The correlation confocal microscope

D.S. Simon[1,2,*] **and A.V. Sergienko**[1,2,3]

[1] *Dept. of Electrical and Computer Engineering, Boston University, 8 Saint Mary's St., Boston, MA 02215, USA*
[2] *Photonics Center, Boston University, 8 Saint Mary's St., Boston, MA 02215, USA*
[3] *Dept. of Physics, Boston University, 590 Commonwealth Ave., Boston, MA 02215, USA*

[*simond@bu.edu](mailto:simond@bu.edu)

Abstract: A new type of confocal microscope is described which makes use of intensity correlations between spatially correlated beams of light. It is shown that this apparatus leads to significantly improved transverse resolution.

© 2010 Optical Society of America

OCIS codes: (180.1790) Confocal microscopy; (180.5810) Scanning microscopy.

References and links

1. M. Minsky, U.S. Patent # 3013467, *Microscopy Apparatus* (1957).
2. M. Minsky, "Memoir on Inventing the Confocal Scanning Microscope," Scanning, **10** 128-138 (1988).
3. R. Pecora, *Dynamic Light Scattering: Applications of Photon Correlation Spectroscopy* (Plenum Press, 1985).
4. K.S. Schmitz, *An Introduction to Dynamic Light Scattering By Macromolecules* (Academic Press, 1990).
5. B.J. Berne and R.J. Pecora, *Dynamic Light Scattering with Applications to Chemistry, Biology, and Physics*, (Dover Publications, 2000). (Reprint of 1976 Wiley edition.)
6. D. Magde, E. Elson, W.W. Webb, "Thermodynamic Fluctuations in a Reacting System-Measurement by Fluorescence Correlation Spectroscopy," Phys. Rev. Lett. **29**, 705–708 (1972).
7. W.W. Webb, "Fluorescence Correlation Spectroscopy: Inception, Biophysical Experimentations, and Prospectus," Appl. Opt. **40**, 3969–3983 (2001).
8. W. Denk, J. Strickler, W.W. Webb, "Two-photon laser scanning fluorescence microscopy," Science **248**, 73–76 (1990).
9. W. Denk, K. Svoboda, "Photon Upmanship: Why Multiphoton Imaging is More than a Gimmick," Neuron **18**, 351–357 (1997).
10. R. Hanbury Brown, R.Q. Twiss, "A Test of a New Type of Stellar Interferometer on Sirius," Nature **178**, 1046–1048 (1956).
11. R. Hanbury Brown, R.Q. Twiss, "Interferometry of the intensity fluctuations in light. I. Basic theory: the correlation between photons in coherent beams of radiation," Proc. Roy. Soc. Lond. A **242**, 300–324 (1957).
12. R. Hanbury Brown, R.Q. Twiss, "Interferometry of the intensity fluctuations in light. II. An experimental test of the theory for partially coherent light," Proc. Roy. Soc. Lond. A **243**, 291–319 (1958).
13. A.f. Abouraddy, K.C. Toussaint, A.V. Sergienko, B.E.A. Saleh, M.C. Teich, "Ellipsometric Measurements Using Photon Pairs Generated by Parametric Down-Conversion," Opt. Lett. **26**, 1717–1719 (2001).
14. A.f. Abouraddy, K.C. Toussaint, A.V. Sergienko, B.E.A. Saleh, M.C. Teich, "Entangled-Photon Ellipsometry," J. Opt. Soc. Am. B **19**, 656–662 (2002).
15. K.C. Toussaint, G.Di Giuseppe, K.J. Bycenski, A.V. Sergienko, B.E.A. Saleh, M.C. Teich, "Quantum Ellipsometry Using Correlated Photon Beams," Phys. Rev. A **70**, 023801 (2004).
16. T.B. Pittman, Y.H. Shih, D.V. Strekalov, A.V. Sergienko, "Optical Imaging by Means of Two-Photon Quantum Entanglement," Phys. Rev. A **52** R3429–R3432 (1995).
17. J.D. Franson, "Nonlocal Cancellation of Dispersion," Phys. Rev. A **45** 3126–3132 (1992).
18. A.M. Steinberg, P.G. Kwiat, R.Y. Chiao, "Dispersion cancellation in a measurement of the single-photon propagation velocity in glass," Phys. Rev. Lett. **68** 2421–2424 (1992).
19. O. Minaeva, C. Bonato, B.E.A. Saleh, D.S. Simon, A.V. Sergienko, "Odd- and Even-Order Dispersion Cancellation in Quantum Interferometry," Phys. Rev. Lett. **102** 100504-1 (2009).
20. C. Bonato, A.V. Sergienko, B.E.A. Saleh, S. Bonora, P. Villoresi, "Even-Order Aberration Cancellation in Quantum Interferometry," Phys. Rev. Lett. **101** 233603 (2008).
21. C. Bonato, D.S. Simon, P. Villoresi, A.V. Sergienko, "Multiparameter Entangled-state Engineering using Adaptive Optics," Phys. Rev. A **79** 062304 (2009).

22. D.S. Simon, A.V. Sergienko, "Spatial-dispersion cancellation in quantum interferometry," Phys. Rev. A **80**, 053813 (2009).
23. R.S. Bennink, S.J. Bentley, R.W. Boyd, "Two-Photon Coincidence Imaging with a Classical Source," Phys. Rev. Lett. **89**, 113601 (2002).
24. R.S. Bennink, S.J. Bentley, R.W. Boyd, and J.C. Howell, "Quantum and Classical Coincidence Imaging," Phys. Rev. Lett., **92**, 033601 (2004).
25. A. Gatti, E. Brambilla, M. Bache, L.A. Lugiato, "Correlated imaging, quantum and classical," Phys. Rev. A **70**, 013802 (2004).
26. Y.J. Cai, S.Y. Zhu, "Ghost imaging with incoherent and partially coherent light radiation," Phys. Rev. E **71**, 056607 (2005).
27. A. Valencia, G. Scarcelli, M. D'Angelo, Y.H. Shih, "Two-Photon Imaging with Thermal Light," Phys. Rev. Lett. **94**, 063601 (2005).
28. G. Scarcelli, V. Berardi, Y.H. Shih, "Can Two-Photon Correlation of Chaotic Light Be Considered as Correlation of Intensity Fluctuations?," Phys. Rev. Lett. **96**, 063602 (2006).
29. F. Ferri, D. Magatti, A. Gatti, M. Bache, E. Brambilla, and L. A. Lugiato, "High-Resolution Ghost Image and Ghost Diffraction Experiments with Thermal Light," Phys. Rev. Lett. **94**, 183602 (2005).
30. D. Zhang, Y.H. Zhai, L.A. Wu, X.H. Chen, "Correlated two-photon imaging with true thermal light," Opt. Lett., **30**, 2354–2356 (2005).
31. A. Muthukrishnan, M.O. Scully, M.S. Zubairy, "Quantum microscopy using photon correlations," J. Opt. B **6** S575–S582 (2004).
32. R.H. Webb, "Confocal Optical Microscopy," Rep. Prog. Phys. **59**, 427–471 (1996).
33. J.C. Mertz, *Introduction to Optical Microscopy* (Roberts and Company Publishers, 2009).
34. P.F. Carcia, R.H. French, M.H. Reilly, M.F. Lemon, D.J. Jones, "Optical superlattices-a strategy for designing phase-shift masks for photolithographys at 248 and 193 nm: Application to AIN/CrN", Appl. Phys. Lett **70**, 2371-2372 (1997)
35. P.F. Carcia, G. Hughes, R.H. French, C.Torardi, G.Reynolds, L. Dieu, "Thin Films for Phase-Shift Masks", Vacuum and Thin Film, Sept. 14-21, 1-10 (1999)
36. K.W.C. Chan, M.N.O. Sullivan, R.W. Boyd, "Optimization of thermal ghost imaging: high-order correlations vs. background subtraction", Optics Express **18**, 5562-5573 (2010).
37. Y. Bai and S. Han, Ghost imaging with thermal light by third-order correlation, Phys. Rev. A. **76**, 043828 (2007).
38. L.-H. Ou and L.-M. Kuang, Ghost imaging with third-order correlated thermal light, J. Phys. B: At. Mol. Opt. Phys. **40**, 18331844 (2007).
39. D.-Z. Cao, J. Xiong, S.-H. Zhang, L.-F. Lin, L. Gao, and K. Wang, Enhancing visibility and resolution in Nthorder intensity correlation of thermal light, Appl. Phys. Lett. **92**, 201102 (2008).
40. I. N. Agafonov, M. V. Chekhova, T. Sh. Iskhakov, and A. N. Penin, High-visibility multiphoton interference of Hanbury Brown-Twiss type for classical light, Phys. Rev. A **77**, 053801 (2008).
41. Q. Liu, X.-H. Chen, K.-H. Luo, W. Wu, and L.-A. Wu, Role of multiphoton bunching in high-order ghost imaging with thermal light sources, Phys. Rev. A **79**, 053844 (2009).
42. K. W. C. Chan, M. N. OSullivan, and R. W. Boyd, High-Order Thermal Ghost Imaging, Opt. Lett. **34**, 33433345 (2009).
43. T.M. Cover, J.A. Thomas, *Elements of Information Theory* (Wiley, 1999).
44. K.D. Mielenz, "Algorithms for Fresnel Diffraction at Rectangular and Circular Apertures," J. Res. Nat. Inst. Stand. Technol., **103**, 497-508 (1997).

---

1. Introduction

The development of the confocal microscope [1, 2] led to a revolution in microscopy. The insertion of source and detection pinholes leads to improved resolution and contrast and the ability to image thin optical sections of a sample in a noninvasive manner. This has made the confocal microscope ubiquitous in modern biomedical optics research, where it is vital not only for imaging, but also for dynamic light scattering ([3–5]), fluorescent correlation spectroscopy ([6, 7]), and other types of experiments. Because all of these experiments rely on achieving the smallest confocal volume (the overlap of the images in the sample of the source and detector pinholes), a great deal of effort has gone into improving the resolution of confocal microscopes in order to minimize this volume. One common approach, which exploits the idea of correlated excitation, is two photon microscopy [8, 9], in which a pair of photons must be absorbed by a fluorescent molecule simultaneously; since this only happens with appreciable probability where the photon density is very high, only the central part of the confocal volume contributes,

leading to a reduced *effective* confocal volume.

Separately, Hanbury Brown and Twiss [10–12] in the 1950's showed that the use of intensity correlations on the detection side could improve the resolution of radio and optical measurements in astronomy. In recent years, similar intensity correlation methods have been used extensively in quantum optics, leading to a number of developments including quantum ellipsometry [13–15], correlated two-photon or ghost imaging [16], and aberration and dispersion cancellation [17–22]. All of these effects were first discovered using quantum entangled photon pairs produced via parametric downconversion, but some have since been reproduced using beams of light with classical spatial correlations; see for example [23–30].

In this paper we investigate the question of whether the advantages of intensity correlation methods at detection and of confocal microscopy can be combined. We will show that use of transverse (lateral) spatial correlations and of coincidence detection can significantly improve the resolution of a confocal microscope. Some care needs to be taken in how this is done since following the naive method of simply sending pairs of spatially correlated photons into a standard confocal microscope will not work; the pinhole destroys all spatial correlations. So instead, we send in uncorrelated photons and use a form of postselection to enforce correlations among the photons we choose to detect.

It is to be noted that the correlation method described here shares in a sense a common underlying philosophy with two-photon microscopy, since the two-photon microscope also uses spatial correlations, but at the *excitation stage*: uncorrelated photons are inserted into the microscope at the source, but the requirement that the two photons interact with the same fluorescent molecule effectively enforces spatial correlations among the photon pairs that contribute to the detected signal. We do something similar here, but use a different method in order to enforce the correlations at the *detection stage*. This will reduce the need for high intensities, thus allowing the use of less powerful lasers, as well as reducing possible damage to the sample.

It should also be pointed out that the idea of using intensity correlations with entangled photon pairs has been been applied before to obtain subwavelength microscopic resolution [31], though not in conjunction with confocal microscopy. This previous work, known as quantum microscopy, required entangled states, in contrast to the method here which will work with a completely classical light source.

The outline of the paper is as follows. In Section 2, confocal microscopy is briefly reviewed. Section 3 discuss the problem of combining confocality with correlation. In Section 4 we introduce the setup for a generalized version of the correlation confocal microscope, and show that it does indeed lead to significant improvement in resolution over the standard (uncorrelated) confocal microscope. In order to make the principles of operation clearer, we will initially consider in Section 4 the unrealistic case of a generalized microscope that requires two identical copies of the object being viewed; in Section 5, we then show how to reduce the apparatus to the realistic case of a single object. Section 6 looks at the axial resolution, with conclusions following in Section 7.

2. The Standard Confocal Microscope.

The basic setup of a standard confocal microscope is shown in Fig. 1. (For a more detailed review see [32]). The two lenses are identical. In real setups they are in fact usually the *same* lens, with reflection rather than transmission occurring at the sample. (In this paper we will for simplicity always draw the transmission case, but most of the considerations will apply equally to the reflection case.) This lens serves as the objective; it has focal length f and radius a, and serves to focus the light going in and out of the sample. The sample is represented at point $\mathbf{y}$ by a function $t(\mathbf{y})$; depending on the setup, $t(\mathbf{y})$ will represent either the transmittance or reflectance
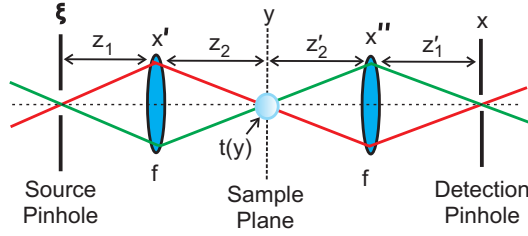
Fig. 1. (Color online) Schematic diagram for standard confocal microscope.

of the sample. At the first lens, the distances are chosen so that the imaging condition

$$\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f} \tag{1}$$

is satisfied; as a result, light entering the microscope through the source pinhole is focused to a small diffraction-limited disk (actually a three-dimensional ellipsoid) centered at a point P in the sample. Any stray light not focused to this point is blocked by the pinhole, thus providing the first improvement in contrast between between P and neighboring points. The distances at the second lens also satisfy the imaging condition, so the second lens performs the inverse of the operation carried out by the first one, mapping the diffraction disk in the sample back to a point at the detection plane. The pinhole in this plane blocks any light not coming from the immediate vicinity of P, thus providing further contrast. Together, the two pinholes serve to pass light from a small in-focus region in the sample and to block light from out-of-focus regions. The in-focus point is then scanned over the sample. The end result is a significant improvement in contrast over the widefield microscope.

The double passage through the lens also leads to improved resolution. To quantify the resolution improvement, we need to look at the impulse response function $h(\mathbf{y})$ and transverse point-spread function (PSF) of the microscope. Let $h_i(\xi, y)$ $(i = 1, 2)$ be the impulse response functions for the first and second lenses individually (including the free space propagation before and after the lens). Up to multiplication by overall constants, these are of the form

$$h_i(\xi, \mathbf{y}) = e^{\frac{ik}{2}\left(\frac{y^2}{z_2} + \frac{\xi^2}{z_1}\right)} \tilde{p}\left(k\left(\frac{\mathbf{y}}{z_2} + \frac{\xi}{z_1}\right)\right), \tag{2}$$

where $\tilde{p}(\mathbf{q})$ is the Fourier transform of the aperture function $p(\mathbf{x}')$ of the lens. $\mathbf{q}$ and k respectively denote the transverse and longitudinal momenta of the incoming photon. We assume that $q << k$. From now on, we will also assume a circular, abberation-free lens of radius a, in which case:

$$\tilde{p}(\mathbf{q}) = 2\pi a^2 \frac{J_1(aq)}{(aq)}, \tag{3}$$

where q is the magnitude of $\mathbf{q}$ and J_1 is the Bessel function of first order. Applying a pinhole at one end, we may also define

$$h_i(\mathbf{y}) \equiv h_i(0, \mathbf{y}) = e^{\frac{iky^2}{2z_2}} \tilde{p}\left(\frac{ky}{z_2}\right) \tag{4}$$

Imagine that a sample is being scanned by the microscope. The amplitude impulse response for the microscope while focused at sample point $\mathbf{y}$ is

$$h(\mathbf{y}) = \int d^2y'\, h_1(\xi, \mathbf{y}')t(\mathbf{y}' + \mathbf{y})h_2(\mathbf{y}', \mathbf{x})\big|_{\mathbf{x}=\xi=0} \tag{5}$$

$$= \int d^2y' \, h_1(\mathbf{y}')t(\mathbf{y}'+\mathbf{y})h_2(\mathbf{y}'). \tag{6}$$

If we insert a sample which is nontransmitting except at a single point, $t(\mathbf{y}) = \delta(\mathbf{y})$, the impulse response becomes the coherent spread function

$$CSF(\mathbf{y}) = h_1(-\mathbf{y})h_2(-\mathbf{y}), \tag{7}$$

with the corresponding point-spread function given by

$$PSF(\mathbf{y}) = |h_1(-\mathbf{y})h_2(-\mathbf{y})|^2. \tag{8}$$

Setting the distances at the first and second lenses to be equal for simplicity ($z_i' = z_i$ for $i = 1, 2$), use of Eq. (3) gives us ([32, 33]):

$$PSF(\mathbf{y}) \propto \left[\tilde{p}\left(\frac{k\mathbf{y}}{z_2}\right) \right]^4 \propto \left(\frac{J_1\left(\frac{kay}{z_2}\right)}{(kay/z_2)} \right)^4, \tag{9}$$

to be compared with the widefield microscope, which has

$$PSF_{wide}(\mathbf{y}) \propto \left[\tilde{p}\left(\frac{k\mathbf{y}}{z_2}\right) \right]^2 \propto \left(\frac{J_1\left(\frac{kay}{z_2}\right)}{(kay/z_2)} \right)^2. \tag{10}$$

Since $\tilde{p}$ is a sharply-peaked function, the higher power in the confocal result leads to a further increase of sharpness and a resulting improvement in resolution.

3. Correlations versus Confocality.

We now wish to introduce spatial correlations into the confocal microscope. However, we immediately run into a problem: the source pinhole is ideally a delta function in position space, which means that its Fourier transform is a constant in momentum space. So, regardless of the spectrum of transverse spatial momentum $\mathbf{q}$ entering the pinhole, the spectrum leaving the pinhole is approximately flat; all transverse momentum correlations are lost. Thus it seems that we must choose between keeping either the correlations or the source pinhole, but not both.

However, the problem may easily be avoided in several ways. One category of solution involves removing the source pinhole and preserving the confocality by other means. Notice that the only purpose the source pinhole serves is to make sure that all of the light entering the microscope is focused at the same point in the object plane. But this can be achieved without a pinhole and therefore without destroying spatial correlations. This can be done in several ways:

(1) One method, often used in the standard confocal microscope, is to have the beam hit the lens parallel to the axis and focus to a point one focal length f away [Fig. 2(a)]. We can then introduce correlations by arranging for pairs of narrow, well-localized beams of light to strike the lens at equal distances from the axis.

(2) A second method is to use pairs of photons produced by spontaneous parametric down-conversion at correlated angles, so that if they are traced backward they both seem to emanate from the same point [Fig. 2(b)]; that point would be then be analogous to the pinhole. The crystal would have to be far enough from the lens for the cross-section of the pump beam to appear to be pointlike.

(3) A third method is to have a very narrow beam reflecting from a fixed point on a rotating mirror, then pass through a beam splitter. The two beams then have anticorrelated directions,
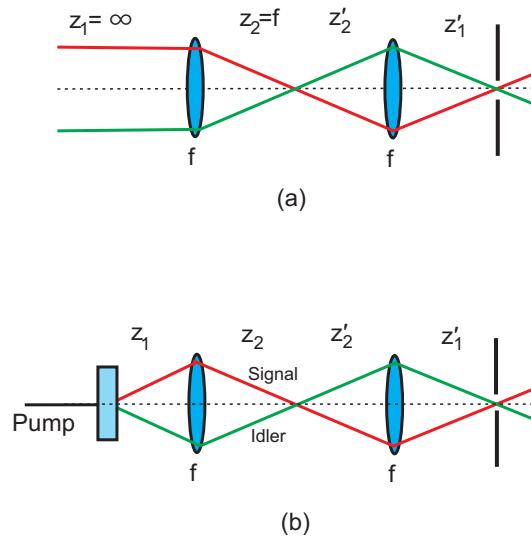
Fig. 2. (Color online) Two methods for making a confocal microscope without a source pinhole.

and both trace back to the same illuminated point on the mirror. The mirror, as it rotates, fills the entire lens aperture with light over time.

The latter two methods are used to enforce spatial correlations in quantum [16] and classical [23] versions, respectively, of correlated-photon or "ghost" imaging. In all three methods, the light would be focused at a point satisfying $\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f}$. (In method 1, we would have $z_1 = \infty$ and $z_2 = f$.) In all of these setups we can think of the arrangement as providing an *effective* or *virtual pinhole*, even though there is no pinhole physically present. In each case, we may proceed formally in the same manner that we would if there was an actual source pinhole.

There is a problem with these approaches, though. Although correlation is maintained up to the location of the first lens, diffractive effects at the lens then destroy a portion of the correlation before the sample is reached. Taking into account the possibility of aberration in the lens as well, the benefits deriving from these methods seem to be limited.

However a completely different class of solution seems more promising: we can take the light to be initially uncorrelated when it enters the microscope, but then select out pairs of photons which happen to be at the same spatial distance from the axis, thus reducing the likelihood of detecting pairs which are at different transverse distances. This effectively enforces correlation at the location of the sample, where it is needed.

In order to enforce the correlation, we need to tag the photons in some manner according to where they happen to intersect the sample, selecting the value of some property of the photon in a manner that depends on position in the sample plane. One way to do this would be to construct filters with narrow, position-dependent pass bands; then, if broadband light is used for illumination, the light leaving the sample will be segregated, with photons of different frequency leaving different points of the sample. Thus, the frequency of the detected light provides an indicator of position and can be used to enforce spatial correlation.

This frequency-tagging method will be discussed in more detail elsewhere. In the current paper, we instead discuss in detail an apparatus in which we tag our photons, not by frequency, but by phase. The two beams are then combined at a beam splitter, converting the phase-tagging into correlation. Effectively, we have two confocal microscopes in the two branches of a Mach-
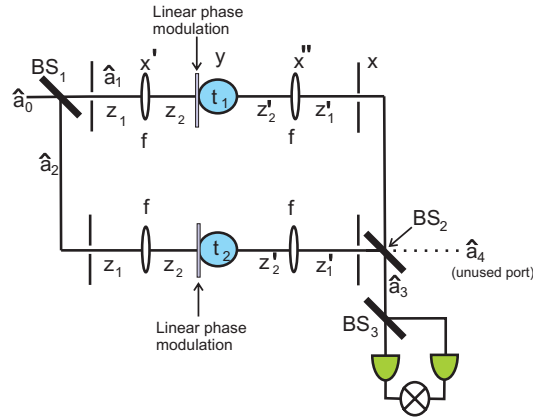
Fig. 3. (Color online) Schematic diagram for correlation confocal microscope with two branches, comprised of two confocal microscopes whose output is combined in a correlated manner.

Zehnder interferometer; the interferometer serves to compare the photon phases and to suppress pairs that differ by a large phase. A pair of detectors and a coincidence counter then count the number of pairs that survive this comparison.

Note that interferometry is not an essential element of the basic correlation confocal scheme, but only of the particular embodiment of it that we discuss in the following sections; it appears only because we chose to tag the photons by phase. If we use frequency-tagging instead, no interferometry would be required. The frequency scheme could be implemented, for example, by spatially separating different frequencies at the output with a diffraction grating; then a coincidence count would occur when the pixels at the same relative position are simultaneously triggered in two CCD cameras.

In the next section, we introduce correlations into a confocal microscope via matching of phases and then examine the use of the resulting correlation confocal microscope to scan over a sample.

4. The Correlation Confocal Microscope.

Here we introduce the correlation confocal microscope. To make the operating principle clearer, we initially work out the formulas for an "unfolded" version of the microscope in which the apparatus has two branches and two identical objects, one in each arm; later we show how to reduce this apparatus to the more useful version with a single arm and single sample. The basic setup is shown in Fig. 3. The locations in the transverse planes at the first lens, second lens, and image plane are represented respectively by $\mathbf{x}'$, $\mathbf{x}''$, and $\mathbf{x}$. Note that $\mathbf{x}'$ and $\mathbf{q}$ are related by $\frac{\mathbf{q}}{k} = \frac{\mathbf{x}'}{z_2}$. Subscripts 1 and 2 on $\mathbf{q}$, $\mathbf{x}'$, or any variable other than z will denote which branch (upper or lower) of the apparatus is being referenced. t_1 and t_2 represent the effect of the samples S_1 or S_2 in the two arms. In the object plane, $\mathbf{y}'$ will denote an arbitrary point in the plane, and $\mathbf{y}$ will be the point actually being viewed in the object; in other words, $-\mathbf{y}$ is the displacement vector of the object during the scan. (The minus sign appears in our convention in order to keep the image from being inverted; alternatively, $+\mathbf{y}$ could be thought of the displacement of the *microscope*, with the sample is held fixed.)

The arguments $\mathbf{y}_1$ and $\mathbf{y}_2$ of the sample functions t_1 and t_2 will be partially linked by the spatial correlation that we will impose; we will find that this leads to an improvement of the lateral or transverse resolution.

In front of the sample in each branch, introduce a position-dependent phase shift

$$e^{i\phi_1(\mathbf{y})} = e^{i(c-\mathbf{y}\cdot\mathbf{b})}, \qquad e^{i\phi_2(\mathbf{y})} = e^{-i(c-\mathbf{y}\cdot\mathbf{b})}, \tag{11}$$

for some constant c and vector field $\mathbf{b}(\mathbf{y})$. We assume that $\mathbf{b}(\mathbf{y})$ is a radial vector of constant magnitude pointing out from the axis, so that the $\phi_i(\mathbf{y})$ depend only on radial distance $y = |\mathbf{y}|$ and are linear in y. The necessary phase shifts may be produced by a graded index material; for a material $.1\ mm$ to $1\ mm$ thick, the gradient of index versus radial distance required is of the order of $\frac{\Delta n}{\Delta r} \sim 10^{-4} - 10^{-3}\ mm^{-1}$, which is well within the range of what is currently technologically feasible. The refractive index gradient could also be achieved by inserting conically-shaped optical wedges in front of each sample. Alternatively, and probably more practically, a thin film such as those designed for use in phase shifting masks could achieve the same effect with a thickness on the order of 100 nm or less [34,35].

We now go through the operation of this apparatus. Uncorrelated photons are input at the left. Suppose for simplicity that they are all in the same state, represented by the creation operator $\hat{a}_0^\dagger$. The first beam splitter (BS_1) splits the beam between the upper arm (1) or lower arm (2) with equal probabilities. The creation operators in these branches are related to the initial state creation operator by

$$\hat{a}_0^\dagger = \frac{1}{\sqrt{2}} \left(\hat{a}_1^\dagger - i\hat{a}_2^\dagger \right). \tag{12}$$

Given photons in each of the two branches (1 and 2), the creation operators are multiplied by the impulse response functions for passage through the confocal microscope in that branch. The creation operators for the upper ($j = 1$) and lower ($j = 2$) branch just before the second beam splitter are then given by

$$\hat{b}_j^\dagger = \hat{a}_j^\dagger \int \tilde{p}^2 \left(\frac{k\mathbf{y}_j}{z_2} \right) t_j(\mathbf{y}_j + \mathbf{y}) e^{i\phi_j(\mathbf{y}_j)} e^{\frac{ik\mathbf{y}_j^2}{2z_2}} d^2 y_j. \tag{13}$$

So if two photons are input at BS_1, then by using Eqs. (12)–(13) we see that the state incident on the second beam splitter is proportional to

$$\left(\frac{1}{\sqrt{2}} \left(\hat{b}_1^\dagger - i\hat{b}_2^\dagger \right) \right)^2 |0\rangle = \frac{1}{2} \left\{ \left[\int \tilde{p}^2 \left(\frac{k\mathbf{y}_1}{z_2} \right) t_1(\mathbf{y}_1 + \mathbf{y}) e^{i\phi_1(\mathbf{y}_1)} e^{\frac{ik\mathbf{y}_1^2}{2z_2}} d^2 y_1 \right]^2 \hat{a}_1^{\dagger 2} \right. \tag{14}$$

$$- \left[\int \tilde{p}^2 \left(\frac{k\mathbf{y}_2}{z_2} \right) t_2(\mathbf{y}_2 + \mathbf{y}) e^{i\phi_2(\mathbf{y}_2)} e^{\frac{ik\mathbf{y}_2^2}{2z_2}} d^2 y_2 \right]^2 \hat{a}_2^{\dagger 2}$$

$$\left. - 2i\hat{a}_1^\dagger \hat{a}_2^\dagger \int \tilde{p}^2 \left(\frac{k\mathbf{y}_1}{z_2} \right) \tilde{p}^2 \left(\frac{k\mathbf{y}_2}{z_2} \right) t_1(\mathbf{y}_1 + \mathbf{y}) t_1(\mathbf{y}_2 + \mathbf{y}) \, e^{i[\phi_1(\mathbf{y}_1) + \phi_2(\mathbf{y}_2)]} e^{\frac{ik(\mathbf{y}_1^2 + \mathbf{y}_2^2)}{2z_2}} d^2 y_1 d^2 y_2 \right\} |0\rangle.$$

We may then substitute into this result the fact that

$$\hat{a}_1^\dagger = \frac{1}{\sqrt{2}} \left(\hat{a}_3^\dagger - i\hat{a}_4^\dagger \right) \tag{15}$$

$$\hat{a}_2^\dagger = \frac{1}{\sqrt{2}} \left(\hat{a}_4^\dagger - i\hat{a}_3^\dagger \right), \tag{16}$$

where 3 and 4 represent respectively the used and unused output ports of BS_2. Taking the inner product of the resulting expression with the state having two photons leaving BS_2 through port

3, we have the two-photon amplitude in branch 3:

$$A_3(\mathbf{y}) = \int d^2y' d^2y'' \tilde{p}^2\left(\frac{k\mathbf{y}'}{z_2}\right)\tilde{p}^2\left(\frac{k\mathbf{y}''}{z_2}\right)t(\mathbf{y}'+\mathbf{y})t(\mathbf{y}''+\mathbf{y})$$
$$\times e^{\frac{ik(\mathbf{y}'^2+\mathbf{y}''^2)}{2z_2}}\left(e^{i[\phi(\mathbf{y}')+\phi(\mathbf{y}'')]} + e^{-i[\phi(\mathbf{y}')+\phi(\mathbf{y}'')]} - 2e^{i[\phi(\mathbf{y}')-\phi(\mathbf{y}'')]}\right). \quad (17)$$

In the last expression, we have set $\phi_1 \equiv \phi$, $\phi_2 \equiv -\phi$, and assumed that the two objects or samples are identical ($t_1 = t_2 \equiv t$). The amplitude may be put into the form (up to an overall multiplicative constant):

$$A_3(\mathbf{y}) = \int d^2y' d^2y'' \tilde{p}^2\left(\frac{k\mathbf{y}'}{z_2}\right)\tilde{p}^2\left(\frac{k\mathbf{y}''}{z_2}\right)t(\mathbf{y}'+\mathbf{y})t(\mathbf{y}''+\mathbf{y})e^{\frac{ik(\mathbf{y}'^2+\mathbf{y}''^2)}{2z_2}}$$
$$\times \left[\cos\left(\phi\left(\mathbf{y}'\right)+\phi\left(\mathbf{y}''\right)\right) - \cos\left(\phi\left(\mathbf{y}'\right)-\phi\left(\mathbf{y}''\right)\right) - i\sin\left(\phi\left(\mathbf{y}'\right)-\phi\left(\mathbf{y}''\right)\right)\right] \quad (18)$$

Note that the expression in the square brackets equals -1 when $\phi(\mathbf{y}') = \phi(\mathbf{y}'') = \frac{\pi}{4}$ and vanishes when $\phi(\mathbf{y}') = \phi(\mathbf{y}'') = 0$. So if we arrange for $\phi(\mathbf{y})$ to drop from $\frac{\pi}{4}$ at the axis to zero at the edge of the Airy disk, this expression will strongly suppress values of $\mathbf{y}'$ and $\mathbf{y}''$ that fall far from the axis, near the edge of the Airy disk (i.e. at the first zero of the Airy function). This is the key to the resolution enhancement. (In the notation introduced earlier, this means $c = \frac{\pi}{4}$ and $\mathbf{b} = \frac{\pi}{4R_{airy}}\hat{r}$.) If the light is reflecting from the sample (as opposed to being transmitted through it), then it will pass through the phase modulation twice, so the modulation should be half as large in this case.

When two photons simultaneously emerge from BS_2 into arm 3, the final beam splitter BS_3 simply routes them (50% of the time) to two different detectors, so that a coincidence count may be measured. The result is that, up to overall constants, the coincidence rate will be:

$$R_c(\mathbf{y}) = |A_3(\mathbf{y})|^2, \quad (19)$$

where A_3 is given by Eq. (18). Note that the coincidence rate does not actually give the image of $|t(\mathbf{y})|^2$, as would normally be obtained from a microscope, but rather it gives the image of $|t(\mathbf{y})|^4$. So the square root of the coincidence count must be taken before comparison with the images from other microscopes.

As a special case, we can obtain the PSF by taking $t(\mathbf{y})$ to be a delta function. Keeping in mind the square root mentioned above, this then gives us:

$$PSF(\mathbf{y}) = [1 - \cos(2\phi(y))]\,\tilde{p}^4\left(\frac{k\mathbf{y}}{z_2}\right) = \sin^2(\phi(y))\,\tilde{p}^4\left(\frac{k\mathbf{y}}{z_2}\right), \quad (20)$$

where y is the magnitude of $\mathbf{y}$. We see that the factor $\sin^2(\phi(y))$ is responsible for the improvement of resolution compared to the standard confocal microscope. This factor suppresses the counting of photon pairs with values of $\mathbf{y}'$ and $\mathbf{y}''$ that differ significantly from each other; this, when combined with the factors of $\tilde{p}$, reduces the contribution to the coincidence rate of points in the outer part of the Airy disk. *Note that all of this comes about because the factors involving the trigonometric functions in Eq. (18) provide a coupling between the integration variables* $\mathbf{y}'$ *and* $\mathbf{y}''$. Without this coupling, which amounts to a spatial correlation between the photons detected from the two branches, Eq. (18) factors into two independent integrals, each of which looks like the amplitude of a standard confocal microscope. *Thus, without the correlation, nothing is gained that could not be obtained from simply using a standard confocal microscope and squaring the output intensity.*
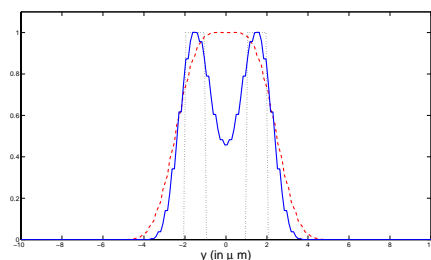
Fig. 4. (Color online) Comparison of images produced by correlated (blue solid) and standard uncorrelated (red dashed) confocal microscopes for two square objects (black dotted) separated by a gap. The correlated version shows a substantially improved ability to distinguish objects with small separation.

We have emphasized the lack of factorability in order to point out what is apparently a fairly common error. The possibility of using third and higher order correlation functions has received much attention recently, and it has been shown both theoretically and experimentally that such higher correlations may be used to improve visibility of interference patterns [36–42]. But on occasion, resolution improvements have been claimed through the use of higher order correlation functions that upon closer examination turn out to factor into products of lower order correlations. Resolution improvements of this sort are illusory. The same effect can be achieved by simply measuring the lower order correlation functions and then multiplying them afterwards. The noise and measurement uncertainties of the individual measurements combine in the product, so that ultimately no benefit is gained; this is consistent with the well-known result of information theory that postprocessing of a measured signal can not increase the amount of information available [43].

The improvement introduced by the induced correlations may be seen from numerical simulations shown in Figs. 4–6. The dotted curves are the transmittance of an object: two square functions separated by a gap in Fig. 4, three square functions in Fig. 5, and a more complicated shape in Fig. 6. The solid curve (blue online) gives the square root of the coincidence rate of the correlation confocal microscope, while the dashed curve (red online) gives the output intensity of a standard confocal microscope. It can be seen clearly in these figures that the correlation microscope produces output that matches the original object significantly more closely than the standard confocal microscope. In particular, in Figs. 4 and 5 the correlation microscope clearly distinguishes the separate square objects and correctly gives their relative heights, whereas the standard confocal microscope simply produces a single blurred peak. Similarly, in Fig. 6 the correlation microscope clearly detects the presence of the "shoulder" on the left-hand curve; the same feature is invisible to the standard confocal microscope. Setting the sample equal to a delta function, we obtain Fig. 7, which shows the transverse intensity PSF. It exhibits a roughly 60% decrease in width, compared to the standard confocal microscope.

Despite the substantial improvement in resolution, one problem with this proposed microscope is also apparent: any aberrations in the lenses or any aberrations induced by the sample will disrupt the phase-matching, and so could reduce the benefits of this scheme. If we chose to use the frequency-tagging scheme mentioned earlier, instead of the phase-tagging version discussed here, a similar degradation could be caused by frequency dispersion in the sample or apparatus. Thus, it may be advantageous to attempt finding a variation of this setup which could incorporate dispersion cancellation ([17–19]) or aberration cancellation ([20–22]) techniques.
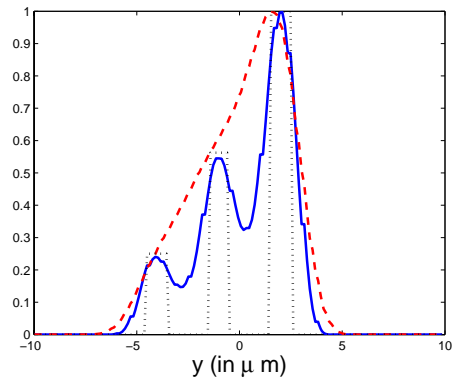
Fig. 5. (Color online) Comparison of images produced by correlated (blue solid) and standard uncorrelated (red dashed) confocal microscopes for three square objects (black dotted) separated by gaps.
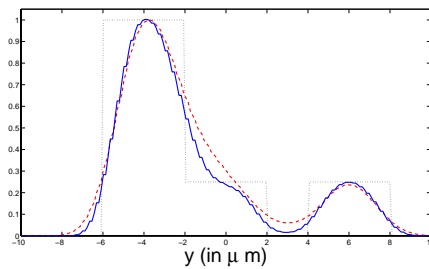

Fig. 6. (Color online) Comparison of images produced by correlated (blue solid) and standard uncorrelated (red dashed) confocal microscopes for a more complicated object (black dotted).
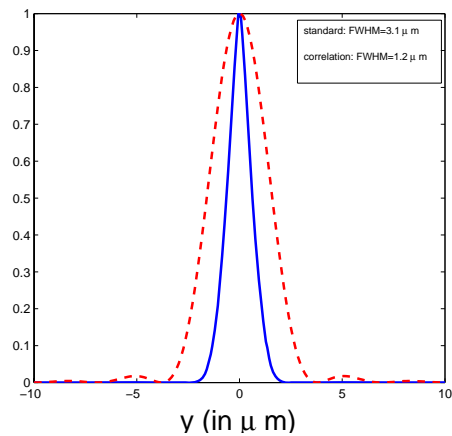

Fig. 7. (Color online) Comparison of transverse intensity point spread functions for standard correlated (blue solid) and uncorrelated (red dashed) confocal microscopes.
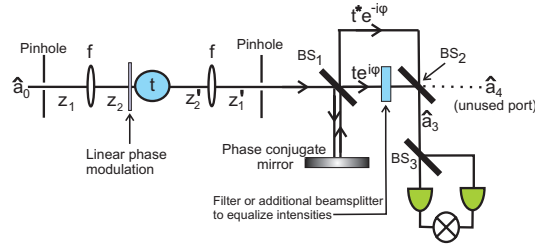
Fig. 8. (Color online) Correlation confocal microscope with a single branch.

5. Reduction to One Sample.

The setup in the previous section is unrealistic in the sense that it requires two identical copies of the object or sample in order to function. Here we go to the realistic case: we show how the apparatus may be altered without affecting the basic principle of operation, in order to require only a single copy of the sample.

The new setup is shown in Fig. 8. All the light now passes through the same sample, gaining a phase shift $e^{+i\phi(\mathbf{y})}$. But then the beam is split at BS_1, after passing through the detection pinhole; half of the beam continues onward to the second beam splitter BS_2 unaltered, while the other half is deflected downward to a phase conjugating mirror. The mirror reverses the sign of the phase and deflects the beam back to BS_1. The half of this beam that is transmitted through BS_1 at this second encounter then recombines with the unaltered beam at BS_2; from this point on, all is as it was for the two-branch version of the previous section, assuming $t(\mathbf{y})$ is real. In the lower branch between BS_1 and BS_2, either a high-density optical filter or another 50/50 beam splitter must be inserted in order to equalize the intensity in the upper and lower branches.

If $t(\mathbf{y})$ is complex, one copy of it will also be complex conjugated at the mirror. The factor $t(\mathbf{y}'+\mathbf{y})t(\mathbf{y}''+\mathbf{y})$ inside the integrals of Eq. (18) now becomes $t(\mathbf{y}'+\mathbf{y})t^*(\mathbf{y}''+\mathbf{y})$; since sample-induced aberrations may be viewed as extra position-dependent phases appearing in t, this change, when combined with the suppression of pairs $(\mathbf{y}',\mathbf{y}'')$ that are far apart, will partially cancel the sample-induced aberrations, but aberrations introduced by the lenses are unaffected.

6. Thick Samples and Axial Resolution.

One of the primary advantages of the standard confocal microscope is its sharp resolution in the axial or longitudinal direction, allowing the ability to optically section materials, i.e. to focus on thin slices in the interior of the material without the view being obscured by the surrounding matter. In this section we investigate the axial resolution of the correlation confocal microscope.

Up until now, we have assumed that the plane in which the phase shift $\phi(\mathbf{y})$ occurs is the same as the object plane; in other words, we have taken the sample to be of negligible thickness and the phase modulation plane to be right up against it. But for the more realistic case of a sample of finite thickness, with the plane to be imaged in the sample's interior, the phase modulation obviously can not be done in the object plane. Instead, the phase modulation plane must be moved a small distance ζ_2 away from the object plane, to the exterior of the sample. The distance from the lens to this plane is $\zeta_1 = z_2 - \zeta_2$ (see Fig. 9). We apply a phase shift $\phi'(\mathbf{x}'') = c - \mathbf{b}' \cdot \mathbf{x}''$ in this plane, which we still assume to vary only in the radial direction. Then the propagation factor from the source pinhole to the object plane [Eq. (4)] can be shown to now be

$$h_1(\xi = 0, \mathbf{y}) = e^{\frac{ik}{2z_2}y^2} e^{i(c-\mathbf{b}''\cdot\mathbf{y})} \tilde{p}\left(\frac{k}{z_2}\left(y + \frac{\zeta_2}{k}b'\right)\right), \qquad (21)$$
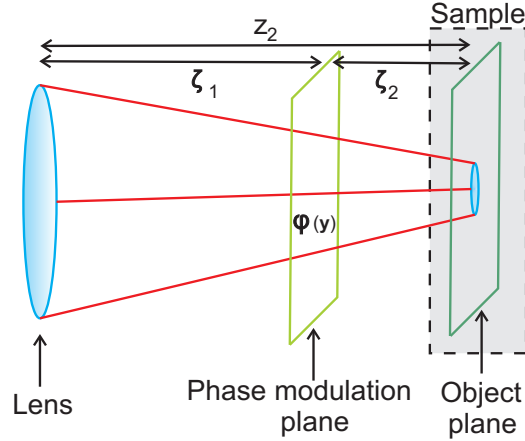
Fig. 9. (Color online) To view thick samples, the phase modulation plane needs to be moved away from the object plane by distance ζ_2.

where $\mathbf{b}'' = \frac{\zeta_1}{z_2}\mathbf{b}'$. Note that this is of the same form as before, except that the phase shift parameter $\mathbf{b}'$ has been altered and the position of each point in the diffraction spot has been shifted in the argument of $\tilde{p}$. Since $\mathbf{b}'$ is a radial vector, each point of the Airy disk moves radially a distance $|\Delta y| = \frac{b'\zeta_2}{k}$, where $b' = |\mathbf{b}'|$. The result is that the Airy disk expands in radius by $|\Delta y|$. This is clearly undesirable, so it is necessary to choose ζ_2 such that $|\Delta y| << R_{airy}$, making the radial shift negligible. Inserting realistic values, this means that we ζ_2 must be at most of order 10^{-4} m. Thus, our ability to view the interiors of thick samples will be limited to depths of less than $100\mu m$.

So, neglecting Δy compared to y, we have

$$h_1(\xi = 0, \mathbf{y}) = e^{\frac{ik}{2z_2}y^2} e^{i(c-\mathbf{b}''\cdot\mathbf{y})} \tilde{p}\left(\frac{k}{z_2}y\right), \tag{22}$$

which is identical to Eq. ((4) if we choose $\mathbf{b}'' = \mathbf{b}$ (or, equivalently, $\mathbf{b}' = \frac{z_2}{\zeta_1}\mathbf{b}$). Here, $\mathbf{b}$ and c are the original parameter values used in the previous sections.

Suppose we view a point on the axis ($\mathbf{y} = 0$) at a distance $z = z_2 + \delta z$ from the lens. z_2 is the distance to the confocal plane ($\frac{1}{z_1} + \frac{1}{z_2} = \frac{1}{f}$), so that δz is the defocusing distance. We assume that $|\delta z| << z_2$ and study how the axial point spread function varies as a function of defocus. It is well known [32] that for the standard confocal microscope the axial PSF is of the form

$$PSF_{standard}(\delta z) = \text{sinc}^4(\kappa \delta z), \tag{23}$$

where $\kappa = \frac{ka^2}{4z_2^2}$ and $\text{sinc}(x) = \sin(x)/x$.

In the case of the correlation confocal microscope, each measurement involves two photons, which may interact with the sample at different points, so there will be two defocus variables, δz_a and δz_b, one for each photon. Using Eq. (22) in place of Eq. (4) and allowing for defocus in each arm, a calculation similar to that of Section 4 leads to the amplitude

$$A_3(\mathbf{y}, \delta z_1, \delta_2) = \int d^2y' d^2y'' t(\mathbf{y}'+\mathbf{y}) t(\mathbf{y}''+\mathbf{y}) e^{\frac{ik(y'^2+y''^2)}{2z_2}} \tag{24}$$

$$\times E(-4\kappa\delta z_a, -\frac{ka}{2z_2}y') E(+4\kappa\delta z_a, -\frac{ka}{2z_2}y')$$

$$\times E(-4\kappa\delta z_b, -\frac{ka}{2z_2}y'')E(4\kappa\delta z_b, -\frac{ka}{2z_2}y'')$$
$$\times \left[\cos\left(\phi\left(\mathbf{y}'\right)+\phi\left(\mathbf{y}''\right)\right) - \cos\left(\phi\left(\mathbf{y}'\right)-\phi\left(\mathbf{y}''\right)\right) - i\sin\left(\phi\left(\mathbf{y}'\right)-\phi\left(\mathbf{y}''\right)\right)\right].$$

Here, we have defined

$$E(u,v) \equiv [L(u,v) + iM(u,v)] = 2\int_0^1 J_0(v\rho)e^{\frac{i}{2}u\rho^2}\rho d\rho, \tag{25}$$

where $J_0(x)$ is the Bessel function of zeroth order. L and M are known as Lommel functions and can be calculated by means of series expansions [44].

As before, the absolute square of amplitude A_3 gives us the coincidence rate. If the object t is taken to be constant in the z direction and a delta function in the transverse (x and y) directions, then this coincidence rate at $\mathbf{y} = 0$ will give us the axial PSF. Up to an overall normalization constant, we then have

$$
\begin{aligned}
PSF_{axial}(\delta z_a, \delta z_b) &= |E(-4\kappa\delta z_a, 0)E(4\kappa\delta z_a, 0)E(-4\kappa\delta z_b, 0)E(4\kappa\delta z_b, 0)|^2 \\
&= \text{sinc}^4(\kappa\delta z_a)\text{sinc}^4(\kappa\delta z_b).
\end{aligned}
\tag{26}
$$

The axial PSF thus factors into the product of the axial PSFs of the separate standard confocal microscopes in the two arms. So, the phase modulation has no effect on the axial resolution.

7. Conclusions.

In this paper we have shown that by incorporating intensity correlations and coincidence measurements into a generalized version of the confocal microscope, it should be possible to achieve large improvements in resolution in the transverse direction. This leads to a great improvement in imaging ability.

There are a number of variations of this scheme that may be investigated, some of which may lead to further improvements. For example, we have assumed that the phase shifts at the sample vary linearly with radial distance; instead, a nonlinear radial dependence could be used to provide further suppression of photon pairs that fall far from the center of the Airy disk. Another variation is that instead of looking for pairs of photons both leaving BS_2 by the same output port (port 3), we could look for coincidences *between* the two possible output ports (one photon each in ports 3 and 4). This would simplify the apparatus; however it can be shown that the resolution in this variation is not as good as the one we have described in the earlier sections, although it is still better than the standard confocal microscope. To be specific, the improvement factor of $[1 - \cos(2\phi(y))] = \sin^2(\phi(y))$ in Eq. (20) would be replaced by $\sin(2\phi(y))$, which decreases more slowly away from the axis for the same choice of phase function.

It is conceivable that for some types of special applications it may be desirable to tailor the shape of the resolvable spot to some other form. The specific implementation that we have described in this paper uses a phase modulation that is radially symmetric, and the result is a radially symmetric focused spot. However, a phase shift that is constant along one axis and linearly varying along the perpendicular axis could be used to induce improved resolution along a particular direction. Similarly, in addition to using a phase shift to improve the radial resolution, the simultaneous use of frequency matching in the tangential direction could give angular resolution *within* the focused spot.

Some disadvantages of the correlation confocal microscope compared to the standard confocal microscope should be mentioned. First, although the improved transverse resolution leads to a smaller confocal volume, the correlation microscope is no better than a standard confocal microscope in fluorescence correlation spectroscopy (FCS) and dynamic light scattering (DLS)

experiments; this is because the random phase shifts from the scattering or fluorescence events will destroy the phase correlations. So for these experiments, the extra complexity involved in the correlation brings no benefits.

Second, most of the photons entering the correlation confocal microscope will be weeded out by the phase-correlation requirement and will not contribute to the image. In this sense, the efficiency of the device is low, and image formation will certainly be slower than with a standard confocal microscope. However, the substantial improvements in resolution should certainly outweigh the loss of speed in many applications.

Acknowledgments