Experimental demonstration of superresolution of partially coherent light sources using parity sorting: supplement

S. A. WADOOD,[1,2] **KEVIN LIANG,**[1,2] **YIYU ZHOU,**[1,2] **JING YANG,**[2,3] **M. A. ALONSO,**[1,2,4] **X.-F. QIAN,**[5] **T. MALHOTRA,**[2,3,6] **S. M. HASHEMI RAFSANJANI,**[7] **ANDREW N. JORDAN,**[2,3,8] **ROBERT W. BOYD,**[1,2,3,9] **AND A. NICK VAMIVAKAS**[1,2,3,10,*]

[1]*The Institute of Optics, University of Rochester, Rochester, New York 14627, USA*

[2]*Center for Coherence and Quantum Optics, University of Rochester, Rochester, New York 14627, USA*

[3]*Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA*

[4]*Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, UMR 7249, 13397 Marseille Cedex 20, France*

[5]*Department of Physics and Center for Quantum Science and Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA*

[6]*Currently with Facebook Reality Labs, Redmond, WA, USA*

[7]*Department of Physics, University of Miami, Coral Gables, Florida 33146, USA*

[8]*Institute for Quantum Studies, Chapman University, Orange, California 92866, USA*

[9]*Department of Physics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

[10]*Materials Science, University of Rochester, Rochester, NY 14627, USA*

[*]*nick.vamivakas@rochester.edu*

Supplementary Material : Experimental demonstration of superresolution of partially coherent light sources using parity sorting

S. A. Wadood,[1,2] **Kevin Liang,**[1,2] **Yiyu Zhou,**[1,2] **Jing Yang,**[2,3] **M. A. Alonso,**[1,2,4] **X.-F. Qian,**[5] **T. Malhotra,**[2,3†] **S. M. Hashemi Rafsanjani,**[6] **Andrew N. Jordan,**[2,3,7] **Robert W. Boyd,**[1,2,3,8] **and A. Nick Vamivakas**[1,2,3,9].

[1]*The Institute of Optics, University of Rochester, Rochester, New York 14627, USA*

[2]*Center for Coherence and Quantum Optics, University of Rochester, Rochester, New York 14627, USA*

[3]*Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627, USA*

[4]*Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, UMR 7249, 13397 Marseille Cedex 20, France*

[5]*Department of Physics and Center for Quantum Science and Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA*

[6]*Department of Physics, University of Miami, Coral Gables, Florida 33146, USA*

[7]*Institute for Quantum Studies, Chapman University, Orange, California 92866, USA*

[8]*Department of Physics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

[9]*Materials Science, University of Rochester, Rochester, NY 14627, USA*

[†]*Currently with Facebook Reality Labs, Redmond, WA, USA*

[*]*Corresponding author: nick.vamivakas@rochester.edu*

Abstract: This document provides supplementary information to 'Experimental demonstration of superresolution of partially coherent light sources using parity sorting' by Wadood *et al.*

S1. Derivation of Eqs. (2-3) of the main text

In this section, we derive Eqs. (2-3) of the main text. In the object plane, consider two monochromatic point source emitters separated by δ, each emitting N_0 photons. The degree of (spatial) coherence between the two sources is γ. In the Coherent Mode Decomposition (CMD), the cross-spectral density (CSD) $W_O(x_1, x_2)$ at the object plane is given by

$$W_O(x_1, x_2) = N_0 \sum_{k=1}^{2} p_k \phi_k^*(x_1) \phi_k(x_2), \qquad (S1)$$

where $p_1 + p_2 = 1$, $p_1 - p_2 = \gamma$, $p_{1,2} = (1 \pm \gamma)/2$, and $\phi_k(x) = f_+(x) - e^{ik\pi} f_-(x)$ are the symmetric (k=1) and antisymmetric (k=2) coherent modes. In the object plane, $f_\pm(x) = \eta(x \pm \delta/2)$, where $\eta(x)$ are localized functions describing the point sources, and can be considered as regularized delta functions with a unit intensity integral, such that each source individually emits N_0 photons. Equation (S1) shows that total object plane photon number $\int dx W_O(x, x) = 2N_0(p_1 + p_2) = 2N_0$ is given by sum of intensities of the symmetric and antisymmetric modes. At the object plane, these coherent modes are orthogonal. In fact unitary propagation of the field preserves their orthogonal nature. However, the modes do not remain orthogonal after passing through the aperture of the imaging system, which is the fundamental cause of the loss of resolution.

Recall that the finite width of the point spread function (PSF) arises due to the nonunitary nature of the imaging system; the aperture behaves as a low pass filter, blocking photons of spatial frequencies higher than those determined by the system's numerical aperture [1]. Consider the canonical $4f$ imaging system with a Gaussian aperture of width σ_{ap} at the Fourier plane. For

simplicity, consider the lenses to be of infinite extent. Since the propagation up to the aperture is unitary, the number of photons hitting the aperture is $2N_0$. There are $2N_0 p_1 = N_0(1 + \gamma)$ photons in the symmetric and $2N_0 p_2 = N_0(1 - \gamma)$ photons in the antisymmetric mode hitting the aperture. At the aperture/Fourier plane, these modes behave like a cosine and sine, each with spatial frequency $g = k\delta/2f$, where k is the wavenumber and f is the focal length [1]. For a large δ such that $\sigma_{ap} g \gg 1$, the aperture samples many periods of the coherent modes. The aperture transmission is then a radiometric factor equal to the ratio σ_{ap}/σ_b, where σ_b is the beam width at the aperture. This factor is same for both the coherent modes in the limit $\sigma_{ap} g \gg 1$. As δ becomes smaller such that $\sigma_{ap} g \ll 1$, the aperture samples less than a fringe period of the coherent modes. This is the sub-Rayleigh regime of our interest. For the cosine modes, the aperture samples the peak of the fringe. For the sine modes, the aperture samples the fringe minimum. The smaller δ gets, the transmission for the sine mode decreases, and the transmission for the cosine mode increases. The transmitted photons $I_{1,2}$ for each mode are then given by

$$N_1 = 2N_0 p_1 \int dx \left(\frac{1}{8\pi\sigma_b^2} \right)^{1/2} 4\cos^2(gx)\, e^{-\frac{x^2}{2\sigma_b^2}}\, e^{-\frac{x^2}{2\sigma_{ap}^2}} = 2N_0 p_1 \frac{\sigma_{ap}}{\sigma_b} \left(1 + e^{-2g^2\sigma_{ap}^2} \right), \quad \text{(S2)}$$

$$N_2 = 2N_0 p_2 \int dx \left(\frac{1}{8\pi\sigma_b^2} \right)^{1/2} 4\sin^2(gx)\, e^{-\frac{x^2}{2\sigma_b^2}}\, e^{-\frac{x^2}{2\sigma_{ap}^2}} = 2N_0 p_2 \frac{\sigma_{ap}}{\sigma_b} \left(1 - e^{-2g^2\sigma_{ap}^2} \right). \quad \text{(S3)}$$

In deriving Eqs. (S2, S3), we have assumed the beam at the aperture is a cosine or sine times a broad Gaussian of width σ_b such that $\sigma_b \gg \sigma_{ap}$. Using $p_1 + p_2 = 1$, we find the total photons *transmitted* by the aperture is given by

$$N_t = N_1 + N_2 = 2N_0 \frac{\sigma_{ap}}{\sigma_b} \left(1 + \gamma e^{-2g^2\sigma_{ap}^2} \right). \quad \text{(S4)}$$

Eq. (S4) does not contradict energy conservation. The transmitted photon number $I_1 + I_2$ is not a strictly conserved quantity since it is the photon number in a local (Gaussian-apodized) portion of a fringe pattern. In essence, Eq. (S4) simply shows that the aperture transmission depends on the spatial coherence among two partially coherent emitters.

The propagation from the aperture to the image plane is also unitary, and hence Eqs. (S2-S4) also give the number of symmetric, antisymmetric, and total image plane photons. At the image plane, the Gaussian PSF's width σ is related to σ_{ap} by $k^2\sigma_{ap}^2/f^2 = 1/4\sigma^2$. We then have, using $g = k\delta/2f$, $\exp[-2g^2\sigma_{ap}^2] = \exp[-\delta^2/8\sigma^2] = d$, where d is the overlap integral of the Gaussian PSFs. The overlap is generated due to the nonunitary nature of the aperture, as is evident from our analysis. Moreover, the image plane CSD given in Eq. (1) of the main text is found by substituting $f \pm (x) = \frac{e^{-(x\pm\delta/2)^2/4\sigma^2}}{(2\pi\sigma^2)^{1/4}}$, and $\kappa = \sigma_{ap}/\sigma_b$ in Eq. (S1). Formally, the image and object plane CSDs can be connected by propagating the coherent modes through the nonunitary imaging system. Finally, by substituting $\kappa = \sigma_{ap}/\sigma_b$ and using $p_{1,2} = (1 \pm \gamma)/2$, Eqs. (S2-S4) reduce to Eqs. (2,3) of the main text.

S2. Derivation of Eqs. (4,5) of the main text

Assuming Poisson statistics and using Eqs. (S2, S3), we derive the FI matrix as

$$
J = \sum_{i=1,2} \begin{bmatrix} \frac{1}{N_i}\left(\frac{\partial N_i}{\partial \delta}\right)^2 & \frac{1}{N_i}\left(\frac{\partial N_i}{\partial \delta}\right)\left(\frac{\partial N_i}{\partial N_0}\right) \\ \frac{1}{N_i}\left(\frac{\partial N_i}{\partial N_0}\right)\left(\frac{\partial N_i}{\partial \delta}\right) & \frac{1}{N_i}\left(\frac{\partial N_i}{\partial N_0}\right)^2 \end{bmatrix},
$$

$$
= \begin{bmatrix} N_0 \frac{\delta^2 d^2}{8\sigma^4}\left(\frac{1-\gamma d}{1-d^2}\right) & -\frac{\gamma \delta d}{2\sigma^2} \\ -\frac{\gamma \delta d}{2\sigma^2} & \frac{2}{N_0}(1+\gamma d) \end{bmatrix}. \tag{S5}
$$

If only δ is unknown, the FI normalized by the object plane photon number $2N_0$ is described by $J(1,1)/2N_0$, which is equivalent to Eq. (4) of the main text and is plotted in Fig. (1a) of the main text.

If both δ and N_0 are unknown, the FI for estimating δ is found by $(C(1,1))^{-1}$, where $C = J^{-1}$, and $C(1,1)$ gives the multiparameter CRB. The matrix C is given as

$$
C = \frac{1}{\frac{\delta^2 d^2}{4\sigma^4}\left(\frac{1-\gamma^2}{1-d^2}\right)} \begin{bmatrix} \frac{2}{N_0}(1+\gamma d) & \frac{\gamma \delta d}{2\sigma^2} \\ \frac{\gamma \delta d}{2\sigma^2} & N_0 \frac{\delta^2 d^2}{8\sigma^4}\left(\frac{1-\gamma d}{1-d^2}\right) \end{bmatrix}. \tag{S6}
$$

The quantity $(C(1,1))^{-1}/2N_0$ is given by Eq. (5) of the main text and plotted in Fig. (1b) of the main text.

S3. Offline Coherence Synthesis and Modal Weights vs. γ

In this section, we describe the offline coherence synthesis and measurements of the normalized image plane modal weights $m_{1,2} = N_{1,2}/N_t$, where $N_t, N_{1,2}$ are given by Eqs. (2,3) of the main text.

Since we are working in the image plane model, we do not know N_0 as explained in the Theory section of the main text. For this section, we therefore denote the symmetric and antisymmetric experimental photon numbers as $N_{sym,asym}$ instead of $N_{1,2}$ as used in Eq. (3) of the manuscript. For a fixed δ, we prepare an antisymmetric mode and record its photon number N_{asym} at the output of the parity sorter. N_{asym} can be arbitrary but is dictated by experimental limitations such as phase noise of the interferometer, electronic noise and dynamic range of the detector etc. We then prepare a symmetric mode for the given δ, and record its flux at the output of the parity sorter over a wide range of input powers. For a particular γ and δ, the corresponding photon number N_{sym} of the symmetric mode is given as $N_{sym} = K N_{asym}$, where $K = (1+\gamma)(1+d)/(1-\gamma)(1-d)$. The modal weights are normalized by $N_t = N_{sym} + N_{asym}$ such that $m_{1,2} = N_{sym,asym}/(N_t)$. We then post-select values for the desired γ. For example, assume $\delta = 0.5\sigma$, and 1 photon in the antisymmetric mode, i.e., $N_{asym} = 1$. For $N_{sym} = \{0, 10, 64, 100\}$ photons in the symmetric mode, the γ would correspond to $\{-1, -0.72, 0, 0.2\}$ respectively. Note that the post-selected N_{sym} and N_{asym} are random variables which include both the quantum shot noise and the classical electronic and optical noise of the system. The post-selection is only applied to the 'mean' values of N_{asym} and N_{sym} satisfying $\langle N_{sym}\rangle = K\langle N_{asym}\rangle$, as explained in the Image processing section. We emphasize that experimentally, in the image plane model and offline coherence synthesis afforded by the CMD, N_t is defined as equal to $N_{sym} + N_{asym}$ (the total photons in the interferometer) and does not depend on δ or γ as in the object plane model of Eq. (2) of the main text.

Figure (S1) shows the measured modal weights for different γ values. The solid red and blue lines in Fig. (S1a-c) are theoretically expected plots of even and odd modal weights respectively. The dashed lines are the theoretically expected plots incorporating effects of cross-talk. The
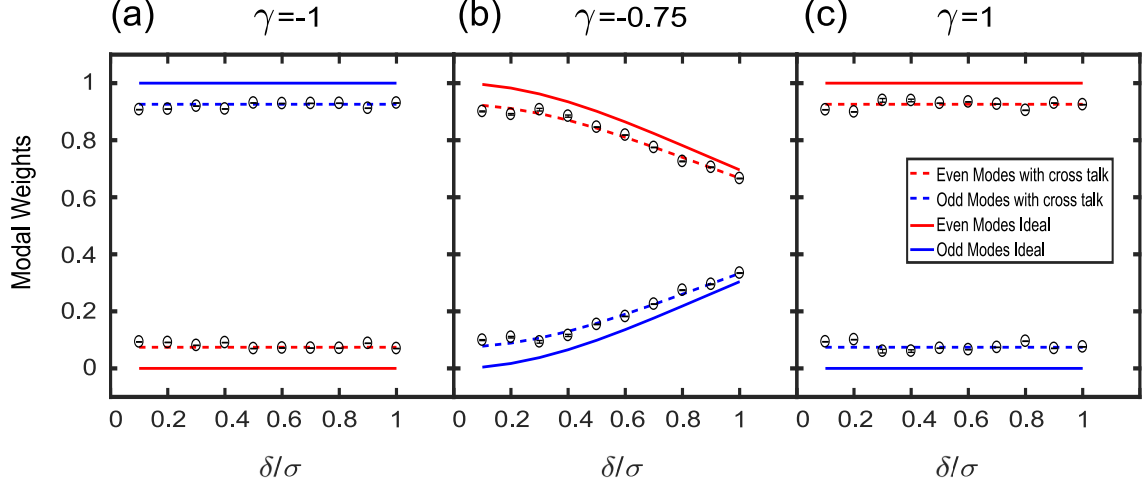
Fig. S1. Measured modal weights for the parity sorter. a-c: Modal weights for $\gamma = -1, 1, -0.75$. Blue and red color indicates the odd and even modal weights respectively. Solid lines indicate the theoretically expected modal weights with zero cross-talk. The dashed lines indicate the expected modal weights for 8% cross-talk. The circles indicate the measured values. All the modal weights are normalized by N_t, the total number of photons in the interferometer. Each point on the graphs represents a mean of 10 measurements, while the error bars are too small to be noticed on the graph. Note that for $|\gamma| = 1$, the modal weights are constant versus δ.

circles indicate the mean values of data recorded for 10 measurements. For $|\gamma| = 1$, all the optical power is directed into a single port; $P_1 = 1$ (solid red line) and $P_2 = 0$ (solid blue line) for $\gamma = 1$. Similarly $P_1 = 0$ and $P_2 = 1$ for $\gamma = -1$. If the optical power in either port does not change as a function of δ, which is the parameter to be estimated, one would expect no information to be gained by parity sorting. The error bars on all figures are too small to be visible on the graph. Finally, Fig. (S1 b) shows the modal weights measured for $\gamma = -0.75$. The modal weights change rapidly with δ, and therefore one would expect a higher FI for $\gamma = -0.75$ than for $\gamma = 1, -1$.

When processing the measured optical powers for the maximum likelihood estimation (MLE), we have subtracted any cross-talk, which was 8% on average, between the two ports. The cross-talk could be attributed to intensity mismatch of the two beams, finite coherence time, and fluctuations of path length and polarization in the two arms. This background subtraction is allowed because we are generating coherence offline by adding post-selected intensities, and only one of the coherent modes is present at a given time. Better alignment of the system can reduce the cross-talk. However, this experimental complexity is irrelevant to the analysis of partial coherence and thus is avoided in our proof-of-principle experiment by using offline coherence generation. Figure (S2) shows the cross-talk subtracted modal weights measured for $\gamma = 0, \pm, 0.25, \pm0.5$. Note that for a given δ, the power in the symmetric (antisymmetric) mode increases (decreases) as γ becomes more positive.

S4. Experimental Details

S4.1. Mode Intensity vs. δ

In our superresolution experiment we directly prepare the image plane CSD, represented by Eq. (1) of the main text, through the SLM and subsequent offline addition of modes. Specifically, this technique allows us to avoid the more complicated procedure of preparing two actual emitters with varying γ and δ, with a CSD given by Eq. (S1), and a physical Gaussian aperture whose
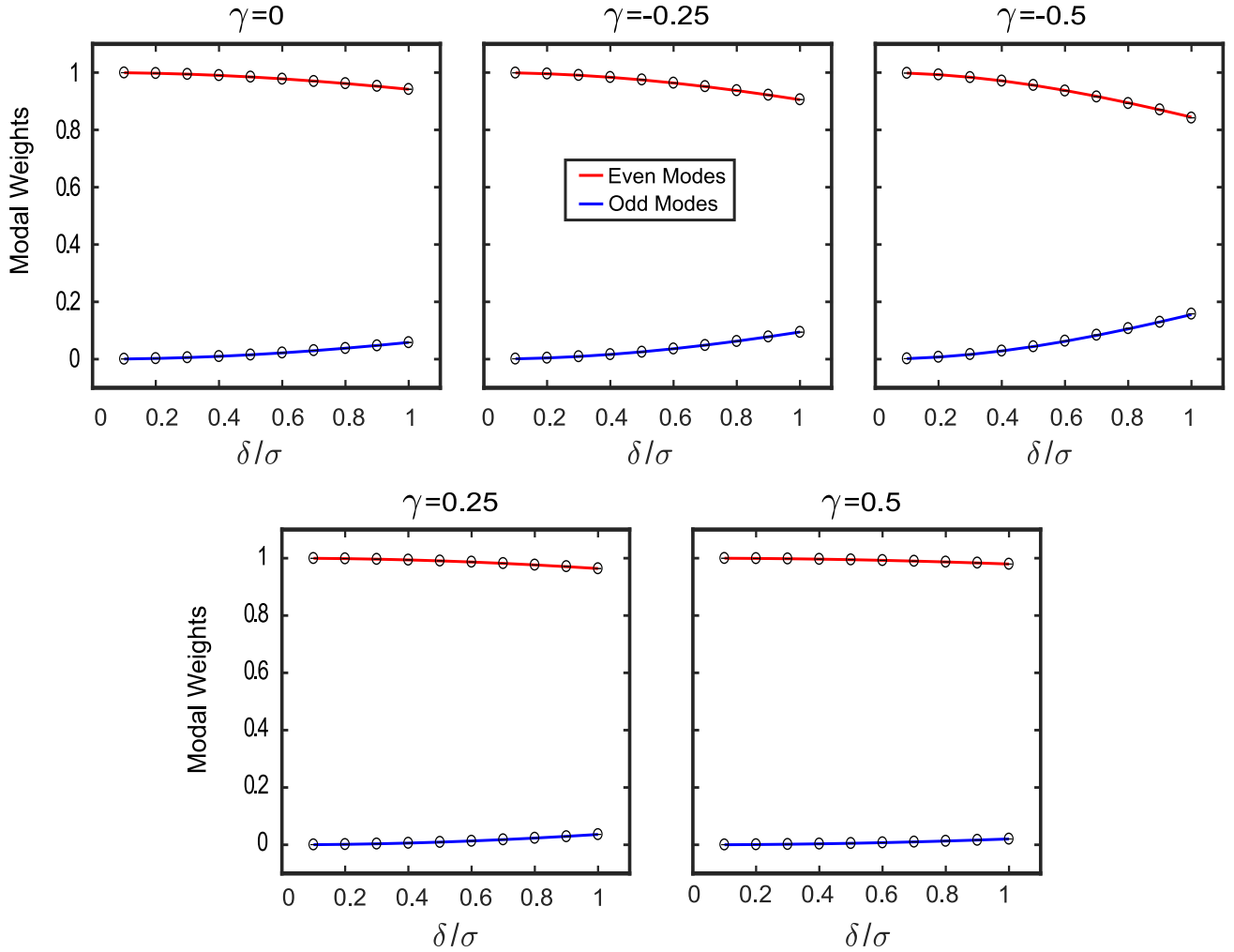
Fig. S2. Measured and expected modal weights for the parity sorter for $\gamma = 0, \pm0.25, \pm0.5$. Blue and red solid lines respectively indicate the theoretically expected modal weights for antisymmetric and symmetric mode. The circles indicate the measured modal weights. Each point on the graphs represents a mean of 10 measurements, while the error bars are too small to be noticed on the graph. Note that cross-talk has been subtracted for this data.

transmission varies according to Eq. (S2-S4). To this end, the power in the coherent modes does not depend on δ. This is because the SLM is fundamentally a phase grating whose first diffraction order contains the desired coherent mode. Ideally, the efficiency of the grating is independent of δ, except for the zero-photon case of $\gamma = -1$ and $\delta = 0$ in which case no grating phase is implemented. As long as the SLM pixel pitch and phase modulation range are not the bottleneck for a given phase grating, the grating efficiency and hence the power in a mode will not change with δ. Figure (S3) shows the power in the symmetric and antisymmetric modes measured for a fixed input power incident on the SLM. We see that the mode power is independent of δ. The grating efficiency for the symmetric mode is actually lower than the antisymmetric mode, which causes the lower power in the symmetric mode. We can control the power going to the interferometer with polarization optics, and can control the powers of the mode *after* the SLM.
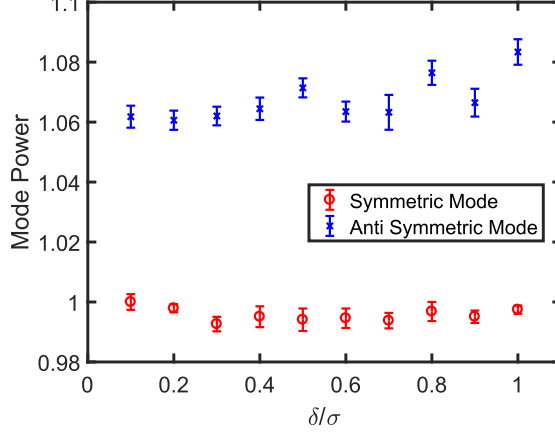
Fig. S3. Power in the Coherent modes versus δ for fixed input laser power impinging on the SLM. The power has been normalized to the power in the Symmetric mode for $\delta = 0.1\sigma$. Note that the power in the antisymmetric mode is consistently higher because of the higher diffraction efficiency of the SLM phase grating required to generate the antisymmetric mode. Ten images were taken for each data point.

S4.2. Image Processing

The basic idea of coherent mode decomposition (CMD) is to add the symmetric and antisymmetric modes on an intensity basis. For a specific δ and γ, the relative intensity of the symmetric and antisymmetric mode is $m_1/m_2 = K = (1+\gamma)(1+d)/(1-\gamma)(1-d)$ as evident from Eq. (3) in the main text. For example, if $\gamma = 0$ and $d = \exp(-\delta^2/8\sigma^2) = 0.5$ (which implies that $(\delta/\sigma)^2 = -8\ln[0.5]$), then $m_1/m_2 = 3$. To 'prepare' the $\gamma = 0$ intensity offline for this δ, one then needs to have a symmetric mode with thrice the power of the antisymmetric mode. As will be explained below, we post-select the required powers for each coherent mode from an array of recorded powers. These post-selected powers are then used to obtain the ML estimate $\hat{\delta}$. This process is repeated 100 times and the variance $\text{Var}[\hat{\delta}]$ is then compared with the CRB.

We measure $\text{Var}[\hat{\delta}]$ for $\gamma = 0$ and $\gamma = -0.75$. For these γ values and $\delta/\sigma < 1$, the power in the symmetric mode is always larger than the power in the antisymmetric modes, as is evident from Figs. (S1,S2). Therefore, a large dynammic range is required to accurately measure these modal weights. To increase the dynamic range of our measured powers, we use the fact that our problem is 1D (x dimension), while the coherent mode is 2D; the coherent modes are a symmetric zeroth order Gaussian in y. The integrated power in each row of the coherent mode then corresponds to an independent parity sorting measurement. N rows in the mode then correspond to N bucket detectors performing parity sorting measurements. Figure (S4) shows an image of the coherent modes and the marginals for $\delta = 0.4\sigma$. Specifically, each point on the Y marginal is used as an independent realization of a parity sorting measurement. This is allowed because we are using an approximately single-mode laser source, for which the pixels on the CCD are uncorrelated in intensity [2]. It is from these Y marginals of the symmetric and antisymmetric modes that we choose the photon numbers $N_{sym,asym}$ such that $N_{sym} = K N_{asym}$.

To saturate the Cramer–Rao bound (CRB), the system should be shot-noise limited. This means that the photon numbers used from the Y marginals should have Poisson statistics. We use a low noise Gigajot CCD with a dark current of less than $0.5\ e-/s/$pixel at room temperature [3]. The photoelectron statistics of each arm of the interferometer are shot noise limited, as can be found by blocking the other arm. We show the noise statistics of the interferometer output in Fig. (S5). For
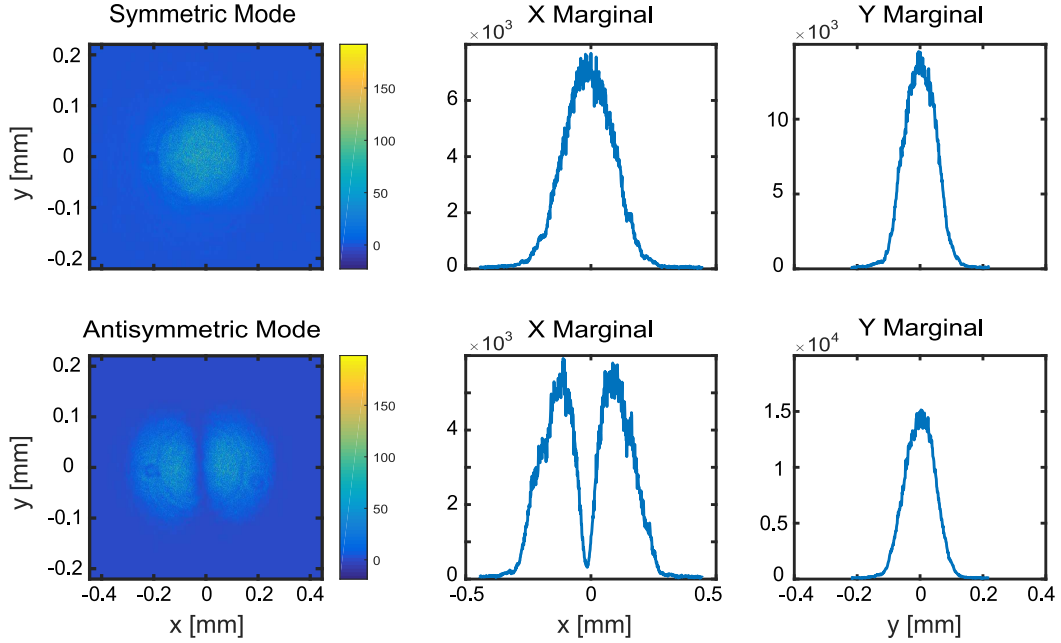
Fig. S4. Sample Image of Coherent Modes and their marginals for $\delta = 0.4\sigma$. Each pixel value on the Y marginal is counted as a bucket detector output in the parity sorter. The colorbar on the mode images and the y-axis on the marginals represent the detected photoelectron number.

every pixel on the Y marginal, we have three numbers: (i) the mean intensity $\langle I(y) \rangle$ (circles and solid blue line), where $I(y) = \int dx |\phi(x, y)|^2$ is the integrated intensity across x and the ensemble average is taken across the 100 acquired images. (ii) The variance $\langle \Delta I^2(y) \rangle = \langle I^2(y) \rangle - \langle I(y) \rangle^2$ (solid orange line), and (iii) the noise metric $M(y) = (\langle \Delta I^2(y) \rangle - \langle I(y) \rangle)/\langle I(y) \rangle$ (cross and solid purple line). $M \approx 0$ corresponds to shot noise statistics, while any classical (electronic or optical) fluctuations will increase the value of M. In Fig. (S5), the reference arm statistics shows that the centre portion of the Y marginal Gaussian is close to being shot noise limited because of the high intensity, as opposed to the tails of the Gaussian. In contrast, the interference images, for which both arms are unblocked, have high M values in the centre of the Gaussian. Note that the classical noise is more pronounced in the antisymmetric mode compared to the symmetric mode.

The major noise source in our interferometer is phase noise due to path length fluctuations in the interferometer. Moreover, the phase noise fluctuations do not affect the coherent modes globally. This is due to the mode inhomogeneities introduced due to phase jumps in the SLM arm of the interferometer [4] that implements the $2f$ system. The peaks in the noise of the interference images in Fig. (S5) are due to the mode inhomogeneities caused mainly due to phase rings of the SLM coupled with the path length fluctuations. However, there is a significant number of rows that are are relatively less affected by phase fluctuations such that $M(y) < 1$. These are the rows we use in the ML estimator. We sort the rows in ascending order according to the metric $M(y)$, and use these sorted rows in the CMD; rows with the lowest absolute $M(y)$ value is picked first for CMD. Note that $M(y) < 0$ is possible because of the finite size of our ensemble of 100 images, and we therefore sort based on $|M(y)|$ rather than $M(y)$. A consequence of our CMD algorithm is that data for a higher γ will have a higher photon number. This is because for a higher γ on the real line, the ratio m_1/m_2 increases as $1/(1-\gamma)$, and hence we require more rows of the symmetric mode to generate the required coherent mode power. Another effect of this
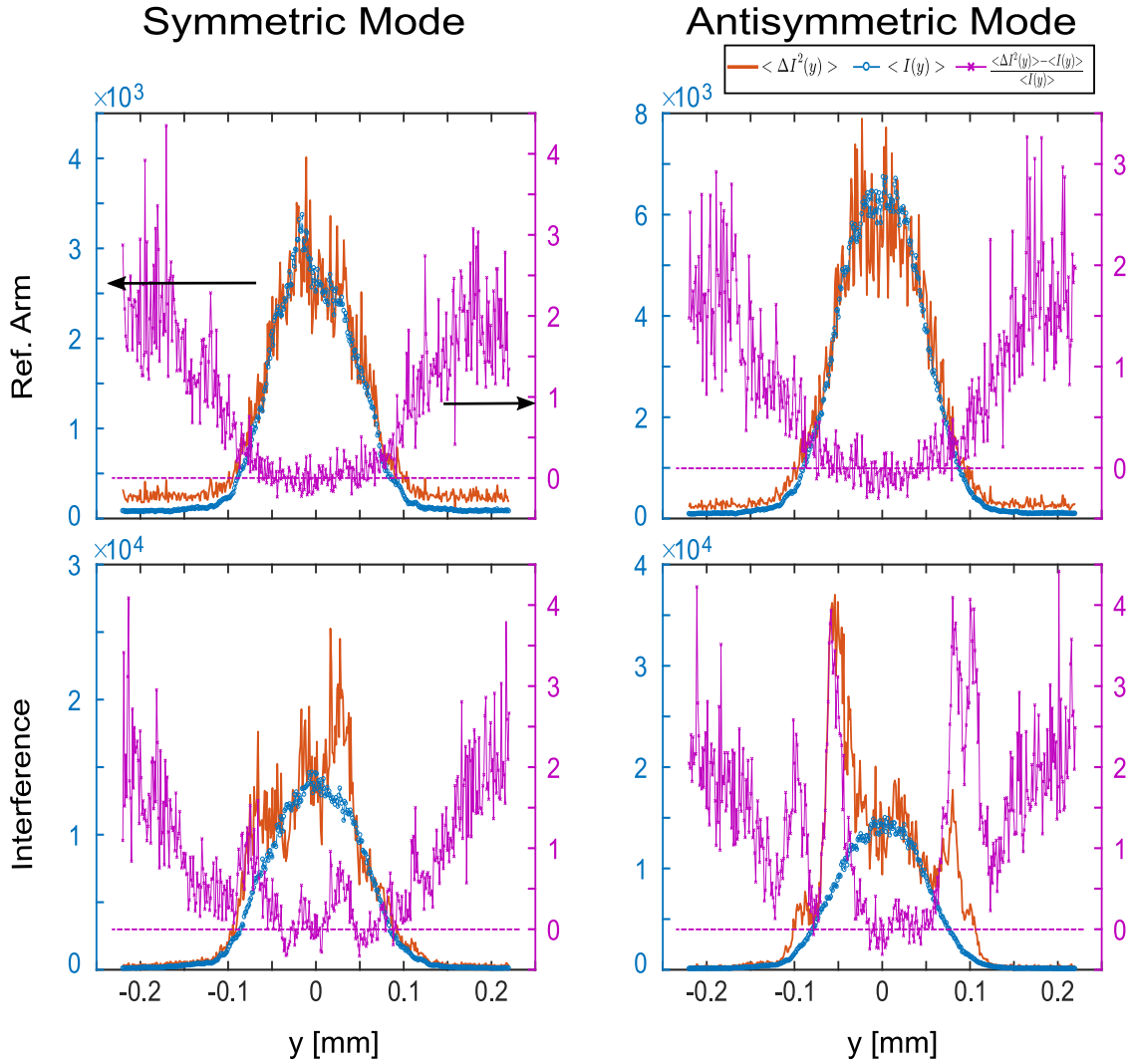
Fig. S5. Noise statistics of the parity sorter. All figures show the statistics of the Y marginal. Top row shows the statistics of the reference arm, with the $2f$ imaging SLM arm blocked. Bottom row shows the statistics of the interference images, with both arms unblocked. Solid blue lines and circles represent the mean value $\langle I(y) \rangle$ over 100 images (blue y-axis on left). Solid orange line represents the variance $\langle \Delta I^2(y) \rangle$. Solid purple line and crosses represent the noise metric $M(y)$ (purple y-axis on right). The dashed purple line represents the shot noise limit $M = 0$. Note that the two reference and interference images were taken separately with different laser powers, so the reference image power is not exactly half that of the interference image power. The text discusses details of the figure.
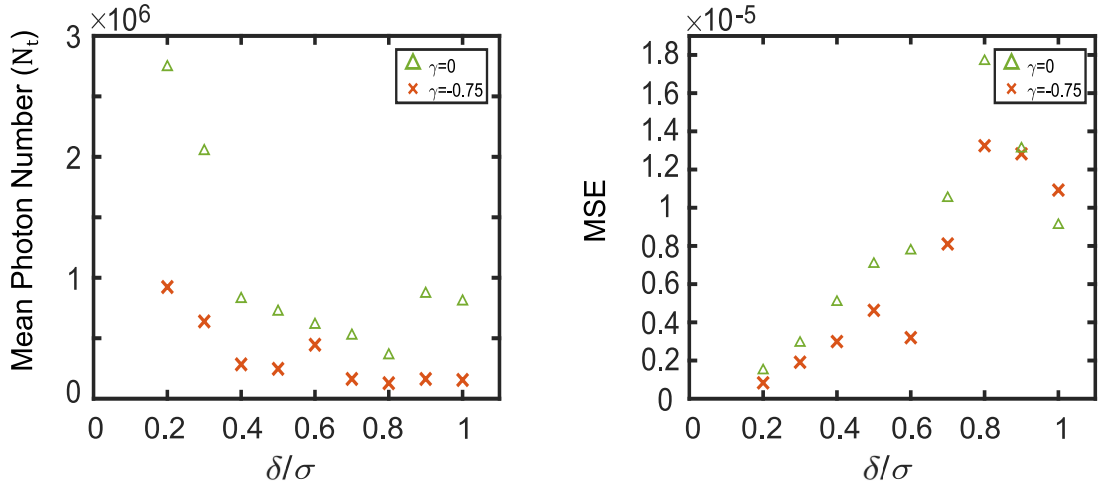
Fig. S6. Left: Mean total image plane photon number N_t per iteration used in estimation of δ. Note that the photon number for $\gamma = 0$ (green triangles) is consistently higher than the photon number for $\gamma = -0.75$. Right: The MSE for $\gamma = 0, -0.75$. The MSE for $\gamma = -0.75$ is less than the MSE for the $\gamma = 0$ case for all but the $\delta = 0.9\sigma$ case.

algorithm is that data for a higher γ will have higher (classical) noise, since more rows are added. The difference in the noise due to the different number of rows is observed to be not significant, however, in our experiment. One way to mitigate this noise difference is to take more images and add them to get a set of 100 images. This will decrease phase fluctuations due to averaging. For $\delta = 0.2 - 0.5$, we took $[500 - 200]$ images and added them together to prepare 100 images each for processing. In Fig. (3) of the manuscript, we compare $\gamma = 0$ and $\gamma = -0.75$, for which the number of rows do not differ significantly and hence the phase noise is approximately the same for both γ cases.

S4.3. Photon number in Fig. (3) of Main text

Figure (3b) of the main text compares the product of the mean squared error ($\text{MSE}/4\sigma^2$) and the mean total image plane photon number (N_t) used for each iteration in the 100 estimates. This was done because the CRB lowerbounds the product $N_t \text{Var}[\hat{\delta}]$ and not just $\text{Var}[\hat{\delta}]$. It is still instructive to look at the actual photon number and the mean squared error separately. As explained in the Image Processing section, the mean photon number is not the same for $\gamma = 0$ and $\gamma = -0.75$. In Fig. (S6) we show the actual values of the photon number and the mean squared error, i.e., $\text{Var}[\hat{\delta}]$ for both $\gamma = 0, -0.75$. Note that the photon number for $\gamma = 0$ is consistently *higher* than the photon number for $\gamma = -0.75$. Despite that, because the FI for $\gamma = -0.75$ is higher than the FI for $\gamma = 0$ case, the variance for partially anticorrelated case is *lower* than the variance for the incoherent case for all but the $\delta = 0.9\sigma$ case. Had the photon number in the $\gamma = -0.75$ case been increased to match the $\gamma = 0$ photon number, the variance in the $\gamma = -0.75$ case would have dropped even further. This shows unambiguously that for the same number of photons, partial anticorrelation among the sources leads to a lower variance than the incoherent scenario.

References

1. J. W. Goodman, *Introduction to Fourier Optics* (Roberts and Company Publishers, 2005).
2. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).
3. J. Ma, S. Masoodian, D. A. Starkey, and E. R. Fossum, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," Optica **4**, 1474–1481 (2017).
4. T. Malhotra, W. E. Farriss, J. Hassett, A. F. Abouraddy, J. R. Fienup, and A. N. Vamivakas, "Interferometric spatial mode analyzer with a bucket detector," Optics express **26**, 8719–8728 (2018).