LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-639499

Signal, noise and resolution in linear and nonlinear structured-illumination microscopy

E. A. Ingerman, R. A. London, M. G. L. Gustafsson

June 19, 2013

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Signal, noise, and resolution in linear and nonlinear structured-illumination microscopy

E. A. Ingerman[¶], R. A. London[†], and M. G. L. Gustafsson[+]

[¶] *Center for Biophotonics Science and Technology, University of California Davis, Davis, California, 95616.*
Present affiliation: Life Technologies, Inc., 180 Oyster Point Blvd., South San Francisco, CA 94080.
[†]*Lawrence Livermore National Laboratory, Livermore, CA 94551, USA.*
[+] *Department of Physiology and Program in Bioengineering, University of California San Francisco, San Francisco, California, 94143.*
Deceased

Correspondence to: Eugene Ingerman. Tel: +1 415 710 6954
e-mail: eugening.ucd@gmail.com

Key words. Super-resolution, illumination patterns, patterned excitation, fluorescence microscopy, photo-switchable molecules, HELM, SPEM, RESOLFT

Summary

Structured-illumination microscopy allows widefield fluorescence imaging with resolution beyond the classical diffraction limit. Its linear form extends resolution by a factor of two, and its nonlinear form by an in-principle infinite factor, the effective resolution in practice being determined by noise. In this article, we analyze the noise properties and achievable resolution of the method from a frequency-space perspective. We develop an analytical theory for a general case of linear or nonlinear fluorescent imaging, and verify the analytical calculations with numerical simulation for a special case where nonlinearity is produced by photo-switching of fluorescent labels. We compare the performance of two alternative implementations, using either two-dimensional illumination patterns or sequentially rotated one-dimensional patterns. We show that one-dimensional patterns are advantageous in the linear case, and that in the nonlinear case two-dimensional patterns provide a slight signal-to-noise advantage under idealized conditions, but perform worse than one-dimensional patterns in the presence of non-switchable fluorescent background.

Introduction

The diffraction limit to the resolution of the light microscope has been well understood for more than a century (Abbe, 1873), and for most of that time has stood

unchallenged. Recently, several techniques have been demonstrated that can go beyond this limit. These techniques include near-field scanning (Betzig & Trautman, 1992); dual-objective-lens methods such as 4-Pi (Hell & Stelzer, 1992) and I^5M (Gustafsson *et al.*, 1995); highly nonlinear phenomena such as stimulated emission depletion (Dyba & Hell, 2002) and other forms of reversible saturable optical transitions (Hell *et al*., 2003; Hofmann *et al*., 2005); and localization of photoactivated labels (Betzig *et al*., 2006; Hess *et al.* 2006; Bates *et al.* 2006). Structured illumination microscopy (SIM) is one such method, which uses patterns of excitation light to encode otherwise unobservable information into the observed image. It has been used for resolution enhancement both in the axial (Bailey et al., 1994) and the lateral (Heintzmann and Cremer, 1998; Gustafsson *et al*., 2000; Frohn *et al*., 2000; Heintzmann *et al*., 2002; Gustafsson, 2005, Chung *et al*., 2007) directions.

The conventional diffraction limit defines a finite range of spatial frequencies that can be transmitted through a microscope. The limiting spatial frequency is determined by the wavelength of the light and the aperture of the objective lens. Structured illumination allows this limit to be extended by an amount equal to the spatial frequency of the illumination pattern (Gustafsson *et al*., 2000; Heintzmann *et al*., 2002; Gustafsson, 2005). Because the spatial frequencies that can be introduced into the illumination pattern are limited by diffraction in the same way as the observable frequencies, the lateral resolution can at most be increased by about a factor of two. This factor-of-two limit, however, applies only to the normal, linear case where the fluorescent emission rate is directly proportional to the excitation intensity. Much larger increases in resolution are possible if the object response depends nonlinearly on the illumination intensity (Heintzmann *et al.*, 2002; Gustafsson, 2005; Schwentker *et al.*, 2007). If the emission rate per fluorophore is a nonlinear function of the illumination intensity, then the pattern of molecular emission rate will contain harmonics of the spatial frequencies of the illumination pattern, and in particular can contain higher spatial frequencies than can exist in any physical far field illumination intensity. Since there is no limit to the spatial frequency of the harmonics, there is no hard limit to the available resolution. The highest resolution achievable in practice is then limited only by the noise in the observed data.

Possible nonlinear object response mechanisms include those used or proposed for resolution enhancement in point-scanning modes, such as stimulated emission depletion (Hell & Wichmann, 1994) and photo-switchable fluorescent molecules (Hell *et al.* 2003; Hell, 2003; Hofmann *et al.*, 2005; Schwentker *et al.*, 2007), as well as other phenomena that are ill-suited for point scanning, such as saturation of the excited state (Heintzmann *et al*., 2002; Gustafsson, 2005).

A one-dimensional (1D) illumination pattern (parallel lines) only extends resolution in one direction. To extend resolution in 2D, there is a choice of approaches: either the 1D pattern is applied sequentially in a series of angular orientations (Gustafsson, 2000; Gustafsson, 2005), or a more complex 2D illumination pattern is used (Frohn *et al*., 2000; Heintzmann, 2003). It is not immediately obvious which choice is preferable.

The main goal of this paper is to analyze the effect of noise on the resulting images in linear and nonlinear SIM of thin planar objects and derive theoretical limit of resolution in nonlinear SIM. Towards this goal we calculate the signal-to-noise ratio (SNR) as a function of spatial frequency, and compare the performance of 1D and 2D illumination patterns. In the end, we confirm the analytical results through computer simulations.

Image formation and reconstruction

Linear Structured Illumination

In conventional linear fluorescent microscopy, the local emission density $Em(\boldsymbol{r})$ at each point $\boldsymbol{r}$ of an object is proportional to the product of the local intensity of excitation light, $I(\boldsymbol{r})$, and the local fluorophore concentration, $S(\boldsymbol{r})$ (i.e. the object structure to be determined):

$$Em(\boldsymbol{r}) = c_e I(\boldsymbol{r}) S(\boldsymbol{r}), \tag{1}$$

where $S(\boldsymbol{r})$ is measured in molecules/area, $Em(\boldsymbol{r})$ and $I(\boldsymbol{r})$ are measured in photons/(area × time), and the emission constant c_e is the product of the fluorophore's absorption cross section and quantum yield. This linear relation holds for conventional fluorescent markers at normal excitation intensities that do not saturate the fluorescence. The observed image, $D(\boldsymbol{r})$, is proportional to the convolution of the emission intensity with the microscope's point spread function, $H(\boldsymbol{r})$:

$$D(\boldsymbol{r}) = c_d T (H \otimes Em)(\boldsymbol{r}) = c_e c_d T \left[H \otimes (I\,S)\right](\boldsymbol{r}), \tag{2}$$

where $\otimes$ denotes the convolution operation; $D(\boldsymbol{r})$ is measured in photons/area; the detection efficiency c_d is the product of the microscope's collection efficiency and the detector's quantum efficiency; T is the exposure time; and the point spread function is normalized to have a unit integral. However, not all spatial frequencies can be transmitted by the microscope. The region of frequency space where the optical transfer function (OTF), denoted by $H(\boldsymbol{k})$ and defined as the Fourier transform of $H(\boldsymbol{r})$, is non-zero is called the "observable region," or the "OTF support." The observable region for a widefield fluorescence microscope (in two dimensions) is the interior of a circle of radius $k_o = 2NA/\lambda_{em}$, where λ_{em} is the fluorescent emission wavelength and NA is the numerical aperture of the objective (Goodman, 2005).

It is impossible to observe directly any information about those spatial frequencies of the light emission that lie outside the observable region. However, the structure of interest is not the emission $Em(\boldsymbol{r})$ but the object $S(\boldsymbol{r})$, and it is possible to observe higher-spatial-frequency components of the object, by rendering them visible in the form of low-spatial-frequency moiré fringes with the structured illumination technique (Gustafsson, 2000). The mathematical principles of structured-illumination microscopy will be outlined in the remainder of this section.

Consider an illumination intensity pattern that is sinusoidal in the lateral direction:

$$I(\mathbf{r}) = \frac{I_0}{2}\left[1 - \cos(2\pi\ \mathbf{p}\cdot\mathbf{r} + \varphi)\right], \tag{3}$$

where $\boldsymbol{r} = (x, y)$ is the position vector, $\boldsymbol{p}$ is the wave vector of the pattern, and φ is its phase. We can determine which spatial frequencies are observable, by Fourier transforming Eqs. (2) and (3) and applying the convolution theorem:

$$\hat{D}(\boldsymbol{k}) = c_e c_d T\ \ \hat{H}(\boldsymbol{k})\left(\hat{I} \otimes \hat{S}\right)(\boldsymbol{k}) \tag{4}$$

$$\hat{I}(\boldsymbol{k}) = \frac{I_0}{2}\left[\delta(\boldsymbol{k}) - \tfrac{1}{2}\delta(\boldsymbol{k} - \boldsymbol{p})e^{-i\varphi} - \tfrac{1}{2}\delta(\boldsymbol{k} + \boldsymbol{p})e^{i\varphi}\right], \tag{5}$$

where $\delta(k)$ denotes the Dirac delta function, carets (^) indicate the Fourier transforms of the corresponding real-space quantities, and $\boldsymbol{k} = (k_x, k_y)$ is the reciprocal space coordinate vector. Substituting Eq. (5) into Eq. (4), and performing the convolution, we obtain the Fourier transform of the observed image:

$$\hat{D}(\boldsymbol{k}) = c_e c_d T \frac{I_0}{2}\hat{H}(\boldsymbol{k})\left[\hat{S}(\boldsymbol{k}) - \tfrac{1}{2}\hat{S}(\boldsymbol{k} - \boldsymbol{p})e^{-i\varphi} - \tfrac{1}{2}\hat{S}(\boldsymbol{k} + \boldsymbol{p})e^{i\varphi}\right] \tag{6}$$

The second and third terms in Eq. (6) correspond to information that has been shifted in reciprocal space by vectors $\boldsymbol{p}$ and $-\boldsymbol{p}$ respectively.

The presence of the factor $\hat{H}(\boldsymbol{k})$ in Eq. (6) indicates that the observed image, $\hat{D}(\boldsymbol{k})$, in reciprocal space can be nonzero only for points within the support of the OTF $\hat{H}$. As $\boldsymbol{k}$ ranges over that region, the first term of Eq. (6) contributes the normally observable information, but the second and third terms contribute information that originates from two other regions of the same shape as the OTF support, but centered at $-\boldsymbol{p}$ and $\boldsymbol{p}$, respectively, shown as light blue circles in Fig. 1*a*. Part of those regions lie outside the OTF support and thus represent normally unobservable high-frequency information that has been made visible in the form of low-frequency moiré fringes.

It is not possible to separate the contributions from the three terms of Eq. (6) based on only a single image. However, three observations with different values of φ will normally supply three independent versions of Eq. (6), enabling the separation of the three contributions.

By the above process, one can obtain extended-resolution information in one direction (Fig. 1*a*, light blue circles). Nearly isotropic high-resolution information can be obtained by repeating the same process for other orientations of the illumination pattern (Fig. 1*b*), or alternatively by using a two-dimensional illumination pattern and shifting it in two directions (Fig. 1*c*). The new information can be used to produce a numerical reconstruction with a resolution that exceeds the classical diffraction limit by an

amount $|\boldsymbol{p}|$. Thus, to obtain the highest possible resolution, one should use an illumination pattern with the highest possible spatial frequency. Unfortunately, the illumination pattern frequency is subject to the diffraction limit in the same way as the observable frequencies are. If the numerical aperture of illumination is the same as that of observation, the illumination pattern frequency $|\boldsymbol{p}|$ cannot exceed $p_{\max} = 2NA/\lambda_{\text{illum}} \approx k_0$. The new resolution limit is then $k_0 + p_{\max} \approx 2k_0$. Hence, by using linear structured illumination we can at most double the resolution of the conventional wide-field fluorescence microscope.

Nonlinear Structured Illumination

Dramatically greater increases in resolution are possible if the fluorophore emission rate can be made to depend nonlinearly on the illumination intensity (Heintzmann *et al.*, 2002; Gustafsson, 2005; Schwentker *et al.*, 2007):

$$Em(\boldsymbol{r}) = F[I(\boldsymbol{r})]S(\boldsymbol{r}) = G(\boldsymbol{r})S(\boldsymbol{r}), \tag{7}$$

where the function $F[I]$ describes the nonlinear response of the fluorophore emission to the illumination intensity, and $G(\boldsymbol{r}) = F[I(\boldsymbol{r})]$ is the local emission rate per molecule. In general, the nonlinearity F can be expanded in a power series:

$$G(\boldsymbol{r}) = F[I(\boldsymbol{r})] = a_0 + a_1 I(\boldsymbol{r}) + a_2 I^2(\boldsymbol{r}) + a_3 I^3(\boldsymbol{r}) + \cdots \tag{8}$$

If the illumination pattern $I(\boldsymbol{r})$ is sinusoidal with a spatial frequency $\boldsymbol{p}$ [Eq. (3)], the quadratic and higher powers in Eq. (8) will cause $G(\boldsymbol{r})$ to contain harmonics at multiples of $\boldsymbol{p}$:

$$G(\boldsymbol{r}) = b_0 + 2b_1\cos(2\pi\boldsymbol{p}\cdot\boldsymbol{r} + \varphi) + 2b_2\cos(4\pi\boldsymbol{p}\cdot\boldsymbol{r} + 2\varphi) + 2b_3\cos(6\pi\boldsymbol{p}\cdot\boldsymbol{r} + 3\varphi) + \cdots, \tag{9}$$

where the coefficients b_m can be calculated from Eqs. (3) and (8) for any given nonlinearity F. When the phase φ of the illumination pattern is equal to zero, the b_m are equal to the coefficients of the cosine Fourier series of the (periodic and even) function G. If the nonlinearity F is non-polynomial, both the power series (8) and the Fourier series (9) will contain an infinite number of terms.

The Fourier transform of the emission rate pattern [Eq. (9)] is

$$\hat{G}(\boldsymbol{k}) = b_0\delta(\boldsymbol{k}) + \sum_{m=1}^{\infty} b_m\left[\delta(\boldsymbol{k} - m\boldsymbol{p})e^{-im\varphi} + \delta(\boldsymbol{k} + m\boldsymbol{p})e^{im\varphi}\right] = \sum_{m=-\infty}^{\infty} b_m\delta(\boldsymbol{k} - m\boldsymbol{p})e^{-im\varphi}, \tag{10}$$

where, for simplicity of notation, the coefficients have been defined also for negative *m* by $b_{-m} = b_m$. The reciprocal-space image becomes [using Eqs. (10) and (4) modified by Eq. (7)]:

$$\hat{D}(\boldsymbol{k}) = c_d T\, \hat{H}(\boldsymbol{k})\left(\hat{G} \otimes \hat{S}\right)(\boldsymbol{k}) = c_d T\, \hat{H}(\boldsymbol{k}) \sum_{m=-\infty}^{\infty} b_m \hat{S}(\boldsymbol{k} - m\boldsymbol{p}) e^{-im\varphi} \ . \tag{11}$$

The image intensity $\hat{D}(\boldsymbol{k})$ is thus a weighted sum of information contributions from an infinite number of spatial frequencies $(\boldsymbol{k} - m\boldsymbol{p})$. If the observations were noiseless, Eq. (11) predicts that we could obtain infinite resolution. In reality, only a finite number of terms in Eq. (11) rise above the experimental noise. Hence this sum can be truncated to a finite number of terms, $|m| \le N$, where N is the index of the last significant term in the sum.

Given that there are only (2N+1) significant terms, we can separate the contributions of (2N+1) different information components $\hat{S}(\boldsymbol{k} - N\boldsymbol{p}), \cdots, \hat{S}(\boldsymbol{k} - \boldsymbol{p}), \hat{S}(\boldsymbol{k})$, $\hat{S}(\boldsymbol{k} + \boldsymbol{p}), \cdots, \hat{S}(\boldsymbol{k} + N\boldsymbol{p})$ by collecting $(2N+1)$ images with different values φ_j of the phase of the illumination pattern, and solving a system of linear equations that have the form of Eq. (11) except that the sum ranges only from -*N* to *N*. A natural choice for the phases φ_j is a set of equidistant values on the interval $[0, 2\pi]$:

$$\varphi_j = j\frac{2\pi}{2N+1} \ , \quad j = 0, \cdots, 2N \ . \tag{12}$$

With equidistant phases, the linear system Eq. (11) becomes very simple: the sum on the right hand side is the discrete Fourier transform of the sequence $b_{-N}\hat{S}(\boldsymbol{k} - N\boldsymbol{p}), \cdots, b_{-1}\hat{S}(\boldsymbol{k} - \boldsymbol{p}), b_0\hat{S}(\boldsymbol{k}), b_1\hat{S}(\boldsymbol{k} + \boldsymbol{p}), \cdots, b_N\hat{S}(\boldsymbol{k} + N\boldsymbol{p})$. Note that this Fourier transform is performed along the index *j* of the phase value φ, not along any physical dimension. We can therefore separate the different information components of the observed object through an inverse discrete Fourier transform of the $2N+1$ images:

$$b_m \hat{H}(\boldsymbol{k})\hat{S}(\boldsymbol{k} - m\boldsymbol{p}) = \frac{1}{c_d T\,(2N+1)} \sum_{j=-N}^{N} e^{imj\varphi_0} \hat{D}_j(\boldsymbol{k}) \ , \tag{13}$$

where $\varphi_0 = 2\pi/(2N+1)$, and $\hat{D}_j$ is the image observed when illuminating with the pattern phase $\varphi_j = j\varphi_0$. We define the concise notation for the separated information components:

$$\hat{R}_m(\boldsymbol{k} - m\boldsymbol{p}) = b_m \hat{H}(\boldsymbol{k})\hat{S}(\boldsymbol{k} - m\boldsymbol{p}), \qquad m = -N, \cdots, N \tag{14}$$

for later use.

Using one-dimensional illumination patterns, as discussed so far, allows us to obtain high resolution in only one direction, that of the pattern wave vector $\boldsymbol{p}$ (Fig. 1*a*). As in the case of linear SIM, effectively isotropic high resolution can be obtained by rotating the illumination pattern to a series of orientations and repeating the reconstruction process (Fig. 1*d*).

The goal is to determine the object information $\hat{S}(\boldsymbol{k})$, the Fourier transform of the object fluorophore density, at each point $\boldsymbol{k}$ of reciprocal space. Ignoring noise for the moment, Eq. (14) yields one estimate $\hat{S}_m(\boldsymbol{k})$ of $\hat{S}(\boldsymbol{k})$ for each m and each pattern orientation:

$$\hat{S}_m(\boldsymbol{k}) = \frac{\hat{R}_m(\boldsymbol{k})}{b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p})} \ . \tag{15}$$

Each such estimate is valid in the circular region $|\boldsymbol{k} + m\boldsymbol{p}| \le k_0$ of reciprocal space where $\hat{H}(\boldsymbol{k} + m\boldsymbol{p}) \neq 0$ (Fig 1*a*). Many of these regions overlap so that there is more than one estimate of $\hat{D}(\boldsymbol{k})$ at the same point $\boldsymbol{k}$ (Fig. 1*d*). The noise-optimal way to combine independent measurements of the same unknown is through a weighted average, in which each measurement is given a weight inversely proportional to its noise variance (Papoulis, 1991). One may also reduce the redundancy of the measurements, and the number of exposures, by setting up an extended system of equations for separating the components (Heintzmann, 2003).

We show in the next section that the noise variance of $\hat{R}_m(\boldsymbol{k})$ is independent of both $\boldsymbol{k}$ and m. The noise variance of $\hat{S}_m(\boldsymbol{k})$ is therefore inversely proportional to $\left|b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p})\right|^2$ (because a constant factor c multiplying a stochastic variable changes the variance by a factor $|c|^2$), and the noise-optimal weighted average becomes

$$\hat{S}_{optimal\_average}(\boldsymbol{k}) = \frac{\sum \left(\frac{\hat{R}_m(\boldsymbol{k})}{b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p})} \right) \left| b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p}) \right|^2}{\sum \left| b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p}) \right|^2} = \frac{\sum b^*{}_m \hat{H}^*(\boldsymbol{k} + m\boldsymbol{p}) \hat{R}_m(\boldsymbol{k})}{\sum \left| b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p}) \right|^2}, \tag{16}$$

where the sums are taken over all pattern orientations as well as over the index *m*.

The weighted average in Eq. (16), a direct linear inverse filter without regularization, is highly unstable in regions where its denominator approaches zero. To regularize the estimate, Eq. (16) can be turned into a generalized Wiener filter (Goodman, 2005) by introducing a Wiener parameter α^2 in the denominator (Gustafsson, 2000, 2005):

$$\hat{S}_r(\boldsymbol{k}) = \frac{\sum b^*{}_m \hat{H}^*(\boldsymbol{k} + m\boldsymbol{p}) \hat{R}_m(\boldsymbol{k})}{\sum \left| b_m \hat{H}(\boldsymbol{k} + m\boldsymbol{p}) \right|^2 + \alpha^2}, \tag{17}$$

where $\hat{S}_r(\boldsymbol{k})$ is the regularized estimate of the object information $\hat{S}(\boldsymbol{k})$. An estimate of the object in real space can be obtained by an inverse Fourier transform of $\hat{S}_r$.

The number of orientations of the 1D illumination pattern is usually chosen to balance the coverage of the observable region against the total sample exposure. With more orientations, the overlap between adjacent information components near the edge of the enlarged observable region (Fig. 1*d*) will be stronger. A weak overlap will mean a small value of the denominator of Eq. (16) in the overlap regions, which leads to a low signal-to-noise level there as described in the *Noise* section. One natural choice would be to make the smallest azimuthal overlaps, those between information components with maximal values of m and consecutive orientations, equal to the radial overlaps between information components with consecutive values of m. This would make the number of illumination pattern orientations, N_{dir}, approximately equal to πN. We instead use $N_{dir} = 2N + 1$, which makes the total number of exposures for a given attempted resolution equal in both methods using rotated 1D patterns and 2D patterns.

An alternative to using rotated 1D patterns is to use a 2D illumination pattern (Heintzmann, 2003). We discuss the possible illumination patterns in Appendix A. In this paper, we will use a cross-polarized 4-beam pattern which we believe to be optimal choice for achieving high and nearly isotropic resolution.

Four beams, symmetrically arranged, can be thought of as two beam pairs, one pair in the x and one in the y direction (Fig. 2*g*). One beam pair by itself would produce a 1D sinusoidal illumination pattern exactly as happens in the 1D case. In order for that interference pattern to have true intensity zeros, its two beams must be s-polarized (Fig. 2*a*). With both pairs s-polarized (Fig. 2*g*), they will not interfere with each other, since their electric fields in the object will be orthogonal. Each pair will therefore produce an independent 1D sinusoidal pattern, as in Eq .(3), and the total pattern will be the incoherent sum of these two orthogonal 1D patterns:

$$I(x,y) = \frac{I_0}{2}\left[2 - \cos(2\pi px + \varphi_x) - \cos(2\pi py + \varphi_y)\right] \tag{18}$$

Here φ_x and φ_y are the phases of the pattern in the two orthogonal lateral directions (Fig. 2*i*), and we have assumed that the x and y patterns have the same spatial frequency p. The parameter I_0 is defined as the maximum intensity for each 1D pattern, in analogy with Eq. (3); the maximum intensity of the 2D pattern is $2I_0$ compared to I_0 for the 1D pattern [Eq. (3)].

The same analysis as was described for the 1D case results in the following expression for the observed image in reciprocal space:

$$\hat{D}(k_x,k_y) = c_d T\, \hat{H}(k_x,k_y) \sum_m \sum_n b_{mn} e^{i(m\varphi_x + n\varphi_y)} \hat{S}(k_x - mp, k_y - np) \tag{19}$$

With a 2D pattern, we can obtain high-resolution information in all directions by using only translations (phase shifts) of the illumination pattern without any need for rotation. As in the 1D case, we only need to consider a finite number of terms in Eq. (19), with $|n| \le N$ and $|m| \le N$. To separate the $(2N+1)^2$ information components, we must acquire

$(2N+1)^2$ images, for all combinations of the phases φ_x and φ_y, each ranging from 0 to 2π in (2*N*+1) steps. The reconstruction process for 2D illumination patterns is otherwise nearly identical to that for 1D patterns, and yields the following expression for the regularized object information estimate:

$$\hat{S}_r(k_x,k_y) = \frac{\sum_{m=-N}^{N}\sum_{n=-N}^{N} b^*_{mn}\hat{H}^*(k_x+mp,k_y+np)\hat{R}_{mn}(k_x,k_y)}{\sum_{m=-N}^{N}\sum_{n=-N}^{N}\left|b_{mn}\hat{H}(k_x+mp,k_y+np)\right|^2+\alpha^2}, \tag{20}$$

where

$$\hat{R}_{mn}(k_x-mp,k_y-np) = \frac{1}{c_d T\,(2N+1)^2}\sum_{j=-N}^{N}\sum_{l=-N}^{N} e^{i(mj\varphi_0+nl\varphi_0)}\hat{D}_{jl}(k_x,k_y) \tag{21}$$

and D_{jl} is the image observed when illuminating with the pattern phases $\varphi_x = j\varphi_0$ and $\varphi_y = l\varphi_0$.

In the next sections, we use analytical derivations and computer simulation to analyze and compare the signal-to-noise properties of one- and two-dimensional illumination patterns.

Noise

In the previous section, the observed image was treated as a continuous function. Actual electronic recording of wide-field microscopy images is typically done with a pixelated detector, such as a CCD camera. The expected number of photons recorded in pixel number *s* is

$$D_s = \int_{\text{pixel } s} D(\boldsymbol{r})d\boldsymbol{r} \approx A_{pix}D(\boldsymbol{r}_s) \tag{22}$$

where A_{pix} is the area of one pixel, and $\boldsymbol{r}_s$ is the position of pixel *s*, both referred to object space. Conveniently the integration over a pixel (with a given pixel form factor) near the sampling coordinates can also be interpreted as a convolution of the continuous function with the form factor kernel prior to being sampled at delta-shaped positions. This convolution translates to a multiplication in Fourier space. As the ideal (noise free) signal is band limited, this extra multiplication in Fourier space can be interpreted as a simple multiplicative change of the optical transfer function. If sampled at or above Nyquist frequency the kernel will have its first zero at or above twice the limit of the discrete Fourier transform. The modification to the OTF is thus minor and not altering the support, but may introduce a slight deviation from circular symmetry and therefore should be included in reconstruction algorithms, if theoretically determined PSFs/OTFs

are used. Thus the results derived for continuous frequency-space quantities in the previous sections can therefore be applied directly to the discrete quantities and we will retain the continuous-variable notation $\hat{f}(\boldsymbol{k})$ for clarity.The measured image is not perfect, but is degraded by noise originating in the camera electronics (readout noise) and in inherent statistical variations in the photon counts (shot noise). Whereas readout noise has a variance that is independent of the measured intensity, the shot noise follows a Poisson distribution with variance equal to the expected number of detected photons in that pixel. We can write the number of photons recorded in the noisy image as:

$$D'_s = D_s + \zeta_s + \xi_s \tag{23}$$

where ζ_s is the shot noise and ξ_s the readout noise in pixel s, and D_s is the ideal, noiseless observation described by Eqs. (2) and (22).

Next, we calculate how noise propagates into the reconstruction (the object estimate) from the noise in the observed images. Because we have defined the image in units of detected photons, the variance of the shot noise in each pixel is simply equal to the expected image value in the same pixel:

$$Var\,(\zeta_s) = D_s \tag{24}$$

In reciprocal space, the variance of the shot noise $\hat{\zeta}(\boldsymbol{k})$ at a given point $\boldsymbol{k}$ (defined to be $Var\,(\varsigma) = E[(\varsigma - \bar{\varsigma})(\varsigma - \bar{\varsigma})^*]$, where '*' denotes complex conjugation) can be calculated by explicitly writing out the definition of $\hat{\zeta}$ as the discrete Fourier transform of ζ:

$$Var\,(\hat{\zeta}(\boldsymbol{k})) = Var\left(\sum_s e^{2\pi i \boldsymbol{k}\cdot\boldsymbol{r}_s}\zeta_s\right) = \sum_s \left|e^{2\pi i \boldsymbol{k}\cdot\boldsymbol{r}_s}\right|^2 Var\,(\zeta_s) = \sum_s D_s = Q \tag{25}$$

where Q is the total expected number of detected photons in the entire image, and we have used the facts that the shot noise is independent from pixel to pixel, that variance is additive in sums of independent variables, and that constant factors in front of stochastic variables enter into the variance of those variables as squared magnitudes.

Eq. (25) is a remarkably simple result: *the shot noise variance of the Fourier transform of each raw data image has the same value at every point, and is equal to the total expected number of detected photons in the whole image.*

The frequency-space variance due to readout noise can be found similarly:

$$Var\,(\hat{\xi}(\boldsymbol{k})) = Var\left(\sum_s e^{2\pi i \boldsymbol{k}\cdot\boldsymbol{r}_s}\xi_s\right) = \sum_s Var\,(\zeta_s) = N_{pix}\,\sigma_{ro}^2 \tag{26}$$

where N_{pix} is the number of pixels, and σ_{ro} is the root-mean-square readout noise, in photon units, of each pixel.

The image reconstruction process (Eqs. (13), (14), (17), and (20)) is linear, so the effect of the noise on the estimated object will also be linear. Applying Eqs. (13) and (14) to the noisy image data D' yields a noisy version of the separated information components $\hat{R}_m$:

$$\hat{R}'_m(\boldsymbol{k}) = \hat{R}_m(\boldsymbol{k}) + \chi_m(\boldsymbol{k}), \tag{27}$$

with the noise term

$$\chi_m(\boldsymbol{k}) = \frac{1}{c_d T\,(2N+1)} \sum_{j=-N}^{N} e^{imj\varphi_0} \left(\hat{\zeta}_j(\boldsymbol{k}) + \hat{\xi}_j(\boldsymbol{k})\right), \tag{28}$$

where $\hat{\zeta}_j(\boldsymbol{k})$ and $\hat{\xi}_j(\boldsymbol{k})$ are the Fourier transforms of the noise parts of the jth raw data image.

In the presence of noise, the object estimate for 1D illumination patterns [Eq. (17)] becomes:

$$\hat{S}_r'(\boldsymbol{k}) = \frac{\sum b^*_m \hat{H}^*(\boldsymbol{k}+m\boldsymbol{p})\left(\hat{R}_m(\boldsymbol{k}) + \chi_m(\boldsymbol{k})\right)}{\sum \left|b_m \hat{H}(\boldsymbol{k}+m\boldsymbol{p})\right|^2 + \alpha^2} = \hat{S}_r(\boldsymbol{k}) + \frac{\sum b^*_m \hat{H}^*(\boldsymbol{k}+m\boldsymbol{p})\chi_m(\boldsymbol{k})}{\sum \left|b_m \hat{H}(\boldsymbol{k}+m\boldsymbol{p})\right|^2 + \alpha^2}, \tag{29}$$

where the sums are taken over orientations as well as over the index m.

The noise variance of $\chi_m(\boldsymbol{k})$ is easily calculated from Eqs. (28), (25), and (26):

$$\begin{aligned} Var\left(\chi_m(\boldsymbol{k})\right) &= \frac{1}{c_d^2 T^2 (2N+1)^2} \sum_{j=-N}^{N} Var\left(\hat{\zeta}_j(\boldsymbol{k}) + \hat{\xi}_j(\boldsymbol{k})\right) \\ &= \frac{1}{c_d^2 T^2 (2N+1)^2} \sum_{j=-N}^{N} \left(Q_j + N_{pix}\sigma_{ro}^2\right) = \frac{Q_{avg} + N_{pix}\sigma_{ro}^2}{c_d^2 T^2 (2N+1)} \end{aligned} \tag{30}$$

for all m. Thus, all the separated image components $\hat{R}'_m$ have the same noise variance. In Eq. (30), Q_j is the total expected number of detected photons in the jth raw data image, and Q_{avg} is the expected average number of detected photons per image [specifically, averaged over the sequence of (2N+1) raw data images for one orientation].

For 2D illumination patterns, the noisy estimate of the object is found in a similar way [see Eq. (20)]:

$$\hat{S}_r'(k_x,k_y) = \hat{S}_r(k_x,k_y) + \frac{\sum_{m=-N}^{N}\sum_{n=-N}^{N} b^*_{mn}\hat{H}^*(k_x+mp, k_y+np)\chi_{mn}(k_x,k_y)}{\sum_{m=-N}^{N}\sum_{n=-N}^{N}\left|b_{mn}\hat{H}(k_x+mp, k_y+np)\right|^2 + \alpha^2}, \tag{31}$$

where

$$\chi_{mn}(k_x,k_y) = \frac{1}{c_d T\,(2N+1)^2} \sum_{j=-N}^{N} \sum_{l=-N}^{N} e^{i(mj+nl)\varphi_0} \left(\hat{\zeta}_{jl}(k_x,k_y) + \hat{\xi}_{jl}(k_x,k_y)\right) \tag{32}$$

is the noise in the 2D separated information components, and $\hat{\zeta}_{jl}(k_x,k_y)$ and $\hat{\xi}_{jl}(k_x,k_y)$ are the Fourier transforms of the noise parts of the raw observed image labeled (*j*,*l*).

The argument used to derive Eq. (25) is valid for any raw data image, and thus applies to the 2D as well as the 1D case:

$$Var\left(\hat{\zeta}_{jl}(k_x,k_y)\right) = \sum_s D_{jl}(\mathbf{r}_s) = Q_{jl}\ , \tag{33}$$

where Q_{jl} is the total expected number of detected photons in image (*j*,*l*). The variance of the noise in the separated information components is then

$$\begin{aligned} Var\left(\chi_{nm}(k_x,k_y)\right) &= \frac{1}{c_d^2 T^2 (2N+1)^4} \sum_{j=-N}^{N} \sum_{l=-N}^{N} Var\left(\hat{\zeta}_{jl}(k_x,k_y) + \hat{\xi}_{jl}(k_x,k_y)\right) \\ &= \frac{1}{c_d^2 T^2 (2N+1)^4} \sum_{j=-N}^{N} \sum_{l=-N}^{N} \left(Q_{jl} + N_{pix}\sigma_{ro}^2\right) = \frac{Q_{2Davg} + N_{pix}\sigma_{ro}^2}{c_d^2 T^2 (2N+1)^2} \end{aligned}\ , \tag{34}$$

where Q_{2Davg} is the average expected number of detected photons per image over the 2D sequence of $(2N+1)^2$ two-dimensionally phase-shifted raw images.

The photon averages Q_{avg} and Q_{2Davg} are related in a simple way to the parameter b_0, which is defined as the zero-frequency term in the Fourier expansion of the emission rate pattern $G(\boldsymbol{r})$ [Eq. (9)], or, in other words, as the average emission rate per molecule. To see this relation, first note that the average image over the phase-shift series is equal to $c_d T\ R_0(\boldsymbol{r})$, by the real-space equivalent of Eqs. (13) and (14) for *m*=0. Then note that the sum of the photon counts across the image is equal to the zero-frequency value of the discrete Fourier transform of that image. Thus Q_{avg} is equal to $c_d T\ \hat{R}_0(\mathbf{0})$, which, by Eq. (14), is equal to $c_d T\ b_0 \hat{H}(\mathbf{0})\hat{S}(\mathbf{0})$, which in turn equals $c_d T\ b_0 \hat{S}(\mathbf{0})$ since the optical transfer function $\hat{H}(\boldsymbol{k})$ is normalized such that $\hat{H}(\boldsymbol{0}) = 1$. The quantity $\hat{S}(\boldsymbol{0})$ has a simple interpretation: it represents the total number of fluorophore molecules in the field of view. The expression $Q_{avg} = c_d T\ b_0 \hat{S}(\mathbf{0})$ is unsurprising; it simply states that the expected number of photons detected in the image (Q_{avg}) equals the product of the number of molecules present [$\hat{S}(\boldsymbol{0})$], the average emission rate per molecule (b_0), the exposure time (T), and the detection efficiency (c_d). Similarly for 2D patterns, $Q_{2Davg} = c_d T\ \hat{R}_{00}(0,0) = c_d T\ b_{00}\hat{S}(\mathbf{0})$. Substituting these values into Eqs. (30) and (34) yields

$$Var\left(\chi(\boldsymbol{k})\right)=\begin{cases}\dfrac{c_d T b_0 \hat{S}(\boldsymbol{0})+N_{pix}\sigma_{ro}^2}{c_d^2 T^2 (2N+1)}, & \text{for 1D patterns}\\[2ex] \dfrac{c_d T b_{00} \hat{S}(\boldsymbol{0})+N_{pix}\sigma_{ro}^2}{c_d^2 T^2 (2N+1)^2}, & \text{for 2D patterns}\end{cases}, \tag{35}$$

where we have suppressed the index *m*, since $Var\left(\chi_m\right)$ is independent of *m*.

From Eqs. (29) and (31) it follows that the variance of the noise in the estimated object in reciprocal space is given by:

$$\sigma^2(\mathbf{k})=Var\left(\hat{S}_r{}'(\mathbf{k})\right)=\begin{cases}\dfrac{Var\left(\chi\right)\sum\left|b^*{}_m \hat{H}^*(\mathbf{k}+m\mathbf{p})\right|^2}{\left(\sum\left|b_m \hat{H}(\mathbf{k}+m\mathbf{p})\right|^2+\alpha^2\right)^2}, & \text{for 1D patterns}\\[2ex] \dfrac{Var\left(\chi\right)\sum\left|b^*{}_{mn} \hat{H}^*(k_x+mp,k_y+np)\right|^2}{\left(\sum\left|b_{mn} \hat{H}(k_x+mp,k_y+np)\right|^2+\alpha^2\right)^2}, & \text{for 2D patterns}\end{cases}, \tag{36}$$

where sums are taken over all the terms involved in the reconstruction of the full image in Eqs. (17) and (20): in the 1D case the terms from all the pattern orientations as well as all values of the index *m* are included, and in the 2D case the sums run over all indices (*m*,*n*).

In practice, the regularized object estimate, Eqs. (17) and (20), may be multiplied by an apodization function before being transformed back to real space, to soften the hard edge of the enlarged observable region, which could otherwise lead to ringing artifacts in the reconstruction. Such apodization has no effect on the signal-to-noise ratio at any given point $\boldsymbol{k}$ of reciprocal space, because signal and noise at that point would be multiplied by the same number. For the same reason, the signal-to-noise ratio in reciprocal space is also independent of the Wiener constant α^2. For purposes of signal-to-noise calculations in reciprocal space we can therefore set α^2=0. The noise in the reconstructed object estimate [Eq. (36)] then simplifies to

$$\sigma^2(\mathbf{k})=Var\left(\chi\right)\frac{\sum\left|b^*{}_m \hat{H}^*(\mathbf{k}+m\mathbf{p})\right|^2}{\left(\sum\left|b_m \hat{H}(\mathbf{k}+m\mathbf{p})\right|^2\right)^2}=\frac{Var\left(\chi\right)}{\sum\left|b_m \hat{H}(\mathbf{k}+m\mathbf{p})\right|^2}, \tag{37}$$

for 1D pattern and similarly for the 2D pattern. Substituting Eq. (35) into Eq. (37) and taking the square root, we obtain the standard deviation of the noise:

$$\sigma(\mathbf{k}) = \begin{cases} \sqrt{\dfrac{b_0 \hat{S}(\mathbf{0}) + N_{pix}\sigma_{ro}^2 / c_d T}{c_d T (2N+1) \sum \left| b_m \hat{H}(\mathbf{k} + m\mathbf{p}) \right|^2}} & \text{for 1D patterns} \\ \sqrt{\dfrac{b_{00} \hat{S}(\mathbf{0}) + N_{pix}\sigma_{ro}^2 / c_d T}{c_d T (2N+1)^2 \sum \left| b_{mn} \hat{H}(k_x + mp, k_y + np) \right|^2}} & \text{for 2D patterns} \end{cases}, \tag{38}$$

where the sum is again taken over the full 2D array of information components in the 2D case, and over both phase shifts and pattern orientations in the 1D case.

The two parts of Eq. (38) differ only (apart from the purely notational distinction between b_0 and b_{00}) in that the factor $(2N+1)$ is replaced by $(2N+1)^2$ in the 2D case. In each case, the factor represents the number of images involved in a single Fourier-based separation operation. We can formulate the result in Eq. (38) as a single expression that covers both the 1D and 2D cases above (as well as other alternative 2D methods to be discussed in the *Linear structured illumination imaging* section) by defining N_{ph} to be the number of phase shifts (and thus raw data images) involved in a single Fourier-based separation and N_{dir} to be the number of repeats, such as at different pattern orientations, of that separation; and adopting the notation b_0 for the zero order coefficient in 2D as well as 1D methods. For example, in the 1D case above $N_{ph} = N_{dir} = 2N+1$, whereas in the 2D case $N_{ph} = (2N+1)^2$ and $N_{dir} = 1$. If we let $SNR(\boldsymbol{k}) = \hat{S}(\boldsymbol{k})/\sigma(\boldsymbol{k})$ be the signal-to-noise ratio at spatial frequency $\boldsymbol{k}$, we get

$$SNR(\mathbf{k}) = \frac{\hat{S}(\mathbf{k})}{\sigma(\mathbf{k})} = \hat{S}(\mathbf{k}) \sqrt{\frac{c_d T N_{ph} \sum_{u=1}^{N_{ph} N_{dir}} \left| b_u \hat{H}(\mathbf{k} + \mathbf{p}_u) \right|^2}{b_0 \hat{S}(\mathbf{0}) + N_{pix}\sigma_{ro}^2 / c_d T}} . \tag{39}$$

Experimentally, readout noise can be largely eliminated by using electron multiplication CCD cameras (albeit at the cost of doubling the shot noise variance (Robbins & Hadwen, 2003), which here can be modeled by decreasing the value of the detection efficiency c_d by half). If readout noise can be neglected, Eq. (39) can be simplified to

$$SNR(\boldsymbol{k}) = \frac{\hat{S}(\boldsymbol{k})}{\sqrt{\hat{S}(\boldsymbol{0})}} \sqrt{c_d T_{tot}} \sqrt{\frac{\sum_{u=1}^{N_{ph} N_{dir}} \left| b_u \hat{H}(\boldsymbol{k} + \boldsymbol{p}_u) \right|^2}{N_{dir} b_0}} . \tag{40}$$

where $T_{tot} = N_{dir} N_{ph} T$ denotes the total exposure time for the entire data set. In the rest of this article, we will assume that readout noise has been eliminated, and thus use Eq. (40).

The first factor in Eq. (40) describes a property of the object itself; it expresses the unsurprising fact that a structure of given strength [i.e. with a given $\hat{S}(\boldsymbol{k})$] is harder to detect if the environment is crowded or contains background (i.e. contains a large total amount $\hat{S}(\boldsymbol{0})$ of fluorophore). The second factor simply states that the photon statistics are improved by detecting light more efficiently, and by exposing longer (assuming that the object tolerates it). The $\boldsymbol{k}$-dependence of the third factor in Eq. (40) can be thought of as an "effective transfer function" that describes the relative signal-to-noise properties of different points of reciprocal space.

The factor N_{dir} in the denominator of the third factor might at first glance seem to imply that, for a given total exposure time T_c, the 1D approach degrades signal-to-noise ratio by a factor $1/\sqrt{N_{dir}}$. That conclusion would be false, however, because the coefficients b_u that also occur in this factor are quite different for 2D than for 1D patterns, as we shall see in the section on switchable labels, and in some situations more than compensate for the factor $1/\sqrt{N_{dir}}$. It is also instructive to consider how the rotationally RMS-averaged SNR depends on N_{dir} (assuming for simplicity that the factor $\hat{S}(\boldsymbol{k})$ is independent of angle). During rotational averaging, the sum $\sum \left| b_u \hat{H}(\boldsymbol{k} + \boldsymbol{p}_u) \right|^2$ accumulates N_{dir} contributions from the N_{dir} pattern orientations and thus becomes proportional to N_{dir}; this factor cancels the factor N_{dir} in the denominator. The rotationally RMS-averaged SNR is thus independent of the number of pattern orientations in the 1D mode. Perfect isotropy could in principle be achieved without noise penalty by letting $N_{dir} \to \infty$, under our assumption that readout noise can be neglected.

The linear reconstruction method described in this article produces the least-squares estimate of the fluorophore density in the object given the assumption of equal variance over space in the data (i.e. a lot of out-of-focus light). Therefore, the results for the signal-to-noise ratio in the object reconstruction applies equally to any other reconstruction method (e.g. iterative) that computes the unconstrained least-squares estimate with equal weights. However, it should be noted, that iterative (e.g. maximum likelihood based) approaches considering the Poisson statistics of photon noise and the spatially varying variance may achieve better reconstruction results with enhanced signal to noise ratios.

Nonlinear imaging with photo-switchable fluorescent labels

A promising method for achieving non-linear response in fluorescent imaging is to use reversibly photo-switchable fluorescent labels (Hell, 2003; Hofman *et al*., 2005). Examples of such labels include chemical fluorophores such as diarylethenes (Irie, 2002) and spirobenzopyrans (Sakata *et al*., 2005); switchable proteins of the GFP (Lukyanov *et al*., 2000, Ando *et al*., 2004) and phytochrome (Tu & Lagarias, 2005) families; and cyanine-dye pairs (Bates *et al*., 2005). Each such label (typically consisting of a single molecule or molecule pair) has two states – the "*on*" state in which the label fluoresces when illuminated at the excitation wavelength λ_{ex}, and the "*off*" state in which the label

does not fluoresce. Labels can be switched from the *on* to the *off* state by illumination with light of wavelength λ_{off} , and switched back from the *off* to the *on* state by illumination at wavelength λ_{on} . For most practical labels, λ_{ex} equals either λ_{on} or λ_{off} (or, more precisely, the spectrum for fluorescence excitation is similar or identical to the activation spectrum for one direction of photo-switching). We will assume three separate wavelengths here, and in a later section discuss issues that arise if λ_{ex} equals λ_{on} or λ_{off} .

An ideal switching process may be modeled in the following way. Consider an object with density $S(\boldsymbol{r})$ of switchable fluorescent labels, which have all been switched to the *on* state initially. Next, the object is illuminated with *off* light with an intensity pattern $I_{off}(\boldsymbol{r})$ for a time interval T . Immediately afterward, the remaining density of *on* labels is:

$$S_{on}(\boldsymbol{r}) = e^{-c_{off} I_{off}(\boldsymbol{r})} S(\boldsymbol{r}) , \tag{41}$$

where the parameter c_{off} is the product of the *off* exposure time T and the *off*-photo-switching cross section. If this sample is now observed under spatially uniform excitation light with intensity I_{ex} , the emission from the object will be

$$Em(\boldsymbol{r}) = c_e I_{ex} S_{on}(\boldsymbol{r}) = c_e I_{ex} e^{-c_{off} I_{off}(\boldsymbol{r})} S(\boldsymbol{r}) = F\left[I_{off}(\boldsymbol{r})\right] S(\boldsymbol{r}) \tag{42}$$

By comparing Eqs. (7) and (42) we see that the switching can play the role of the nonlinear fluorescent response discussed in previous sections.

In the above description, nonlinearity was created by saturating the *on* to *off* transition; alternatively the *off* to *on* transition could be used in the same way, but that choice generally leads to a higher total *on* population, and thereby to higher noise. We will therefore use the *on* to *off* transition in this article.

For the specific case of a sinusoidal pattern of *off* light [Eq. (3)], one can derive closed form approximate expressions for the fraction of the labels that will remain in the *on* state, and for the coefficients b_m . These expressions are valid when the *off* light intensity is strong enough that labels remain in the *on* state only in small neighborhoods around the zeros of the *off* light intensity pattern $I_{off}(x)$. This is the relevant regime for high-resolution imaging, which makes it clear that the pattern must have clean intensity zeros to be useable, as was mentioned earlier. In this regime, we can approximate the *off* light intensity, in a neighborhood around each intensity zero, by taking the first term in the power series expansion of $I_{off}(x)$ around that point:

$$I_{off}(x) = \frac{I_0}{2}\left(1 - \cos\left(2\pi p x + \varphi\right)\right) \approx I_0 \frac{\left(2\pi p x + \varphi\right)^2}{4} \tag{43}$$

for x near the point where $2\pi p x + \varphi = 0$. In that neighborhood, therefore, the local emission rate $G(x)$ (see Eq. (8)) can be well approximated by a Gaussian:

$$G(x) = F[I_{off}(x)] = c_e I_{ex} e^{-c_{off} I_{off}(x)} \approx c_e I_{ex} e^{-\frac{\eta(2\pi px+\varphi)^2}{4}}, \text{ for } -\pi \le 2\pi px + \varphi \le \pi, \quad (44)$$

where we have defined the "saturation level" $\eta = c_{off} I_0$, which measures the factor by which the photo-driven *on*→*off* transition of the fluorophore has been saturated at the peak intensity points of the illumination pattern. Because the illumination pattern is periodic with period 1/*p*, we can extend this approximation to the full space by repeating expression (44) periodically:

$$G(x) \approx c_e I_{ex} \left(\sum_{q=-\infty}^{\infty} \delta\left(x - \frac{q}{p} \right) \right) \otimes e^{-\frac{\eta(2\pi px+\varphi)^2}{4}} = c_e I_{ex} \sum_{q=-\infty}^{\infty} e^{-\frac{\eta[2\pi(px-q)+\varphi]^2}{4}}. \quad (45)$$

Figure 3 illustrates the good agreement between the exact and approximate expressions for $G(x)$, even for a moderate saturation level of $\eta = 10$. In the regime of moderate to high η, the pattern $G(x)$ of emission rate per fluorophore consists of a series of parallel Gaussian stripes (Fig. 4*c*). From Eq. (45), the full width at half maximum of these stripes is:

$$\text{FWHM} = \frac{2\sqrt{\ln(2)}}{\pi p \sqrt{\eta}}, \quad (46)$$

which can be much smaller than the diffraction limit if $\eta >> 1$.

The advantage of the approximate expression for the emission rate, Eq. (45), is that there is a simple explicit expression for its Fourier transform:

$$\hat{G}(k) = c_e I_{ex} \sum_{m=-\infty}^{\infty} \frac{e^{-\frac{m^2}{\eta}}}{\sqrt{\pi\eta}} \delta(k - mp) e^{-im\varphi} \quad (47)$$

By comparing Eq. (47) with Eq. (10) we recognize that

$$b_m = \frac{c_e I_{ex}}{\sqrt{\pi\eta}} e^{-\frac{m^2}{\eta}}, \quad (48)$$

a closed form expression for the strength of the mth Fourier component of the emission rate pattern under 1D illumination.

A similar derivation for 2D illumination patterns gives:

$$b_{mn} = \frac{c_e I_{ex}}{\pi\eta} e^{-\frac{m^2+n^2}{\eta}}. \quad (49)$$

Thus the values of the coefficients b_m and b_{mn} follow Gaussian profiles, in one and two dimensions, for both 1D and 2D illumination patterns. For fair comparison, we have taken the parameter η, and therefore I_0, to be equal for the 1D and 2D cases, in order that the 1D and 2D Gaussians have the same width and thus similar resolution potential. From Eqs. (3) and **(18)** it follows that the total *off* light exposure to the specimen will be twice as large in the 2D case as in the 1D case, at equal resolution potential. The higher total *off*-light exposure may or may not constitute a serious drawback of the 2D approach, depending on to what extent *off* light contributes to photo-degradation of the particular fluorophore used.

Figure 5 compares the effective OTF, in the sense of Eq. (40), for 1D and 2D illumination using switchable labels. Each effective OTF consists of a number of copies of the conventional OTF $\hat{H}$, each located at a position $\boldsymbol{p}_u$ (compare Fig. 1, *d* and *e*) and scaled by a coefficient b_u, and combined in the overlaps by Eq. (40). The origin peak of each $\hat{H}$ copy is easily recognized in the figure. The overall shape of the effective OTF for 2D illumination patterns (Fig. 5*b*) is defined by the Gaussian profile of the coefficients b_{mn}. In the 1D case, even though the coefficients b_m for each orientation follow the same Gaussian profile as in the 2D case, the heavy overlap of many information components from neighboring pattern orientations (Fig. 1*c*) increases the values of the effective OTF strongly at the lowest spatial frequencies (Fig. 5*a*). At most spatial frequencies the two methods differ only slightly (Fig. 5*c*).

In addition to the shot noise of the switchable labels themselves, an observation may contain additional noise sources, such as readout noise, or shot noise from background light. Background light may be due to autofluorescence, or to a fraction of the labels that do not respond to the *off* light. Some switchable fluorophores inherently possess such a *permanent-on* state in which they can get trapped (Lukyanov *et al*., 2002). We include this effect by assuming that a fraction γ of the fluorophore labels are non-switchable (or *permanent-on*). The local emission rate Eq. (42) can then be modeled as:

$$Em(\mathbf{r}) = c_e I_{ex}\left((1-\gamma)e^{-c_{off} I_{off}(\mathbf{r})} + \gamma\right)S(\mathbf{r}) = (1-\gamma)Em_{ideal}(\mathbf{r}) + c_e I_{ex}\gamma\, S(\mathbf{r})\,. \tag{50}$$

where $Em_{ideal}(\boldsymbol{r})$ is the fluorescent emission that the object would produce if all the fluorophores were switchable (i.e $\gamma = 0$).

In reciprocal space, Eq. (50) becomes:

$$\hat{E}m(\mathbf{k}) = (1-\gamma)\hat{E}m_{ideal}(\mathbf{k}) + c_e I_{ex}\gamma\, \hat{S}(\mathbf{k}) = \sum_m (1-\gamma)b_m\, \hat{S}(\mathbf{k} - m\mathbf{k}) + c_e I_{ex}\gamma\, \hat{S}(\mathbf{k})\,, \tag{51}$$

using Eqs. (7) and (10). Here the $\gamma\hat{S}(\boldsymbol{k})$ term effectively contributes to the m=0 term of the sum. Thus, the effect of a fraction γ of *permanent-on* fluorophore labels is to increase the strength of the zero-order (conventional resolution) information component from b_0 to $(1-\gamma)b_0 + c_e I_{ex}\gamma$, and to decrease the strengths of all other components by a factor $(1-\gamma)$. If γ is known and remains constant in time, we can correct for it in the reconstruction process by rescaling the Fourier weights b_m in the above manner.

However, the *permanent-on* fraction will affect the signal-to-noise ratio in two ways: it will decrease the signal at all extended-resolution spatial frequencies by a modest factor $(1-\gamma)$, and, more importantly, it will increase the noise at all frequencies by effectively increasing the coefficient b_0 which determines the noise level [Eq. (40)]. Because the fraction $b_0/c_e I_{ex}$ of the switchable fluorophores that are left in the *on* state after exposure to the *off* light pattern is typically small, even a small fraction γ of *permanent-on* fluorophores can have a large impact on b_0. The relative increase in noise can be particularly severe in the case of 2D patterns (Fig. 5*d*), for which the fraction of *on*-state fluorophores is even smaller than in the 1D case (Fig. 4*d*).

The comparison in Fig. 5 was done at one particular saturation level, $\eta = 20$. To judge how the relative performance of 1D and 2D patterns depends on the saturation level, we consider the noise in the highest reconstructed spatial frequencies. This comparison is simple to do analytically, because there is no overlap between reconstructed components at the highest resolution (Fig. 1*d*,*e*), so that the sums in Eq. (40) reduce to a single term. From Eqs. (48) and (49), the average emitted photon count for the 1D illumination pattern is $b_0/b_{00} = \sqrt{\pi\eta}$ times greater than for 2D pattern, and the same ratio holds for b_N/b_{0N}. Using Eq. (40), the ratio of the SNR at the highest resolution of the reconstructed image between 2D and 1D patterns is then $SNR_{2D}/SNR_{1D} = (b_{0N}/b_N)\sqrt{N_{dir}\, b_0/b_{00}} = \sqrt{N_{dir}}\Big/\sqrt[4]{\pi\eta} = \sqrt{2N+1}\Big/\sqrt[4]{\pi\eta}$, assuming again that the same total number $(2N+1)^2$ of images is used in both cases.

To see how this ratio behaves with increasing saturation level, we must first establish how the truncation point *N* scales with the saturation level η. The number of Fourier components b_m that must be taken into account is determined by the acceptable level of relative truncation error ε. The Gaussian envelope that describes the b_m values has the form $\exp\left(-m^2/\eta\right)$, implying that a particular relative truncation error will occur for a particular ratio m^2/η. As η increases, the truncation point $m = N$ will thus be chosen so that *N* scales as $\sqrt{\eta}$. The SNR ratio $\sqrt{2N+1}\Big/\sqrt[4]{\pi\eta}$ at the highest spatial frequencies will therefore be essentially independent of the saturation level η, as both the numerator and the denominator scale as $\sqrt[4]{\eta}$ for $\eta >> 1$. For the saturation level $\eta = 20$ and N=7 used above, this ratio is approximately 1.37. There is thus no dramatic difference between the noise performance of 1D and 2D patterns at the highest spatial frequencies at any saturation level. By a similar argument, one can show that the relative performance of 1D and 2D patterns is independent of saturation level also at intermediate frequencies (in the absence of permanently fluorescent background). In the case of idealized switchable labels, the 2D method has a slight advantage at high frequencies (as indicated by the factor of ≈1.37 quoted above), whereas at lower spatial frequencies the 1D patterns will provide a higher SNR because of the overlap between the information components coming from multiple orientations of the illumination pattern (Fig. 5*c*). As was mentioned in the discussion after Eq. (51), the performance of the 2D method is more sensitive than the 1D method to background. If even a few percent of the fluorophores are in a *permanent-on* state, or a comparable amount of non-switchable

autofluorescence, background light, or camera readout noise is present, the 1D method becomes superior at all frequencies (Fig. 5*d*).

Numerical results

To verify the analytical expressions derived in the previous section, we have performed numerical experiments, using phantom objects consisting of a random collection of straight rods (Fig. 6*a, d*). Results of the above analytical calculations were compared to direct simulations of the microscopy process. To illustrate how the noise in the reconstructed image depends on the total observed photon count, we used two phantom objects, one consisting of narrow bright rods on a dark background, and the other consisting of dark rods on a bright background. Both test objects were set to zero near the image boundary to avoid any possibility of edge artifacts that might bias the noise analysis. Only shot noise was considered; readout noise was assumed to have been eliminated. To represent background, a fraction $\gamma = 4\%$ of the labels were assumed to be permanently fluorescent.

To calculate signal and noise we assumed a labeling density such that each white pixel in the bright-background object, and each one-pixel length of rod in the dark-background object, would emit an average of 140 detectable photons per exposure, if all fluorophore labels in that pixel were in the *on* state. This number was chosen to be similar to the number of photons a single fluorescent-protein molecule per pixel could emit without bleaching. It was estimated from the best-characterized fluorescent protein, EGFP. Measurements have shown that it takes about 174s to bleach EGFP by 50%, with an illumination intensity that yields an initial emission rate of 1000 photons/s/molecule (Shaner *et al.*, 2005). That photostability result corresponds to each molecule emitting an average of $174 \cdot 1000 / \ln(2) \approx 240{,}000$ photons before bleaching, if the decay is approximated as single-exponential. We account for the several signal-reduction factors: acceptance solid angle of the objective lens of (0.33), detector quantum efficiency (0.9), a factor to avoid complete photo-bleaching (0.5) and losses in lenses and filters (0.9). This leaves 32,000 observable photons. Divided over 225 images (see below), this yields about 140 detectable photons per molecule per image. Since each molecular label will be turned *off* in the majority of images, the total number of photons it emits per data set will actually be much lower than the above numbers. The above calculations should be considered as order-of-magnitude estimates, since the photo degradation properties of real photo-switchable molecules are not yet well characterized and are likely to be complicated. We chose these light levels for our simulations with the intent that the resulting performance estimates should be conservative for real samples.

The simulations and numerical calculations used an analytical widefield OTF for a diffraction-limited optical microscope in the scalar, paraxial approximation (Born & Wolf, 2002):

$$OTF(\boldsymbol{k}) = \left(2b(|\boldsymbol{k}|) - \sin(2b(|\boldsymbol{k}|))\right)/\pi, \quad \text{where } b(k) = \cos^{-1}(k/k_0), \tag{52}$$

and $k_o = 2NA/\lambda_{em}$ is the radius of the normally observable region of reciprocal space. This expression was picked because the particulars of the OTF are unimportant for the general questions we want to address.

The numerical simulations were performed on a grid of 256x256 pixels. The frequency-space pixel size was set equal to 1/10 of the conventional resolution limit, k_0. According to the Nyquist theorem, the highest spatial frequency encodable by the data was 12.8 times k_0. In experimental practice, the raw data would be sampled closer to the conventional Nyquist limit, because excessive oversampling decreases field of view for a detector with a fixed number of pixels, and increases the readout noise. The density of pixels would then be increased during data processing to accommodate the increased resolution. Our choice to use constant, small pixel spacing throughout the computations was based on convenience; the small raw-data pixel size has no deleterious effect because the simulation does not include readout noise. In terms of physical distance, our pixel size would be $[1/(2k_0)]/12.8 \approx 8.1$ nm for an NA 1.2 objective observing at λ_{em} =500 nm. With this pixel size, our assumption of a single label per one-pixel length of rod is well within the range of experimentally achieved GFP labeling densities on fibrous cellular structures (Faire *et al*., 1999), and is several orders of magnitude below the ultimate packing limit set by the size of GFP molecules (Patterson *et al*., 2007).

The saturation parameter η was varied between 10 and 30. To illustrate the practical meaning of these values we note that for $\eta = 10$, approximately 18% of the fluorescent labels remain in the *on* state for 1D illumination pattern, and 3.4% for a 2D pattern. For $\eta = 30$, the corresponding fractions are 10.4% and 1.1%. The truncation value N was selected such that the strength b_{N+1} of the first discarded term was less than 5% of the constant term b_0.

Noiseless observed images were calculated from the phantom objects and the applied illumination patterns by using Eqs. (2), (3), (18), and (42), and noisy observed images were formed from the ideal ones by applying a Poisson distribution. During the reconstruction, we used Eqs. (13) and (14), and their 2D equivalent Eq. (21), to separate the raw data into the information components $\hat{R}_m{}'(\boldsymbol{k} - m\boldsymbol{p})$. These components were then translated to their correct positions in reciprocal space, reassembled using Eqs. (17) and (20), apodized with a triangular window function, and finally inverse Fourier transformed back to real space to obtain a high resolution reconstruction of the object (Fig. 7). We set the cutoff frequency of the apodization function to 90% of the theoretical resolution limit $Np + k_0$, to account for the non-circular shape of the support of the effective OTF (compare Fig. 1*d*,*e*). Because the reciprocal-space translation vectors $(-m\boldsymbol{p})$ did not in general correspond to integral pixels in the discrete data arrays, the translations were performed by multiplying by the corresponding harmonic wave $e^{-2\pi i m \boldsymbol{p}\cdot\boldsymbol{r}}$ in real space.

In order to avoid introducing an additional adjustable parameter, we used no Wiener regularization [i.e., we set $\alpha^2 = 0$ in Eqs. (17) and (20)], and instead relied entirely on the apodization to suppress the divergent noise levels at the edge of the

extended observable region. This choice only affects the visual appearance of the final image and does not affect our analysis of resolution and signal-to-noise ratio.

The 1D pattern was rotated to a sequence of different orientations, as described in the theory section. The number of orientations was chosen to be equal to the number of translations (phase shifts) for each direction (i.e., $N_{dir} = N_{ph} = 2N+1$). The 2D pattern was instead translated in two orthogonal directions through an $(2N+1)\times(2N+1)$ array of positions. The total exposure of *off* light in each image was taken to be twice as large in the 2D case as in the 1D case, to produce the same resolution potential [see the discussion after Eq. (49)]. The 2D exposure pattern is essentially two mutually incoherent copies of the 1D pattern, one oriented in the x and one in the y direction, hence the doubled exposure. The total number of images was the same for the 1D and 2D cases.

It is immediately apparent from the simulation results (Fig. 7) that the dark-background object can be reconstructed with better signal-to-noise ratio than the bright-background object. More detailed comparisons require a quantitative measure. We define $SNR\ (k)$ to be the rotationally averaged SNR in the estimated object (averaged over the annular band of spatial frequencies $\boldsymbol{k}$ whose magnitude $|\boldsymbol{k}|$ lies between k and $k+\Delta k$ for a small increment Δk), or more precisely the root-mean-square (RMS) signal divided by the RMS noise:

$$SNR\ (k) = \frac{\sqrt{\sum_{k<|\boldsymbol{k}|<k+\Delta k} \left|\hat{S}(\boldsymbol{k})\right|^2}}{\sqrt{\sum_{k<|\boldsymbol{k}|<k+\Delta k} \left|\hat{S}(\boldsymbol{k}) - \hat{S}_r(\boldsymbol{k})\right|^2}} \tag{53}$$

where $\hat{S}_r(\boldsymbol{k})$ is the value of the Fourier transform of the reconstructed object estimate at the point (pixel) $\boldsymbol{k}$, $\hat{S}(\boldsymbol{k})$ is the value of the Fourier transform of the corresponding noiseless object estimate, and the sums are taken over all the pixels of the discrete frequency-space data. We used this definition to compare the noise performance for different parameters and methods of illumination, and to compare numerical and analytical results.

The numerical results for the signal-to-noise ratio (SNR) are in excellent agreement with the analytical predictions given by Eq. (40), confirming the correctness of our analysis (Fig. 8). The SNR curves, especially that for the 1D illumination case, show an oscillatory behavior that reflects the rugged landscape of the effective OTF (see Fig. 5): during rotational averaging, some frequency bands contain “peaks,” while other bands sample mostly “valleys” of the effective OTF. As expected, structured illumination with a 1D pattern yields higher SNR at low spatial frequencies, but comparable or slightly lower SNR than the 2D pattern at the highest frequencies. (The slight difference compared to Fig. 5*d* in the relative high-frequency performance of 1D and 2D patterns is due to Fig. 5*d* being a single line profile, whereas Fig. 8 is a rotational average.) Reconstructed images of the test object with bright background have substantially lower SNR than those of the dark-background test object, because the bright background increases the noise

variance at all spatial frequencies in proportion to the increased total number of observed photons, but contributes to the signal only at the smallest spatial frequencies. The maximum resolution at a given SNR for the bright-background object is only about two-thirds of the value for the dark-background object. In all cases, the SNR diminishes rapidly as k increases. At low levels of permanently fluorescent background, either method can achieve a signal-to-noise ratio above unity out to a resolution of about 6 times higher than that of a conventional microscope, even at the low assumed signal level of 140 detected photons from a fully-*on* 8-nm pixel. Higher labeling density or greater photostability would enable even higher resolution.

The dependence of the signal-to-noise ratio on the saturation factor η is indicated in Fig. 9, for particular spatial frequencies. The precise values of the signal-to-noise ratio depend somewhat on the choice of spatial frequency, because of the oscillations apparent in Fig. 8, but the general trends are consistent. The calculation makes the important assumptions that the photo-stability is unaffected by the saturation level (which could be dramatically incorrect for some systems), so that the exposure and emission parameters can be kept constant as η is varied. The *permanent-on* fraction is similarly assumed constant, which may be false if there is a light-induced transition to a *permanent-on* state, as for asCP (Lukyanov *et al.*, 2000). Under these assumptions, we expect the SNR to initially increase with the saturation factor, to reach a maximum at some optimal saturation level, and to decline beyond that. The initial SNR increase is due to the decreasing width of the stripes (or spots, in the 2D case) of remaining fluorescence after the illumination with *off* light [Eq. (46)]. However, as the areas of remaining fluorescence decreases further in size, the decreasing amount of emitted signal light eventually leads to a leveling off and decrease in the SNR at high saturation levels. The value of this optimal saturation level increases with the spatial frequency and decreases with the amount of *permanent-on* background. In Fig. 9*a*, the peak of the SNR curve is apparent for the cases of 4% and 10% *permanent-on* fraction at 3 times the conventional resolution with the 2D pattern; for other parameters the peak occurs at saturation factor values beyond the range plotted.

It was seen from Fig. 5 that in the absence of permanently fluorescent background, the 2D illumination method has better SNR at high and moderate spatial frequencies than the 1D illumination, but that a *permanent-on* fraction of as little as 4% renders the 2D method slightly inferior to the 1D scheme at all spatial frequencies. The same trend is supported by our numerical results (Figs. 9 and 7*c*,*d*). The relative disadvantage of 2D patterns in the presence of *permanent-on* background increases with the saturation level (i.e. with the attempted resolution) (Fig. 9).

The crossover value for the *permanent-on* fraction, at which the SNR of the 1D illumination method begins to outperform that of the 2D method even at high spatial frequencies, is very low, around 2% for a saturation level of 20 (Fig. 10). (This value should only be taken as guidance, because the oscillatory nature of the SNR function (Fig. 8) makes the exact crossover value depend slightly on the particular spatial frequency at which the comparison is made). Figure 10 also confirms our earlier observation that for all methods, objects with bright background yield much lower SNR than objects where the only fluorescent emission comes from the structure of interest.

The 2D method's very low tolerance of permanently fluorescent background stems from the fact that only a very small fraction of the switchable fluorophores are left in the *on* state to fluoresce (Fig. 4*d*); any background therefore rapidly overwhelms the small signal.

Discussion

In this article, the fluorescent labeling of the object has been treated as a continuous density of fluorophore that is continuously switchable. This continuity assumption is valid when there are many labels per resolved area, so that the random switching and bleaching of individual labels averages out. In the extreme resolution regime, the number of labels per resolved area may be small. Addressing this regime accurately would involve modeling the object as a collection of individual labels, and taking into account the stochastic switching response and randomly timed photo bleaching of each label during acquisition, as well as the interaction of the polarization properties of the light with the vector direction of each label's absorption dipole(s). Such a discrete-molecule model is outside the scope of this article.

We have used an idealized model of switchable labels that has stable *on* and *off* states and separate *on*, *off*, and excitation wavelengths, and therefore can be prepared with *off* light and observed at leisure afterwards. Some real photo-switchable molecules such as asCP [also known as asulCP or asFP595 (Chudakov *et al*., 2003)] undergo thermal relaxation from one state to the other, and have an excitation spectrum that is similar to the *on* switching spectrum. The similarity of the spectra means that the act of observation will cause some molecules to turn back on. The degree of *on*-switching caused by an exposure that is sufficiently intense to produce an acceptable number of detected photons will depend on the density of fluorophores and the ratio of fluorescence quantum yield to photo-switching quantum yield. If that degree is small, on the order of a few percent, it will lead to a background similar to that produced when a fraction of the molecules are in a *permanent-on* state. If, on the other hand, the degree of switching on is large, on the order of 100%, our model of switching first and exposing later breaks down. Instead, one would be forced to illuminate with the *off* pattern as well as with excitation light during the observation itself.

For other switchable labels such as Dronpa (Ando *et al*., 2004) and Cy-dye pairs (Bates *et al.*, 2005), the excitation wavelength is instead coincident with the *off* wavelength. In that situation, there is no undesired *on*-switching. The result of the identity of *off*- and excitation wavelengths is instead that the fluorescence is being turned off during observation. If the number of photons that can be detected before the fluorophores have been turned off is insufficient, it may be necessary to repeat the cycle of *on* and *off* light exposure and observation more than once for each phase (and orientation) of the illumination pattern, either as separate images or while integrating the exposures onto a single image. Whether such repeated observation is a problem will depend on whether *on*/*off* cycling contributes to photo-degradation of the photoswitchable compound in question.

For labels that emit very few photons per *on*/*off* cycle, such as some recently developed variants of Dronpa (Ando *et al.*, 2007; Stiel *et al.*, 2007), even the repeated-observation approach may be unworkable. For microscopy with such labels it may be necessary to illuminate with both *on* and *off* light simultaneously, which would cause the labels to cycle rapidly between their *on* and *off* states.

Thus both types of two-wavelength labels can require simultaneous illumination with *on* and *off* light. Such double illumination, which has been used in recent publications (Hofman *et al*., 2005; Schwentker *et al*., 2007), has the undesirable property that the relationship between illumination intensity and emission rate becomes a rational function:

$$Em(\boldsymbol{r}) \propto \frac{I_{on}(\boldsymbol{r})\sigma_{on}}{I_{on}(\boldsymbol{r})\sigma_{on} + I_{off}(\boldsymbol{r})\sigma_{off}} , \tag{54}$$

(where σ_{on} and σ_{off} are the cross sections for on and off-switching), due to competitive equilibrium between on- and off-switching, rather than the exponential relationship of Eq. (41). Because the rational function decays more slowly than the exponential one at high I_{off}, it leads to a significant fraction of the labels being left in the *on* state between the desired peaks (Fig. 11*a*). The effect on noise is similar to that of having a fraction of labels in a *permanent-on* state, or other forms of background: an increased average emission rate (higher coefficient b_0), which decreases the signal-to-noise ratio by lowering the ratios $b_u \big/ \sqrt{b_0}$ in Eq. (40). As with other types of background, the deterioration is more severe for 2D than for 1D patterns. The effect is illustrated in Fig. 11*b* for a 2D illumination pattern, comparing the exponential relationship of Eq. (42) for a maximum saturation factor $\eta = 20$, with the rational relationship of Eq. (54) for a maximum $I_{off} \big/ I_{on}$ ratio of 20. With these parameters, both functions have similar behavior near the peaks. The rational relationship produces significantly lower SNR in this example, except at the lowest spatial frequencies.

Even though this article has discussed signal and noise in the context of two-dimensional imaging of flat specimens, the same theoretical framework can be directly applied to three-dimensional forms of SIM (Gustafsson *et al*., 2000; Frohn *et al*., 2001, Gustafsson *et al.*, 2008, Shao *et al.*, 2008), with obvious modifications for using 3D pattern wave vectors, or separate OTFs for different information orders, if needed.

Any fair comparison of different methods requires knowledge of their performance-limiting factors; for example, bleaching-limited methods should be compared under conditions that cause equal amounts of photobleaching, etc. Nonlinear SIM is likely to be limited by photodegradation, but it is not a priori obvious which processes will dominate the rate of photodegradation of a particular fluorophore. In the case of switchable labels, there could be separate photo-destruction cross sections of labels in their *on* and *off* states for illumination by excitation light, *on* light, or *off* light, and the relative importance among these could vary depending on the type of molecules and on mounting conditions (such as oxygen concentration). The appropriate form of

performance comparison, including which parameters to keep constant, will therefore depend on the specifics of a particular system. In this article we have attempted generality, but have by necessity made some assumptions: that excitation light is damaging to both *on* and *off* labels (we therefore made comparisons at equal total exposure to excitation light, regardless what fraction of labels are in the *on* state), and that further exposure of *off* labels to the *off* light wavelength is not critical (we therefore made comparisons at similar resolution potential, even though it required larger exposure to *off* light with 2D than with 1D patterns). The main point of this article, therefore, is not its detailed results but rather to present a general framework that can be used to analyze and optimize the performance of specific forms of linear or nonlinear SIM, once the limiting factors for those specific cases are known.

Conclusions

We have presented a theoretical framework for analyzing methods of extended-resolution optical microscopy using linear or nonlinear structured illumination. We have applied the general analysis to linear SIM, and to the specific form of nonlinear SIM that is based on photo-switchable fluorescent labels. The effects of photon shot noise on the reconstructed image have been derived analytically and confirmed through simulation.

For linear SIM, the method using sequential 1D illumination patterns provides slightly higher SNR performance and more isotropic resolution (when used with ≥3 pattern orientations) than methods based on 2D patterns.

For nonlinear SIM with switchable labels in the absence of background, the 2D method performs slightly better at high spatial frequencies, whereas the 1D pattern is superior at low spatial frequencies. The 2D scheme deteriorates much more rapidly than the 1D method in the presence of background, however, including background produced when a fraction of the fluorophores are permanently fluorescent. A given amount of resolution extension also requires twice as much exposure to *off* light for the 2D as for the 1D method; this may be a significant disadvantage of the 2D method for molecules where *off* light contributes to photo-degradation.

In calculations and simulations for a dark-background object, emitting at light levels consistent with single GFP-like switchable molecules per 8x8 nm area, both 1D and 2D methods achieved a signal-to-noise ratio of unity out to a resolution of six times the diffraction limit, corresponding to about 33 nm, even when 4% of the fluorescent labels were assumed to be unswitchable.

Acknowledgements

The authors would like to acknowledge John Sedat and David Agard for initiating the project that developed into this article, and thank them and Abraham Szoke for many helpful discussions. The authors are grateful to Rainer Heintzmann for reviewing this paper.

This work was supported in part by the National Science Foundation, the Sandler Family Supporting Foundation, the David and Lucile Packard Foundation, and by the National Institutes of Health through grants GM25101 to John W. Sedat and GM31627 to David A. Agard. The Center for Biophotonics, an NSF Science and Technology Center, is managed by the University of California, Davis, under Cooperative Agreement No. PHY 0120999. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References

Abbe, E. (1873) Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Arkiv. Mikroskop. Anat*. **9**, 413–468.

Ando, R., Mizuno, H. & Miyawaki, A. (2004) Regulated Fast Nucleocytoplasmic Shuttling Observed by Reversible Protein Highlighting. *Science*, **306**, 1370-1373.

Ando, R., Flors, C., Mizuno, H., Hofkens, J., & Miyawaki, A. (2007) Highlighted Generation of Fluorescence Signals Using Simultaneous Two-Color Irradiation on Dronpa Mutants. *Biophys. J.* **92**, L97–L99.

Andresen, M., Wahl, M.C., Stiel, A.C., Gräter, F., Schäfer, L.V., Trowitzsch, S., Weber G., Eggeling, C., Grubmüller, H., Hell, S.W. & Jakobs, S. (2005) Structure and mechanism of the reversible photoswitch of a fluorescent protein. *Proc. Nat. Acad. Sci. USA*, **102**, 13070-13074.

Bailey, B., Krishnamurthi, V., Farkas, D.L., Taylor, D.L. & Lanni, F. (1994) Three-dimensional imaging of biological specimens with standing wave fluorescence microscopy. *Proc. SPIE*, **2184**, 208-213.

Bates, M., Blosser, T.R., & Zhuang, X. (2005) Short-Range Spectroscopic Ruler Based on a Single-Molecule Optical Switch. *Phys. Rev. Lett.* **94,** 108101

Betzig, E, Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., & Hess, H.F. (2006) Imaging Intracellular fluorescent Proteins at Nanometer Resolution. *Science*, ***31*3**, 1642-1645.

Betzig, E. & Trautman, J.K. (1992) Near-Field Optics: Microscopy, Spectroscopy, and Surface Modification Beyond the Diffraction Limit. *Science*, ***25*7**, 189-195.

Born, M. & Wolf, E. (2002) *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light.* Cambridge University Press, Cambridge.

Chudakov, D.M., Feofanov, A.V., Mudrik, N.N., Lukyanov, S. & Lukyanov, K.A. (2003) Chromophore environment provides clue to “kindling protein” riddle. *J. Biol. Chem.* **278**, 7215–7219.

Chung, E., Kim, D., Cui, Y., Kim, Y.-H., & So, P.T.C. (2007) Two-Dimensional Standing Wave Total Internal Reflection Fluorescence Microscopy: Superresolution Imaging of Single Molecular and Biological Specimens. *Biophys. J.* **93** 1747–1757.

Denk, W., Strickler, J.H. & Webb, W.W. (1990) Two-photon laser scanning fluorescence microscopy. *Science,* **248**, 73-76.

Dyba, M & Hell, S.W. (2002) Focal spots size lambda/23 open up far-field fluorescence microscopy at 33 nm axial resolution. *Phys. Rev. Lett*. **88**, 163901-16304.

Faire, K, Waterman-Storer, C.M., Gruber, D., Masson, D., Salmon, E.D., & J. Bulinski, C. (1999) E-MAP-115 (ensconsin) associates dynamically with microtubules in vivo and is not a physiological modulator of microtubule dynamics, *J. Cell Science,* **112**, 4243-4255.

Frohn, J.T., Knapp, H.F. & Stemmer, A. (2000) True optical resolution beyond the Rayleigh limit achieved by standing wave illumination. *Proc. Nat. Acad. Sci. USA,* **97**, 7232-7236.

Frohn, J.T., Knapp, H.F. & Stemmer, A. (2001) Three-dimensional resolution enhancement in f luorescence microscopy by harmonic excitation. *Opt. Lett.* **26**, 828–830.

Goodman, J.W. (2005) *Introduction to Fourier optics,* 3rd ed. Roberts & Co., Englewood, Colorado.

Gustafsson, M.G.L., Agard, D.A. & Sedat, J.W. (1999) I5M: 3D widefield light microscopy with better than 100nm axial resolution. *J. Microscopy*, **195**, 10–16.

Gustafsson, M.G.L. (2000) Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *J. Microscopy*, **198**, 82-87.

Gustafsson, M.G.L. Agard, D.A. & Sedat, J.W. (2000) Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination. *Proc. SPIE*, **3919**, 141–150.

Gustafsson, M.G.L. (2005) Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc. Nat. Acad. Sci. USA,* **102**: 13081-13086.

Gustafsson M.G.L., Shao L., Carlton P.M., Wang C.J., Golubovskaya I.N., Cande W.Z., Agard D.A., Sedat J.W. (2008). Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys J.* **94**:4957-70

Shao L., Isaac B., Uzawa S., Agard D.A., Sedat J.W., Gustafsson M.G.L (2008) I^5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions. *Biophys J.* **94**:4971-83

Heintzmann, R. & Cremer, C. (1998) Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. *Proc. SPIE,* **3568**, 185–195.

Heintzmann, R., Jovin, T.M. & Cremer C. (2002) Saturated patterned excitation microscopy-a concept for optical resolution improvement. *J. Opt. Soc. Am. A,* **19**, 1599-1609.

Heintzmann, R. (2003) Saturated patterned excitation microscopy with two-dimensional excitation patterns. *Micron*, **34**, 283–291.

Hell, S.W. & Stelzer, E.H.K. (1992) Properties of a 4Pi-confocal fluorescence microscope. *J. Opt. Soc. Am. A*, **9**, 2159-2166.

Hell, S.W & Wichmann, J. (1994) Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett*. **19**, 780.

Hell, S.W., Jakobs, S. & Kastrup, L. (2003) Imaging and writing at the nanoscale with focused visible light through saturable optical transitions, *Appl. Phys. A*, **77**, 859-860.

Hell, S.W. (2003) Toward fluorescence nanoscopy, *Nat. Biotechnol.* **21**, 1347-1355.

Hess, S.T., Girirajan, T.P., & Mason, M.D. (2006) Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophys. J.* **91**, 4258-4272.

Hofman, M., Eggeling, C., Jacobs, S. & Hell, S.W. (2005) Breaking the diffraction barrier in fluorescence microscopy at low light intensities by using reversibly photoswitchable proteins. *Proc. Nat. Acad. Sci. USA*, **102**, 17565–17569.

Irie M, Fukaminato T, Sasaki T, Tamai N, & Kawai T (2002) A digital fluorescent molecular photoswitch. *Nature*, **420**, 759–760.

Klar, T.A., Jakobs, S., Dyba, M., Egner, A. & Hell, S.W. (2000) Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Nat. Acad. Sci. USA*, **97**, 8206-8210.

Lukyanov, K.A., Fradkov, A.F., Gurskaya, N.G., Matz, M.V., Labas, Y.A., Savitsky, A.P., Markelov, M.L., Zaraisky, A.G., Zhao, X., Fang, Y., Tan, W. & Lukyanov, S.A. (2000) Natural animal coloration can be determined by a nonfluorescent green fluorescent protein analog. *J. Biol. Chem*. **275**, 25879–25882.

Papoulis, A. (1991) *Probability, random variables, and stochastic processes.* Third edition, McGraw Hill, New York.

Patterson, G.H., Betzig, E., Lippincott-Schwartz, J., & Hess, H.F. (2007) Developing Photoactivated Localization Microscopy (PALM). *Proc. IEEE*, **4244**, 940-943.

Robbins, M.S. & Hadwen, B.J. (2003) The noise performance of electron multiplying charge-coupled devices, *IEEE Trans. Electron Devices* **50**, 1227 – 1232.

Rust, J.M., Bates, M., & Zhuang, X. (2006) Sub-diffraction-limit imaging by stochastic optical resonctruction microscopy (STORM). *Nat. Methods* **3**, 793-795.

Sakata, T., Yan, Y., & Marriott, G. (2005) Optical switching of dipolar interactions on proteins. *Proc. Nat. Acad. Sci. USA*, **102**, 4759–4764.

Schroeder, M.R. (1997) *Number Theory in Science and Communication: With Applications in Cryptography, Physics, Digital Information, Computing, and Self-Similarity*. Springer, Berlin.

Schwentker, M.A., Bock, H., Hofmann, M., Jakobs, S., Bewersdorf, J., Eggeling, C., & Hell, S.W. (2007) Wide-field subdiffraction RESOLFT microscopy using fluorescent protein photoswitching. *Microsc. Res. Tech.* **70**, 269-280.

Shaner, N.C., Steinbach, P.A., & Tsien, R.Y. (2005) A guide to choosing fluorescent proteins. *Nat. Methods*, **2**, 905–909.

Stiel, A.C, Trowitzsch, S., Weber, G., Andresen, M., Christian Eggeling, C., Hell, S.W., Jakobs, S., & Wahl, M.C. (2007) 1.8 Å bright-state structure of the reversibly switchable fluorescent protein Dronpa guides the generation of fast switching variants. *Biochem. J.* **402**, 35–42.

Tu, S.-L., & Lagarias, J.C. (2005) The Phytochromes. *Handbook of Photosensory Receptors.* (ed. by J. Spudich and W. Briggs). Wiley, Weinheim.

Appendix A. Illumination Patterns for Nonlinear Imaging

Several types of patterns can be generated by superposing several beams of light. The 1D sinusoidal pattern in Eq. (3) could be produced by interference between two mutually coherent, collimated beams of uniform and equal intensity. To produce the perfect contrast and true intensity zeros described by Eq. (3), the two beams must be s-polarized relative to each other (Fig. 2*a*). To produce 2D patterns, one would use more than two interfering beams. However, the use of a large number of beams increases the number of images quadratically without any concomitant increase in resolution. Since it is generally undesirable to increase the number of required exposures, only small numbers of interfering beams are interesting. Usable patterns can be produced with three or four beams.

Three beams, symmetrically arranged and azimuthally polarized (Fig. 2*d*) produce a hexagonal intensity pattern with clean zeros (Fig. 2*f*). However, the three-beam pattern has two disadvantages relative to the four-beam pattern. First, its frequency-space unit cell is smaller (by a factor $3\sqrt{3}/8 \approx 0.65$) and therefore it requires a larger number of information components to cover a given area of frequency space (i.e., a larger number of raw data images at a given resolution). Second, its zeros are less sharp (data not shown), thus requiring a higher order of nonlinearity to reach a given resolution.

Additional possible 2D patterns, which provide more isotropic coverage in linear SIM, will be discussed in the context of linear imaging in Appendix B.

Appendix B. Linear structured illumination imaging

We can use analytical results for SNR in nonlinear SIM for analyzing and selecting optimal illumination patterns in linear SIM. The only complications are that the choice of the candidate 2D pattern is less clear than in the nonlinear case. Some candidate patterns are best suited for different phase shift sequences than the rectangular raster scan described above.

To evaluate candidate 2D patterns, we will first compare the corresponding observable regions, in particular the maximum and minimum resolution extension factors in different directions, and then compare their signal-to-noise properties. The values of

these and other parameters are summarized in Table 1. In order to generate simple, easily interpretable numerical values, we here assume that the emission wavelengths are equal, that the beams traverse the pupil at its edge. For comparison, the resolution extension factor for the 1D pattern applied in three orientations has a maximum value of 2 (in the pattern directions), and a minimum of $\sqrt{3} \approx 1.73$ (in the directions furthest removed from any pattern direction), as is clear from the geometry of Fig. 1*b*.

The main 2D pattern that has been discussed above in the nonlinear case is formed by four illumination beams, linearly polarized in two orthogonal pairs (Fig. 2*g*). The light intensity of this pattern has 5 frequency components (Fig. 2*h*), and in the linear case gives access to five information components arranged in a cross (Fig. 1*c*). A drawback of this pattern is that the resolution is extended much less in the diagonal directions than in the x and y directions, extending resolution by a factor of $\sqrt{2} \approx 1.41$, rather than 2 (Fig. 1*c*). The performance in practice can be even more anisotropic, as the extended-resolution information in the diagonal directions is contributed only by peripheral parts of the OTF, which may be quite weak for real-world objective lenses. In fact, implementations of this pattern have explicitly traded off resolution in favor of improved isotropy by using less-than-maximal values of the pattern wave vector $|\boldsymbol{p}|$ (Frohn *et al.*, 2000).

It is therefore attractive to consider alternative 2D patterns that yield more isotropic coverage of frequency space. One alternative is the 3-beam, azimuthally polarized arrangement that has already been mentioned (Fig. 2*d*). Its main drawback is that, because the beam positions in the pupil are not diametrically opposite each other, its largest resolution extension factor is only $1 + \sqrt{3}/2 \approx 1.87$, compared to the maximal resolution extension factor of 2 that the 1D and 4-beam-2D arrangements provide (in their best directions). The 3-beam pattern's smallest resolution extension factor, however, in the direction of the concave "bays" of its observable region, is $(3 + \sqrt{13})/4 \approx 1.65$, larger than the 4-beam cross-polarized arrangement's smallest resolution extension factor of 1.41.

One way to alleviate the anisotropy of the 4-beam arrangement is to choose different polarization states for the beams, such as circular polarization, to make all four beams interfere pairwise. Doing so generates a square array of nine frequency components in the light intensity (Fig. 2*j*,*k*), corresponding to a similar array of nine information components (not shown). The additional components increase the diagonal resolution extension factor to $1 + 1/\sqrt{2} \approx 1.71$, while increasing the required number of images from 5 to 9. The actual minimum resolution factor, $\sqrt{\left(3 + \sqrt{7}\right)/2} \approx 1.68$, occurs at a different angle, 57.1°.

The 2D method described in the *image formation* section separates a rectangular array of information components by using a rectangular array of phase shifts followed by a 2D Fourier transform. Some candidate patterns, however, such as those in Fig. 2*e* and *h*, produce sets of intensity components (and thus information components) that are not arranged in a rectangular (or rhombic) array. In principle, this is not a problem, since any

random set of N_{ph} 2D phase shifts is likely to produce N_{ph} independent equations, and thus allow separation of N_{ph} information components. An arbitrary set of phase shifts, however, often yields an ill-conditioned system of equations, which leads to high and uneven noise levels in the separated components. Fourier-transform-based schemes are well-conditioned and produce equal noise levels in all components. They have the additional nice property that their illumination pattern sequences (with equally-spaced phases from 0 to $2n\pi$) sums to uniform illumination, which means that the total photo-bleaching of the object after one pattern orientation (in the 1D method) or one complete data set (in 2D) is uniform. Illumination sequences that do not sum to uniform intensity risk bleaching patterns into the specimen, which could interfere with later observations. One way to keep the attractive aspects of Fourier-based separation in cases like those in Fig. 2*e* and *h* is to add fictitious dummy components to fill out a rectangular or rhombic array, and to use the larger set of phase shifts appropriate for that extended array. For example, the 5 components in Fig. 2*h* could be separated by using a 3x3 array of phase shifts appropriate for the 9 components of Fig. 2*k*; and the 7 components of Fig. 2*e* could be extended to a 3x3 rhombic array by adding dummy components at the far left and right on the horizontal axis, and could be separated with a 3x3 array of phase shifts. The dummy components could possibly serve a diagnostic function, as they should contain only noise after a perfect separation. This dummy-component approach has the obvious drawback that it increases the number of raw data images required.

In some cases, including those in Fig. 2*e* and *h*, there is an alternative solution, which maintains the Fourier-transform form of the separation while not increasing the number of exposures (or increasing it less than the approach described above), and is still compatible with being generated by rigid-body shifting of a diffraction mask. This idea is to find an oblique direction vector $\boldsymbol{v}$ along which the intensity components are equally spaced. For example, the five components in Fig. 2*h* are equally spaced along the direction $\boldsymbol{v} = (2,1)$, and the seven components in Fig. 2*e* are equally spaced along the direction $\boldsymbol{v} = \left(5/\sqrt{3},1\right)$. In other words, the scalar products of the vector $\boldsymbol{v}$ with the list of positions of the intensity components forms an equally-spaced sequence of numbers; the spacing can be made unity by suitable normalization of $\boldsymbol{v}$. By shifting the real-space pattern in a sequence of equal-sized steps of $\boldsymbol{v}/N_{ph}$, the phase shift of component u in image j becomes $\varphi_{u,j} = ju2\pi/N_{ph}$, which is exactly the form of the phase that appears in a 1D discrete Fourier transform. We can therefore separate the 2D components using a 1D inverse Fourier transform, as was done in Eq. (13) for 1D components. In this approach, phase shifting would involve translating the mask through a large distance (several pattern periods) in a single oblique direction, which would use different hardware than the 2D scan over a single unit cell that is typically used for 2D patterns.

So far, we have assumed that the pattern is produced by a physical diffraction mask, and phase shifted by moving this mask. That is a very attractive situation in practice, though it places constraints on the possible combinations of phase shifts, since there are only two or three free parameters (the mask shifts in the x and y, and possibly z, directions). If, instead, the hardware allows the phase of each of the illumination beams to be controlled independently, it becomes possible to generate a Fourier-type phase

sequence for any arbitrary arrangement of beams. The optimal way to apply such phase shifts to B beams is to assign to beam number *s* one of the integers w_s from an optimal Golomb ruler of order B, and to acquire $1+2G_B$ images with the phase of beam *s* in image *j* set to $\varphi_{s,j} = jw_s 2\pi/(1+2G_B)$, where G_B is the length of the Golomb ruler. [A Golomb ruler of order *n* is a set of *n* integers among which no differences are equal; its length is the largest difference, and an optimal Golomb ruler is one with minimal length (Schroeder ,1997).] The intensity component that corresponds to interference of beams *s* and *t* will thereby be phase shifted by $j(w_s - w_t)2\pi/(1+2G_B)$, where no differences $w_s - w_t$ are equal. This set of phase shifts again has the form of a discrete Fourier transform, and thus allows separation through a 1D inverse Fourier transform (possibly with a few dummy components, if there are gaps in the list of differences). Curiously, for example, this approach makes it possible to perform SIM with five beams arranged pentagonally in the pupil, even though the corresponding intensity pattern in real space is not periodic, but is more akin to a quasi-crystal; the resulting 21 information components could be separated using a 1D Fourier transform from 23 phase shifts (the length of the optimal Golomb rulers of order 5 being 11). (This pattern would be ill suited for many types of nonlinear SIM, however, in that it lacks pure intensity zeros.)

The signal-to-noise ratio produced by the different illumination patterns can be compared by using Eq. (40), but we must first determine the coefficients b_m for each pattern. For linear SIM, the response function F is a direct proportionality: $F(I) = c_e I$ and $G(\boldsymbol{r}) = c_e I(\boldsymbol{r})$ [Eqs. (1) and (7)]. The strengths b_m of the frequency components of the emission rate G are therefore proportional to the strengths of the corresponding frequency components of the illumination intensity I. The factor that limits the available signal is photobleaching, as in most fluorescence microscopy. In the case of linear fluorescence, the degree of photobleaching can be taken to be proportional to the integrated exposure to excitation light. Comparison of different techniques should therefore be done under conditions of equal total exposure time, and equal average intensity over the image. In our notation, the latter condition is equivalent to equal b_0. The relevant quantities, therefore, are the ratios b_m/b_0 for each pattern. The values of these ratios for some illumination schemes are displayed next to the corresponding intensity components in Fig. 2*b*, *e*, *h*, and *k*. For the 1D illumination pattern (Fig. 2*a*–*c*), Eq. (5) immediately implies that $b_1/b_0 = 1/2$. Similarly, for 2D illumination with the cross-polarized 4-beam pattern [Eq. (18), Fig. 2*g*–*i*], it is clear that $b_{01}/b_{00} = 1/4$. With 3-beam illumination, the strength of the side components is affected by the non-parallel polarizations, by a factor of cos(60°)=1/2. In the case of circularly polarized 4-beam illumination, there will be a contribution from axially polarized light in the focal plane, stemming from the radial polarization components in the pupil plane. This axial contribution is 180° out of phase with the laterally polarized pattern, and therefore decreases the b_m/b_0 ratios. The relative amplitude of the axially oriented field, and hence the ratios b_m/b_0, depends on the polar angle β at which the beams pass the focal plane; the values displayed in Fig. 2 and Table 1 were calculated for $\beta = 60°$. Because the lateral and axial polarization patterns combine destructively, it could be attractive to separate them temporally, by recording two images with linearly polarized illumination,

one x- and one y-polarized, in place of the one image with circularly polarized illumination that is in effect an average of the two linearly polarized ones. With linearly polarized light, the b_m/b_0 ratios for the highest spatial frequency components take the values $b_x = 1/4$ and $b_y = -1/8$ for the directions perpendicular and parallel to the polarizations, respectively, both an improvement over the 1/16 value for circularly polarized light (shown in Fig. 2*k*) which equals $(b_x + b_y)/2$. The total exposure time would be divided among the single-polarization exposures, leading to an effective b_m/b_0 ratio of $\sqrt{(b_x^2 + b_y^2)/2}$, or 0.197.

To compare the SNR properties of the different 2D linear SI methods, we have used two figures of merit: the *relative peak SNR at maximum resolution*, which compares the strengths of the highest-resolution information components (i.e. it compares the SNR within those parts of the highest-resolution components that do not overlap any adjacent components), and the *relative average SNR at maximum resolution*, which is proportional to the rotationally RMS-averaged SNR, and takes into account the effect of overlaps. It follows from Eq. (40) that the peak SNR at maximum resolution is proportional to $1/\sqrt{N_{dir}}$ and to $\sqrt{|b_m|^2/b_0} = \sqrt{b_0}\,(|b_m|/b_0) \propto |b_m|/b_0$ for fixed b_0. The average SNR at maximum resolution is proportional to the peak SNR times the square root of the number of information components that reach maximum resolution (in different directions). The values of these figures of merit are tabulated in Table 1 for different illumination modes. As the table illustrates, the 1D method compares well with the 2D approaches, providing the highest SNR and good isotropy while requiring only a modest number of raw data images.

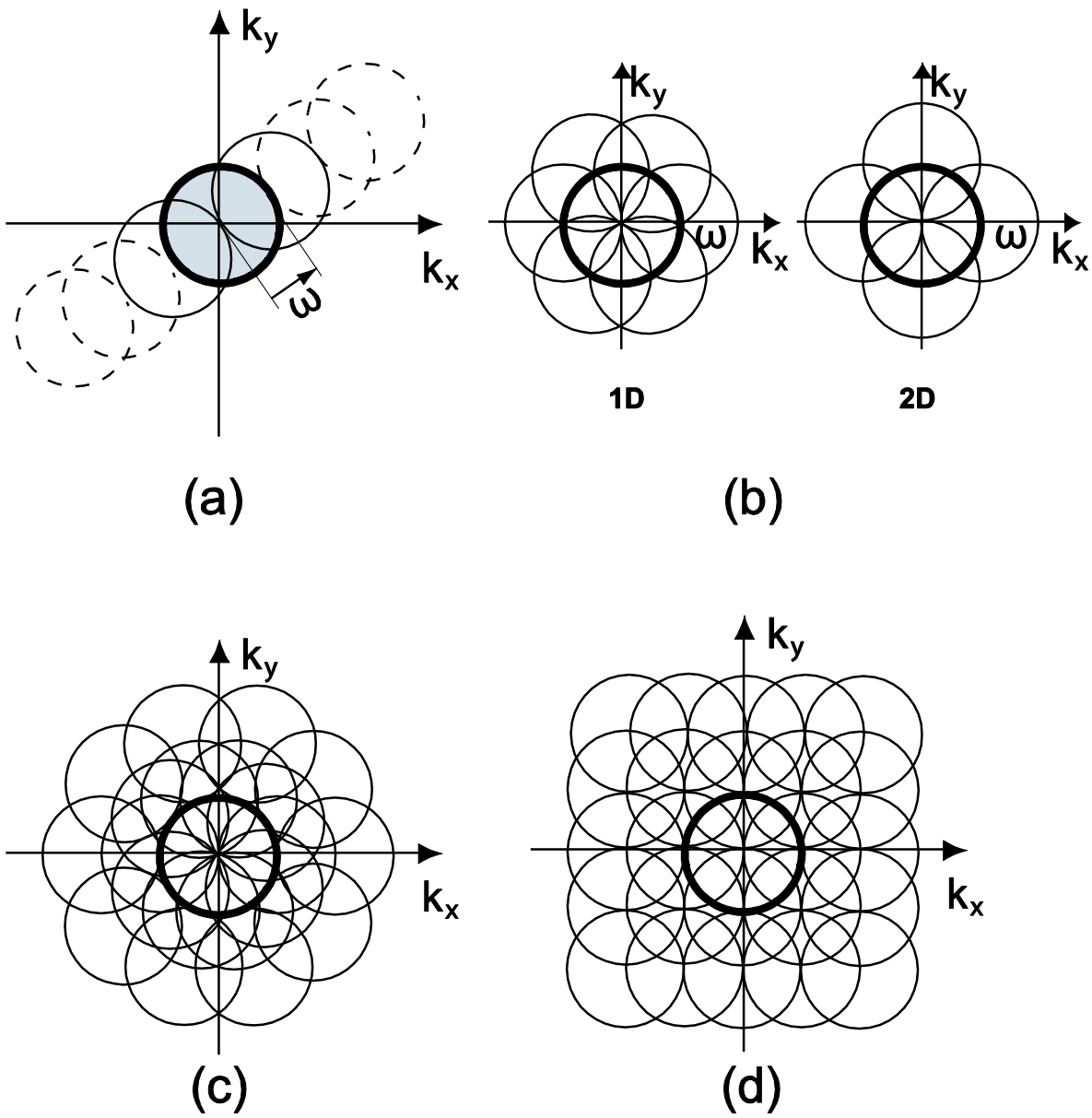

Fig. 1. Enlargement of the observable region by structured illumination microscopy (SIM). (a) The set of spatial frequencies that is observable in conventional microscopy defines a circular region, shown in dark blue. Additional information that can be observed using linear SIM with pattern frequency $\boldsymbol{p}$ is shown as light blue solid circles. Dashed red circles denote further spatial frequencies that can be made observable by nonlinear SIM. (b–e): Spatial frequencies observable by linear structured illumination with a rotating 1D illumination pattern (b), or a 2D pattern (c), and by nonlinear SIM with a rotating 1D pattern (d), or a 2D pattern (e). In (a), (d) and (e) only the two lowest orders are shown, out of in-principle infinite series of information components.

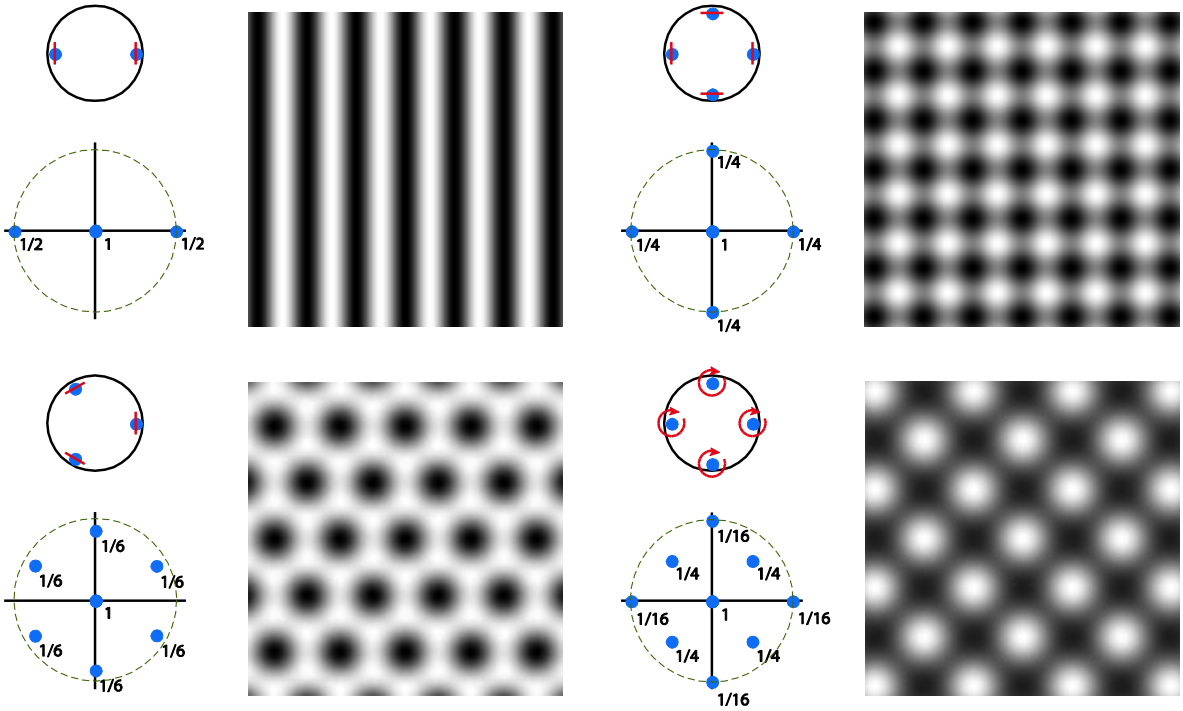

Fig. 2. Positions (blue dots) and polarizations (red lines) of illumination beams in the objective aperture (a,d,g,j), frequency components of the illumination intensity (b,e,h,k), and the corresponding intensity pattern (c,f,i,l), for 1D (a–c), 3-beam 2D (d–f), and 4-beam 2D illumination patterns with crossed linear (g–h) or circular (j–l) polarization. The green dashed circle indicates the maximum possible spatial frequency, corresponding to maximum resolution extension. Numbers next to the blue dots indicate the relative strengths of the frequency components

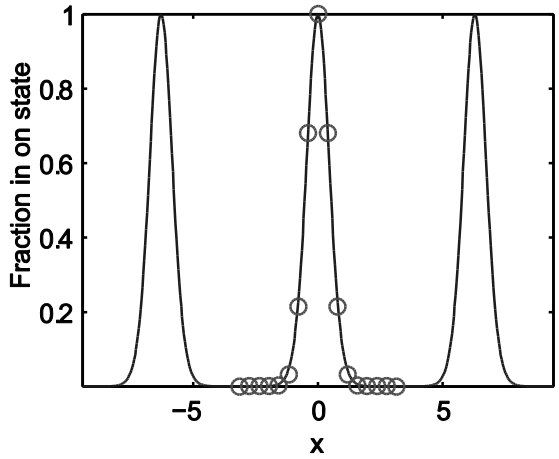

Fig. 3. Fraction of the fluorescent labels that remain in the *on* state after illumination with a sinusoidal pattern of *off* light at a saturation level $\eta = 10$. The solid curve follows the exact expression Eq. (42), and the circles the approximate expression Eq. (45).

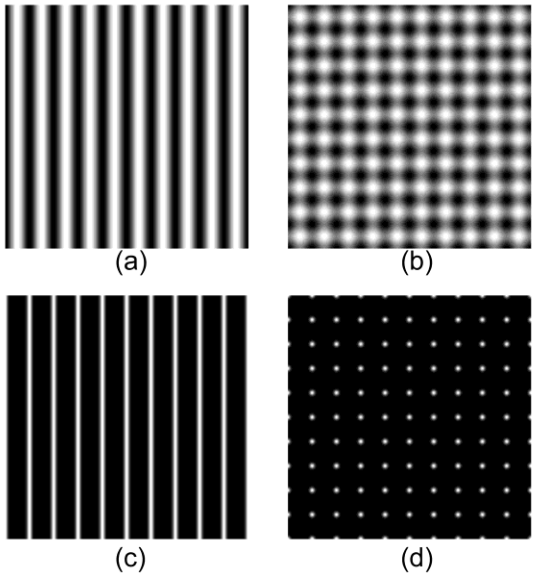

Fig. 4. The *off*-light illumination intensity pattern used in the 1D case (a), and the 2D case (b). The pattern of remaining fluorescent emission rate after illumination with the *off* light pattern for the 1D case (c) and the 2D case (d), at a saturation level η of 10. Fluorescent labels in the *on* state remain only near the intensity zeros of the *off* light pattern. At a given attempted resolution (width of the peaks), a much smaller fraction of the fluorophores remain *on* in the 2D case (d) than for 1D patterns (c).

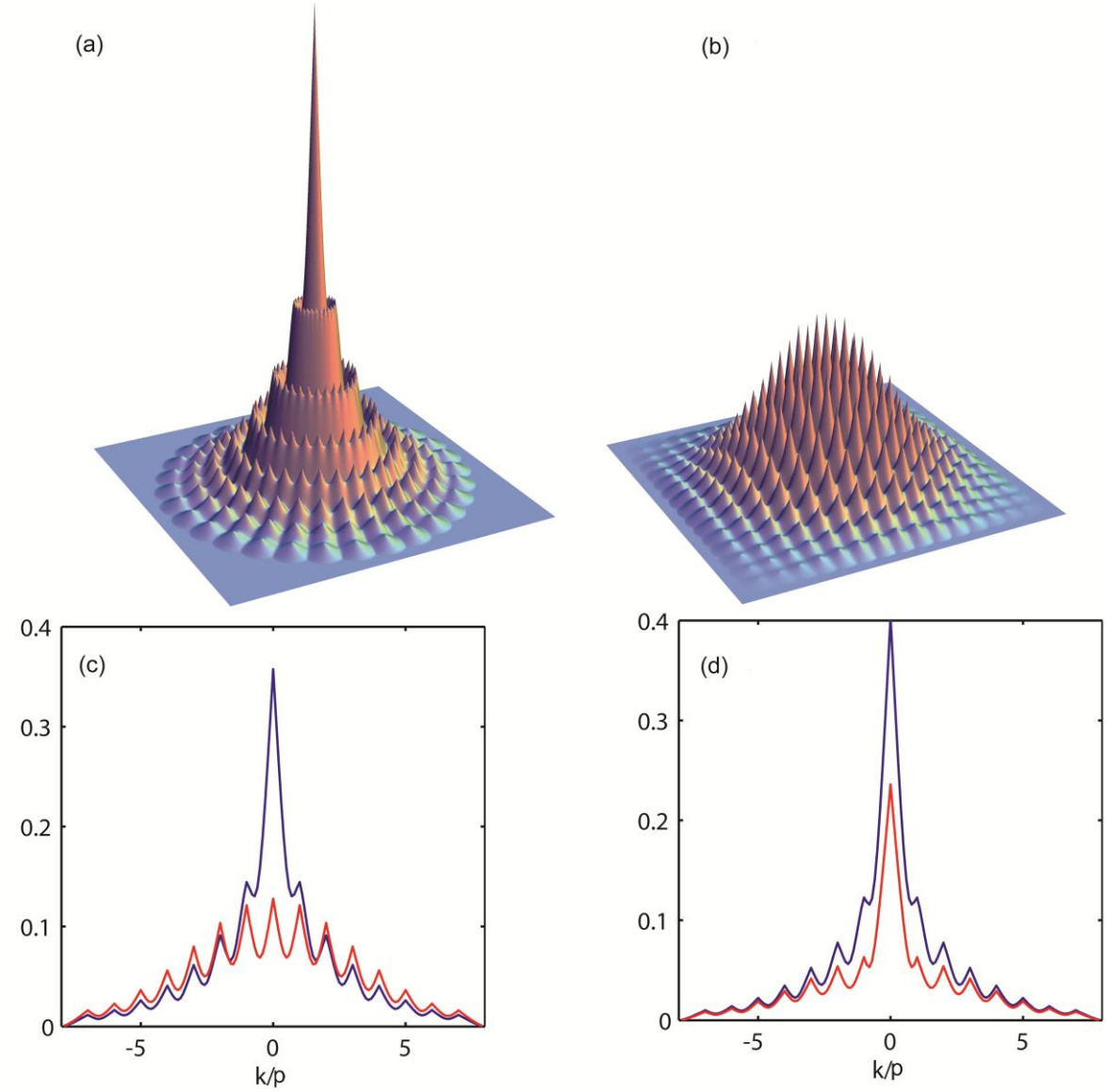

Fig. 5. Effective OTF (i.e., the dependence of the signal-to-noise ratio on spatial frequency for a point object) for nonlinear SIM with switchable labels, for a saturation factor $\eta = 20$ and a total of 225 images, using sequential 1D illumination patterns (a), or a 2D illumination pattern (b). Profiles through the center of the effective OTF for 1D illumination (blue line) and 2D illumination (red line) without background (c), and with 4% permanently fluorescent background (d). (The units on the vertical axes are normalized to the SNR at zero spatial frequency of the conventional microscopy that would result in the absence of *off* light.)

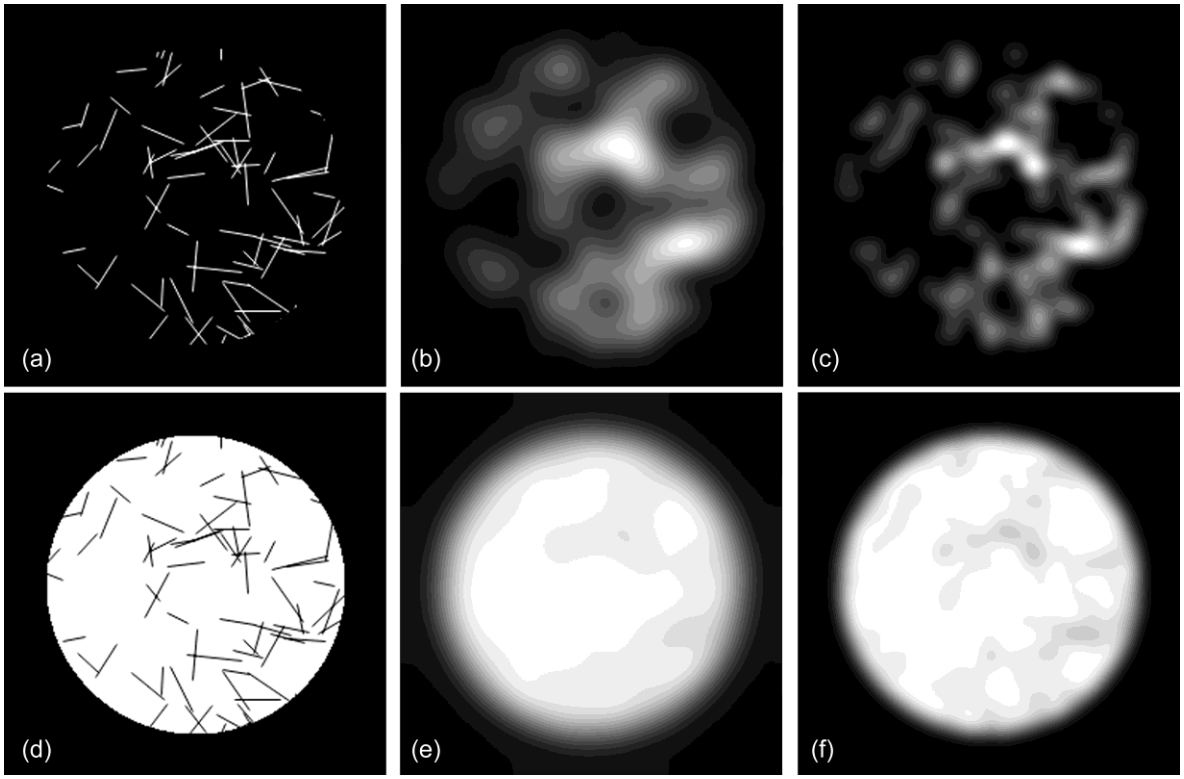

Fig. 6. Simulation of linear SIM. Test object with a dark background (a) and bright background (d). Simulated images of the test objects as observed by conventional microscopy (b, e), and by linear SIM with 1D patterns (c, f).

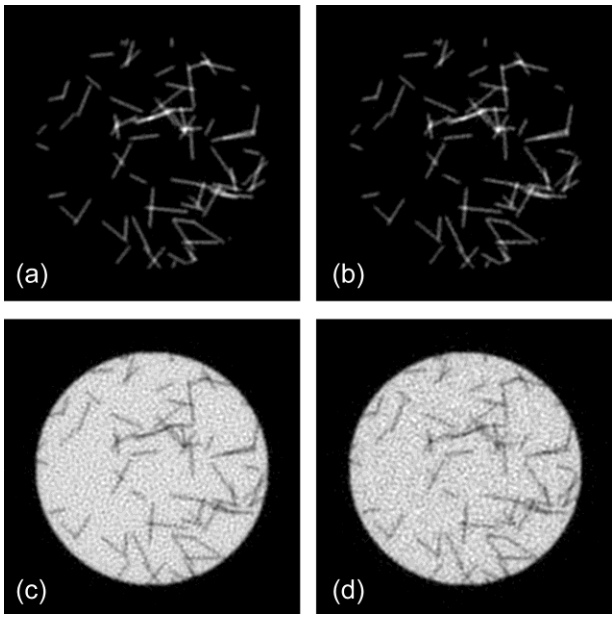

Fig. 7. Simulation of nonlinear SIM with the same test objects as in Fig. 6. Reconstructions of the test objects with dark background (a,b) and bright background (c,d), using 1D (a,c) and 2D (b,d) illumination patterns. The simulation used 225 raw images, a saturation factor η of 20, and a fraction of 4% *permanent-on* fluorophores. Each white 8-nm pixel in the object emitted 140 detectable photons per exposure when fully in the *on* state.

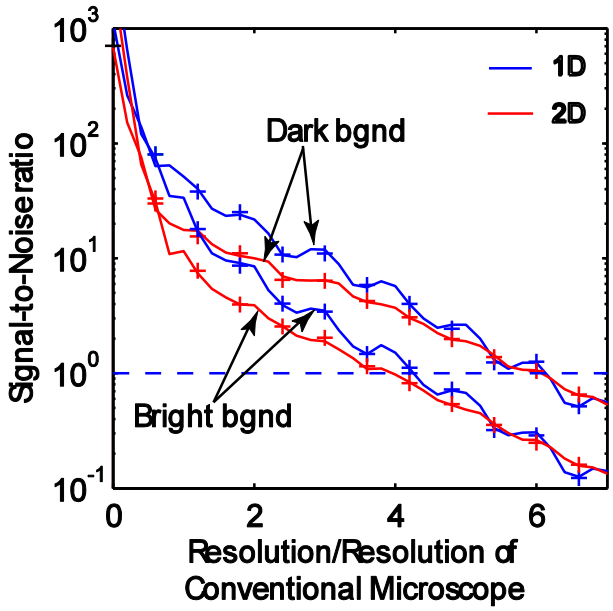

Fig. 8. Rotationally averaged signal-to-noise ratio as a function of spatial frequency. Results of numerical simulations are denoted by '+' marks, and analytical results by solid lines. The two top lines show the SNR for the test object with a dark background (Fig. 6a), the two bottom lines show the SNR for the test object with bright background (Fig. 6d). The blue lines show results for 1D illumination, the red lines for 2D illumination. Simulation parameters as in Fig. 7.

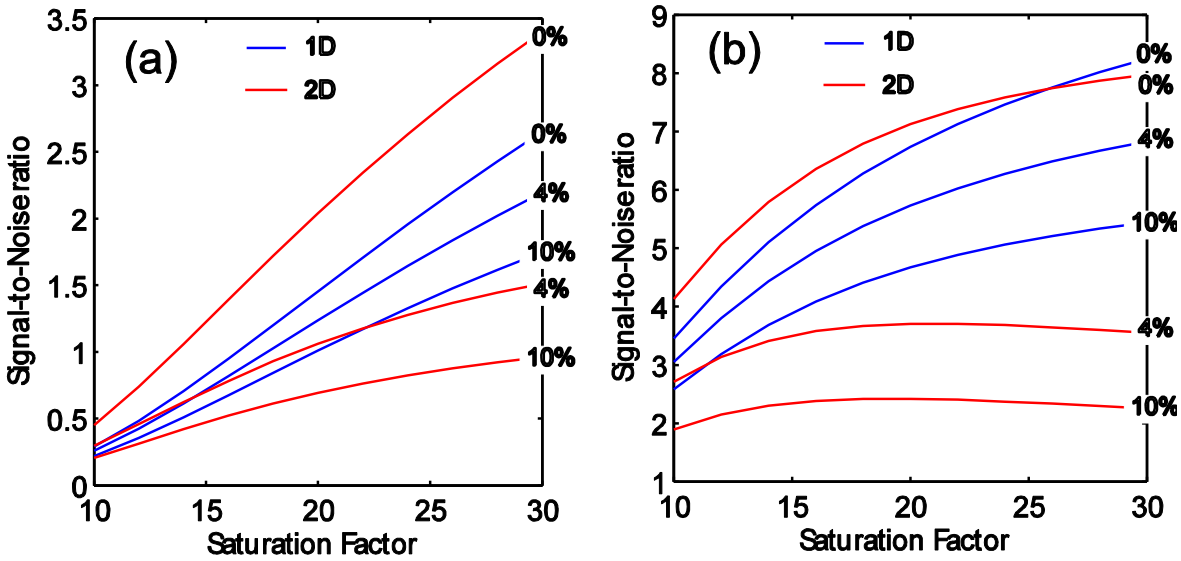

Fig. 9. Rotationally averaged signal-to-noise ratio for simulated reconstructions of the dark-background test object as a function of the saturation factor η, at a spatial frequency equal to (a) 3 times the resolution limit of conventional microscopy and (b) 6 times the resolution limit. Three values of the *permanent on* fluorophore fraction were used: 0%, 4% and 10% (as indicated); other parameters were as in Fig. 7. Blue lines denote 1D and red lines 2D illumination patterns.

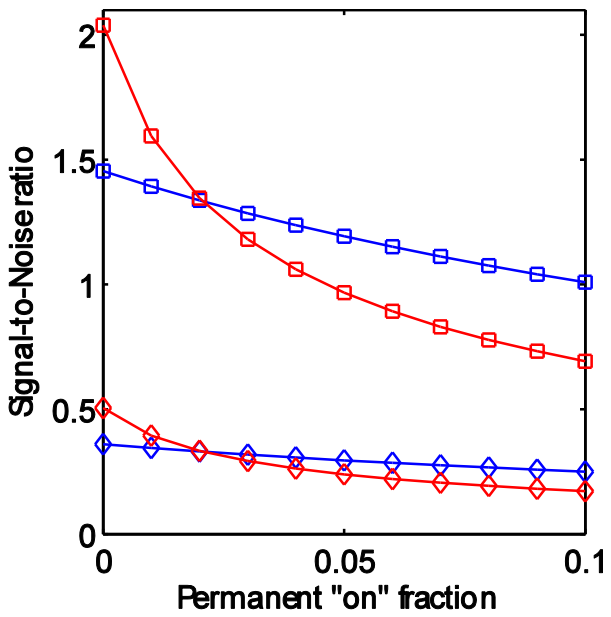

Fig. 10. Rotationally averaged signal-to-noise ratio as a function of the fraction of permanently fluorescent labels, at a spatial frequency equal to 6 times the resolution limit of the conventional microscope. Other simulation parameters were as in Fig. 7. Blue lines show SNR for the 1D illumination method, red lines for the 2D illumination method. Two upper lines represent the dark-background test object (Fig. 6a), the two lower lines the bright-background test object (Fig. 6d).

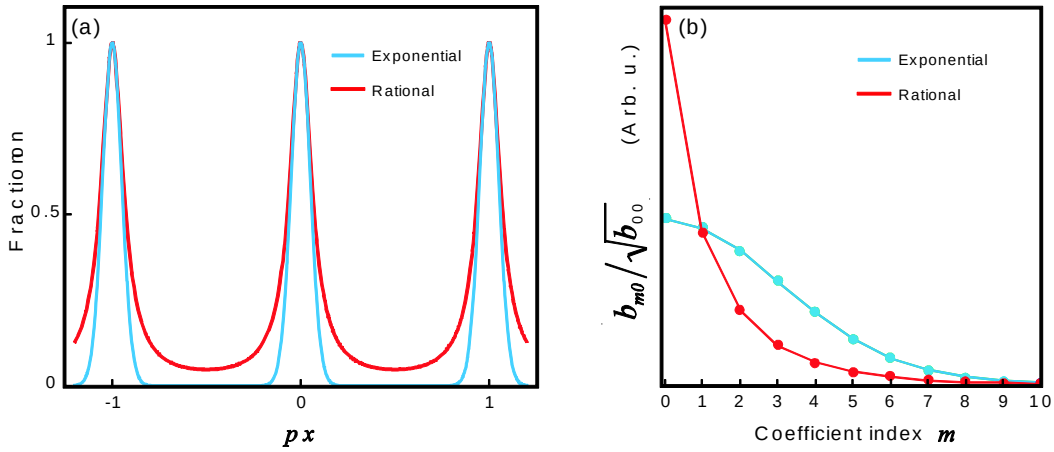

Fig. 11. Comparison of exponential [Eq. (42)] and rational [Eq. (55)] turn-off functions for photo-switchable labels. (a) The fraction of labels that remain in the *on* state as a function of position, under illumination by a sinusoidal *off*-light pattern with a maximum saturation factor η of 20 in the exponential case (cyan curve), and a maximum I_{off}/I_{on} of 20 in the rational case (red curve). (b) The ratio $b_{mn}/\sqrt{b_{00}}$ that determines the signal-to-noise ratio [see Eqs. (39) and (41)] for a 2D pattern with the above parameters, plotted vs. m for $n=0$. The rational function effectively leaves a background of labels in the *on* state, which raises the overall light level b_{00} and thereby decreases the signal-to-noise ratio at higher spatial frequencies. Lines are drawn between points only to guide the eye.