AIP | Review of Scientific Instruments

Apertureless scanning near field optical microscope with sub-10nm resolution

Alpan Bek, Ralf Vogelgesang, and Klaus Kern

Citation: Rev. Sci. Instrum. **77**, 043703 (2006); doi: 10.1063/1.2190211
View online: http://dx.doi.org/10.1063/1.2190211
View Table of Contents: http://rsi.aip.org/resource/1/RSINAK/v77/i4
Published by the American Institute of Physics.

Apertureless scanning near field optical microscope with sub-10 nm resolution

Alpan Bek,[a] Ralf Vogelgesang, and Klaus Kern[b]
Max Planck Institut für Festkörperforschung, D-70569 Stuttgart, Germany

(Received 19 December 2005; accepted 5 March 2006; published online 4 April 2006)

We report on the implementation of a versatile dynamic mode apertureless scanning near field optical microscope (aSNOM) for nanoscopic investigations of optical properties at surfaces and interfaces. The newly developed modular aSNOM optomechanical unit is essentially integrable with a multitude of laser sources, homemade scanning probe microscopes (SPMs) as well as commercially available SPMs as demonstrated here. The instrument is especially designed to image opaque surfaces without a restriction to transparent substrates. In the description of the instrument we draw frequent attention to various possible artifact mechanisms, how to overcome them, and we present effective checks to ensure true near field optical contrast. Lateral optical contrast in optical amplitude and phase images below 10 nm is demonstrated. © *2006 American Institute of Physics.* [DOI: 10.1063/1.2190211]

I. INTRODUCTION

Nano-optical investigations suffer from the mismatch of far field collection optics and detectors with the near field nature of optical interactions at small material volumes. In the detection of subdiffraction optical interactions, spatial information are lost by the convolution with the instrumental response of diffraction-limited imaging systems. Often, the minute optical power of interest is buried in a large parasitic signal background.

Several avenues have been proposed to overcome the discrimination dilemma: essentially transparent samples such as biological material may be investigated with 4π microscopes that achieve spatial resolution beyond the 2π diffraction limit through the simultaneous illumination with two conventional microscope objectives in confocal opposition.[1,2] Surface nanostructures have been made accessible by scanning near field optical microscopy (SNOM),[3–8] which uses a nanoscopic aperture—scanned through the near field of the surface—to reduce the illumination/collection volume. Throughput issues and finite penetration depths of optical radiation into the aperture material, however, limit the lateral resolution of SNOM to about a tenth of the wavelength. More recently, field enhancing near field optical microscopy (FENOM) techniques have been developed. They use highly localized optical fields at sharp geometric features of (usually metallic) surfaces or in nanoscopic gaps between them, often in conjunction with plasmonic or other resonances. Here, the achievable lateral resolution does not as much depend on the wavelength of the excitation as on the sharpness of the field gradient generating surface topography, which makes FENOM very interesting especially for longer wavelengths.

In this article we present in detail our implementation of a FENOM system of the scattering type or apertureless SNOM [aSNOM (Refs. 9–16)]. Our setup is particularly flexible, with a broad range of applications, yet it provides lateral optical resolution down to the sub-10 nm range. The working principle of aSNOM is a single, well controlled hot spot caused by a mechanism somewhat reminiscent of the lightning rod effect.[17–20] Upon external illumination of the tip of an atomic force microscope (AFM) or a scanning tunneling microscope (STM),[21,22] an enhanced optical near field forms at its apex,[23,24] whose interaction with the sample surface results in scattered far field radiation. The local field enhancement varies from tip to tip, though, being a function of tip material and more importantly of the exact geometry at the apex.[25–31] Higher field enhancement is generally expected for sharper tips with smaller effective apex radii. Metallic tips or metal-coated Si tips have given good results. Resonant particles may also be attached individually.[32] Occasionally, a resonant structure may form at the apex, which yields extraordinarily high field enhancement. This behavior is still not well understood, though, and the average enhancement factors in aSNOM are orders below, for instance, those found in the stochastically forming hot spots responsible for surface-enhanced Raman scattering. In a recent study it was reported that the field intensities at the apex were so high that even nonlinear effects such as second harmonic generation and its polarization selection rules at the tip apex could be analyzed.[33]

aSNOM can be used with evanescent fields, for example, in the vicinity of waveguide structures or at totally internally reflecting interfaces. Being nonpropagating, such fields can only be observed in the near field. Here, the aSNOM tip acts as a rescatterer that converts them into propagating fields detectable in the far field.[34–36]

Thin film samples on transparent substrates may be directly illuminated through the bottom of the substrate by strongly focused excitation with high numerical aperture (NA) objectives. Given sufficient local enhancement and a

[a]Present address: Ludwig Maximilians Universität, München, Germany; electronic mail: bek@lmu.de
[b]URL: www.fkf.mpg.de/kern

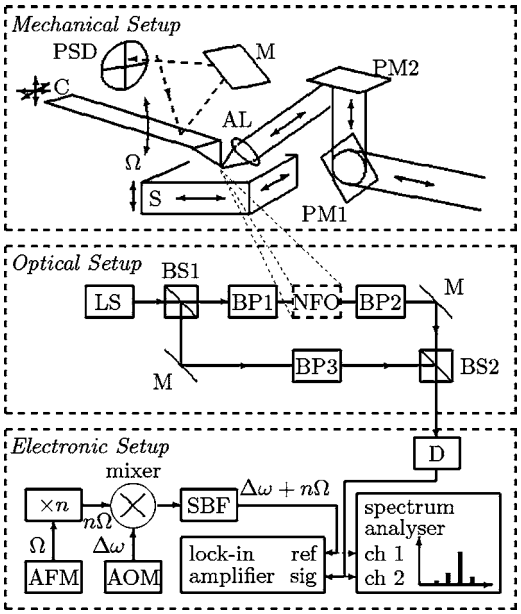

FIG. 1. Overview of the three major systems of the aSNOM. *Mechanical setup*: S: sample stage, C: AFM cantilever, M: mirror, PSD: position sensitive detector, Ω: noncontact mode oscillation frequency of the cantilever, PM: periscope mirror, and AL: aspheric focusing lens. *Optical setup*: LS: laser source, BS: beam splitter, BP: beam preparation, M: mirror, NFO: near field optical interaction region, and D: detector. *Electronic setup*: Ω: oscillation frequency of the AFM cantilever, n: harmonic index, $\Delta\omega$: AOM drive frequency, and SBF: sideband filter.

preferably rarefied sample material background (e.g., in the study of one-dimensional objects) the local optical signal may even exceed the background and enable direct local optical spectroscopy without the need for further discrimination.[30,37,38]

In the more general case, however, some scheme of discrimination needs to be applied. This holds particularly in the study of surfaces of opaque samples, where far field optics have to be used for both excitation and collection and the parasitic background signals are considerably high. Quite successful in this context is the harmonic modulation of the tip-to-sample distance with subsequent demodulation by lock-in detection, referred to as dynamic aSNOM.[10,24,39] Demodulation of the scattered field at higher order cantilever harmonics with lock-in detection has proven to reveal complex optical material constants at surfaces at subwavelength scale.[40] Exciting demonstrations of this technique have been given, including subsurface imaging at subwavelength scales,[41] investigations of phonon or molecular vibrational resonances[42,43] at mid-IR wavelengths, or the mapping of optical field mode distributions and optical phase in integrated optical waveguide circuits.[44]

In view of the broad variety of possible applications of aSNOM, we followed the design principles of flexibility and modularity as much as possible. This regards exchangeable laser sources, transparent as well as opaque sample classes, different choices of AFM tips, and a versatile optical signal detection scheme, which is set up as a unit on a portable breadboard. Our aSNOM consists of three main subsystems, displayed in Fig. 1, viz., mechanics, optics, and electronics. The mechanical motion control is done in the vertical direction by a commercial AFM and a sample stage in the horizontal. The optical arrangement comprises fiber-delivered laser radiation, a Mach-Zehnder interferometer, operated in heterodyne or homodyne mode, and a fast sensitive detector. Central to the electronic setup is a fast lock-in amplifier, which filters the desired signal with an appropriately synthesized reference frequency.

In the following, we describe these three subsystems with some emphasis on various design considerations. We outline a quantitative alignment procedure and describe a number of tests that ensure true near field optical contrast. Next, we demonstrate the instrument's capability of mapping amplitude and phase of optical fields with spatial resolution below 10 nm and how such maps help us to derive material contrast with the same resolution. Along the way, we make frequent reference to possible artifact mechanisms and how to avoid them. Finally, we summarize.

II. DESCRIPTION OF THE aSNOM

A. Mechanical setup

1. AFM

An ideal AFM system for dynamic aSNOM operation generates a true noncontact harmonic cantilever oscillation. It is crucial to avoid touching the sample surface with the tip, which induces undesired higher harmonics of the cantilever drive frequency in the tip motion. Such higher harmonics of the cantilever motion degrade the near field contrast obtained from the sample surface in connection with the specific detection scheme that we employ, because it relies on optical higher harmonics generation.[45]

In our aSNOM setup we use a commercial amplitude modulation AFM operated in noncontact mode (AutoProbe M5 from Veeco; formerly Park Scientific Instruments). Its 100 μm scanner head offers a travel range of approximately 120 μm in both of the x and y scan axes and 9 μm along the z axis. The AFM is equipped with an imaging camera unit for monitoring the top of the sample area of interest. The beam from a 670 nm diode laser is focused on the reflective side of an AFM cantilever and the reflected beam is positioned on a quadrant diode for beam deflection detection. Because this deflection signal is used directly to synthesize the reference signal for lock-in demodulation, it is crucial to suppress any signals due to other light sources. A simple field-of-view reduction cover proved effective to limit interference from, e.g., too strong room illumination. The cantilever is attached at 13° with respect to the sample surface and a piezocrystal slab drives it with a sinusoidal excitation of angular frequency Ω to let the tip swing along the z axis. A conventional amplitude feedback mechanism controls the z position of the cantilever, giving the usual topography information as the tip is raster scanned over the sample surface.

The AFM scan head is located about 10 cm above the granite base and its spatially generous design allows for simple prototyping of the optical and mechanical modules of the aSNOM with convenient access to the tip region. Below the AFM head a sample stage is mounted, which provides the sample-scanning acquisition mode in addition to the tip-scanning mode afforded by the AFM.

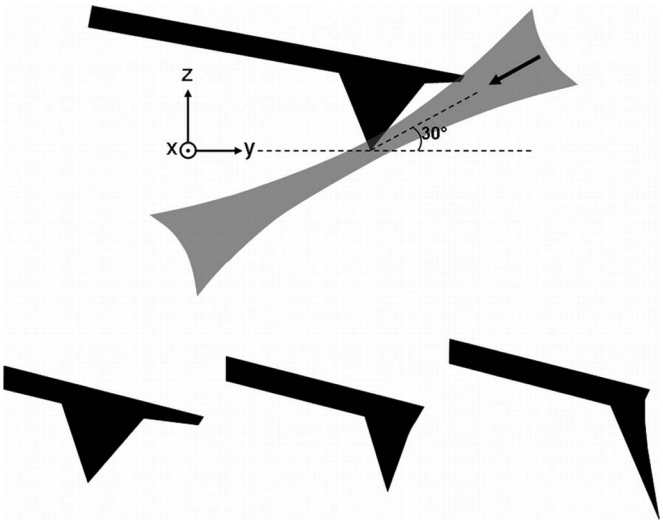

FIG. 2. Upper half: Oblique incidence illumination of the tip apex with a focused laser beam. The cantilever body limits the incidence angle. Lower half: Cantilever tip profiles from three different suppliers.

2. Tip shapes

Empirically, we found commercially available noncontact mode AFM tips of any shape, pyramidal or conical, to work for aSNOM operation. Uncoated tips may not provide enough local field enhancement, but a 10–20 nm thick gold or silver coating generally allows us to generate recordable near field optical signals. Concerning convenience and reliability of tip alignment in the optical focus, what matters most for the tip selection is the geometrical accessibility of the tip apex for illumination. Ideally, the exciting optical field is centered at the tip apex and should illuminate as little of the tip body and sample surface as possible. The upper part of Fig. 2 shows how geometry constraints can pose a challenge, as protruding parts of the cantilever may partially obstruct the optical beam or even lead to spurious scattering signals. Typical side-view profiles of AFM tips are displayed in the lower row of Fig. 2. Of these, the rightmost profile is best suited for aSNOM optical alignment. As a bonus, the tip apex alignment is easily observed even when viewed from the top as the tip protrudes outward from the upper cantilever edge.

3. Sample stage

For constant optical alignment conditions the sample is scanned under the AFM tip with a digitally controlled three-axis piezostage (Physik Instrumente P-517.3CD scanner, E-710.3CD controller) operated hysteresis free in closed loop mode. The travel range and positioning precision are 100 μm$\pm$1 nm in the lateral (x/y) and 20 μm$\pm$0.1 nm in the vertical (z) axis. A miniature z translation stage on top of the piezoscanner provides manual coarse approach of the sample.

4. Sample versus tip scanning

Our aSNOM allows us to obtain optical images scanning either the tip or the sample, with the optical focus being fixed in space. The tip-scanning option of the commercial AFM platform is already optimized for fast acquisition of internal and external data input channels. In tip scans of areas larger than the optical focus, however, the near field image clearly exhibits an intensity envelope proportional to the focal intensity (see also Fig. 7). This variation of the optical excitation is fixed in sample-scanning aSNOM operation due to the constant spatial relation of optical focus and local field enhancing tip. To achieve optimal alignment, the tip apex is first centered at the optical focus, then the sample is brought in engagement with the AFM tip from underneath and scanned laterally.

Sample scanning a rectangular area is done in a line-by-line fashion. Compared to a pixel-by-pixel data acquisition, this is about ten times faster, provides better line-to-line alignment, and increases the image quality by reducing long term drift effects considerably. The sample stage is scanned through the fast direction at constant speed and a number of instruments simultaneously record various data channels at a given rate. For synchronization, the sample stage provides a trigger pulse at the beginning of the linear motion section of each line. After completion of a line, all channels are read out from the respective instruments, which slows the data acquisition somewhat down in comparison to tip scanning with its optimized all-internal data acquisition as designed by the AFM manufacturer. Typically, the acquired data channels are the interferometric amplitude s_n and phase ϕ_n demodulated at the nth order harmonic of the tip frequency, and/or auxiliary input voltages, such as the topography signal generated by the AFM.

5. Mechanical stability

A Mach-Zehnder interferometer requires a high degree of mechanical stability in the optical setup. In order to minimize temperature related mechanical drifts, our aSNOM is set up on a 60×80 cm^2 optical breadboard of very low thermal expansion coefficient material at room temperature (Invar36). The whole AFM unit attached to the aSNOM optomechanical breadboard setup is placed on a table floated on a passive vibration damping system (Newport) decoupling the system from ground vibrations. In addition, an acoustical isolation dome around the entire instrument provides shielding from sound waves and convection of the ambient air. The dome also acts as a thermal buffer around the setup letting the different parts of the setup thermalize evenly. These improvements have proven essential, particularly in homodyne interferometry.

B. Optical setup

To illustrate the situation and the typical power levels involved in aSNOM, consider an ensemble of closely spaced $\sim$10 nm sized objects observed in backscattering geometry with a conventional NA=0.3 objective which has an $\sim$1 μm focal spot in the visible. Using the geometric cross section as an estimate, illumination with 1 mW results in only a few nanowatts of detectable backscattered radiation from each object; realistically it is more likely in the picowatts to nanowatts range which is at or below the noise level of available linear detectors. Notice that the detectable power per object (or lateral resolution spot in continuous media) scales with the square of the object size and a 1 nm sized object yields

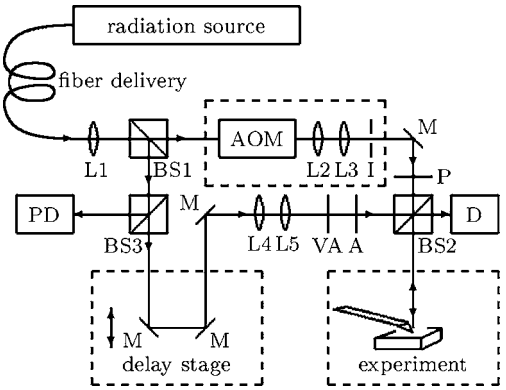

FIG. 3. Scheme of the interferometric optical detection. L: lens, BS: beam splitter, AOM: acousto-optic modulator, I: iris stop, M: mirror, PD: photodiode, VA: variable attenuator, P: polarizer, A: analyzer, and D: detector.

correspondingly another hundred times lower intensity.

With single-photon detectors available, the low scattering signal in itself poses no difficulty. However, the discrimination against the background signal is problematic. Integrated over all scatterers in the diffraction-limited collection volume of the objective, the background is essentially independent of the focus size and may easily reach several percent of the illumination power, here many microwatts, about six orders of magnitude higher than the desired signal. Such a dynamic range is again at the limit of current detector technology. To boost the level of the weak near field signal above the noise level and also to reduce the necessary dynamic range of the detector by a factor of 2, we employ interferometric detection.[11] As a bonus, the optical phase becomes accessible in addition to the amplitude.

From the three main interferometer arrangements—Mach-Zehnder, Michelson, and Sagnac—we chose the Mach-Zehnder type because it affords the greatest flexibility in manipulating the reference arm separately from the signal arm to implement, for instance, a heterodyne version. The price paid is comparatively high requirements for mechanical stability. If desired, our aSNOM can easily be rearranged as a Michelson interferometer which is less demanding in this respect. A Sagnac interferometer, perhaps the most stable interferometer type, might also be an interesting option in the context of magneto-optical imaging.[46]

1. Laser sources and fiber-optic delivery

A fiber-optical delivery scheme allows the lasers and the aSNOM unit to be installed on different tables. The minute aperture of a single mode fiber-optic core acts essentially as a spatial filter providing a near-Gaussian optical beam profile. In addition, it offers a "plug and play" interface for changing laser sources without the need for complete optical realignment. With a careful compensation of the optical path length difference in the two interferometer arms aSNOM operation can be achieved even with relatively broadband, nonsingle frequency lasers.

2. Homo/heterodyne detection

Figure 3 shows the optical detection scheme of our aSNOM, capable of both homodyne and heterodyne operations. It consists of a Mach-Zehnder interferometer for continuous wave (cw) radiation with one of the interferometer arms illuminating the near field optical experimental volume in a backscattering configuration. Here, we briefly outline the strategies for interferometric data acquisition.

Let ω_L be the frequency of the exciting optical radiation, possibly shifted by $\Delta\omega$ for heterodyne interferometry. The optical field in the signal beam arriving at the detector is modulated with $\cos(n\Omega t)$ at the nth overtone of the tip vibration frequency Ω, appearing in two sidebands,

$$\mathbf{E}_{sn} \exp i[(\omega_L + \Delta\omega \pm n\Omega)t + \phi_n]. \tag{1}$$

It has suffered an overall phase shift ϕ_n due to the optical interaction and the path length difference in the interferometer. When let to interfere with the reference field $\mathbf{E}_{r0} \exp(\omega_L t)$, the detector registers the two interference phasors

$$s_n \exp i[(\Delta\omega \pm n\Omega)t + \phi_n], \tag{2}$$

whose amplitude is given by

$$s_n = \mathbf{E}_{sn} \cdot \overline{\mathbf{E}_{r0}}. \tag{3}$$

In heterodyne operation ($\Delta\omega \neq 0$) a dual phase lock-in amplifier filtering at one of the two distinct frequencies $\Delta\omega \pm n\Omega$ can measure both the nth optical amplitude and phase simultaneously. In homodyne mode ($\Delta\omega = 0$), the two sidebands coincide in frequency space, forming the signal

$$2s_n \cos(n\Omega t) \exp i\phi_n \tag{4}$$

from which only one part of the phasor can be extracted by single lock-in detection at frequency $n\Omega$. To achieve quadrature, it is convenient to separately measure the signal again with the phase shifted by 90°, for example, by changing the interferometric path length difference by $\lambda/4$ with the help of a spatial delay stage or an electro-optic retarder in the reference arm.

3. Interferometer

All optical elements are mounted such that the beam propagates in a plane 50 mm above the Invar36 base plate. The single mode fiber emits a nearly Gaussian beam with NA ~0.11 which is collimated by L1 to a beam width of $\varnothing = 1.7$ mm matching the clear aperture of the acousto-optic modulator. The 50:50 nonpolarizing beam splitter cube BS1 separates the beam into a "reference" and a "signal" arm.

In the reference arm, BS3 splits off a few percent of the power for amplitude normalization with the dc light level detected by the photodetector (PD). A delay stage, consisting of two mirrors in retroreflector arrangement mounted on a linear translation stage, is used to coarse adjust the optical path difference of the interferometer arms to zero. In addition, a piezodriver allows us to change the path difference up to a few micrometers to adjust the phase of the reference beam, which is useful in achieving quadrature for homodyne interferometry. Though not critical for achieving reliable aSNOM operation, a two-lens telescope expands the beam width slightly to 5 mm, matching that of the signal arm. The intensity in the reference arm is set with variable attenuator (VA) to about 0.1 mW to keep the detector in the linear response region. The analyzer (A) controls the polarization

state of the reference arm and thus—due to the interferometric nature of the signal, Eq. (3)—also the polarization of the scattered radiation recorded.

The signal arm of the interferometer is used to irradiate the near field optical interaction region. Generally, the power reaching the sample is controlled at the radiation source. For homodyne operation only the polarization of the beam needs to be set with the polarizer (P). For heterodyne operation the signal beam is frequency upshifted by an acousto-optical modulator (AOM) and an iris stop I selects the proper beam from the AOM. L2 and L3 form a beam expander matching the acceptance width of the objective.

Our heterodyne interferometer in conjunction with the spectrum analyzer allows us to check for zero-path-difference (ZPD) alignment in a convenient manner. Replacing the sample with a mirror, the interference spectrum from the detector is recorded for a number of path differences around the ZPD. A clear strong beat signal at the shift frequency of $\Delta\omega \approx 80$ MHz is detected only within the coherence length of the radiation source employed. For our Ar^+ ion laser, for instance, we verified a coherence length of only a few centimeters around the ZPD corresponding with its ~20 GHz linewidth.

4. AOM

In heterodyne interferometry, the signal and/or reference radiation is shifted in frequency, conveniently achieved with an AOM in one (or both) of the beams. The effect of the sound wave inside the AOM on the incoming laser beam is similar to that of a grating. It results in multiple orders of diffracted beams, propagating in characteristic directions. However, as the grating is not stationary, the diffracted orders also suffer a Doppler shift with frequency differences given as integer multiples of the sound frequency $\Delta\omega$. For sufficient angular separation of the diffracted orders, it is useful to drive the AOM crystal at frequencies greater than ~40 MHz. In our setup, the AOM (AA Opto-Electronic AA.ST.80/A1-vis-01) is driven by a combination of a frequency tunable driver and a rf amplifier(DRF.70-90.B4. 1 μs, AMP.50-110.B6) at $\Delta\omega \approx 80$ MHz.

AOMs exhibit notorious deficiencies worthy of discussion at this point. Concerning beam profiles, the first order beam is still reasonably Gaussian, whereas that of the unshifted beam is considerably non-Gaussian, which is one reason why we opted to use a separately split-off beam for the reference arm of the interferometer. A possible improvement is the use of a high quality beam expanding spatial filter instead of the L2/L3/I combination. The second undesired property of AOMs, even more severe for heterodyne interferometry, is frequency cross-talk between the various orders. That is, in each diffraction order direction the beam contains sidebands spaced at integer multiples of $\Delta\omega$ away from the central frequency, leading to additional, spurious interference intensities at the detector. To minimize this effect, we split off a frequency-pure reference beam *before* the AOM. The detectable interference of this radiation with the signal backscattered from the sample is thus nearly perfectly heterodyne, as desired. A possible alternative is the use of an AOM in both the reference arm and the signal arm, operating at different $\Delta\omega_r$ and $\Delta\omega_s$, respectively. Detection at the heterodyne frequency difference $\Delta\omega_r - \Delta\omega_s$ should then yield equivalently pure interference.

5. Objective periscope

The expanded beam reaches the experiment volume at an ~30° oblique angle to the sample plane by a homemade periscope unit with two flat mirrors and an aspheric lens (BP1 in Fig. 1). The signal from the experiment travels through the same periscope unit (BP2) before it is joined with the reference beam at a second beam splitter (BS2). The joined beams are focused with an $f=100$ mm lens onto the detector (D). One aspect regarding the visibility of the interferometric fringes should be noted. If the spatial coherence of signal and reference wave fronts is not very high across the entire beam cross section, it may be advantageous to insert a suitable mask into the beam in front of the detector to select that part of the cross section where signal and reference wave fronts match.

C. Electronic setup

As seen in Fig. 1 the signal detector converts the incident optical radiation into an electronic signal, which is processed with the help of a homebuilt electronics. Before we describe the generation of the reference signal for the use with a fast lock-in detector, we underline the importance of using strictly linear elements in the signal path. A slight nonlinear admixture to the detected optical signal can hide the near field optical information generated at the higher harmonics of the tip oscillation frequency. This concerns any amplifier used behind the detector but most notably the detector itself, which must be operated in a linear response regime. We note that this requirement sets the maximum allowable level for the reference beam intensity, specific to the detector used.

1. Reference signal generation

The block diagram of our reference signal generation electronics is depicted in the lower panel of Fig. 1. To obtain a truthfully phase-following electronic reference signal at $\Delta\omega \pm n\Omega$, use of the actual physical actions of the AOM and AFM as signal sources is imperative. In the case of the AOM it is safe to assume that the phase relation between drive and generated sound wave is constant, and consequently we may use the drive signal. For amplitude modulated AFMs, however, the phase between drive and the actual cantilever deflection varies significantly. We use the measured deflection directly and thus avoid this mechanical phase shift. A viable alternative would be to use the AFM drive signal and separately measure and subtract this mechanical phase.

The deflection signal from the quadrant detector is first amplified. We implemented two options of the $\times n$ block used to obtain a signal with a constant phase relation at an integer multiple of the fundamental frequency. In the first design, a highly nonlinear amplifier operating near saturation converts the fundamental sine to a near-rectangular wave, in which all harmonics are present. Sixth order *LC* bandpass filters of suitable band width and center frequency can efficiently isolate any desired overtone $n\Omega$ of the original signal.

While this passive approach works reliably, the necessary manual selection of proper bandpass filters is inconvenient. Therefore, we recently developed a phase-locked-loop (PLL) based synthesizer to generate the higher harmonics. At the time of writing, however, we have not yet accumulated enough experience to fully judge the relative performance of these options.

The resulting $n\Omega$ signal (several hundreds of kilohertz) is mixed with the 80 MHz AOM drive signal, taken from the 0 dB reference port of the rf driver. After the mixer, the signal contains besides the center frequency mainly two sidebands at $\Delta\omega \pm n\Omega$. We explored only passive options for the sideband filter. *LC* filters, with their relatively low Q factors (typically a few tens), require forbiddingly high cascade orders that eventually diminish the throughput of the filter to intolerable levels. Quartz filters feature such large Q factors (up to millions) that it would be extremely difficult to meet and/or keep the necessary center frequency, because rf circuitry is notoriously difficult to stabilize thermally. Therefore, the use of surface acoustic wave (SAW) filters is an interesting option. With medium Q factors in the hundreds to thousands they filter a single sideband efficiently, yet tolerate frequency drifts of several tens of kilohertz. We employ two cascaded SAW filters (TSF80B by VI Telefilter) with a 20 dB amplifier in between. For tuning the AOM frequency to meet the fixed 3 dB passband at $\Delta\omega + n\Omega \approx 80 \pm 0.1$ MHz, a spectrum analyzer (HP 8568A) is a valuable tool. The same instrument can also help us to analyze the avalanche photodiode (APD) signal as well as monitor cantilever harmonics directly.

2. Lock-in detection

The detector is a fast, sensitive, low-noise avalanche photodiode (APD, model C5331-02 by Hamamatsu) operated in the linear response regime. The heterodyne interferometer generates an optical field at the APD that beats at $\Delta\omega \approx 80$ MHz. This beat signal is further modulated due to the oscillation of the AFM tip at the cantilever frequency Ω and its harmonics. To observe the relevant spectral components at $\Delta\omega \pm n\Omega$, the detected signal is demodulated at the corresponding frequencies with a fast, dual phase lock-in amplifier (Stanford Research Systems SR844). The synthesis of the reference signal has been discussed above. Furthermore, the lock-in amplifier is used in ratio mode, i.e., the detected amplitude is normalized to an external voltage proportional to the instantaneous power in the reference beam. Low- to midfrequency laser power oscillation and settling effects are thus canceled in the amplitude outputs of the lock-in amplifier.

An important issue that needs special attention is the possible cross-talk of other electronic signal sources at the reference frequency. A poorly rf shielded line from the mixer to the reference input might couple into the signal input of the lock-in amplifier. Most notably, though, the coherent pickup inside the lock-in amplifier adds a small, experiment-independent background phasor to the real signal. Thus, especially for very low level signal detection, this effect mimics an optical interference not actually present in the signal. While not entirely avoidable, it is at least relatively easy to check for this effect by simply selecting different amplifier settings that alter the coherent pickup, while keeping the experimental conditions fixed.

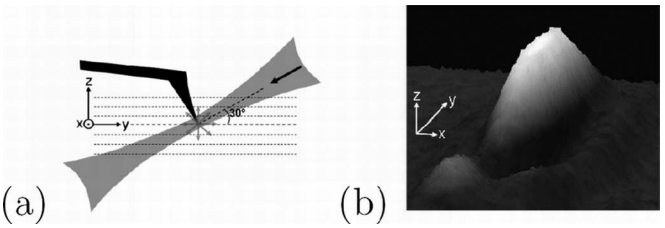

FIG. 4. (a) Easy alignment of an AdvancedTEC tip with the optical focus. Accumulating x/y s_1 scans at several z values allows us to map the entire focal volume. (b) Focal spot s_1 scan of a 23×23 μm^2 area at the optimum z position.

III. ALIGNMENT AND NEAR FIELD OPTICAL CONTRAST

This section discusses alignment and operation of the instrument. For alignment, say, after a tip exchange, the tip apex is first centered in the optical focus. A by-product of our alignment procedure is the three-dimensional characterization of the focus volume. A following sample-tip approach yields immediate feedback on the quality of local near field optical information and tests the suppression of background far field signals at higher harmonics of the cantilever vibration. Other essential tests of true near field optical contrast are off-apex illumination and successive imaging at different demodulation orders.

A. Optomechanical alignment

All field-enhanced near field optical microscopic experiments require three-dimensional overlap of four individual volumes: the sample area of interest, the far field foci for illumination and collection, and the near field enhancement region at the tip apex. If the same objective is used for illumination and collection in backscattering geometry, the nontrivial task remains to align the centers of three independent volumes with an accuracy of better than ~ 10 nm.

In our setup, we first position the AFM tip apex at the center of the far field focus while the sample is retracted. For tips of general shape, side-view observation of the tip during coarse alignment helps us to avoid unintended illumination at places other than the apex. This problem is much alleviated by cantilevers of suitable shape (Nanosensors, AdvancedTEC® silicon SPM sensors), as shown in Fig. 4. The protruding tip design results in faster and more reliable alignment even with imaging from the top.

For fine alignment, the tip-scanning capability of the instrument comes in handy to image planar sections of the focal volume and to position the tip subsequently at the center. Figures 4(a) and 4(b) illustrate how scanning the tip through the illumination beam focus while recording s_1 allows us to identify the focal center as the peak position. The peak appears extended in the beam propagation direction as a result of the 30° angle between the beam field depth and the scan plane. Indeed, several such scans at various heights, accompanied by modeling of the light scattering due to the

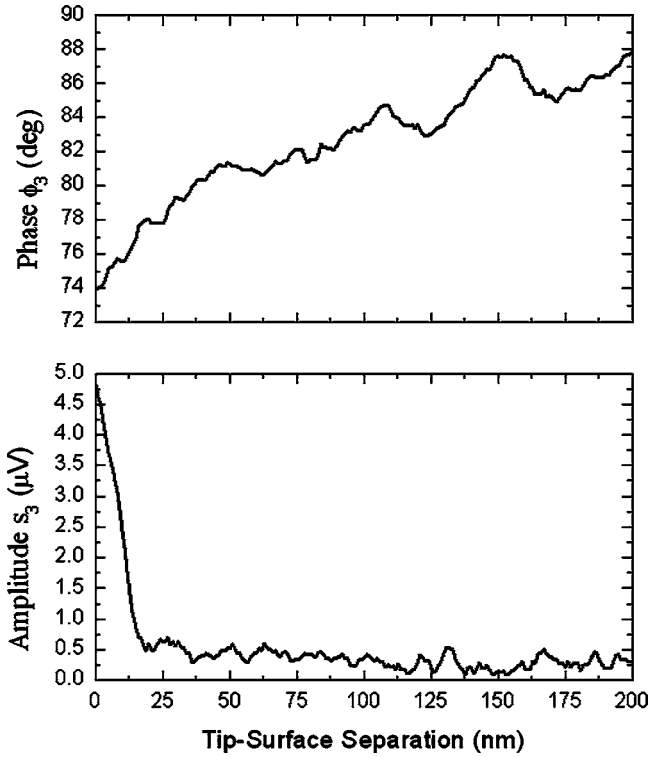

FIG. 5. Amplitude s_3 and phase ϕ_3 scattered by a Au coated Si tip while a thick Au film surface is approached from below. An optical near field contribution is detected at distances $\lesssim$20 nm from the Au surface.

tip shape, allow us to reconstruct both amplitude and phase of the optical fields in the focus volume with nanometer precision. As the procedure for initial alignment of tip apex and optical focus we choose the three-dimensional maximum intensity location.

B. Approach curves

Having centered the AFM tip at the illumination focus, the sample surface must be approached to engage the AFM topography feedback. Acquiring the optical amplitude and phase during such an approach provides an important first evaluation of their near field optical origin, because not all AFM tips and samples provide sufficient field enhancement. If they do, the characteristic signature is a marked increase of the higher harmonics signal within a few nanometer of distance from the sample like the example shown in Fig. 5. A blunted or oxidized tip may yield no detectable near field signal at all, while the far field scattered amplitude (s_1) usually still is substantial. Just as well it is possible that the particular sample surface does not exhibit much near field optical contrast.

Approach curves are an essential and easily performed test for true near field optical contrast. They have proven a valuable tool for detecting faulty experimental conditions and correcting them. One important caveat, however, is the possibility of cross-talk between mechanical anharmonicities and the near field optical nonlinearities.[45] The mechanical anharmonicities are largely controlled by two parameters: the free space cantilever amplitude and the sample engaged cantilever amplitude setpoint. On the one hand, if the AFM cantilever is overdriven, one registers strong higher harmonics

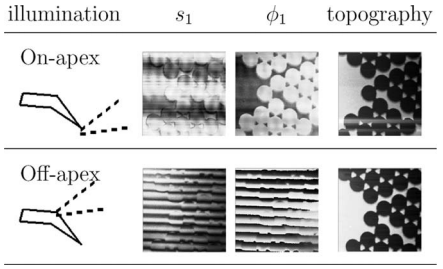

FIG. 6. aSNOM scans of 3×3 μm^2 sample area at 1 Ω demodulation. The upper line is the set of 1 Ω images with normal tip apex centered illumination. The lower line is obtained under *identical* experimental conditions, except for the excitation beam being focused at the top of the tip shaft, not its apex.

signals even in the absence of any sample. Not exciting the cantilever enough, on the other hand, results in too little demodulation signal. In general one has to find a cantilever-specific optimal free space drive amplitude.

Any form of anharmonic motion due to mechanical tip-sample interaction should be avoided, such as snap-in or intermittent contact mode. It was found empirically that snap-in is much reduced if the product of amplitude and cantilever force constant is greater than $\sim$200 nN.[41] Another empirical rule helps us to keep the cantilever motion in the attractive regime by demanding the ratio of AFM amplitude setpoint to free space oscillation amplitude to be at least 90%.

Together, these conditions on the AFM settings should ensure nearly perfect harmonic cantilever oscillations. However, an independent verification is desirable. The worst cases of strong mechanical anharmonicities are quickly identified by monitoring the cantilever beam deflection signal. But it likely does not reach the level of sensitivity afforded by the interferometric aSNOM detection itself. A very useful intrinsic performance check is therefore the intentional off-apex tip illumination test.

C. Off-apex tip illumination test

Intentional misalignment of the excitation to focus not at tip apex but elsewhere on the tip shaft must lead to drastic changes in the detected signal, if it actually originates from the near field optical interaction at the tip apex. Otherwise, the aSNOM merely detects changes in the anharmonic motion of the cantilever. Figure 6 presents such a comparison of images taken with on-apex and off-apex illuminations.

The sample consists of nanosphere lithographically defined gold patches on glass. A 10 nm Au coated AdvancedTEC Si tip with a resonance frequency of Ω =320 kHz is used as the apertureless probe. The dynamic mode tip vibration amplitude is 15 nm at the AFM feedback condition. The free space tip vibration amplitude is 16 nm, meaning that the feedback amplitude to free space amplitude has a $>$90% ratio which is crucial for operation with minimal mechanical anharmonicity. The scan speed is 1.5 s/line and the scan size is 3×3 μm^2.

While the *on-apex* illuminated images exhibit the normal aSNOM operation mode, with *off-apex* illumination—all other parameters being equal—the images show drastic

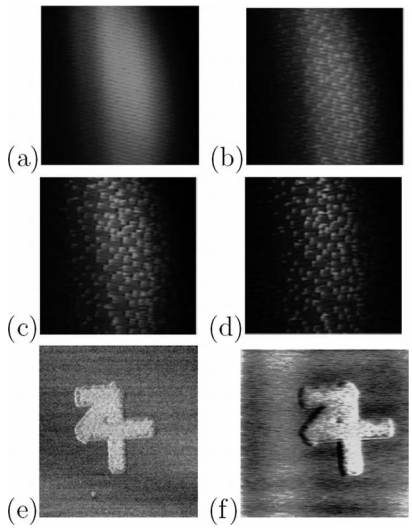

FIG. 7. (a)–(d) Optical amplitude images of a 10 μm^2 area of Ag colloids on a Si substrate acquired in high-speed overview tip-scanning mode. (a) s_1, with the sample retracted, shows the field strength envelope in the optical focus. (b)–(d), with the sample in contact, show the s_1, s_2, and s_3 optical amplitudes recorded at demodulation frequencies of 1, 2, and 3 Ω. (e) and (f) 5×5 μm^2 optical amplitude images (s_3) of a 60 nm thick Au pattern on a 100 nm thick oxide substrate on Si, obtained by scanning (e) the sample and (f) the tip.

changes because the near field optical tip-sample interaction is essentially turned off.

D. Sample scanning versus tip scanning

Sample scanning with its undisturbed tip apex/optical focus alignment is favorable in aSNOM. Nevertheless, for a quick overview and an initial test of near field optical contrast, the tip-scanning mode of the commercial AFM is an interesting option. In this operation mode the optical signals are fed from the lock-in amplifier into the auxiliary input channels of the AFM electronics.

In Figs. 7(a)–7(d), the AFM tip scans an area of 10×10 μm^2. The sample is a Si surface hosting Ag colloids of 20 nm diameter with low coverage. Since the tip is being scanned, the characteristic amplitude envelope of the optical focus as measured in (a) with the sample removed is observed also with the sample in place [(b)–(d)]. In (b), the s_1 amplitude is recorded. The colloids are not clearly distinguished from the Si substrate in the predominantly far field optical signal background. (c) and (d) show the s_2 and s_3 images, respectively, and demonstrate the improving background suppression with successive higher harmonic demodulation.

Two images of the same area on another sample surface, acquired by scanning the sample and the tip, respectively, are displayed in Figs. 7(e) and 7(f). Clearly, the intensity distribution in the sample-scanned image is more homogenous, whereas the tip-scanned image exhibits an intensity variation following the focal amplitude envelope with a waist of about 3 μm.

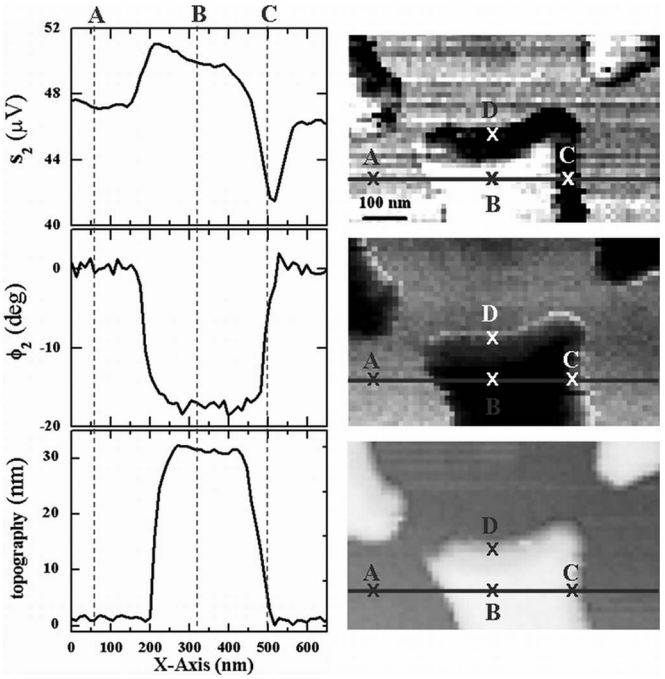

FIG. 8. A 650×390 nm^2 scan of nanosphere lithographically patterned Au on glass sample. On the left, optical amplitude, optical phase, and topography information are displayed for the horizontal line cut indicated on the right images.

IV. OPTICAL IMAGING AT SUB-10 nm RESOLUTION

A. Material contrast: Nonoptical artifacts

The possibility to learn about the material composition at the nanometer scale through the use of local optical probes is one of the most exciting prospects of aSNOM. At the same time, a thorough understanding of the possible artifact mechanisms of this technique has to be developed to avoid misinterpretations. In Fig. 8, a detailed view of the same sample used for Fig. 6 is presented showing how the *optical* interaction of tip and sample surface gives rise to a material contrast in the scattered radiation [(A) and (B)] while at the same time artifacts of *topographical* origin appear, especially at steep edges [(C) and (D)].

The simultaneously acquired topography and optical maps demonstrate the clear distinction of gold and glass in both optical amplitude and phase. Compared to the substrate (A), the flat parts of the Au islands (B) exhibit an increase in near field optical amplitude, confidently observed even against a relatively large background, and a negative shift in the phase by some 17°. At the edges of the islands, that is, near steep slopes, a darker rim appears, which may be related to an instantaneous change of contact point from the tip's apex to its side.[48] In addition, at such steep slopes, two other kinds of artifacts are discernible.

Along edges parallel to the fast scan direction, the dark rim is much stronger for the upper edges (D). As the illumination/detection is from/to the bottom of the images into the image plane as displayed in Fig. 8, a kind of shadowing effect appears to play a role. Of those edges of the islands that lie across the fast scan direction (C) one side appears darker than the other. In the reverse scan direction (not shown here), the opposite side appears darker. This type

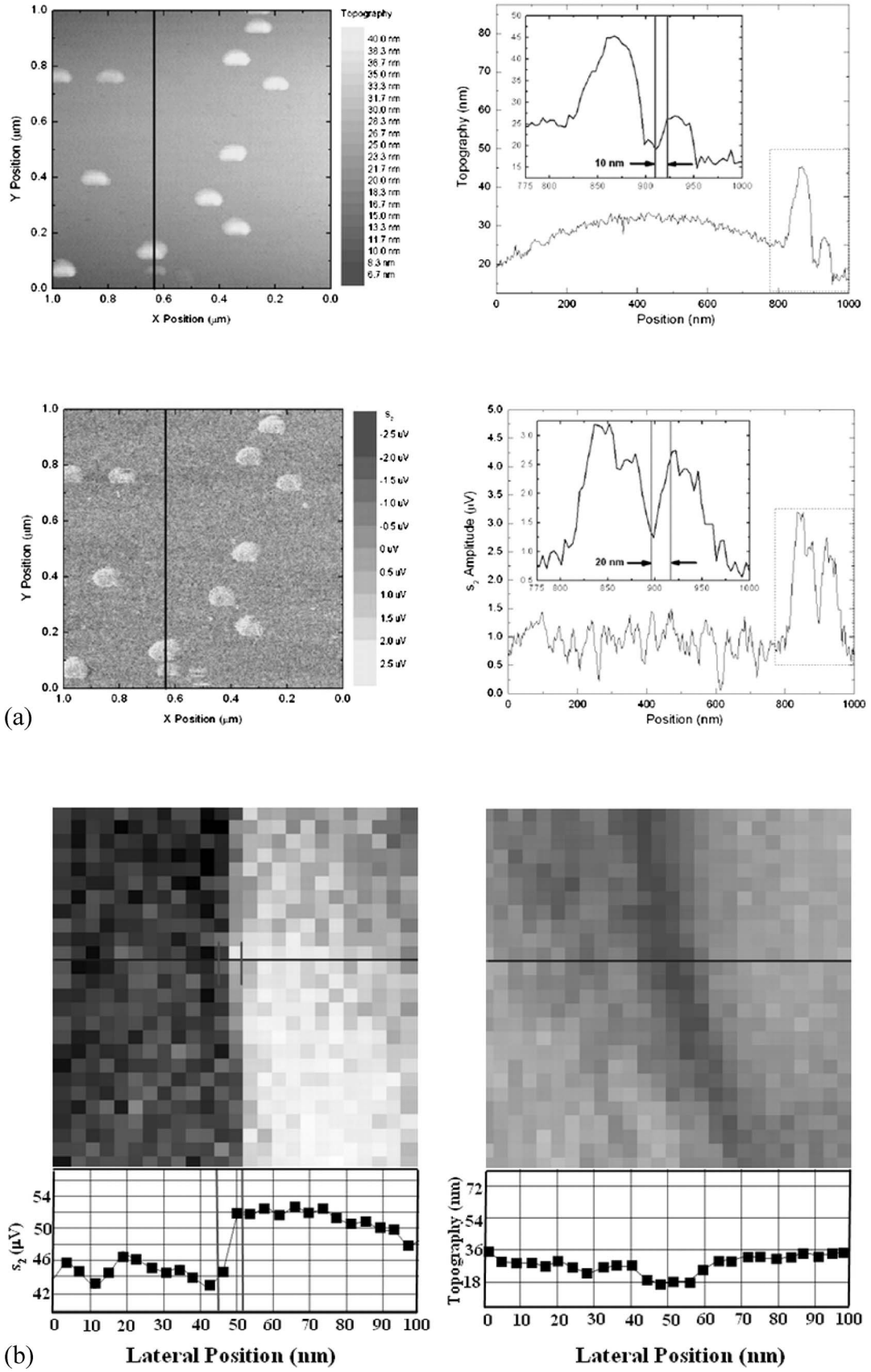

FIG. 9. Spatial resolution studies with aSNOM. (a) Topography (top) and s_2 (bottom) of a 1×1 μm^2 scan of nominally 20 nm diam Ag colloids deposited on Si. Plots along the lines marked in the images are displayed on the right. (b) Optical s_2 (left) and topography (right) of a 100×100 nm^2 scan of SiO_x overgrown Si substrate edge. Line plots given at the bottom of each image belong to the marked positions in the images.

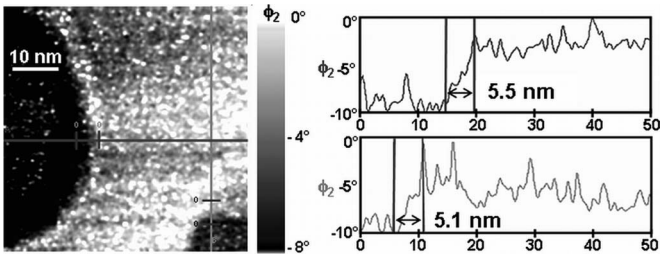

FIG. 10. A 50×50 nm^2 optical phase image of the same sample as in Fig. 9(a). On the right are the line scan data belonging to the marked positions on the image.

of artifact is generated by suboptimal AFM scan parameter. Notably, reducing the scan speed helps us to eliminate it. Apparently, as the AFM tip encounters a steep slope, its topography-following feedback electronics are not quite fast enough to ensure purely harmonic tip oscillations and for a brief time, higher harmonics in the mechanical motion of the tip contribute to the optically recorded signal.

B. Complex near field optical contrast and resolution

In Fig. 9(a) the optical amplitude lateral resolution is studied with Ag colloids deposited on Si. The imaging tip in this case is an uncoated Si AdvancedTEC tip that evidently has been slightly degraded, predominantly along the x direction. Therefore, we analyze the resolution along y with cuts along the indicated vertical lines displayed on the right. Of the two closely spaced colloids the one at ~860 nm is about twice as high as that at ~930 nm. In the s_2 image, however, both of them appear similarly strong. Evidently, the optical signal identifies the material of both colloids as silver, despite their different sizes. The spatial resolution achieved is slightly worse in the optical signal than in the topographical around 20 nm or $\approx\lambda/31$ for the HeNe 633 nm laser line used.

In a further optical amplitude resolution study we observe the contrast between SiO_x and Si at the cut and polished edge of a silicon wafer overgrown with a 300 nm oxide layer. Figure 9(b) shows the optical amplitude s_2 (left) and topography (right) of a 500×500 nm^2 area scanned with a 5 nm Cr +10 nm Au coated Si AdvancedTEC tip. Line plots of optical amplitude and topography are given at the bottom corresponding to the marked lines in the images. A clear material interface is recognized as a sharp step in the optical amplitude change with no equivalent feature in the topography. One finds an optical s_2 amplitude lateral resolution of about 2 pixels or ≈8 nm corresponding to $\approx\lambda/80$.

A 50×50 nm^2 phase image is presented in the left part of Fig. 10. The dark areas correspond to two Ag colloids. From the line scans displayed on the right, it is evident that the near field optical phase of Ag colloids develops more abruptly than the associated optical amplitudes within approximately 5 nm or $\approx\lambda/100$. At this length scale, however, the morphological convolution of sample and tip geometry with the optical image has to be taken into account.

For objects of diameter ≈20 nm, such as these colloids, this convolution first of all results in a larger effective diameter imaged. Here, a cross sectional diameter of the scanning AFM tip of about 25–30 nm (at the 10 nm height where the sides of tip and colloids meet) results in an image diameter of slightly less than 50 nm. Secondly, and more importantly for optical near field images, as the AFM tip scans over a colloid the closest proximity point jumps from the tip apex to its side. The ensuing abrupt rearrangement of the optical near fields is the likely cause of the sharp step in optical phase observed in Fig. 10. Therefore an optical phase resolution of 5 nm cannot be claimed with confidence, but it is fair to estimate it as considerably less than the imaged object's size, i.e., <20 nm.

V. SUMMARY

The modular design of our aSNOM implementation allows for a variety of possible applications with minimal changes to the setup. The instrument works with any kind of opaque or transparent samples. A metallic or dielectric AFM tip with a sharp apex acts as a local probe. Usually it is convenient to boost the scattered fields originating from the local interaction volume near the tip apex to detectable levels by means of heterodyne or homodyne interferometrical amplification. Direct detection is enabled simply by blocking the reference arm. Parasitic scattering generated at other illuminated surfaces is effectively suppressed by demodulating the detected scattering amplitude at higher harmonics of the vibration frequency of AFM cantilever.

The aSNOM setup allows for plug and play switching between different laser sources. The optics is modifiable to a variety of optical characterization schemes by simple inclusion of additional optical elements. For magneto-optical studies, for instance, a polarization rotator could be included in the reference arm. Lateral scanning of images can be done with either the tip or the sample stage. The performance tests of the aSNOM show superior properties like an easy, rational tip-focal spot alignment. This instrument images amplitude and phase of optical near fields with a spatial resolution better than 10 nm, essentially limited only by the apex size.

ACKNOWLEDGMENTS

The authors thank R. Esteban for many stimulating and useful discussions. The authors are grateful to Professor Leiderer from the University of Konstanz for providing the samples used for the measurements presented in Figs. 6 and 8.

[1] S. W. Hell and J. Wichmann, Opt. Lett. **19**, 780 (1994).
[2] C. Hettich, C. Schmitt, J. Zitzmann, S. Kühn, I. Gerhardt, and V. Sandoghdar, Science **298**, 385 (2002).
[3] E. H. Synge, Philos. Mag. **6**, 356 (1928).
[4] E. A. Ash and G. Nicholls, Nature (London) **237**, 510 (1972).
[5] D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984).
[6] A. Lewis, M. Isaacson, A. Harootunian, and A. Muray, Ultramicroscopy **13**, 227 (1984).
[7] E. Betzig, J. K. Trautmann, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251**, 1468 (1991).
[8] M. A. Paesler and P. J. Moyer, *Near-Field Optics* (Wiley, New York, 1996).
[9] R. Bachelot, P. Gleyzes, and A. C. Boccara, Microsc. Microanal. Microstruct. **5**, 389 (1994).
[10] F. Zenhausern, M. P. Oboyle, and H. K. Wickramasinghe, Appl. Phys. Lett. **65**, 1623 (1994).
[11] F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, Science **269**, 1083

(1995).
[12] B. Knoll, F. Keilmann, A. Kramer, and R. Guckenberger, Appl. Phys. Lett. **70**, 2667 (1997).
[13] B. Knoll and F. Keilmann, Appl. Phys. A: Mater. Sci. Process. **66**, 477 (1998).
[14] F. Keilmann, B. Knoll, and A. Kramer, Phys. Status Solidi B **215**, 849 (1999).
[15] B. Knoll and F. Keilmann, Appl. Phys. Lett. **77**, 3980 (2000).
[16] R. Hillenbrand, Ph.D. thesis, Technische Universität München, 2001.
[17] I. S. Averbukh, B. M. Chernobrod, O. A. Sedletsky, and Y. Prior, Opt. Commun. **174**, 33 (2000).
[18] E. Oesterschulze, G. Georgiev, M. Müller-Wiegand, A. Vollkopf, and O. Rudow, J. Microsc. **202**, 39 (2001).
[19] H. F. Hamann, M. Kuno, A. Gallagher, and D. J. Nesbitt, J. Chem. Phys. **114**, 8596 (2001).
[20] K. B. Crozier, A. Sundaramurthy, G. S. Kino, and C. F. Quate, J. Appl. Phys. **94**, 4632 (2003).
[21] L. A. Bottomley, Anal. Chem. **70**, 425R (1998).
[22] H. K. Wickramasinghe, Acta Mater. **48**, 347 (2000).
[23] Y. Inouye and S. Kawata, Opt. Lett. **19**, 159 (1994).
[24] R. Bachelot, P. Gleyzes, and A. C. Boccara, Opt. Lett. **20**, 1924 (1995).
[25] K. Dickmann, J. Jersch, and F. Demming, Surf. Interface Anal. **25**, 500 (1997).
[26] S. Nie and S. R. Emory, Science **275**, 1102 (1997).
[27] L. Novotny, E. J. Sánchez, and X. S. Xie, Ultramicroscopy **71**, 21 (1998).
[28] B. Pettinger, G. Picardi, R. Schuster, and G. Ertl, Electrochemistry (Tokyo, Jpn.) **68**, 942 (2000).
[29] M. S. Anderson, Appl. Phys. Lett. **76**, 3130 (2000).
[30] R. Stockle, Y. Suh, V. Deckert, and R. Zenobi, Chem. Phys. Lett. **318**, 131 (2000).
[31] N. Hayazawa, Y. Inouye, Z. Sekkat, and S. Kawata, Opt. Commun. **183**, 1 (2000).
[32] T. Kalkbrenner, M. Ramstein, J. Mlynek, and V. Sandoghdar, J. Microsc. **202**, 72 (2001).
[33] C. C. Neacsu, G. A. Reider, and M. B. Raschke, Phys. Rev. B **71**, 201402 (2005).
[34] M. L. M. Balistreri, J. P. Korterik, L. Kuipers, and N. F. V. Hulst, Appl. Phys. Lett. **79**, 910 (2001).
[35] S. Aubert, A. Bruyant, S. Blaize, R. Bachelot, G. Lerondel, S. Hudlet, and P. Royer, J. Opt. Soc. Am. B **20**, 2117 (2003).
[36] J. Renger, S. Grafström, and L. M. Eng, J. Opt. Soc. Am. A **21**, 1362 (2004).
[37] A. Hartschuh, E. J. Sanchez, X. S. Xie, and L. Novotny, Phys. Rev. Lett. **90**, 095503 (2003).
[38] T. Ichimura, N. Hayazawa, M. Hashimoto, Y. Inouye, and S. Kawata, Appl. Phys. Lett. **84**, 1768 (2004).
[39] G. Wurtz, R. Bachelot, and P. Royer, Rev. Sci. Instrum. **69**, 1735 (1998).
[40] R. Hillenbrand and F. Keilmann, Phys. Rev. Lett. **85**, 3029 (2000).
[41] T. Taubner, F. Keilmann, and R. Hillenbrand, Opt. Express **13**, 8893 (2005).
[42] R. Hillenbrand, T. Taubner, and F. Keilmann, Nature (London) **418**, 159 (2002).
[43] M. B. Raschke, L. Molina, T. Elsaesser, D. H. Kim, W. Knoll, and K. Hinrichs, ChemPhysChem **6**, 2197 (2005).
[44] I. Stefanon, S. Blaize, A. Bruyant, S. Aubert, G. Lerondel, R. Bachelot, and P. Royer, Opt. Express **13**, 5553 (2005).
[45] A. Bek, R. Vogelgesang, and K. Kern, Appl. Phys. Lett. **87**, 163115 (2005).
[46] A. Bauer, *Far-Field and Near-Field Magneto-Optical Microscopy of Ultrathin Films* (Freie Universität Berlin, Germany, 2000).
[47] F. J. Gießibl, *Progress in Atomic Force Microscopy* (Universität Augsburg, Germany, 2000).
[48] T. Taubner, R. Hillenbrand, and F. Keilmann, J. Microsc. **210**, 311 (2003).