Beating Rayleigh's Curse by Imaging Using Phase Information

Weng-Kian Tham,[1] Hugo Ferretti,[1] and Aephraim M. Steinberg[1,2]
[1]Centre for Quantum Information & Quantum Control and Institute for Optical Sciences, Department of Physics, University of Toronto, 60 St. George St, Toronto, Ontario, Canada, M5S 1A7
[2]Canadian Institute For Advanced Research, 180 Dundas St. W., Toronto, Ontario, Canada, M5G 1Z8
(Received 21 July 2016; published 15 February 2017)

Every imaging system has a resolution limit, typically defined by Rayleigh's criterion. Given a fixed number of photons, the amount of information one can gain from an image about the separation between two sources falls to zero as the separation drops below this limit, an effect dubbed "Rayleigh's curse." Recently, in a quantum-information–inspired proposal, Tsang and co-workers found that there is, in principle, infinitely more information present in the full electromagnetic field in the image plane than in the intensity alone, and suggested methods for extracting this information and beating the Rayleigh limit. In this Letter, we experimentally demonstrate a simple scheme that captures most of this information, and show that it has a greatly improved ability to estimate the distance between a pair of closely separated sources, achieving near-quantum-limited performance and immunity to Rayleigh's curse.

DOI: 10.1103/PhysRevLett.118.070801

Any imaging device such as a microscope or telescope has a resolution limit, a minimum separation it can resolve between two objects or sources; this limit is typically defined by "Rayleigh's criterion" [1], although in recent years there have been a number of high-profile techniques demonstrating that Rayleigh's limit can be surpassed under particular sets of conditions [2–4]. As an electromagnetic wave, light is characterized by both an amplitude and a phase. Traditional imaging systems use lenses or mirrors to refocus this wave and project an image of the source onto a screen or camera, where the intensity (or rate of photon arrivals) is recorded at each position. (We refer to all such techniques as "image-plane counting" or IPC). Although the phase of the wave at the position of the optics plays a central role during the focusing, any information about the phase in the image plane is discarded. When light passes through finite-sized optical elements, diffraction smears out the spatial distribution of photons so that point sources map (via the point spread function or PSF) onto finite-sized spots at the image plane. Thus, our ability to resolve the point sources is inhibited when their separation in the image plane, δ, is comparable to, or less than, the width σ of the PSF.

The typical response to diffraction limits has been to build larger (or higher numerical aperture) optics, thereby making the PSF sharper or narrower. In recent years, techniques have been developed in specific cases that address these limits in more novel ways [2–11]. Despite their success, these techniques require careful control of the source of illumination, which is not always possible in every imaging application (e.g., astronomy). In order to beat the diffraction limit for fixed, mutually incoherent sources, a paradigm shift arising from the realization that there is a huge amount of information available in the phase discarded by IPC may prove revolutionary.

Inspired by ideas in quantum information and quantum metrology [12–16], Tsang *et al.* [17] showed that whereas in IPC the Fisher information, I_f [18], vanishes quadratically with the separation δ between two equal-intensity incoherent point sources of weak thermal light with Gaussian PSF, it remains undiminished when the full electromagnetic field is considered. Later, these results were extended to more general types of sources [19,20]. I_f is related to the performance of a statistical estimator by

$$\mathrm{Var}(\delta_{\mathrm{est}}) \geq \frac{1}{I_f}\left(1+\frac{\partial(\mathrm{bias})}{\partial\delta_{\mathrm{actual}}}\right), \tag{1}$$

where δ_{est} is some estimator of δ_{actual} and $\mathrm{bias} \equiv \langle\delta_{\mathrm{est}}\rangle - \delta_{\mathrm{actual}}$ [18]. In the case of an unbiased estimator, this limit, known as the Cramer-Rao lower bound (CRLB) reduces to $(1/I_f)$. The vanishing of I_f as $\delta \to 0$ suggests that for closely separated sources, the variance in an IPC-based estimate of δ is cursed to diverge. That it is independent of δ for the full field, on the other hand, appears to suggest that this divergence can be averted by using phase as well as intensity information.

One natural way to do this would be to use spatial mode demultiplexing (SPADE) [17,21], in which incoming light is decomposed into its Hermite-Gauss (HG) [22] components and the amplitude of each is measured. This HG basis is centered exactly between the two PSFs, and the width of the 0th mode(TEM_{00}) matches the width of the Gaussian PSF. It can be shown that the full set of HG amplitudes contains the same I_f as the full EM field. A reduced version called binary SPADE prescribes discriminating only between the TEM_{00} mode and the sum of all other modes. For small δ, only one other mode acquires significant amplitude in any case, so the I_f available to binary SPADE becomes essentially equal to the full Fisher information. The method can be understood as follows: the projection always succeeds ($P_{00} = 1$) when the two point sources are overlapped ($\delta = 0$), but has a

failure probability $1 - P_{00}$ which grows quadratically with δ. Knowing the TEM_{00} component as a proportion of all HG amplitudes (i.e., P_{00} and $1 - P_{00}$) allows one to deduce δ. All the results above and the experimental work below deal with Gaussian PSFs, which are of interest because they are frequently used as approximations for the Airy rings produced by circular apertures. However, recent results [23] have shown that in the small separation regime, a small number of suitable spatial projections of the electromagnetic field capturing the full Fisher information can also be generated for any PSF.

In this proof-of-principle experiment, we continue with Gaussian PSFs. The spatial wave function of the EM field of the two sources is given by $\psi_{1/(2)}(x, y) = \exp\{[-(x \pm \delta/2)^2 + y^2]/4\sigma^2\}$. Experimentally, merely capturing the TEM_{00} component (say, by coupling into a single mode fiber) without a normalization factor (which allows us to deduce $1 - P_{00}$) provides no advantage over IPC. Practically speaking, the crucial information comes from a projection onto some mode *orthogonal* to TEM_{00} in order to estimate $1 - P_{00}$. While a mode such as TEM_{10} would contain all the information (for a separation in the x direction in that example), the same scaling can be obtained by projecting onto any spatially antisymmetric field mode. As a proof of principle, we have designed and implemented an experimentally convenient method, super-resolved position localization by inversion of coherence along an edge (SPLICE), which instead carries out one single technically straightforward projection onto the mode function $\psi_\perp(x, y) = \exp[-(x^2 + y^2/4\sigma^2)]\mathrm{sgn}(x)$. This function is constructed such that its inner product with TEM_{00} vanishes. The probability that such a projection succeeds is

$$P_\perp = \frac{1}{2}(|\langle\psi_1|\psi_\perp\rangle|^2 + |\langle\psi_2|\psi_\perp\rangle|^2) = e^{-2\Delta}\mathrm{erf}^2\sqrt{\Delta}, \qquad (2)$$

where $\Delta = \delta^2/32\sigma^2$, δ is the separation between point sources on the image plane, and $\psi_{1/(2)}$ is the field from each source.

The per-photon Fisher information can be written as

$$I_f = \frac{(e^{-\Delta}\sqrt{\pi\Delta}\mathrm{erf}\sqrt{\Delta} - e^{-2\Delta})^2}{2\pi\sigma^2} + \frac{\left[e^{-\Delta}\sqrt{\pi\Delta}\mathrm{erf}^2(\sqrt{\Delta}) - e^{-2\Delta}\mathrm{erf}\sqrt{\Delta}\right]^2}{2\pi\sigma^2(e^{2\Delta} - \mathrm{erf}^2\sqrt{\Delta})}, \qquad (3)$$

where the first term comes from $P_\perp$ and the second from $1 - P_\perp$. Crucially, as $\Delta \to 0$, $1 - P_\perp$ vanishes, meaning that an experimentally simple scheme for projecting only onto $\psi_\perp$ does as well as a more complicated scheme which could measure multiple projections simultaneously. In Fig. 1, we plot the Fisher information for SPLICE in comparison with other methods. It is easy to see that it remains nonzero as $\delta \to 0$, evading Rayleigh's curse, and extracting nearly 2/3 of the total information available to full SPADE using an experimentally simple technique. More sophisticated methods relying on wave guides or cavities could be designed to approach 100% of the optimal I_f.

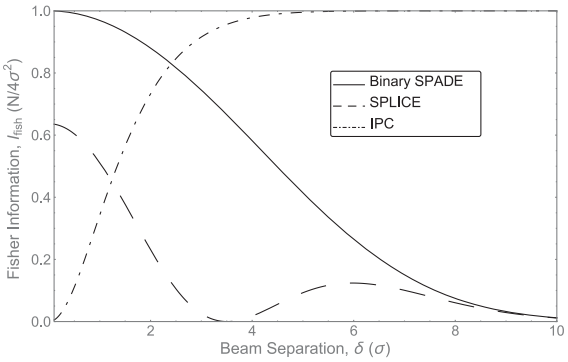

FIG. 1. Theory plot of the Fisher information for SPLICE, binary SPADE and IPC vs beam separation δ, normalized to units of $N/4\sigma^2$ and σ, respectively.

In order to experimentally demonstrate improved performance over IPC, we used two mutually incoherent collimated TEM_{00} Gaussian beams in place of distant point sources and an imaging optical setup. The beams were directed through a Sagnac-like beam displacer shown in Fig. 2. By moving a mirror on a motorized translation stage as shown, we precisely control the separation δ between the otherwise parallel beams. The separation is induced symmetrically, such that the geometrical centroid (x_0, y_0) remains static.

At zero separation, the beams are overlapped and are both coupled into single-mode TEM_{00} fiber (coupling efficiencies are 90% and 85%, respectively). Collimation of the beams is such that their waists are closely matched immediately before the fiber coupler ($\sigma = 434 \pm 3$ and

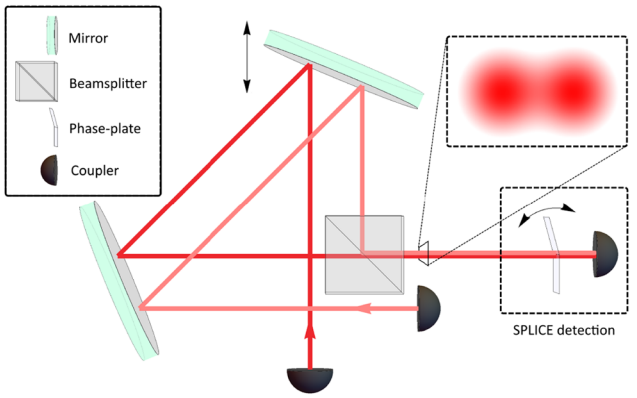

FIG. 2. Cartoon of experiment. Shown is the experimental apparatus. In the lower right-hand box is a representation of SPLICE, the measurement scheme tested in this experiment. In the upper right-hand box is a sketch of the spatial profile of the electromagnetic field before the measurement. The rest of the figure depicts the device used to simulate the two light sources, which can be displaced around their centroid by the displacement of the top mirror.

420 ± 7 μm) in order to emulate Gaussian point spread functions of distant sources. The projection onto $\psi_\perp$ (the SPLICE measurement) is achieved by inserting a phase plate immediately in front of the coupler such that when $\delta = 0$, a semicircular cross section of the beams undergoes a π-phase shift whereas the other half experiences none. The phase plate consists of two transparent glass flat microscope cover slips, connected along one sharp rectangular edge. They are mounted such that one glass slip tilts relative to the other by pivoting along that edge. We then position this contraption such that the connected edge of the glass slips bisects the circular beam cross sections when both beams are overlapped (i.e., $\delta = 0$). Thus, we can impart different phase shifts onto opposite halves of the beams by tilting one glass slip relative to the other (which we do, in order to minimize coupling into an otherwise well-aligned coupler). We typically achieve an extinction of $\geq$99%.

The light source used to create the two mutually incoherent beams is an 805-nm heralded single-photon source which relies on type-I spontaneous parametric down-conversion (SPDC) in a 2 mm-thick BBO crystal. The crystal is pumped by 402.5 nm light obtained from a frequency-doubled 100-fs Ti:sapphire laser. One photon from the SPDC pair is used to herald the presence of a signal photon as a means of rejecting spurious background light and detector dark counts (our accidental coincidences average 2 ± 1 counts/sec). Our SPDC source has a very low probability of producing more than one photon per coherence time ($\approx$150 fs). This resembles the regime investigated in the original theory proposal [17,24]. Furthermore, the low intensity of the source allows us to directly compare our experimental performance with the quantum limit shown in Fig 1. It must be noted that the use of heralded single photons is not necessary to this scheme, which is independent of the photon statistics of the point sources. In order to emulate two point sources, the signal photon is split at a 50/50 fiber splitter and out-coupled to free space. The two resulting beams are incoherent; they have splitter-to-coupler distances that differ by 5 cm, whereas the SPDC photons are filtered to $\Delta\lambda = 3$ nm (i.e., coherence length ≈ 45 μm). ND filters were used to reduce the intensity imbalance between the beams to $\approx(3 \pm 3)$%.

To compare the performance of our method (SPLICE) with a more traditional imaging setup relying on IPC, we replaced the phase plate with a 200 μm slit that served as the image plane, coupling all the light transmitted through the slit into a multimode fiber. Scanning the slit, we were able to perform one-dimensional IPC.

With SPLICE, the separation of the incoherent beams was scanned, with the detectors counting for 1 sec at each step. Two sets of SPLICE scans were performed, one at coarse intervals of δ (spanning -1.16 to $+1.16$ mm, in steps of 0.08 mm). Another scan at finer intervals (-0.43 mm $\leq \delta \leq +0.37$ mm in steps of 0.032 mm) was performed to provide more data points in the region of low separation, where SPLICE provides an advantage. Data from nine repetitions of the coarse scan and fifteen of the fine scan were recorded.

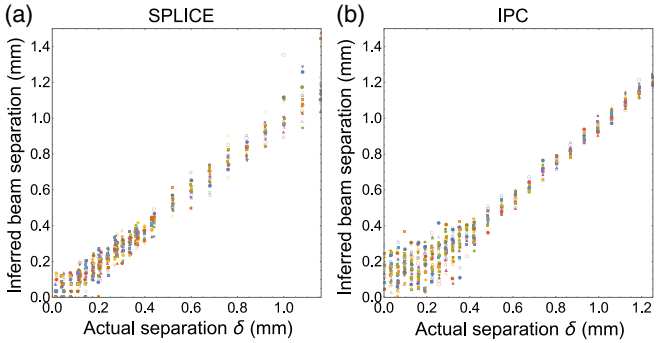

FIG. 3. Inferred separation vs known actual separation for (a) SPLICE (from lookup on calibration curve) and (b) IPC. Note that the spread in the IPC estimates grows drastically as $\delta \to 0$, while the spread for SPLICE remains essentially constant.

Whereas the ideal functional form for the resulting count rate vs separation δ is proportional to Eq. (2), we add a constant γ to account for residual background counts:

$$\text{SPLICE counts} = \alpha \exp\left(-\frac{(\delta-\delta_0)^2}{16\sigma_f^2}\right)\text{erf}^2\left(\frac{\delta}{4\sqrt{2}\sigma_f}\right) + \gamma. \tag{4}$$

A calibration curve was obtained from a least squares fit of this function to data from a longer run (2 sec count time instead of 1 sec for each δ). From a fit to the singles counts (counts which are not conditioned on the simultaneous detection of a heralding photon), the beam waist σ_f (0.46 mm) and δ_0 were extracted. Then, the normalization $\alpha(1206\text{s}^{-1})$ was extracted from a fit to coincidence counts. The background γ was fixed to an average of multiple values detected at a separation of δ_0 (2.73 mm). This step might be thought of as being analogous to characterizing one's imaging optics before use. One might then use the curve as a "lookup table" from which δ can be estimated from count rates. We performed such a lookup with the remainder of our data. The resulting estimates for δ are plotted versus their actual values (as reported by motorized translation stage controllers) in Fig. 3(a).

The traditional image plane counting data were acquired using the configuration described above, scanning the 200 μm slit between -1 and $+1$ mm of the centroid of the two beams, counting for 4 sec at each step. Again, we repeated this for various separations δ and, in turn, repeated the whole scan several times. As before, a set of coarse scans (-0.028 mm $\leq \delta \leq 1.25$ mm in steps of 0.064 mm, 16 repetitions) and a set of fine scans (-0.42 mm $\leq \delta \leq 0.35$ mm in steps of 0.032 mm, 17 repetitions) were performed. Estimation of δ in this IPC comparison was done by least-squares fitting the resulting image plane intensity profile to

$$\text{IPC counts} = \alpha[\exp_- + \exp_+ + \gamma], \quad (5)$$

where $\exp_\pm = \exp[-(x \pm \delta/2)^2/2\sigma^2]$. Again, a calibration waist σ and background γ were obtained beforehand, leaving the scale α and separation δ as the only fitting parameters. In practice, the fitting procedure used to fit IPC data for small δ was different from the one used to treat data for large δ. For the latter, we simply used built-in numerical algorithms in *Mathematica* and NumPy. For small δ's, however, the routines exhibited convergence and stability issues, forcing us to Taylor expand Eq. (5) to 2nd order in δ and manually invert the resulting 2×2 design matrix. The resulting estimated separations are plotted against actual separations in Fig. 3(b). As is immediately apparent, for separations below about 0.25 mm (approximately 0.6σ), the spread of the IPC data begins to grow, while that of the SPLICE data remains essentially constant.

Two key metrics for the performance of either method are the standard deviation (SD) (i.e., "spread") and root-mean-square error (RMSE) of the estimated beam separation. The SD measures the *precision* of a data set but not necessarily its accuracy, while the RMSE is sensitive to the accuracy since it quantifies the error relative to a known actual value and not simply the reported result. In Fig. 4, SD and RMSE are plotted versus known actual separations.

In order to ensure a reasonably even-footed comparison between IPC and SPLICE, the spreads in inferred separation plotted in [Fig. 4(a)] are scaled by $\sqrt{N}$. For IPC, N is simply the total photons that comprise an "image" on the image plane, which in our case is actually a set of photon counts, one at each position of the 200 μm slit. We measured $N = 2900$ and $N = 2800$ for the coarse and fine scan respectively. For SPLICE, during a calibration run, we estimate N by counting at our detector over a 1 sec window while both beams are centered (i.e., $\delta = 0$) on the coupler into TEM_{00} fiber with the phase plate removed. Since our source intensity is stable, this gives us an estimate of the number of incident photons for subsequent measurements when $\delta \neq 0$. For SPLICE, we measured $N = 1500$ and $N = 1200$ for the coarse and fine scans respectively.

The RMSE plotted in Fig. 4(b). is *not* similarly normalized because in addition to possible systematics, the inferred separation is biased relative to the actual separation when δ is small (see Fig. 5 and the Supplemental Material [25]). *A priori*, there is no reason to suspect either bias or systematics to scale as $\sqrt{N}$. Despite *not* normalizing and despite using approximately twice as many photons, the IPC method performs noticeably *worse* than SPLICE when $\delta < 0.6$ mm.

The attentive reader will note that while the spread is greater for IPC, it does not diverge as $\delta \to 0$. In fact, it would be implausible for the uncertainty on δ to ever exceed σ [as is clear from inspection of Fig. 4(b) at small δ]. The apparent discrepancy with the vanishing of the Fisher information can be understood by recognizing that the practically implemented IPC estimator is not unbiased. To better understand

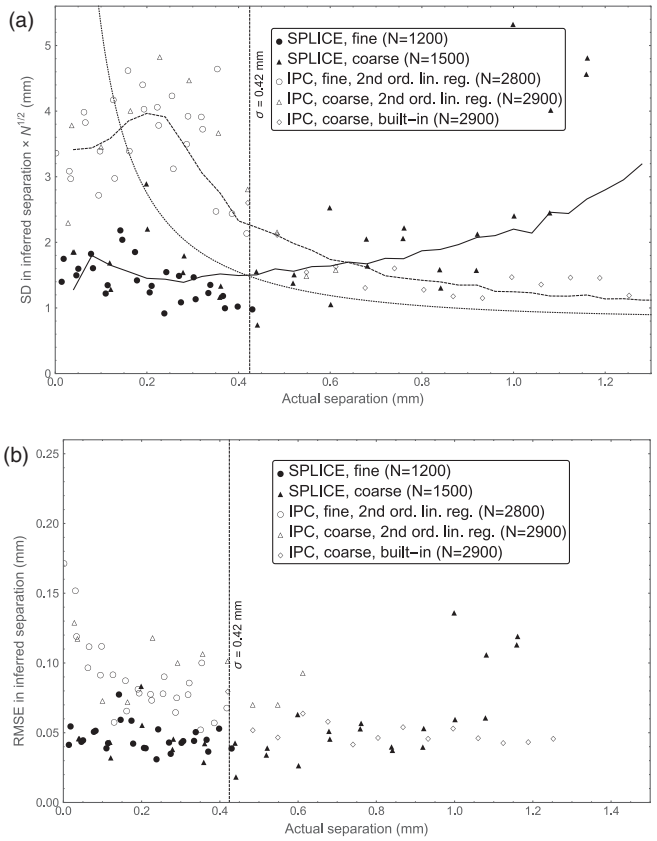

FIG. 4. (a) Renormalized standard deviation (SD) and (b) unnormalized root mean-square error in the estimated separation plotted as functions of actual separation for both IPC and SPLICE. We plot the computed SD multiplied by $\sqrt{N}$ to compensate for the scaling of the uncertainty with the size of the data set. The solid and dashed curves are the corresponding Monte Carlo simulations. The dotted curve is the CRB for IPC and the dashed horizontal line represents the absolute fundamental limit of $2\sigma/\sqrt{N}$. The RMSE (unlike SD) is not similarly rescaled. It allows us to gauge absolute error relative to the known value of the parameter being estimated so that biases are accounted for. Note that two methods were used in the fitting of IPC data to Eq. (5); for small $\delta(<0.65$ mm), Eq. (5) was expanded to 2nd order and linear regression was performed whereas for large $\delta(>0.4$ mm), a nonlinear fitting routine built into *Mathematica* was used.

the bounds on the advantage that one can expect of SPLICE over IPC, we return to Eq. (1). Clearly, one needs to know the bias to evaluate the RHS. For SPLICE, the only potential source of bias is the lookup procedure. If, for example, a less-than-perfect visibility results in a calibration curve that does not vanish at $\delta = 0$, then one might obtain "unphysical" data points that fall *under* the minima of the calibration curve, thereby resulting in a bias when a lookup is attempted. In our case, this is negligible since our visibility exceeds 99%. The CRLB is therefore just the reciprocal of I_f, implying a $1/\sqrt{N}$ scaling in the spread of δ_{est}.

With IPC, the least-squares estimate of δ is heavily biased at small δ. An intuitive way to understand this is to

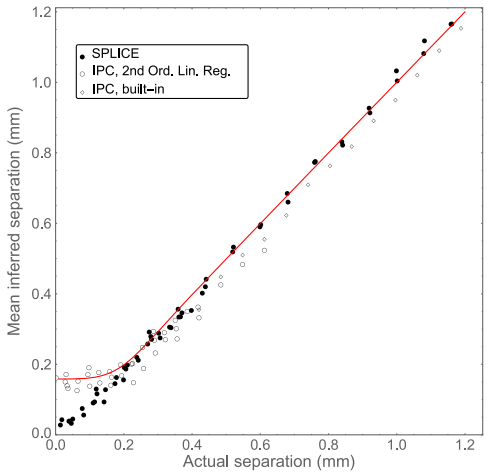

FIG. 5. Averaged measured separations. Mean estimated δ for IPC and SPLICE plotted against known actual δ. Two methods were used in the fitting of IPC data to equation (5); for small $\delta(<0.65$ mm), equation (5) was expanded to 2nd order and linear regression was performed whereas for large $\delta(>0.4$ mm), a nonlinear fitting routine built into *Mathematica* was used.

note that since the problem being addressed is the resolving of two *equal intensity* sources, the $+\delta$ and $-\delta$ cases are physically indistinguishable; therefore, what is really being estimated is the absolute value $|\delta|$. But as long as the spread in the estimated δ is nonzero, the *mean* estimated $|\delta|$ is never zero. Figure 5 shows a plot of mean inferred δ (averaged across all our data sets) vs actual δ. Overlayed is a theory curve for IPC, which takes into account an expected bias at small δ. The vanishing of the slope of this curve at low separations means that any attempt to invert it, generating an unbiased estimator, will introduce a diverging uncertainty.

In Ref. [25], we present theory showing that the bias term for IPC falls to -1 sufficiently quickly that the RHS of inequality (1) tends to a finite value as $\delta \to 0$. That finite value is shown to scale as $N^{-1/4}$, which is in stark contrast to the behavior of the spread at large δ (for IPC) as well as for SPLICE (at all δ), where the standard scaling of $N^{-1/2}$ is obeyed. This scaling is further substantiated with Monte Carlo simulations shown in a figure in Ref. [25]. Thus, while SPLICE does not offer an infinite advantage over IPC as a naive analysis might have us believe, it does nevertheless offer a substantial improvement in the absolute error and the scaling with photon number, while simultaneously eliminating the problem of bias.

In summary, we have developed and demonstrated a simple technique that surpasses traditional imaging in its ability to resolve two closely spaced point sources. For $\delta < 0.2$ mm(0.47σ), the average spread in the measured separation was approximately twice the quantum limit. Nearing zero separation, SPLICE has outperformed IPC by reducing the normalized standard deviation by a factor of 2 and the unnormalized total error by a factor of 3 despite the higher photon number used in IPC. Furthermore, unlike existing superresolution methods, ours requires no exotic illumination with particular coherence or quantum properties, and is applicable to classical incoherent sources. Crucially, as a proof of principle, this technique highlights that the fundamental limits on the precision with which one can estimate the separation between two point sources $(2\sigma/\sqrt{N})$ are independent of the separation itself. In traditional imaging techniques discarding the phase information (IPC), for separation below Rayleigh's criterion$(\delta \lesssim \sigma)$, the standard deviation in the measurement of separation goes as $2\sqrt{2}\sigma^2/(\delta\sqrt{N})$ for the best unbiased estimator or as approximately $\alpha\sigma/N^{1/4}$, where α is a numerical factor of the order of unity, when the bias becomes dominant. We expect that SPLICE and other related techniques that do not discard the phase information will be developed in the future for a broad range of imaging applications.

We would like to thank Mankei Tsang for useful discussions and the Facebook post which alerted us to their discovery and triggered this work. This work was funded by NSERC, CIFAR, and Northrop-Grumman Aerospace Systems *NG Next*.

W.-K. T. and H. F. contributed equally to this work.

Note added.—While preparing this manuscript, we learned that similar work was being pursued by Yang *et al.* [26] using inversion interferometry, Sheng, Durak, and Ling [27] using heterodyne detection, and Paúr *et al.* [28] using digital holography.

[1] L. Rayleigh, Philos. Mag. **8**, 261 (1879).
[2] S. W. Hell and J. Wichmann, Opt. Lett. **19**, 780 (1994).
[3] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, Science **313**, 1642 (2006).
[4] S. T. Hess, T. P. Girirajan, and M. D. Mason, Biophys. J. **91**, 4258 (2006).
[5] P. R. Hemmer and T. Zapata, J. Opt. **14**, 083002 (2012).
[6] V. Giovannetti, S. Lloyd, L. Maccone, and J. H. Shapiro, Phys. Rev. A **79**, 013827 (2009).
[7] M. Tsang, Phys. Rev. Lett. **102**, 253601 (2009).
[8] H. Shin, K. W. C. Chan, H. J. Chang, and R. W. Boyd, Phys. Rev. Lett. **107**, 083603 (2011).
[9] L. A. Rozema, J. D. Bateman, D. H. Mahler, R. Okamoto, A. Feizpour, A. Hayat, and A. M. Steinberg, Phys. Rev. Lett. **112**, 223602 (2014).
[10] F. Tamburini, G. Anzolin, G. Umbriaco, A. Bianchini, and C. Barbieri, Phys. Rev. Lett. **97**, 163903 (2006).
[11] O. Schwartz, J. M. Levitt, R. Tenne, S. Itzhakov, Z. Deutsch, and D. Oron, Nano Lett. **13**, 5832 (2013).
[12] V. Giovannetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).
[13] V. Giovannetti, S. Lloyd, and L. Maccone, Nat. Photonics **5**, 222 (2011).
[14] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg, Nature (London) **429**, 161 (2004).

[15] P. Walther, J.-W. Pan, M. Aspelmeyer, R. Ursin, S. Gasparoni, and A. Zeilinger, Nature (London) **429**, 158 (2004).
[16] T. Nagata, R. Okamoto, J. L. O'Brien, K. Sasaki, and S. Takeuchi, Science **316**, 726 (2007).
[17] M. Tsang, R. Nair, and X.-M. Lu, Phys. Rev. X **6**, 031033 (2016).
[18] G. Casella and R. L. Berger, *Statistical Inference* (Duxbury, Pacific Grove, 2002), Vol. 2.
[19] R. Nair and M. Tsang, Phys. Rev. Lett. **117**, 190801 (2016).
[20] C. Lupo and S. Pirandola, Phys. Rev. Lett. **117**, 190802 (2016).
[21] D. Richardson, J. Fini, and L. Nelson, Nat. Photonics **7**, 354 (2013).
[22] A. E. Siegman, *Lasers* (University Science Books, Mill Valley, 1986), p. 798.
[23] J. Rehacek, M. Paur, B. Stoklasa, L. Motka, Z. Hradil, and L. Sanchez-Soto, arXiv:1607.05837.
[24] M. Tsang, https://sites.google.com/site/mankeitsang/news/rayleigh/faq.
[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.118.070801 for more information on the $N^{-1/4}$ scaling of the error in δ with Image Plane Counting and further discussion.
[26] F. Yang, A. Taschilina, E. S. Moiseev, C. Simon, and A. I. Lvovsky, Optica **3**, 1148 (2016).
[27] T. Z. Sheng, K. Durak, and A. Ling, Opt. Express **24**, 22004 (2016).
[28] M. Paur, B. Stoklasa, Z. Hradil, L. L. Sanchez-Soto, and J. Rehacek, Optica **3**, 1144 (2016).