiometric and congruent materials. The general form of the expansion of $LiNbO_3$ is that the **a**–axis expands with an increasing rate as the temperature becomes higher, while the expansion of the **c**–axis reaches a maximum at about 925K and decreases thereafter. Taylor [157] summarized the experiment results from all the previous contributions and fit the data by the method of least squares to a third–order polynomial

Temperature range	a	b×10^6	c×10^9	d×10^{12}
60—373K	13.8637	4.09	4.40	-20.2
373—1373K	13.8633	4.40		-3.69

Table 5.2: The coefficients for **c**–axis expansion in congruent $LiNbO_3$.

Temperature range	a	b×10^6	c×10^9	d×10^{12}
220—523K	5.1490	13.43	17.55	-16.3

Table 5.3: The coefficients for **a**–axis expansion in congruent $LiNbO_3$.

as shown below.

$$x = a(1 + bT + cT^2 + dT^3), \tag{5.34}$$

where T is the temperature in Celsius. x is the mean cell size in Angstrom. The coefficients a, b, c, and d are listed in Table 5.2 and 5.3.

We evaluated Equation 5.34 based on the parameters in the two tables. The percentage contraction of the **c**–axis at room temperature ($27°C$) with respect to that at $150°C$ is found to be 0.050%. The percentage contraction along the **a**–axis is 0.20%. The thermal contraction on the **a**–axis is more significant. In transmission–geometry recording, this **a**–axis contraction has the same effect as the shrinkage problem in photopolymers. However, the latter seems to be much more serious. For example, the shrinkage is 3% in DuPont polymers [158].

5.3.2 Bragg shift and image distortion

The main effect of the change in the refraction index on the reconstruction of thermally fixed holograms recorded at high temperatures is the change in the incident angle, hence the phase–matching condition required for Bragg–matching. In general, the effect of index change could be compensated by adjusting the incident angle of the readout beam to re–match the hologram. Thermal contraction, on the other hand,

imposes a more serious problem. Since the expansion is not homogeneous, one can not solve the problem by simply adjusting the incident angle or the wavelength of the readout beam. Significant asymmetry (or simply **a**–axis contraction) in the axial expansion of the material will result in distortion in the reconstructed images [113,158].

Observed Bragg shift: In a special case of transmission–geometry recording where the two writing beams are symmetric about the normal of the entrance face on the crystal, the change of the magnitude of the grating vector and the incident angle cancel out each other. Therefore, the same incident angle used for recording is capable of Bragg–matching the hologram. Furthermore, when two plane–wave writing beams are used, the grating vector formed is parallel to the **c**–axis. As a result, the effects of index change and **a**–axis expansion are both eliminated. Only the **c**–axis contraction affects the phase–matching condition and Bragg–match can be satisfied by adjusting the incident angle of the readout beam. To see this, we summarized the shift in the selectivity curves from all the previous M/# measurements at $120°C$ and $150°C$ in Figure 5.19. Another measurement at $100°C$ is also included in the same figure. The solid line in the figure is derived from theory based on Equation 5.31, 5.33, and the phase–matching condition of the transmission–geometry hologram. The filled circles represent the measured Bragg shift in terms of frequency detuning in the RF frequency of the AOD. This figure shows a good agreement between the predicted and measured Bragg Shift. Most of the deviation and scattering is caused by the measurement of weak holograms and the errors in determining the incident angles.

Image distortion due to thermal contraction: When information–bearing signal beam is used, Bragg–matching condition is not satisfied with all the spatial frequencies in the signal spectrum due to thermal contraction. As a result, only a portion of the signal is Bragg–matched in Fourier–transform–plane recording, resulting in the distortion similar to that due to the dual–wavelength effect [159]. To evaluate the influence of thermal contraction on the image quality in the reconstruction, we chose a simplified case where the **c**–axis contraction was ignored. In Figure 5.20(a), the

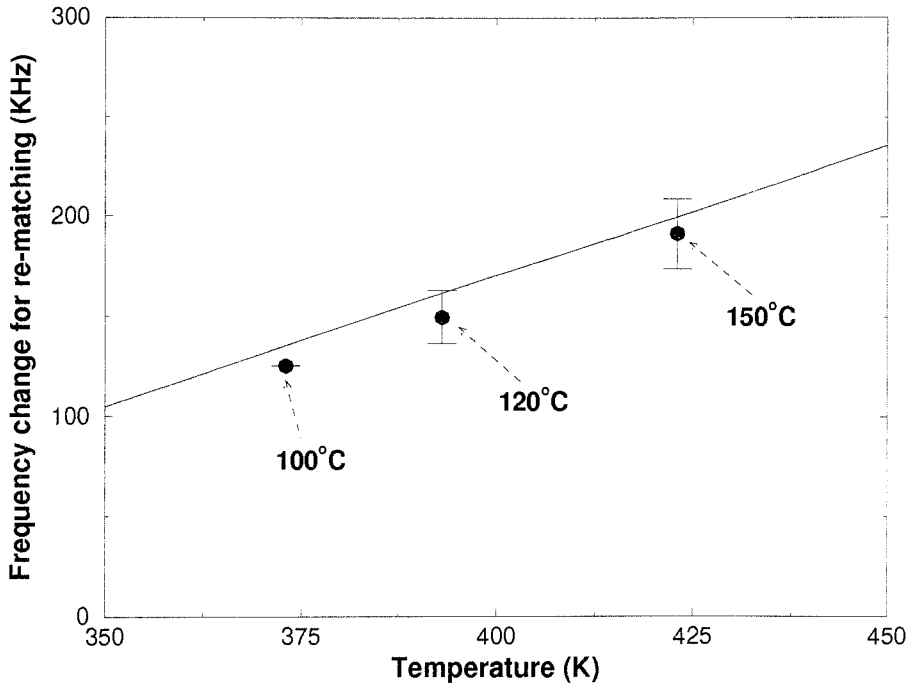

Figure 5.19: Bragg shift in the selectivity curves due to **c**–axis contraction.

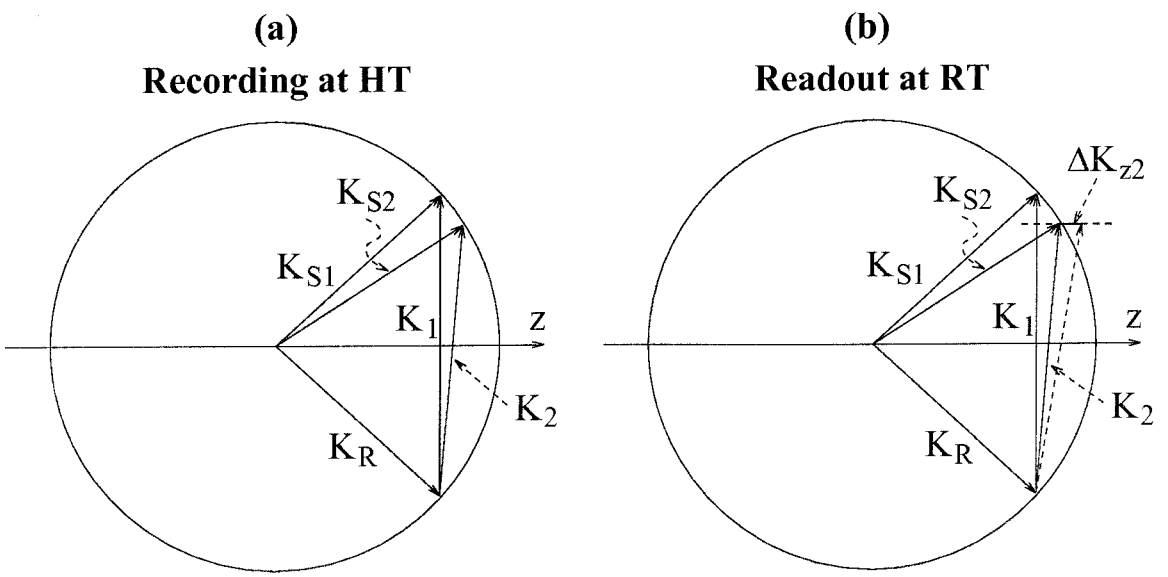

Figure 5.20: K–space description of phase–mismatching due to thermal contraction along the **a**-axis.

grating $\vec{K}_1$ formed between the center signal beam $\vec{K}_{S1}$ and the reference $\vec{K}_R$ is parallel to the **c**–axis. Therefore, it is not affected by the "shrinkage" problem due to **a**–axis (or in this case, z–axis) contraction. At the edge of the signal spectrum, $\vec{K}_{S2}$ writes a grating $\vec{K}_2$ with $\vec{K}_R$. This grating has a component K_{2z} along the z axis. After recording at elevated temperature, the hologram is revealed and readout at room temperature. Because of the thermal contraction, K_{2z} experiences a change which can be related to the percentage contraction $\Delta a/a$ in the z axis by

$$\Delta K_{2z} = -K_{2z}\frac{\Delta a}{a}, \tag{5.35}$$

where ΔK_{2z} determines the phase–matching condition for the reconstruction of $\vec{K}_{S2}$ as shown in Figure 20(b). For a quick assessment of phase–mismatching as a result of thermal contraction, we use the experimental parameters of the multiple–hologram system. The angle between $\vec{K}_R$ and $\vec{K}_{S1}$ is 90° for the 90° geometry. The angle between $\vec{K}_{S1}$ and $\vec{K}_{S2}$ is 1°, determined by the size of the SLM, the demagnifying ratio of the 4F imaging system, as well as the focal length of the FT lens in the signal arm. The percentage contraction in the z axis is 0.20%. From Equation 5.35, the phase–mismatching factor ΔK_{2z} is found to be 7.46cm^{-1}. This is equivalent to a half–null shift in the reference beam with an 5mm interaction length [12]. As a result, the edge of the reconstruction is about 30% weaker than the center. This is shown in Figure 5.21 where one 90°–geometry hologram written at 150°C is revealed and reconstructed at room temperature.

It should be noted that the degree of distortion due to thermal contraction mostly depends on the bandwidth of the signal (or the size of the image in Fourier–transform–plane recording), not on its complexity.

[12]The selectivity function $sinc(\frac{\Delta K_{2z} L}{2\pi}) \approx sinc(0.59) \approx 0.5$

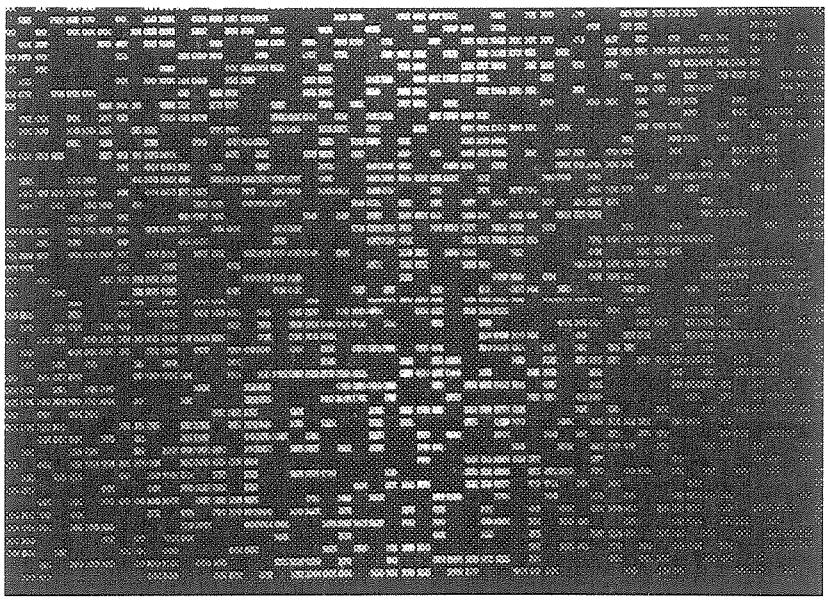

Figure 5.21: Reconstruction of one hologram written at 150°C.

5.4 Appendix A: Derivation of the extended Kukhtarev equations

The extended set of equations describing both the light–induced and thermal–induced charge transport are

$$\frac{\partial N_D^+}{\partial t} = SI(N_D - N_D^+) - \gamma_R N_D^+ n_e, \tag{5.36}$$

$$\frac{\partial n_e}{\partial t} = \frac{\partial N_D^+}{\partial t} + \frac{1}{q}\frac{\partial J_e}{\partial x}, \tag{5.37}$$

$$J_e = q\mu_e n_e E + K_B T \mu_e \frac{\partial n_e}{\partial x} + P(N_D - N_D^+)I, \tag{5.38}$$

$$\frac{\partial n_p}{\partial t} = -\frac{1}{q}\frac{\partial J_p}{\partial x}, \tag{5.39}$$

$$J_p = q\mu_p n_p E - K_B T \mu_p \frac{\partial n_p}{\partial x}, \tag{5.40}$$

$$\epsilon\frac{\partial E}{\partial x} = q(N_D^+ - N_A + n_p - n_{p0} - n_e). \tag{5.41}$$

By assuming that $m \ll 1$, these equations can be linearized and separated into two sets of equations for the DC components and first–order spatial perturbations of

all the variables. The DC response is listed below.

$$\frac{\partial N_{D0}^{+}}{\partial t} = SI_0(N_D - N_{D0}^{+}) - \gamma_R N_{D0}^{+} n_{e0}, \tag{5.42}$$

$$\frac{\partial n_{e0}}{\partial t} = \frac{\partial N_{D0}^{+}}{\partial t}, \tag{5.43}$$

$$J_{e0} = q\mu_e n_{e0} E_0 + P(N_D - N_{D0}^{+}) I_0, \tag{5.44}$$

$$\frac{\partial n_{p0}}{\partial t} = 0, \tag{5.45}$$

$$J_{p0} = q\mu_p n_{p0} E_0, \tag{5.46}$$

$$0 = N_{D0}^{+} - N_A - n_{e0}, \tag{5.47}$$

We also assume that $N_D, N_A \gg n_e$. Then from the DC Poisson's relationship, we have $N_{D0}^{+} = N_A$. From Equation 5.42, 5.43 and 5.44, the buildup of the DC free electrons is

$$\frac{1}{\gamma_R N_A}\frac{\partial n_{e0}}{\partial t} = -n_{e0} + \frac{SI_0(N_D - N_A)}{\gamma_R N_A}, \tag{5.48}$$

which is the same as in normal room-temperature recording. So it is still valid to make the quasi-steady state assumption to treat n_{e0} as a constant.

To solve the evolution of the DC field during recording, a set of boundary conditions have to be used. They are

$$\frac{\partial \rho}{\partial t} = -J, \tag{5.49}$$

$$E_0 = \frac{\rho}{\epsilon}, \tag{5.50}$$

which relate the current density to the accumulation of surface charge on the boundaries of the illuminated region. J is the combination of J_{e0} and J_{p0} in Equation 5.44 and 5.46. The dynmaics of DC field buildup is found to be

$$E_0 = -E_{0ss}(1 - e^{-t/\tau_{dc}}), \tag{5.51}$$

where

$$E_{0ss} = -\frac{1}{1+\sigma_p/\sigma_e}E_{0ph}, \tag{5.52}$$

$$\tau_{dc} = \frac{1}{1+\sigma_p/\sigma_e}\tau_{di}, \tag{5.53}$$

where σ_p, σ_e are the conductivity of the protons and electrons, respectively. τ_{di} is the dielectric time constant for the electrons.

The set of equations for the first–order spatial harmonics are

$$\frac{\partial N_{D1}^+}{\partial t} = SI_1(N_D - N_A) - SI_0N_{D1}^+ - \gamma_R(n_{e0}N_{D1}^+ + n_{e1}N_A), \tag{5.54}$$

$$\frac{\partial n_{e1}}{\partial t} = \frac{\partial N_{D1}^+}{\partial t} + j\frac{K}{q}J_{e1}, \tag{5.55}$$

$$J_{e1} = q\mu_e(n_{e0}E_1 + n_{e1}E_0) + jK_BTK\mu_e n_{e1} + P[(N_D - N_A)I_1 - N_{D1}^+I_0], \tag{5.56}$$

$$\frac{\partial n_{p1}}{\partial t} = -j\frac{K}{q}J_{p1}, \tag{5.57}$$

$$J_{p1} = q\mu_p(n_{p0}E_1 + n_{p1}E_0) - jK_BTK\mu_p n_{p1}, \tag{5.58}$$

$$E_1 = -j\frac{q}{K\epsilon}(N_{D1}^+ + n_{p1} - n_{e1}). \tag{5.59}$$

By assuming that $\partial n_{e1}/\partial t = 0$, Equation 5.54, 5.55 and 5.56 yield the steady–state value of the first–order free electron concentration.

$$n_{e1} = \frac{m(N_D - N_A)(SI_0 + jPI_0K/q) - (SI_0 + \gamma_R n_{e0} + jPI_0K/q) + jK\mu_e E_1 n_{e0}}{\gamma_R N_A + \mu_e K E_d - j\mu_e K E_0}. \tag{5.60}$$

The rate of change in the electronic grating is obtained by substituting the expression of n_{e1} in Equation 5.60 into Equation 5.54.

$$\frac{\partial N_{D1}^+}{\partial t} = -j\mu_e K n_{e0} \times \frac{m\left[(E_0 + E_{0ph}) + jE_d\right] - R\left[(E_0 + \frac{N_A}{N_D}E_{0ph}) + jE_d\right]N_{D1}^+ + E_1}{1 + \frac{E_d}{E_{ue}} - j\frac{E_0}{E_{ue}}} \tag{5.61}$$

The rate of change in the protonic grating is easily obtained by combining Equation 5.57 and 5.58.

$$\frac{\partial n_{p1}}{\partial t} = -j\mu_p K n_{p0}\left[(E_0 - jE_d)\frac{n_{p1}}{n_{p0}} + E_1\right]. \tag{5.62}$$

The above two rate equations are coupled by the first-order Poisson's relationship

$$E_1 = -j\frac{q}{\epsilon K}(N_{D1}^+ + n_{p1}), \tag{5.63}$$

which is obtained from Equation 5.59 by simply throwing away n_{e1} which is orders of magnitudes smaller.

5.5 Appendix B: Derivation of the simplified model

When it is valid to assume that complete compensation of the electronic charge grating by the protons takes place instantaneously, the electric field inside the material is zero. So the first-order perturbations in the ionized donor and proton concentration are decoupled. The equations describing the formation of electronic charge grating N_{D1}^+ become

$$\frac{\partial N_{D1}^+}{\partial t} = SI_1(N_D - N_A) - SI_0 N_{D1}^+ - \gamma_R(n_{e0}N_{D1}^+ + n_{e1}N_A), \tag{5.64}$$

$$\frac{\partial n_{e1}}{\partial t} = \frac{\partial N_{D1}^+}{\partial t} + j\frac{K}{q}J_{e1}, \tag{5.65}$$

$$J_{e1} = jK_B T K \mu_e n_{e1} + P(N_D - N_A)I_1 - PN_{D1}^+ I_0. \tag{5.66}$$

To solve this set of simplified equations, we still start with the first-order free electron concentration using the quasi-state assumption—$\partial n_e/\partial t = 0$.

$$n_{e1} = \frac{1}{\gamma_R N_A}\left[-\frac{\partial N_{D1}^+}{\partial t} + SI_1(N_D - N_A) - SI_0\frac{N_D}{N_A}N_{D1}^+\right]. \tag{5.67}$$

Substituting this expression of n_{e1} in Equation 5.64, the dynamics of N_{D1}^{+} is found as

$$\tau_{di}\frac{1+E_d/E_u}{E_d/E_q - jN_A E_{0ph}/N_D E_q}\frac{\partial N_{D1}^{+}}{\partial t} = -N_{D1}^{+} + m\frac{\epsilon K}{e}\frac{E_d - jE_{0ph}}{E_d/E_q - jN_A E_{0ph}/N_D E_q}. \tag{5.68}$$

Chapter 6 Inter–pixel grating noise

Contents

6.1 What is inter–pixel grating noise? **162**

6.2 Why is inter–pixel grating noise important? **164**

6.3 Temporal development of inter–pixel gratings **168**

6.4 Formation of inter–pixel grating noise **175**

6.4.1 Bragg degenerate "connections" 175

6.4.2 Strength of inter–pixel gratings 180

6.4.3 Bragg selectivity 186

6.5 Random–phase modulation in the signal **190**

6.6 Experiment results **191**

6.6.1 Experiment with plane–wave signal 192

6.6.2 Experiment with random–bit patterns 197

Like in any other storage technology, error performance is a very important issue in the application of large–scale holographic memories. The noise sources that affect the error performance of a large–scale holographic memory could be divided into two classes:

- Holographic–recording–induced noise: These noise sources include crosstalk that arises from imperfect Bragg–matching condition of complex signals [60–67], band–pass effects due to nonuniform exposure intensity profiles, two–beam coupling effects caused by the interaction between the writing beams and their strong scattering off the holographic gratings and impurity sites inside the storage material (the latter is also called fanning), inter–pixel grating noise, as well as those from non–uniform recording and erasure due to the absorption in the material [41, 72].

- Intrinsic noise—that does not depend on holographic recording. These include burst errors due to material surface defect and dust particles on the optical components, laser power fluctuation, phase variations due to the instability of the optical setup, non-uniformity of the Spatial Light Modulator (SLM), the original bulk scattering of the recording material, and the inter-pixel noise which is the crosstalk between pixels due to the point-spread function of the optical systems [160, 161].

The second class of noise sources is almost deterministic. They can be eliminated by careful engineering of the optical system or greatly suppressed by "post-recording" processing such as illumination profile compensation described in Chapter 2 and in [72, 161]. On the other hand, most of the holographic-recording-induced noise sources in the first class are truly random noise. They depend on the recording sequence, exposure times, 2-D data structure in the images, and/or the position of the holograms. Among them, crosstalk noise is mostly local—significant crosstalk seen by a specific hologram comes from its neighborhood. Strongest crosstalk occurs when the holograms are separated by the first-null angle (Bragg selectivity) determined by the center spatial frequency of the signal beam and the reference beam. To avoid the influence of crosstalk, one can always space the holograms further apart. For example, the inter-hologram spacing we chose in our large-scale holographic memory was almost five times wider than the Bragg selectivity. Therefore, we did not observe significant crosstalk. Two-beam coupling effects involving the reference beam only occurs at the beginning of the recording as a result of the interaction between the writing beams and their strong diffractions. So it is also a local effect and can be eliminated by simply throwing away the first several "bad" holograms, or reducing the exposure times. Because the exposure times for the holograms in the recording sequence decreases exponentially, a slight change in the last exposure time (which determines the diffraction efficiency of the equalized holograms) will result in significant reduction of the first exposure time. Fanning noise, which arises from the interaction between the main writing beams and their scattering, is very hard to suppress unless the recording materials are of high optical quality.

In this chapter, we investigate the formation and influence of the inter–pixel grating noise. In Section 6.1—6.4, we present theoretical treatment to describe its origin and effect. In the reminder of this chapter, we experimentally demonstrate the buildup and influence of the inter–pixel grating noise. We propose and demonstrate the use of random–phase modulation to suppress the inter–pixel grating noise.

6.1 What is inter–pixel grating noise?

Inter–pixel grating noise is a very important yet largely ignored form of holographic noise. It is a secondary effect—the secondary diffraction (rediffraction) of the diffracted signal upon reconstruction, via the gratings formed between the multiple plane–wave (spatial–frequency) components of the signal. Inter–pixel grating noise can be considered as crosstalk noise between the pixels (bits) within a page of information, similar to a class of higher–order crosstalk noise in volume holographic interconnections [162]. This crosstalk occurs directly between the pixels when Fourier–transform–plane recording is used where individual pixels on the Spatial Light Modulator (SLM) are converted into plane waves inside the storage medium. During holographic recording, these plane waves interfere with the reference beam to form the "useful" information hologram. In the meantime, they interfere with each other to create noise gratings. Upon readout, the same reference beam is used to reconstruct the plane waves, which are converted back to the corresponding pixels for data retrieval by imaging optics. In addition, these reconstructed plane waves give rise to secondary diffraction (rediffraction) via the noise gratings, resulting in crosstalk noise. This is the reason that we name this noise source inter–pixel grating noise [1].

Figure 6.1 shows a simple example of the occurrence of the inter–pixel grating noise. $\vec{K}_i$ and $\vec{K}_j$ are the wave–vectors of two plane waves corresponding to two pixels i, j in the image. $\vec{K}_r$ is the reference beam wave–vector. The gratings $\vec{K}_{ri}$ and $\vec{K}_{rj}$ are the "useful" ones, responsible for information storage and retrieval. Upon

[1] A similar term—"inter–pixel noise" is used to indicate the crosstalk among the pixels due to their finite size and limited optical aperture of the system [160, 161].

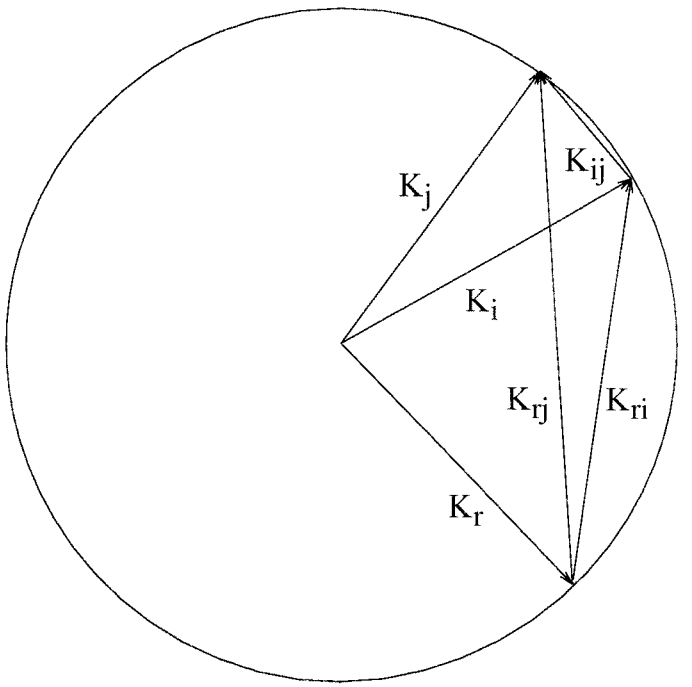

Figure 6.1: K–space description of inter–pixel grating noise.

readout, the "output" pixel j receives light from the reference via grating $\vec{K}_{rj}$. In addition, it receives light from pixel i through the intermediate step of diffraction by grating $\vec{K}_{ri}$ followed by rediffraction via grating $\vec{K}_{ij}$. This additional light is the inter–pixel grating noise.

It should be noted that the inter–pixel gratings noise exists, though in different ways, when other recording configuration such as Fresnel–plane or image–plane recording is used. In such occasions, it arises from the crosstalk between the spatial–frequency components of the signal.

6.2 Why is inter–pixel grating noise important?

Inter–pixel grating noise is often neglected in the literature since the noise gratings normally can not be written effectively because of two reasons: their large grating periods and small modulation depth. The first is the result of the small angles between the plane–wave components. The second one is due to the simple fact that the power a pixel carries is orders of magnitudes less than that of the reference beam. However,

in a large–scale holographic memory using $LiNbO_3$:Fe crystals, the inter–pixel grating noise is indeed a very important noise source because of

- the presence of strong photovoltaic field in photovoltaic materials like $LiNbO_3$:Fe, which enables effective holographic recording of gratings even with very large periods;
- the prolonged exposure sequence for multiple hologram recording which allows the same inter–pixel grating to be strengthened many times;
- high degree of Bragg degeneracy and weak Bragg selectivity among the pixels because of the small angles between their corresponding plane–wave components.

As a result, the influence of the inter–pixel grating noise should become stronger as the total exposure time or the total number of holograms stored increases. In addition, one must note that the inter–pixel grating noise is a global effect—any hologram shall see the same level of noise despite of its temporal position in the recording sequence and its physical address in the memory. Nonetheless, the amount of noise that an individual pixel sees also depends on its spatial location in the image and its neighborhood.

We conducted a series of experiments to characterize the system error performance in an attempt to track down the effect of individual noise sources (for similar demonstrations, refer to Chapter 2 and 4). We analyzed the SNR of the reconstruction from a single, 100, 300, 1,000, 5,000, and 10,000 holograms stored in the same volume inside a $LiNbO_3$:Fe crystal by angle and fractal multiplexing. All the holograms are recorded to the same strength using the same exposure schedule with the same last exposure time. To avoid significant buildup of two–beam coupling effects (including fanning) at the beginning due to long exposures, we did not use the full dynamic range of the storage system [2]. The inter–hologram spacing was set at five times wider than the measured Bragg selectivity to reduce the crosstalk noise.

[2]The last exposure time 0.48 seconds $< \tau_e/10000 = 0.6$ seconds.

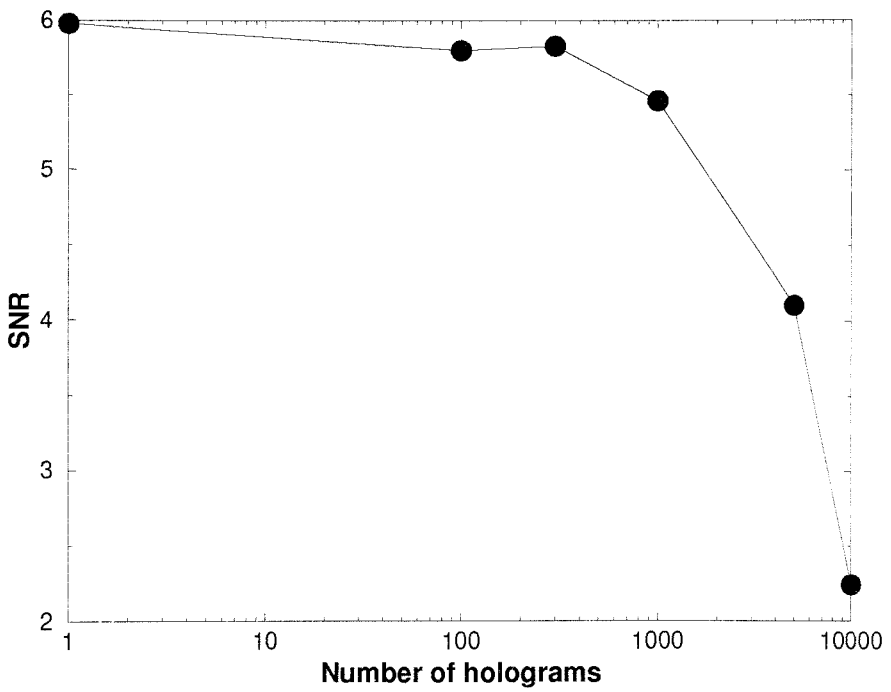

Figure 6.2: SNR degradation as a function of the number of holograms.

The measured SNR as a function of the number of holograms stored is plotted in Figure 6.2. It shows that the SNR initially remains nearly constant as the number of holograms stored increases and drops only after a large number of holograms are recorded. This phenomenon suggests that "inter-hologram" crosstalk noise is very unlikely to be responsible for the degradation of system error performance since its effect is mainly local. The crosstalk noise shall initially buildup as the number of holograms becomes larger and reach a "plateau" when holograms further recorded are too far away to contribute noticeable noise [3].

The significant drop in the SNR as the number of holograms becomes larger indicates that the dominant noise source depends strongly on the total exposure time. We notice that any noise source that depends on the exposure time during holographic recording is associated with the buildup of noise gratings between various sources—the writing beams, their scattering off the holographic gratings and the impurity centers inside the material. There are several noise sources falling into this category.

[3]The drop in the SNR of reconstructions from a few hundred holograms should be closely related to the crosstalk.

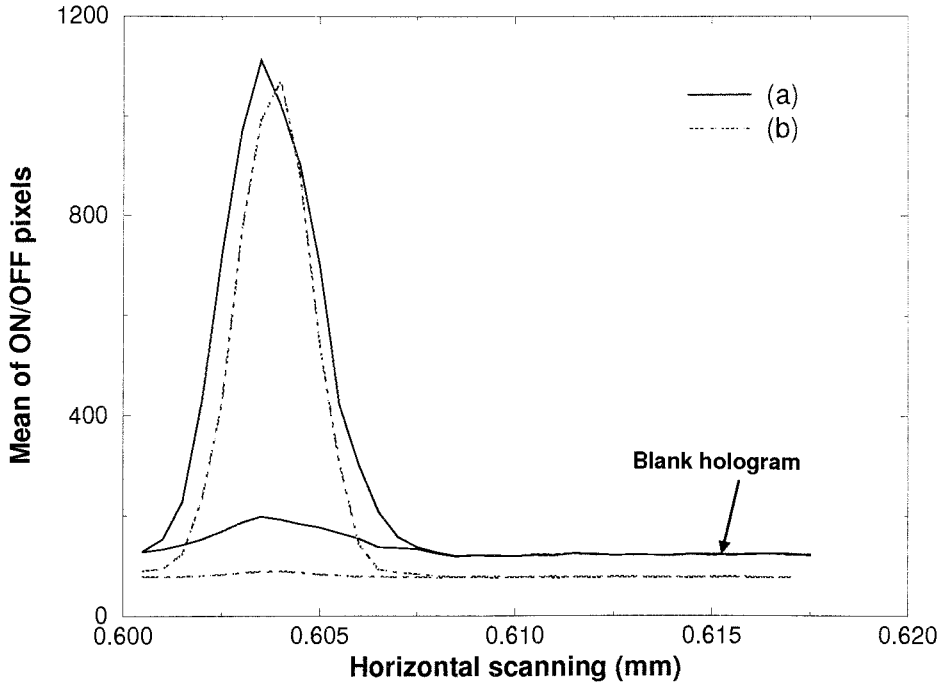

Figure 6.3: Mean of ON/OFF pixels as a function of reference beam scanning: (a) a "blank" hologram and its neighbor from 10,000 stored holograms; (b) a single hologram.

Among them, some are related to the reference beam. They include

- Two–beam coupling effect between the reference and signal beam. This effect is local—it only affects the first several holograms. It was greatly suppressed in our experiments by reducing the exposure times.

- Fanning which arises from the interaction between the reference beam and its scattering. It results in the enhanced scattering due to prolonged reference beam illumination. This noise source should not be very prominent because of 90°–geometry recording and continuous scanning of the reference beam for multiplexing.

Others are related to the illumination of the signal:

- Inter–pixel grating noise formed between the plane–wave components of the signal.

- Fanning which is the result of the interaction between the signal and its scattering. This noise is much weaker than the inter–pixel grating noise because the bulk scattering of a fresh crystal is at least 10 times weaker than a normal hologram (at pixel level). Therefore, the modulation depth responsible for writing this type of noise gratings shall be much smaller than that of the inter–pixel gratings.

Since fanning noise is the enhanced scattering which affects the entire background, it is easy to differentiate it from the inter–pixel grating noise which tends to pinpoint its effect directly on the reconstructions [4]. In our experiments, we always left a "blank" hologram in the recording sequence (the signal beam was not turned on during recording). During reconstruction, we scanned the reference beam across a range that covers this "blank" hologram and its neighbors and took images at each increment of the scanning. To process the data, we first located the ON and OFF signals in the Bragg–matched reconstruction of the neighboring hologram and took their average values. We then computed the average values of the pixels falling in the same ON/OFF locations from every reconstruction and obtained two curves as a function of the angle change of the reference beam. The result from the 10,000 hologram experiment is shown in Figure 6.3(a). Also in the same figure, we show the result from a single stored hologram (which is very "clean") for comparison (Figure 6.3(b)). Ideally, the ON curve reflects the Bragg selectivity, while the OFF curve is a flat floor. However, it is interesting to notice that, even with only one hologram stored, the OFF curve at the Bragg–matching location is slightly higher. We attribute this to the limited contrast ratio of the SLM, and the "inter–pixel noise" due to the point–spread function. After 10,000 holograms were stored, the background (the part of the curves where the "blank" hologram is) rises slightly (by $< 50\%$). In contrast, the OFF signal at the Bragg–matched location increases by a factor of more than 2. This suggests a significant contribution from another noise source—the inter–pixel grating noise.

[4]Here, the "background" refers to the reference beam angles that do not Bragg–match any of the stored holograms.

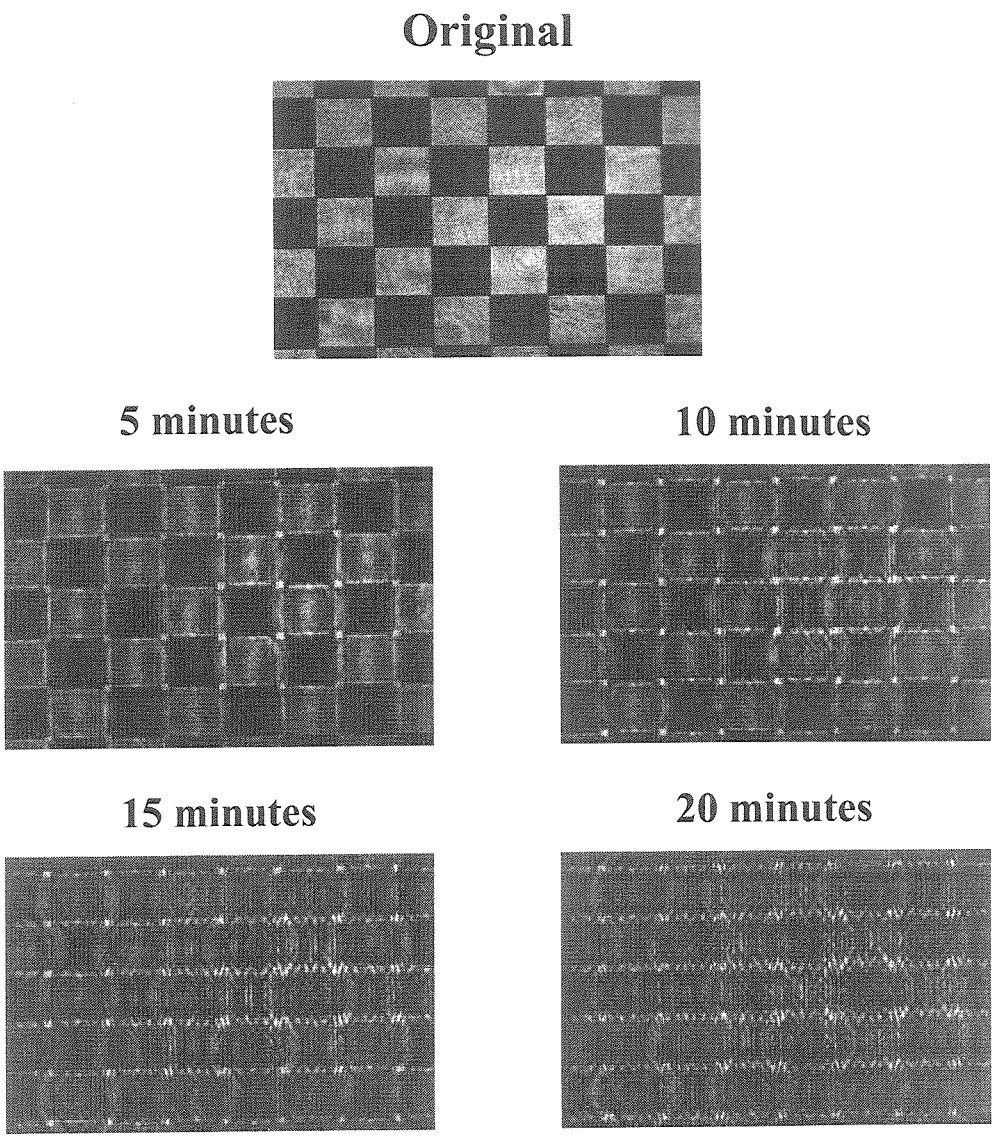

Figure 6.4: Strong two–beam coupling effect via the inter–pixel gratings.

It is important to note that the inter–pixel gratings can buildup very quickly to be strong enough to cause two–beam coupling effect of their own. An example is shown in Figure 6.4. A fixed pattern signal with about 18mw/cm^2 intensity continuously illuminated the crystal. The reference beam was not used. Therefore, there were no holograms recorded except for those within the signal. After a few minutes, dark lines appear in the bright regions, indicating a strong non–linear effect which depletes part of the pixels via the strong inter–pixel gratings. The reason for the exchange of energy being along the horizontal direction (**c**–axis) is that only the gratings parallel or at small angles to the **c**–axis grow much faster since the photovoltaic effect is along the **c**–axis. Shortly after, the dark areas become more noisy as part of the energy is spread over from the bright pixels. This is the direct result of strong gratings written by the bright pixels and the dark ones. And all of this happened within minutes!

In conclusion, the above experimental results provide evidence showing that the inter–pixel grating noise is a very significant noise source in the application of a large–scale holographic memory. This explains why we are interested in the inter–pixel grating noise and justifies further investigation on this noise source.

6.3 Temporal development of inter–pixel gratings

In this section, we describe the time evolution of the inter–pixel gratings. The questions that we would like to address here are:

- How strong is an inter–pixel grating at the end of multiple hologram recording?
- How does the number of holograms affect the strength of the inter–pixel gratings?

In answering these questions, we would treat the buildup of an inter–pixel grating as a linear process, ignoring the non–linear effects that are also likely to happen.

Normally, the period of a grating formed in between two plane–wave components of the signal is very large. Therefore, diffusion of the excited electrons is not sufficient to separate the charges to form a holographic grating effectively. However, in $LiNbO_3$:Fe

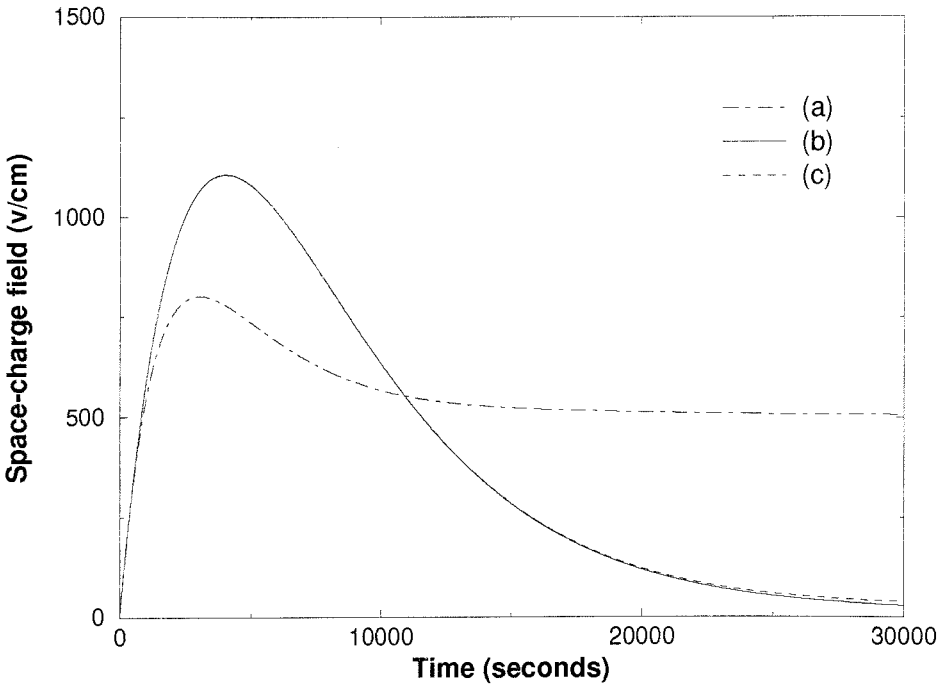

Figure 6.5: Recording of a 90°–geometry hologram and inter–pixel gratings (a) 90°–geometry, (b) transmission–geometry $\theta_{out} = 0.1°$, (c) transmission–geometry $\theta_{out} = 1°$.

or any other photovoltaic material, the presence of a strong photovoltaic effect alone is capable of generating modulated photovoltaic current to effectively transport the free electrons to form the space–charge grating. The time evolution in holographic recording of a 90°–geometry "information" hologram and two inter–pixel gratings with the same modulation depth are shown in Figure 6.5. The angles between the two plane waves writing the inter–pixel gratings are 0.1° and 1° (outside the medium), respectively. It shows that the buildup of the inter–pixel gratings is as fast as the "information" hologram. We also learn from the figure that the saturation values of the inter–pixel gratings are much less than that of the 90°–geometry hologram. This is because of the screening of the photovoltaic effect by the DC field. As a result, if the total exposure for the recording of multiple holograms is long enough for the DC field to fully buildup, the inter–pixel gratings can be greatly suppressed. However, the buildup of the DC field is usually very slow (whose characteristic time constant is τ_{di}). It would take much more exposure for significant suppression in the inter–pixel

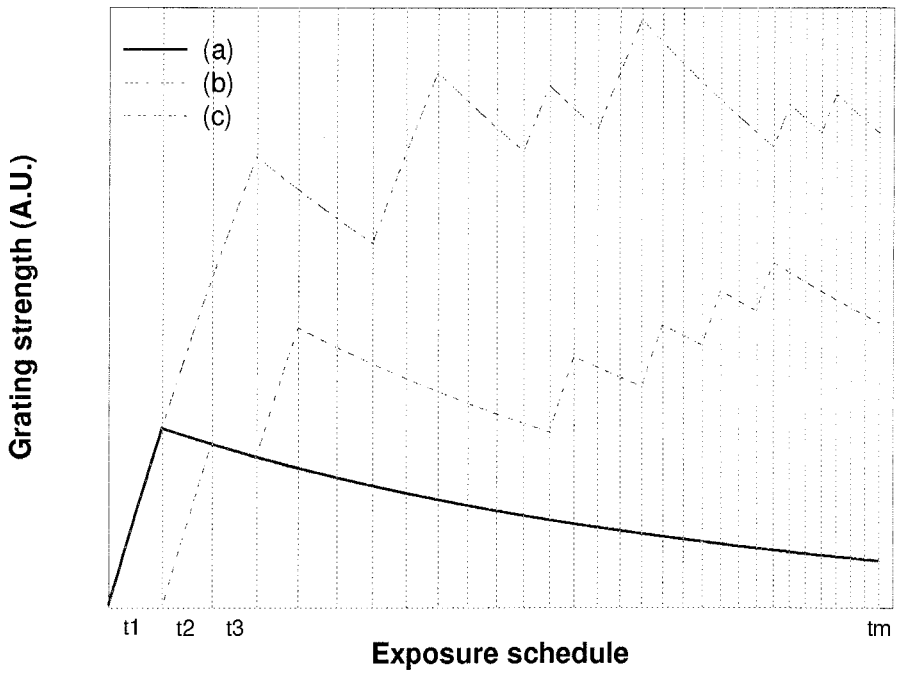

Figure 6.6: Different recording and erasure behaviors of (a) an information grating and (b)(c) inter–pixel gratings during multiple hologram recording.

gratings to take place than what is needed for hologram recording [5]. Alternatively, "pre–exposure" by UV light can be employed to enable the DC field to buildup before recording. By doing this, the inter–pixel grating noise can be reduced, however, at the expense of the system dynamic range (Section 2.8).

For multiple hologram recording in $LiNbO_3$, an exposure schedule is used to equalize the strength of all the holograms. Given the erasure time constant of the holographic memory system, the exposure time of an individual hologram in the recording sequence and the total exposure time are determined by the exposure schedule. For one "useful" (information) grating in a hologram (e.g., $\vec{K}_{rj}$ in Figure 6.1), it is recorded for a certain amount of time according to the schedule, and then erased during the recording of the rest of the holograms with different reference $\vec{K}_r$'s. This is shown in Figure 6.6(a). To the contrary, an inter–pixel grating formed by two plane–wave components (e.g., $\vec{K}_{ij}$ in Figure 6.1) is recorded for many times, whenever both pixels i, j are ON. If we assume that the images to be stored in the memory

[5] For the recording of 10,000 holograms, it only takes about 8,000 seconds (Chapter 2).

are uncorrelated pages of binary data, the possibilities of a pixel to be either ON or OFF are the same across the entire image and from image to image. Therefore, for holographic recording of M holograms, one inter–pixel grating is strengthened repeatedly, for $M/4$ times on average (Figure 6.6(b)(c)). As a result, the final diffraction efficiencies of the inter–pixel gratings strongly depend on the total number of holograms and can be conceivably large compared to that of one information grating.

The strength of an inter–pixel grating also depends on the number of pixels in the image. The power of an individual pixel is inversely proportional to the total number of pixels in an image. Therefore, the modulation depth of an inter–pixel grating decreases as the number of pixels becomes larger. We define the modulation depth for the information grating $\vec{K}_{ri}$ and the inter–pixel grating $\vec{K}_{ij}$ to be m_{ri} and m_{ij}.

$$m_{ri} = \frac{|A_i||A_r|}{I_r + I_s}, \tag{6.1}$$

$$m_{ij} = \frac{|A_i||A_j|}{I_r + I_s}, \tag{6.2}$$

where A_i, A_j, A_r are the amplitudes of the plane–wave components in the signal, and the reference beam. I_r and I_s are the intensities of the reference and the signal beam, respectively. By assuming that the number of pixels on the SLM is $N \times N$ and half of them are ON at the same time in a random–bit pattern, I_s can be written as

$$I_s \equiv \sum_{i'=1}^{N^2} I_s^{i'} = \frac{N^2}{2} I_s^I, \tag{6.3}$$

where $I_s^i = |A_i|^2$ is the intensity of a pixel. Under further assumptions that the intensities of the signal and reference beam are the same ($I_s = I_r$), we have (when pixel i and j are ON),

$$|A_i| = |A_j| = \frac{\sqrt{2}}{N}|A_r|. \tag{6.4}$$

Combining Equation 6.1, 6.2 and 6.4, the modulation depths of the information ("useful") and the inter–pixel gratings can be related by the following expression:

$$\frac{m_{ij}}{m_{ri}} = \frac{\sqrt{2}}{N}. \tag{6.5}$$

Normally, N, the number of pixels on the SLM in one dimension, is very large. Therefore, the modulation depth of an inter–pixel grating is much less than that of an information grating.

In general, the time evolution of the information and the inter–pixel grating can be described by the following equations.

$$\tau^{info}(t)\frac{\partial E_1^{info}}{\partial t} = -E_1^{info}(t) + E_\infty^{info}, \tag{6.6}$$

$$\tau^{ip}(t)\frac{\partial E_1^{ip}}{\partial t} = -E_1^{ip}(t) + E_\infty^{ip}, \tag{6.7}$$

where E_1^{info} and E_1^{ip} are the space–charge fields of the information and inter–pixel grating, respectively. They change with their own complex time constants τ^{info}, τ^{ip}, and saturation values $E_\infty^{info}, E_\infty^{ip}$. During recording, the DC field also builds up, following

$$E_0 = -E_{0ph}(1 - e^{-t/\tau_{di}}), \tag{6.8}$$

where τ_{di} is the dielectric time constant.

Based on Equation 6.6—6.8, we conducted numerical evaluation on the temporal development of the two types of gratings using the 4^{th}–order Rouge–Kutta method. The parameters we used in the evaluation are the same as in Chapter 4. An exposure schedule was used to simulate multiple hologram recording. The strength of an information grating $\vec{K}_{ri}$ or $\vec{K}_{rj}$ at the end of the recording sequence was determined by the exposure schedule from Equation 6.6 and 6.8. The same schedule (the sequence of exposures) was applied to the evolution of the inter–pixel grating $\vec{K}_{ij}$ as well. We assume that the pixels have equal possibilities of being 0 (OFF) or 1 (ON). At the

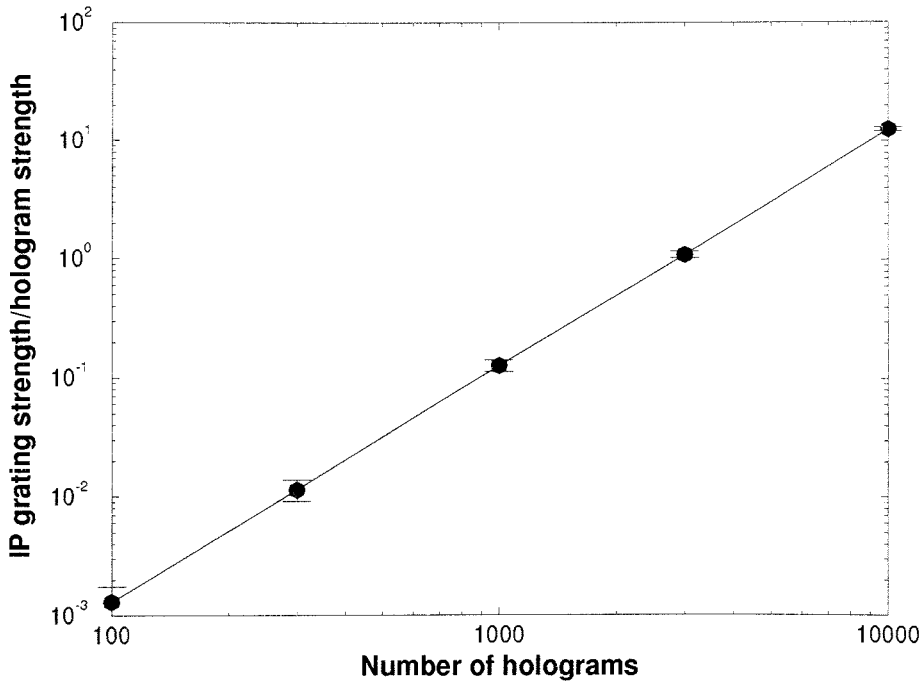

Figure 6.7: Ratio of the diffraction efficiencies of the inter–pixel grating $\vec{K}_{ij}$ and the information grating $\vec{K}_{ri}$ as a function of the number of holograms recorded with the same schedule.

beginning of an exposure, two independent binary random–sequence generators were used to simulate the states of the two pixels i, j, determining whether the inter–pixel grating $\vec{K}_{ij}$ should be strengthened or erased. When i, j are both ON, the modulation depth determined by Equation 6.5 was used to simulate recording. Otherwise, a zero modulation depth was assigned in the simulation for uniform erasure. At the end of the simulation, the ratio of the diffraction efficiencies of the inter–pixel grating $\vec{K}_{ij}$ and the information grating $\vec{K}_{ri}$ was calculated. We simulated the recording of 100, 300, 1,000, 3,000, and 10,000 holograms assuming $N = 1000$ and the angle between $\vec{K}_i$ and $\vec{K}_j$ is 0.5°. The results are summarized in Figure 6.7.

We learn from the figure that, although the modulation depth of an inter–pixel grating is orders of magnitude smaller than that of an information grating, the prolonged, repeated strengthening during multiple hologram recording results in a grating which is much stronger.

6.4 Formation of inter–pixel grating noise

The above demonstration provides the temporal picture of how an inter–pixel grating is formed during the recording of multiple holograms. It shows that an inter–pixel grating can be much stronger than an information grating after a large number of holograms are recorded. However, since the inter–pixel grating noise is a secondary effect (the noise comes from the rediffraction), the crosstalk noise at a pixel from another one via the "linking" inter–pixel grating is orders of magnitudes less than the signal (inversely proportional to the diffraction efficiency of the inter–pixel grating).

What makes the inter–pixel grating noise a significant noise source is that it arises from the collective behavior of multiple pixels in the image via multiple noise gratings. A pixel writes inter–pixel gratings with all the other $(N^2 - 1)$ pixels in the same image during recording. Upon readout, it receives crosstalk noise from the other $(N^2 - 1)$ pixels via the "direct connections." In addition, crosstalk between pixels can take place via "indirect connections" that are the result of Bragg degeneracy and weak Bragg selectivity. In this section, we discuss this collective effect in detail. We describe how the pixels are "connected" and how well they are "connected."

6.4.1 Bragg degenerate "connections"

For Fourier–transform–plane recording as shown in Figure 6.8, the crystal is placed at the focal plane of the Fourier–transform lens. The pixels on the SLM are converted into plane waves by this lens. The plane waves form a "signal" cone on the K–sphere. Normally, the size of the SLM is much smaller than the focal length of the FT lens. Therefore, by paraxial approximation, this cone on the surface of the K–sphere becomes a plane as shown in Figure 6.9. As a result, the inter–pixel gratings formed in between the plane waves all lie in this plane (with no z–components). When all the inter–pixel gratings are on the same plane, the connections between the pixels can only be either direct, or degenerate. In other words, two pixels only interfere with each other via the grating they form directly and other degenerate gratings with the same grating vector (same orientation and magnitude).

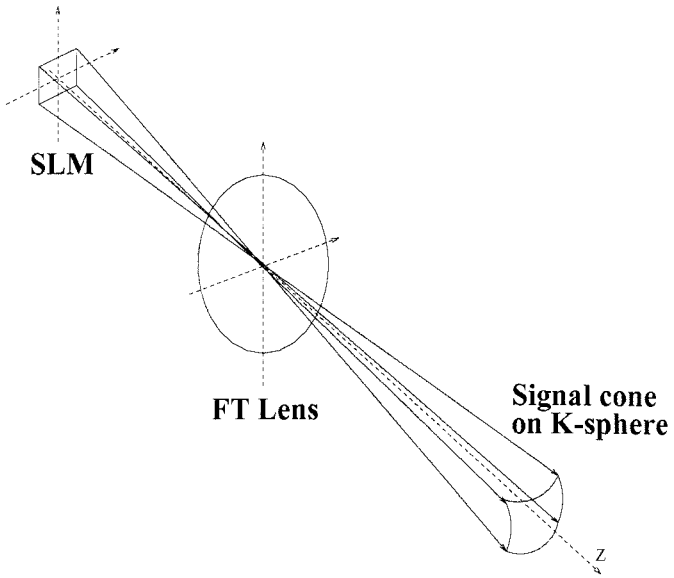

Figure 6.8: Signal cone in Fourier–transform–plane recording.

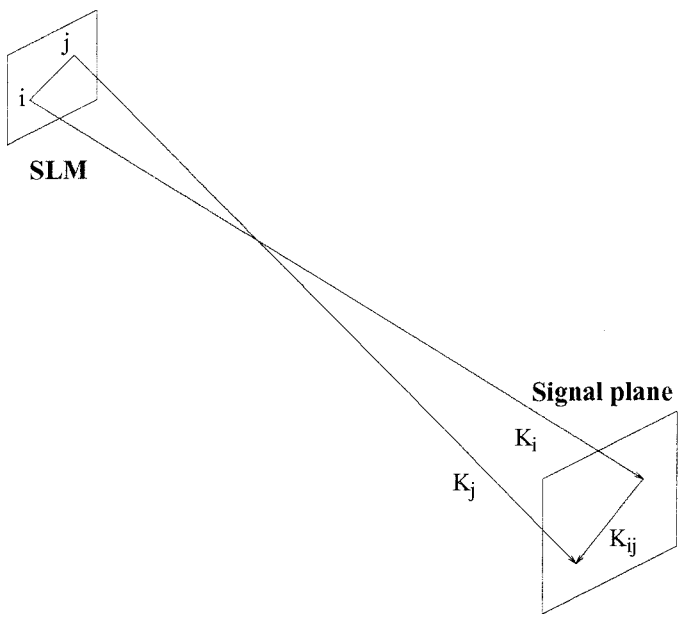

Figure 6.9: Signal plane in Fourier–transform–plane recording.

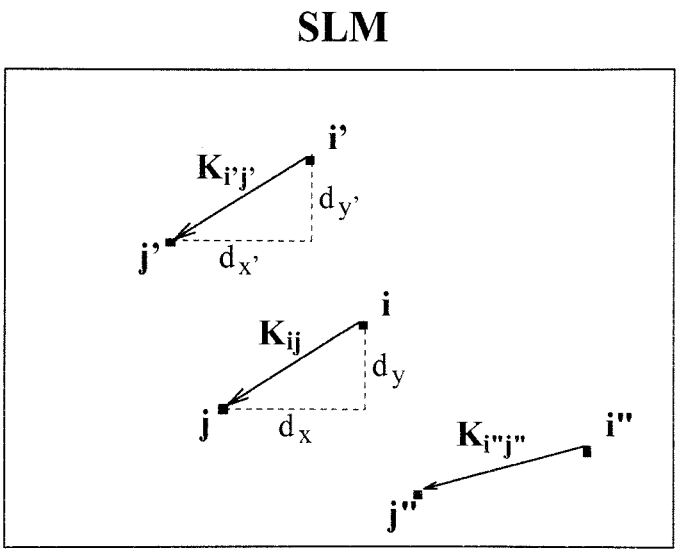

Figure 6.10: Links inducing crosstalk between a pair of pixels separated by horizontal and vertical distances d_x, d_y.

Figure 6.9 shows that the signal plane formed by the tips of all the plane–wave components of the signal can be treated as the "image" of the SLM (from "pixel space" to "grating space"). The grating vector of one inter–pixel grating in the signal plane can be determined by the relative position of the two pixels on the SLM that write this grating. As a result, the analysis of the "connection" between the pixels can be carried out directly on the SLM plane. Figure 6.10 shows the SLM plane where there are three "links" [6] between three pairs of pixels. The relative position of pixel j with respect to pixel i is the same as that of j' with respect to i':

$$d_x = d_{x'}\,;\;\; d_y = d_{y'}.$$

Therefore, the links $\vec{K}_{ij}$ and $\vec{K}_{i'j'}$ are degenerate. As a result, crosstalk occurs between pixel i and j via the direct link $\vec{K}_{ij}$, as well as the indirect "link" $\vec{K}_{i'j'}$.

To find a group of "indirect" (degenerate) links that a pair of pixels can use to interfere with each other, only the horizontal and vertical distances between them, d_x and d_y, are needed [7]. To assess the number of links in this group, we need to compute

[6] We prefer to use "links" instead of "gratings" because this is not the signal plane where the inter–pixel grating vectors lie. However, they represent the same concept in our analysis.

[7] All the pairs of pixels that are separated by the same horizontal and vertical distances can utilize

the number of all the pairs of pixels with the same horizontal and vertical distances. This can be done in three steps: First, the number of pairs of columns on the SLM that are separated horizontally by a distance d_x is given by

$$N_c = \sum_{i=0}^{d_x-1} floor(\frac{N-i-1}{d_x}) \approx N - \frac{d_x+1}{2}, \tag{6.9}$$

where function "floor" returns the integer value of its variable. In the second step, the number of pairs of pixels separated vertically by a distance d_y on a pair of columns with distance d_x is found to be

$$N_p = N - d_y. \tag{6.10}$$

The total number of links in the group specified by (d_x, d_y) is then

$$N_g = N_c N_p, \tag{6.11}$$

where N_g indicates the number of degenerate links any pair of pixels separated by (d_x, d_y) can use to interact with each other. A pixel at location (x, y) on the SLM receives crosstalk from another one at either of the four locations $(x \pm d_x, y \pm d_y)$ via the N_g links in this group.

The total number of links from or to a pixel at (x, y) is the summation of all the links (direct and indirect) between it and all the other pixels in the same image. It can be written as

$$N_{total}(x,y) = \sum_{\substack{x'=0 \\ x' \neq x}}^{N-1} \sum_{\substack{y'=0 \\ y' \neq y}}^{N-1} (N - \frac{|x-x'|+1}{2})(N - |y-y'|), \tag{6.12}$$

where $N_{total}(x, y)$ provides the information about how well a pixel is "connected" to others, in which each "link" accounts for an increment in the crosstalk noise the pixel sees. Therefore, it shows how "noisy" it is at the pixel at (x, y).

The total number of links of each pixel in an image consisting of 100× 100 pixels

this same group of degenerate links to interfere with each other.

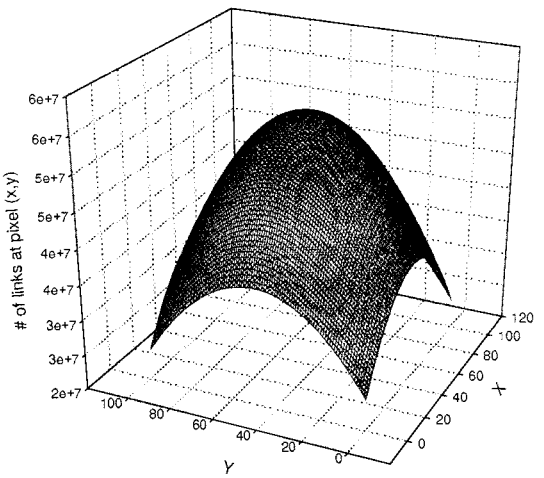

Figure 6.11: The number of links of every pixel in a 100 × 100 pixel image.

was computed and the result is summarized in Figure 6.11. If we assume that the strength of the inter–pixel gratings are the same regardless of their orientations and magnitudes, the amount of crosstalk noise at each pixels is then proportional to the number of links. This figure shows that the inter–pixel grating noise is not uniform across the image, and the worst crosstalk occurs at the center pixel. The number of links to or from the center pixel is about 6×10^7 which is much larger than the number of direct links between it and all the other pixels ($\approx N^2=10^4$). We plot the number of links of the center pixel as a function of the number of pixels N in one dimension on the SLM in Figure 6.12. It shows that the number of links to or from the center pixel in an image of 1000×1000 pixels is nearly 1×10^{12}. We learn from previous discussions that for a grating formed in between two pixels on an image of the same size, the inter–pixel grating strength at the end of recording of 10,000 holograms is about 12 times stronger than an information grating. For a quick assessment of the diffraction efficiency of the inter–pixel grating, we assume that the M/# of the memory is 1. The diffraction efficiency of a hologram of $1000\times$ 1000 pixels is then 1×10^{-8}. Therefore, the diffraction efficiency of an inter–pixel grating is about 2×10^{-14}. As a result, the diffraction efficiency of the "connection" consisting of all the 1×10^{12} links is 24%!

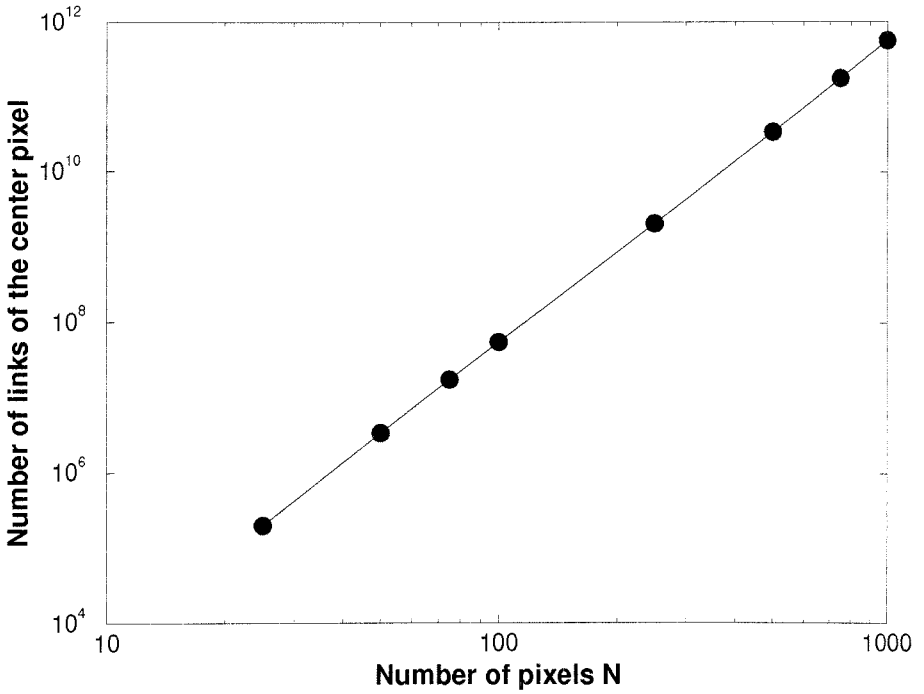

Figure 6.12: The number of "links" of the center pixel as a function of the number of pixels N in one dimension on the image.

The inter–pixel grating noise is much stronger than the crosstalk or any other noise sources.

6.4.2 Strength of inter–pixel gratings

In the above discussion, we show the worst–case scenario where all the inter–pixel gratings are treated as the same and the crosstalk arises from Bragg degenerate connections. In the remainder of this section, we modify the assumptions we made to reach a more realistic picture of the effect of the inter–pixel grating noise.

First, we notice that the strength of an inter–pixel grating depends strongly on its orientation with respect to the **c**–axis of the material (Figure 6.13). This angle–dependence comes from two effects—the photovoltaic and the electro–optic effect.

Photovoltaic effect: The linear photovoltaic effect in $LiNbO_3$ can be described by the following expression relating the light intensity and the photovoltaic current

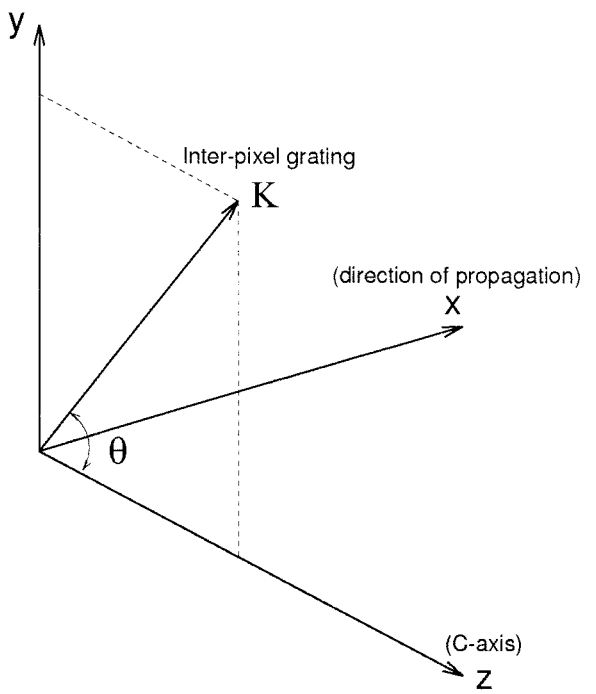

Figure 6.13: The grating vector of the inter–pixel grating.

inside the material, along the **c**–axis.

$$j^{pv} = \alpha G I, \tag{6.13}$$

where α is the absorption coefficient of the material. G is the constant of proportionality called the Glass constant [76, 77, 163].

When the grating vector is not parallel to the **c**–axis, the tensor description of the photovoltaic effect should be used to evaluate its contribution to the formation of the grating. It can be written as

$$\begin{aligned} j_i^{pv} &= G_{ijl}\hat{e}_j\hat{e}_l I, \\ i &= 1,2,3 \ \ or \ \ x,y,z. \end{aligned} \tag{6.14}$$

G_{ijl} is the element in a 3^{rd}–rank Glass tensor **G**. $\hat{e}_j$ and $\hat{e}_l$ are the unit vectors in the input light polarization. For ordinarily polarized light, the photovoltaic currents

along the x, y and z axis are (for detailed discussion, see Reference [163])

$$\begin{aligned} j_x &= \alpha G_{15} I, \\ j_y &= \alpha G_{22} I, \\ j_z &= \alpha G_{31} I. \end{aligned} \tag{6.15}$$

Since $G_{31} \gg G_{22} \gg G_{15}$ in $LiNbO_3$, a significant photovoltaic effect only takes place along the z (**c**)–axis. As a result, the photovoltaic current contributing to the buildup of an inter–pixel grating whose grating vector is not along the **c**–axis is simply the projection of j_z—$j_z \cos\theta$ (Figure 6.13). Therefore, the inter–pixel gratings whose grating vectors are parallel or at small angles to the **c**–axis are much stronger because of the significant contribution from the photovoltaic effect. This explains the horizontal energy exchange shown in Figure 6.4.

Electro–optic effect: $LiNbO_3$ is a uni–axial material whose index ellipsoid is written as [164]

$$\frac{x^2}{n_o^2} + \frac{y^2}{n_o^2} + \frac{z^2}{n_e^2} = 1, \tag{6.16}$$

where n_o, n_e are the ordinary and extraordinary index of refraction. In a typical arrangement for holographic recording, the writing beams are ordinarily polarized, and the grating vector of a hologram is parallel to the **c**–axis. Therefore, the electro–optic coefficient used to translate the space–charge field to the spatial modulation of the refraction index is Γ_{31}. When the grating vector of the inter–pixel grating is not parallel to the **c**-axis, but pointing at an arbitrary direction as shown in Figure 6.13, the eigen-axes of the crystal are rotated, and the electro–optic effect takes a more complicated form. By neglecting the x–component of the grating vector, the changes

in the index ellipsoid can be written as

$$\begin{aligned}
\Delta(\frac{1}{n^2})_1 &= -\Gamma_{22}E_y + \Gamma_{13}E_z, \\
\Delta(\frac{1}{n^2})_2 &= \Gamma_{22}E_y + \Gamma_{13}E_z, \\
\Delta(\frac{1}{n^3})_3 &= \Gamma_{33}E_z, \\
\Delta(\frac{1}{n^2})_4 &= \Gamma_{51}E_y,
\end{aligned} \tag{6.17}$$

where E_y and E_z are the two components of the space–charge field E_{sc} along the y, z axes.

$$E_y = E_{sc}\sin\theta; \quad E_z = E_{sc}\cos\theta. \tag{6.18}$$

The presence of a non–zero $\Delta(\frac{1}{n^2})_4$ results in the rotation of the y, z axes about the x axis. The angle of rotation is

$$\phi = \arctan\frac{2C}{B-A}/2, \tag{6.19}$$

where

$$\begin{aligned}
A &= \frac{1}{n_o^2} + \Delta(\frac{1}{n^2})_2, \\
B &= \frac{1}{n_e^2} + \Delta(\frac{1}{n^3})_3, \\
C &= \Delta(\frac{1}{n^2})_4.
\end{aligned} \tag{6.20}$$

Typically, this angle is very small [8], so we could neglect the rotation of the axes. As a result, the index modulation seen by the ordinarily polarized beams can be written as

$$\Delta n_o = -\frac{n_o^3}{2}(\Gamma_{22}E_y + \Gamma_{13}E_z). \tag{6.21}$$

[8]For $\theta \approx 90°$ and a space–charge field of 500v/cm, ϕ is less than 0.006°.

It depends on the angle θ between the grating vector and the z (**c**)–axis.

Different spatial frequencies: The above discussion revealed the dependence of the strength of an inter–pixel grating on the its orientation. Here we describe the effect of the spatial frequency of a grating. The purpose is to find the spatial–frequency–dependence of the grating strength. We repeated the evaluation of the time evolution of the gratings as described in Section 6.3. In the simulation, we chose two extreme cases:

- The angle between the two plane waves was 1°. This corresponds to the highest spatial frequency of the inter–pixel gratings if an SLM of dimension 20mm×15mm and a lens of focal length 300mm were used.
- The angle between the two plane wave was 0.003° which creates almost the lowest spatial frequency under our experiment conditions (for adjacent pixels on TVT6000 with 480× 440 pixels).

The results only showed a less than 5% difference in the grating strengths in the two cases. This suggests that at very low spatial frequencies, the diffusion is overwhelmed by the photovoltaic effect. As a result, the difference in the spatial frequencies does not make a big difference in our theoretical treatment.

In conclusion, the variations in the grating strength only come from the orientation of the grating vector, not its magnitude.

Bragg degenerate "connections" via gratings of different strengths: The general numerical evaluation of the inter–pixel grating noise is very time–consuming, since the crosstalk arises from the collective behavior of all the N^2 pixels via the N^4 possible inter–pixel gratings, after the storage of many random–bit pattern holograms. As we learn from previous discussions, the worst inter–pixel grating noise occurs at the center of the image. Therefore, in the following analysis, we only consider the noise statistics of the center pixel. First, we would analyze the "contribution" a pixel at (x, y) makes to the noise buildup at this center pixel. If we neglect the variations of

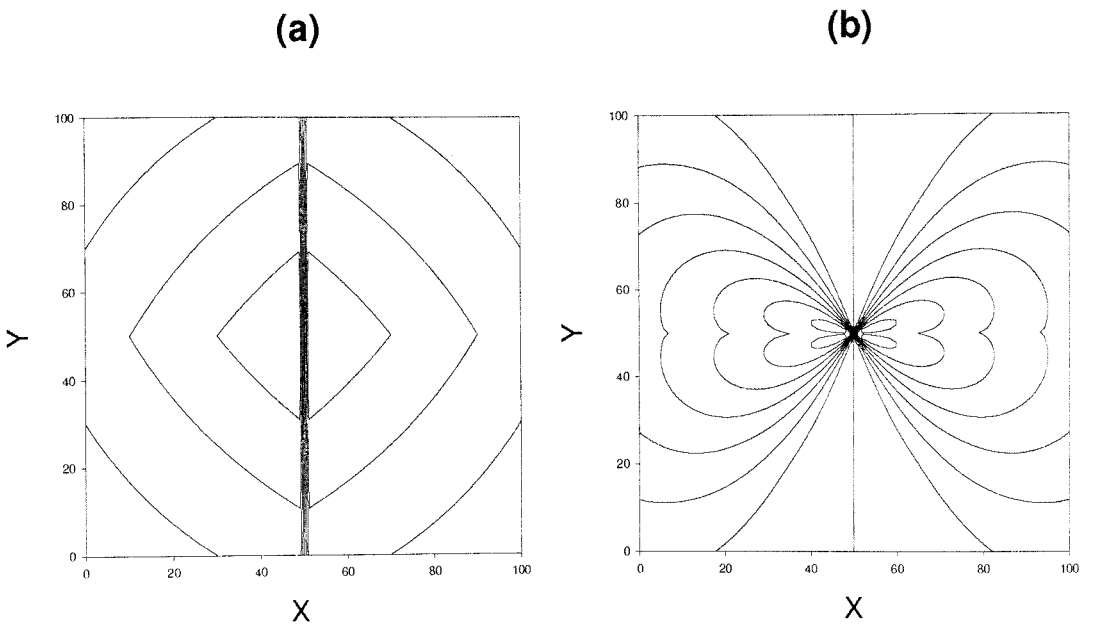

Figure 6.14: The "contribution" the pixel at (x, y) makes to the noise buildup at the center pixel of a 100×100 image: (a) worst case where all the inter–pixel gratings are of the same strength; (b) variations in the grating strength are considered.

the grating strength, it can be described by the number of links N_g (Equation 6.11) and shown in the form of a contour plot in Figure 6.14(a). The number of links ("contribution" to the noise at the center) is the same for the pixels on the same contour line.

Figure 6.14(b) describes the case where the variations in the grating strength are included. In order to compare with the worst–case scenario where the number of links represents the amount of contribution a pixel makes to the center pixel, we used the "normalized" number of links. To obtain the "normalized" number of links, we first evaluated the strength of the inter–pixel grating after the recording of 10,000 holograms. This grating strength was then normalized by the strength of a grating that is parallel to the **c**–axis (see Figure 6.7). The normalized grating strength was used to "weigh" the number of links N_g to generate the normalized number of links. By doing this, the normalized number of links from a pixel is directly related to the diffraction efficiency that leads to crosstalk at the center from this pixel, using the simulation results we obtained earlier. This figure shows that for those pixels with

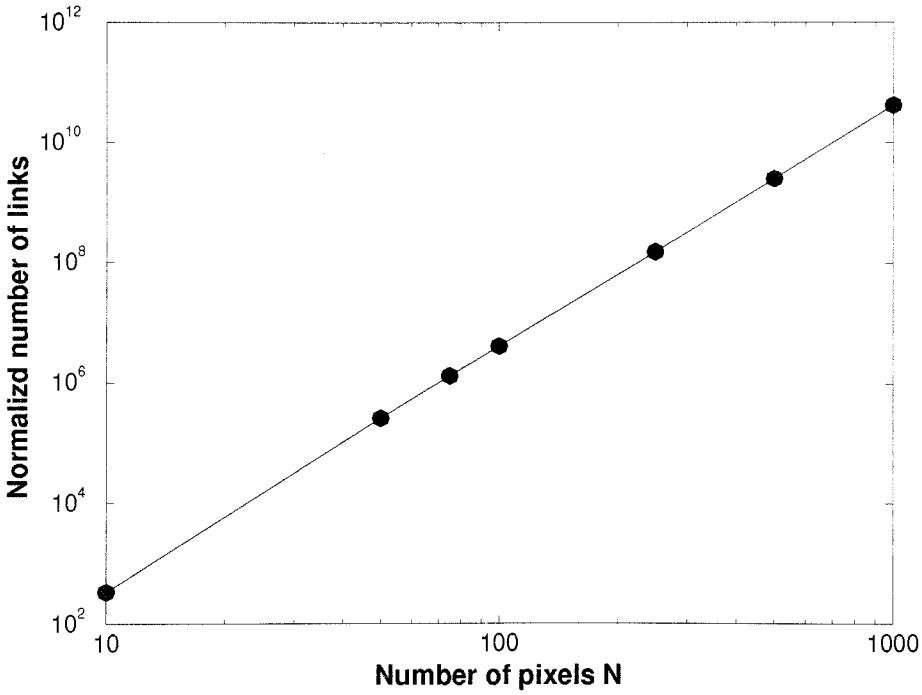

Figure 6.15: The normalized number of "links" of the center pixel as a function of the number of pixels N in one dimension on the image.

links at very large angles to the **c**–axis, their contribution to the noise at the center becomes much less.

The normalized number of links at the center pixel from all the other pixels in the same image as a function of the number of pixels N in one dimension on the image is plotted in Figure 6.15. As shown in this figure, the number of links for an image of 1000×1000 is reduced to about 2×10^{11} because of the variations in their strengths. In a similar way as in Section 6.4.1, we can evaluate the diffraction efficiency of the noise in this case. Using the same conditions, the estimated value is about 7%.

6.4.3 Bragg selectivity

When we consider the z–components of the grating vectors, most of the links in group specified by the relative positions of a pair of pixels are no longer degenerate. However, since the angles between the plane–wave components in the signal are very small, these gratings are still capable of giving rise to crosstalk between the two pixels. This is shown in Figure 6.16. In Figure 6.16(a), two pairs of plane waves

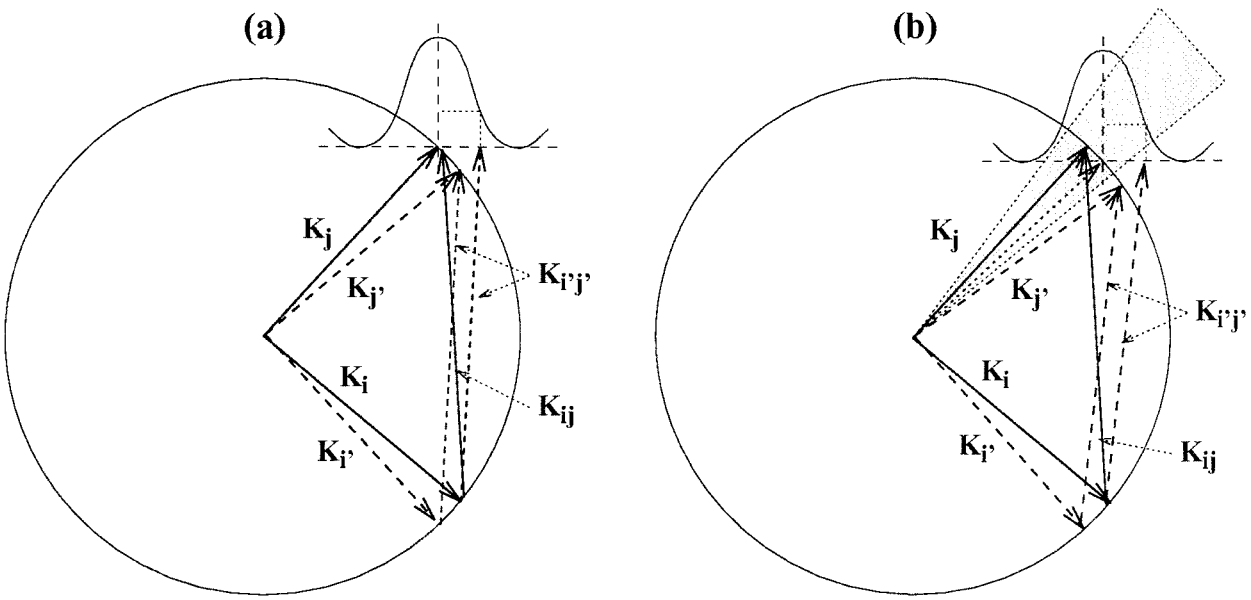

Figure 6.16: "Connections" arising from wide Bragg selectivity.

corresponding to pixel (i, j) and (i', j') write the two inter–pixel gratings $\vec{K}_{ij}, \vec{K}_{i'j'}$. Upon readout, pixel j receives noise from pixel i via grating $\vec{K}_{ij}$ (the direct link). In the meantime, though gratings $\vec{K}_{ij}$ and $\vec{K}_{i'j'}$ are no longer degenerate, additional crosstalk at pixel j from i via $\vec{K}_{i'j'}$ exists as long as the vector $\vec{K}_i + \vec{K}_{i'j'}$ ends on the horizontal line passing through $\vec{K}_j$ and falls in the width of the selectivity curve. In addition, because the pixels are not perfect point sources and the aperture of the optical system is finite, a "cone" of wave–vectors instead of a single one is collected at one pixel [9]. As a result, the reconstructions which are not along K_j but fall in this cone which is mainly determined by the diffraction limit would contribute to the noise at pixel j as well (Figure 6.16(b)).

To evaluate the total amount of noise at a pixel, we assume that the fill factor of the SLM and CCD is 1. The rediffraction by a pixel will always contribute to the noise at other pixels. The crosstalk between pixels is a mutual effect—the noise at pixel i from pixel j via any noise grating is the same as that at pixel j from pixel i via the same grating. In addition, if we neglect the phase difference in the diffraction introduced by the photorefractive phase shift, the noise at one pixel is the coherent

[9]The overlapping of the "cones" associated with neighboring pixels results in "inter–pixel noise."

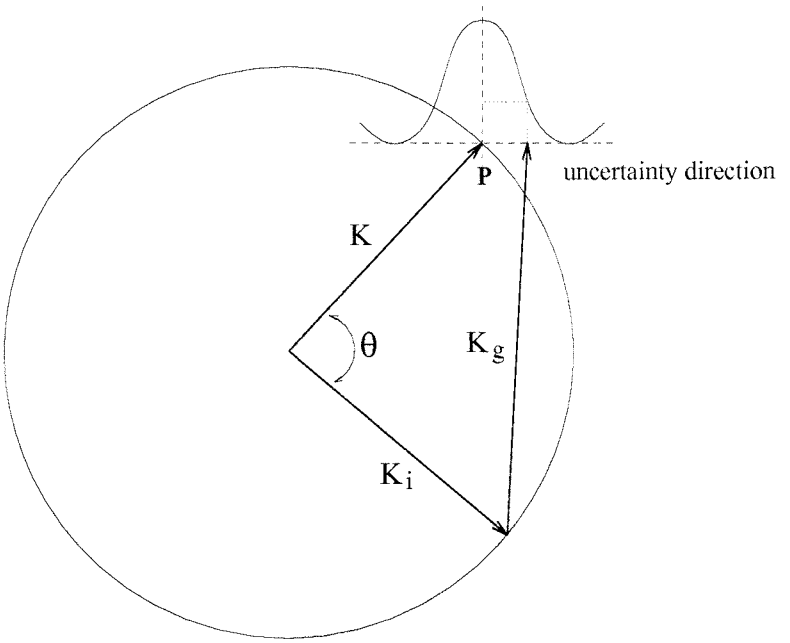

Figure 6.17: The rediffraction by pixel i via noise grating $\vec{K}_g$.

summation of that from every pixel in the same image [10]. As a result, when we consider the crosstalk noise at a pixel, we can simplify the analysis by computing the noise the pixel gives away by rediffraction.

To compute the amount of rediffraction a pixel i gives rise to via a specific grating $\vec{K}_g$, only the angle between the plane wave corresponds to this pixel and the rediffraction is needed. This angle can be determined in a way as shown in Figure 6.17. First, a horizontal line is drawn from vector $\vec{K}_i + \vec{K}_g$. It intersects the K–sphere at point P. The wave vector $\vec{K}$ pointing at P is then the rediffraction from pixel i via noise grating $\vec{K}_g$. The angle θ between $\vec{K}$ and $\vec{K}_i$ determines the width of the Bragg selectivity. The distance from P to the tip of vector $\vec{K}_i + \vec{K}_g$ is the phase–matching factor. It determines the strength of this rediffraction by pixel i.

We numerically evaluated the crosstalk noise at the center pixel which is believed to be the worst across the image. The number of pixels on the SLM was assumed to be 500 × 500. The dimensions of the SLM was chosen as 1.5cm× 2cm, which is similar to the Epson TVT6000 Liquid Crystal TV we used in our large–scale memory

[10]The phase difference caused by the path difference is cancelled out by the phase in the inter–pixel gratings.

system. The focal length of the Fourier–transform lens was chosen as 300mm. The exposure sequence in the evaluation was based on the recording of 10,000 holograms with a time constant determined by equalizing the hologram strength numerically (Section 6.3). The computer simulation was carried out multiple times to provide enough samples to show the variation in the crosstalk due to the random recording sequence as described in Section 6.3. The data was then processed to give the mean and the variance of the crosstalk and an SNR was calculated.

The SNR at the center pixel was found to be 29.5 after the recording of 10,000 holograms. It is much worse than what is determined by crosstalk noise when holograms are recorded at the first null of the angle selectivity curve of the neighboring hologram [165]. In addition, one should note that this is the result based on the assumption that the inter–pixel grating noise is a linear effect. In many applications where significant non–linear effect exists, the SNR would be much lower.

6.5 Random–phase modulation in the signal

We learn from previous discussions that the inter–pixel grating noise is a very important noise source in a holographic storage system. In a memory storing large amount of holograms, it becomes the dominant factor in limiting the error performance and hence the storage capacity of the system. Inter–pixel grating noise is a global effect since it arises from the secondary diffraction among the pixels in the reconstruction. This makes it hard to reduce [11]. On the other hand, the inter–pixel grating noise is not uniform across the entire image. Therefore, position–dependent threshold has to be applied to retrieve the binary information with minimal errors. However, position–dependent thresholding complicates detection and data processing. In addition, since the amount of noise at a pixel depends strongly on the data structure of its neighborhood, effective elimination can not be achieved by "post–recording" compensation

[11] As mentioned earlier, "bad" holograms normally appear at the beginning of the recording sequence because of two–beam coupling effect due to long exposures. Since this kind of noise depends on the reference beam, it only exists at the first several addresses (reference beam angles). Thus we can always eliminate this "local" noise by throwing the "bad" holograms away.

(Chapter 2 or [72,161]).

One of the reasons for an inter–pixel grating to be strong enough to give rise to noise during readout is that it is being strengthened repeatedly in multiple hologram recording (whenever the two pixels on the SLM are ON at the same time). Therefore, if the phase difference between the planes waves corresponding to the pixels is randomly and continuously modulated at either zero or π during exposure, the final grating strength as a result of "positive" and "negative" recording shall be greatly reduced. In the meantime, a real hologram is not affected except for the extra phase added to each pixel which would not make a difference at the CCD because of the intensity detection. Furthermore, since random–phase modulation is capable of stopping significant growth of the inter–pixel grating, the non–linear effect as the result of interaction between the writing beams and their strong diffraction shall be greatly suppressed as well.

We carried out an experiment to demonstrate the effect of random–phase modulation on the buildup of inter–pixel grating noise. We repeated the demonstration in Section 6.2 to show the deterioration of the image quality due to continuous signal illumination (refer to Figure 6.4). Then we inserted an SLM which was converted to operate in phase modulation in the signal path. The fixed chessboard pattern was imaged onto this SLM and focused at the crystal by an FT lens. While the crystal was illuminated continuously by roughly the same amount of light intensity, a sequence of random–bit patterns were displayed on the SLM at about 5Hz. The transmitted image was sampled at 5, 10, 20 minutes. They are shown in Figure 6.18. Comparing the image quality of the transmission with and without random–phase modulation, a conclusion can be drawn that, with phase modulation, the buildup of inter–pixel grating noise becomes slower. The reason that it was not completely eliminated is that the SLM was not operating in pure phase modulation.

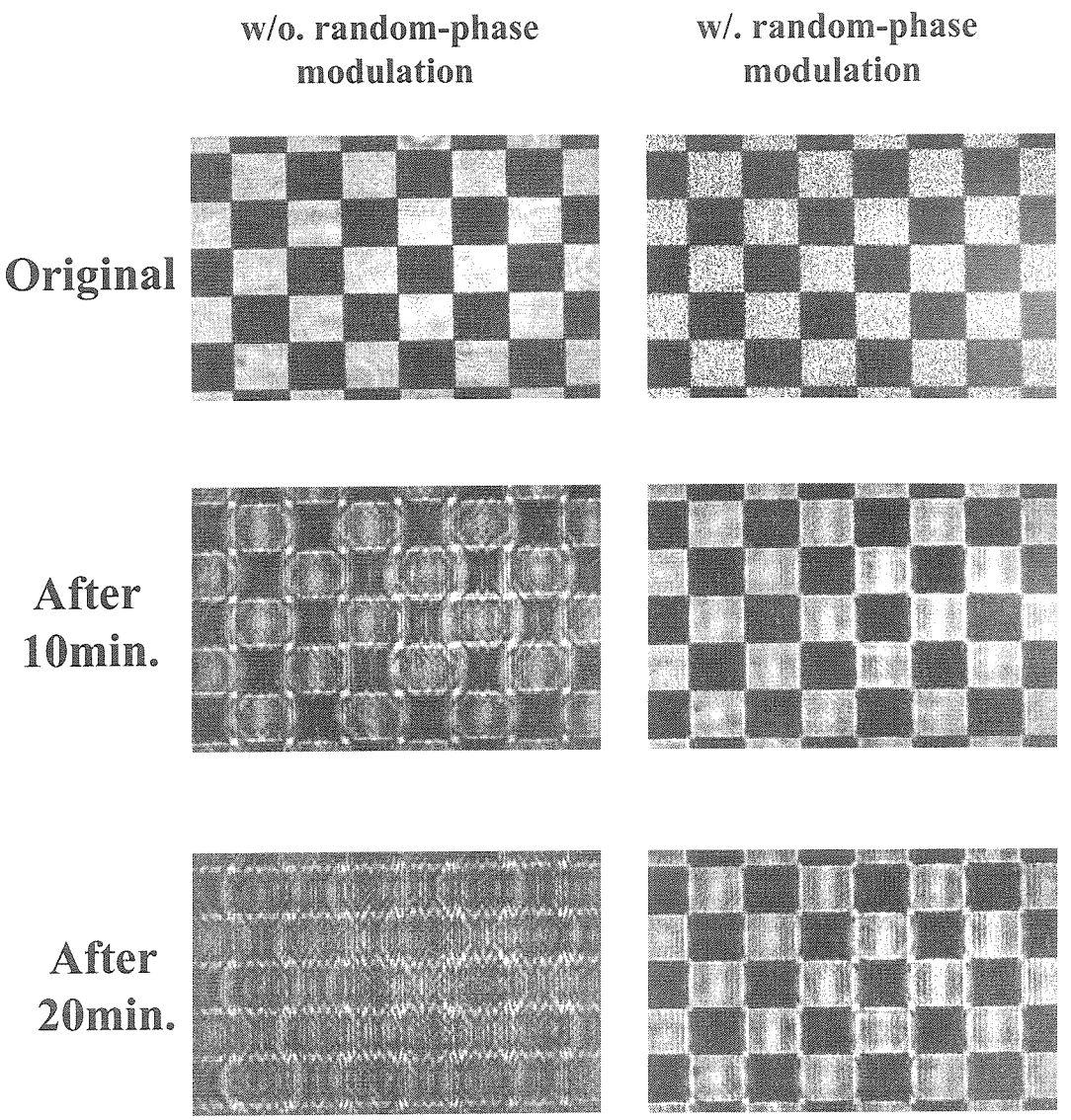

Figure 6.18: Random–phase modulation to reduce the strong two–beam coupling effect.

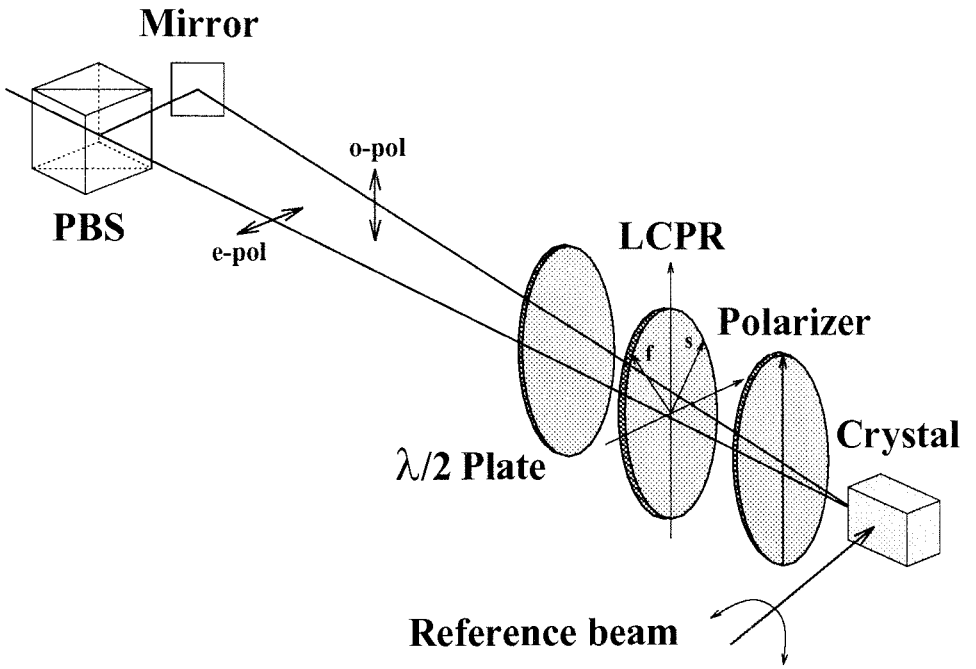

Figure 6.19: Experimental setup for the plane–wave signal experiment.

6.6 Experiment results

In this section, we describe experimental demonstrations that show the buildup of inter–pixel gratings as well as the deterioration in the SNR as a result of the buildup of the inter–pixel gratings. We also demonstrate experimentally the application of random–phase modulation in the signal to improve the system error performance by suppressing the inter-pixel grating.

6.6.1 Experiment with plane–wave signal

The main purpose of this experiment is to reveal the temporal evolution of the inter–pixel gratings and to demonstrate the use of random–phase modulation to prevent them from being strengthened repeatedly. The experimental setup is shown in Figure 6.19. It was based on our 160,000 hologram system with a storage capacity of 10,000 holograms using angle and fractal multiplexing. In the signal arm, the SLM and following imaging and Fourier–transform optics were removed. Instead, we used two plane waves to represent two pixels at the Fourier plane. The two plane waves

were set apart horizontally by an outside angle of 1°. Two shutters were used to control the ON/OFF states of the two beams independently.

To randomly control the phase difference between the two plane–wave beams, a Liquid Crystal Phase Retarder (LCPR) (Meadowlark D1040) was used in the signal path. The fast and slow axes of the liquid crystal device were at 45° to the lab axes. An alternating signal was applied to the device to induce a change in the extraordinary refraction index. As a result, the phase difference between the two eigen–modes travelling through the device can be controlled electronically. A half–wave plate was placed in front of the LCPR to align the two linearly polarized plane–wave beams with the fast and slow axes of the LCPR. After the device, a phase difference of either 0 or π was imposed onto the two beams. A polarizer followed the LCPR to filter out unwanted polarization.

To find the electronic control signals that create a π phase shift between the two eigen–modes, the LCPR was first calibrated. It was placed in between a pair of crossed or parallel polarizers, and operated in intensity–modulation mode. This is shown in Figure 6.20. The difference between the applied signals that correspond to the maximal and minimal transmission would result in a phase change of π.

To simulate the recording of binary random–bit patterns, we randomly modulated the ON/OFF states of the two beams at the same time to form a series of random 2–bit patterns. The crystal was exposed by each pattern according to an exposure schedule. Ideally, since the inter–pixel grating is written by the plane-wave components in the signal beam, to investigate their buildup, we can simply illuminate the crystal with the signal beam for a certain amount of time. This would save a lot of trouble because holographic recording is not involved and the noise caused by the reference beam illumination is ruled out. However, because we did not have a large amount of plane waves (pixels), the ON/OFF states of the two beams significantly affected the total light intensity, thus the buildup rate of the inter–pixel grating. Therefore, a strong reference beam was used to actually record holograms using the exposure schedule. The intensity ratio of the reference beam and each of the plane–wave signal was set at more than 10. This guaranteed that the total exposing light intensity,

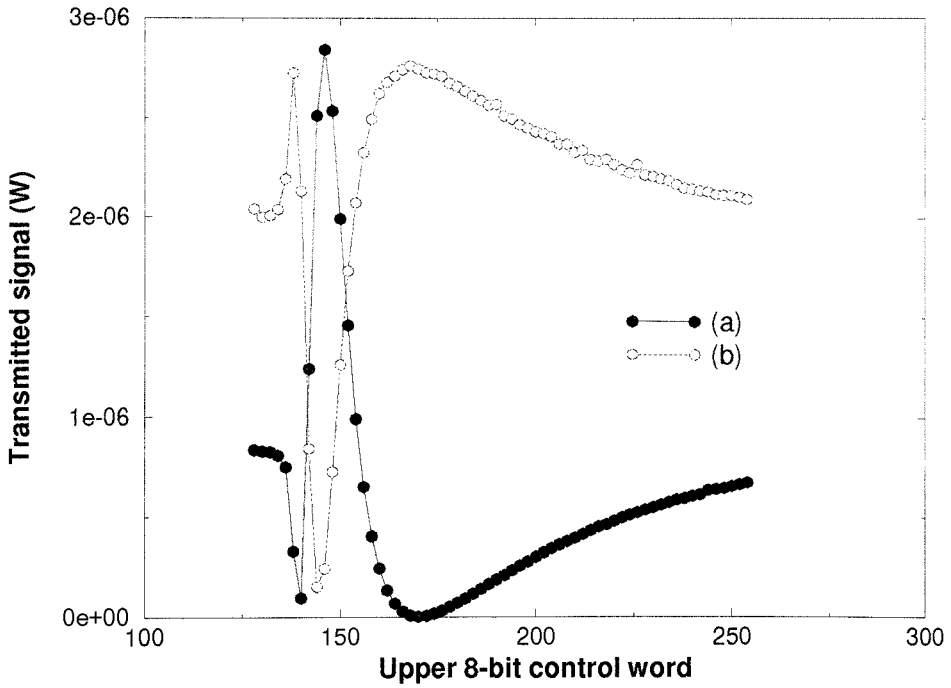

Figure 6.20: Liquid Crystal Phase Retarder operating in intensity–modulation mode: (a) in between crossed polarizers; (b) in between parallel polarizers.

hence the time constant, remained nearly the same regardless of the ON/OFF states of the two signal beams. We simulated the case for the storage of 10, 100, 300, 1,000, 2,000, 3,000, and 5,000 holograms. The same exposure schedule with the same last exposure time was used to record equalized holograms with the same strength despite the different number of holograms stored for each experiment. At the end of the recording, one of the signal beams was turned on and its diffraction into the other one via the inter–pixel grating formed during the prolonged exposure was monitored by a photodetector. The readout was used to calculate the diffraction efficiency of the inter–pixel grating and compared to that of the hologram reconstructed by the reference beam. In the demonstration with random–phase modulation, we used a random–number generator to determine whether the phase difference between the two beams is to be set at 0 or π before recording a new hologram. At the end of recording, the same procedure was followed to yield the ratio of the strength of the inter–pixel grating and the holograms as a function of the number of holograms. The results from the experiments with and without random–phase modulation are shown

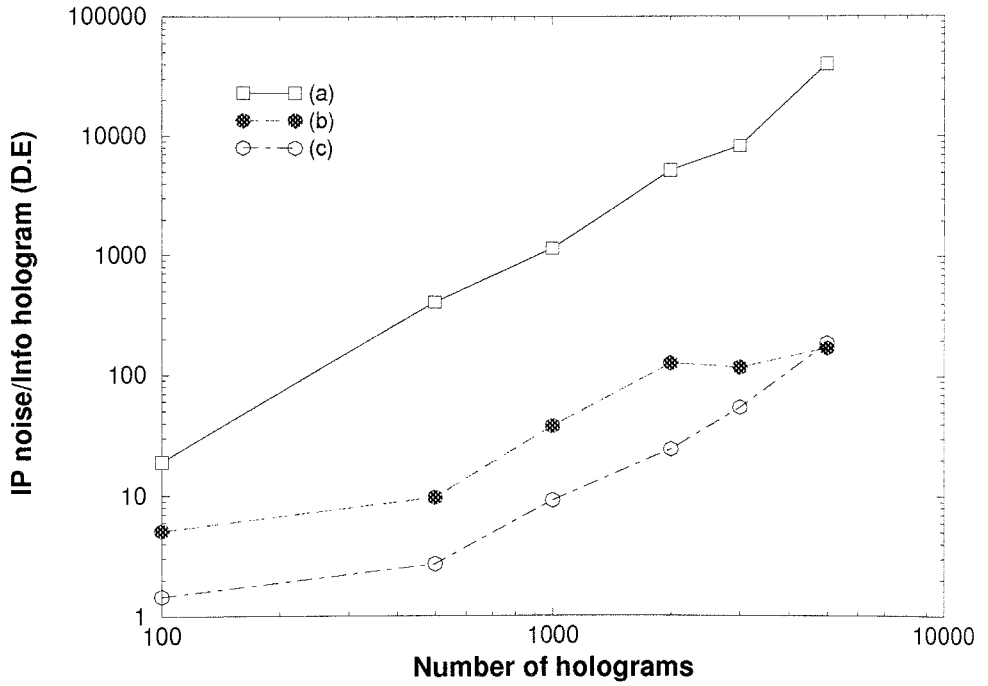

Figure 6.21: Ratio of the diffraction efficiencies of the inter–pixel grating and the hologram as a function of the number of holograms (a) without phase modulation; (b) random–phase modulation all the time; (c) π phase shift only when the two beams were ON at the same time.

in Figure 6.21(a) and (b).

From Figure 6.21, the inter–pixel grating is about 4×10^4 times stronger than the 5,000 (information) holograms stored. Since the modulation depth of this inter–pixel grating is much larger than that in a practical case where there are a large amount of pixels, the relative change or the rate of increase provides more realistic information about its buildup. Compared to the recording of 100 holograms, the inter–pixel grating is strengthened by a factor of more than 1,000! This shows how serious the inter–pixel grating noise could be in a large–scale storage system. We also learn from Figure 6.21 that the strength of the inter–pixel grating can be greatly reduced by random–phase modulation. However, the reduction is not significant when the number of holograms is small. This could be attributed to the fact that the ON/OFF states of the two plane waves and their phase difference were controlled randomly and independently. When there are not a large amount of holograms, there are not enough samples to average out the grating buildup by "positive" and "negative" recordings.

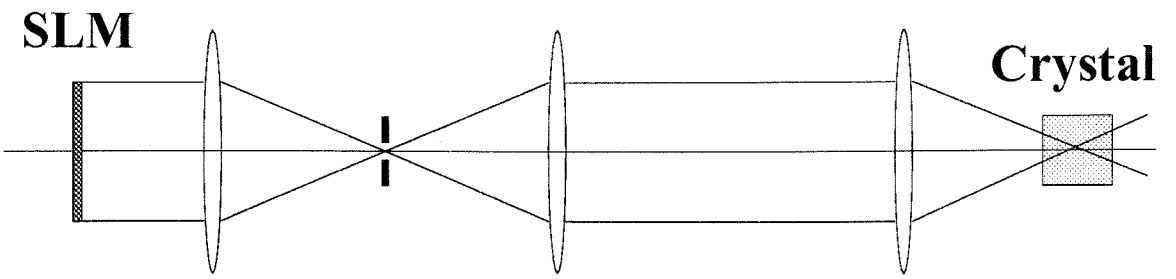

Figure 6.22: Experimental setup for random-bit patterns.

In order to observe the "ideal" effect of phase modulation, we chose a different way to control the phase difference in another set of experiments. Before the recording of a new hologram, the ON/OFF states of the two beams were compared. If they were to be both ON, a new phase difference was set which was π phase-shifted with respect to the old one. This guaranteed that whenever the inter-pixel grating was to be strengthened effectively, the exposing pattern underwent a π phase shift relatively, which "nullified" the previously written one. The result from this set of experiments is also shown in Figure 6.21. As indicated by Curve (c), the strength of the inter-pixel grating was further weakened. It is interesting to note that curve (b) and (c) seem to converge at a non-zero value. This means that even with random-phase modulation, the inter-pixel grating can not be completely eliminated. Intuitively, for repeated recording of a hologram at the same angle location with a "perfect" exposure schedule, the alternating "positive" and "negative" recording as used in our experiments would leave nothing recorded at all. However, in a holographic memory, the exposure schedule was determined by the dynamics of the buildup of the information holograms (in 90°-geometry recording in our case), while the buildup of an inter-pixel grating is different in the time constant. As a result, the "perfect" exposure schedule for the information holograms is not perfect for the inter-pixel gratings. After the repeated strengthening and weakening by phase modulation, a residual grating remains, capable of causing noise.

6.6.2 Experiment with random–bit patterns

In this section, we investigate the effect of the inter–pixel grating noise in a holographic memory system storing binary random–bit patterns. With random–bit patterns, there are enough samples in the reconstruction that allow us to use the Signal–to–Noise Ratio (SNR) to evaluate the system error performance statistically. In addition, the formation and influence of the inter–pixel grating noise as the result of the collective effect among many pixels can be revealed. Another merit is the simplified experiment procedure—we can use the complex signal alone to illuminate the crystal to keep track of the evolution of the inter–pixel grating noise.

To generate the random–bit patterns, we used the finest graphics mode in BASICA which has 640× 200 pixels on VGA display. The SLM we used was an Epson TVT6000 Liquid Crystal TV with 480× 440 pixels. Since it takes a long time for the computer to generate the random–bit patterns in real time, we prepared a library of 100 patterns before the recording. These 100 patterns were cycled many times for the recording of a large amount of holograms. Therefore, the recording sequence was not strictly random.

To randomly phase–modulate the signal, we first tried to use another SLM which was converted to phase–modulation mode. The SLM which operated in intensity modulation to display the random–bit patterns was imaged onto the phase–modulation SLM (PM SLM) by a 4F system. The phase–modulated image was then focused at the crystal by a Fourier–transform lens. During the exposure corresponded to the recording of a random–bit pattern, the PM SLM displayed a series of random–bit patterns continuously to make the phase difference between the pixels time–varying and random. The CCD detector only detected the random–bit patterns displayed by the intensity–modulation SLM (IM SLM) to retrieve the stored information. However, this method had several drawbacks:

- The SLMs are very lossy because of their pixelation and small fill factor (usually, the light obtained in the filtered DC portion is only 10%); a large amount of light was wasted by stacking two SLMs together.

- For optimal random-phase modulation, the PM SLM needs to display the random-bit patterns very rapidly.

- It complicates the system hardware including precision imaging optics, additional SLM and control electronics, etc.

Alternatively, we can use only one SLM for both displaying and random-phase modulating the signal. This could be carried out with a phase-modulation SLM (PM SLM) which encodes the information as phase shift 0 or π at the pixels. All the pixels on the PM SLM are "ON" all the time regardless of the image to be stored since the information is embedded in the phase, not the amplitude. As a result, all the inter-pixel gratings are recorded every time when a new hologram is added to the memory. However, since the possibilities of one pixel being either 1 or 0 (phase-shifted by 0 or π) are the same in a random-bit pattern, the possibility of two pixels being in the same state is 1/2. In other words, the two pixels are the same for half of the time, and in the other half, they are different. In the PM SLM, when two pixels are in the same state, the phase difference between them is zero. On the other hand, the phase difference is π when the two pixels are in different states. Therefore, although a pair of pixels continuously write the "linking" inter-pixel grating during hologram recording, during half of the exposures on average, the recording is "positive" (strengthening). While in the other half, the recording is "negative" (weakening). As a result, the inter-pixel grating noise could be greatly reduced by the random-phase modulation that arises from the random occurrence of the binary random-bit patterns to be stored in the memory. This random-phase modulation comes naturally as the side-product with a PM SLM as the displaying device. This method significantly simplifies the system hardware. Other advantages of using a PM SLM for holographic storage include the improved light efficiency (all the pixels on the PM SLM are ON all the time) and well distributed Fourier spectrum in Fourier-transform-plane recording [12]. However, the use of PM SLM raises problems in the detection. Since the information is embedded in the phase, special techniques are required to translate the phase information

[12]In Chapter 2, we had to use a separate random-phase plate to diffuse the DC portion of the spectrum for optimal performance with Fourier-transform-plane recording.

into intensity information for data detection and retrieval. This can be done with an interferometer. But the interferometers are very sensitive to the environment such as air current and mechanical vibrations. Alternatively, edge-enhancement of the reconstructed holograms can be applied at the expense of weaker signal and lower capacities. A better approach might be the application of a specially designed phase plate similar to the phase-contrast microscope invented by Zernike 60 years ago. The basic idea is to introduce different phase excursions in different portions of the Fourier spectrum of the reconstruction which will translate the phase-modulated signal back to intensity modulation. This would not cost a great loss in the light intensity but might result in a reduction in the capacity.

We simulated the recording of a single hologram, 10, 30, 100, 300, 1,000, 2,000, 3,000, 5,000, and 10,000 holograms by illuminating the crystal with random-bit patterns according to an exposure schedule. At the end of every simulation, a random-bit pattern was displayed on the SLM. The transmitted signal was then sensed by a CCD camera (Photometric Imagepoint CCD) and sent to a computer for data analysis. By doing this, the signal was the direct transmission while the noise was its diffraction via the inter-pixel gratings. The SNR from the transmitted images is summarized in Figure 6.23(a). The SNR drops from 10.3 from a single hologram down to 5.2 when 10,000 holograms were "stored," indicating significant influence of the inter-pixel grating noise as the total exposure increases.

In the experiment with the phase-modulation SLM, we did not use any of the special techniques in the detection. Instead, after the simulation of the storage of a certain number of holograms, the PM SLM was reconfigured back in intensity-modulation mode. The SLM displayed a random-bit pattern, and the transmission through the crystal was detected and processed. The SNR as a function of the number of holograms is also shown in Figure 6.23. As shown by Curve (b) in Figure 6.23, the system error performance was improved significantly by random-phase modulation.

The means and variances of the OFF pixels in the two sets of experiments are shown in Figure 6.24. It clearly shows the reduction in the OFF (dark) pixels as the inter-pixel grating noise was suppressed.

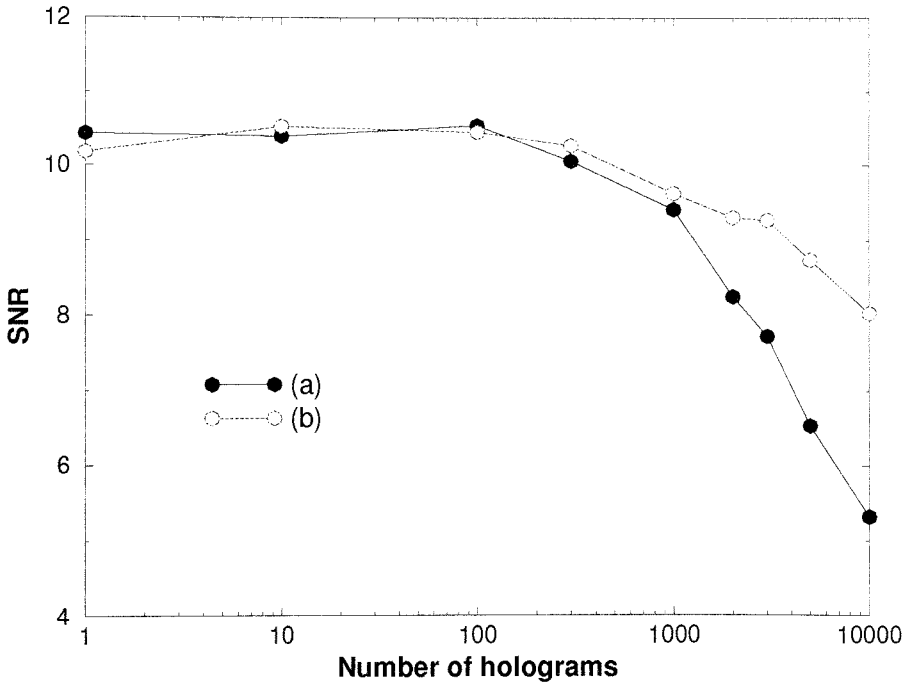

Figure 6.23: SNR as a function of the number of holograms: (a) without random-phase modulation; (b) with random-phase modulation.

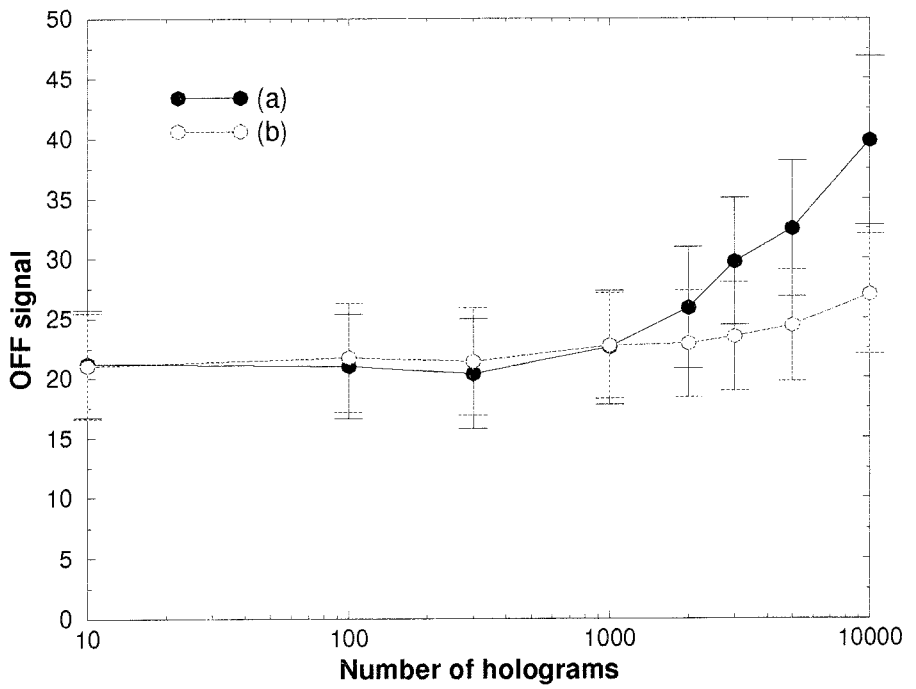

Figure 6.24: OFF signal as a function of the number of holograms: (a) without random-phase modulation; (b) with random-phase modulation.

We did not compare the experiment result to theory which predicts the deterioration in the system error performance in terms of the SNR. The major reason is that the SNR of simply the transmitted signal through a "fresh" crystal is lower than what is predicted by the theory after the storage of 10,000 holograms. This is because of the limited contrast ratio of the SLM ($\approx$ 30 overall) and the "inter–pixel noise" as shown in Figure 6.3. In addition, the drop in the SNR we observed from the experiments without random–phase modulation suggests a more dramatic increase of the noise than what is predicted by the theory. We suspect the reason to be the non–linear effect which is ignored in our theoretical treatment. Nonetheless, the significant suppression of the noise by using the random–phase modulation is in good support, though indirect, of the key point in our theory—the inter–pixel grating noise that arises from the noise gratings formed during multiple hologram recording is the dominant holographic noise source in a large–scale memory.

Bibliography

[1] D. Gabor, "A new microscope principle," *Nature*, vol. 161, pp. 777–778, 1948.

[2] D. Gabor, "Microscopy by reconstructed wavefronts," *Proceeding of Royal Society A*, vol. 197, pp. 454–487, 1949.

[3] E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory," *Journal of the Optical Society of America*, vol. 52, pp. 1123–1130, 1962.

[4] E. N. Leith and J. Upatnieks, "Wavefront reconstruction with diffused illumination and three–dimensional objects," *Journal of the Optical Society of America*, vol. 54, pp. 1295–1301, 1964.

[5] S. A. Benton, "Hologram reconstruction with extended incoherent sources," *Journal of the Optical Society of America*, vol. 59, p. 1454A, 1969.

[6] P. J. van Heerden, "Theory of optical information storage in solids," *Applied Optics*, vol. 2, no. 4, pp. 393–400, 1963.

[7] E. N. Leith, A. Kozma, J. Upatnieks, J. Marks, and N. Massey, "Holographic data storage in three–dimensional media," *Applied Optics*, vol. 5, no. 8, pp. 1303–1311, 1966.

[8] E. G. Ramberg, "Holographic information storage," *RCA Review*, vol. 33, pp. 5–53, 1972.

[9] A. V. Lugt, "Signal detection by complex signal filtering," *IEEE Transaction on Information Theory*, pp. 139–145, April 1964.

[10] J. Upatnieks, A. V. Lugt, and E. N. Leith, "Correction of lens aberrations by means of holograms," *Applied Optics*, vol. 5, no. 4, pp. 589–593, 1966.

[11] J. Upatnieks, A. V. Lugt, and E. N. Leith, "Wavefront–reconstruction imaging through random media," *Applied Physics Letters*, vol. 8, pp. 311–315, 1966.

[12] L. J. Cutrona, "On the application of coherent optical processing techniques to synthetic–aperture radar," *Proceedings of the IEEE*, vol. 54, p. 1026, 1966.

[13] W. L. Bragg, "The X–ray microscope," *Nature*, vol. 149, pp. 470–475, 1942.

[14] H. Kogelnik, "Coupled wave theory for thick hologram gratings," *The Bell System Technical Journal*, vol. 48, no. 9, pp. 2909–2947, 1969.

[15] B. Benlarbi, D. J. Cooke, and L. Solymar, "Higher order modes in thick phase gratings," *Optica Acta*, vol. 27, no. 7, pp. 885–895, 1980.

[16] B. Benlarbi and L. Solymar, "Higher order modes in non–sinusoidal volume phase gratings," *International Journal of Electronics*, vol. 48, no. 4, pp. 351–359, 1980.

[17] M. G. Moharam, T. K. Gaylord, and R. Magnusson, "Criteria for bragg regime diffraction by phase gratings," *Optics Communications*, vol. 32, no. 1, pp. 14–18, 1980.

[18] 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*, vol. 70, no. 4, pp. 437–442, 1980.

[19] 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*, vol. 9, no. 1, pp. 72–74, 1966.

[20] J. J. Amodei, "Electron diffusion effects during hologram recording in crystals," *Applied Physics Letters*, vol. 18, no. 1, pp. 22–24, 1971.

[21] J. J. Amodei, "Analysis of transport processes during holographic recording in insulators," *RCA Review*, vol. 32, pp. 185–197, 1972.

[22] 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*, vol. 24, no. 6, pp. 264–265, 1974.

[23] G. A. Alphonse, R. C. Alig, D. L. Staebler, and W. Phillips, "Time dependent characteristics of photoinduced space-charge field and phase holograms in lithium niobate and other photorefractive media," *RCA Review*, vol. 36, pp. 213–229, 1975.

[24] V. L. Vinetskii and N. V. Kukhtarev, "Theory of the conductivity induced by recording holographic gratings in nonmetallic crystals," *Soviet Physics—Solid State*, vol. 16, no. 12, pp. 2414–2415, 1975.

[25] D. W. Vahey, "A nonlinear coupled-wave theory of holographic storage in ferroelectric materials," *Journal of Applied Physics*, vol. 46, no. 8, pp. 3510–3515, 1975.

[26] S. F. Su and T. K. Gaylord, "Unified approach to the formation of phase holograms in ferroelectric crystals," *Journal of Applied Physics*, vol. 46, no. 12, pp. 5208–5213, 1975.

[27] N. V. Kukhtarev, "Kinetics of hologram recording and erasure in electrooptic crystals," *Soviet Technical Physics Letters*, vol. 2, pp. 438–440, December 1976.

[28] M. G. Moharam and L. Young, "Hologram writing by the photorefractive effect with gaussian beams at constant applied voltage," *Journal of Applied Physics*, vol. 47, no. 9, pp. 4048–4051, 1976.

[29] R. Magnusson and T. K. Gaylord, "Use of dynamic theory of describing experimental results from volume holography," *Journal of Applied Physics*, vol. 47, no. 1, pp. 190–199, 1976.

[30] D. M. Kim, R. R. Shah, T. A. Rabson, and F. K. Tittel, "Nonlinear dynamic theory for photorefractive phase hologram formation," *Applied Physics Letters*, vol. 28, no. 6, pp. 338–340, 1976.

[31] K. Blotekjaer, "Theory of hologram formation in photorefractive media," *Journal of Applied Physics*, vol. 48, no. 6, pp. 2495–2501, 1977.

[32] M. G. Moharam and L. Young, "Hologram writing by the photorefractive effect," *Journal of Applied Physics*, vol. 48, no. 8, pp. 3230–3236, 1977.

[33] M. G. Moharam and L. Young, "Recording and optical erasure of holograms stored by the photorefractive effect in lithium niobate," *Applied Optics*, vol. 17, no. 17, pp. 2773–2778, 1978.

[34] 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*, vol. 22, pp. 949–960, 1979.

[35] F. S. Chen, J. T. LaMacchia, and D. B. Fraser, "Holographic storage in lithium niobate," *Applied Physics Letters*, vol. 13, no. 7, pp. 223–225, 1968.

[36] J. J. Amodei and D. L. Staebler, "Holographic recording in lithium niobate," *RCA Review*, vol. 33, pp. 71–93, 1972.

[37] J. Bordogna, S. A. Keneman, and J. J. Amodei, "Recyclable holographic storage media," *RCA Review*, vol. 33, pp. 227–247, 1972.

[38] D. L. Staebler, J. J. Amodei, and W. Philips, "Multiple storage of thick phase holograms in $LiNbO_3$," in *VII International Quantum Electronics Conference*, (Montreal), May 1972.

[39] F. H. Mok, "Angle–multiplexed storage of 5000 holograms in lithium niobate," *Optics Letters*, vol. 18, pp. 915–917, June 1991.

[40] F. H. Mok, G. W. Burr, and D. Psaltis, "System metric for holographic memory systems," *Optics Letters*, vol. 21, no. 12, pp. 896–898, 1996.

[41] G. Burr, *Volume holographic storage using the* 90° *geometry.* PhD thesis, California Institute of Technology, 1996.

[42] P. J. van Heerden, "A new optical method of storing and retrieving information," *Applied Optics*, vol. 2, no. 4, pp. 387–392, 1963.

[43] F. T. S. Yu, S. Wu, A. W. Mayers, and S. Rajan, "Wavelength multiplexed reflection matched spatial filters using $LiNbO_3$," *Optics Communications*, vol. 81, no. 6, pp. 343–347, 1991.

[44] G. A. Rakuljic, V. Leyva, and A. Yariv, "Optical data storage by using orthogonal wavelength–multiplexed volume holograms," *Optics Letters*, vol. 17, no. 20, pp. 1471–1473, 1992.

[45] J. Rosen, M. Segev, and A. Yariv, "Wavelength–multiplexed computer–generated volume holography," *Optics Letters*, vol. 18, no. 9, pp. 744–746, 1993.

[46] S. Yin, H. Zhou, F. Zhao, M. Wen, Z. Zhang, and F. T. S. Yu, "Wavelength–multiplexed holographic storage in a sensitive photorefractive crystal using a visible–light tunable diode laser," *Optics Communications*, vol. 101, pp. 317–321, 1993.

[47] 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 Communications*, vol. 103, pp. 59–62, 1993.

[48] J. T. LaMacchia and D. L. White, "Coded multiple exposure holograms," *Applied Optics*, vol. 7, no. 1, pp. 91–94, 1977.

[49] T. F. Krile, M. O. Hagler, W. D. Redus, and J. F. Walkup, "Multiplex holography with chirp–modulated binary phase–coded reference–beam masks," *Applied Optics*, vol. 18, no. 1, pp. 52–56, 1979.

[50] J. E. Ford, Y. Fainman, and S. H. Lee, "Array interconnection by phase–coded optical correlation," *Optics Letters*, vol. 15, no. 19, pp. 1088–1090, 1990.

[51] C. Denz, G. Pauliat, G. Roosen, and T. Tschudi, "Volume hologram multiplexing using a deterministic phase encoding method," *Optics Communications*, vol. 85, pp. 171–176, 1991.

[52] C. Alves, G. Pauliat, and G. Roosen, "Dynamic phase–encoding storage of 64 images in $BaTiO_3$," *Optics Letters*, vol. 19, no. 22, pp. 1894–1896, 1994.

[53] 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*, vol. 19, no. 14, pp. 1079–1081, 1994.

[54] J. F. Heanue, M. C. Bashaw, and L. Hesselink, "Encrypted holographic data storage based on orthogonal–phase–code multiplexing," *Applied Optics*, vol. 34, no. 26, pp. 6012–6015, 1995.

[55] D. Psaltis, M. levene, A. Pu, G. Barbastathis, and K. Curtis, "Holographic storage using shift multiplexing," *Optics Letters*, vol. 20, no. 7, pp. 782–784, 1995.

[56] D. Psaltis, D. Brady, and K. Wagner, "Adaptive optical networks using photorefractive crystals," *Applied Optics*, vol. 27, no. 9, pp. 1752–1759, 1988.

[57] D. Psaltis, D. Brady, X. G. Gu, and S. Lin, "Holography in artificial neural networks," *Nature*, vol. 343, no. 6526, 1990.

[58] K. Curtis, A. Pu, and D. Psaltis, "Method for holographic storage using peristropic multiplexing," *Optics Letters*, vol. 19, no. 13, pp. 993–994, 1994.

[59] D. Psaltis, "Class notes: EE133 Optical computing, Caltech," 1994.

[60] W. Burke and P. Sheng, "Crosstalk noise from multiple thick–phase holograms," *Journal of Applied Physics*, vol. 48, no. 2, pp. 681–685, 1977.

[61] H. Lee, "Cross–talk effects in multiplexed volume holograms," *Optics Letters*, vol. 13, no. 10, pp. 874–876, 1988.

[62] H. Lee, X.-G. Gu, and D. Psaltis, "Volume holographic interconnections with maximal capacity and minimal cross talk," *Journal of Applied Physics*, vol. 65, pp. 2191–2194, March 1989.

[63] 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*, vol. 9, no. 11, pp. 1–6, 1993.

[64] C. Gu, A. Chiou, and J. Hong, "Cross–talk noise in photorefractive interconnection," *Applied Optics*, vol. 32, no. 8, pp. 1437–1440, 1993.

[65] G. P. Nordin and O. Asthana, "Effects of cross talk on fidelity in page–oriented volume holographic optical data storage," *Optics Letters*, vol. 18, no. 18, pp. 1553–1555, 1993.

[66] K. Curtis and D. Psaltis, "Cross talk for angle– and wavelength–multiplexed image plane holograms," *Optics Letters*, vol. 19, no. 21, pp. 1774–1776, 1994.

[67] X. Yi, P. Yeh, and C. Gu, "Statistical analysis of cross–talk noise and storage capacity in volume holographic memory," *Optics Letters*, vol. 19, no. 19, pp. 1580–1582, 1994.

[68] G. W. Burr, F. H. Mok, and D. Psaltis, "Spatial and angle multiplexing using the Long Interaction Length Architecture," in *OSA Annual Meeting, Paper TuH6*, (Toronto), 1993.

[69] G. W. Burr, F. H. Mok, and D. Psaltis, "Angle and space multiplexed holographic storage using the 90° geometry," *Optics Communications*, vol. 117, pp. 49–55, 1995.

[70] F. H. Mok, D. Psaltis, and G. W. Burr, "Spatially– and Angle– multiplexed holographic random access memory," *SPIE Proceedings*, vol. 1773c, p. 1, 1992.

[71] H. S. Li, *Photorefractive 3–D disks for optical data storage and artificial neural networks*. PhD thesis, California Institute of Technology, 1994.

[72] X. An and D. Psaltis, "Experimental characterization of an angle–multiplexed holographic memory," *Optics Letters*, vol. 20, no. 18, pp. 1913–1915, 1995.

[73] C. Gu, J. Hong, H. Y. Li, D. Psaltis, and P. Yeh, "Dynamics of grating formation in photovoltaic media," *Journal of Applied Physics*, vol. 69, no. 3, pp. 1167–1172, 1991.

[74] R. A. Rupp, R. Sommerfeldt, K. H. Ringhofer, and E. Kratzig, "Space charge field limitations in photorefractive $LiNbO_3$:Fe crystals," *Applied Physics B*, vol. 51, pp. 364–370, 1990.

[75] 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*, vol. 27, no. 3, pp. 509–515, 1991.

[76] A. M. Glass, D. von der Linde, and T. J. Negran *Applied Physics Letters*, vol. 25, p. 233, 1974.

[77] A. M. Glass, D. von der Linde, D. H. Auston, and T. J. Negran *Journal of Electronics Matterts*, vol. 24, p. 915, 1975.

[78] I. McMichael, W. Christian, and J. Hong, "Compact volume holographic memory system with rapid acoustooptic addressing," in *OSA Annual Meeting, Paper THB2*, (Portland, OR), 1995.

[79] X. An and D. Psaltis, "Angle–multiplexed holographic storage using acousto–optic deflectors," in *OSA Annual Meeting, Paper MP6*, (Portland, OR), 1995.

[80] F. Ito and K. Kitayama, "Real–time holographic storage of a temporal bit sequence by using angular multiple recording of spectral components," *Optics Letters*, vol. 17, pp. 1152–1154, August 1992.

[81] A. Yariv, *Optical Electronics*, ch. 12, pp. 464–465. Saunders College Publishing, 4 ed., 1991.

[82] A. Yariv, *Optical Electronics*, ch. 12, pp. 466–473. Saunders College Publishing, 4 ed., 1991.

[83] A. Yariv, *Optical Electronics*, ch. 9, pp. 330–332. Saunders College Publishing, 4 ed., 1991.

[84] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*, ch. 5, pp. 273–281. Pergamon Press, Oxford, NY, 6 ed., 1989.

[85] G. Burr, X. An, D. Psaltis, and F. Mok, "Large–scale rapid access holographic memory," in *1995 Optical Data Storage Meeting, SPIE Technical Digest Series*, vol. 2514, (Portland, OR), pp. 363–371, 1995.

[86] A. D. Kersey, D. A. Jackson, and M. Corke *Electronics Letters*, vol. 18, pp. 392–395, 1982.

[87] J. Frejlich, "Fringe–locked running hologram and multiple photoactive species in $Bi_{12}TiO_{20}$," *Journal of Applied Physics*, vol. 68, pp. 3104–3109, October 1990.

[88] A. A. Freschi and J. Frejlich, "Stabilized photorefractive modulation recording beyond 100% diffraction efficiency in $LiNbO_3$:Fe crystals," *Journal of the Optical Society of America B*, vol. 11, pp. 1837–1841, September 1994.

[89] A. A. Freschi and J. Frejlich, "Adjustable phase control in stabilised interferometry," *Optics Letters*, vol. 20, pp. 635–637, March 1995.

[90] J. Frejlich, P. M. Garcia, and A. A. Freschi, "Deeply modulated stabilized photorefractive recording in $LiNbO_3$:Fe," *Optical Materials*, vol. 4, pp. 410–413, January 1995.

[91] S. Bian and J. Frejlich, "Photorefractive response time measurement in GaAs crystal by phase modulation in two–wave mixing," *Optics Letters*, vol. 19, pp. 1702–1704, November 1994.

[92] E. A. Barbosa, L. S. Troncoso, and J. Frejlich, "Actively stabilised electro–optic phase modulator," *Electronics Letters*, vol. 31, pp. 127–128, January 1995.

[93] J. P. Huignaid, J. P. Herriau, and F. Micheron, "Coherent selective erasure of superimposed volume holograms in $LiNbO_3$," *Applied Physic Letters*, vol. 26, pp. 256–258, March 1975.

[94] R. Magnusson, X. Wang, T. D. Black, and L. N. Tello, "Diffraction efficiency dependence of holographic subtraction interferometry in Fe:$LiNbO_3$," *Applied Optics*, vol. 31, pp. 3350–3353, June 1992.

[95] P. Yeh, T. Y. Chang, and P. H. Beckwith, "Real–time optical image subtraction using dynamic holographic interference in photorefractive media," *Optics Letter*, vol. 13, pp. 586–588, 1988.

[96] H. Sasaki, J. Ma, Y. Fainman, S. H. Lee, and Y. Taketomi, "Fast update of dynamic photorefractive optical memory," *Optics Letters*, vol. 17, pp. 1468–1470, October 1992.

[97] Y. Qiao and D. Psaltis, "Sampled dynamic holographic memory," *Optical Materials*, vol. 4, pp. 410–413, January 1995.

[98] A. Roy and K. Singh, "Higher order self–diffraction in absorptive photorefractive medium; transmission geometry," *Optics Communications*, vol. 87, pp. 351–357, 1992.

[99] S. Tao, Z. H. Song, and D. R. Selviah, "Bragg–shift of holographic gratings in photorefractive Fe:$LiNbO_3$ crystals," *Optics Communications*, vol. 108, pp. 144–152, 1994.

[100] K. Buse, S. Kamper, J. Frejlich, R. Pankrath, and K. H. Ringhofer, "Tilting of holograms in photorefractive $Sr_{0.61}Ba_{0.39}Nb_2O_6$ crystals by self–diffraction," *Optics Letters*, vol. 20, pp. 2249–2251, November 1995.

[101] D. G. Gray, M. G. Moharam, and T. M. Ayres, "Heterodyne technique for the direct measurement of the amplitude and phase of photorefractive space-charge fields," *Journal of the Optical Society of America B*, vol. 11, pp. 470–475, March 1994.

[102] W. Lawler, C. J. Sherman, and M. G. Moharam, "Direct measurement of the amplitude and the phase of photorefractive fields in $KNbO_3$:Ta and $BaTiO_3$," *Journal of the Optical Society of America B*, vol. 8, pp. 2190–2195, October 1991.

[103] J. J. Amodei and D. L. Staebler, "Holographic pattern fixing in electro-optic crystals," *Applied Physics Letters*, vol. 18, no. 12, pp. 540–542, 1971.

[104] W. Bollmann and H. J. Stohr, "Incorporation and mobility of OH^- ions in $LiNbO_3$ crystals," *Physics Status Solidi (a)*, vol. 39, pp. 477–484, 1977.

[105] H. Vormann, G. Weber, S. Kapphan, and E. Kratzig, "Hydrogen as origin of thermal fixing in $LiNbO_3$:Fe," *Solid State Communications*, vol. 40, pp. 123–125, August 1981.

[106] S. Klauer, M. Wohlecke, and S. Kapphan, "Isotope effect of protonic conductivity in $LiNbO_3$," *Radiation Effects and Defects in Solids*, vol. 119-121, pp. 699–704, 1991.

[107] W. Meyer, P. Wurfel, R. Munser, and G. Muller-Vogt, "Kinetics of fixation of phase holograms in $LiNbO_3$," *Physics Status Solidi (a)*, vol. 53, pp. 171–180, February 1979.

[108] P. Hertel, K. H. Ringhofer, and R. Sommerfeldt, "Theory of thermal hologram fixing and application to $LiNbO_3$:Cu," *Physics Status Solidi (a)*, vol. 104, pp. 855–862, 1987.

[109] 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*, vol. 7, pp. 2317–2322, December 1990.

[110] A. Yariv, S. Orlov, G. Rakuljic, and V. Leyva, "Holographic fixing, readout, and storage dynamics in photorefractive materials," *Optics Letters*, vol. 20, pp. 1334–1336, June 1996.

[111] S. Orlov, *Holographic storage dynamics, phase conjugation, and nonlinear optics in photorefractive materials.* PhD thesis, California Institute of technology, 1996.

[112] 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*, vol. 26, pp. 182–184, February 1975.

[113] J. F. Heanue, M. C. Bashaw, A. J. Daiber, R. Snyder, and L. Hesselink, "Digital holographic storage system incorporating thermal fixing in lithium niobate," *Optics Letters*, vol. 21, pp. 1615–1617, October 1996.

[114] X. An and D. Psaltis, "Thermal fixing of 10,000 holograms in $LiNbO_3$:Fe," in *OSA Annual Meeting, Paper MAAA5*, (Rochester, NY), 1996.

[115] W. Bollmann, "Diffusion of hydrogen (OH^- ions) in $LiNbO_3$ crystals," *Physics Status Solidi (a)*, vol. 104, pp. 643–648, 1987.

[116] R. Matull and R. A. Rupp, "Microphotometric investigation of fixed holograms," *Journal of Physics D: Applied Physics*, vol. 21, no. 11, pp. 1556–1565, 1988.

[117] A. Yariv, *Optical Electronics*, ch. 18, pp. 640–649. Saunders College Publishing, 4 ed., 1991.

[118] 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*, vol. 3, pp. 5399–5406, March 1991.

[119] W. Phillips and D. L. Staebler, "Control of the Fe^{2+} concentration in iron-doped lithium niobate," *Journal of Electronic Materials*, vol. 3, no. 2, pp. 601–617, 1974.

[120] G. A. Alphonse and W. Phillips, "Iron-doped Lithium Niobate as a read-write holographic storage medium," *RCA Review*, vol. 37, pp. 184–205, 1976.

[121] R. S. Weis and T. K. Gaylord, "Lithium niobat: summary of physical properties and crystal structure," *Applied Physics A*, vol. 37, pp. 191–203, 1985.

[122] W. Bollmann, K. Schlothauer, and J. Zogal *Krist. & Technik (Germany)*, vol. 11, no. 12, pp. 1327–1332, 1976.

[123] L. Kovacs and K. Polgar, "Conductivity of lithium niobate," *EMIS Datareview RN=16037*, pp. 109–114, August 1989.

[124] K. Buse, S. Breer, K. Peithmann, S. Kapphan, M. Gao, and E. Kratzig, "Origin of thermal fixing in photorefractive lithium niobate crystals," *Physical Review B*, vol. 56, pp. 1225–1236, July 1997.

[125] R. Sommerfeldt, R. A. Rupp, H. Vormann, and E. Kratzig, "Thermal fixing of volume phase holograms in $LiNbO_3$:Cu," *Physics Status Solidi (a)*, vol. 99, pp. K15–K19, 1987.

[126] G. Montemezzani and P. Gunter, "Thermal hologram fixing in pure and doped $KNbO_3$ crystals," *Journal of the Optical Society of America B*, vol. 7, pp. 2323–2328, December 1990.

[127] G. Montemezzani, M. Zgonik, and P. Gunter, "Photorefractive charge compensation at elevated temperatures and application to $KNbO_3$," *Journal of the Optical Society of America B*, vol. 10, pp. 171–185, February 1993.

[128] D. L. Staebler and J. J. Amodei, "Thermally fixed holograms in $LiNbO_3$," *Ferroelectrics*, vol. 3, pp. 107–113, February 1972.

[129] N. B. Anghert, V. A. Pashkov, and N. M. Solov'yeva *Zh. Eksp. Teor. Fiz*, vol. 26, pp. 1666–1672, 1972.

[130] K. G. Belabaev, V. B. Markov, and S. G. Odulov *Ukr. Fiz. Zh.*, vol. 21, p. 1820, 1976.

[131] Y. Ohmori, M. Yamaguchi, K. Yoshino, and Y. Inuishi *Japan Journal of Applied Physics*, vol. 15, pp. 2263–2264, 1976.

[132] W. Josch, R. Munser, W. Ruppel, and P. Wurfel, "The photovoltaic effect and the charge transport in $LiNbO_3$," *Ferroelectrics*, vol. 21, pp. 623–625, 1977.

[133] I. B. Barkan, E. M. Baskin, and M. V. Entin *Physics Status Solidi (a)*, vol. 59, pp. K97–K102, May 1980.

[134] D. F. Balsa and A. V. Bobyl *Soviet Physics: -Solid State*, vol. 27, pp. 313–316, 1985.

[135] R. Shah, D. M. Kim, T. A. Rabson, and F. K. Tittel, "Characterization of iron–doped lithium niobate for holographic storage applications," *Journal of Applied Physics*, vol. 47, pp. 5421–5431, December 1976.

[136] S. Klauer, M. Wohlecke, and S. Kapphan, "Influence of H-D isotopic substitution on the protonic conductivity of $LiNbO_3$," *Physical Review B*, vol. 45, pp. 2786–2799, February 1992.

[137] F. Jariego and F. Agullo-Lopez, "Monotonic versus oscillatory behavior during holographic writing in photorefractive photovoltaic materials," *Optics Communications*, vol. 76, no. 2, pp. 169–172, 1990.

[138] B. Tatian *Applied Optics*, vol. 23, pp. 4477–4485, December 1984.

[139] M. V. Hobden and J. Warner *Physics Letters*, vol. 22, pp. 243–244, August 1966.

[140] D. F. Nelson and R. M. Mikulyak *Journal of Applied Physics*, vol. 45, pp. 3688–3689, August 1974.

[141] C. J. G. Kirkby, "Refractive index of $LiNbO_3$: wavelength dependence," *EMIS Datareview RN=16002*, July 1988.

[142] N. N. Sirota, V. P. Yarunichev, and I. T. Bodnar *Journal of Applied Spectroscopy*, vol. 31, pp. 1165–1167, September 1979.

[143] G. Bayer *Proc. Br. Ceram. Soc.*, vol. 22, pp. 39–53, 1973.

[144] Y. S. Kim and R. T. Smith *Journal of Applied Physics*, vol. 40, pp. 4637–4641, 1969.

[145] F. C. Lissalde and J. C. Peuzin *J. Appl. Crystallogr.*, vol. 7, pp. 572–576, December 1974.

[146] S. C. Abrahams, H. J. Levinstein, and J. M. Reddy *J. Phys. Chem. Solids*, vol. 27, pp. 1019–1026, 1966.

[147] Z. I. Shapiro, S. A. Fedulov, Y. N. Venevtsev, and L. G. Rigerman *Sov. Phys. Crystallogr.*, vol. 10, pp. 725–728, 1966.

[148] I. G. Ismailzade, V. I. Nesterenko, and F. A. Mirishli *Sov. Phys. Crystallogr.*, vol. 13, pp. 25–28, 1968.

[149] P. K. Gallagher and H. M. O'Bryan *J. Am. Ceram. Soc.*, vol. 68, pp. 147–150, March 1985.

[150] P. K. Gallagher, H. M. O'Bryan, E. M. Gyorgy, and J. T. Krause *Ferroelectrics*, vol. 75, no. 1-2, pp. 71–77, 1987.

[151] V. V. Zhdanova, V. P. Klyeuv, V. V. Lemanov, I. A. Smirnov, and V. V. Tikhonov *Soviet Physics:–Solid State*, vol. 10, pp. 1360–1362, 1968.

[152] G. A. Smolenski, N. N. Krainik, N. P. Khuchua, V. V. Zhdanova, and I. E. Mylnikova *Physics Status Solidi.*, vol. 13, pp. 309–314, 1966.

[153] N. N. Krainik, N. P. Khuchua, V. V. Zhdanova, I. E. Mylnikova, and N. N. Parfenova *Proceeding of International Meeting on Ferroelectrics*, vol. 1, pp. 377–386, 1966.

[154] T. J. Hsu, H. M. Sun, and Y. N. Wang *Ko Hsueh Tung Pao (China)*, vol. 25, pp. 586–588, 1980.

[155] J. S. Browder and S. S. Ballard *Applied Optics*, vol. 16, pp. 3214–3217, December 1977.

[156] J. S. Ballard, S. E. Brown, and J. S. Browder *Applied Optics*, vol. 17, pp. 1152–1154, April 1978.

[157] D. Taylor, "Thermal expansion coefficient of lithium niobate," *EMIS Datareview RN=16023*, pp. 44–48, September 1988.

[158] A. Pu and D. Psaltis, "High-density recording in photopolymer-based 3-dimensional disks," *Applied Optics*, vol. 35, pp. 2389–2398, May 1996.

[159] E. Chuang and D. Psaltis, "Storage of 1000 holograms with use of a dual-wavelength method," *Applied Optics*, vol. 36, pp. 8445–8454, November 1997.

[160] M. P. Bernal, G. W. Burr, H. Coufal, and M. Quintanilla, "Optimizing the CCD fill factor to balance inter-pixel crosstalk and thermal noise in holographic data storage," in *OSA Annual Meeting, Paper ThB2*, (Long Beach, CA), 1997.

[161] G. W. Burr, "Elimination of nonrandom noise in holographic data storage," in *OSA Annual Meeting, Paper ThB5*, (Long Beach, CA), 1997.

[162] L. Hyuk, X. Gu, and D. Psaltis, "Volume holographic interconnections with maximal capacity and minimal cross talk," *Journal of Applied Physics*, vol. 65, pp. 2191–2194, March 1989.

[163] B. Sturman and V. Fridkin, *The Photovoltaic and Photorefractive Effects in Noncentrosymmetric Materials*, vol. 8 of *Ferroelectricity and related phenomena*. Gordon and Breach Science Publishers, 1992.

[164] A. Yariv, *Optical Electronics*, ch. 9. Saunders College Publishing, 4 ed., 1991.

[165] 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.