Analysis of detection limit to time-resolved coherent anti-Stokes Raman scattering nanoscopy*

Liu Wei(刘 伟), Liu Shuang-Long(刘双龙), Chen Dan-Ni(陈丹妮)†, and Niu Han-Ben(牛憨笨)‡

Institute of Optoelectronics, Key Laboratory of Optoelectronic Deviced and Systems of Education Ministry, Shenzhen University, Shenzhen 518060, *China*

(Received 26 December 2013; revised manuscript received 14 March 2014; published online 10 August 2014)

In the implementation of CARS nanoscopy, signal strength decreases with focal volume size decreasing. A crucial problem that remains to be solved is whether the reduced signal generated in the suppressed focal volume can be detected. Here reported is a theoretical analysis of detection limit (DL) to time-resolved CARS (T-CARS) nanoscopy based on our proposed additional probe-beam-induced phonon depletion (APIPD) method for the low concentration samples. In order to acquire a detailed shot-noise limited signal-to-noise (SNR) and the involved parameters to evaluate DL, the T-CARS process is described with full quantum theory to estimate the extreme power density levels of the pump and Stokes beams determined by saturation behavior of coherent phonons, which are both actually on the order of $\sim 10^9$ W/cm^2. When the pump and Stokes intensities reach such values and the total intensity of the excitation beams arrives at a maximum tolerable by most biological samples in a certain suppressed focal volume (40-nm suppressed focal scale in APIPD method), the DL correspondingly varies with exposure time, for example, DL values are 10^3 and 10^2 when exposure times are 20 ms and 200 ms respectively.

Keywords: break through the diffraction limit, coherent anti-Stokes Raman scattering, nonlinear optics, detection limit

PACS: 42.25.Fx, 42.65.Dr, 42.65.–k, 87.64.–t **DOI:** 10.1088/1674-1056/23/10/104202

1. Introduction

Coherent anti-Stokes Raman scattering (CARS) microscopy has proven to be a promising nonlinear optical technique that is capable of obtaining high-sensitivity and three-dimensional images based on molecular vibrational spectrum without labeling.[1–3] As a nonlinear process, CARS enables three-dimensional imaging with high spatial resolution, but it is still limited by diffraction and thus cannot preferably visualize the cellular interior. Several successful resolution enhancement techniques in fluorescence microscopy open up possibilities for imaging cells with spatial resolution significantly beyond the diffraction limit.[4–6] In comparison, even though several schemes have been suggested recently for breaking through the diffraction barrier in CARS microscopy, current CARS implementations are yet unable to achieve nanoscale resolution.[7–10] Among them, a theoretical approach termed additional probe-beam-induced photon depletion (APIPD)[11] would make CARS microscopy extremely powerful, because not only nanoscale resolution, but also broadband CARS spectral signals of the interested molecules can be achieved simultaneously. However, APIPD method increases the resolution by re-engineering the point spread function (PSF) of the system, which implies that the useful signal photons correspondingly decrease with the squeezed focal spot size decreasing. In other words, resolution enhancement is at the expense of sensitivity. Besides, when CARS is used in biology, the interested molecules are usually at a low concentration, which makes the signal detection more difficult. For the low concentration samples, the minimum concentration of an analyte that can be identified, defined as the detection limit (DL), is a crucial issue to be considered before the implementation of CARS nanoscopy. To our knowledge, the analysis of DL for CARS nanoscopy has not yet been discussed in publications, and so far, the existing studies mainly aim at analyzing the sensitivity about CARS and stimulated Raman scattering (SRS) for comparing their sensitivity disparity.[12–14] In this paper, we analyze the detailed shot-noise limited signal-to-noise (SNR) and several parameters relevant to evaluating the DL. We believe that this analysis will be essential to the current development of CARS nanoscopy.

In this work, we first analyze the shot-noise-limited SNR for time-resolved CARS (T-CARS) nanoscopy based on the proposed APIPD method, but signal strength in the suppressed focal volume is still unknown. The semiclassical theory can provide a concrete expression of the signal strength but fails to define the intensity domains of the pump and Stokes beams. Therefore, we turn to analyze the T-CARS process with full quantum theory, which not only causes the CARS process

*Project supported by the National Basic Research Program of China (Grant No. 2012CB825802), the Major Scientific Instruments Equipment Development of China (Grant No. 2012YQ15009203), the National Natural Science Foundation of China (Grant Nos. 60878053 and 11004136), and the State Key Laboratory of Precision Measurement Technology and Instruments, Tsinghua University, China (Grant No. DL12-01).

†Corresponding author. E-mail: dannyc007@163.com

‡Corresponding author. E-mail: hbniu@szu.edu.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd

to be remarkably simplified, but also successfully defines the domains of the excitation intensities. Saturation of coherent phonons imposes restrictions on the intensities of the pump and Stokes beams, and the further increased intensities will not generate any more coherent phonons. Photo-damage threshold for living cells determines the total intensities of excitation beams injected on the sample. In addition, the intensity relation between the two probe beams determines the attainable lateral resolution in APIPD. Besides the intensities of the excitation beams, shot-noise-limited SNR is mainly determined by the exposure time, the suppressed focal volume, and the concentration of the sample. Under the above conditions, DL of T-CARS nanoscopy correspondingly varies with the exposure time by taking an attainable resolution of about 40 nm in three dimensions, for example. Theoretically speaking, longer exposure time can progressively increase the shot-noise-limited SNR. Nevertheless, longer exposure time leads to the face that the SNR will not be shot-noise-limited, and application of CARS in imaging biological samples restricts the employed exposure time to a level of ten milliseconds, thus the DL in such a nano-voxel will be thousands of target molecules.[15] If number of the interested molecules exceeds this limit, the signal can be well detected from the noise fluctuation and the construction of T-CARS nanoscopy will be certain.

2. Shot-noise-limited SNR of T-CARS nanoscopy

CARS is not background-free. The inherent non-resonance severely restricts the detection sensitivity and complicates spectral assignment. To suppress the non-resonant background signal, a time-resolved scheme can be employed.[16–19] Without the non-resonant background, and omitting the intensity noise of the laser pulses, the SNR of the T-CARS nanoscopy is closely related to various noise contributions, including shot noise, dark current noise, read-out noise, etc., and expressed as

$$\mathrm{SNR_e} = \frac{n_\mathrm{e}}{\sqrt{\sigma_\mathrm{shot}^2 + \sigma_\mathrm{dark}^2 + \sigma_\mathrm{readout}^2}}. \tag{1}$$

Both dark current and read noise follow Poisson statistical distribution as

$$\sigma_\mathrm{dark} = G \cdot F \cdot \sqrt{I_\mathrm{d} \Delta t}, \tag{2}$$

$$\sigma_\mathrm{shot} = G \cdot F \cdot \sqrt{n_\mathrm{e}}. \tag{3}$$

Here, G is the gain coefficient of the EMCCD, F is the noise factor, I_d is the dark current, and Δt is the integration time. The anti-Stokes photons generated from the suppressed focal volume hit the detector's photoreactive surface that produces photoelectrons. The number of photoelectrons n_e, is given by $n_\mathrm{e} = n_\mathrm{sr} \cdot Q_\mathrm{E} \cdot G$, where n_sr is the useful signal's photon number generated from the suppressed focal volume and Q_E is the detector's quantum efficiency. As a result, the SNR, ratio between the signal strength and the background noise, can be written in the following form:

$$\mathrm{SNR_e} = \frac{n_\mathrm{sr} \cdot Q_\mathrm{E}}{\sqrt{F^2 \cdot I_\mathrm{d} \Delta t + F^2 \cdot n_\mathrm{sr} \cdot Q_\mathrm{E} + \left(\frac{\sigma_\mathrm{readout}}{G}\right)^2}}. \tag{4}$$

High-performance detectors are usually cooled to a temperature at which dark current noise σ_dark as well as read noise σ_readout is negligible. Therefore, for the non-resonant background-free T-CARS system, the only remaining noise source is shot noise σ_shot, which follows the Poisson statistical distribution and is equivalent to square-root of the signal strength

$$\mathrm{SNR_e} \approx \frac{1}{F} \sqrt{n_\mathrm{sr} \cdot Q_\mathrm{E}}. \tag{5}$$

Here, the shot-noise limit is reached in the ideal situation in which the other various noise sources are omitted and only the inherent shot noise remains. Consequently, n_noise approximately equals the shot noise of the system and scales as $\sqrt{n_\mathrm{e}}$. $\mathrm{SNR_e}$ is the shot-noise-limited SNR, which can be used to estimate the detective ability of a T-CARS nanoscopy. Usually, CARS signal lies in the visible region, where detectors with Q_E greater than 90% are available.[20]

In order to achieve a concrete expression of $\mathrm{SNR_e}$ and the involved parameters, expression of the anti-Stokes signal strength in the suppressed focus is desired. The semiclassical theory treats the coupling of quantum-mechanical atoms with the classical electromagnetic field. In a semiclassical picture of CARS, the response originates from the third order nonlinear susceptibility, which is the sum of a resonant contribution and a non-resonant electronic component. The detected signal strength, given by $I_\mathrm{as} \propto |\chi^{(3)}|^2 I_\mathrm{L} I_\mathrm{S} I_\mathrm{P}$, is proportional to the square modulus of the total nonlinear susceptibility $\chi^{(3)}$ and depends nonlinearly on incident intensity. There are no restrictions on the adopted laser intensities and three excitation beams play the same role in the signal generation process. So the signal strength, as well as the sensitivity can infinitely increase with the growth of laser intensities. In fact, the signal strength of CARS is at a higher laser power density level. Excessively high power will not contribute to the effective signal generation any more but introduce photo-damage to the sample. The semiclassical theory fails to define the intensity domains of the three excitation beams. Accordingly, in order to circumvent the above shortcomings of the semiclassical theory description, the full quantum mechanism model, in which both the radiation field and the molecular system are quantum-quantized, is provided in T-CARS. It can cause T-CARS process to have a remarkably simplified result and define the intensity domains of the pump and Stokes beams at the same

time. It must be pointed out that the full quantum mechanism description is effective only in T-CARS, which can be divided into two individual steps, and is not suitable for the other four-wave mixing processes.

In the following section, we analyze the T-CARS process under the full quantum picture to obtain a detailed expression of $\mathrm{SNR_e}$ and several parameters in it.

3. Full quantum description of T-CARS

The full quantum picture of T-CARS process is depicted in Fig. 1(a). $|e\rangle$ and $|g\rangle$ correspond to the vibration excited state and ground state respectively. For simplicity of analysis, a single frequency model is provided. The laser fields-the pump, Stokes, probe, and phonons (or molecules) are all quantized. The whole CARS process can be divided into two individual steps. In the first step, a coherent phonon population is directly created by coherent Raman scattering when a pump beam at frequency ω_{L} and a Stokes beam at frequency ω_{S}, with their frequency difference $\omega_{\mathrm{L}} - \omega_{\mathrm{S}}$ tuned to a Raman-active molecular vibration ω_{v} arrive at the sample simultaneously via a high numerical aperture objective. The process induces impulsively stimulated Raman scattering (ISRS), which is considered as the main mechanism of coherent phonon generation.[21,22] The coherent phonons are bosons and do not obey an exclusion principle. They are in a non-equilibrium state and in phase. In the second step, the coherent phonons interact with a time-delayed probe beam at frequency ω_{P} to generate a blue-shift anti-Stokes signal at $\omega_{\mathrm{AS}} = \omega_{\mathrm{P}} + \omega_{v}$, leading to the separation of the CARS signal from the incident laser beams conveniently and efficiently rejecting fluorescence. Delay time between coherent phonons and probe beam is certainly much shorter than the vibrational dephasing time.

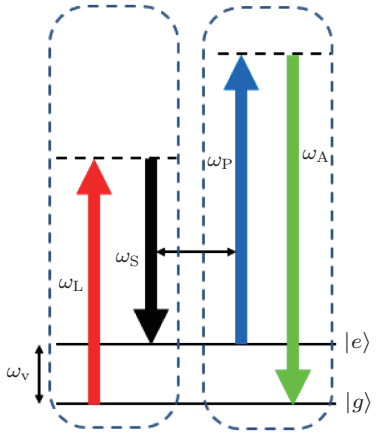


Fig. 1. (color online) Energy level diagram of the two individual steps in time-resolved coherent anti-Stokes Raman scattering (T-CARS). Left, the first step, two excitation beams at frequency ω_{L} and ω_{S} interacting with the molecules. A great number of coherent phonons are generated when $\omega_{\mathrm{L}} - \omega_{\mathrm{S}}$ coincides with the frequency difference between the molecular vibrational levels $|e\rangle$ and $|g\rangle$. Right, the second step, coherent phonons interacting with the time-delayed probe photons, resulting in a strong directional anti-Stokes signal.

Considering a molecule with just a single vibrational transition, the coherent phonon number per time unit in the first ISRS step can be written as [23]

$$n_{\mathrm{v}} = \frac{I_{\mathrm{S}}\mu_{\mathrm{S}}V}{\hbar\omega_{\mathrm{S}}c} \cdot I_{\mathrm{L}} g l, \tag{6}$$

where n_{v} indicates the coherent phonon number; the interaction volume V, which contains N molecules, is defined by the diffraction-limited focal spot area A_{foc} and the interaction length l parallel to the beam propagation direction; ω_{S} is the frequency of the Stokes light; I_{L} and I_{S}, in the form of Gaussian distribution, are average intensities of the pump and Stokes beams respectively; g is the gain coefficient; μ_{S} denotes the refractive index; $\hbar$ is the Plank's constant. For $g \propto N_0$ (N_0 denotes concentration of the interested molecules), it is apparent that the coherent phonon number is linearly dependent on N_0.

The phonon number is subjected to a saturation effect and will not increase infinitely with the growth of incident light power.[24] The maximal coherent phonon number per pulse is the molecule number in the focal volume. In other words, all of the molecules in the focus are excited to a vibrational state.

$$n_{\mathrm{v}}^{\mathrm{pulse}} = \frac{I_{\mathrm{S}}\mu_{\mathrm{S}}V}{\hbar\omega_{\mathrm{S}}c} \cdot I_{\mathrm{L}} g \tau \leqslant N. \tag{7}$$

Here, $n_{\mathrm{v}}^{\mathrm{pulse}}$ is the coherent phonon number per pulse. The magnitude of the spot is roughly estimated to be 10^{-9} cm^2, corresponding to a focal volume V of 8×10^{-14} cm^3. The gain coefficient is $g \sim N_0 \times 10^{-31}$ cm/W. From Eq. (3), we see that saturation of coherent phonons causes the product of the pump and Stokes intensities not to exceed about 10^{18} $\mathrm{W}^2/\mathrm{cm}^4$. Therefore in the first step of a T-CARS process, the pump power density and Stokes power density at focus are both on the same order of $\sim 10^9$ $\mathrm{W/cm}^2$. In principle, if the pump intensity and Stokes intensity exceed such an extreme value, the further increase of intensity will not contribute to the generation of the coherent phonons any more. In practice, however, energies of the pump laser and Stokes laser may be higher than that theoretical estimation due to the spectral broadening of femtosecond laser pulses.

Hereafter, we analyze the second step in T-CARS process. In the spontaneous Raman anti-Stokes process, the involved lattice vibrations called incoherent phonons are thermally excited with a density function n_{v} given by the Bose–Einstein distribution. In this case, oscillatory phases of the phonons are completely random. Thus, individual oscillatory motions are canceled out by each other. Spontaneous Raman scattering is an incoherent linear process and the incoherent signal is randomly scattered to 4π solid angle, then the collected signal is[25]

$$P_{\mathrm{A}} = N \left(\frac{\partial\sigma}{\partial\Omega}\right)_A n_{\mathrm{v}} I_{\mathrm{L}} \Omega_{\mathrm{collection}}, \tag{8}$$

where P_A represents the anti-Stokes power; I_L is the incident laser intensity; $\Omega_{collection}$ is the effective collection angle, and $(\partial\sigma/\partial\Omega)_A$ is the spontaneous Raman differential scattering cross section. The scattered signal is proportional to the incident laser intensity and the occupation number of the upper Raman level. Unlike the spontaneous anti-Stokes Raman scattering, the second step in T-CARS is via introducing probe photons to interact with the coherent phonons, which are in a non-equilibrium state and in phase. Delay time between coherent phonons and probe beam is certainly much shorter than the vibrational dephasing time. Therefore, the relaxation of the phonons is omitted, that is, all phonons are considered to have chances to interact with the probe photons and the influence of the delay time on the signal strength as well as the SNR can be readily omitted. An important advantage of CARS microscopy over spontaneous Raman scattering microscopy is the fact that CARS signal is emitted into a 10^{-4} sr range compared with spontaneous Raman signal randomly scattered into 4π solid angle, which provides an important information that scattering cross section $\partial\sigma_C$ for the interaction between coherent phonons and probe photons is estimated to be enhanced by about five orders of magnitude.[26] In some sense, coherent process is equivalent to increasing the scattering cross section. The coherent anti-Stokes photons are confined to such a small solid angle that a condenser can almost completely collect the forward propagating signal. For low concentration samples, the distance between adjacent solute molecules is so great that the energy of the coherent phonons cannot effectively transmit from one molecule to another. Accordingly, in the second step, the anti-Stokes signal strength is exclusively proportional to the number of coherent phonons and thus linearly dependent on the concentration of solute molecules. By comparing with incoherent (spherical) spontaneous Raman process, we calculate the number of the detected photons for coherent (directional) CARS in the repetition mode as

$$n_{CARS}^{low} = n_v^{pulse} \cdot n_P \cdot \frac{1}{A_{foc}} \partial\sigma_C \cdot \Delta t f_{rep}. \quad (9)$$

In contrast, for high concentration samples, the coherent phonon energy can freely transmit among solute molecules in the focal volume, therefore in the second step, anti-Stokes signal strength is not simply proportional to the coherent phonon number but also concerns with the molecule number. The overall T-CARS process depends quadratically on concentration, hence the anti-Stokes photon number in the repetition mode is

$$n_{CARS}^{high} = N \cdot n_v^{pulse} \cdot n_P \cdot \frac{1}{A_{foc}} \partial\sigma_C \cdot \Delta t f_{rep}, \quad (10)$$

where n_{CARS}^{low} and n_{CARS}^{high} denote anti-Stokes photon numbers in the low and high concentration samples, respectively; n_P is the photon number of the probe beam per pulse; Δt is exposure time; N is the molecule number in the foci. The three pulsed near-infrared excitation beams are tightly focused with a 1.2-NA water-dipping lens for deeper penetration depth to the spot area A_{foc} In short, for $n_v^{pulse} \propto N_0$, the anti-Stokes signal strength is strictly linearly and quadratically dependent on concentration in the low and high concentration samples respectively, which conflicts with the popular statements of CARS signal, i.e., the anti-Stokes signal strength is consistent to be quadratically dependent on concentration. The linear concentration dependence is especially beneficial to the quantitative analysis of low concentration samples.[27–30] In the cross area between the low concentration and the high concentration, the observed concentration dependence from the relative experiments is neither quadratic nor linear, i.e., $n_{CARS} \propto N^{\alpha}$, where $1 < \alpha < 2$. An experimental proof in Fig. 2 shows the dependence of T-CARS signal strength on the concentration of alcohol in pure water at C–H stretch band (2900 cm^{-1}).[31,32] We record the signal intensity at each volume fraction 20 times to exclude the influence of signal fluctuation. When volume fraction of alcohol gradually increases, concentration dependence transforms from strictly linear to sub-quadratic, until quadratic manner. It is apparent that when volume fraction of alcohol in water is lower than 20%, the concentration dependence is strictly linear. We prepare the aqueous solutions of alcohol with smaller volume fraction intervals to obtain a more accurate fitting result. This experiment result consists with the theoretical analysis, and further different samples will be prepared to verify the universality of the theoretical analysis.

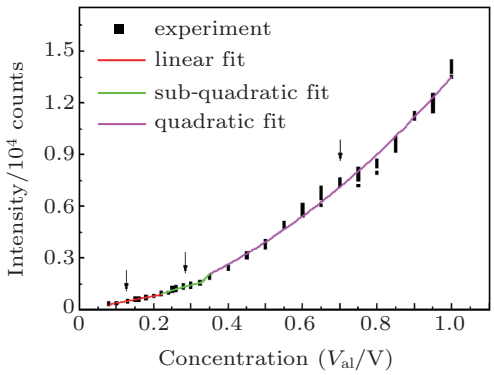


Fig. 2. (color online) T-CARS intensities for increasing volume fractions of alcohol in water at different volume fractions from 7% to 100%.

4. Results and discussion

From Eqs. (1) and (9), shot-noise-limited SNR of T-CARS nanoscopy in a low concentration sample, SNR_e^{low}, is expressed as

$$SNR_e^{low} = \frac{1}{F}\sqrt{n_v^{sr,low} \cdot n_P^{sr} \cdot \frac{1}{A_{sr}} \partial\sigma_C \cdot \Delta t f_{rep} \cdot Q_E}, \quad (11)$$

where $n_{\mathrm{v}}^{\mathrm{sr,low}}$ is the number of coherent phonons per pulse in the suppressed focal volume V_{sr}, which corresponds to a squeezed spot area $A_{\mathrm{sr}} \approx \Delta r^2$ and axial resolution improvement l_{sr} and $n_{\mathrm{P}}^{\mathrm{sr}}$ is the probe photon number in V_{sr}. An approximation is made that $n_{\mathrm{P}}^{\mathrm{sr}}$ is proportional to product of the peak intensity (twice the average intensity) and the squeezed spot area. The starting point of the approximation is the squeezed spot area located in the very center of Airy disk.

To evaluate the shot-noise-limited SNR of T-CARS nanoscopy in the low concentration samples, there are several major parameters that must be taken into account. The first one is the coherent phonon number that is determined by intensities of the pump and Stokes beams. The second is the probe photon number from intensity of the Gaussian probe beam, and the third is the suppressed focal volume. The last one is exposure time Δt. On the basis of the above discussion, we can work out these parameters easily. Theoretically speaking, longer exposure time and higher excitation intensities could increase the CARS signal strength, and hence its SNR. However, as discussed above, the phonon saturation defines the maximum pump intensity and Stokes intensity. Furthermore, photo-damage concerns brings with a practical restriction on how much power can be applied to the sample. The ultimate size of the suppressed focal volume decides the reduced level of the useful anti-Stokes signal and plays an important role in SNR of T-CARS nanoscopy. In our previous work, the APIPD method is proposed to effectively suppress the extension of the focal spot and the full width at half maximum (FWHM) of PSF, denoted by Δr, is worked out as

$$\Delta r = \sqrt{2}\frac{\lambda}{\pi n \sin\alpha\sqrt{3+K}} \approx \frac{0.9}{\sqrt{3+K}} \cdot \frac{\lambda}{2n\sin\alpha}. \qquad (12)$$

Here, the effects of the differences in refractive index and in wavelength among the pump, Stokes, and probe beams are negligible, and the wavelength and refractive index both are uniformly expressed as λ and n. K describes the size of the doughnut probe beam. If K tends to large enough, extent of the focal spot can be substantially reduced, even down to the size of a molecule dimension. Since CARS signal intensity is proportional to the product of the three excitation beam intensities, the lateral resolution is $\sim\sqrt{3}$ times better than that of the linear optical microscopy. In practice, imaging with a focal power of 10 mW from a near-infrared 80-MHz excitation produces excellent CARS signal levels while photo-damage effects are kept at a minimal level for biological samples.[33,34] Taking $K \sim 50$ to obtain a lateral resolution of $\sim$40 nm for example and considering the maximum additional probe beam intensity tolerable by most biological samples, the maximum intensity of the Gaussian probe beam injected on the sample is approximately $\sim 2\times10^{10}$ W/cm^2. The pulse repetition rate f_{rep} is 80 MHz and pulse duration τ is 100 fs. For the attainable spatial resolution of $\sim$ 40 nm in three dimensions, the spot area is roughly $A_{\mathrm{sr}} \sim 10^{-11}$ cm^2 under a tight focus condition, corresponding to a focal volume $V_{\mathrm{sr}} \sim 6.7\times10^{-17}$ cm^3. For the representative C–H stretch vibration around 2900 cm^{-1} in biological molecule, particularly in lipids, the spontaneous Raman differential scattering cross section is 2.5×10^{-29} cm^2/molecule $\cdot$ sr, and then the coherent scattering cross section $\partial\sigma_{\mathrm{C}}$ is estimated as 2.5×10^{-24} cm^2/molecule. Clearly, when intensities of the three excitation beams reach their maxima to guarantee a detectable signal (SNR$_{\mathrm{e}} \geqslant \sqrt{2}Q_{\mathrm{E}} \sim 90\%$) for C–H stretch band at 2900 cm^{-1}, the theoretical DLs of T-CARS nanoscopy with a spatial resolution $\sim$ 40 nm in three dimensions vary with exposure time in different manners correspondingly, as listed in Table 1. Theoretically speaking, longer exposure time can progressively increase the shot-noise-limited SNR. Nevertheless, the dark current noise is proportional to the integration time, and longer exposure time results in the SNR that will not be shot-noise-limited. Meanwhile, much too long exposure time not only induces the photo-damage to sample, but also restricts the ability of CARS to follow the dynamics of live cells. Thus, application of CARS restricts the exposure time to a level of ten milliseconds and the DL will be thousands of target molecules in such a squeezed volume. The signal can be well detected from the noise fluctuation only if the number of the interested molecules exceeds this limit.

Table 1. Relationship between DL with exposure time.

Exposure time/ms	DL
20	$\sim 10^3$
200	$\sim 10^2$

5. Conclusion

In this work, a theoretical analysis of the DL for T-CARS nanoscopy is reported based on our proposed APIPD method. Via analyzing the T-CARS process with full quantum theory, the detailed shot-noise-limited SNR and several major parameters in T-CARS nanoscopy are obtained. While the intensities of the three excitation beams reach their maxima and imaging living cells and organism restricts the exposure time to a level of ten milliseconds, the DL will be thousands of target molecules in a squeezed focal volume with spatial resolution $\sim$ 4 nm in three dimensions. The signal can be well detected from the noise fluctuation only if the number of the molecules of interest exceeds this limit. It is believed that this theoretical work will be essential to the current development of CARS nanoscopy.

References

[1] Evans C L and Xie X S 2008 *Ann. Rev. Anal. Chem.* **1** 883
[2] Yuan J H, Xiao F R, Wang G Y and Xu Z Z 2005 *Chin. Phys.* **14** 935

[3] Li X, Zhang H, Zhang X Y, Zhang S A, Wang Z G and Sun Z R 2008 *Chin. Phys. Lett.* **25** 2062
[4] Hajek K M, Littleton B, Turk D, McIntyre T J and Rubinsztein-Dunlop H 2010 *Opt. Express* **18** 19263
[5] Betzig E, Patterson G H, Sougrat R, Lindwasser O W, Olenych S, Bonifacino J S, Davidson M W, Lippincott-Schwartz J and Hess H F 2006 *Science* **313** 1642
[6] Hell S W and Wichmann J 1994 *Opt. Lett.* **19** 780
[7] Beeker W P, Gro P, Lee C J, Cleff C, Offerhaus H L, Fallnich C, Herek J L and Boller K J 2009 *Opt. Express* **17** 22632
[8] Nikolaenko A, Krishnamachari V V and Potma E O 2009 *Phys. Rev. A* **79** 13823
[9] Beeker W P, Lee C J, Boller K J, Gro P, Cleff C, Fallnich C, Offerhaus H L and Herek J L 2010 *Phys. Rev. A* **81** 12507
[10] Cleff C, Groß P, Fallnich C, Offerhaus H L, Herek J L, Kruse K, Beeker W P, Lee C J and Boller K J 2012 *Phys. Rev. A* **86** 23825
[11] Liu W and Niu H B 2011 *Phys. Rev. A* **83** 23830
[12] Ozeki Y, Dake F, Kajiyama S, Fukui K and Itoh K 2009 *Opt. Express* **17** 3651
[13] Ozeki Y and Itoh K 2010 *Laser Phys.* **20** 1114
[14] Min W, Freudiger C W, Lu S and Xie X S 2011 *Ann. Rev. Phys. Chem.* **62** 507
[15] Potma E O and Xie X S 2003 *J. Raman Spectrosc.* **34** 642
[16] Potma E O, Evans C L and Xie X S 2006 *Opt. Lett.* **31** 241
[17] Jurna M, Korterik J P, Otto C, Herek J L and Offerhaus H L 2008 *Opt. Express* **16** 15863
[18] Cheng J X, Volkmer A, Lewis D and Xie X S 2001 *J. Phys. Chem. B* **105** 1277
[19] Volkmer A, Book L D and Xie X S 2002 *Appl. Phys. Lett.* **80** 1505
[20] Everall N J 2000 *Appl. Spectrosc.* **54** 1515
[21] Vallée F and Bogani F 1991 *Phys. Rev. B.* **43** 12049
[22] Waltner P, Materny A and Kiefer W 2000 *J. Appl. Phys.* **88** 5268
[23] Loudon R 2000 *The Quantum Theory of Light*, 2nd edn. (New York: Oxford University Press) pp. 166–332
[24] Diasty F E 2011 *Vib. Spectrosc.* **55** 1
[25] Portnov A, Rosenwaks S and Bar I 2008 *Appl. Phys. Lett.* **93** 41115
[26] Begley R F, Harvey A B and Byer R L 1974 *Appl. Phys. Lett.* **25** 387
[27] Cui M, Bachler B R and Ogilvie J P 2009 *Opt. Lett.* **34** 773
[28] Zumbusch A, Holtom G R and Xie X S 1999 *Phys. Rev. Lett.* **82** 4142
[29] Bergner G, Chatzipapadopoulos S, Akimov D, Dietzek B, Malsch D, Henkel T, Schlücker S and Popp J 2009 *Small* **5** 2816
[30] Bergner G, Albert C R, Schiller M, Bringmann G, Schirmeister T, Dietzek B, Niebling S, Schlücker S and Popp J 2011 *Analyst* **136** 3686
[31] Begley R F, Harvey A B and Byer R L 2003 *Appl. Phys. Lett.* **25** 387
[32] Zimmerley M, Oertel D C and Ward J L 2009 *J. Biomed. Opt.* **14** 44019
[33] Nan X, Potma E O and Xie X S 2006 *Biochem. J.* **91** 728
[34] König K 2001 *J. Microsc.* **200** 83