Superresolved multiphoton microscopy with spatial frequency-modulated imaging

Jeffrey J. Field[a,b,c,d,1,2], Keith A. Wernsing[a,1], Scott R. Domingue[a], Alyssa M. Allende Motz[e,f], Keith F. DeLuca[b,c,d], Dean H. Levi[f], Jennifer G. DeLuca[b,c,d], Michael D. Young[e], Jeff A. Squier[e], and Randy A. Bartels[a,c,d,g]

[a]Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523; [b]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523; [c]Microscope Imaging Network Foundational Core Facility, Colorado State University, Fort Collins, CO 80523; [d]Institute for Genome Architecture and Function, Colorado State University, Fort Collins, CO 80523; [e]Department of Physics, Colorado School of Mines, Golden, CO 80401; [f]National Renewable Energy Laboratory, Golden, CO 80401; and [g]School of Biomedical Engineering, Colorado State University, Fort Collins, CO 80523

Superresolved far-field microscopy has emerged as a powerful tool for investigating the structure of objects with resolution well below the diffraction limit of light. Nearly all superresolution imaging techniques reported to date rely on real energy states of fluorescent molecules to circumvent the diffraction limit, preventing superresolved imaging with contrast mechanisms that occur via virtual energy states, including harmonic generation (HG). We report a superresolution technique based on spatial frequency-modulated imaging (SPIFI) that permits superresolved nonlinear microscopy with any contrast mechanism and with single-pixel detection. We show multimodal superresolved images with two-photon excited fluorescence (TPEF) and second-harmonic generation (SHG) from biological and inorganic media. Multiphoton SPIFI (MP-SPIFI) provides spatial resolution up to 2η below the diffraction limit, where η is the highest power of the nonlinear intensity response. MP-SPIFI can be used to provide enhanced resolution in optically thin media and may provide a solution for superresolved imaging deep in scattering media.

superresolution | harmonic generation | multiphoton microscopy

Of the numerous far-field image formation techniques, multiphoton microscopy (MPM) has proved particularly valuable in many biological studies, providing multimodal images from deep within tissues (1), as well as label-free images of complex biological structures (2–4) and inorganic harmonic probes (5, 6). Deep imaging is achieved twofold in laser-scanning microscopy with nonlinear excitation. First, inherent optical sectioning provided by the nonlinear intensity dependence (7) permits removal of the confocal pinhole that is required in linearly excited microscopy, and highly scattered signal light can be collected with a single-pixel detector. Second, excitation wavelengths in the near infrared (NIR) (8, 9) enhance penetration depth of the ballistic photons responsible for nonlinear excitation (10). Because of the diffraction limit of light, these longer excitation wavelengths yield lower spatial resolution than wavelengths in the visible region of the spectrum, such as those required for confocal fluorescence microscopy. An imaging method that could meet or surpass the spatial resolution achievable with confocal fluorescence microscopy while maintaining the penetration depth of multiphoton microscopy could dramatically impact biological imaging.

The most successful superresolved microscopy techniques reported to date have exploited the real energy states of probe molecules to generate image contrast. These real energy states provide a means to manipulate fluorescent light emitted from probe molecules, offering a pathway to reducing the effective size of the point-spread function (PSF), which sets the spatial resolution of the microscope. Some examples include saturated structured-illumination microscopy (SSIM) (11), photoactivated localization microscopy (PALM) (12, 13), stimulated emission depletion (STED) microscopy (14–16), ground state depletion (GSD) microscopy (17, 18), saturation of transient absorption in

electronic states (19), and saturation of scattering due to surface plasmon resonance (20, 21).

Conversely, harmonic generation (HG) results in no net energy transfer to the contrast generating molecules. In this case, the photokinetics are often described via virtual energy states that drive instantaneous coherent nonlinear scattering. The application of most superresolution imaging methods to virtual energy states is not possible, including structured illumination microscopy (SIM) (22). Despite this challenge, examples of HG microscopy with subdiffraction-limited resolution have been reported (23, 24). In both cases, the polarization state of an illuminating laser pulse was manipulated to reduce the size of the effective PSF. Although both methods succeeded in collecting HG images with subdiffraction limited resolution, these methods are unlikely to be robust to imaging at depth in tissue because of effects such as birefringence, circular dichroism, and optical scattering, which disrupt the input polarization state of the excitation laser pulse.

Until now, a technique that provides superresolved imaging of both incoherent and coherent contrast mechanisms in complex media has not yet emerged. Here, we report the use of multiphoton spatial frequency-modulated imaging (MP-SPIFI) (25–28) for superresolved imaging of second-harmonic generation (SHG) and two-photon excited fluorescence (TPEF). We show

Significance

Superresolution microscopy is indispensable in biological sciences. The vast majority of superresolution imaging techniques exploit real energetic states of fluorescent molecules to break the diffraction limit. To date, superresolved imaging of second- and third-harmonic generation has been limited to specific sample preparations where the polarization state of the excitation laser can be manipulated to overcome the diffraction limit. Here, we describe a method for multiphoton superresolved imaging that does not place such restrictions on the sample and allows for simultaneous superresolved imaging of both coherent and incoherent signal light. Combined with single-element detection, this technique may allow for significant advances in multimodal multiphoton imaging of highly scattering biological tissues.

that MP-SPIFI can be used to collect images of both SHG and TPEF simultaneously and with spatial resolution beyond that of a conventional laser-scanning multiphoton microscope. By combining the relative scattering immunity offered by illumination wavelengths in the NIR with single-pixel detection of signal light, MP-SPIFI has the potential to collect images from unlabeled biological specimens that exhibit harmonic generation.

The physical principles of image formation in SPIFI microscopy have been rigorously studied in previous reports (25, 27). However, no reports to date have described imaging beyond the diffraction limit with SPIFI. Here, we describe the image formation process in SPIFI for both linear and nonlinear excitation of signal light generation.

The microscope configuration used for this work forms 1D images, similar to the systems used in refs. 25 and 27. Consequently, we constrain the theoretical analysis that follows to a 1D imaging system that collects image data in the x dimension. However, there is no physical constraint that prevents applying the principles described here to 2D SPIFI imaging systems such as the configurations described in refs. 26 and 29.

The spatial resolution of a microscope is determined by the set of spatial frequencies that the system can capture. In a conventional imaging system, the range of spatial frequencies that can be collected is limited by the numerical aperture (NA) of the objective lens and the wavelength of light. For a coherent illumination source of wavelength λ, the largest spatial frequency passed by the objective lens is $f_{x,c} = NA/\lambda$. Assuming that all spatial frequencies are passed by the objective lens with equal amplitude, the spatial resolution for a coherent imaging system is $\delta x = \lambda/(2\,NA) = (2\,f_{x,c})^{-1}$. Acquiring superresolved images is equivalent to collecting specimen information beyond this range of spatial frequencies. Thanks to the spatiotemporally varying illumination intensity used to generate signal light, spatial frequency information beyond the conventional limit is measured with SPIFI under both linear and nonlinear excitation.

One-dimensional images are formed in SPIFI microscopy by illuminating the specimen with a light sheet containing a temporally varying spatial interference pattern along the line focus (i.e., in the x dimension for a light sheet focused in the y dimension). The intensity of a simulated SPIFI light sheet at a snapshot in time is shown in Fig. 1A, whereas the intensity squared, which

represents the probability of contrast light generation for a second-order nonlinear process, is shown in Fig. 1B. To form a 1D image of the object, the spatial frequency of the fringes in the illumination intensity, which is inversely proportional to the separation of adjacent fringes (Fig. 1C), is varied in time through the range of spatial frequencies supported by the objective lens, i.e., $f_x(t) \in [-NA/\lambda, +NA/\lambda]$ (Movie S1). Signal light generated by the interaction of the illuminating light sheet and the specimen is then collected with a single-pixel detector.

The measured signal from the detector, $S(t)$, is represented by integrating the signal light intensity, $\beta(x,t)$, over all space: $S(t) = \int_{-\infty}^{\infty} dx\,\beta(x,t) = \langle\beta(x,t)\rangle_x$. In linearly excited SPIFI, the signal light is proportional to the product of the illumination intensity, $I(x,t)$, and the spatial distribution of contrast-generating molecules, $C(x)$. Note that we have assumed $C(x)$ is stationary and does not photobleach during the measurement of a single 1D image. Absorbing proportionality constants such as the absorption cross section and quantum efficiency into $C(x)$, we can write $\beta(x,t) = I(x,t)\,C(x)$.

The spatiotemporally varying light sheet, $I(x,t)$, is generated by forming a line focus on a spinning modulation mask and reimaging the mask plane to the object region with a tube lens and an objective lens (25). Schematics of the illumination system are shown in Fig. S1 A and B. As the spinning disk (Fig. S1C) rotates, the illuminating line focus is diffracted into multiple directions at an angle with respect to the optic axis of the microscope that depends on the rotation angle of the mask. In the pupil plane of the objective lens, the diffracted beams scan across the pupil horizontally with time, whereas the undiffracted beam propagates colinearly with the optic axis of the objective lens (Fig. S1D). The microscope is configured such that the diffracted beams scan across the entire pupil of the objective lens, thereby "filling" the pupil of the objective lens in the x direction through a single rotation of the modulation mask (Movie S2).

In the ideal case, the spatially coherent illuminating light sheet is formed by the spatial interference of three light sheets—one being the undiffracted beam and the other two being the temporally varying diffracted beams. At the focus of the light sheet, the interference of these three beams generates a spatially modulated line focus of the form $I(x,t) = H_0(t) + H_1(t)\cos[2\pi\,f_x(t)x] + H_2(t)\cos[2\pi\,(2f_x(t))x]$, where the time-varying amplitudes arise from vignetting caused by the circular pupil, as well as depolarization (Supporting Information), and $f_x(t)$ is the spatial frequency of the interference fringes in the object plane. The first spatially modulated component of the intensity, $I_1(x,t) = H_1(t)\cos[2\pi\,f_x(t)x]$, results from spatial interference of the undiffracted beam with each diffracted beam, both of which propagate at a time-dependent angle $\theta(t)$ with respect to the optic axis (Fig. S1A). Note that the maximal crossing angle between each diffracted beam and the undiffracted beam is the semiaperture angle of the lens, α, and that $NA = n_2 \sin\alpha = \lambda\,f_{x,c}$, where n_2 is the index of refraction in the object region. The second spatially modulated term, $I_2(x,t) = H_2(t)\cos[2\pi\,(2f_x(t))x]$, originates from the interference of the two diffracted beams with one another, where the crossing angle is always $2\theta(t)$, and as a result, the spatial frequency of $I_2(x,t)$ is twice that of $I_1(x,t)$ at all times (Movie S3). Here, the maximal crossing angle of the beams is 2α.

In SPIFI, spatial frequencies greater than $2f_{x,c}$ are produced in the specimen by nonlinearities in the optical response of contrast generating molecules (e.g., fluorophores). These nonlinearities could arise from a number of physical effects, such as saturation of electronic absorption (11). In general, the intensity of the signal light is proportional to a function of the illumination intensity, i.e., $\beta(x,t) = f[I(x,t)]\,C(x)$ (30). In this work, we concentrate on multiphoton-excited SPIFI where the signal light intensity is of the form $\beta(x,t) = I^\eta(x,t)\,C(x)$. In this case, the effective illumination intensity driven by the nonlinear specimen
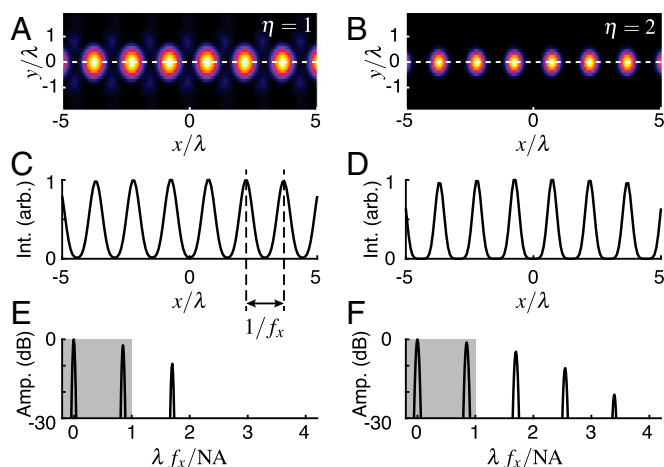
Fig. 1. The intensity in the (x,y) plane at the focus of the light sheet for linear (A) and second-order nonlinear (B) excitation. One-dimensional line-outs of the intensity at $y = 0$ are shown in C and D, respectively. Fourier transforms of the 1D intensities in C and D are shown in E and F, respectively. A Fourier transform of the 1D intensity for $\eta = 2$ reveals harmonics up to $4f_x$. The shaded gray areas represent the spatial frequency region for conventional imaging with coherent illumination.

response is a sum of $2\eta + 1$ terms, consisting of a background intensity that contains no interference fringes, and 2η interference terms whose spatial frequencies are harmonics of one another, increasing from $f_x(t)$ through $2\,\eta\,f_x(t)$ (Movie S4).

By projecting spatial frequencies onto the object up to 2η times the conventional limit, it follows that MP-SPIFI can, in principle, provide spatial resolution up to 2η smaller than the diffraction limit of a conventional imaging system under linear excitation. To illustrate how SPIFI encodes spatial frequency information from a single-pixel detector, we first examine the signal $S_1(t)$, which can be expressed as $S_1(t) = \langle I_1(x,t)\,C(x)\rangle_x = (1/2)[\tilde{S}_{1+}(t) + \tilde{S}_{1-}(t)]$, where

$$\tilde{S}_{1\pm} = \int_{-\infty}^{\infty} dx\, H_1(t)\, e^{\pm i 2\pi f_x(t)x}\, C(x) = H_1(t)\tilde{C}[\pm f_x(t)], \qquad [1]$$

and $\tilde{C}(f_x)$ is the Fourier transform of $C(x)$. Eq. **1** shows the key feature of SPIFI microscopy, namely that at each snapshot in time, an intensity pattern with spatial frequency $f_x(t)$ is projected onto the specimen, and collection of the signal light with a single pixel detector forms a projection of that spatial frequency onto the object. The signal measured from the detector thereby encodes a measurement of the spatial frequency content of the object modulated by the temporal transfer function $H_1(t)$. A 1D image in space is recovered by digitally isolating either $\tilde{S}_{1+}(t)$ or $\tilde{S}_{1-}(t)$ and applying an inverse Fourier transform.

Because the spatial frequency $f_x(t)$ is varied through the range supported by the objective lens, $\tilde{S}_{1\pm}(t)$ represents a measurement of $\tilde{C}(f_x)$ over this conventional range. However, higher-order MP-SPIFI signals $S_{q>1}(t)$ correspond to measurements of the spatial frequency content of the specimen over a range q-times greater than the conventional limit. In general, we can write the complex q^{th}-order signal as

$$\tilde{S}_{q+}(t) = H_q(t)\int_{-\infty}^{\infty} dx\, e^{i 2\pi q f_x(t)x}\, C(x) = H_q(t)\tilde{C}[q f_x(t)]. \qquad [2]$$

Because $f_x(t) \propto t$, this expression can be recast into the spatial frequency domain to find $\tilde{S}_{q+}(f_x) = H_q(f_x)\tilde{C}(q f_x)$, where we recognize $H_q(f_x)$ as the optical transfer function (OTF) of the microscope. In the absence of aberrations, $H_q(f_x)$ is real-valued and is therefore a modulation transfer function (MTF) that determines the amplitude of each spatial frequency component. $\tilde{S}_{q+}(f_x)$ has the form of a linear shift invariant (LSI) imaging system, and the qth image can be expressed as a convolution in

the spatial domain as $S_{q+}(x) = h_q(q\,x)\otimes C(x)$, where $h_q(x)$ is an order-dependent PSF that decreases in spatial extent with increasing SPIFI order q.

The real-valued SPIFI signal collected from the photodetector is a linear combination of the $2\eta + 1$ signals arising from the various interference terms projected onto the specimen by the three illumination beams: $S(t) = \sum_{q=0}^{2\eta} S_q(t)$. Each component $S_{q\neq0}(t)$ corresponds to a 1D image in the x direction with varying spatial frequency content, whereas $S_0(t)$ is a measure of the DC spatial frequency content of the specimen. To reconstruct the qth-order image, the corresponding temporal signal $S_q(t)$ must be isolated from $S(t)$. Fortunately, the design of the spinning modulation mask provides a simple means to isolate each image without requiring multiple measurements of $S(t)$.

The mask is designed to modulate the intensity of the incident line focus at a temporal frequency that depends linearly on x (25). We denote the centroid of the line focus incident on the mask as x_c, which corresponds to a carrier frequency ν_c. The beams diffracted from the modulation mask acquire a time-dependent linear phase, $\exp[i 2\pi f_x(t)x_c]$. This linear phase is accumulated in each harmonic order of the illumination intensity, such that the qth-order component of the intensity contains a linear phase shift: $I_q(x,t) = H_q(t)\cos[2\pi q f_x(t)\,x + 2\pi q f_x(t)x_c]$. Accordingly, the interference fringes in the focal plane of the light sheet shift laterally as a function of disk rotation (Movie S2). Because spatial frequency is linearly proportional to scan time, we can recast this linear phase in the temporal domain and write the intensity as $I_q(x,t) = H_q(t)\cos[2\pi q f_x(t)x + 2\pi q \nu_c\, t]$. Physically, this order-dependent modulation frequency arises from the density of the fringes in each intensity component as they shift across the lateral field of view (Movie S4). Practically, this means that a point emitter at any point along the line focus will be modulated at harmonics of the carrier frequency (Movie S5), and thus each complex component of the measured signal can be separated with standard Fourier decomposition methods.

Including the temporal carrier frequency, the signal measured from the photodetector is

$$S(t) = S_0(t) + \sum_{q=1}^{2\eta} H_q(t)\mathrm{Re}\{\exp(i 2\pi q \nu_c\, t)\tilde{C}[q f_x(t)]\}. \qquad [3]$$

Because $S_q(t)$ is real-valued, the Fourier transform of the signal displays conjugate symmetry and thus two redundant images are
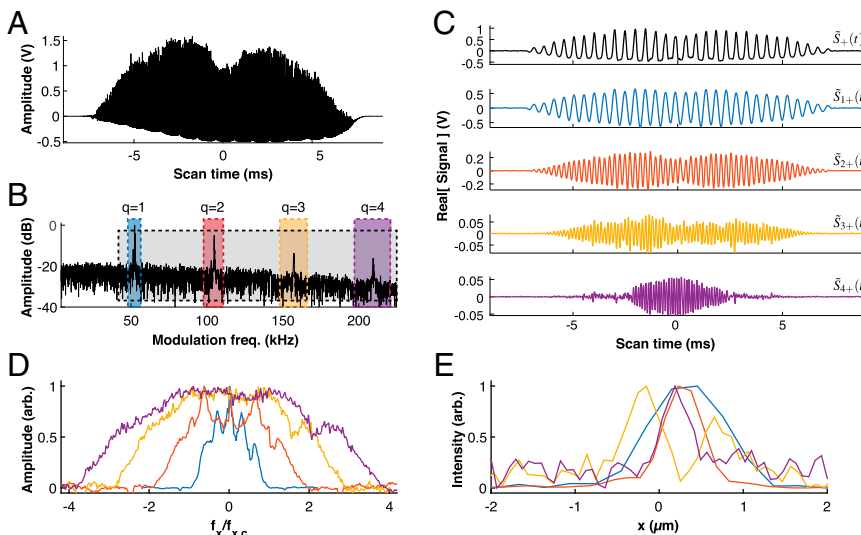
Fig. 2. MP-SPIFI provides increased lateral spatial frequency support. (*A*) The average of 1,000 temporal measurements of SHG from a 200-nm-diameter BTO particle. Each temporal scan was collected at 31.6 Hz for a total acquisition time of 31.6 s. Note that there is no DC offset S_0 (t) because the amplifier was set to AC coupling. (*B*) A Fourier transform of the mean signal with respect to scan time reveals each MP-SPIFI image encoded at a different carrier frequency. (*C*) The complex temporal trace, $\tilde{S}_{q+}(t)$, was recovered for each SPIFI order by digitally filtering each harmonic in the modulation frequency domain with a square window, as shown in *B*. To obtain $\tilde{S}_+(t)$, the digital filter was set to retain all four image orders. All five complex traces were demodulated such that the fundamental carrier frequency is 2 kHz as opposed to ∼52 kHz. (*D*) Experimental MTFs for each MP-SPIFI image measured from a 100-nm-diameter FND. The qth-order image spans the spatial frequency range $[-q\,f_{x,c}, q\,f_{x,c}]$. (*E*) PSFs for each MP-SPIFI order. The spatial resolution computed from each PSF are listed in the main text. FND data were collected with the same temporal acquisition parameters as the SHG measurements.

Fig. 3. TPEF and SHG images collected from biological media with MP-SPIFI. (*A* and *B*) First-order (*A*) and second-order (*B*) MP-SPIFI images of TPEF from a mitotic HeLa cell immunostained with primary antibodies against alpha tubulin and secondary antibodies tagged with Alexa 546. (Scale bar: 3 µm.) (*C–F*) First- through fourth-order images of SHG from fixed, 16-µm-thick rabbit tendon. (Scale bar: 10 µm.) All images were collected at 0.8 NA with laser pulses centered at 1,065 nm. TPEF was collected in the epidirection and SHG was collected in the forward-scattered direction. Both datasets were captured at 31.6 Hz. The HeLa image is the average of 41 images for a total collection time of 58.4 s, whereas the images of rabbit tendon are averages formed from 1,000 images for a total collection time of 665 s.

encoded, centered on $\pm q\,\nu_c$. In practice, we select the image at positive modulation frequencies, for which the qth-order image is $\tilde{S}_{q+}(t) = \exp(\mathrm{i}\,2\pi\,q\,\nu_c\,t)\,H_q(t)\,\tilde{C}[q\,f_x(t)]$. Isolation of each image is possible thanks to the carrier frequency, which separates the image information in the modulation frequency domain, as we show in Fig. 2.

To demonstrate that MP-SPIFI provides images with frequency support beyond the diffraction limit for contrast mechanisms with both real and virtual energy states, we imaged TPEF from 100-nm-diameter fluorescent nanodiamonds (FNDs) and SHG from 200-nm-diameter barium-titanate oxide (BTO) crystals. The objects were illuminated with femtosecond laser pulses centered at a wavelength of 1,065 nm and focused by a 0.8 NA objective lens. For both TPEF and SHG, the predicted spatial resolution for MP-SPIFI order $q = 1$ through $q = 4$ are 666 nm, 333 nm, 222 nm, and 167 nm, although these estimates assume a uniform MTF.

Because the BTO crystals were larger than the predicted spatial resolution of the fourth-order MP-SPIFI image, we used FNDs to determine the spatial resolution of the imaging system. Fig. 2*D* shows the recovered MTF for each SPIFI order, scaled to the lateral spatial frequency domain (*Materials and Methods*). It is clear that lateral spatial frequency information beyond the conventional limit is contained in all MP-SPIFI orders for $q > 1$. The corresponding PSFs are displayed in Fig. 2*E*. Using the Raleigh criterion for spatial resolution, we determined the spatial resolution for the first-, second-, and fourth-order PSFs to be 949 nm, 491 nm, and 362 nm, respectively (Fig. S2). We note that the third-order PSF suffers from a dual-lobe structure, where each lobe of the third-order PSF has spatial resolution of 468 nm. This modulated PSF arises when higher-order diffracted beams from the binary modulation mask interfere with the first- and zero-order beams in the object plane and can be controlled with careful design of the binary modulation mask (31).

We compared the experimental resolution performance to theoretical PSFs computed with a vector-focusing model of MP-SPIFI that accounts for only first-order diffraction from the modulation mask (*Supporting Information*). This model accounts for vignetting of the beams that scan across the pupil of the objective lens, as well as depolarization of the scanning beams,
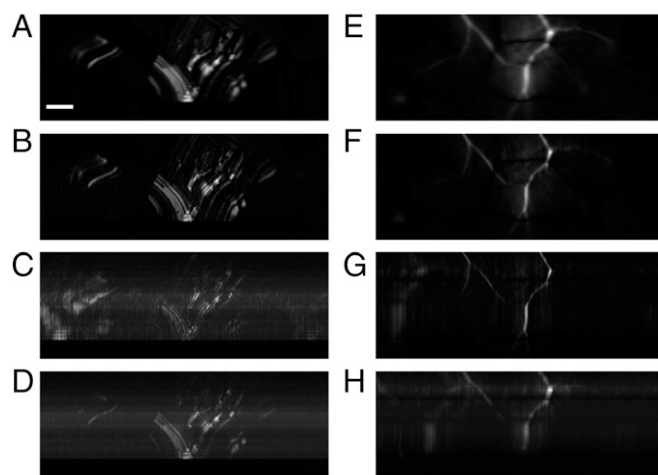
which are responsible for deviations from a uniform MTF. Using this model, we computed the spatial resolution at the half-width at 10.4% maximum for first- through fourth-order MP-SPIFI images to be 818 nm, 472 nm, 300 nm, and 240 nm, respectively.

We demonstrated superresolved multiphoton imaging of biological media by imaging TPEF from polymerized tubulin fibers (i.e., microtubules) tagged with Alexa 546 in a mitotic HeLa cell and SHG from collagen fibers in fixed rabbit tendon (Fig. 3). One-dimensional MP-SPIFI images were collected in the horizontal dimension as the specimen was scanned vertically to form a 2D image. Resolution enhancement occurs only in the horizontal dimension in this illumination scheme. Lateral tomography has been used for 2D image capture with SPIFI and could be applied to MP-SPIFI for superresolved 2D imaging (26).

To demonstrate multimodal superresolved imaging with MP-SPIFI, we simultaneously imaged photoluminescence (PL) and SHG from CdTe solar cells (Fig. 4). Because of the opacity of the CdTe cells, both contrast mechanisms were collected in the epidirection. Fig. 4 *A–D* display the first- through fourth-order SHG-SPIFI images, and the corresponding PL-SPIFI images are shown in Fig. 4 *E–H*. Comparison of the images for each MP-SPIFI order makes it clear that the spatial resolution is enhanced in both SHG and PL simultaneously, although the resolution appears to be different for these two contrast modes beyond the second-order data. The reason for this discrepancy is unclear at this time, although we note that the relatively long lifetime of the PL contrast mechanism, as well as diffusion of excited carriers, may be responsible for these differences.

Although the spatial frequency support attainable for TPEF and SHG imaging with MP-SPIFI is four times the conventional limit, images shown in this work do not reach this theoretical limit. For example, TPEF-SPIFI images obtained from tubulin in HeLa cells were limited to a resolution enhancement of 2×. The loss of spatial frequency information in higher-order MP-SPIFI images was attributed to a reduction in signal-to-noise ratio (SNR) with increasing MP-SPIFI order. This reduction in SNR results from decreased contrast in the fringes of the illuminating light sheet at high spatial frequencies due to the MTF for each order and shot noise contributions from all MP-SPIFI orders that raises the noise floor of the measured signal.

The order-dependent MTF results primarily from vignetting of the illumination beams due to the circular objective pupil (Movie

Fig. 4. Multimodal superresolved images of SHG and PL from CdTe solar cells. (*A–D*) First- through fourth-order SHG images. (*E–H*) Corresponding PL images. Images presented here are averages of 500 images acquired at 31.6 Hz for a total acquisition time of 1,282 s. (Scale bar: 10 µm.)

S5). Tight focusing in the y dimension requires that the objective lens be filled by the illumination light in the pupil plane. Conversely, the spatially modulated line focus in the x dimension necessitates tightly focused beams in the pupil plane. Thus, highly elliptical beams are needed in the pupil plane to form the spatiotemporally modulated light sheet (Fig. S3). Because the beams are highly elliptical, vignetting of the two diffracted beams in the pupil plane of the objective lens causes a reduction in both transmitted energy and spatial frequency support in the vertical (y) dimension. We investigated the effect of the circular objective pupil for each MP-SPIFI order using numerical simulations.

First, the throughput energy of the beams with scan time was computed with an overlap integral (Fig. S4). The energy loss experienced by the diffracted beams causes a reduction in the modulation depth of the illumination intensity pattern, reducing the amplitude of each MP-SPIFI order with respect to the background DC signal.

We also investigated the effect of the circular pupil aperture on the spatial frequency support of the diffracted beams in the vertical dimension with a numeric calculation of the electric fields in the object region using the angular spectrum formalism (32, 33) (Supporting Information). The numeric model confirms reduced high spatial frequency support due to vignetting of the diffracted beams in the pupil plane (Fig. S5).

Finally, depolarization as a function of scan time can cause a reduction in the contrast of the illumination intensity fringes and is dependent upon the polarization state of the beams in the pupil plane and the polarization dependence of the contrast mechanism. We used the angular spectrum formalism to investigate depolarization in the focal plane of the illumination light sheet for linearly polarized input light (Fig. S6). We find not only that the depolarization can cause reduction in the amplitude of the MTF with scan time but that the shape of the MTF depends on the input polarization state.

Along with vignetting, shot noise reduces the SNR of each MP-SPIFI order. The noise floor in the modulation frequency (spatial) domain is determined by the shot noise from the time-averaged total light intensity measured on the photodetector, $\langle \beta(x,t) \rangle_x$. Conversely, each MP-SPIFI order represents only a portion of this total light intensity. Because the shot-noise-limited noise floor scales as $\overline{\langle \beta(x,t) \rangle_x}$, the SNR of higher-order MP-SPIFI images scales as $\langle \beta_q(x,t) \rangle_x / \overline{\langle \beta(x,t) \rangle_x}$. This effect has also been observed in fluorescence imaging using radiofrequency-tagged emission (FIRE) (34), another imaging method that uses spatial frequency-modulated illumination to form images, and the spatial analog has been analyzed in detail for off-axis SHG holography (35). Images of TPEF collected from HeLa cells truncated at second-order due to a large noise floor that restricted our ability to collect higher-order images. On the other hand, images of SHG from rabbit tendon had a significantly larger SNR due to increasing the power of the illumination laser by ∼ 50%, making the fourth-order signal detectable above the shot noise floor. A possible route to enhance the SNR in higher-order MP-SPIFI images is the use of a phase modulator in place of the amplitude modulating disk (Fig. S1). This modification can reduce the overall noise floor by allowing the fourth-order MP-SPIFI image to be recovered using only the two diffracted beams.

The ability to resolve fine spatial features in an image depends on both the shape and the extent of the spatial frequency support of the OTF. To evaluate the relative imaging performance of MP-SPIFI compared with conventional laser-scanning microscopy (LSM) techniques, the OTFs of these methods were computed using the vector-focusing numerical model for two-photon SPIFI (2P-SPIFI), and two- and three-photon LSM. Intensities were computed for an illumination wavelength of 1,065 nm, 0.8 NA, and linearly polarized light aligned with the vertical (y) direction.

For 2P-SPIFI, the widths of the intensity distributions in the pupil plane were modeled after the MP-SPIFI system reported here (Supporting Information). For the multiphoton LSM calculation, the beam size in the vertical dimension of the 2P-SPIFI microscope was used as the radius of the circularly symmetric Gaussian distribution, resulting in a fill factor of ∼2 (33). The lateral spatial frequency support was then computed from the illumination intensity for each imaging type. The resulting spatial frequency support is shown in Fig. 5, where the shaded gray regions indicate the spatial frequency support of each LSM mode, whereas solid lines indicate the spatial frequency support for each 2P-SPIFI order.

What is clear in Fig. 5 is that MP-SPIFI imaging collects significantly more information content (high SNR) at high spatial frequencies than multiphoton LSM, thus providing the ability to extract significantly more spatial information. As the scan time progresses in 2P-SPIFI, the focusing of the transmitted diffracted orders is perturbed by the circular objective pupil, which shapes the spatial frequency support of the OTF. The MTFs associated with 2P-SPIFI extend to greater spatial frequency extent for the second-order nonlinear interaction than either two- or three-photon LSM. Moreover, the amplitude of the high spatial frequency content for 2P-SPIFI is significantly higher than for multiphoton LSM. The result is that 2P-SPIFI captures significantly more fine spatial information about an image than multiphoton LSM.

In summary, we have described a previously unidentified approach to superresolved imaging in multiphoton excited microscopy that is capable of measuring spatial frequency information beyond the diffraction limit for both multiphoton fluorescence and coherent nonlinear scattering. The combination of excitation wavelengths in the NIR and single-pixel detection should allow MP-SPIFI to form images with enhanced spatial resolution at penetration depths well beyond those attainable with linearly excited fluorescence. MP-SPIFI enables superresolved imaging of new contrast mechanisms, which may lead to new insights in both biological and material science studies.

Materials and Methods

SPIFI Microscope. The microscope system was illuminated by pulses from a nonlinear fiber amplifier centered at 1,065 nm (36), seeded by an all-normal dispersion fiber oscillator using Yb-doped gain fiber (37). Temporal compression in a folded Martinez compressor resulted in near transform-limited durations of ∼150 fs at an average power ranging from 900 mW to 1.3 W and a repetition rate of 52.5 MHz. The beam from the laser was collimated and expanded to a beam size of 8.85 mm (full-width at half-maximum of the intensity). The beam was brought to a line focus on the modulation mask with an achromatic cylindrical lens (ACY-254-150-B; ThorLabs) oriented such that the beam is focused in the vertical (y) direction only. The mask plane was reimaged to the object plane in two stages, first with a 1:1 image relay system composed of two achromatic lenses (AC254-100-C-ML; ThorLabs) in a 4-f configuration and again with an image relay system consisting of a tube lens (AC254-150-C-ML; ThorLabs) and an infinity-corrected objective lens. Data presented in this work were collected with a 0.8-NA objective lens (N-Achroplan 50×/0.8 NA Pol; Zeiss), such that the mask is imaged to the object region with de-magnification of 46.5×. The SPIFI modulation mask was described in previous work (25, 26) and consists of a pattern defined in
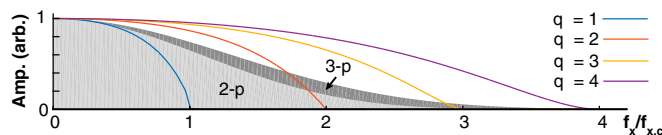

Fig. 5. Comparison of the MTF in multiphoton laser-scanning microscopy and 2P-SPIFI. The shaded gray regions indicate the frequency support for two- and three-photon laser-scanning microscopy. Solid lines indicate the frequency support in 2P-SPIFI. All OTFs were computed with the Debye integral using $\lambda = 1,065$ nm and 0.8 NA.

polar coordinates by the expression $m(r, \varphi) = 1/2 + 1/2 \, \text{sgn}[\cos(\Delta k \, r \, \varphi)]$. The modulation mask used in this work had a density parameter $\Delta k = 70/\text{mm}$. Imperfections in disk mounting lead to an additional phase accumulated in $S(t)$ as a function of mask rotation. The process used to remove disk aberration phase has been discussed at length in a previous study (38). All images collected in the epidirection used a long-pass dichroic beam-splitter placed between the tube lens and the objective lens (FF875-Di01; Semrock). Fluorescent light from the HeLa specimen was collected in the epidirection through a short-pass filter (FF720/SP; Semrock) and a 500-nm long-pass filter (ET500lp; Chroma).

Samples. Samples used for point spread function measurements were 200-nm BaTiO$_3$ particles (1148DY; Nanostructured & Amorphous Materials) for SHG and 100-nm FNDs (ND-400NV-100nm-10mL; Adamas Nanotechnologies) for TPEF. In both cases the insoluble nanoparticles, suspended in water, were diluted, disaggregated in an ultrasonic bath, drop cast onto a microscope slide, and cover-slipped after the water had evaporated. HeLa cells were seeded onto glass coverslips and fixed when they reached ∼70% confluency. Cells were rinsed rapidly with 37 °C 1× PHEM buffer (60 mM Pipes, 25 mM Hepes, 10 mM EGTA, 4 mM MgSO$_4$, pH 7.0), followed by lysis at 37 °C for 5 min in freshly prepared lysis buffer (1× PHEM plus 0.5% Triton X-100). Cells were then fixed on the bench top for 3 min using ice-cold methanol (95% methanol plus 5 mM EGTA), followed by an additional 20 min of methanol fixation at 20 °C. At room temperature, cells were then rehydrated with 1× PHEM, then rinsed 3× 5 min in PHEM-T (1× PHEM plus 0.1% Triton X-100) and blocked in 10% (vol/vol) boiled donkey serum (BDS) in PHEM for 1 h at room temperature. Microtubules were labeled with anti-alpha tubulin primary antibodies (Sigma-Aldrich) diluted in 5% (vol/vol) BDS for 12 h at 4 °C. Following primary antibody incubation, cells were rinsed 3× 5 min in PHEM-T and then incubated for 45 min at room temperature with secondary antibodies conjugated to Alexa 546 (Life Technologies). Cells were then rinsed 3× 5 min in PHEM-T, incubated in a solution of 2 ng/mL 4,6-diamidino-2-phenylindole (DAPI) diluted in PHEM, rinsed again (3× 5 min), and then mounted onto microscope slides in an anti-fade solution containing 90% (vol/vol) glycerol and 0.5% N-propyl gallate. Coverslips were sealed with nail polish to affix them to the slides. The polycrystalline, large-grained CdTe sample was grown by close-spaced sublimation (CSS), treated with CdCl$_2$, and polished, effectively "sealing" small grains (∼2 µm) after growth. The sample was polished using an ion beam technique.

Data Analysis. Data analysis was performed in MATLAB (MathWorks). Each 1D line image was reconstructed by computing the fast Fourier transform (FFT) of the measured temporal data $S(t)$ to obtain the MP-SPIFI images in the modulation frequency domain, i.e., $\bar{S}(\nu_t)$. Each MP-SPIFI order was recovered by numerically filtering at the positive carrier frequency. The filtered data, $\bar{S}_q(\nu_t)$, was inverse transformed to obtain $S_{q+}(t)$. The complex temporal trace was then demodulated by the relative carrier frequency, $q\,\nu_c$, and the temporal axis was calibrated according to the relation $f_{x,2}(t) = q\,\kappa\,t$ to obtain $S_{q+}(f_{x,2})$. The coupling coefficient κ between modulation frequency and space was calibrated by measuring light transmitted through a 2-µm slit in the object plane as a function of lateral position (25). The centroid of the measured distribution was computed as a function of lateral position. This measurement allowed the quantity describing the relationship between lateral position and carrier frequency to be recovered empirically.

1. Helmchen F, Denk W (2005) Deep tissue two-photon microscopy. *Nat Methods* 2(12): 932–940.
2. Zoumi A, Yeh A, Tromberg BJ (2002) Imaging cells and extracellular matrix in vivo by using second-harmonic generation and two-photon excited fluorescence. *Proc Natl Acad Sci USA* 99(17):11014–11019.
3. Olivier N, et al. (2010) Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science* 329(5994):967–971.
4. Campagnola PJ, Dong C-Y (2011) Second harmonic generation microscopy: Principles and applications to disease diagnosis. *Laser Photonics Rev* 5(1):13–26.
5. Pantazis P, Maloney J, Wu D, Fraser SE (2010) Second harmonic generating (SHG) nanoprobes for in vivo imaging. *Proc Natl Acad Sci USA* 107(33):14535–14540.
6. Grange R, Lanvin T, Hsieh C-L, Pu Y, Psaltis D (2011) Imaging with second-harmonic radiation probes in living tissue. *Biomed Opt Express* 2(9):2532–2539.
7. Denk W, Strickler JH, Webb WW (1990) Two-photon laser scanning fluorescence microscopy. *Science* 248(4951):73–76.
8. Kobat D, et al. (2009) Deep tissue multiphoton microscopy using longer wavelength excitation. *Opt Express* 17(16):13354–13364.
9. Horton NG, et al. (2013) In vivo three-photon microscopy of subcortical structures within an intact mouse brain. *Nat Photonics* 7(3):205–209.
10. Dunn AK, Wallace VP, Coleno M, Berns MW, Tromberg BJ (2000) Influence of optical properties on two-photon fluorescence imaging in turbid samples. *Appl Opt* 39(7): 1194–1201.
11. Gustafsson MGL (2005) Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proc Natl Acad Sci USA* 102(37):13081–13086.
12. Betzig E, et al. (2006) Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313(5793):1642–1645.
13. Betzig E (1995) Proposed method for molecular optical imaging. *Opt Lett* 20(3): 237–239.
14. Hell SW (2007) Far-field optical nanoscopy. *Science* 316(5828):1153–1158.
15. Hell SW, Wichmann J (1994) Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Opt Lett* 19(11): 780–782.
16. Hell SW (2003) Toward fluorescence nanoscopy. *Nat Biotechnol* 21(11):1347–1355.
17. Hell SW, Kroug M (1995) Ground-state-depletion fluorescence microscopy: A concept for breaking the diffraction resolution limit. *Appl Phys B* 60(5):495–497.
18. Bretschneider S, Eggeling C, Hell SW (2007) Breaking the diffraction barrier in fluorescence microscopy by optical shelving. *Phys Rev Lett* 98(21):218103.
19. Wang P, et al. (2013) Far-field imaging of non-fluorescent species with subdiffraction resolution. *Nat Photonics* 7(6):449–453.
20. Chu S-W, et al. (2014) Measurement of a saturated emission of optical radiation from gold nanoparticles: Application to an ultrahigh resolution microscope. *Phys Rev Lett* 112(1):017402.
21. Chu S-W, et al. (2014) Saturation and reverse saturation of scattering in a single plasmonic nanoparticle. *ACS Photonics* 1(1):32–37.
22. Wicker K, Heintzmann R (2014) Resolving a misconception about structured illumination. *Nat Photon* 8(5):342–344.
23. Masihzadeh O, Schlup P, Bartels RA (2009) Enhanced spatial resolution in third-harmonic microscopy through polarization switching. *Opt Lett* 34(8):1240–1242.
24. Liu J, Cho I-H, Cui Y, Irudayaraj J (2014) Second harmonic super-resolution microscopy for quantification of mRNA at single copy sensitivity. *ACS Nano* 8(12):12418–12427.
25. Futia G, Schlup P, Winters DG, Bartels RA (2011) Spatially-chirped modulation imaging of absorption and fluorescent objects on single-element optical detector. *Opt Express* 19(2):1626–1640.
26. Schlup P, Futia G, Bartels RA (2011) Lateral tomographic spatial frequency modulated imaging. *Appl Phys Lett* 98(21):211115.
27. Higley DJ, Winters DG, Futia GL, Bartels RA (2012) Theory of diffraction effects in spatial frequency-modulated imaging. *J Opt Soc Am A Opt Image Sci Vis* 29(12): 2579–2590.
28. Hoover EE, et al. (2012) Eliminating the scattering ambiguity in multifocal, multimodal, multiphoton imaging systems. *J Biophotonics* 5(5-6):425–436.
29. Winters DG, Bartels RA (2015) Two-dimensional single-pixel imaging by cascaded orthogonal line spatial modulation. *Opt Lett* 40(12):2774–2777.
30. Heintzmann R, Gustafsson MGL (2009) Subdiffraction resolution in continuous samples. *Nat Photon* 3(7):362–364.
31. Wernsing KA, et al. (2016) Point spread function engineering with multiphoton SPIFI. *Proc of SPIE* 9713:971304.
32. Born M, Wolf E (1999) *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge Univ Press, Cambridge, UK), 7th Ed.
33. Novotny L, Hecht B (2012) *Principles of Nano-Optics* (Cambridge Univ Press, Cambridge, UK), 2nd Ed.
34. Diebold ED, Buckley BW, Gossett DR, Jalali B (2013) Digitally synthesized beat frequency multiplexing for sub-millisecond fluorescence microscopy. *Nat Photon* 7(10): 806–810.
35. Smith DR, Winters DG, Bartels RA (2013) Submillisecond second harmonic holographic imaging of biological specimens in three dimensions. *Proc Natl Acad Sci USA* 110(46): 18391–18396.
36. Domingue SR, Bartels RA (2014) Nonlinear fiber amplifier with tunable transform limited pulse duration from a few 100 to sub-100-fs at watt-level powers. *Opt Lett* 39(2):359–362.
37. Chong A, Buckley J, Renninger W, Wise F (2006) All-normal-dispersion femtosecond fiber laser. *Opt Express* 14(21):10095–10100.
38. Field JJ, Bartels RA (2016) Digital aberration correction of fluorescent images with coherent holographic image reconstruction by phase transfer (CHIRPT). *Proc of SPIE* 9713:97130B.
39. Field JJ, Winters DG, Bartels RA (2015) Plane wave analysis of coherent holographic image reconstruction by phase transfer (CHIRPT). *J Opt Soc Am A Opt Image Sci Vis* 32(11):2156–2168.

Supporting Information

Field et al. 10.1073/pnas.1602811113

I. MP-SPIFI in One Spatial Dimension

As discussed in the main text, images are formed in the SPIFI microscope by projecting illumination intensity patterns into the specimen and collecting signal light emitted from the object on a single-element detector. The illumination intensity is formed by the spatial interference of multiple light sheets propagating at varied angles with respect to the optic axis. The density (spatial frequency) of the interference fringes in the illumination intensity is precisely controlled by use of a circular glass disk on which is printed an amplitude modulation mask. A spatially coherent illumination beam is brought to a line focus on the modulation mask, and the mask plane is then imaged to the object region with a tube lens and an objective lens (compare Fig. S1 *A* and *B*).

During a single rotation of the disk, the fundamental spatial frequency of the interference fringes is varied through the entire range of spatial frequencies supported by the image relay system (Movie S1). As the disk rotates, the signal light, which is generated by the interaction of the illumination intensity and the distribution of molecules in the object, is collected on a single-element detector. The temporal signal measured from the photodetector is expressed by integrating the contrast light over all space. As discussed in the main text, we only consider the case where the contrast intensity depends on an integer power of the illumination intensity. The contrast signal, $\beta(x,t)$, is then the product of the illumination intensity pattern raised to the power of the nonlinearity, η, and the contrast distribution function, $C(x,t)$, which describes how the contrast-producing molecules within the object are spatially distributed as a function of time:

$$S(t) = \int_{-\infty}^{\infty} dx\, \beta(x,t) = \langle [I(x,t)]^{\eta}\, C(x,t) \rangle_x. \qquad \textbf{[S1]}$$

The signal measured from the photodetector, $S(t)$, represents the projection of the object spatial frequency content onto the spatial frequencies of the illumination structure. At a given scan time, t_s, the principle spatial frequency projected onto the object is $f_x(t_s)$, and $S(t_s)$ encodes the relative amplitude of $f_x(t_s)$ in the object. The whole temporal signal $S(t)$ thereby encodes a 1D image of the object in the spatial frequency domain.

Here, we wish to derive the form of the signal from the photodetector, $S(t)$, for a second-order multiphoton process, such as SHG and TPEF. Ignoring polarization effects and assuming a static contrast function, the signal light is then $\beta(x,t) = I^2(x,t)\, C(x)$, where $C(x) = |\chi^{(2)}(x)|^2$ for SHG and where $\chi^{(2)}(x)$ is the second-order nonlinear susceptibility tensor for the molecules in the specimen. For TPEF, $C(x)$ describes the spatial distribution of fluorescent molecules. In both cases we have absorbed proportionality constants such as two-photon cross-section into $C(x)$ for simplicity. Because the illumination beam is spatially coherent, the modulation mask causes the incident light beam to be diffracted into multiple beams propagating at differing angles with respect to the optic axis. Although the amplitude modulation masks used in this work are binary, and thus have numerous diffracted orders occurring at odd-order harmonics, we analyze a system in which only first-order diffraction occurs. Formally, this approximation is equivalent to considering an amplitude modulation mask for which the printing process is not binary but instead can faithfully produce sinusoidal amplitude modulation.

Using the results of ref. 39 to compute the electric field of each beam in the object region, we computed the 1D spatiotemporal

illumination intensity at the focal plane of the illumination light sheet for a second-order nonlinear process:

$$
\begin{aligned}
I^2(x,t) = {}& \left[\frac{1}{16} + \frac{3}{\pi^2} w^2(t) + \frac{6}{\pi^4} w^4(t) \right] \\
& + \left[\frac{1}{\pi} w(t) + \frac{12}{\pi^3} w^3(t) \right] \cos[2\pi f_x(t)\, x + 2\pi\, \nu_c\, t] \\
& + \left[\frac{3}{\pi^2} w^2(t) + \frac{8}{\pi^4} w^4(t) \right] \cos[4\pi f_x(t)\, x + 4\pi\, \nu_c\, t] \qquad \textbf{[S2]} \\
& + \frac{4}{\pi^3} w^3(t) \cos[6\pi f_x(t) x + 6\pi\, \nu_c\, t] \\
& + \frac{2}{\pi^4} w^4(t) \cos[8\pi f_x(t) x + 8\pi\, \nu_c\, t],
\end{aligned}
$$

where $w(t)$ is the so-called temporal window function, which describes the amplitude of the diffracted beams in the object plane as a function of scan time. In general, this temporal window function can vary for each of the diffracted beams. However, for a well-aligned system, the temporal window function is identical for both diffracted beams, as we have assumed in the preceding expression. We also assume an imaging system with no aberrations, causing $w(t)$ to be a real-valued function. Numerical prefactors were determined using the diffraction efficiency from a perfect square grating. We can write the intensity compactly as

$$I^2(x,t) = \sum_{q=0}^{4} H_q(t) \cos[2\pi\, q\, f_x(t) x + 2\pi\, q\, \nu_c\, t], \qquad \textbf{[S3]}$$

where $H_q(t)$ is determined by the temporal window function and the image order, q. We have explicitly written in the carrier frequency, ν_c, which results from the optic axis being aligned at a lateral position $x_c \neq 0$, as shown in Fig. S1*A*. The lateral position relative to the center of the modulation mask can be written as $x' = x_c + x$, where x is the lateral position relative to the optic axis. The carrier frequency can then be written as $\nu_c = (n_2/n_1)\, M\Delta k\, \nu_r\, x_c$.

This analysis can be generalized for an arbitrary integer nonlinearity to find

$$[I(x,t)]^{\eta} = \sum_{q=0}^{2\eta} H_q(t) \cos[2\pi\, q\, f_x(t)\, x + 2\pi\, q\, \nu_c\, t]. \qquad \textbf{[S4]}$$

Inserting this expression of the nonlinear illumination intensity into the general form of the MP-SPIFI signal in Eq. **S1**, we find

$$
\begin{aligned}
S(t) = {}& \int_{-\infty}^{\infty} dx \sum_{q=0}^{2\eta} H_q(t) \cos[2\pi q f_x(t) x + 2\pi q \nu_c t] C(x) \\
= {}& H_0(t)\tilde{C}(0) + \frac{1}{2} \sum_{q=1}^{2\eta} e^{i2\pi q \nu_c t} H_q(t) \tilde{C}[q f_x(t)] \qquad \textbf{[S5]} \\
& + \frac{1}{2} \sum_{q=1}^{2\eta} e^{-i2\pi q \nu_c t} H_q(t) \tilde{C}^*[q f_x(t)].
\end{aligned}
$$

The contrast function $C(x)$ is real, so it follows that the Fourier transform of the contrast function displays conjugate symmetry in the spatial frequency domain.

Eq. S5 illustrates how 2η unique images of the object are encoded in the spatial frequency domain with a single rotation of the modulation mask. The spatial frequency support of each image varies with harmonic order q, allowing spatial frequency information up to q times the conventional limit to be encoded. Moreover, single-shot collection of all images is possible thanks to the carrier frequency that allows for each image to be isolated in the modulation frequency domain by standard Fourier methods.

We can also express the signal measured in time as a function of spatial frequency by noting that spatial frequency is directly proportional to time. Then, we can express the measured image in spatial frequency space for a given order q as

$$\tilde{S}_{q+}(f_x) = e^{i2\pi q \nu_c t} H_q(f_x) \tilde{C}[q f_x] = e^{i2\pi q f_x x_c} H_q(f_x) \tilde{C}[q f_x]. \quad \text{[S6]}$$

As the illumination fringes shift across the line focus with scan time, a given point in the focal region will experience a modulation frequency that increases with the density of the fringe pattern (Movie S4). In the spatial frequency domain, the carrier frequency simply imparts a linear phase ramp, i.e., $q f_x x_c = q \nu_c t$. Consequently, in the spatial domain, which is directly proportional to the modulation frequency domain, all 2η images are separated by a linear shift in x. In practice, we extract the image that is shifted to $2\pi q f_x x_c$:

$$\tilde{S}_{q+}(f_x) = e^{i2\pi q f_x x_c} H_q(f_x) \tilde{C}(q f_x), \quad \text{[S7]}$$

By removing the residual linear phase, which is equivalent to de-modulation by the carrier frequency, we can express the qth order image in the spatial domain as

$$S_{q+}(x) = \frac{1}{q} h_q(qx) \otimes C(x). \quad \text{[S8]}$$

II. Limitations of Resolution Enhancement

The circular pupil of the objective lens imposes limitations on the resolution enhancement attainable by apodizing the OTFs for each SPIFI order. This apodization is intuitively understood by observing that the shape of each diffracted beam in the pupil plane is highly elliptical. Because each of the diffracted beams scans laterally across the x dimension of the pupil (Movie S2), the vertical spatial frequency support and the throughput energy of the diffracted beams varies with scan time. Both of these effects cause a reduction in the peak intensity in the object plane and impose strict limitations on the spatial frequency support. Moreover, a time-dependent depolarization of the diffracted beams leads to further reductions in the contrast of the illumination fringes. In the following sections, we describe the numerical models used to compute the theoretical MTFs for each MP-SPIFI order.

A. Energy Transmission vs. Scan Time. We computed the energy transmission as a function of scan time by numerically integrating

the amplitude of the diffracted fields over the circular pupil of the objective lens:

$$e_j(t) = \int\limits_0^{2\pi} d\phi \int\limits_0^{\alpha} d\theta \sin\theta \, P_j(\theta, \phi, t). \quad \text{[S9]}$$

Here, $P_j(\theta, \phi, t)$ is the time-dependent pupil function, describing the distribution of the electric field amplitude in the pupil plane. The relative energy transmission curves are shown in Fig. S4, where it is clear that the diffracted modes ($j = 1$) contribute less to the overall energy of the illumination intensity as the beam traverses the pupil of the objective lens.

B. Vignetting of Diffracted Illumination Beams. The circular pupil of the objective lens reduces the amplitude of high spatial frequencies in measured data. For example, the fourth-order MP-SPIFI image is formed by the second-order nonlinear response of the intensity pattern created by interference between the two diffracted beams. In the notation of ref. 39, we can write the fourth-order component of the illumination intensity as a product of the electric fields, $v_{\pm1}(x, t)$, and their conjugates, $u_{\pm1}(x, t)$:

$$S_4(t) = \langle I_4(x,t) C(x,t) \rangle_{\mathbf{r}} = \left\langle [v_1(x,t) \, u_{-1}(x,t)]^2 \, C(x,t) \right\rangle_x$$
$$+ \left\langle [v_{-1}(x,t) \, u_1(x,t)]^2 \, C(x,t) \right\rangle_x. \quad \text{[S10]}$$

As the two diffracted beams scan off the optic axis of the imaging system, vignetting by the circular pupil of the objective lens causes a loss of the vertical spatial frequency support, f_y, which is responsible for a further decrease in the peak intensity in the object region.

We studied the effect of vignetting on lateral spatial frequency support by computing the illumination intensity in the object region with the angular spectrum representation (32, 33). To compute $S(t)$, we must compute the illumination intensity in the object region. The vectorial electric field of each diffracted beam in the object region is

$$\mathbf{v}_{j,2}(\mathbf{r}, t) = v_{x,j,2}(\mathbf{r}, t) \mathbf{x} + v_{y,j,2}(\mathbf{r}, t) \mathbf{y} + v_{z,j,2}(\mathbf{r}, t) \mathbf{z}, \quad \text{[S11]}$$

and the illumination intensity becomes

$$I_{\text{ill}}(\mathbf{r}, t) = \sum_{k=-N}^{N} \sum_{j=-N}^{N} \mathbf{v}_{j,2}(\mathbf{r}, t) \cdot \mathbf{u}_{k,2}(\mathbf{r}, t). \quad \text{[S12]}$$

The measured MP-SPIFI signal is then

$$S(t) = \left\langle \left[\sum_{k=-N}^{N} \sum_{j=-N}^{N} \mathbf{v}_{j,2}(\mathbf{r}, t) \cdot \mathbf{u}_{k,2}(\mathbf{r}, t) \right]^{\eta} C(\mathbf{r}, t) \right\rangle_{\mathbf{r}}. \quad \text{[S13]}$$

To obtain the illumination intensity at the focal plane, the Debye integral must be solved numerically for each diffracted order:

$$\begin{pmatrix} v_{x,j,2}(\mathbf{r},t) \\ v_{y,j,2}(\mathbf{r},t) \\ v_{z,j,2}(\mathbf{r},t) \end{pmatrix} = \int\limits_0^{2\pi} d\phi \int\limits_0^{\alpha} d\theta \sin\theta \, P_j(\theta,\phi,t) \begin{pmatrix} \cos\phi \cos\theta \cos(\phi-\gamma) + \sin\phi \sin(\phi-\gamma) \\ \sin\phi \cos\theta \cos(\phi-\gamma) - \cos\phi \sin(\phi-\gamma) \\ \sin\theta \cos(\phi-\gamma) \end{pmatrix}$$
$$\times \sqrt{\frac{n_1}{n_2}} \cos\theta \, \exp(i \, k_2 \, z \cos\theta) \exp[i \, k_2 \, \rho \sin\theta \cos(\phi-\varphi)]. \quad \text{[S14]}$$

We assume linearly polarized light, with the angle γ representing the angle of the polarization vector with respect to x. The variable φ is the azimuthal angle in the pupil plane of the objective lens, measured with respect to x.

In Cartesian coordinates, the pupil function has the form of a 2D Gaussian distribution:

$$P_j\left(x_p, y_p, t\right) = a_j \exp\left\{-\left[\frac{x_p - x_{0,j}(t)}{w_x}\right]^2\right\} \exp\left\{-\left[\frac{y_p - y_{0,j}(t)}{w_y}\right]^2\right\},$$

[S15]

where (x_p, y_p) are coordinates in the pupil plane, $x_{0,j}(t)$ and $y_{0,j}(t)$ are the lateral and vertical shifts of the jth diffracted order in the pupil plane, and w_x and w_y are the widths of the Gaussian distributions. To solve the integral equation in Eq. S14, the pupil function must be converted to spherical polar coordinates, (F_o, θ, ϕ). Note that the radial coordinate is equivalent to the focal length of the objective lens because the pupil function is to be described on the surface of an aplanatic focal sphere (33). Applying the coordinate change, we find

$$x_p = F_0 \sin\theta \cos\phi$$

[S16]

$$y_p = F_0 \sin\theta \sin\phi,$$

[S17]

and the pupil function in Eq. S15 becomes

$$P_j(\theta, \phi, t) = a_j \exp\left\{-\left[\frac{F_o \sin\theta \cos\phi - x_{0,j}(t)}{w_x}\right]^2\right\} \\ \times \exp\left\{-\left[\frac{F_o \sin\theta \sin\phi - y_{0,j}(t)}{w_y}\right]^2\right\}.$$

[S18]

The lateral shifts of the diffracted orders in the pupil plane are determined by the spatial frequency of the modulation mask as a function of scan time, the wavelength of the illumination light, and the focal length of the tube lens (39):

$$x_{0,j}(t) = F_t\, j\, \frac{\lambda}{n_1} f_{x,1}(t) = F_t\, j\, \frac{\lambda}{n_1} \Delta k\, \nu_r\, t = j\, x_{0,1}(t)$$

[S19]

$$y_{0,j} = F_t\, j\, \frac{\lambda}{n_1} f_{y,1} = F_t\, j\, \frac{\Delta k}{k_1} = j\, y_{0,1}.$$

[S20]

The widths of the distributions are set by the focal lengths of the cylindrical and tube lenses, as well as the input beam size, w_{in}, and the illumination wavelength:

$$w_x = \frac{\lambda\, F_t}{\pi\, w_{\text{in}}}$$

[S21]

$$w_y = \frac{F_t}{F_c}\, w_{\text{in}},$$

[S22]

where F_c is the focal length of the cylindrical lens.

We computed the illumination intensity in the transverse plane located in the focus of the illumination objective ($z = 0$) for several scan times, considering only the $j = 0$ and $j = \pm 1$ beams. From Fig. S5, it is clear that as the first-order diffracted beams scan away from the optic axis in the lateral (x) dimension, the vertical spatial frequency support decreases and causes a corresponding broadening of the illumination beams in the vertical dimension.

Finally, the vector-focusing formalism was used to compute the OTF of each MP-SPIFI order using the parameters of the microscope described in *Materials and Methods* ($\lambda = 1{,}064$ nm; $\Delta k = 70$/mm; 150-mm focal length tube lens; 50×/0.8 NA Zeiss air-immersion UIS series objective lens). To compute the MP-SPIFI signal, we assumed a point emitter located at the focal point of the objective lens, $\mathbf{r}_0 = (0,0,0)$. The point emitter was modeled by a Dirac-δ distribution, $\delta(\mathbf{r} - \mathbf{r}_0)$. Because the contrast distribution we consider is a delta function, we need only compute the intensity at $\mathbf{r}_0$ to obtain the MP-SPIFI signal:

$$S(t) = \left\langle \left[I_{\text{ill}}(\mathbf{r}, t)\right]^2\, \delta(\mathbf{r} - \mathbf{r}_0) \right\rangle_{\mathbf{r}} = \left[I_{\text{ill}}(\mathbf{r}_0, t)\right]^2.$$

[S23]

For the simulated TP-SPIFI data presented in Fig. 5, we considered a second-order nonlinearity ($\eta = 2$) and included only first-order diffraction from the modulation mask, where $a_0 = 1/2$ and $a_1 = 1/\pi$ were set by the diffraction efficiency of a square grating.
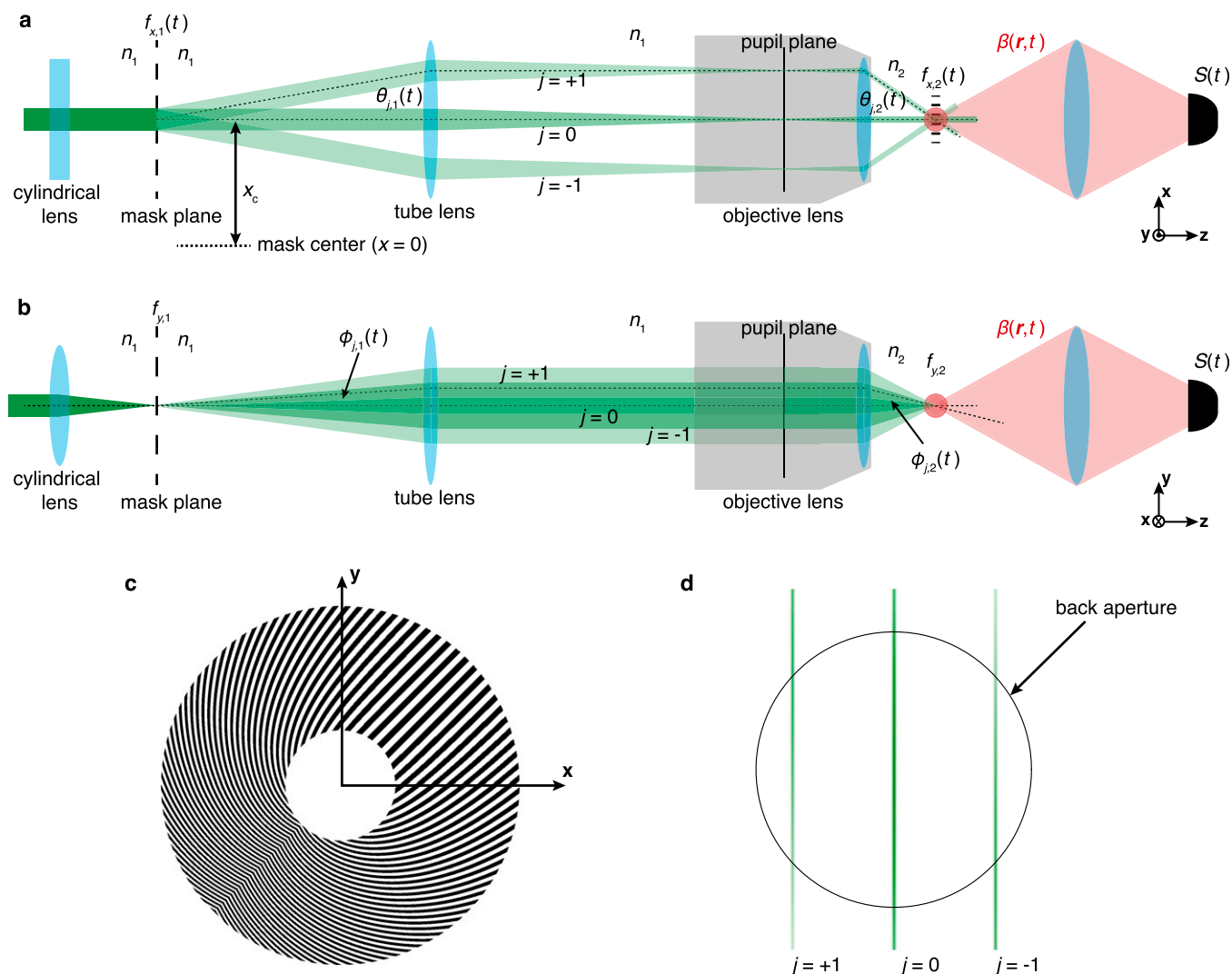
Fig. S1. MP-SPIFI system configuration. (*A* and *B*) System schematics in the (x, z) (*A*) and (y, z) (*B*) planes. The time-dependent angle in the mask region, $\theta_{1,2}(t)$, depends on the instantaneous lateral frequency experienced by the line focus $\theta_{1,2}(t) = \sin^{-1}[\lambda\, f_x(t)]$. (*C*) The SPIFI modulation mask with $\Delta k = 2/$mm. (*D*) The spatial distribution of diffracted beams in the pupil plane, computed using the parameters of the imaging system used in this work ($\lambda = 1{,}064$ nm, NA = 0.8, tube lens focal length of 150 mm). The back aperture diameter was computed using the magnification of the objective lens, M_o, and the design tube lens focal length $F_t^l = 164.5$ mm (Zeiss UIS series) with the expression $d = 2\, F_t^l\, \text{NA}/M_o$. The color gradients on each beam indicate the relative intensity of each beam. Because the beam is Gaussian in intensity profile on the mask, the line cursors scanning in the pupil plane are Gaussian in intensity profile.
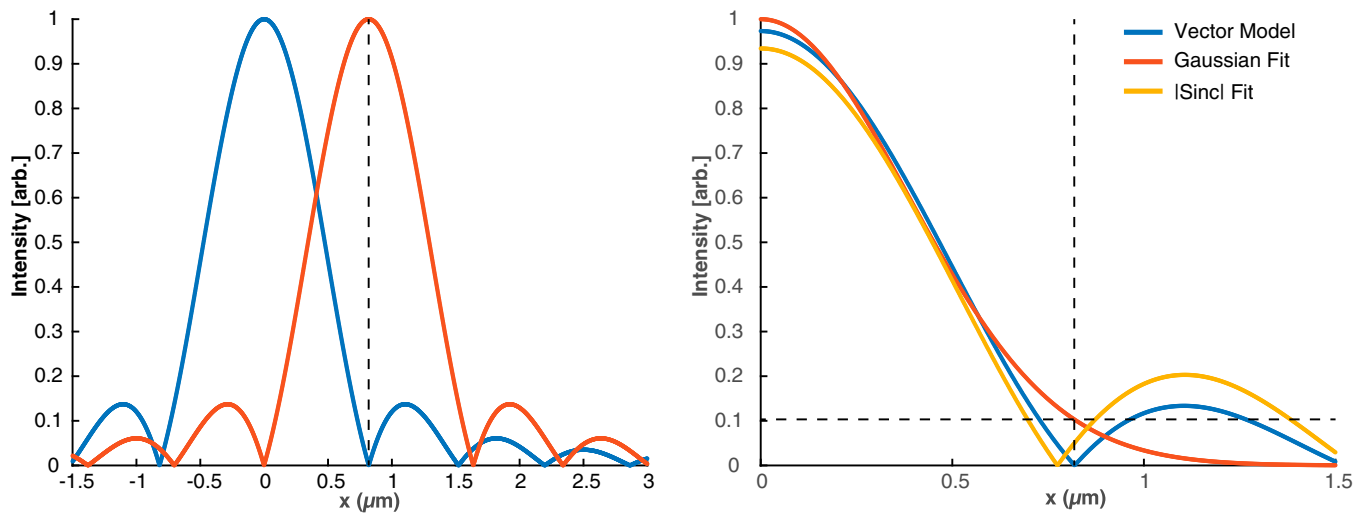
Fig. S2. The Raleigh resolution criterion was used to benchmark the spatial resolution from the PSFs shown in Fig. 2*E*. The left graph shows two copies of the simulated PSF for the first-order ($q = 1$) MP-SPIFI image, computed with the vector-focusing model described in the *Supporting Information*. Raleigh resolution is defined as the distance that a copy of the PSF must be shifted in order for its peak to coincide with a zero in the unshifted PSF. In the measured PSFs data, the shot noise background makes it difficult to interpret where the first zero of the PSF is located. Instead of directly detecting this zero, we fit a Gaussian profile to the measured PSF. In the ideal case of a uniform MTF, the measured PSF would be proportional to $|\text{sinc}(x)|$. However, the MTF is neither uniform, nor is it a Gaussian, so fitting a $\text{sinc}(x)$ or a Gaussian to the PSF results in deviations from the measured PSF. This fit is shown explicitly in the right graph, where we have fit both a Gaussian and a $|\text{sinc}(x)|$ to the simulated PSF. We find that for the numerically simulated PSF, the half-width at 10.4% of the maximum of the Gaussian profile corresponds to the first zero in the computed PSF. We used this metric to compute the spatial resolution of our experimental data.
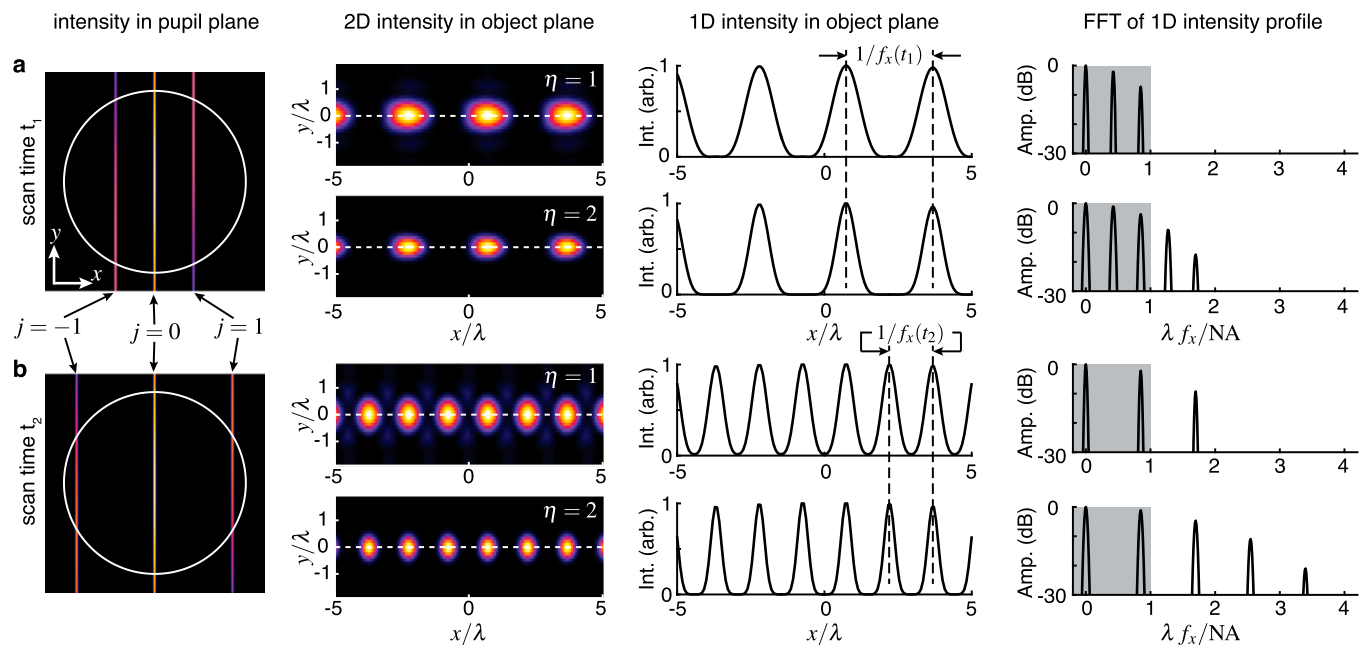
Fig. S3. The illumination intensity pattern at two snapshots in time, t_1 (*A*) and t_2 (*B*), during a single temporal scan of the MP-SPIFI signal, $S(t)$. The left column shows the intensity at the rear pupil of the objective lens, where the white circle represents the pupil. The second column displays the illumination intensity in the focal plane of the objective lens. The illumination intensity for both linear and second-order nonlinear contrast generation are shown for each snapshot in time. The third column displays the intensity along the x axis in the object plane, as indicated by the dashed lines in the second column. Finally, the fourth column shows the FFT of the 1D intensity line-outs, where the gray boxes represent the passband of the illumination objective lens. Note that the amplitude information in the frequency domain is symmetric about $f_x = 0$, but we display only the positive spatial frequencies to save space. The spacing between the $j = \pm 1$ beams and the $j = 0$ beam dictates the fundamental spatial frequency of the illumination intensity in the object plane. At time t_1, the $j = \pm 1$ beams are relatively close to the $j = 0$ beam in the pupil plane, and the resulting intensity in the object plane has a low spatial frequency, $f_x(t_1)$. At time t_2, the $j = \pm 1$ beams are separated from the $j = 0$ beam by twice the distance as at time t_1, thereby producing an intensity pattern with twice the spatial frequency, $f_x(t_2)$. The illumination intensities in the focal plane (second column) were computed using the Debye formalism outlined in the *Supporting Information*.
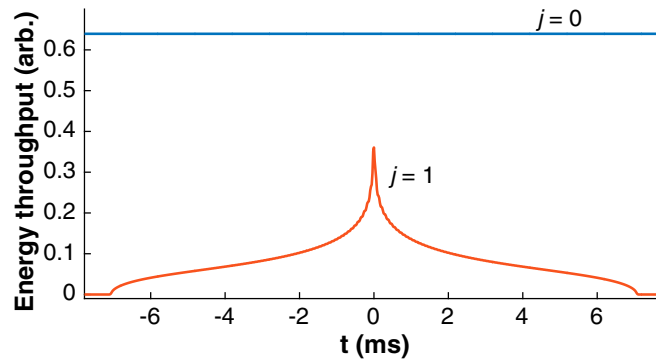
Fig. S4. Simulated energy transmission of the diffracted orders through the objective lens. Relative energies are normalized by the total transmitted energy at scan time $t = 0$, when both diffracted orders are overlaid in the lateral dimension.
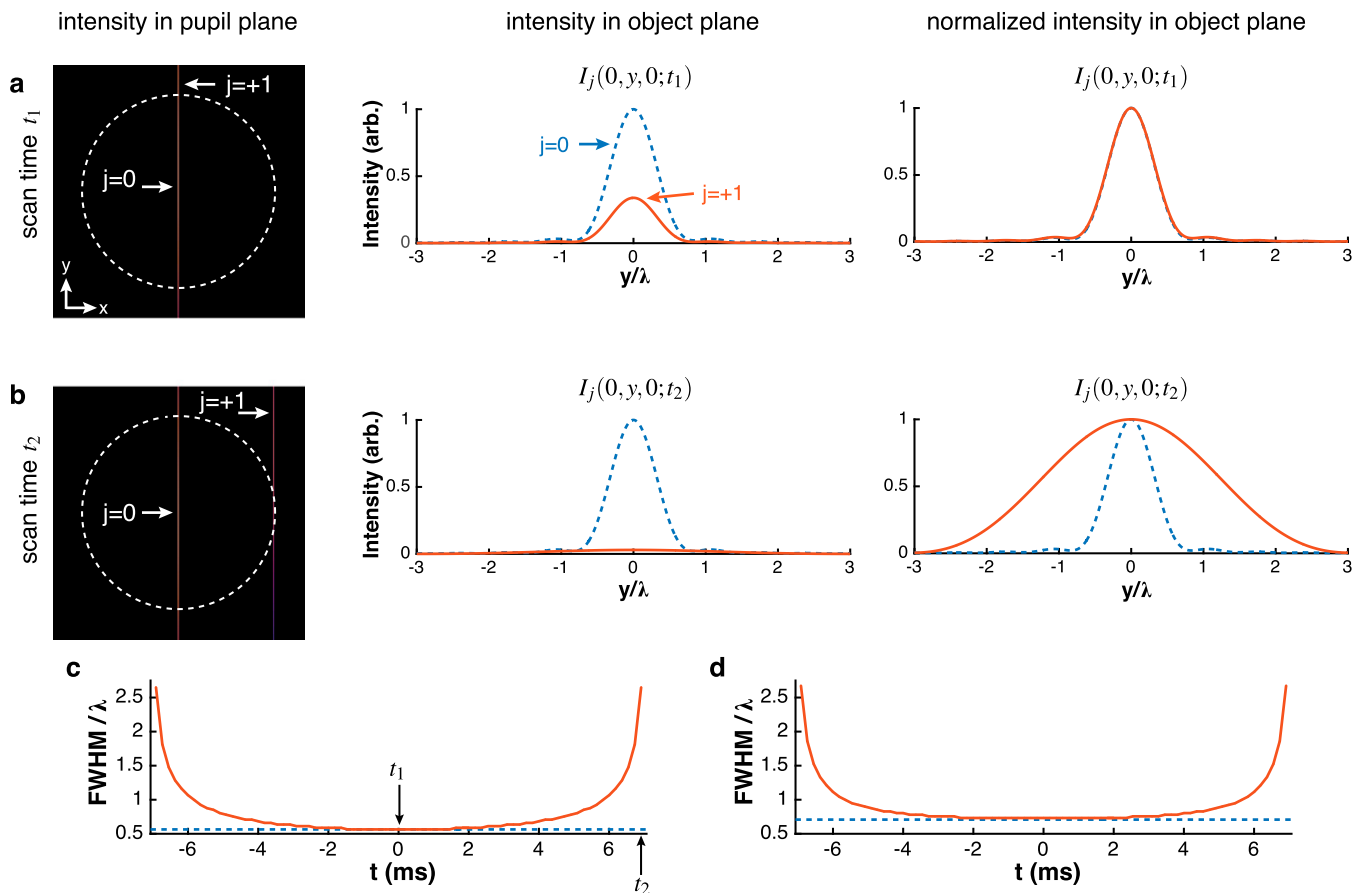
Fig. S5. Vertical spatial frequency support vs. scan time. (*A*) At zero scan time, the diffracted beams and the undiffracted beam fill the objective lens pupil in the vertical dimension. In the focal plane, the width of the light sheet in the vertical dimension is diffraction limited for both the diffracted and undiffracted beams. (*B*) At higher scan times, when the diffracted order approaches the maximum spatial frequency supported by the lens in the horizontal dimension, the vertical spatial frequency content is significantly reduced by the circular pupil. Consequently, the $j = 1$ light sheet does not focus as tightly in the vertical dimension as the $j = 0$ light sheet. This apodization reduces the intensity of the diffracted beam in the focal plane, thus reducing the contrast of the illumination intensity fringes. The width of the light sheet changes with scan time and varies slightly for x-polarized (*C*) and y-polarized (*D*) input light.
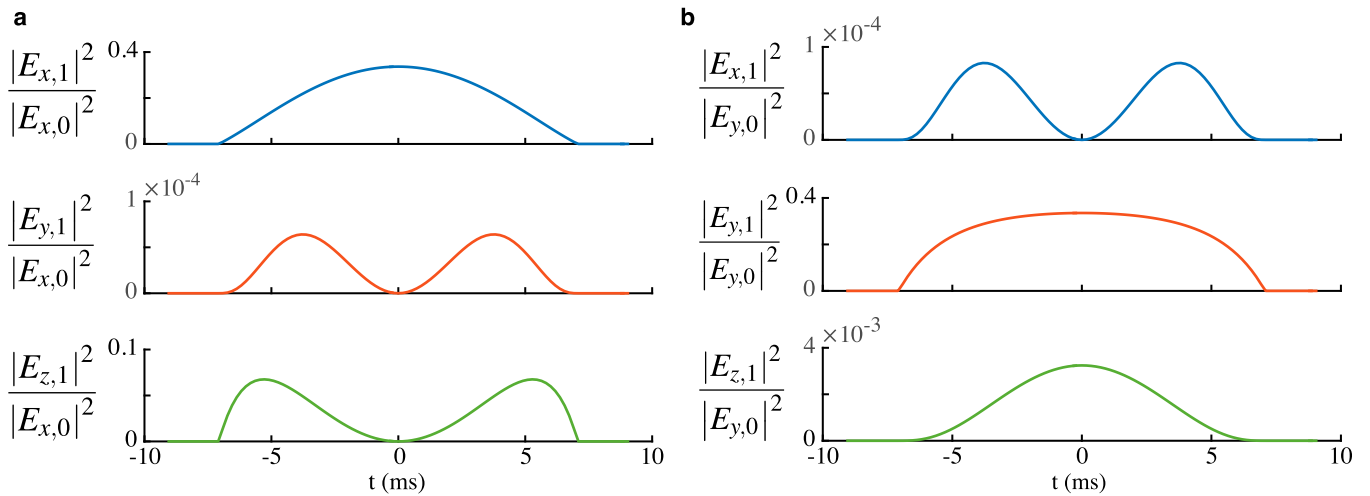
Fig. S6. Depolarization of linearly polarized input electric fields vs. scan time. The polarization state in the pupil plane was aligned in the *x* direction (*A*) and *y* direction (*B*). Depolarization of the first diffracted order ($j = 1$) is shown by plotting the ratio of the intensities in each spatial dimension to the intensity of the input polarization state for the zero order beam. Depolarization at high spatial frequencies causes a reduction in the contrast of the illumination intensity, which can reduce the SNR.
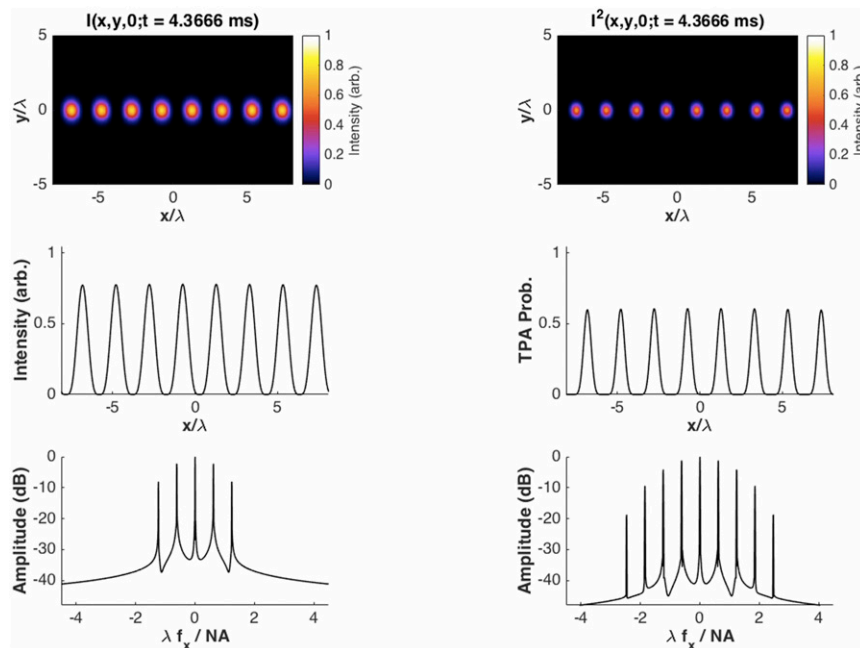
Movie S1. Simulation of the intensity and spatial frequency content of the illuminating light sheet at the focal plane of the objective lens during a single rotation of the modulation mask. The left column represents the illumination intensity ($\eta = 1$), whereas the right column corresponds to the intensity squared ($\eta = 2$). The first row shows the intensity in the lateral (x, y) plane at the focal plane of the objective lens ($z = 0$). The second row shows a 1D line-out of the intensity at $y = 0$. An FFT of the 1D intensity for linear excitation (*Left*) shows a peak at zero spatial frequency and peaks corresponding to the fundamental spatial frequency, $f_x(t)$, as well as $2f_x(t)$. (*Lower Right*) The FFT of the 1D intensity for $\eta = 2$, representing a two-photon process such as SHG. In this case, harmonics of the fundamental spatial frequency are present up to $4f_x(t)$. The temporal window function, $w(t)$, was computed for *y*-polarized light with the vector-focusing model described in the *Supporting Information* and applied to the simulated data. The simulation used the parameters of the imaging system we describe in the main text [i.e., $\lambda = 1,065$ nm, a 46.5×/0.8 NA imaging system, and a mask with $\Delta k = 70$/mm (*Materials and Methods*)].

Movie S1

Movie S2. Visualization of the intensity in conjugate image planes of the MP-SPIFI microscope system. (*Upper Left*) The modulation mask, which rotates at a constant angular velocity as a function of scan time. Note that for visual clarity, we have plotted a mask with a maximum density of $\Delta k = 2/\text{mm}$, although in the experiment we used a mask with $\Delta k = 70/\text{mm}$ (*Materials and Methods*). The red ellipse on the right side of the modulation mask indicates the location of the line focus that is modulated by the mask. This schematic representation of the laser beam is not drawn to scale to facilitate visual clarity. (*Upper Right*) The intensity in the pupil plane of the objective lens, which has diameter $d = 2 f_{obj}$ NA, where f_{obj} is the apparent focal length of the objective lens. The undiffracted beam is stationary at $x/d = 0$ for the duration of the scan, whereas the two diffracted beams scan across the pupil in the horizontal (x) dimension. (*Lower Left*) The illumination intensity in the focal plane of the objective lens, which is a conjugate image plane to the mask plane. (*Lower Right*) The intensity squared in the same plane, corresponding to the probability of exciting a two-photon process. The temporal window function, $w(t)$, was computed for y-polarized light with the vector-focusing model described in the *Supporting Information* and applied to the simulated data. The simulation used the parameters of the imaging system that we describe in the main text [i.e., $\lambda = 1,065$ nm, a 46.5×/0.8 NA imaging system, and a mask with $\Delta k = 70/\text{mm}$ (*Materials and Methods*)].

[Movie S2](#)

Movie S3. Visualization of the intensity in the pupil plane and the object plane of the MP-SPIFI microscope as a function of scan time for different beam combinations. The upper row displays the interference of the undiffracted and the positive-diffracted ($j = 0, +1$) beams, whereas the lower row displays the interference due to the two diffracted beams ($j = \pm 1$). The spatial frequency of the illumination pattern at *Lower Right* is twice the spatial frequency of the pattern at *Upper Right* at all times. The temporal window function, $w(t)$, was computed for y-polarized light with the vector-focusing model described in the *Supporting Information* and applied to the simulated data. The simulation used the parameters of the imaging system we describe in the main text [i.e., $\lambda = 1,065$ nm, a 46.5×/0.8 NA imaging system, and a mask with $\Delta k = 70/\text{mm}$ (*Materials and Methods*)].

[Movie S3](#)

Movie S4. Decomposition of the effective illumination intensity due to a second-order nonlinearity into harmonics of the spatial frequency, $f_x(t)$. The upper left graph displays the illumination intensity squared, $I^2(x, y, 0; t)$, whereas each successive graph shows a component of the pattern, $I_q(x, y, 0; t)$, where $I^2(x, y, 0; t) = \sum_{q=0}^{4} I_q(x, y, 0; t) = \sum_{q=0}^{4} H_q(t) \cos[2\pi q f_x(t)(x + x_c)]$. The temporal window function, $w(t)$, was computed for y-polarized light with the vector-focusing model described in the *Supporting Information* and applied to the simulated data. The simulation used the parameters of the imaging system we describe in the main text [i.e., $\lambda = 1,065$ nm, a 46.5×/0.8 NA imaging system, and a mask with $\Delta k = 70/$mm (*Materials and Methods*)].

[Movie S4](#)

Movie S5. A simulation of the illumination intensity in the object plane and the signal collected from a point emitter located at $\mathbf{r} = (0,0,0)$. The left column shows each harmonic component of the illumination intensity for $\eta = 2$, similar to Movie S4. In each graph, we have plotted a small green beacon to indicate the instantaneous intensity of light emitted from the point emitter by a two-photon process. The center column displays the full measured signal $S(t)$ (*Top*), whereas each successive graph corresponds to a single harmonic component of the signal, $S_q(t)$. The right column displays the absolute value of the OTF as a function of scan time after recalibrating the horizontal axis to the spatial frequency domain by the relation $f_{x,q}(t) = q\, f_x(t) = q\, \kappa t$ (*Materials and Methods*). The temporal window function, $w(t)$, was computed for y-polarized light with the vector-focusing model described in the *Supporting Information* and applied to the simulated data. The simulation used the parameters of the imaging system we describe in the article [i.e., $\lambda = 1,065$ nm, a 46.5×/0.8 NA imaging system, and a mask with $\Delta k = 70/$mm (*Materials and Methods*)].

[Movie S5](#)