Far-field background suppression in tip-modulated apertureless near-field optical microscopy

Pietro Giuseppe Gucciardi[a)]
CNR-Istituto per i Processi Chimico-Fisici, Sezione di Messina, Via La Farina 237, I-98123 Messina, Italy

Guillaume Bachelier
Laboratoire de Spectrométrie Ionique et Moléculaire (LASIM-CNRS UMR 5579), Université Claude Bernard-Lyon I, Bâtiment A. Kastler, 43 Boulevard du 11 Novembre 1918, F-69622 Villeurbanne, France

Maria Allegrini[b)]
Dipartimento di Fisica "E. Fermi," Università di Pisa, and polyLAB-CNR, Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

(Received 13 January 2006; accepted 4 April 2006; published online 29 June 2006)

In apertureless near-field optical microscopy the vertical dithering of the tip, associated with demodulation at higher harmonics ($n>1$), allows us to suppress the far-field background, providing artifact free elastic scattering images. This paper analyzes, both theoretically and experimentally, the physical origin of the background signal at the different harmonics and the mechanisms underlying its rejection for the general case of propagative-field illumination. We show that Fourier components of the background must be expected at every harmonic, evidencing why demodulation at higher harmonics is not an inherently background-free technique, and assessing the experimental conditions in which it becomes like that. In particular, we put forward the fundamental roles of both the harmonic order and the tip oscillation amplitude in the background suppression mechanisms. Furthermore, we outline how the lock-in detection of the signals amplitude can enhance the nonlinear dependence of the background on the tip-sample distance. Such effect provides a more subtle source of topography artifacts since the optical maps become qualitatively uncorrelated from the topographic counterpart, requiring an upgrade of the criteria to assess the absence of artifacts from the optical maps. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2208527]

I. INTRODUCTION

Scanning near-field optical microscopy[1] (SNOM) overcomes the diffraction limit ($\sim\lambda/2$) by scanning a nanometric probe close to the sample surface. The superior resolution of SNOM arises from the exploitation of the evanescent components of the electromagnetic field, the so-called near field, in the probe-sample optical interaction. Due to the evanescent nature of the near fields, the tip must follow the surface at close distance (1–10 nm) during the scanning process [the so-called constant gap operation mode (CGM)]. Aperture-SNOM employs the aperture present at the apex of a tapered, metal coated-optical fiber as a source of evanescent waves.[2] The development of *apertureless*-SNOM (*a*-SNOM) techniques (see Ref. 3, for example) has demonstrated that the use of a sharp tip ending with a nanometer-scale radius of curvature can provide a local enhancement of the field by several orders of magnitude.[4] Optical imaging with spatial resolution in the 1–10 nm range is thus achievable.[5,6] In an ideal elastic scattering SNOM experiment, the radiation reaching the detector should originate only from the nanometric region of the probe-sample optical interaction. In practice, conversely, the total signal is dominated by a large background (BKG) contribution due to stray light scattered by other sources that interfere at the detector. The probe shaft, the optical elements, or portions of the sample "far" from the probe are typical examples of undesirable sources of stray light. The major inconvenience arises from the fact that, since the phase of the stray fields depends on the mutual distance among these sources, the BKG is not a constant plateau, but is sensitive to any position variation of each source. In CGM, in particular, the continuous displacement of the sample, driven by the feedback loop, induces a strong cross talk between the topographical and the optical BKG, whose variations are orders of magnitude more intense than any near-field scattering. Fictitious optical maps are thus obtained, which do not contain information about the optical properties of the sample, but rather representing an optical readout of the topography.[7,8] This is what is called a topography, or *z*-motion artifact.

The most important challenge in *a*-SNOM is, therefore, the extraction of the tiny near-field scattering signal from the huge far-field background. Moreover, due to the fact that the scattering amplitude scales down with the sixth power of the sample's dimensions,[9] such a task becomes more challenging as the structures we want to investigate become smaller. In order to extract the near-field signal from the BKG two steps are required. Firstly, the near-field scattering must be amplified. This is done through heterodyne,[10,11] or homodyne[12–14] interferometric techniques or, when available, by exploiting material-dependent phonon-polariton resonances.[15,16] Sec-

[a)]Electronic mail: gucciardi@me.cnr.it
[b)]Present address: National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8424.

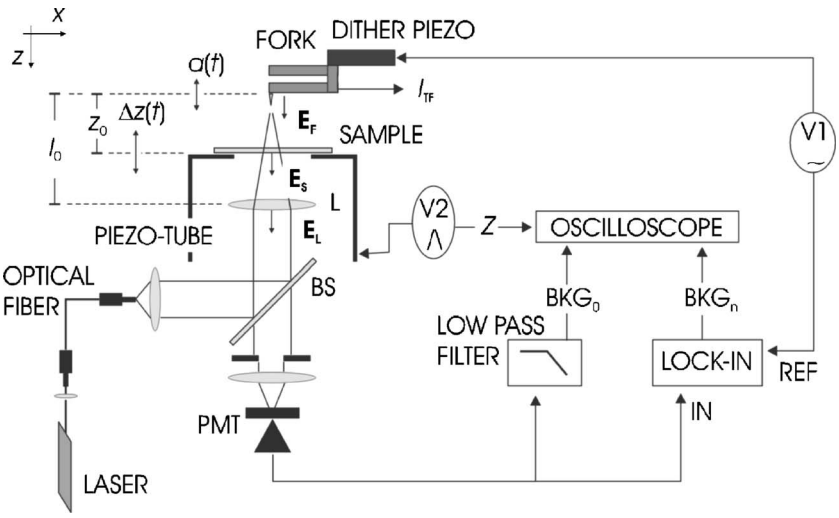

FIG. 1. Sketch of the experimental setup.

ondly, the BKG must be suppressed exploiting the nonlinear part of the tip-sample optical interaction.[9,17] Experimentally, the latter task is accomplished by vertically vibrating the tip at frequency ω and demodulating the optical signal at higher harmonics ($n\omega$, $n>1$).[10,12] This *modus operandi* induces, in fact, a high-pass filtering of the spatial frequencies of the electromagnetic fields.[18–21] The result is that the low-spatial-frequency propagating fields, such as the BKG, are largely suppressed, while the evanescent near-field components are enhanced. Demodulation at higher harmonics is thus considered to be an artifact-free imaging technique.

In this paper we precise under which conditions this statement can be considered true. The paper is organized as follows. In Sec. II we characterize experimentally the BKG signal, in particular its dc, first, and second harmonic components. The results are thus interpreted in a theoretical framework (Sec. III) which allows us to point out the crucial role of the tip oscillation amplitude in the suppression of the BKG at the higher harmonics. Finally, in Sec. IV, we outline how the nonlinear dependence of the BKG can yield fictitious optical images quite different from the topographical counterpart, allowing us to discuss and update the criteria for artifacts identification.

II. EXPERIMENT

Approach curves, in which the optical signal is monitored as a function of the tip-sample distance, represent a powerful instrument to characterize the different far-field and near-field regimes in the tip-sample optical interaction. In particular, they have already allowed us to evidence the occurrence of harmonic components of the far-field BKG at the first[10,22,23] and the second harmonics,[11,24] superimposed to the genuine near-field scattering whose signature shows up at close tip-sample distances (<100 nm). Similar to the dc component, the harmonic components of the BKG (BKG_n, $n=1,2,\ldots$) depend both on the tip's and on the sample's position, representing a source of topography artifacts. Approach curves permit to quantitatively assess the *degree of artifact content* in the optical signal through the ratio $\varepsilon_n = BKG_n/NF_n$ between the oscillation amplitude of the nth-harmonic BKG_n (measured in the far-field region) and the near-field signal enhancement NF_n demodulated at the same harmonic (measured when the tip goes in contact with the sample). In the visible spectrum ($\lambda=633$ nm), first harmonic demodulation leads to values $\varepsilon_1 \sim 1$ for tip oscillation amplitudes ranging from a few nanometers to a few tens of nanometers.[10,11] The consequent optical maps will be very likely affected by severe topography artifacts. Demodulation at the second harmonic demonstrates a stronger BKG rejection power ($\varepsilon_2 < 10^{-1}$).[11,12] Demodulation at the third harmonic guarantees artifact-free imaging[11,12,24] since BKG_3 is always rejected below the noise level ($\varepsilon_3=0$). The situation is better in the IR ($\lambda=10.6$ μm) where the first harmonic demodulation already provides a good BKG suppression power ($\varepsilon_1 \sim 0.2$), even for tip oscillation amplitudes in the 100 nm range.[23,25]

In order to quantitatively assess such results, we have experimentally studied the interferential nature of the far-field BKG with an apparatus working in transmission mode, as shown in Fig. 1. The tip is held in the far field, at a distance Z_0 of several micrometers from a glass coverslip, used as sample, which is scanned in the vertical direction with an excursion Δz of a few microns. The scan allows us to monitor the BKG dependence (both the dc component and all the harmonics) on the tip-sample distance. Our tips are made of gold or silver by means of electrochemical etching techniques[26,27] and are glued orthogonally to one prong of a quartz tuning fork (TF).[28] The assembly is oscillated at resonance ($f_{res}=\omega/2\pi \sim 30$ kHz) by a dither piezo, driven by a function generator (V1) which allows us to change the tip oscillation amplitude a_0. A measure of the effective tip oscillation amplitude is retrieved either interferometrically[29] or by monitoring the piezocurrent I_{TF} delivered by the fork.[30,31] The light is provided by a HeNe laser ($\lambda=632.8$ nm). The beam is coupled to a single mode optical fiber whose output is collimated, reflected by a 50% beam splitter (BS), and finally focused on the tip shaft by means of a lens (L) ($f=8$ mm). The sample is mounted on a piezotube, scanned up and down by means of a linear voltage ramp (V2), never touching the tip. The field backscattered by the tip shaft (E_F) is collected by the same lens and focused on the detector [photomultiplier tube (PMT)]. Due to the high coherence de-

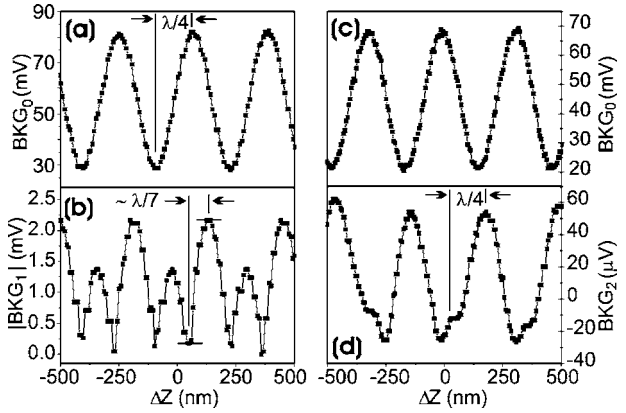

FIG. 2. Plot of the background signals BKG_0 [(a) and (c)], $|BKG_1|$ (b), and BKG_2 (d) vs the sample position.

gree of the light source, E_F will interfere with the back reflection of the sample (E_S), and with part of the light reflected back by the lens (E_L). The total intensity measured by the PMT will thus be

$$\begin{aligned} \mathrm{BKG} &= |E_F + E_S + E_L|^2 \\ &= |E_F|^2 + |E_S|^2 + |E_L|^2 + 2\,\mathrm{Re}(E_F^{*} E_S) \\ &\quad + 2\,\mathrm{Re}(E_F^{*} E_L) + 2\,\mathrm{Re}(E_L^{*} E_S). \end{aligned} \tag{1}$$

As we will see in more detail in the next section, this signal results in a sum of harmonic terms $\mathrm{BKG} = \Sigma_n \mathrm{BKG}_n(a_0, \Delta z)\cos(n\omega t)$, which represent the optical background in a typical *a*-SNOM experiment with propagative-field illumination. In order to measure both the dc part (BKG_0) and the harmonics BKG_n, the PMT output is split in two channels and fed, respectively, to a low-pass filter ($f_{cut}=1.6$ kHz$\ll f_{res}$), and to a lock-in amplifier synchronous with the *n*th harmonic of the TF oscillation frequency ω. We set the lock-in in amplifier to acquire either the in-phase signal $\mathrm{BKG}_n(a_0,z)$, or its modulus $|\mathrm{BKG}_n(a_0,z)|$. A digital oscilloscope is used to monitor the optical signals, simultaneously with the voltage V2 providing the sample vertical position.

Figures 2(a) and 2(b) show, respectively, the behavior of BKG_0 and of the modulus of BKG_1($|BKG_1|$), simultaneously acquired as a function of the sample position, in a scan of 1 μm wide (scan speed of 2 μm/s). In Figs. 2(c) and 2(d) we report two analogous plots for the BKG_0 and the in-phase component of BKG_2, simultaneously acquired. All the signals show an oscillatory dependence on the sample excursion Δz, with periodicity $\lambda/2$, as a result of the interference of the fields scattered by the stray light sources mentioned above. The signals BKG_n are characterized by a modulation amplitude A_n superimposed to a constant offset C_n due, as we will see in the next section, to the backscattering of the lens. Such an offset causes the height asymmetry of the bumps in the $|BKG_1|$ signal [Fig. 2(b)]. The tip oscillation amplitude in this experiment is $a_0=(3.4\pm0.1)$ nm,[29] equivalent to a ratio $a_0/\lambda \approx 5\times10^{-3}$.

To quantitatively assess the BKG rejection power of the demodulation technique at higher harmonics, we can define the quantity $R_n=A_n/A_0$, namely, the *n*th-harmonic-to-dc

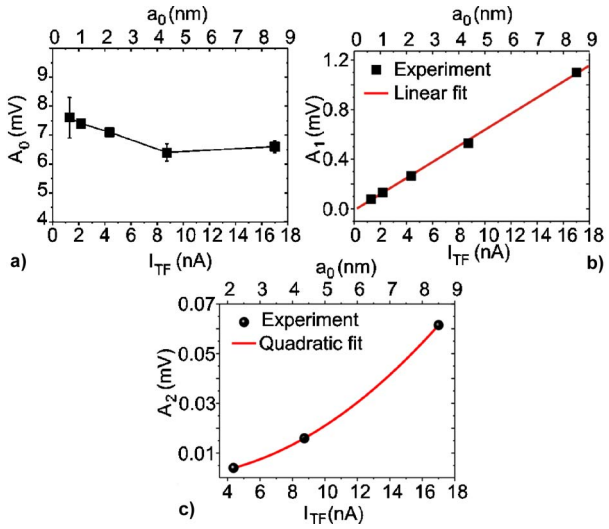

FIG. 3. (Color online) Dependence of the BKG modulation amplitudes $A_{0,1,2}$ [(a)–(c) respectively] on the tuning-fork current (bottom axes) and on the tip oscillation amplitude (top axes). The linear behavior of A_1 is pointed out by a linear fit [(b), red line)]. A power fit ($y=Ax^n$, $n=2.02\pm0.02$) of the second harmonic BKG amplitude [(c), red line)] evidences the quadratic dependence on a_0.

background ratio, representing the fraction between the modulation amplitude of the background demodulated at the *n*th harmonic with respect to the dc one. The stronger the BKG suppression, the smaller the expected values of R_n. In the experiment of Fig. 2, for example, we find decreasing values of $R_1=6.5\times10^{-2}$ and $R_2=1.2\times10^{-3}$, assessing that the second harmonic demodulation provides a BKG suppression superior of a factor of 50 with respect to the first harmonic demodulation. Even though the BKG_2 is still well above the noise. To study how the BKG_n depends on the tip oscillation amplitude we have monitored the modulation amplitudes $A_{0,1,2}$ varying a_0 in the 0.5–10 nm range. In Figs. 3(a)–3(c) we can note the different behaviors of the three signals. While A_0 is almost constant, the first and the second harmonics increase with a_0, respectively, in a linear and a quadratic fashion [red line in Fig. 3(c)]. In Fig. 4 we plot the corresponding values of R_1 [Fig. 4(a)] and R_2 [Fig. 4(b)]. These graphics allow us to conclude that for tip oscillation amplitudes in the 0.5–10 nm range first harmonic demodulation reduces the BKG by one-two orders of magnitude. For tip oscillation amplitudes between 2 and 10 nm, second harmonic demodulation reduces the BKG down to 10^{-2}–10^{-4}. In our experiment the second harmonic signal is rejected below the noise level for $a_0<1.5$ nm.

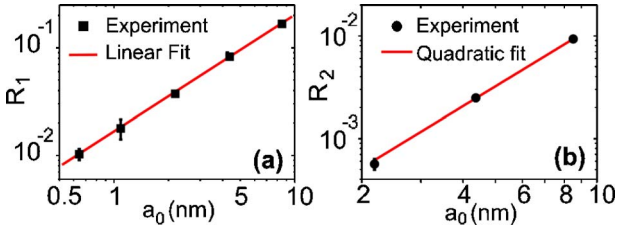

FIG. 4. (Color online) Plot of the *n*th-harmonic-to-dc-BKG ratio R_n as a function of the tip oscillation amplitude for first (a) and second harmonic (b) demodulations. The red lines reproduce the linear (a) and quadratic (b) fits of the experimental data.

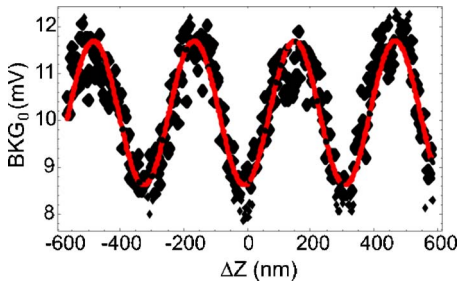

FIG. 5. (Color online) (black symbols) Plot of the dc background observed when removing the tip from the optical path as a function of the sample position. (red line) Sinusoidal fit of the experimental data [$y = A\sin(4\pi/\lambda \cdot \Delta z + \phi)$, $A = (1.5 \pm 0.1)$ mV].

It is also important to evaluate the contribution of the field backscattered by the focusing lens to the total BKG signal. We expect that, since the average tip-lens distance is constant, the contribution of E_L to the harmonic terms BKG_n will be a mere constant offset, i.e., not depending on Δz. Vice versa, the dc term BKG_0, due to the interference between E_L and E_S ($\mathrm{BKG}_0 = |E_S + E_L|^2$), is expected to depend on the sample's position Δz sinusoidally. For these measurements we have retracted the tip by several millimeters and swept the sample vertically. In Fig. 5 (black symbols) we clearly see the interference pattern due to the variation of the lens-sample distance. From a measure of the oscillation amplitude (the red line is from a sinusoidal fit), we assess that the sample-lens interference contribution amounts to a maximum of 10% of the total dc background (BKG_0) measured in the presence of the tip [Fig. 2(a)].

III. THEORETICAL INTERPRETATION

Understanding the physical origin of the background and, in particular, of its harmonic components is of fundamental importance to model and quantitatively evaluate the rejection power of higher harmonics demodulation. The presence of a modulated background field using evanescent wave illumination has been pointed out by Hudlet *et al.*[22] Concerning the illumination with propagative fields, Formanek *et al.*,[25] within the dipole approximation,[9] have shown that the BKG modulated at the first harmonic arises from the interplay of the fields reflected by the sample and the tip shaft. In the following, we will see that such stray fields, interfering with the back reflection of the focusing optics, provide a source of BKG components modulated at every harmonic of the tip vibration frequency. Following the scheme of Fig. 1, let us define a coordinate system in which the z-axis points downward and the zero coincides with the tip position at rest. In this system we call $a(t) = a_0\cos(\omega t)$ the tip position, l_0 the position of the focusing lens, $z(t) = z_0 + \Delta z(t)$ the sample position, where z_0 is the position at rest, and $\Delta z(t)$ the vertical excursion as a function of the time. As we have observed experimentally, the BKG originates from the interference on the detector of three fields backscattered by the tip shaft $E_F = E_F^0 \exp(-i\Delta\phi_{FL})$, the sample's bottom surface $E_S = E_S^0 \exp(-i\Delta\phi_{SL})$, and the focusing optics $E_L = E_L^0$. Where, given $k = 2\pi/\lambda$, the phases $\Delta\phi_{FL} = 2k[l_0 - a(t)]$ and $\Delta\phi_{SL} = 2k[l_0 - z(t)]$ are related to the optical path differences accumulated by the fields scattered, respectively, by the tip and by the sample, with respect to the one scattered by the lens, in their travel from the laser to the detector. Substituting in Eq. (1), the expression reduces to

$$\begin{aligned}\mathrm{BKG} = {} & C_0 + B_{SL}\cos[\phi_1 - 2k\Delta z(t)] + B_{FL} \\ & \times\cos[\phi_0 - 2ka_0\cos(\omega t)] + B_{FS}\cos[\phi_2 + 2k\Delta z(t) \\ & - 2ka_0\cos(\omega t)],\end{aligned} \tag{2}$$

where the constant parameters $C_0 = |E_F|^2 + |E_S|^2 + |E_L|^2$, $B_{SL} = 2|E_L E_S|$, $B_{FL} = 2|E_L E_F|$, and $B_{FS} = 2|E_F E_S|$ depend on the amount of light backscattered by the single sources. The phases $\phi_0 = 2kl_0$, $\phi_1 = 2k(l_0 - z_0)$, and $\phi_2 = 2kz_0$ depend on the distances of the lens and of the sample at rest from the tip. The latter quantities can be assumed as constants, as well, unless mechanical or thermal drifts take place moving apart the scattering sources. Expanding the cosines in Eq. (2) we find

$$\begin{aligned}\mathrm{BKG} = {} & C_0 + B_{SL}\cos[2k\Delta z(t) - \phi_1] + \{B_{FL}\cos(\phi_0) \\ & + B_{FS}\cos[\phi_2 + 2k\Delta z(t)]\}\cos[2ka_0\cos(\omega t)] \\ & + \{B_{FL}\sin(\phi_0) + B_{FS}\sin[\phi_2 + 2k\Delta z(t)]\} \\ & \times\sin[2ka_0\cos(\omega t)],\end{aligned} \tag{3}$$

from which we evidence the presence of terms having the form $\cos(\alpha\cos\omega t)$ and $\sin(\alpha\cos\omega t)$. The latter expressions can be developed into a series of terms oscillating at the harmonic frequencies $n\omega$, whose amplitudes are given by the nth-order Bessel functions of first kind J_n.[32] This is formally the reason why we must expect a BKG signal at every harmonic order. In particular, the expansion of Eq. (3) yields

$$\begin{aligned}\mathrm{BKG} = {} & C_0 + B_{SL}\cos[2k\Delta z(t) - \phi_1] + \{B_{FL}\cos(\phi_0) \\ & + B_{FS}\cos[\phi_2 + 2k\Delta z(t)]\}\cdot J_0(2ka_0) \\ & + 2\{B_{FL}\cos(\phi_0) + B_{FS}\cos[\phi_2 + 2k\Delta z(t)]\} \\ & \times\sum_{n=1}^{\infty}(-1)^n J_{2n}(2ka_0)\cos[(2n)\omega t] \\ & + 2\{B_{FL}\sin(\phi_0) + B_{FS}\sin[\phi_2 + 2k\Delta z(t)]\} \\ & \times\sum_{n=0}^{\infty}(-1)^n J_{2n+1}(2ka_0)\cos[(2n+1)\omega t],\end{aligned} \tag{4}$$

which is an expression of the form $\mathrm{BKG}_0(a_0, \Delta z) + \Sigma_{n\geq 1}\mathrm{BKG}_n(a_0, \Delta z)\cdot\cos(n\omega t)$, consisting of a dc signal plus a series of harmonics whose amplitudes depend both on the sample position and on the tip vibration amplitude. The amplitude of the dc component is

$$\begin{aligned}\mathrm{BKG}_0(a_0, \Delta z) = {} & C_0 + B_{SL}\cos[2k\Delta z(t) - \phi_1] \\ & + J_0(2ka_0)\cdot\{B_{FL}\cos(\phi_0) \\ & + B_{FS}\cos[\phi_2 + 2k\Delta z(t)]\},\end{aligned} \tag{5}$$

while the amplitudes of the even and of the odd harmonics are, respectively,

$$\mathrm{BKG}_{2n}(a_0,\Delta z) = 2\cdot(-1)^n J_{2n}(2ka_0)\cdot\{B_{FL}\cos(\phi_0) + B_{FS}\cos[\phi_2 + 2k\Delta z(t)]\}, \quad (6)$$

and

$$\mathrm{BKG}_{2n+1}(a_0,\Delta z) = 2\cdot(-1)^n J_{2n+1}(2ka_0)\cdot\{B_{FL}\sin(\phi_0) + B_{FS}\sin[\phi_2 + 2k\Delta z(t)]\}. \quad (7)$$

These expressions represent the transfer functions between the average vertical position of the sample (Z) and the optical far-field background. Let us analyze them in more detail. The dc component is composed of a constant offset $C_0 + J_0(2ka_0)\cdot B_{FL}\cos(\phi_0)$ slightly dependent on a_0 (usually $C_0 \gg B_{FL}$), superimposed to a term

$$B_{SL}\cos[2k\Delta z(t) - \phi_1] + [J_0(2ka_0)B_{FS}]\cdot\cos[\phi_2 + 2k\Delta z(t)], \quad (8)$$

oscillating with the sample position, with period $\Delta z = \lambda/2$. This term arises from the linear superposition of the sample-lens and of the tip-sample interferences, respectively. The two contributions add coherently (amplitude and phase) yielding the total BKG_0 signal experimentally observed in Figs. 2(a) and 2(c). In particular, in absence of the tip [$B_{FS} = 0$ in Eq. (8)], the light scattered by the focusing lens leads to a BKG_0 signal whose modulation amplitude B_{SL} is much smaller than B_{FS}, as reported in Fig. 5. Therefore, since the second term of Eq. (8) is prevalent, we can conclude that the dc background depends on the tip-oscillation amplitude as $\mathrm{BKG}_0 \propto J_0(2ka_0)$. The structure of the BKG at the higher harmonics is similar. The even harmonics [Eq. (6)] are characterized by an offset $2(-1)^n J_{2n}(2ka_0)B_{FL}\cos(\phi_0)$, induced by the backscattering of the lens ($B_{FL} = 2|E_L E_F|$), plus a term $2(-1)^n J_{2n}(2ka_0)B_{FS}\cos[\phi_2 + 2k\Delta z(t)]$ oscillating with the sample vertical position, whose amplitude A_n scales proportionally to $J_n(2ka_0)$. The odd harmonics [Eq. (7)] show the same behavior, but the oscillatory dependence on the sample position is shifted by 90° with respect to the even ones. For small oscillations, $ka_0 << 1$, we can expand the Bessel functions in power series, finding that the background modulation amplitude depends on the tip oscillation amplitude through a polynomial law of the dimensionless quantities $(ka_0)^n \propto (a_0/\lambda)^n$.[29] To the first order in ka_0, R_n is expected to be

$$R_n = \frac{|A_n|}{|A_0|} = \frac{2J_n(2ka_0)}{J_0(2ka_0)} \approx \frac{2}{n!}\left(\frac{2\pi a_0}{\lambda}\right)^n. \quad (9)$$

That is, the background at the nth harmonic scales down as the nth power of the ratio a_0/λ multiplied by a further factor $1/n!$. In particular, the advantage of increasing by 1 the harmonic order, in terms of BKG reduction, is $R_{n+1}/R_n \approx ka_0/(n+1)$. This value must be compared with the near-field signal loss due to the spatial filtering effect[17,18] to assess the best experimental parameters. In Fig. 6 we plot the BKG amplitude at the various harmonics A_n as a function of (a_0/λ) (in the top axis we report the values of a_0 for $\lambda = 633$ nm). In the visible range (e.g., for HeNe illumination), and for a tip oscillation of 20 nm peak to peak ($a_0 = 10$ nm), increasing the harmonic order $n = 1$, 2, and 3 leads to an increased rejection of the background, whose modulation amplitude will be reduced to values of 10^{-1}, 5×10^{-3}, and 2×10^{-4} with respect to the dc amplitude. Decreasing the tip oscillation by one order of magnitude, from 10 to 1 nm (thicker lines in Fig. 6), has different consequences depending on the harmonic order of the BKG. While the dc signal (black line) is almost unaffected, the amplitude of the signal demodulated at the first, second, and third harmonics (red, blue, and green lines) is reduced, respectively, by further one, two, and three orders of magnitude. In the mid IR ($\lambda \sim 10\ \mu$m) the technique is even more powerful, since for $a_0 = 10$ nm [namely, $(a_0/\lambda) \approx 10^{-3}$], first, second, and third harmonic demodulations will reduce the BKG amplitude of factor of 6×10^{-3}, 2×10^{-5}, and 4×10^{-8} with respect to the dc amplitude, respectively. The higher rejection power in the IR part of the electromagnetic spectrum explains why far-field oscillations are clearly visible in the second harmonic approach curves at in the visible ($\lambda = 633$ nm) reported by Hillenbrand and Keilmann[11] and Roy *et al.*,[24] while there is just an outline in the 2ω curve at $\lambda = 10.6\ \mu$m reported by Formanek *et al.*[25]

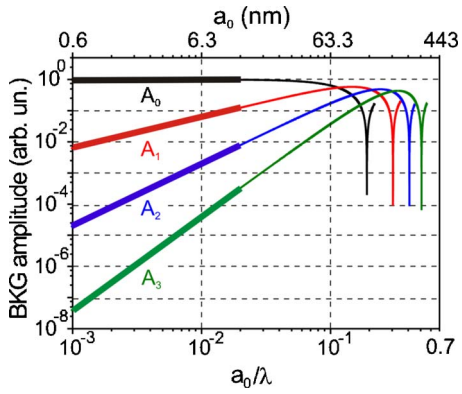

FIG. 6. (Color online) Plot of the background modulation amplitudes at the various harmonics [dc (black), first (red), second (blue), and third (green)] as a function of the ratio (a_0/λ) (bottom axis) and of a_0 assuming $\lambda = 633$ nm (top axis).

Decreasing the oscillation amplitude is, however, not the only way to reduce the BKG. In fact, when we plot the correct expression $|J_n(2ka_0)|$ for the amplitude of the BKG harmonics, as in Fig. 6, it emerges that there are some particular values of the tip oscillation amplitude $\tilde{a}_0$ that strongly reduce the BKG. Such values represent the first zeros of the Bessel functions for which $J_n(2k\tilde{a}_0) = 0$. The corresponding values of the ratios $\tilde{a}_0/\lambda$ for $n = 0, \ldots, 4$ fall in the interval $a_0/\lambda \sim 0.2$–0.7, as reported in Table I. In order to reject the BKG at the nth harmonic we can thus either decrease the oscillation amplitude to smaller and smaller values, or tune the oscillation amplitude a_0 to well defined values $\tilde{a}_0$ which

TABLE I. Theoretical values of the ratios $\tilde{a}_0/\lambda$, corresponding to the zeros of the J_n Bessel functions, that yield complete background suppression at each harmonic n. The values $\Delta a_0/\tilde{a}_0$ indicate the precision needed on $\tilde{a}_0$ to have a BKG rejection better than 10^3 at every harmonic.

Order	$n=0$	$n=1$	$n=2$	$n=3$	$n=4$
$\tilde{a}_0/\lambda$	0.191	0.305	0.409	0.508	0.604
$\Delta a_0/\tilde{a}_0$	0.0016	0.0013	0.0013	0.001	0.001

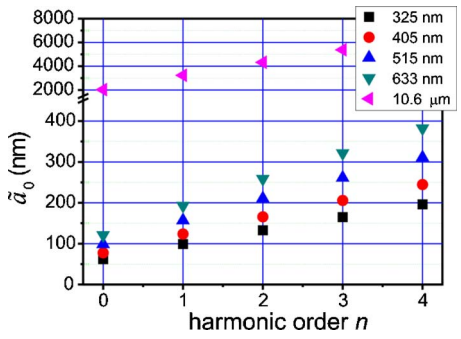

FIG. 7. (Color online) Values of the tip oscillation amplitude $\tilde{a}_0$ corresponding to the first zeros of the nth Bessel functions for various wavelengths λ.

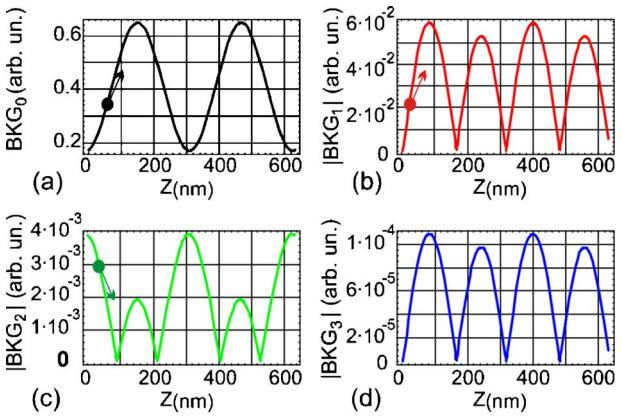

FIG. 8. (Color online) Approach curves predicted for the dc background (a) and the modulus of the first three harmonics [(b)–(d), respectively] base on Eqs. (5)–(7), for $a_0=10$ nm and $\lambda=633$ nm. The colored circles and the arrows indicate how the optical signal changes as a function of an increased sample topography.

depend on n and λ. The numerical values of $\tilde{a}_0$ for some typical laser wavelengths are shown in Fig. 7. In particular, in the NUV-visible region (black, red, blue, and green symbols in Fig. 7), the spectrum of values of $\tilde{a}_0$ which allow to null the BKG at the first three harmonics falls in the 100–200 nm range. These values are certainly reasonable for experimental configurations employing freestanding tips or atomic force microscopy (AFM) cantilevers,[33,34] but quite unusual for systems based on quartz tuning forks, suited for low oscillation amplitudes. On the contrary, oscillation amplitudes of several microns (pink symbols in Fig. 7) would be required to accomplish such a task in the mid-IR ($\lambda=10.6$ μm). In order to establish how fine the tuning of a_0 must be, we note that the Bessel functions are almost linear around their first zeros. This implies that the maximum discrepancy interval $\Delta a_0/\tilde{a}_0$ allowed to have a BKG rejection of three orders of magnitude is of the order of 0.1%. More precise calculations for the different harmonics are reported in Table I.[35]

IV. NONLINEAR TOPOGRAPHY ARTIFACTS

In near-field optical microscopy an artifact can generally be defined as information recorded in the optical map which has not an optical origin. In particular, topography artifacts are induced by the vertical motion of the sample that modulates the intensity of the far-field BKG from point to point. Such signal, originating from an interference process, is thus expected to vary on tip-sample distances of the order of the laser wavelength. In particular, if the topography excursions are much smaller than λ, the coupling can be assumed as linear,[7] leading to optical maps very similar to the topographical counterpart, and easily identifiable as affected by artifacts. If the height of the topography structures becomes comparable with the wavelength, or larger, the sinusoidal character of the interference profile will provide optical maps qualitatively different from the topography.[8] From the physical point of view there are almost no differences between topography artifacts in aperture- and apertureless-SNOM. In aperture-SNOM (illumination mode), the far-field emission from the probe interferes with the light reflected by the sample. In apertureless-SNOM it is the light scattered by the tip shaft that interferes with the light reflected by the sample (plus a smaller contribution from the focusing lens which, as we have seen, is negligible). As a consequence, the BKG transfer functions are expected to be analogous. Nevertheless, the different experimental techniques employed to detect the optical signals can yield different results. In aperture-SNOM, the tip is not vertically dithered and the dc component $S \propto A + B\cos[K\Delta z(t)+\phi]$ is the only signal acquired. It consists of an offset A superimposed to a smaller modulation ($B \ll A$). K varies from $k=2\pi/\lambda$ for grazing collection to $k=4\pi/\lambda$ for a collection geometry parallel to the tip axis.[8,36,37] In apertureless-SNOM higher harmonics demodulation requires the use of a lock-in amplifier whose output is usually set to provide the modulus of the optical signal, i.e., $S_n \propto |A+B\cos[K\Delta z(t)+\phi]|$. Since the modulation amplitude B is much greater than the offset A ($B \gg A$), the resulting approach curves will be similar to the ones in Fig. 2(b).[10,25] In particular, monitoring the modulus rather than the in-phase component of the optical signal, accentuates its nonlinear dependence on the sample position. The signal, in fact, will vary on length scales which are about halved, passing from the minimum to the maximum on distances of $\sim\lambda/7$ [Fig. 2(b)], rather than $\lambda/4$ [Fig. 2(c)].[38] As a consequence, at least in the visible range, topographic structures as short as 60 nm are already expected to produce optical artifacts nonlinearly related with the topography. To better visualize such phenomenon, starting from Eqs. (5)–(7), we have calculated the approach curves that would be expected for typical experimental conditions: $a_0=10$ nm, $\lambda=633$ nm, 0.5% of light back reflected from the lens, 4% from the sample, and 50% from the tip shaft. In Fig. 8 we show the behavior of the dc background (a) and the modulus of the first three harmonics [(b)–(d), respectively] as a function of the distance between the tip apex and the average sample plane Z. We assume random phases ϕ_0 and ϕ_2 (since they depend on the tip length and on the lens position, these parameters are not really controllable in a experiment). The position $Z=0$ corresponds to the contact point of the tip with the sample surface. In CGM operation, any topographic feature will cause a drawdown of the sample, and thus an increase of Z. As a consequence, the BKG signal will vary according to the curves in Fig. 8. It must be noted that, in the absence of a clear fingerprint of the near-field optical interaction in the contact region ($Z\sim 0$), the optical signal mea-

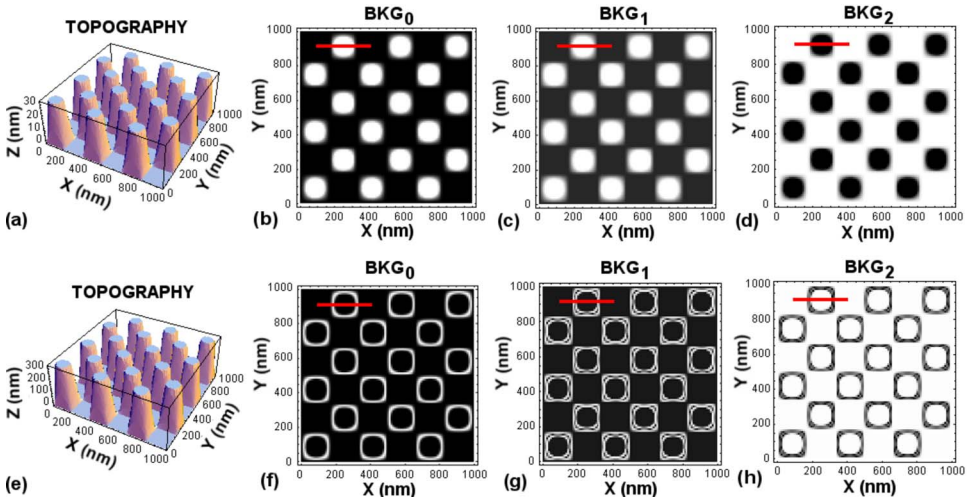

FIG. 9. (Color online) [(b) and (f)] dc, [(c) and (g)] first harmonic, and [(d) and (h)] second harmonic BKG maps simulated for a topography [(a) and (e)] consisting of pillars on a horizontal substrate, having heights of 30 and 300 nm, respectively.

sured will be due entirely to the BKG contribution. And even the presence of such a fingerprint does not guarantee that the optical signal, and the corresponding optical image, will be artifact-free. We must, in fact, assure that the changes of the optical signal are not entirely due to the BKG modulations, but do originate from the near-field scattering of the sample. As a consequence, we should check that at least one of the following conditions is satisfied: (i) the BKG is actually below the noise level, (ii) the excursions of the recorded optical signal are larger than the BKG modulations observed in the far-field region of the approach curves, and (iii) topography structures located at the same absolute position Z produce different optical responses (as in Ref. 11). From Eq. (7) we note that the first and the third harmonic signals have the same functional dependence on Z [Figs. 8(b) and 8(d)]. Therefore, in the presence of artifacts, the corresponding optical maps are expected to be qualitatively alike (the only difference relying on the magnitude of the signal). More generally, we expect the images acquired at the even (or at the odd) harmonics to be all qualitatively identical, independently the specific harmonic order. This would not be the case, indeed, if some near-field contribution was present, due to high-pass filtering effect on the image spatial frequencies.[18]

It is interesting to visualize the fictitious optical maps expected for an experiment like the one mentioned above, assuming the transfer functions plotted in Fig. 8, calculated from Eqs. (5)–(7). In Figs. 9(b)–9(d) we display the images calculated for the dc, the modulus of the first and of the second harmonics, respectively, for a sample consisting of set of truncated cone pillars, having a base diameter of 170 nm and height of 30 nm grown on a flat glass coverslip with plane parallel faces [Fig. 9(a)]. The total sample's scattering is assumed to be due entirely to the back reflection of the bottom coverslip surface, i.e., the scattering of the pillars is assumed negligible. For small topography excursions optical images qualitatively similar to the topography are retrieved, but in the second harmonic map (d) the contrast is reversed. The different behaviors can be understood considering the slope of the approach curves at the contact point (Z=0), which is positive for the dc and the first harmonic [Figs. 8(a) and 8(b)], and negative the second harmonic [Fig.

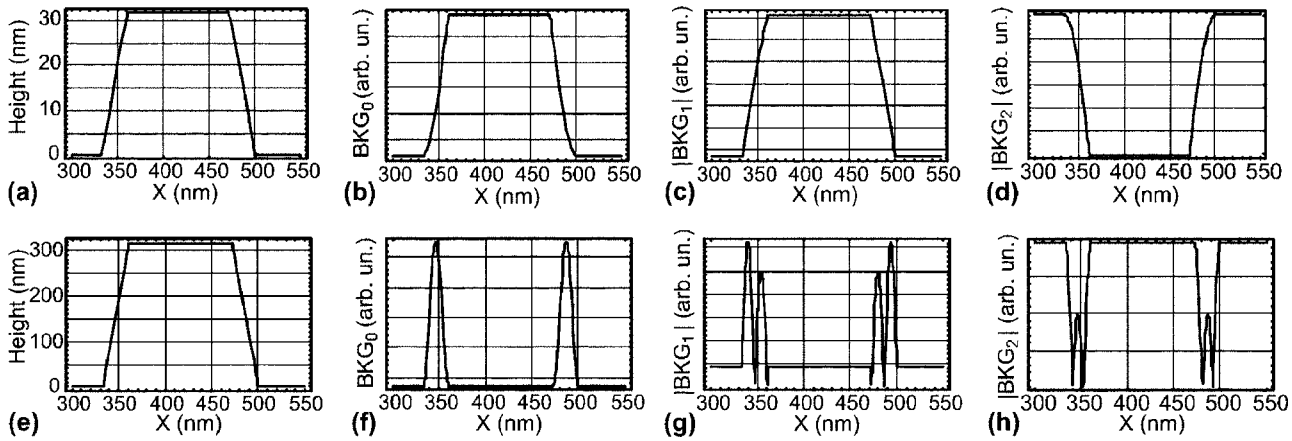

FIG. 10. [(a)–(h)] Line profiles drawn in correspondence of the red lines marked in Fig. 9. The optical profiles reproduce the topography only for the 30 nm high pillars.

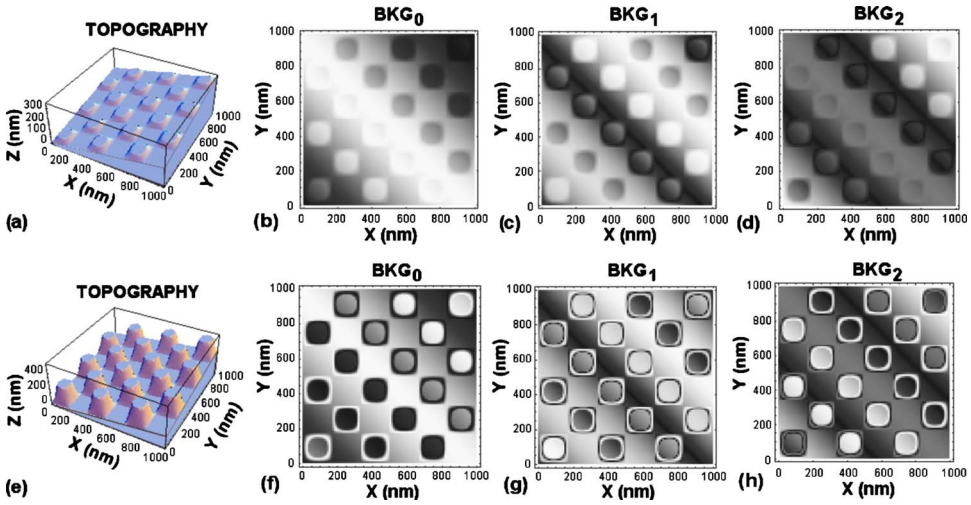

FIG. 11. (Color online) [(b) and (f)] dc, [(c) and (g)] first harmonic, and [(d) and (h)] second harmonic BKG maps simulated a topography [(a) and (e)] consisting of pillars on an wedgelike substrate (inclination 10°), having heights of 30 and 300 nm, respectively.

8(c)]. For structures sensibly higher, as in Fig. 9(e) (300 nm, i.e., $\sim\lambda/2$), the optical maps become different from the topography. Annular structures appear in the dc image [Fig. 9(f)], double rings show up in the $|BKG_1|$ map [Fig. 9(g)], and contrast reversal is observed in the $|BKG_2|$ image [Fig. 9(h)]. The qualitative difference with respect to the previous structures is better highlighted in the line profiles of Fig. 10. For pillar heights of $\sim\lambda/20$ the transfer functions in Figs. 8(a) and 8(c) behave as linear. Therefore, the optical signals in Figs. 10(b) and 10(d) perfectly reproduce the topography profile [Fig. 10(a)], apart from the contrast inversion in the second harmonic. In particular, no lateral shifts are expected between the topography and the optical maps. For pillar heights of $\sim\lambda/2$ [Fig. 10(e)] the nonlinearity of the transfer function yields modulations of the optical signals [Figs. 10(f)–10(h)] difficult to attribute to artifacts *a priori*. The signal enhancement at the structure edges [Figs. 10(g) and 10(h)] could, in fact, be confused with a high-pass spatial frequency filtering,[18] and be attributed to a genuine near-field effect. In particular, the shape and position of such modulations will depend on several experimental parameters such as the harmonic number (experimentally controllable), or the phases ϕ_0, and ϕ_2 (hard to control). This phenomenon appears in a more intriguing way if we suppose that the pillars have grown a coverslip whose faces are not perfectly parallel, i.e., on a wedge. In such case the different local thicknesses of the coverslip will induce a further optical path difference between the light reflected by the bottom sample's surface, and the light reflected by the tip, which will follow the top coverslip surface. As a consequence, a long range sinusoidal modulation is expected in the optical signal, as evidenced in Fig. 11. Here we have simulated the fictitious optical maps (dc, modulus of the first, and of the second harmonics) that would arise for pillars of 30 nm [(b)–(d)] and 300 nm [(f)–(h)] grown onto a coverslip whose top surface is inclined of $\sim 10°$ with respect to the bottom surface, assumed horizontal [Figs. 11(a) and 11(e)]. In particular, we see [Fig. 11(b)] that the pillars can appear either brighter or darker with respect to the neighboring substrate, depending on their spatial position within the map. It can also happen that half of the single pillar looks bright and half dark, as in Fig. 11(d). Moreover, if the pillars height is comparable to the wavelength, they can appear as bright rings with a dark interior or vice versa within the same map, as in Figs. 11(g) and 11(h). A similar situation has already been encountered experimentally for the demodulated first harmonic signal.[3] The presence of artifacts in similar cases can, however, be pointed out noting that an increase of the harmonic order is not expected to suppress the long range modulations. The disappearance of such a feature from the images at higher harmonics can thus be held as a proof of the genuine nature of the optical signal.[3,19]

V. CONCLUSIONS

In conclusion, the far-field background in apertureless-SNOM has been studied both theoretically and experimentally. The occurrence of a signal modulated at frequencies multiple of the tip vibration, evidenced experimentally, has been explained in terms of the interference process taking place between the stray fields scattered by the sample and the tip shaft. The field backscattered by the focusing optics has been observed to take part only in the dc background. The power-law dependence of the background modulations on the harmonic order has been put forward, explaining quantitatively the advantages obtainable in terms of background rejection when increasing the harmonic order. The possibility to reject the background decreasing the tip oscillation amplitude has been quantitatively studied, highlighting at the same time the opportunity to exploit the first zeros of the Bessel functions to accomplish such a task. Finally, we have simulated different optical images induced by topography artifacts in correspondence to structures having different heights, deposited on flat or inclined substrates, pointing out the effect of the nonlinear BKG-topography coupling, and of the substrate inclination on the optical images. This, in particular, has allowed us to state a precise criterion for artefacts identification: in presence of artefacts, the optical maps acquired at the even harmonics are expected to be all qualitatively identical, independently from the harmonic order (the

same states for the odd harmonics). Therefore, any qualitative difference between the optical images acquired at different even (or odd) harmonics, such as the disappearance of the long range modulations induced by the sample's inclination, can be reasonably ascribed to genuine near-field spatial filtering effects.[3,12,19]

ACKNOWLEDGMENTS

A. Mlayah is acknowledged for carefully reading the manuscript and M. Labardi for fruitful discussions. One of the authors (G.B.) acknowledges the CNR-CNRS bilateral project "Diffusione Raman e Brillouin risonante e localizzazione spaziale di stati elettronici" for partial financial support. P.G.G. is grateful to D. Arigò and G. Spinella for their skill ful expertise in the SNOM setup manufacturing.

[1]D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984); A. Lewis, M. Isaacson, A. Harootunian, and A. Murray, Ultramicroscopy **13**, 227 (1984).
[2]E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251**, 1468 (1991).
[3]S. Patanè, P. G. Gucciardi, M. Labardi, and M. Allegrini, Riv. Nuovo Cimento **27**, 1 (2004).
[4]L. Novotny, R. X. Bian, and X.-S. Xie, Phys. Rev. Lett. **79**, 645 (1997).
[5]F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, Science **269**, 1083 (1995).
[6]A. Lahrech, R. Bachelot, P. Gleyzes, and A. C. Boccara, Opt. Lett. **21**, 1315 (1996).
[7]B. Hecht, H. Bielefeldt, Y. Inouye, D. W. Pohl, and L. Novotny, J. Appl. Phys. **81**, 2492 (1997).
[8]P. G. Gucciardi and M. Colocci, Appl. Phys. Lett. **79**, 1543 (2001).
[9]B. Knoll and F. Keilmann, Opt. Commun. **182**, 321 (2000).
[10]R. Hillenbrand and F. Keilmann, Phys. Rev. Lett. **85**, 3029 (2000).
[11]R. Hillenbrand and F. Keilmann, Appl. Phys. Lett. **80**, 25 (2001).
[12]M. Labardi, S. Patanè, and M. Allegrini, Appl. Phys. Lett. **77**, 621 (2000).
[13]T. Taubner, F. Keilmann, and R. Hillenbrand, Nano Lett. **4**, 1669 (2004).
[14]L. Stebunova, B. B. Abramitchev, and G. C. Walker, Rev. Sci. Instrum. **74**, 3670 (2003).
[15]R. Hillenbrand, T. Taubner, and F. Keilmann, Nature (London) **418**, 159 (2002).
[16]N. Ocelic and R. Hillenbrand, Nat. Mater. **3**, 606 (2004).
[17]M. Labardi, S. Patanè, and M. Allegrini, in *Proceedings of the International School of Physics "E. Fermi" Course CXLIV, Varenna, Italy, 2001*, edited by M. Allegrini, N. Garcia, and O. Marti (IOS, Amsterdam, 2001), p. 425.
[18]J. N. Walford *et al.*, J. Appl. Phys. **89**, 5159 (2001).
[19]N. Maghelli, M. Labardi, S. Patanè, F. Irrera, and M. Allegrini, J. Microsc. **202**, 84 (2001).
[20]I. Stefanon, S. Blaize, A. Bruyant, S. Aubert, G. Lerondel, R. Bachelot, and P. Royer, Opt. Express **13**, 5553 (2005).
[21]T. Taubner, F. Keilmann, and R. Hillenbrand, J. Korean Phys. Soc. **47**, S213 (2005).
[22]S. Hudlet *et al.*, Opt. Commun. **230**, 245 (2004).
[23]F. Formanek, Y. De Wilde, and L. Aigouy, J. Appl. Phys. **93**, 9548 (2003).
[24]D. Roy, S. H. Leong, and M. E. Welland, J. Korean Phys. Soc. **47**, S140 (2005).
[25]F. Formanek, Y. De Wilde, and L. Aigouy, Ultramicroscopy **103**, 133 (2005).
[26]B. Ren, G. Picardi, and B. Pettinger, Rev. Sci. Instrum. **75**, 837 (2004).
[27]K. Dickmann, F. Demming, and J. Jersch, Rev. Sci. Instrum. **67**, 845 (1996).
[28]E. Cefalì, S. Patanè, P. G. Gucciardi, M. Labardi, and M. Allegrini, J. Microsc. **210**, 262 (2003).
[29]P. G. Gucciardi, G. Bachelier, A. Mlayah, and M. Allegrini, Rev. Sci. Instrum. **76**, 036105 (2005).
[30]K. Karrai and R. D. Grober, Appl. Phys. Lett. **66**, 1842 (1995).
[31]K. Karrai, in *Le Champ Proche Optique*, edited by D. Courjon and C. Bainier (Springer-Verlag, Paris, 2001).
[32]M. R. Spiegel and J. Liu, *Mathematical Handbook of Formulas and Tables*, 2nd ed. (McGraw-Hill, New York, 1999).
[33]R. Bachelot, P. Gleyzes, and A. C. Boccara, Opt. Lett. **20**, 1924 (1995).
[34]P.-M. Adam, P. Royer, R. Laddada, and J.-L. Bijeon, Ultramicroscopy **71**, 327 (1998).
[35]Here R_n is considered as the ratio between the BKG amplitude A_n and the amplitude of the dc component A_0 calculated in the small oscillations approximation ($ka_0 \ll 1$).
[36]P. G. Gucciardi, M. Labardi, S. Gennai, F. Lazzeri, and M. Allegrini, Rev. Sci. Instrum. **68**, 3088 (1997).
[37]M. Labardi, P. G. Gucciardi, M. Allegrini, and C. Pelosi, Appl. Phys. A: Mater. Sci. Process. **66**, S397 (1998).
[38]The periodicity is exactly halved only if the offset is null.