Review article

Nan Zhou, Xianfan Xu*, Aaron T. Hammack, Barry C. Stipe, Kaizhong Gao, Werner Scholz and Edward C. Gage

Plasmonic near-field transducer for heat-assisted magnetic recording

Abstract: Plasmonic devices, made of apertures or antennas, have played significant roles in advancing the fields of optics and opto-electronics by offering subwavelength manipulation of light in the visible and near infrared frequencies. The development of heat-assisted magnetic recording (HAMR) opens up a new application of plasmonic nanostructures, where they act as near field transducers (NFTs) to locally and temporally heat a sub-diffraction-limited region in the recording medium above its Curie temperature to reduce the magnetic coercivity. This allows use of very small grain volume in the medium while still maintaining data thermal stability, and increasing storage density in the next generation hard disk drives (HDDS). In this paper, we review different plasmonic NFT designs that are promising to be applied in HAMR. We focus on the mechanisms contributing to the coupling and confinement of optical energy. We also illustrate the self-heating issue in NFT materials associated with the generation of a confined optical spot, which could result in degradation of performance and failure of components. The possibility of using alternative plasmonic materials will be discussed.

Keywords: field confinement; field enhancement; heat-assisted magnetic recording; low loss materials; near field transducer; optical antenna; plasmonics.

DOI 10.1515/nanoph-2014-0001
Received January 20, 2014; accepted May 8, 2014

*Corresponding author: Xianfan Xu, School of Mechanical Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47906, USA, e-mail: xxu@purdue.edu
Nan Zhou: School of Mechanical Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47906, USA
Aaron T. Hammack and Barry C. Stipe: HGST, San Jose Research Center, 3403 Yerba Buena Road, San Jose, CA 95135, USA
Kaizhong Gao and Werner Scholz: Seagate Technology, 7801 Computer Ave. S., Bloomington, MN 55435, USA
Edward C. Gage: Seagate Technology, 1280 Disc Drive, Shakopee, MN 55379, USA

Edited by Vlad Shalaev

© 2014 Science Wise Publishing & De Gruyter

1 Introduction

As projected by the International Data Corporation (IDC), the worldwide data storage need will continue to grow more than 40% annually. This demand is met mostly by an increase in the areal density, often expressed in bits per square inch [1]. The hard disk drive (HDD) industry is the primary industry to satisfy this demand. During the last decades, by replacing longitudinal magnetic recording (LMR), perpendicular magnetic recording (PMR) was able to continue increasing areal density, where the magnetic elements are aligned perpendicular to the disk surface [2]. However, there is a limitation with the PMR technology. When the bits are more closely packed, the grain volume \( V \) in the recording medium must shrink to maintain the signal to noise ratio (SNR) under the presence of thermal fluctuation \( k_B T \), where \( k_B \) is the Boltzmann constant. As this occurs, the ability to store information degrades, which is known as the superparamagnetic limit [3] that must satisfy:

\[
\frac{K V}{k_B T} \geq 70, \tag{1}
\]

where \( K \) is the uniaxial anisotropy energy density. Increasing the anisotropy in the media regains the thermal stability, but at the cost of increased media coercivity. Due to limitations on magnetic write fields that can be produced using current writer materials and designs in HDDs, this leads to the inability to record information using conventional PMR. To keep increasing the areal density, new physics and technologies are needed.

Heat assisted magnetic recording (HAMR) is one of the new technologies for advancing disk drive areal density beyond 1 Tb/in\(^2\) [4, 5], which is the estimated limit of PMR. It removes the switching limitation by applying local heating to the recording media to lower its coercivity. This allows the use of very high anisotropy materials such as FePt to maintain data thermal stability and the ability to record information. In fact, significant progress in HAMR
has already been made, and each company in the HDD industry has plans to introduce HAMR technology into product within the next few years. In March 2012, Seagate Technology (Bloomington, MN, USA) has demonstrated a HAMR areal density of 1 Tb/in² [6], with a linear bit density of around 2 million bpi (bits per inch) [7]. This is about 55% higher than today’s 620 Gb/in² in 3.5-inch hard drives. Extension to beyond 1 Tb/in² can be achieved by increasing the magnetic anisotropy and reducing the grain size [8]. More recently, Seagate Technology introduced a prototype of a fully integrated and functioning HAMR drive [9, 10] and proposed that the next generation HAMR technology will be incorporated into 2.5-inch enterprise HDDs. Western Digital (Irvine, CA, USA) also demonstrated its HAMR technology at the 2013 China International Forum on Advanced Materials and Commercialization, where a PC powered by a 2.5-inch HAMR hard drive was presented [11].

Figure 1 shows the industry projected areal density growth and the timeline when the HAMR technology is to be introduced into production.

The key component in HAMR is a near field transducer (NFT) for applying heat through the use of a laser on the medium. Details of the light delivery system to the NFT vary depending on the specific implementation [12–14]. A general schematic of the HAMR head is given in Figure 2A and the common elements of a light delivery system are shown in Figure 2B. The laser diode, with the wavelength near 800 nm, is coupled to a waveguide using a mode coupler. The light then propagates down the waveguide where it couples into the NFT which then radiates into the recording medium where it is converted into thermal energy. The waveguide could have a parabolic shape, the planar solid immersion mirror (PSIM), with a dual offset grating to focus the waveguide mode [12]. A thin film dielectric waveguide with a high refractive index core [13], or a metallic surface plasmon (SP) waveguide [14] are another two waveguide designs. The optical spot from NFT has to be localized to a very small dimension to achieve areal densities in the 1-5 Tb/in² range. For 1 Tb/in² at a bit aspect ratio of 4.0, the track density is 500 ktpi (kilo tracks per inch) or a track pitch of 51 nm which is about 16 times smaller than the wavelength of a diode laser (~800 nm). The NFT design is largely based on excitation of surface plasmons of a nanostructure, which re-radiate and produce a sub-diffraction-limited light spot. The much enhanced field produced by the plasmonic nanostructure is only confined in the near field and has a large divergence. However, this is not a concern in HDD, as the distance between the head and the medium is only a few nanometers during operation.

In this review, we focus on plasmonic NFT designs that can produce sub-diffraction-limited optical spots to increase areal density. In Section 2, the mechanisms that contribute to the energy confinement and enhancement of coupling efficiency in NFT will be discussed. A number of designs will be described and compared using a proper figure of merit (FOM). Associated with the localization of the optical spot dimensions is self-heating in NFT materials that will degrade the NFT performance. The influence of optical properties on the performance of NFT, i.e., power delivery vs. self-heating will be discussed in Section 3.

### 2 Designs of plasmonic NFT

Various designs of NFT have been proposed to localize light onto the recording medium [5, 15]. The fundamental capability of an NFT is to break the diffraction limit by concentrating the optical energy into a spot much smaller than the incident laser wavelength. Accompanying with the localization, there is also a requirement of large field enhancement within the optical spot. Apart from producing a cross-track full-width at half-maximum (FWHM) spot of <50 nm as required by the areal density, the NFT must simultaneously deliver enough power to the recording medium with as small as possible incident laser power to reduce the possible self-heating of the NFT. As such, the efficiency of the NFT is a key figure of merit in determining the quality for a given transducer design.

The NFT usually takes the form of an antenna including aperture type antenna, and many designs are based on localized surface plasmon (LSP) resonance. Different from propagating SPs, LSPs are oscillations of surface charges.
bounded to a finite structure such as a metallic (nano) particle or a dielectric particle surrounded by metal. Some simple shapes include nanoscale spheres, disks, holes, or rectangular apertures. At the wavelength and polarization for plasmonic resonance, the incident power is coupled to the structure to the maximum extent and produces a field enhanced spot comparable to the structure dimension, which will couple energy to the recording medium on the same spatial scale. The simple circular and rectangular shape apertures have the limitation that in order to obtain a small spot the dimension must reduce, resulting in transmission or coupling efficiency too low to be useful for HAMR. For example, based on Bethe’s theory [16], the transmission through a 50 nm diameter circular aperture is <0.04% at 800 nm. Numerical simulation results shown in Figures 3A and B demonstrate that energy cannot penetrate the small aperture and the transmitted spot size is larger than the aperture. It is known that the problem of low transmission of a single aperture can be resolved by using an array of holes [19, 21, 22] or adding grooves

Figure 2  (A) Schematic of a HAMR head. (B) Common blocks of the light delivery system in a HAMR head.

Figure 3  (A) Cross-sectional intensity and (B) transmitted electric field intensity |E|^2 (5 nm from the aperture exit plane, in air) distributions for a 50 nm diameter circular aperture in a 50 nm gold film (optical properties are taken from [17]) on glass. A y-polarized plane wave at 800 nm illuminates the aperture from the glass side. Simulations are performed using a frequency-domain finite-element method (FEM) solver [18]. (C) A hole array (from [19]) and (D) groove structure (from [20]) used to achieve transmission enhancement.
around the aperture [20, 23, 24], as shown in Figures 3C and D respectively. This phenomenon of extraordinary optical transmission (EOT) essentially results from a combination of propagating SP, grating effect, and scattering evanescent fields [19], and has received significant attention. However, the overall size of array or grating structures is of the order of several wavelengths [25, 26], which makes EOT designs questionable to be integrated into a recording head. In order to improve the transmission or power output of a single nanostructure, many variants of simple circular and rectangular structures have been investigated, where sharp, nanoscale tips, pins, and notches are intentionally used to take advantage of the lightning rod effect. Charges accumulate at the sharpest areas of the object to produce strongest electric field. Different from LSP, the lightning rod effect is a non-resonant phenomenon, also called “non-resonant amplification” to better describe the essence of the process [27]. On the other hand, the lightning rod effect is readily combined with LSP resonance to further increase the field enhancement and confinement. The dimensions of both the resonator and the sharp feature need to be optimized to achieve the best energy coupling efficiency. Common designs include triangle antenna [28, 29] and triangle aperture [30, 31], C aperture [32–36], bowtie antenna [37–40] and bowtie aperture [15, 23, 41–44].

The bowtie and C apertures are good examples for utilizing the combined effect of resonance and non-resonant amplification by nanostructures. The low transmission of a regularly shaped aperture can be understood as a result of the cutoff of propagating waveguide modes. For example for a cylindrical waveguide, the cutoff occurs for a diameter <0.55 λ [23]. From a waveguide point of view, an efficient approach to enhance the transmission is to increase the cutoff wavelength [32] so a propagating mode can be supported. One type of aperture that can be explored is the ridge aperture, which adopts the concept of a ridge waveguide in microwave engineering that has been widely used to increase the bandwidth [45]. Both bowtie and C apertures are ridge apertures and have been extensively studied. Some numerical and experimental results related to bowtie apertures are shown in Figure 4. A bowtie aperture is a counterpart of the bowtie antenna, and can be formed by loading a rectangular aperture with a pair of conducting triangular ridges, forming a narrow gap in the center. Under the illumination of a light polarized across the gap, an LSP resonance will be excited in both ridges, driving charges to the two apexes where the lightning rod effect occurs. In a modal study [43] shown in Figure 4A, the large field intensity near the entrance and exit surfaces of the aperture demonstrates the LSP excitation and non-resonant amplification. In addition, both a characteristic TE_{10} waveguide mode and an SP mode can be observed in the gap between two metallic walls, where the TE_{10}-like mode is not cut off as in small rectangular apertures. The coupling of the two modes efficiently delivers photon energy to the other side [43], leading to an enhanced transmission. To illustrate this, near field imaging in Figure 4B shows a peak for the bowtie aperture with a 36 nm gap at 633 nm, while the small square apertures do not transmit [43]. The larger square, even with the same opening area as the bowtie aperture, allows almost no light transmission. The 450×50 nm² rectangular aperture results in a propagation mode, but without field confinement. These measurements directly confirm that the bowtie aperture, as a type of ridge aperture, is capable of enhancing the optical transmission at a subwavelength scale. Figure 4C shows the transmitted field enhancement as a function of wavelength for a bowtie aperture on glass. With a 20 nm gap, a 105 nm aperture in a 60 nm-thick gold film resonates at 800 nm, with a full-width at half-maximum (FWHM) spot size of 36.5×36.5 nm². When a bowtie aperture is placed right above a media stack with a small air gap, a sub-diffractive hot spot can be produced in the recording layer [46], Figure 4D shows the FWHM at the surface of an 8 nm-thick FePt layer, which is separated from the aperture by a 4 nm gap. For a small 5 nm gap, the optical spot in FePt is only 19×19 nm² [46]. Evidently, the spot size in the recording media is most influenced by the gap size of the bowtie aperture.

The ability of C aperture to support the TE_{10}-like guided mode have been studied for various materials and with different illumination methods [34, 35]. It was pointed out in [34] that the TE mode in an aluminum C aperture hybridizes a TM character that originates from the SPs along metal boundaries. The wavelength dependent peak intensity and the intensity profile at resonance of a gold C aperture with a 20×20 nm² gap are shown in Figures 5A and B, respectively. Similar as the results in Figure 4D, the simulation results shown in Figure 5 also include the recording medium [47]. At the resonant wavelength, the FWHM in the recording medium is 39×34 nm² for an air-filled C aperture. It is noted that if the gap region is very small and the aperture is very wide, there is a chance to produce unwanted elongated spot as shown in Figure 5C, because of the propagation of SP along the ridges called channeled SP. One way to mitigate the channeled SP is to have a flare angle to open the channel, resulting in a half-bowtie shaped aperture as shown in Figure 5D. For this half bowtie aperture, the near field distribution is confined. It is noted that for bowtie, C, and half-bowtie apertures, it is quite straightforward to alter the gap dimensions s and d.
Figure 4  (A) |\(E_y|^2\) (left column) and |\(E_z|^2\) (right column) intensity distributions in a 160 nm gold film illuminated by a 633 nm y-polarized light from the substrate side. Top: \(yz\)-plane; bottom: \(xy\)-plane. The simulated bowtie aperture has a dimension of 190×230 nm\(^2\), with a gap of 36 nm. (B) NSOM images of the sample in the inset. Small square: 36×36 nm\(^2\); bowtie aperture: 190×230 nm\(^2\); larger square: 136×136 nm\(^2\); rectangular aperture: 450×50 nm\(^2\). The last three apertures have about the same opening area. Adapted with permission from Ref. [43] (A and B). Copyright 2006, Springer. (C) Wavelength dependency of the field enhancement (at a point 10 nm from the center of aperture exit) for a 105 nm bowtie aperture in a 60 nm gold film with a gap size of 20 nm. Also shown is the intensity |\(E|^2\) distribution at 800 nm with a peak intensity 61 times of the incident intensity. (D) Dependence of FWHM in \(x\) and \(y\) directions on aperture gap size \(d\) for a 200 nm bowtie aperture in a 100 nm silver film. From Ref. [46].

to generate elongated spots to match the bit aspect ratio on the recording track [46]. As an example, Figure 5E shows a spot produced by a 345 nm half-bowtie aperture with \(s=15\) nm, \(d=5\) nm. The spot size is about 37×16 nm\(^2\), with an aspect ratio of approximately 2.3.

One of the NFTs designed by HGST (San Jose, CA, USA) is very similar to a C aperture antenna [13], irradiate by light polarized in the horizontal direction in Figure 6A. The orange colored part is made of gold and forms an E-shape; therefore, this NFT is also called E antenna. The notch at the center concentrates in a small volume the surface charges generated through a plasmonic resonance in the body. Figure 6B compares the absorption profiles produced by the antenna with and without the notch at a wavelength of 780 nm. It is seen that without the notch, the strong absorption around the left surface of the body

One of the NFTs designed by HGST (San Jose, CA, USA) is very similar to a C aperture antenna [13], irradiate by light polarized in the horizontal direction in Figure 6A. The orange colored part is made of gold and forms an E-shape; therefore, this NFT is also called E antenna. The notch at the center concentrates in a small volume the surface charges generated through a plasmonic resonance in the body. Figure 6B compares the absorption profiles produced by the antenna with and without the notch at a wavelength of 780 nm. It is seen that without the notch, the strong absorption around the left surface of the body
corresponds to a large surface charge density, indicating a plasmonic resonance supported by the body. With the assistance of the lightning rod effect produced by the notch, the FWHM spot size was reduced from more than 200 nm to <40 nm, along with increased peak intensity within the spot by 7 times [13]. It is noted that practically a magnetic pole is integrated to the open area opposite to the notch as illustrated in Figure 6A. Therefore, the E-antenna is essentially a C aperture with a small ridge/notch surrounded by gold and pole material. The optical near field of this E-antenna has been characterized using a scattering type near field scanning optical microscopy (s-NSOM). This method was based on oscillating the AFM tip and analyzing the collected scattering signals at higher harmonics of the tapping frequency to suppress the background noise that comes from the tip shaft and sample surface [48, 49]. Figure 6C shows an optical near field spot size of 60×42 nm² is produced by this E antenna.

Another type of design utilizes plasmonic resonance in a metallic structure (instead of an aperture in bowtie or C aperture antenna) and a smaller nanostructure for further field localization and enhancement. Examples are a “lollipop” design by Seagate Technology and a “nanobeak” design by HGST. The lollipop design indicated by the red dashed outline in Figure 7A [12] consists of a 200 nm diameter gold disk and a 15 (length)×50 (width) nm² gold peg. The design of the waveguide is such that it results in a vertically polarized net field at the focal point to correctly excite the NFT [12]. For numerical modeling, it is placed 7.5 nm above a 12.5 nm Fe recording layer. The larger circular disk acts as the LSP resonator and the smaller peg further localizes the optical energy via the lightning rod effect. In addition, the design takes into account the effect of the recording medium which produces an imaging effect and introduces an additional enhancement to the energy confinement and coupling.
Figure 7B shows the simulated spectral coupling efficiency (will be discussed later), which indicates that the NFT is designed specifically for the laser wavelength near 800 nm. Figure 7C shows the optical energy profile produced by this NFT in the central plane of the Fe layer, with an FWHM of about 70 nm. The SP resonances in lollipop both without and with the presence of the medium have been experimentally characterized via a pump-probe photothermal measurement and by taking AFM topography of illuminated devices [51]. In addition, the disk resonator can be replaced by a rectangular resonator to enhance the excitation efficiency [52]. The nanobeak antenna [50, 53, 54] is essentially a triangular antenna with a 3D beaked apex, as shown in Figure 7D. With the lighting rod effect taking place along both in-plane and out-of-plane directions, field enhancement occurs at the tip of the beak, and 40 nm marks were successfully written onto the medium [50]. Numerical results in Figure 7E show that a flat triangle without the beak produces a FWHM of about 25×25 nm², compared to a 15×20 nm² spot with the use of the beaked design. The intensity profiles are computed on the surface of the recording medium. The nanobeak antenna can also be integrated with a thin film wing to form a SP waveguide [14].

Apart from the mechanisms used in the designs discussed above, there are other effects that can also be applied in NFT design, such as the dual-dipole effect existing in two closed spaced nanoparticles [5, 15] and the Fabry-Perot effect in relatively thick films [15]. Other methods that manipulate the shape of NFTs include canted antennas or apertures [15, 55], butted-grating structure [56, 57], and tapered plasmonic waveguide [58, 59]. The last is a 3D tapered metal-insulator-metal (MIM) multilayered structure terminated in a nanometer sized cross section which determines the cross-track spot size in the recording media. The fundamental mode in the MIM waveguide is supported without cutoff [58]. The idea of plasmonic taper that produces a hot spot with a significant fraction of energy deposited at the tip was first established in [60]. In [14], the thin film waveguide with a nanobeak antenna integrated at the end can be understood as a variant of a tapered plasmonic waveguide. Ultimately, the spot size in the medium are complex results that determined by the smallest structure dimension, which has a direct impact on the manufacturing requirements and the overall process capability for HAMR.

The optical spot size in the recording medium is often used as a figure of merit (FOM) to characterize NFT, as has...
been discussed above for several designs. It is also commonly used for other nanofocusing devices when evaluated in free space [61]. The power transmitted or scattered by the nanostructures to the desired region is another important criteria or an FOM for evaluating NFT’s performance. The total diffracted power could exceed the irradiation on the open area of the aperture or the area of a nanoparticle, i.e., EOT [19–24]. It should be noted that an issue associated with the transmittance FOM is that it is often evaluated in the absence of the recording medium. The optical and thermal properties of the recording medium affect the heating of the medium by NFT. Therefore, a more proper FOM based on the coupling efficiency, should be defined as the percentage of the total focused power dissipated as heat in the recording medium within a confined spot, as has been reported by Seagate Technology [12] and HGST [13]. However, establishing an FOM based on heating can be difficult since the specific recording materials vary among different companies.

In [15], a standard geometry, including a solid immersion lens (SIL) and the recording stack (10 nm cobalt layer and 105 nm gold heat sink), is used for simulating and comparing different NFT designs illuminated by a focused laser beam. The coupling efficiency is computed as the fraction of the incident optical power coupled into a 50×50 nm² area in the cobalt layer, at the resonant wavelength of each NFT. The results indicate that canted bowtie antenna (4.1%), beaked triangle antenna (3.4%), C aperture (2.8%), and bowtie aperture (2.3%) are promising choices for NFT. In other calculations, the coupling efficiency of a lollipop NFT is found to reach 8% as shown in Figure 7B at the resonance, considering a 70×70 nm² region of a 12.5 nm thick Fe medium [12]. HGST’s E-antenna is modeled to couple about 11.7%–14%, depending on different pole materials, of the waveguide optical power to the 50 nm cobalt medium within a 50×50 nm² footprint [13]. For the nanobeak NFT, an 8% efficiency is estimated with a region of 50×50 nm² in an 8 nm thick FePt recording layer [14]. It also needs to be noted that this coupling efficiency depends strongly on the NFT-medium separation distance because of the evanescent nature of local fields [12, 62]. For a lollipop NFT, It has been reported that the coupling efficiency reduces rapidly from 8% to only 1% if the NFT-medium separation increases from 8 nm to 20 nm [12]. Concerning the difficulty of designing and modeling media and making fair comparisons between head designs, the Advanced Storage Technology Consortium (ASTC) provided a standard media stack and FOMs for modeling HAMR NFTs [63]. In addition to the thermal spot size in the media and the coupling efficiency, other FOMs including thermal efficiency, thermal gradient in the media, and normalized peak temperature [63], which are related to the heating effect in NFT which will be discussed in Section 3.

Another aspect worth highlighting is the different wave guiding and coupling schemes applied in different configurations. These are important for the overall head design that aims at efficient light delivery from the laser to the recording medium. The laser could be directly focused onto the NFT by a conventional objective lens [35, 53] or a SIL [15, 64]. This resembles the Otto technique for exciting SP [65]. A broad wave vector spectrum is provided by the total internal reflection and couples to the NFT film by generating SPs. As introduced earlier and illustrated in Figure 2, the waveguide linking the light condenser and the NFT takes various forms. In [12], the laser light is coupled to a transverse electric (TE) mode supported by the PSIM via a dual offset grating with an efficiency of about 50%, and then focused by the parabolic mirror to generate a vertical-polarized net field at the focal point where the lollipop locates. Laser can be coupled to dielectric waveguides in an endfire manner, as demonstrated in [13] where the power is then guided by the lowest order transverse magnetic (TM) mode and ~40% arrives at the E-antenna. It has been pointed out that the evanescent coupling is efficient for exciting SPs with a strong confinement and is suitable for integration of plasmonic components with photonic networks [66]. This is also widely adopted in HAMR and a perfect transfer of power from the dielectric to the plasmonic waveguide is achievable [67]. The plasmonic waveguide, in a shape of rectangular [68], needle with a triangular cross section [67] or taper [14], acts as a NFT at the same time by guiding optical waves to the medium surface. As a modification of the needle design, a magnetic core antenna was proposed in [69], where the interior of the plasmonic waveguide is replaced by magnetic materials and forms an extension of the magnetic pole. This improves the overlap of the magnetic field from the pole with the heating profile from the NFT [4, 5, 70], leading to a better performance of the head. Other techniques that guide the waves to a NFT include using a tapered MIM waveguide [59] or an aperture surrounded by grooves [25, 26].

From a circuit theory point of view, the low coupling efficiency in a HAMR system indicates an impedance mismatch between NFT and the media. The media, together with the air gap, are equivalent to terminating loads [34, 59, 68]. When modeled together with the recording medium, the tapered MIM waveguide design turns out to have a very large impedance that better matches the load (air gap + media) and thus outperforms the lollipop and the E-antenna [59]. An optimization of the media
properties was carried out in [68]. Similarly, it is found that the maximum coupling efficiency is achieved when the resistance (real part of the impedance) of the plasmonic waveguide matches that of the load. At the same time, the capacitive impedance of the air gap cancels the inductive part from medium. Although the circuit analysis neglects possible higher order modes and simplifies the complex geometry, it provides a direction for qualitatively optimizing the system.

3 NFT self-heating and material choice

Because of the introduction of thermal energy into HDD in HAMR, the performance of all elements will be impacted by thermal effects. Prime concerns include the instability of the slider and the failure of materials [4, 5, 12, 61, 70]. Thermal expansion in NFT can cause an NFT protrusion to the recording media surface, which will require a better control of the air gap and the surface roughness to avoid contact between NFT and medium surface [12]. Temperature rise also changes the NFT-writhing pole separation, which may cause variations in coupling efficiency by attenuating the resonance [12] since the writing pole is made of metal and is part of the resonance structure. Thermal modeling [71–73] has been carried out to study dependence of NFT temperature rise on absorbed power, its size, and the NFT-pole separation, as shown in Figure 8 [72]. The temperature increases linearly with the heat dissipation in NFT and could rise by several hundred degrees. A recent proposal is that the medium could be directly illuminated by the waveguide first to get a moderate background temperature rise before being heated by a NFT locally [74]. This two-stage heating scheme reduces the local thermal load to NFT and thus could possibility prolong its lifetime, but it also increases the possibility of interference and even erasure between adjacent tracks. Here we focus on plasmonic heating in NFT only and the investigation for alternative plasmonic NFT materials.

To get a better understanding of the effects of optical properties on NFT performance, we start with the equation for computing dissipated power density in NFT [75].

\[ P = \frac{1}{2} \Re(\sigma) |E|^2 + \frac{1}{2} \epsilon_0 \omega \Im(\epsilon) |E|^2 \]  

where \( \epsilon_0 \) is the vacuum permittivity and \( \omega \) is the angular frequency of the laser. The relationship between the relative permittivity \( \epsilon \) and the conductivity \( \sigma, \epsilon = 1 + i\sigma/\epsilon_0 \omega \) [66], has been applied. Eq. (2) indicates that the dissipation in a lossy medium is determined by not only the imaginary part of the permittivity \( \Im(\epsilon) \), but also the peak field intensity \( |E|^2 \). Therefore the heating effect in NFT can be minimized by reducing \( \Im(\epsilon) \) and \( |E|^2 \), the latter is largely determined by the real part of the permittivity \( \Re(\epsilon) \). However, in most cases, a large magnitude of \( \Re(\epsilon) \) indicates a strong resonance, large field enhancement, and lateral energy confinement, which is favorable for NFT in HAMR application. Therefore apparently there is a compromise between improving field localization/enhancement and minimizing heat dissipation. As a first order estimation, consider a sphere in the quasistatic approximation, an FOM can be defined by combining both parts of the permittivity to evaluate the overall performance of LSP systems [76] as:

\[ Q_{LSP} = \frac{-\Re(\epsilon)}{\Im(\epsilon)} \]  

In other words, a desirable NFT material should have a minimized \( \Im(\epsilon) \) and a high \( \Re(\epsilon) \), and thus a relatively large \( Q_{LSP} \). Real and imaginary parts of permittivities of a number of metals [17, 77], gold, silver, aluminum, chromium, and titanium, are shown in Figures 9A and B. The FOM \( Q_{LSP} \) as defined in Eq. (3), was also plotted in Figure 9C for comparison. For a diode laser at near-IR wavelengths, silver has the smallest \( \Im(\epsilon) \) in the interested range and a relatively large \( \Re(\epsilon) \), thus the best \( Q_{LSP} \) but suffers from chemical stability against, for example, possible decomposition of lubricants on the medium surface.

![Figure 8](image-url)  

**Figure 8** Temperature rise in NFT as a function of absorbed power at different transducer sizes (W×L) and different NFT-pole distances under 50 nm input laser power, 4 nm fly height and 6.5 m/s fly speed. A 15×10⁻⁶ m²K/W boundary thermal resistance (BTR) between the recording layer and heat sink is included in the model. A C-aperture NFT in gold film was used in this simulation. Figure from [72].
Aluminum is naturally excluded as an NFT material because of the interband transition around 800 nm as can be seen from the peak of Im(ε) spectrum and a low melting temperature. Chromium and titanium have similar properties, and they can hardly support plasmonic modes in the near-IR. At present, gold is widely used as the NFT material because of its chemical stability, melting point much greater than the Curie point of popular recording medium (~750 K for FePt, [5]), and high thermal conductivity. However, nano-structured gold suffers from poor thermal stability (high ductility) at temperatures much below its melting point. For example, the stress was found to start relaxing at a low temperature of 100°C for gold, which could be a result of the highly mobile grain boundaries [79]. It needs to be pointed out that Eq. (3) is only exact for spheres in the quasistatic limit. For particles that are large and take complex geometries, as for NFT designs in HAMR technology, Eq. (3) and Figure 9C could be unreliable. This issue has been discussed in [80], which suggests a generalized form of the scattering efficiency, called the near-field intensity efficiency, as a more comprehensive FOM for large scatterers in LSP applications. Additionally, a recent work [81] demonstrated the absorption efficiency of particles as the critical FOM for local heating applications.

Recently, there is a significant interest in searching for alternative low-loss plasmonic materials [76, 82–86], and applying these materials in SP, LSP, transformation optics, and metamaterials. Some of these alternative plasmonic materials for visible and near infrared frequencies may offer a possibility of achieving high performance for NFT devices. As discussed in [84], the reported alternative plasmonic materials can be loosely categorized as metallic alloys [78, 79, 82], semiconductor-based [76], ceramic materials [85], 2D materials such as graphene [76, 86] and organic materials [87]. Among these materials, metallic alloys, semiconductor-based transparent conducting oxides (TCOs), and transition-metal nitrides [84, 86, 88] can be promising for HAMR application near a wavelength of 800 nm. To tune standard semiconductors such as silicon and germanium into metallic materials near this wavelength, an ultrahigh doping larger than 10²¹ cm⁻³ is required, which challenges their usage as alternative plasmonic materials because of additional concerns such as the solid solubility limit, crystal defects created that limit the carrier concentration, and the difficulty in maintaining a high carrier mobility [76, 86]. Alloying metals with different proportions of each element will create a unique band structure that shifts the interband transition to a less critical spectral range. As shown in Figure 10A, the original bands I (centered around 667 nm) and II (centered around 333 nm) of gold can overlap at about 500 nm in the alloy, leaving other part of spectrum less lossy [82]. For TCOs and metal-nitrides, optical properties of thin films were characterized and fitted with a Drude+Lorentz oscillator model [84]. Figures 10B and C illustrate a comparison of optical properties of TCOs and metal-nitrides with those of gold and silver [84]. It shows that TCOs become metallic and have lowest loss in the near-IR compared to gold and silver. Its drawback is its relatively low -Re(ε). Metal-nitrides have comparable properties as gold and provide potential usage in visible frequencies. The general guideline behind the two approaches described above is to reduce the free electron density in metals since loss in conventional noble metals is closely associated with the large free electron density as indicated by Drude model, or increase the free electron density in semiconductors and ceramics by heavy doping [83].

Comparative numerical and experimental studies have also been carried out for specific structures. Figure 11 shows the maximum field enhancement, absorption, and extinction cross sections of LSP resonance modes, which
are related to the HAMR technology. The geometry used is a nanosphere with a diameter of wavelength/10 (in the quasistatic limit), surrounded by a host material of refractive index 1.33 [86]. The materials investigated are noble metals, metal nitrides, and TCOs. The peak field enhancement shown in Figure 11A is the same as the FOM given by Eq. (3) for spheres. Similar as the conclusions drawn from Figure 10B and C, metal-nitrides such as TiN and ZrN provide LSP resonances between 700 nm and 1000 nm. The peak absorption and extinction cross sections of TiN and ZrN are a little larger than those of gold, as shown in Figure 11B and C. Overall, the nitrides are comparable with gold at the wavelength range from about 500 nm to 1000 nm. Although they do not outperform noble metals, the optical properties can be improved by affecting deposition processes. In addition, metal-nitrides can be attractive because of controllability, superior thermo-mechanical properties, and chemical stability [84, 88]. For example, TiN has an extremely high melting point which is larger than 2900°C, making it potentially useful in plasmonic thermal applications.

The search for alternative plasmonic materials should always be associated with specific applications, the interested spectrum range, and a proper choice of FOMs. The alternative materials discussed above are investigated to be applied particularly in the visible and near-IR ranges for plasmonic and metamaterial applications. The optical properties could be significantly affected by fabrication processes and experimental conditions, for example, substrate material and temperature, processing parameters, and thickness of deposited film [78, 79, 86, 88]. Figure 12 shows a parametric study of optical constants of NFT material on the absorption rate and the coupling efficiency [72]. The coupling efficiency is defined in the
same way as that discussed in Section 2, with a medium volume of 50×50×10 nm³. The NFT used is a C aperture, same as that for Figure 8, and both recording media and waveguide were included in this model [72]. Since the optical constants are related to permittivity by 

\[
Re(\varepsilon) = n^2 - k^2 \quad \text{and} \quad Im(\varepsilon) = 2nk,
\]

as n reduces, Re(\varepsilon) remains almost unchanged since n is typically much smaller than k, and Im(\varepsilon) decrease. With a reduced n, absorption reduces rapidly and the coupling efficiency increases as expected. Smaller k leads to a smaller Im(\varepsilon), but higher absorption is observed, which could be caused by the existence of a stronger resonance at which the absorption increases with |E|^2 [72]. The dependency of coupling efficiency on k is somewhat complicated because both Re(\varepsilon) and Im(\varepsilon) vary simultaneously. In [88], a nanorod NFT made of TiN (n=0.99, k=3.6 at 830 nm) was modeled, showing a 1.1% coupling efficiency which is about one third of that of a gold nanorod NFT.

Because of the complexity of the HAMR system, optical properties alone are not sufficient to determine the NFT performance. Other properties such as mechanical properties also need to be considered. Researches have been performed to characterize alternative plasmonic materials, for example, gold alloys [78] and silver alloys [79]. The hardness of gold can be enhanced by ~32% if doped with copper at a concentration of 10.3%. The corresponding FOM, defined as 

\[
3*|Re(\varepsilon)|/|Im(\varepsilon)|,
\]

is about 30 at the wavelength of 830 nm, while the FOM of pure gold is about 43 at the same wavelength [78]. In [79], it was found that the resistance to grain growth in an AgPd alloy improves with an increasing palladium concentration, which helps to prevent the plastic deformation. A 100–150 nm thick AgPd (5.8 at% Pd) film provides about a two-fold increased hardness (normalized to pure gold); a thermal conductivity of 160 W/(mK) that is larger than gold and silver under the same conditions; and more importantly, a FOM close to 30 at 830 nm [79]. Thus, gold and silver alloys are promising NFT materials by providing improved hardness and higher stress relaxation and creep resistances, with still acceptable optical properties. To find the best alternative NFT material for HAMR, both optical and thermo-mechanical properties, as well as fabrication and integration issues should be brought together into consideration.

To conclude this section, we note that the widely used multi-physics model for simulating the electromagnetic coupling in HAMR is questionable, as has been examined in [89], and the origin lies in the failure of macroscopic Maxwell equations and constitutive relations in materials for nanoscale systems. The interaction of the highly focused laser beam with metallic materials will induce non-linear and non-local effects. As a result, different zones for energy penetration have to be considered. In addition, the conventional Joule’s law expressed by Eq. (3) and the heat conduction equation are not applicable to the transducer surface. SP oscillations in NFTs improve the coupling efficiency for a HAMR system, but also accelerate components failure. It is pointed out in [89] that a rethink of the local heating process in NFT helps to explain the short, lower than expected lifetime of NFTs.

4 Conclusions

An areal density of 1 Tb/in² is estimated to be the limit for the present HDD products using the PMR technology, due to the requirement of thermal stability and available magnetic write fields. By incorporating thermal energy into the head to locally reduce the coercivity of the medium in a sub-diffraction-limited area, HAMR becomes one of the most promising technologies to keep increasing the areal density. The key component in HAMR is a NFT that needs to deliver sufficient fraction of incident optical energy into the recording medium within a region far below the diffraction limit. NFTs that based on nanoantennas and nanoapertures take advantage of various underlying physics, including resonant and non-resonant amplifications to achieve sufficient spatial resolution and coupling efficiency. On the other hand, self-heating in NFT is a concern for the NFT performance. Less dissipation in NFT and more power coupled to the recording medium are desirable. The discovery of low-loss plasmonic materials can open up possibilities for better devices for the development of the HAMR technology.
**Acknowledgments:** N.Z. and X.X. acknowledge the support from the Defense Advanced Research Projects Agency (Grant No. N66001-08-12037), the National Science Foundation (Grant No. CMMI-1120577), and the Advanced Storage Technology Consortium (ASTC).

**References**


