

High-density data storage using proximal probe techniques

by H. J. Mamin
B. D. Terris
L. S. Fan
S. Hoen
R. C. Barrett
D. Rugar

We describe some of the achievements and problems associated with proximal probe-based approaches to high-density data storage. While STM-based methods have demonstrated spectacular areal densities dwarfing anything achievable with today's storage technologies, reliability and data rate issues present serious obstacles. These problems have led us to focus on techniques based on AFM and near-field optics. First, we have developed a thermomechanical writing scheme using an AFM tip. We have addressed many of the practical issues involved, including data rate. With custom low-mass cantilevers, we have demonstrated readback on real data with a data rate of 1.2 Mb/s. We have also pursued nontopographic storage techniques based on charge storage in nitride-oxide semiconductor structures and near-field optical storage. These techniques should be able to achieve densities comparable to those reached with the AFM scheme, with the added advantage that they are fast and reversible. Although it is not yet clear whether any of these probe-based approaches can ever be made practical, they do represent potential

pathways to the higher densities that will be needed in the decades ahead.

Introduction

Scanning probe microscopy has long been seen as a possible basis for an ultrahigh-density data storage device [1]. Numerous means have been demonstrated with both the scanning tunneling microscope (STM) [2-13] and the atomic force microscope (AFM) [14-18] to write and detect surface features on the nanometer or subnanometer scale. While many of these modification techniques are worthy of study in their own right, the potential for application to data storage has undoubtedly helped to generate interest in this area. One of the key forces driving the rapid advances in data storage has been the ever-increasing areal density. Scanning probes, with their ability to modify surfaces all the way down to the atomic scale, offer a path for this trend to continue far beyond the limits of conventional technologies.

This approach has some formidable obstacles to overcome in order to become practical, however. One of the key difficulties is the extremely low data rates typically achieved with scanning probe techniques. Overall reliability is the other main concern. In our laboratory, we

©Copyright 1995 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

0018-8646/95/\$3.00 © 1995 IBM

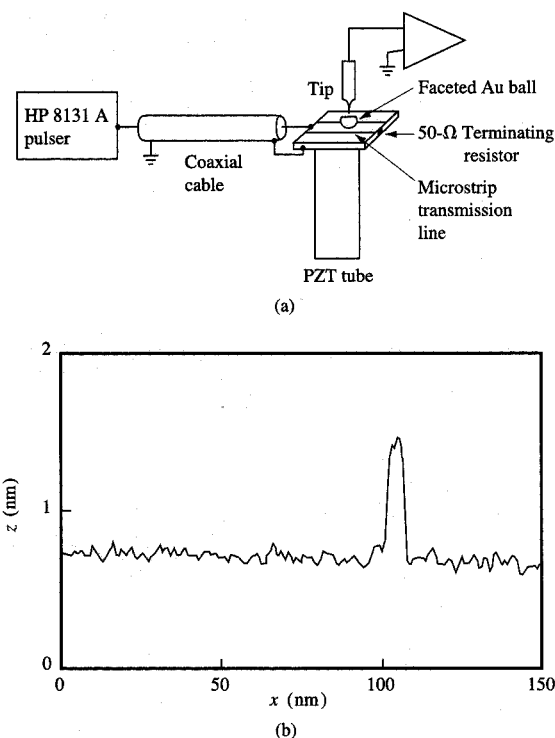


Figure 1

(a) Setup for applying sub-ns voltage pulses between a Au STM tip and Au sample. The microstrip transmission line impedance was verified to be 50 Ω out to 3 GHz. (b) Cross section of Au mound formed with a 5-V, 900-ps voltage pulse. Typical mound dimensions were 100–200 Å across and 10 Å high; x is the lateral distance in the plane of the surface, and z is the height of the surface [21].

have investigated a number of proximal probe techniques with these concerns in mind.

In this paper we present an overview of some of these approaches. We do not attempt to cover the entire field, but rather focus only on the work undertaken at the IBM Almaden Research Center. We first examine a particular STM-based recording scheme, deposition from a gold STM tip. That leads to some general comments about the difficulties with STM-based recording. We then discuss several other proximal probe techniques: AFM thermomechanical recording, charge storage in nitride-oxide-semiconductor structures, and near-field optical storage with a solid immersion lens. The discussion progresses from schemes which give the highest density but are least practical, to those which offer a more modest improvement in density but are most viable.

STM gold deposition

Perhaps the ultimate demonstration of scanning probe data storage is STM writing with single-atom “bits,” as performed at low temperatures by Eigler and Schweizer [12]. Such single-atom writing was not seriously proposed as a data storage technology; it has the obvious limitation that it is many orders of magnitude too slow. From a practical standpoint, it also has the drawback that it operates under rather special conditions: ultrahigh vacuum and low temperature. Even at this exploratory stage, it would be desirable to find a process which is at least potentially fast and operates under ambient conditions.

We have developed one such technique, writing with a gold STM tip, that does operate in air at room temperature, and is potentially very fast [8]. It is among the more reliable STM writing mechanisms: With a good tip, a writing probability of greater than 95% can be obtained, and a probability greater than 50% can be achieved with most tips. The marks are also reasonably long-lived. We have observed that even after exposure to mildly elevated temperature (up to 65°C), the marks remained over a period of weeks, with at most a 10% reduction in height [19].

The writing process itself is extremely fast. We have devised a method for applying sub-ns voltage pulses to the faceted Au sample, as shown in Figure 1(a). We successfully deposited mounds with pulses as short as 700 ps. The cross section of a mound deposited with a 900-ps pulse is shown in Figure 1(b). We believe the mechanism is a form of field evaporation, though that remains the subject of some debate [20, 21].

Although the time to write a single mark is quite fast, the overall data rate is limited by the time it takes to move the tip over the surface to read or write the next mark. A pattern containing a few dozen marks, for example, typically takes 5–10 minutes to read and write. This arises from the inherent limitation in the speed with which the STM can maintain the tip-sample separation constant. If the separation becomes too large, the tunnel current is lost, while if the separation is too small, the tip crashes. This separation is normally maintained with the STM servo to within better than 1 Å. If one started with an atomically flat sample and could write marks which were extremely uniform in height, it is possible to think of flying the tip over the tops of the marks without the need to servo the tip at high frequencies. Barring this situation, however, the bandwidth of the servo will always limit the data rate of the STM. Partly for this reason, we have explored the degree to which the bandwidth of the STM servo loop can be increased. With the use of some novel mechanics and a two-stage servo, we were able to obtain a servo with a response time of less than 10 μ s [22]. This allowed us to scan the tip over stepped gold surfaces as fast as 1 mm/s,

suggesting that this STM should be capable of data rates of the order of 100 kHz.

With this fast STM, we were able to write the letters "IBM" in roughly 1 s, as shown in **Figure 2**. The substrate for this writing was a faceted gold ball, which has rather large atomically flat terraces. Because the sample was more massive than the Au-on-mica samples we otherwise used, we were not able to get the full bandwidth in this case. Nonetheless, at 1 s, this writing was considerably faster than had previously been achieved.

Although we made significant improvements in data rate with the STM, it is still far from where it has to be in order to be a viable storage technology. While some further improvements in servo bandwidth are no doubt possible, they are not easy to obtain. They require ever smaller piezoelectric actuators and faster electronics. In addition, the reliability still leaves something to be desired, as seen by the missing bits in **Figure 2**.

AFM thermomechanical writing

• Principle

In many ways the AFM presents a more attractive alternative to high-density data storage. Because the tip is on the end of a weak cantilever, any forces generated by a tip crash are greatly reduced. In fact, with a sufficiently soft cantilever one can simply operate with the tip in contact, thereby completely eliminating the need for a high-speed servo. This greatly simplifies the device and should allow for much higher data rates.

We have recently developed a means for reading and writing with an AFM cantilever which achieves densities of about $2\text{--}3 \times 10^{10}$ bits/in.² [15]. This density, though far below the roughly 10^{12} bits/in.² achieved with the gold writing, is still well above the roughly 6×10^8 bits/in.² found in today's magnetic recording products. In this technique, the tip of an AFM cantilever is placed in contact with a rotating surface, as shown in **Figure 3**. The substrate material is polycarbonate, and the writing is done by heating the tip with a pulsed infrared laser. With the tip in contact, the heated tip softens the polymer surface, and the pressure from the tip creates a small pit. Typical pulse parameters are 20-mW and 1- μ s pulses, with a loading force on the tip of roughly 10^{-7} N. Pits as small as 100 nm across and 10 nm deep have been written in this way [15].

In **Figure 4** we show an AFM micrograph of a pattern written into a test sample with this heating technique. The sample, which was a section of a grooved optical disk, was raster-scanned under the tip using piezoelectric scanners, and pits were written by pulsing the laser at the desired locations. This pattern represents a density of about 25 Gb/in.², a more than 20 $\times$ increase in areal density compared to conventional optical recording.

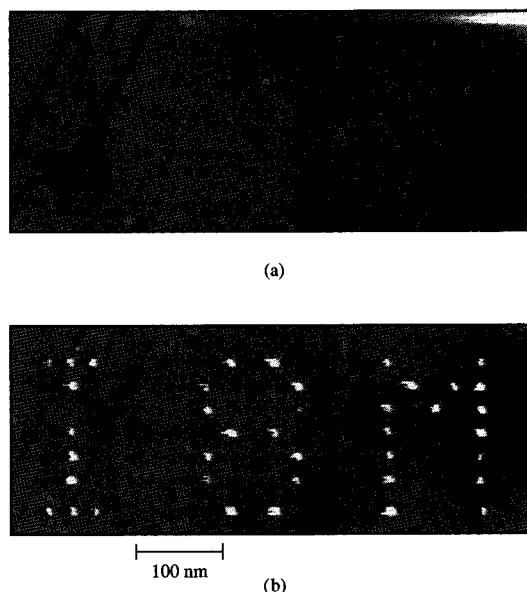


Figure 2

(a) 7000-Å $\times$ 4000-Å region of a faceted Au surface. (b) Same region with the letters IBM written on it. Writing was performed in a single 1-s pass using deposition from a gold STM tip [22].

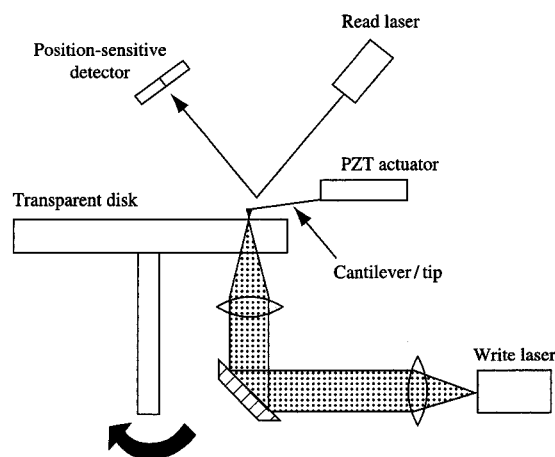


Figure 3

Experimental setup. Focused laser beam propagates through transparent polycarbonate sample and heats the optically absorbing AFM tip. The heated tip softens the substrate, and the local tip pressure creates an indentation. The sample is placed on a precision air-bearing spindle to allow for sample rotation. The PZT actuator is used to maintain generally constant loading force [15].

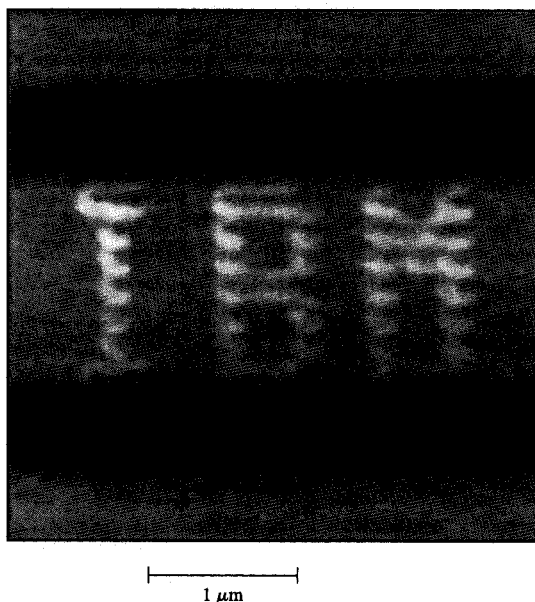


Figure 4

The letters IBM written between the grooves of an optical disk. Each letter would nearly fit inside a typical CD ROM mark [24].

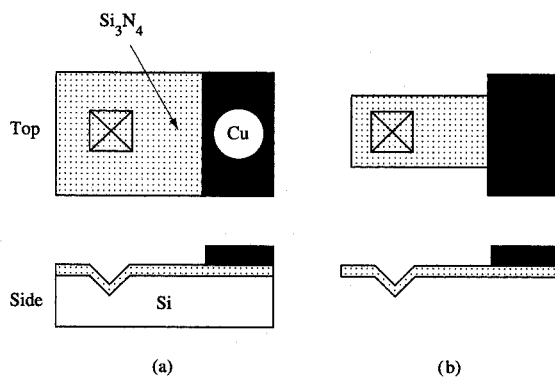


Figure 5

(a) Formation of Si_3N_4 structure. The tip is formed by depositing onto a (111) etch pit in $\text{Si}(100)$. A Cu backing layer is plated before removal of the Si. (b) A freestanding cantilever is formed upon removal of the Si. The 12- μm -thick Cu film provides rigid support over its 0.5-mm dimensions [24].

We detect the marks on the rotating sample with mechanical readback. As the tip rides over the surface of the sample, a topographic feature such as a pit causes a

deflection of the cantilever. This deflection is then detected using a standard optical sensor, as shown in Figure 3. The maximum data rate in this mechanical scheme is related to f_{max} , the maximum frequency at which the tip can track the surface. For a rectangular beam, a simple model which neglects higher-order modes gives

$$f_{\text{max}} = f_0 \left(\frac{2z_L}{h} \right)^{1/2} = 0.46 \left(\frac{F_L}{mh} \right)^{1/2},$$

where f_0 is the resonant frequency of the free cantilever, h is the pit depth, z_L is the load displacement of the cantilever, F_L is the loading force, and m is the mass of the cantilever [15]. Because of wear and signal-to-noise considerations, the loading force and pit depth cannot be varied arbitrarily. Thus, the primary means to improve performance is to reduce the cantilever mass. This is easily understood, as it is the acceleration imparted to the tip that determines the response. Our efforts have therefore been to fabricate smaller cantilevers as a means to improve data rate.

• Low-mass cantilever fabrication

Low-mass cantilevers were custom-fabricated from silicon nitride and had integrated pyramidal tips, following a process similar to that of Albrecht et al. [23, 24]. This process uses anisotropic etching of Si to form pyramidal pits. A low-stress LPCVD Si_3N_4 film is then deposited onto the Si substrate with its pits, as shown in Figure 5. When the Si is removed, this releases the Si_3N_4 film, which becomes a free-standing beam with a pyramidal tip on the end.

One obvious difficulty with this process is that a support is needed at the base of the cantilever. Unfortunately, the underlying Si cannot be retained to provide the support, as that would leave the tips pointed in the wrong direction. That is, the support would be on the same side as the tip, where it would severely interfere with tip access to the sample. In commercial cantilevers, this problem is solved by anodically bonding a glass substrate to the nitride layer before removal of the Si [23]. Precision saw cuts are made in the glass to define an edge for each cantilever. As the glass and wafer are pressed into contact, the nitride cantilevers are positioned to overhang the saw cut. After bonding is complete, the Si substrate is then completely etched away, leaving the glass support. The precision with which one can do the mechanical glass-wafer alignment can affect how well the cantilever length is defined. Slight delamination of the nitride film at the base of the cantilever also can cause variations in the effective cantilever length. With cantilevers of 100 μm or longer, these concerns are usually not significant.

Because we were making shorter cantilevers, however, the length tolerance had to be tighter. It was therefore desirable to have a more precise means to define their

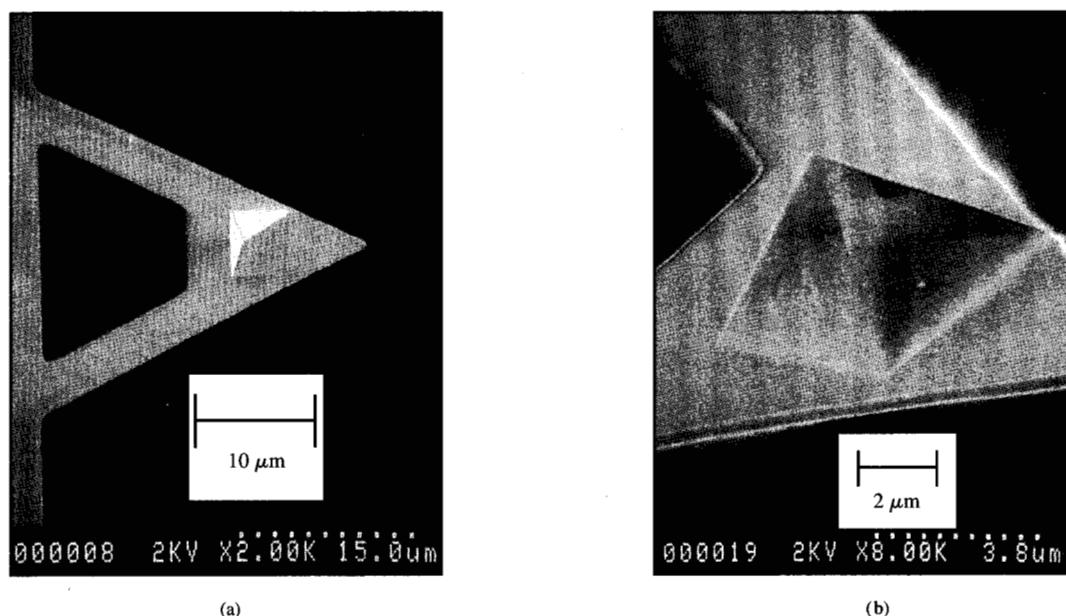


Figure 6

(a) Scanning electron micrograph of a triangular tip of a Si_3N_4 cantilever. Length is roughly $25\text{ }\mu\text{m}$. (b) Scanning electron micrograph of the pyramidal tip. Base of tip is roughly $5\text{ }\mu\text{m}$ [24].

length, one that did not rely on the physical positioning of two objects in contact. Accordingly, a process was developed which defines the length of the cantilever with a photolithographically aligned edge [24]. A low-stress, plated copper film is used as the support in this process. First, a copper seed layer is deposited. Then, a $20\text{-}\mu\text{m}$ -thick photoresist is coated and patterned, which in turn is used as a plating stencil. The side-wall profile of this thick resist must be nearly 90° ; otherwise, the plated film overhangs the cantilever and blocks the laser beam used to detect tip movement. Copper is electrochemically plated, and then the photoresist and seed layer are removed. Finally, the bulk silicon is selectively dissolved in a KOH solution, leaving a freestanding cantilever on a suitably rigid Cu support, as shown in Figure 5. (Copper was used for its good tolerance to KOH etching.) The Cu support and its cantilever are left attached to a silicon frame, ready to be detached and placed on the AFM tester.

Both triangular and rectangular cantilevers were fabricated. The lateral dimensions were chosen so as to make the cantilevers straightforward to fabricate and reasonable to work with in our setup. The thickness was then chosen to give a desired spring constant, of the order of 1 N/m . Using these criteria, we chose the cantilevers to be $25\text{ }\mu\text{m}$ long and $0.35\text{ }\mu\text{m}$ thick, giving a mass of $0.3\text{--}0.6\text{ ng}$,

depending on geometry. For comparison, commercial silicon nitride cantilevers have a mass of at least 10 ng . Figure 6 shows scanning electron micrographs of one of the cantilevers. The cantilever is roughly $25\text{ }\mu\text{m}$ long. The base of the pyramidal tip is about $5\text{ }\mu\text{m}$ across. The point of the pyramid has a radius of well under 100 nm .

The vibrational noise spectrum of a soft cantilever shows a distinct peak at its resonant frequency due to its thermally driven noise [25, 26]. On a solid triangular cantilever, we observed this peak at 1.3 MHz , in excellent agreement with the value of 1.35 MHz calculated for a cantilever without a tip. (We assumed the modulus E to be $1.6 \times 10^{11}\text{ N/m}^2$ and density ρ to be 3300 kg/m^3 [23].) The spring constant k was calculated from the geometry to be 2 N/m . This can also be estimated from the amplitude and width of the thermal noise peak, using the fact that the integrated noise energy is of order $k_B T$ [25]. With this method we obtained 3 N/m , in reasonable agreement. The rectangular cantilevers had calculated spring constants closer to the design point of 1 N/m , with frequencies at about 600 kHz . For comparison, the resonant frequency of a commercial cantilever is typically about 90 kHz , with a spring constant of roughly 0.4 N/m .

Overall, we see that we have increased resonant frequencies by factors of 7 to 14 over the commercial

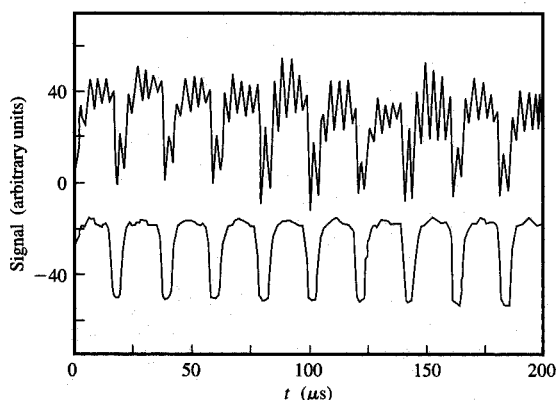


Figure 7

Readback signal on a grooved glass test structure. Upper curve, obtained using a commercial cantilever, shows unwanted ringing, degrading the signal-to-noise ratio. Lower curve, obtained with a new low-mass cantilever, gives an accurate representation of the surface topography [24].

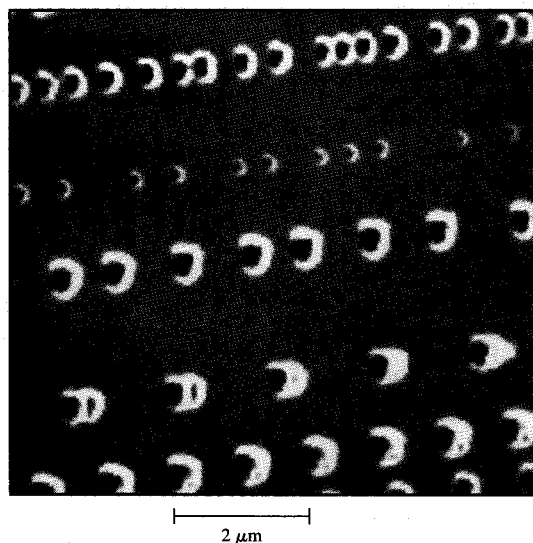


Figure 8

AFM micrograph of tracks written thermomechanically on the rotating sample. Data are encoded by the spacing between the pits. Data are encoded using a (2,7) run-length-limited code, meaning that there are two to seven 0s between every pair of 1s [24].

rectangular levers, and somewhat more for the triangular cantilevers. For stiffer cantilevers, the issue of wear does become a concern. In fact, we originally aimed for spring constants of 1 N/m for the stiffer cantilevers, but deposited a thicker nitride layer than intended. Nonetheless, we found that with a few exceptions, we could generally operate at low enough load that wear was not a major problem, as discussed below. With the same basic design, a thinner nitride layer would of course give softer cantilevers and retain a significant frequency advantage compared to the commercial cantilevers.

• Recording performance

The new low-mass cantilevers gave significantly enhanced performance, as shown in **Figure 7**. A grooved glass disk was used as a test sample, where the groove width was 0.8 μm . In the upper trace, taken with the commercial cantilever, the basic groove periodicity was seen, along with some unwanted ringing, corresponding to higher-order cantilever modes. With a low-mass triangular cantilever, however, the tip tracked the surface cleanly without ringing, as shown in the lower trace.

To measure the data rate, we first wrote a series of tracks containing a known data pattern. This data pattern was a pseudorandom 127-bit sequence of 1s and 0s which repeated. A 1 is represented by the presence of a pit, and a 0 by its absence. This leads to a series of pits with variable spacing, as shown in **Figure 8**. The writing was performed at a rate of about 200 Kb/s; its speed was limited by the thermal response time of the tip, currently about 5 μs . The smallest marks for this set of experiments were about 0.3 μm across.

The readback signal from a triangular cantilever on such a track is shown in **Figure 9**, along with the corresponding data pattern. In this case we have pushed the data rate to the limit of this cantilever: One can see that the tip is not always making it all the way down to the bottom of each trough. Nonetheless, the peaks are easily distinguished, and occur where they should whenever there is a 1 in the data pattern. The data rate from this waveform corresponds to 1.25 Mb/s. This is a rate roughly an order of magnitude higher than was previously possible; it is equivalent to standard (1 $\times$) CD-ROM rates.

We can use standard signal processing algorithms to detect the location of the peaks, and then characterize the readback in a number of ways. First, we can simply decode the peaks to get back to our original data, and compare to look for errors. For data rates below 800 Kb/s, which give waveforms with more uniform peak heights, we typically find about 1–3 errors per 1000 bits; i.e., the raw-error rate is of the order of 10^{-3} . This value is high by data storage standards, but is excellent for a scanning probe technique. It also undoubtedly includes contributions from

cantilevers. This comes at the expense of only a modest increase in the spring constant: about 2–3 $\times$ for the

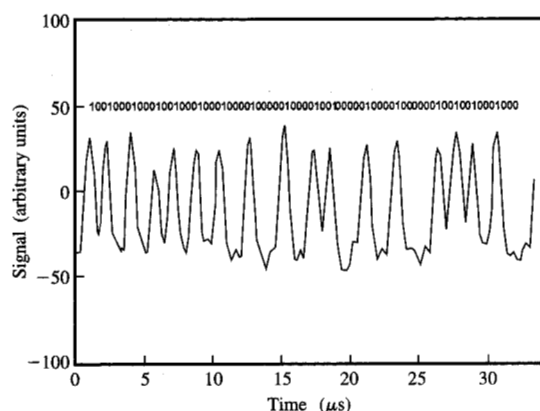


Figure 9

Readback signal on data tracks shown above. The corresponding data pattern is shown above the waveform. The clock period is $0.4 \mu\text{s}$, which with a rate $1/2$ (2,7) code corresponds to a data rate of 1.25 Mb/s [27]. This data rate is an order of magnitude higher than that achieved with conventional cantilevers [24].

pre-existing defects on the disk, which could in principle be reduced or mapped out in advance. Of course, standard error-correction schemes [27] can be used to lower the error rate even further, at the expense of some reduction in effective areal density.

We can also analyze how accurately the peaks are located in time or space to determine the timing error ("jitter"). **Figure 10** shows a histogram of peak spacings for a data pattern containing roughly 20000 marks. Distinct histograms are observed at the expected spacings, with quite narrow widths, suggesting that somewhat higher-density recording is possible. A histogram of the timing error in units of the clock cycle is shown in **Figure 11**. (This is essentially the difference between each mark's intended and actual position.) The width of this histogram indicates how accurately the marks can be both positioned during writing and detected during reading. This value, along with the minimum mark separation, determines the ultimate data density. The standard deviation of this histogram corresponds to about $40 \text{ \AA}$, a small fraction of the mark size, indicating that the positioning is very precise.

• Track servo and wear studies

In magnetic and optical disk drives, tracking servos are used to maintain the head or optical spot on a given data track. Because of the small marks possible in tip-based recording, with track widths well under a micron, the issue of tracking is even more critical. We have largely avoided this problem through the use of the air-bearing spindle.

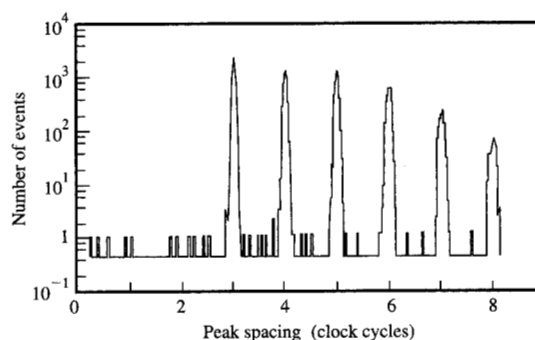


Figure 10

Histogram of peak spacings for data written and read at 170 Kb/s . The peaks of the histograms occur at the integer values between 3 and 8, as they should for a (2,7) code. The events not occurring at integer values are errors; the raw-error rate in this case was roughly 10^{-3} . The histograms are non-overlapping and fairly narrow, indicating low timing jitter. One clock cycle corresponds to $3 \mu\text{s}$ [24].

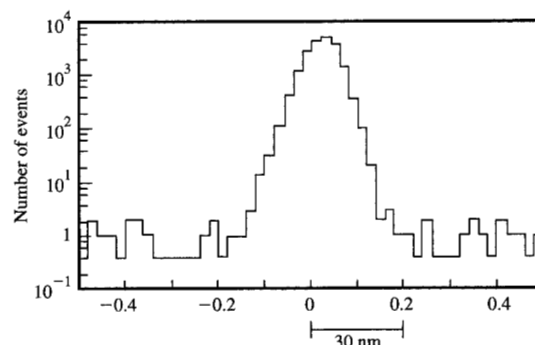


Figure 11

Histogram of phase error determined from the AFM readback. Note that the vertical scale is logarithmic. The standard deviation is roughly $40 \text{ \AA}$. The fact that the mean does not occur precisely at zero is due to the choice of clock and is not significant [24].

The sample rotation on this spindle is very repeatable, with nonrepeatable runout well below $1000 \text{ \AA}$, so that a servo is generally not needed. Both for demonstration purposes and to counter long-term drift, however, we have implemented a tracking servo.

We have studied a number of approaches to this problem. Initially, we hoped that the twist signal from the cantilever could be used to track on the data itself: If the tip were on one side of a mark, the cantilever would twist

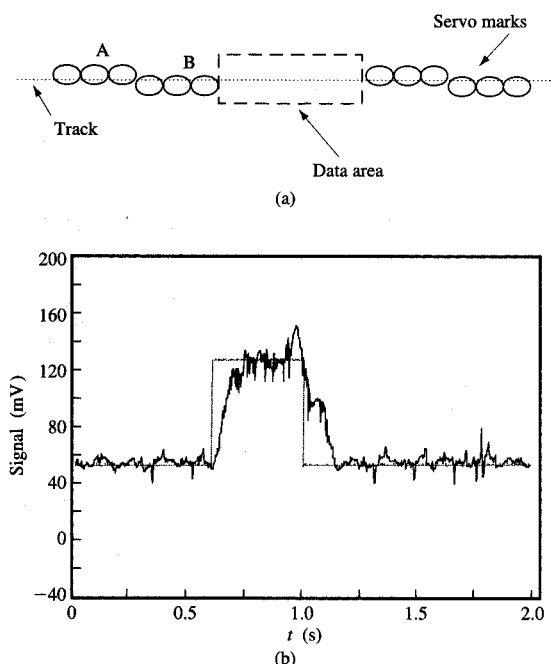


Figure 12

(a) Sector-servo scheme for keeping the tip on a data track, borrowed from magnetic recording. If the tip wanders off-track in the up direction, a signal from the A marks is observed. If it wanders off in the other direction, the signal is from the B marks, which have a different timing. A servo can be used to keep the two signals equal, which keeps the tip in the middle of the track. (b) Servo response (solid) to a square pulse (dotted) fed into the servo loop. The tip position tracks the input with a roughly 100-ms response time.

in one direction, and if it were on the other side, it would twist in the opposite direction. While we were able to verify this general behavior, we found that the signal-to-noise ratio on the twist signal was rather marginal. In principle, narrower rectangular cantilevers (for smaller torsional spring constant) and longer tips (for higher torque) could be fabricated to enhance the signal.

We eventually devised a successful track servo using a scheme borrowed from magnetic recording, called sector-servo. Pre-written servo marks at specific locations around the disk were used to give feedback as to whether the tip was to one side of the track or the other, as shown in **Figure 12(a)**. The data were then written in between the servo marks. We used 100 sets of marks around the track, covering roughly 15% of the track length. **Figure 12(b)** shows the response of the servo to a square pulse input, simulating a sudden shift in the track position. The tip position tracks the input, with a roughly 100-ms response

time. This time is limited primarily by the large mass attached to the piezoelectric actuators. The servo easily compensated for the slow drift, and was able to keep the tip on the data track for a period of days.

With the servo in place, we were able to make much more extended wear studies than would otherwise have been possible. We found that when the tip remained on a single track over roughly two days, with a load of 5×10^{-7} N, a wear track formed. The depth of the wear track was roughly 60 Å. Interestingly, however, the wear made almost no difference in the AFM readback: The readback signal at the end of the two days was almost identical to that at the beginning. When the same experiment was performed at a load of 1×10^{-7} N, no wear was observed. Thus, wear-free operation is possible over this time scale, but only if the load is sufficiently low.

Generally the servo mechanism used to control cantilever loading force had no difficulty maintaining the force below the threshold for wear. At the high velocities needed to measure the highest data rates, however, the servo was not always capable of this, particularly with the stiffer cantilevers. For this reason, wear tracks were occasionally observed for these cantilevers after operation at high velocity.

Finally, it can be noted that a first-order requirement for any type of tracking servo is a uniform envelope to the signal; i.e., the data or servo signal amplitude must be fairly constant around the circumference of the disk. This in turn requires that the writing effectiveness be very uniform around the disk. Unfortunately, this was often quite difficult to achieve because of drifts and variations in sample properties which interfered with the critical alignment between the tip and the focused write laser. Having stable mechanics and samples with minimal tilt or thickness variations improved this situation somewhat. Integrating a heater onto the cantilever would undoubtedly help considerably to produce a more uniform envelope.

• Substrate smoothing

One source of noise which undoubtedly contributes to the errors is the media noise, which in this case is from the intrinsic topographic roughness of the surface of the polycarbonate substrate. AFM studies of the injection-molded substrates reveal that they contain fine structure with an rms roughness of at best about 6 Å. In addition they contain occasional defects, either occlusions or pits in the polymer, of up to a few thousand angstroms across. We have developed a technique to reduce the topographic roughness of the surface as a means to reduce the media contribution to the noise. The smoothing technique is successful at reducing the fine-scale features as well as removing some of the gross defects left by the injection molding process. The technique is based on vapor treatment with the solvent methylene chloride.

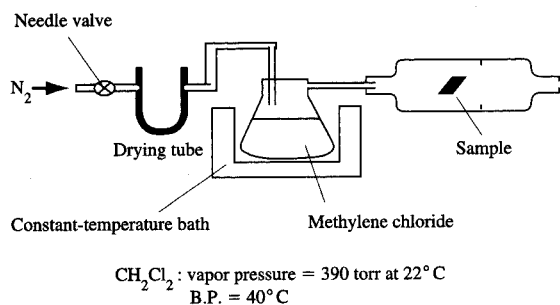
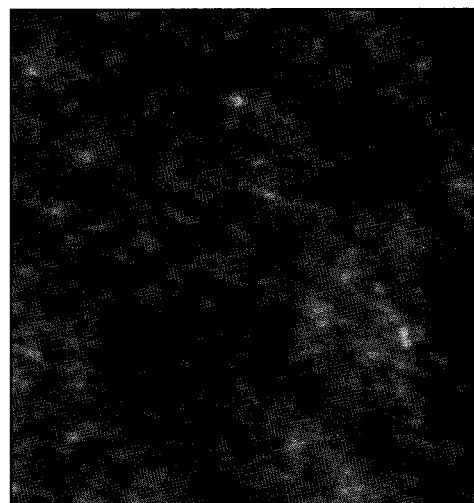


Figure 13

Schematic of vapor transport apparatus for smoothing of polycarbonate substrates. Methylene chloride vapor acts to partially dissolve the surface, which then anneals.

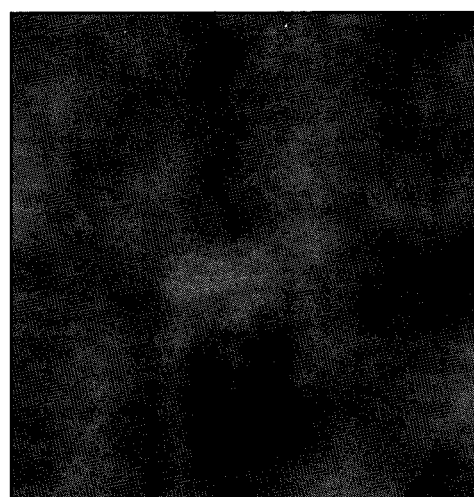
It is a common machinist's trick to place machined pieces of plastic over boiling methylene chloride to smooth the rough edges. To first approximation, the solvent condenses on the surface and dissolves it, and the surface then anneals. When we tried this on a sample of our polycarbonate, we found we could reduce the roughness to 2 Å rms. However, we found that after several days, the samples had developed large craters in the surface. In many cases the samples had crazed as well. Clearly this treatment was too drastic and uncontrolled.

To give a more controlled treatment, we built the apparatus shown in **Figure 13**. Dry nitrogen was flowed over liquid methylene chloride and then over the polymer sample. A hot plate was needed to maintain the solvent at the desired temperature, since the gas flow caused considerable cooling of the liquid. Using this vapor transport system, we could independently control the temperature, flow rate, and time of the treatment. We found that smoothing could take place even without condensation. We propose a fairly simple model for the smoothing process: Solvent molecules from the vapor diffuse a certain distance into the surface of the polymer. Once there, they interpose themselves between the tangled polymer chains and reduce the van der Waals interaction between them. This then has the effect of lowering the glass transition temperature to the point where significant motion is possible at room temperatures. Whatever the process, we find that under typical conditions of five minutes at 26°C and 800 ml/min, we were able to reduce the rms roughness of a smooth polycarbonate sample from 6 Å rms to 3 Å rms, or about a factor of 2 (**Figure 14**). With this treatment, no craters or crazing behavior was observed. As expected, the highest signal-to-noise ratios upon readback were obtained with the treated samples.



(a)

5 μm



(b)

Figure 14

(a) AFM micrograph of untreated polycarbonate. At 6 Å rms roughness, this represents the smoothest injection-molded plastic we could obtain. Gray scale is 40 Å full scale. (b) AFM micrograph of treated polycarbonate. The treatment reduces both small- and large-scale roughness.

• Alternative heating schemes

One of the most troublesome aspects of the AFM thermomechanical writing was making the critical alignment between the focused laser spot and the end of

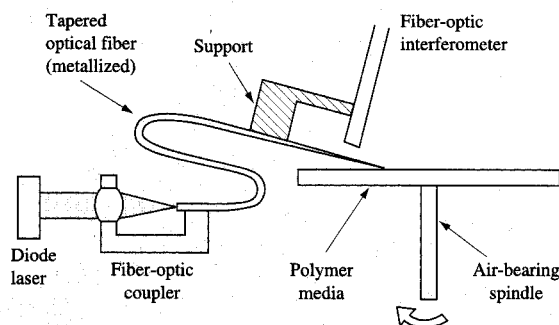


Figure 15

Schematic diagram for writing and reading nanoindentations with a tapered optical fiber. To write, a pulse from the laser diode propagates down the fiber and is absorbed at the tip of the metallized fiber. As this region heats, it softens the underlying polymer and creates a nanoindentation. To read, the interferometer detects the motions of the tapered fiber as it rides over the indentations. The taper metallization conditions are described in the text [28].

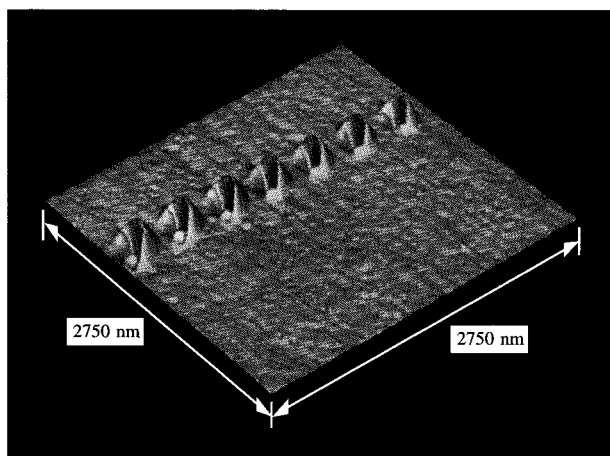


Figure 16

AFM image of a track written on polycarbonate using a tapered optical fiber with a series of 30- μ s laser pulses at 7.5 mW internal to the fiber. The pit size is roughly $300 \times 450 \times 30$ nm³ and reflects the shape of the tip [28].

the tip. We found that to heat the tip to sufficiently high temperature with a reasonable laser power, it was necessary to align the spot on the tip to within about 1 μ m in three dimensions. This alignment required bulky optical components and was imperfect at best. Any alternative

which eliminates the need for such alignment has significant advantages.

One possibility would be to use the cantilever itself as a waveguide to guide light to the tip. It is not out of the question to use a conventional silicon nitride cantilever for this, as it has a fairly high index of refraction of about 2. It is not clear how to launch the light into the cantilever, however, or how it will reach the end of the pyramidal tip. Another approach is to use a single-mode optical fiber and taper it to form a cantilever/tip. We have undertaken this latter approach.

Figure 15 shows our apparatus for reading and writing with this approach [28]. A tapered single-mode fiber was used as the cantilever; the sharp end forms the stylus. We produced the taper by heating and pulling the fiber in the same manner used to make probes for near-field scanning optical microscopy (NSOM) [29]. We chose narrow taper angles of 3–5° to achieve a low spring constant, and achieved tip diameters of 30–200 nm. We then metallized the last 5 μ m of the fiber to absorb the laser light, so as to heat the stylus and in turn the underlying material. In contrast to NSOM, there was no need to leave an aperture at the end for the light. To write, we used an 80-mW infrared diode laser to launch optical heating pulses into the fiber. To read, we used a separate fiber-optic interferometer to measure the motion of the fiber [30].

Figure 16 shows an AFM image of some relatively large indentations, $\sim 300 \times 450$ nm² and 30 nm deep. In general, the fiber produced marks which were comparable in size to those produced by silicon nitride tips. Readback was achieved at a few hundred kHz for the shorter fibers, which had lengths below 200 μ m. Surprisingly, even at loading forces greater than 5×10^{-6} N, we saw little evidence of wear, in contrast to the pyramidal tips at these loads. We believe this is related to the much shallower attack angle ($\sim 20^\circ$) between the disk surface and the leading edge of the fiber tip compared to the pyramidal tip. High spatial resolution is still obtained, since the trailing edge is nearly vertical.

The main limitations to this approach are threefold. First, the thermal time constant for cooling is fairly long because of the large thermal mass which is heated. We found evidence that heating occurs all along the metallized edge of the taper, which is why we metallized only the last few microns. Even then, the time constant was roughly 20 μ s. Second, it is still necessary to have a separate sensor to measure the fiber deflection. The very narrow fiber does not lend itself to a simple integrated deflection sensor as easily as does the planar cantilever structure. We had hoped that the light from the fiber itself could be used to measure deflection, but the signal-to-noise ratio of this approach was inadequate. Finally, the fibers have the

disadvantage that they must be pulled and mounted on an individual basis, as opposed to the batch fabrication possible with cantilevers.

Another approach to heating would be to integrate an electrical heater onto a batch-fabricated AFM cantilever. We are currently working to fabricate silicon nitride cantilevers with integrated NiCr heaters. Finite element calculations¹ indicate that the thermal response time for our design should be of the order of 2 μ s. While this work is underway, we have also undertaken a proof of principle with commercially available piezoresistive cantilevers² [31]. These U-shaped cantilevers are made to allow current to flow through them so that the resistance can be monitored as they flex. In our case, we simply used them as resistors which could be used to heat the tip. They are fairly massive, and it turns out that most of the power is dissipated in the legs of the U, over 100 μ m from the tip. For these reasons, one expects the thermal time constant to be extremely long. Still, we found that with enough power (roughly 0.25 W) and a 100- μ s pulse, we were able to write with them. **Figure 17(a)** shows a mark written in this way, and **Figure 17(b)** shows the associated readback signal when this was performed on a spinning sample. The mark is elongated because of the long cooling time, which was found to be roughly 300 μ s. Given the unfavorable heating geometry, however, it is encouraging that it works as well as it does. With a smaller thermal mass heater located much closer to the tip, greatly improved performance should be possible. Of course, until actually demonstrated, the overall performance and reliability are unknown. Nonetheless, an integrated heater could be a solution which is fairly fast, compact, and capable of being batch-fabricated.

• Future prospects

Sample smoothing and low-mass cantilevers have led to significant improvements in data rate and signal-to-noise ratio. The feasibility of track servoing has been demonstrated, as has wear-free operation over a period of days. The data rate remains an issue, as do contamination, long-term reliability, and implementation. While it is possible to go to even smaller cantilevers for higher data rates, it will probably be desirable at some point to use an array of tips operated in parallel for further improvements. In this case, it will be essential to integrate some of the functionality onto the cantilever. Ideally, one would like to have both a heater and a deflection sensor integrated onto the cantilever. This would certainly simplify the overall implementation, make the device more compact, and eliminate the troublesome alignments of the tip and laser. Possible sensor schemes include capacitance detection [32]

¹ Performed by MCR Associates, Inc., Sunnyvale, CA; private communication.
² Courtesy of Park Scientific Instruments, Inc., Sunnyvale, CA.

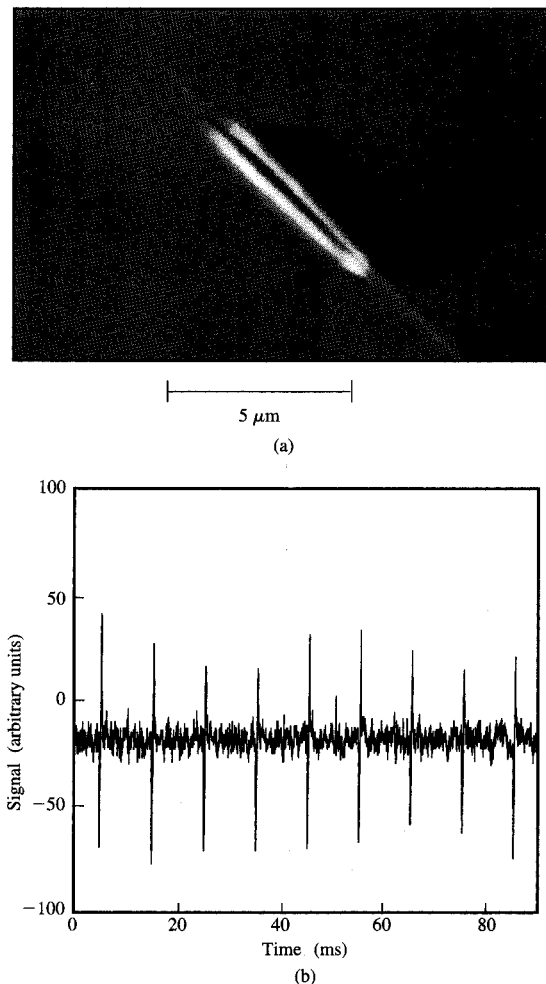


Figure 17

(a) AFM micrograph of mark written with a piezoresistive cantilever. A 0.1-ms, 11-mA current pulse was applied to the conducting cantilever ($R = 2 \text{ k}\Omega$) to heat the lever to enable the tip to write. The large thermal mass and separation between the tip and region of maximum power dissipation give rise to a long cool-down time, resulting in elongated marks. (b) Readback signal from a series of such marks. Standard optical beam detection was used to measure the deflection of the cantilever.

or piezoresistive cantilevers [31]. As mentioned, work is currently underway at our laboratory to implement an integrated heater.

Clearly, a great deal of work must be done to combine these functions and fulfill all of the stringent requirements. At this point, one can say that the technique appears promising, and we have made significant progress, but it is still too early to say whether it can be made practical.

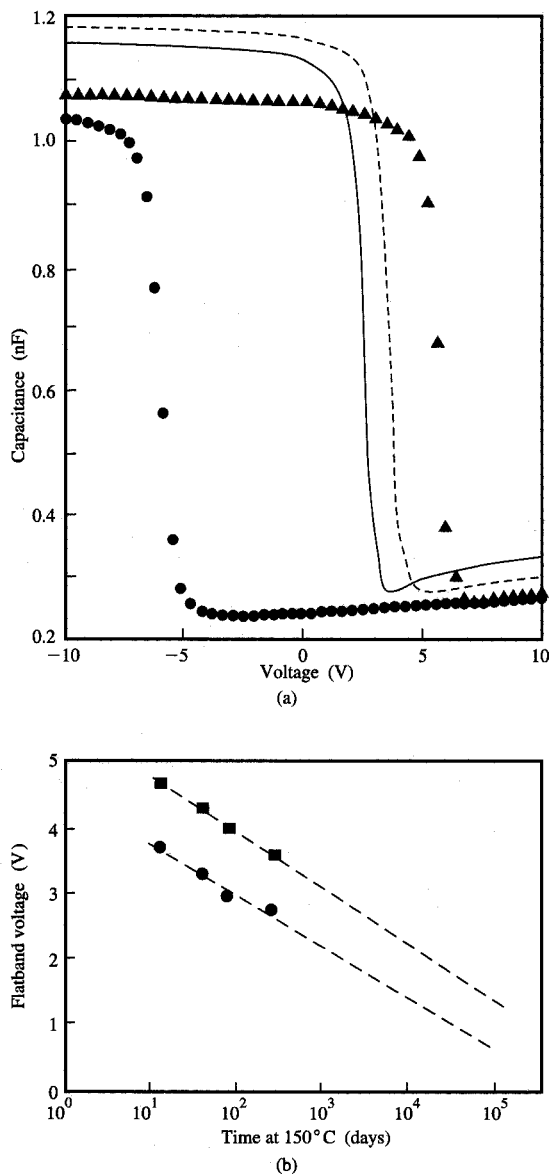


Figure 18

(a) Capacitance-voltage characteristic for a NOS sample with a 1-mm-diameter electrode deposited on the nitride surface. The solid circles show the initial curve before application of a write voltage pulse. After application of a -40-V pulse to the Si, the curve is shifted along the voltage axis, as shown by the solid triangles. The pulse width was $100\text{ }\mu\text{s}$. The curves represented by the dashed and solid lines were taken after writing with the -40-V pulse and aging the sample at 150°C for 13 days and 9 months, respectively. (b) The flatband voltage vs. time at 150°C for two samples with differing oxide thickness and initial charge levels. The solid circles are for the same $39\text{-}\text{\AA}$ -thick oxide sample as (a). The solid squares are for a $66\text{-}\text{\AA}$ -thick oxide sample. The dashed lines are logarithmic fits to the data with slopes of 0.33 V/decade (solid circles) and 0.36 V/decade (solid squares) [37].

Charge storage in NOS structures

Both the AFM thermomechanical writing and the STM gold-writing schemes discussed above use changes in topography (pits or bumps) to record information. Nontopographic alternatives exist, including altering the substrate's electrical or optical properties. These alternatives lend themselves more easily to nonmechanical readback, which may offer a speed advantage. In addition, it is usually not straightforward to undo a topographic change, while it may be feasible to reverse an electrical or optical change. Certainly it would be desirable to have a recording scheme which is erasable, as opposed to the write-once schemes discussed above. One nontopographic method that we have considered is charge storage in nitride-oxide-semiconductor (NOS) structures.

The technique of using a stylus to store localized charge in $\text{Si}_3\text{N}_4\text{-SiO}_2\text{-Si}$, or NOS, was first demonstrated by Iwamura et al. [33] and later refined by Barrett and Quate [34, 35]. By placing a movable metal electrode on the Si_3N_4 surface and applying a voltage pulse between the metal electrode and the Si substrate, charge carriers can be made to tunnel through the oxide and be trapped in the nitride. The charge-trapping capabilities of NOS can be seen by measuring the capacitance-voltage characteristic, as shown in Figure 18(a). As the voltage on the nitride side is swept from negative to positive, the majority carriers in the p-type Si are driven away from the oxide and a depletion region is formed. The voltage at which depletion occurs can be shifted by trapping charge in the nitride, as shown by the difference between the solid circles and triangles in Figure 18(a). This charge was trapped by applying a -40-V pulse to the silicon. Barrett and Quate [34], using a modified atomic force microscope, demonstrated that the stored charge could be localized to 75-nm -sized regions and predicted a lower limit of several tens of nm. Iwamura et al. [33] also investigated the data storage applications of this technology and demonstrated reasonable data rates using a sample mounted on a rotating spindle. We have extended the work of these investigators and examined the issues of projected data lifetime and signal-to-noise ratio [36, 37].

• Stored-charge lifetime

One issue in using NOS for data storage is the long-term stability of the trapped charge. Using an Ir-coated Si tip and a $44\text{-}\text{\AA}$ -oxide NOS structure, we have written a pattern of marks and monitored the pattern as the sample aged at 150°C . The marks were produced using -40-V pulses. As shown in Figure 19, there was no appreciable change in the mark size after nine months at 150°C . The six rows of marks were written with different pulse widths. Starting from the top of the image, the write pulse widths for the rows were $90, 70, 50, 30, 10,$ and $1\text{ }\mu\text{s}$. Note the decreasing mark size for the shorter pulses, with the

smallest marks measuring approximately 150 nm in diameter. This dependence of mark size on pulse duration is a result of the decrease in the field strength with increasing distance from the tip. The shortest -40-V pulses with which we have been able to write were 40 ns in duration, the limit of the high-voltage pulse generator.

The signal detected in these experiments is dC/dV at a fixed bias voltage. The contrast in the image reflects the absence/presence of a nonzero slope in the C - V characteristic at the bias voltage. The bias voltage is typically set at the transition to depletion when no charge is trapped, -6 V for the sample shown in Figure 18(a). This high dC/dV value is depicted by white in the gray-scale image (Figure 19). When charge is stored, the transition will be shifted in voltage. The shift is typically about 12 V, as seen in Figure 18(a). In this case, the slope dC/dV at the bias voltage will be small; this is shown as black in Figure 19.

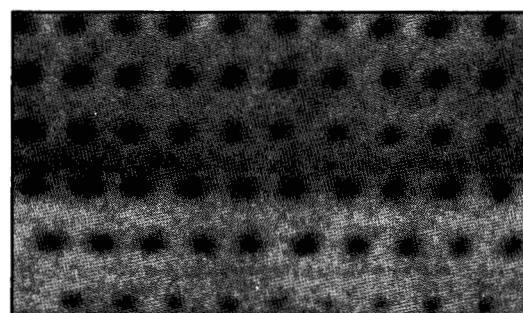
From these data it cannot be determined whether charge is decaying during the aging experiments. It can only be said that there is still some trapped charge, and that no appreciable spreading of the charge is seen. Small changes in the apparent mark size can occur because of differences in the tip size or differences in the signal level and gray-scale settings between the images.

To quantify the rate of decay of the charge, C - V characteristics were also measured as a function of time stored at 150°C. Figure 18(a) shows the C - V curves at various times for the 39-Å-oxide NOS structure with a 1-mm Au electrode deposited on the surface. After nine months at 150°C, a decrease in the trapped charge level of approximately 3 V is seen, a relatively small fraction of the total 12-V shift. Similar charge decay rates were seen with the two thicker oxide samples.

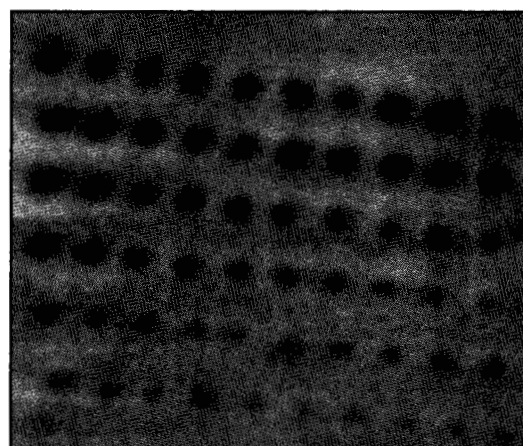
Shown in Figure 18(b) is the flat-band voltage (V_{FB}) as a function of time at 150°C for two samples with differing oxide thickness and initial charge level. The solid circles are for the same 39-Å-thick oxide sample in Figure 18(a). The solid squares are for a 66-Å-thick oxide sample. On the basis of data from the thinner oxide sample, we would expect a stored-charge lifetime in excess of 30 years at 150°C, assuming a reasonable minimum detectable V_{FB} of 1.5 V. In the thicker oxide sample, the stored charge lifetime at 150°C is estimated to be as long as 300 years!

• Carrier-to-noise ratio

By placing the NOS sample on an air-bearing spindle and using an etched tungsten tip, we are able to achieve relatively high data rates and make carrier-to-noise ratio (CNR) measurements. A CNR of nearly 60 dB (in a 3-kHz bandwidth) was achieved on a 500-kHz carrier, written and read at a linear disk velocity of 2.3 m/s [36]. From the linear velocity and the data frequency, we estimate the bit size to be 2 μ m. This is quite large and is a result of the



(a)



(b)

Figure 19

Localized charge regions written with an Ir-coated Si tip in NOS, (a) just after writing, (b) after aging at 150°C for 9 months. The six rows were written with successively shorter -40-V pulses. Starting from the top, the pulse widths were 90, 70, 50, 30, 10, and 1 μ s [37].

tip size. The requirement of electrical contact between the tip and nitride results in a rapid wearing of the tip, as was also seen by Iwamura et al. [33]. This wearing is a significant problem for this technology and will have to be overcome for it to be successful.

This measured CNR is much less than the value we expect. The measured signal is due to the difference in capacitance between regions where no charge is stored and those where charge is stored. On the basis of the measured depletion capacitance for large (1-mm-diameter) areas and the RCA sensor noise level [34, 38], a CNR of 105 dB in a 3-kHz bandwidth should have been achievable, or 45 dB more than the measured CNR [37]. While some of this

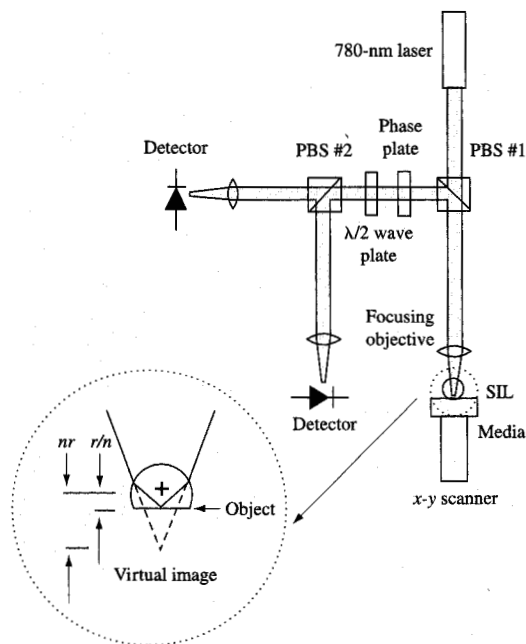


Figure 20

Schematic of the optical storage test setup with the SIL inserted between a focusing objective and the media. PBS is a polarizing beam splitter; r is the radius of curvature of the SIL [47].

difference can be attributed to extraneous noise sources [36], we cannot account for most of this difference as an increase in the noise. This suggests that there is also a smaller signal than expected, i.e., that $\Delta C/A$ for the $2\text{-}\mu\text{m}$ bits is less than that found for the 1-mm areas and less than that calculated from theory.

We have found that this discrepancy in measured CNR, along with an unexpected dependence of the capacitance on the measurement frequency and electrode size, the measured dC/dV curves of Barrett and Quate, and the horizontal shift in our C - V curves of -5 to -7 V, can all be explained by the presence of a uniform layer of trapped positive charge in the oxide film [37]. This layer of trapped charge is unrelated to the discrete areas of trapped charge which are used for our data storage. Since our experiments were done using p-type Si, the uniform positive charge layer biases the entire wafer so as to create an inversion layer at the surface. This inversion layer provides a source of minority carriers and a conduction path for them to travel across the silicon surface, and thus reduces the change in capacitance measured when the Si is biased to depletion. By eliminating this layer of positive charge, or by using n-type Si, it should be possible to eliminate this

effect and to increase the measured signal. Preliminary experiments on n-type Si show this to be the case.³

The future of NOS for high-density data storage is still uncertain. It has been demonstrated that bits as small as 75 nm can be made, although the predicted CNR of such bits, using the RCA sensor, would only be of the order of 38 dB in a 30-kHz bandwidth. Using n-type Si, it is possible that such a CNR could be achieved, although that has not been proven, and media noise at such data densities has yet to be measured. The biggest challenge, however, is probably tip wear, because the sharp, conducting tip must be kept in good contact with the media. Because silicon nitride is very hard, the tip tends to wear faster than the disk. The challenge is to design the tip to wear in a manner so that probe and bit sizes remain constant.

Near-field optical recording with a solid immersion lens

Another approach to nontopographic recording is to use an optical, rather than an electrical, probe. The way to circumvent the diffraction limit and achieve high density is to use a near-field scanning optical microscope (NSOM) [29, 39–41]. Betzig and coworkers used an NSOM based on a tapered optical fiber to demonstrate writing in a magneto-optic medium and claimed densities as high as 45 Gb/in.^2 [42]. Their technique currently has the drawbacks that it operates in transmission mode, which makes it bulkier and more complicated than reflection mode, and it has very low optical efficiency.

Kino and co-workers [43, 44] have demonstrated a second approach to decreasing the optical spot size. Their approach is to increase the numerical aperture (N.A.) above the theoretical upper limit of 1 which is obtainable in air. This can be accomplished by using an immersion lens made from a high-index-of-refraction material, the solid immersion lens (SIL). The lens is formed by placing a truncated sphere between a focusing objective and the sample of interest. The wavelength inside the sphere is reduced by the high index of the glass, leading to a reduction in the diffraction-limited spot size. In addition, if the incident rays are refracted at the sphere's surface, θ_{max} can be increased, leading to a further reduction in spot size. The large-angle rays, however, may be at angles greater than the critical angle for exiting the high-index glass at the base of the SIL. Thus, both the reduced wavelength and large-angle rays exist only within the high-index SIL and must be coupled via their evanescent field to the sample located at the base of the SIL. By placing the substrate within the evanescent decay length, the small spot can be transmitted across the air gap. This is, therefore, a form of near-field optics.

³ B. D. Terris and R. C. Barrett, IBM Almaden Research Center, private communication, 1995.

There are two special truncated spheres for which converging incident rays are focused without aberration at the base of the partial sphere. One truncated sphere for which such aberration-free focusing occurs is the hemisphere. In this case the rays enter normal to the sphere surface and are therefore not refracted. The wavelength, however, is reduced by a factor of n , and the effective N.A. is increased by n . Mansfield and Kino [43] demonstrated imaging with such a hemispherical glass lens and resolved a 200-nm period in a photoresist grating using 436-nm light.

In our work we used the second type of partial sphere for which aberration-free focusing occurs [45, 46]. This occurs when the sphere is truncated to a thickness $r(1 + 1/n)$, where r is the radius of the sphere, as shown in the lower left corner of **Figure 20**. A focused spot is obtained at the base of the SIL when the incident rays are converging toward a point located a distance nr below the center of the sphere. Because the rays are refracted at the spherical surface, $\theta_{\max}$ is increased. This increase in $\theta_{\max}$, together with the reduction in λ , results in an increase in N.A. by a factor of n^2 as compared to the incident N.A. The maximum possible N.A., however, cannot exceed n . This is identical in principle to an oil immersion lens. By using a focusing objective with N.A. = 0.55, typical of those used in optical recording, and a solid immersion lens made of $n = 1.9$ index glass, an effective N.A. of 1.9 can be obtained. The focused spot is thereby reduced by 3.5 $\times$ and the potential bit areal density increased by an order of magnitude.

• Static testing results

We have recently used the apparatus shown in **Figure 20** to demonstrate optical recording in magneto-optic (MO) media [47]. Standard optical recording optics, including differential detection of the Kerr rotation, were used except for the addition of a SIL placed between the focusing objective and the MO media. The focusing objective was a N.A. = 0.6 microscope objective, and the laser was a 780-nm diode laser. Piezoelectric scanners were used to scan the MO sample. Two types of SILs were tested. A 2-mm-diameter SIL was made by lapping and polishing a $n = 1.83$ fiber-optic coupling sphere. This SIL was placed on the sample and scanned with it. The second SIL consisted of a 300- μm -diameter, $n = 1.9$, sphere which was glued into a thin, rectangular holder. The holder/sphere assembly was then lapped and polished until the exposed face of the sphere was the desired diameter. This holder was suspended so that it was just touching the surface and could be tilted so that the lens bottom was parallel to the surface. This allowed for moving the media with respect to the optics, as required for data storage applications.

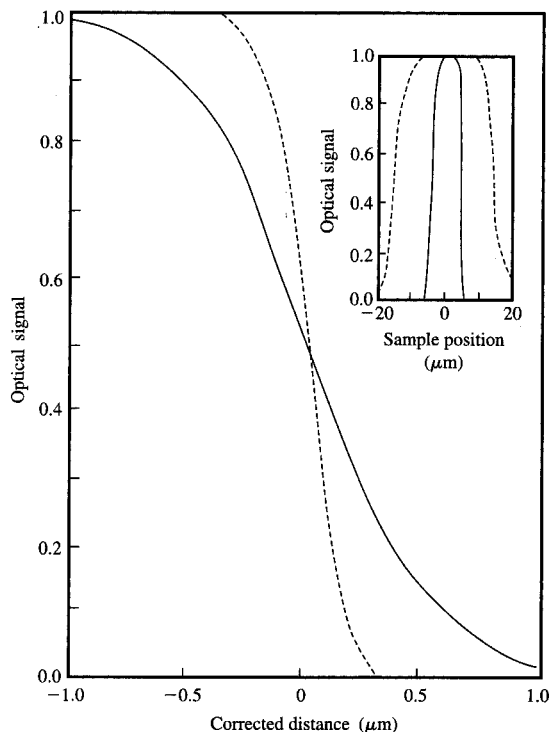


Figure 21

The line scan across one edge of a metal line using a 0.6-N.A. objective with (dashed) and without (solid) the $n = 1.83$ SIL. The x scale has been corrected for the SIL magnification. The inset shows the line scan across a 10- μm -wide line using the objective alone (solid line) and with a $n = 1.83$ SIL placed on top of the sample (dashed line). This demonstrates the SIL magnification of a factor of n^2 [47].

The SIL performance was first characterized by imaging an optical microscope calibration slide. The area imaged consisted of alternating opaque and transparent lines. The sample and 2-mm SIL were mechanically scanned together with respect to the focusing objective using piezoelectric scanners. Scans across a full line and smaller scans across just one edge of a line, with and without the SIL in place, are shown in **Figure 21**. Note that when the SIL is used, the imaged distance is n^2 times smaller than the physically scanned distance. Thus, comparing the line width measured with and without the SIL, we find that the scanners moved 3.3 times farther when using the SIL in order to image the same 10- μm line. This is in good agreement with the value of 3.35 calculated from the index of the glass of 1.83.

Using the $n = 1.83$ SIL, a focused spot size of 317 nm was achieved, as determined from the scans of the single

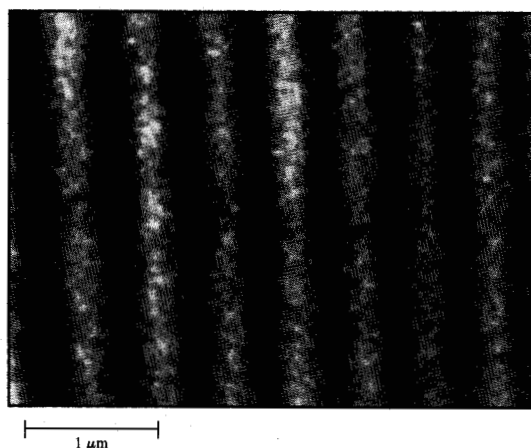


Figure 22

A 500-nm-period grating as imaged with the $n = 1.83$ SIL [47].

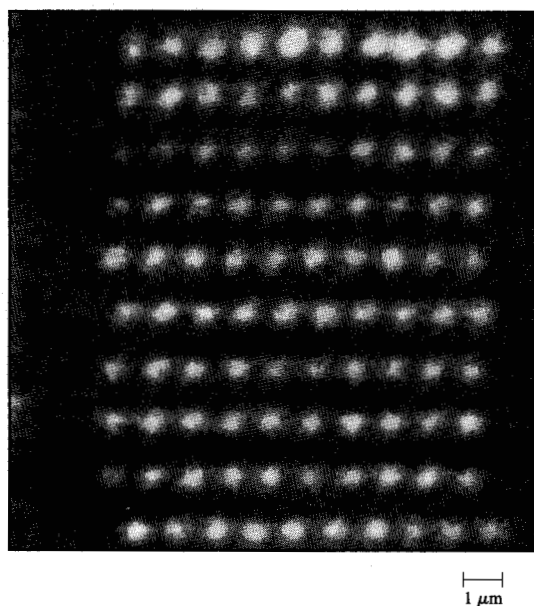


Figure 23

Domains in a magneto-optic sample written and read through a $n = 1.9$ SIL. The domains were written by pulsing the laser with the sample in an applied magnetic field. The signal is due to the Kerr rotation of the linearly polarized light by the magnetic sample. The smallest domains are ~ 350 nm in diameter. The domains are reversible by writing in the presence of the reversed magnetic field [47].

edges shown in Figure 21. Assuming a Gaussian beam profile, the half-width is given by

$$\sigma = \frac{1}{\sqrt{2\pi}} \frac{I_0}{dI/dx},$$

where I_0 is the full height of the transition and dI/dx is the slope at the midpoint of the transition. The full-width-half-maximum spot size is then given by $d = 2\sqrt{2 \ln 2} \sigma$. From Figure 21, we estimate that for the microscope objective alone $d = 990$ nm. Following the scalar diffraction analysis of Jipson and Williams [48] and using the measured ratio of the laser beam $1/e^2$ width to the microscope lens aperture of 0.49, we expect a spot size of $d = 0.76(\lambda/\text{N.A.}) = 1020$ nm. This is in good agreement with the measured value. The addition of the SIL will reduce the spot size by n^2 . One factor of n arises from the wavelength reduction and one factor from the increase in θ_{max} inside the lens due to refraction at the lens surface. By geometric construction it can be shown that $\sin \theta$, measured in air without the SIL, is increased to $n \sin \theta$ with the addition of the SIL. We therefore expect $d = 0.76(\lambda/n^2 \text{N.A.}) = 304$ nm, in good agreement with the measured value from Figure 21 of 317 nm. Note that under the same conditions a hemispherical SIL would have given a spot size of 557 nm, considerably larger than the 317 nm achieved with the SIL. As one further test of the SIL, we have imaged lithography features of varying periodicities. Using 780-nm light, we were able to resolve a 500-nm-period grating, as shown in Figure 22.

Figure 23 shows MO domains written and read back through the SIL using the 300- μm lens. The MO sample consisted of $\text{Si}_3\text{N}_4/\text{TbFe}/\text{Si}_3\text{N}_4$ sputter deposited onto a glass substrate. The Si_3N_4 is 80 nm thick and serves as an antireflection coating. The SIL was placed directly on top of the upper Si_3N_4 so that it would be as close as possible to the MO layer. The domains, or bits, were written by pulsing the laser while scanning the sample and applying a dc magnetic field. Prior to writing the bits, the sample was uniformly magnetized in the opposite direction. The bits were erasable by writing in the presence of the opposite-polarity magnetic field. Since the sample is scanned using piezoelectric translators, the scan rate is very slow, and care must be taken not to erase while reading. This was accomplished by pulsing the read beam at a low ($\sim 5\%$) duty cycle and using lock-in detection. The smallest domains which we have written and read using the SIL have a full-width-half-maximum width of approximately 350 nm.

• Flying solid immersion lens

We have also flown a SIL over a spinning MO disk to demonstrate high data rate [49]. Except for the addition of the SIL placed between the focusing objective and the MO media, we have used the same type of optical recording

setup shown in Figure 20. An 830-nm laser diode was used, along with a N.A. = 0.5 objective lens. The medium was $\text{Si}_3\text{N}_4/\text{TbFeCo}/\text{Si}_3\text{N}_4$ sputtered onto a glass disk. The laser light was incident from the media side of the disk. The SIL was an $n = 1.5$, 1-mm-diameter glass sphere which was polished and mounted onto a slider. The slider had an air bearing patterned onto it in the same manner as used in magnetic recording, allowing it to fly. **Figure 24** shows the read signal, with and without the SIL, from a track written without the SIL. The signal with the SIL clearly shows a more square shape and steeper transitions at the edges of the bit, indicative of the smaller spot size. A least-squares fit to the transition width indicates a spot size of roughly $0.6 \mu\text{m}$.

It is clear from Figure 24 that data rates of at least several MHz can easily be achieved with this technique. Although the spot size was still fairly large in this preliminary flying demonstration, that can be explained by a number of factors, including the SIL height above the MO disk of more than $1 \mu\text{m}$. With lower heights to get the full near-field effect, and higher-index spheres, it should be possible to achieve results comparable to those from the static tester. Even those results are not yet at the limit of the SIL. By using a low-flying SIL with shorter-wavelength light (blue), it should be possible to achieve 125-nm spots with very little loss of signal. Such a spot size would correspond to storage densities approaching 40 Gb/in.^2 , which compares favorably to those reached with AFM and NSOM techniques. Because the SIL works in reflection mode and has very high optical efficiency, it might be more easily implemented into a data storage technology than other near-field optical techniques. By combining the SIL with current optical data storage technology, it may be possible to dramatically increase the areal density while retaining erasability and high data rate.

Conclusions

The techniques we have discussed cover a wide range of densities and data rates. It is instructive to make a plot of these parameters and see where the various techniques lie in this phase space. Such a plot is shown in **Figure 25**. At the top of the plot is the herding of individual atoms with the STM. It has the highest density by far, but the lowest data rate. At the other extreme corner we have placed the real-world technology of magnetic recording. It has much lower density, but achieves data rates approaching 100 MHz. The thermomechanical recording as well as the SIL and NOS recording lie in between on both axes. We have also shown other demonstrations on this plot, such as the AT&T NSOM work [42], and STM [10] and AFM [18] phase change recording. Inevitably one finds that as one moves to the right on this plot, by developing techniques with higher data rate, one tends to give up some data density. Nonetheless, we believe that the AFM

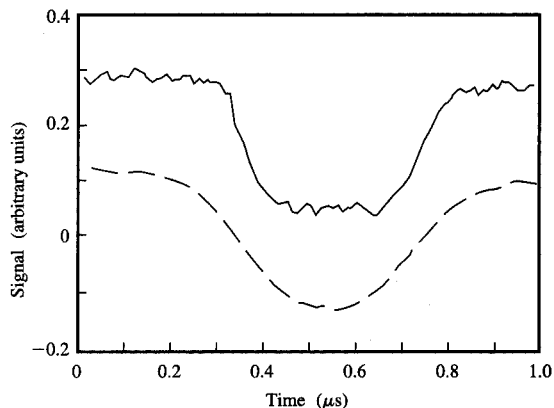


Figure 24

Waveform of a bit written without the SIL and read without (dashed) and with (solid) the SIL [49].

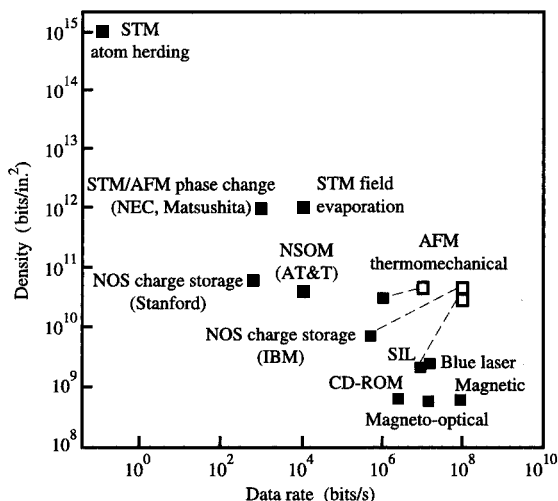


Figure 25

Plot of areal density vs. data rate for various tip-based demonstrations and conventional storage devices. In general, the highest-density demonstrations involve the lowest data rates. The AFM and SIL approaches offer potentially interesting combinations of density and data rate, particularly if the projected improvements (dashed lines) can be achieved.

thermomechanical recording still has room for improvement in data rate. The NOS recording also has room for improvement in density if the wear issue can be

resolved. The SIL work has just begun, and we see no reason that one cannot achieve data rates comparable to those in conventional optical recording. With higher index lenses and blue lasers, the SIL density should also be able to reach much higher values. Some reasonable extrapolations for these techniques are also indicated on the plot. If they become reality, these techniques should occupy extremely attractive regions of density/data rate phase space.

In conclusion, proximal probe techniques offer the potential for extremely compact, high-density data storage devices. In some sense, these are all forms of near-field recording, as is conventional magnetic recording. In any such technique, the density is ultimately limited by the size of the probe and the media-probe spacing. Tip-based techniques have the advantage that one can readily fabricate probes with dimensions of the order of 50 nm or less, a feat difficult to match with magnetic heads. While there are myriad technical obstacles to overcome in these approaches, the promise of tiny, inexpensive drives containing gigabytes of data is a powerful motivation for pursuing this course.

Acknowledgments

The authors gratefully acknowledge H. Birk for help in building the fast STM, Ragnar Erlandsson for performing AFM wear studies, Park Scientific Instruments for supplying the piezoresistive cantilevers, P. Wimmer for assistance in assembling the test stands, M. Melas for the digital channel used in analyzing the AFM waveforms, G. Kino for initiating the work in the solid immersion lens, and B. McChesney for providing the magneto-optic media.

References and notes

- C. F. Quate, U.S. Patent 4,575,822, 1986.
- M. Ringger, H. R. Hidber, R. Schlogl, P. Oelhafen, and H. J. Guntherodt, *Appl. Phys. Lett.* **46**, 832 (1985).
- D. W. Abraham, H. J. Mamin, E. Ganz, and J. Clarke, *IBM J. Res. Develop.* **30**, 492 (1986).
- M. A. McCord and R. F. W. Pease, *J. Vac. Sci. Technol. B* **4**, 86 (1986).
- J. S. Foster, J. E. Frommer, and P. C. Arnett, *Nature* **331**, 324 (1988).
- T. R. Albrecht, M. M. Dovek, M. D. Kirk, C. A. Lang, C. F. Quate, and D. P. E. Smith, *Appl. Phys. Lett.* **55**, 1727 (1989).
- J. Schneir, R. Sonnenfeld, O. Marti, P. K. Hansma, J. E. Demuth, and R. J. Hamers, *J. Appl. Phys.* **63**, 717 (1988).
- H. J. Mamin, P. H. Guethner, and D. Rugar, *Phys. Rev. Lett.* **65**, 2418 (1990).
- Y. Utsugi, *Nature* **347**, 747 (1990).
- A. Sato and Y. Tsukamoto, *Nature* **363**, 431 (1993).
- R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, *Nature* **325**, 419 (1987).
- D. M. Eigler and E. K. Schweizer, *Nature* **344**, 524 (1990).
- See, for example, review articles by U. Staufer in *Scanning Tunneling Microscopy II*, R. Weisendanger and H.-J. Güntherodt, Eds., Springer, Berlin, p. 273, and C. F. Quate in *Highlights in Condensed Matter Physics and Future Prospects*, L. Esaki, Ed., Plenum, New York, p. 573.
- T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, and U. D. Schwarz, *Ultramicrosc.* **42**, 1446 (1992).
- H. J. Mamin and D. Rugar, *Appl. Phys. Lett.* **61**, 1003 (1992).
- R. Barrett and C. F. Quate, *J. Appl. Phys.* **70**, 2725 (1991).
- S. Hosaka, H. Koyanagi, and A. Kikukawa, *Jpn. J. Appl. Phys.* **32**, L464 (1993).
- H. Kado and T. Tohda, *Appl. Phys. Lett.* **66**, 2961 (1995).
- H. J. Mamin, S. Chiang, H. Birk, P. H. Guethner, and D. Rugar, *J. Vac. Sci. Technol. B* **9**, 1398 (1991).
- J. I. Pascual, J. Mendez, J. Gomez-Herrero, A. M. Baro, and N. Garcia, *Phys. Rev. Lett.* **71**, 1852 (1993).
- H. J. Mamin and D. Rugar, *Phys. Rev. Lett.* **72**, 1128 (1994).
- H. J. Mamin, H. Birk, P. Wimmer, and D. Rugar, *J. Appl. Phys.* **75**, 161 (1994).
- T. R. Albrecht, S. Akamine, T. E. Carver, and C. F. Quate, *J. Vac. Sci. Technol. A* **8**, 3386 (1990).
- H. J. Mamin, L. S. Fan, S. Hoen, and D. Rugar, *Sensors & Actuators* **48**, 215 (1995).
- Y. Martin, C. C. Williams, and H. K. Wickramasinghe, *J. Appl. Phys.* **61**, 4723 (1987).
- G. M. McClelland, R. Erlandsson, and S. Chiang, in *Review of Progress in Quantitative Nondestructive Evaluation*, D. O. Thompson and D. E. Chimenti, Eds., Plenum, New York, 1987, p. 307.
- C. D. Mee and E. D. Daniel, *Magnetic Recording* McGraw-Hill Book Co., Inc., New York, 1987.
- S. Hoen, H. J. Mamin, and D. Rugar, *Appl. Phys. Lett.* **64**, 267 (1994), and references therein.
- E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, *Science* **251**, 1468 (1991).
- D. Rugar, H. J. Mamin, and P. Guethner, *Appl. Phys. Lett.* **55**, 2588 (1989).
- M. Tortorone, R. C. Barrett, and C. F. Quate, *Appl. Phys. Lett.* **62**, 834 (1993).
- J. Brugger, N. Blanc, P. Renaud, and N. F. de Rooij, *Sensors & Actuators A* **43**, 339 (1994).
- S. Iwamura, Y. Nishida, and K. Hashimoto, *IEEE Trans. Electron Devices* **ED-28**, 854 (1981).
- R. C. Barrett and C. F. Quate, *J. Appl. Phys.* **70**, 2725 (1991).
- R. C. Barrett and C. F. Quate, *Ultramicrosc.* **42-44**, Pt. A, 262 (1992).
- B. D. Terris, R. C. Barrett, and H. J. Mamin, *Proc. SPIE* **1855**, 195 (1993).
- B. D. Terris and R. C. Barrett, *IEEE Trans. Electron Devices* **42**, 944 (1995).
- R. C. Palmer, E. J. Denlinger, and H. Kawamoto, *RCA Rev.* **43**, 194 (1982).
- D. W. Pohl, W. Denk, and M. Lanz, *Appl. Phys. Lett.* **44**, 651 (1984).
- A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, *Ultramicrosc.* **13**, 227 (1984).
- E. Betzig, M. Isaacson, and A. Lewis, *Appl. Phys. Lett.* **51**, 2088 (1987).
- E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, P. L. Finn, M. H. Kryder, and C. H. Chang, *Appl. Phys. Lett.* **61**, 142 (1992).
- S. M. Mansfield and G. S. Kino, *Appl. Phys. Lett.* **57**, 2615 (1990).
- S. M. Mansfield, W. R. Studenman, G. S. Kino, and K. Osatoo, *Opt. Lett.* **18**, 305 (1993).
- M. Born and E. Wolf, *Principles of Optics*, Pergamon Press, Oxford, England, 1980, p. 253.
- M. P. Davidson, *Proc. SPIE* **1926**, 84 (1993).
- B. D. Terris, H. J. Mamin, and D. Rugar, *Appl. Phys. Lett.* **65**, 388 (1994).

48. V. B. Jipson and C. C. Williams, *Appl. Opt.* **22**, 2202 (1983).
49. H. J. Mamin, B. D. Terris, and D. Rugar, *Proceedings of the Magneto-Optical Recording International Symposium '94*, Magnetics Society of Japan, Vol. 19, Suppl. S1, p. 409 (1995).

Received November 22, 1994; accepted for publication July 28, 1995

H. Jonathon Mamin *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (mamin@almaden.ibm.com).* Dr. Mamin received his B.S. degree in physics from Stanford University in 1978, and a Ph.D. degree in physics from the University of California at Berkeley in 1984. After receiving his doctorate, he remained at Berkeley as a postdoctoral fellow in the then-emerging field of scanning tunneling microscopy. In 1987 Dr. Mamin joined IBM as a research staff member at the Almaden Research Center in the Storage Systems and Technology Department. His research interests have included magnetic force microscopy, STM surface modification, and high-density data storage. Dr. Mamin is a member of the American Physical Society.

Bruce D. Terris *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (TERRIS at ALMADEN, terris@almaden.ibm.com).* Dr. Terris received the B.S. degree in applied physics from Columbia University in 1979 and the M.S. (1981) and Ph.D. (1983) degrees in physics from the University of Illinois at Urbana-Champaign. After receiving his doctorate, he was a postdoctoral fellow at the Argonne National Laboratory. In 1985, Dr. Terris joined IBM as a research staff member at the Almaden Research Center. His current research interests include scanning probe and near-field optical microscopies and their application to high-density data storage. Dr. Terris is a member of the American Physical Society.

Long-Sheng Fan *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (LSFAN at ALMVMA).* Dr. Fan is a research staff member in the area of Interface Tribology and Mechanics at the IBM Almaden Research Center. He did his thesis at UC Berkeley on material and technology issues of surface-micromachined, *in-situ*-assembled micromechanisms. In 1989 he joined the IBM Research Division at Almaden, doing micromechanics research for information storage. Dr. Fan initiated the building of a Si CVD laboratory and set up a micromechanics characterization laboratory at the Research Center. He developed an integrated multilayer high-aspect-ratio actuator technology, and fabricated AFM tips that can read more than a Mb/s. He is currently working on MEMS technologies for high-density recording. Dr. Fan is on the editorial board of the international journal *Sensors and Actuators* as a section editor; he is a member of the IEEE and the American Physical Society.

Storrs Hoen *Solid State Technologies Laboratory, Hewlett-Packard Laboratories, 3500 Deer Creek Road, Palo Alto, California 94304 (hoen@saiph.hpl.hp.com).* Dr. Hoen is a research intern in the Solid State Technologies Laboratory of the Hewlett-Packard Laboratories. He received his B.E. in electrical engineering, materials science, physics, and mathematics from Vanderbilt University in 1984, and an M.A. in physics and philosophy from Oxford University in 1986. Dr. Hoen's doctoral research at the University of California at Berkeley was completed in 1992 and focused on the elastic properties of high-temperature superconductors. From 1992 until 1994, he was a member of technical staff at the Almaden Research Center, working on exploratory data storage. Dr. Hoen is a member of the American Physical Society.

Robert C. Barrett *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (BARRETT at ALMADEN, barrett@almaden.ibm.com).* Dr. Barrett is a research staff member in the User Ergonomics Research Department in the Computer Science function at the IBM Almaden Research Center. He received B.S. degrees in physics and electrical engineering and an M.S. degree in physics from Washington University (St. Louis) in 1987, and M.S. and Ph.D. degrees in applied physics from Stanford University in 1989 and 1991, respectively. He subsequently joined IBM at the Almaden Research Center, where he has worked on magnetic storage, scanned probe storage, computer pointing devices, and information retrieval technologies. Dr. Barrett is a member of the Association for Computing Machinery.

Daniel Rugar *IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120 (RUGAR at ALMADEN, rugar@almaden.ibm.com).* Dr. Rugar is manager of the Exploratory Storage Studies group at the IBM Almaden Research Center. He received a B.A. degree in physics from Pomona College in 1975, and a Ph.D. degree in applied physics from Stanford University in 1982. In 1984 he joined the IBM Research Division, where he has worked on magneto-optical data storage, magnetic resonance force microscopy, and the application of scanned probe techniques for high-density data storage. Dr. Rugar is a member of the American Physical Society and the Institute of Electrical and Electronics Engineers.