Supplemental Document

OPTICA

Full Poincaré polarimetry enabled through physical inference: supplement

Chao He,[1,*] Jianyu Lin,[2,3] Jintao Chang,[4] Jacopo Antonello,[1] Ben Dai,[5] Jingyu Wang,[1] Jiahe Cui,[1] Ji Qi,[2,3] Min Wu,[6] Daniel S. Elson,[2,3] Peng Xi,[7] Andrew Forbes,[8] and Martin J. Booth[1,9]

[1]*Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK*
[2]*Hamlyn Centre for Robotic Surgery, Imperial College London, London SW7 2AZ, UK*
[3]*Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, UK*
[4]*Department of Physics, Tsinghua University, Beijing 100084, China*
[5]*Department of Statistics, The Chinese University of Hong Kong, Shatin, HK SAR, China*
[6]*Department of Computer Science, University of Oxford, Parks Road, Oxford OX1 3QD, UK*
[7]*Department of Biomedical Engineering, College of Engineering, Peking University, Beijing 100871, China*
[8]*School of Physics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa*
[9]*e-mail: martin.booth@eng.ox.ac.uk*
[*]*Corresponding author: chao.he@eng.ox.ac.uk*

Supplement DOI: https://doi.org/10.6084/m9.figshare.20715868

Parent Article DOI: https://doi.org/10.1364/OPTICA.452646

Full Poincaré polarimetry enabled through physical inference: supplemental document

Chao He[1,9], Jianyu Lin[2,3], Jintao Chang[4], Jacopo Antonello[1], Ben Dai[5], Jingyu Wang[1], Jiahe Cui[1], Ji Qi[2,3], Min Wu[6], Daniel S. Elson[2,3], Peng Xi[7], Andrew Forbes[8] and Martin J. Booth[1,10]

[1]*Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK*
[2]*Hamlyn Centre for Robotic Surgery, Imperial College London, London SW7 2AZ, UK*
[3]*Department of Surgery and Cancer, Imperial College London, London SW7 2AZ, UK*
[4]*Department of Physics, Tsinghua University, Beijing 100084, China*
[5]*Department of Statistics, The Chinese University of Hong Kong, Shatin, HK SAR, China*
[6]*Department of Computer Science, University of Oxford, Parks Road, Oxford, OX1 3QD, UK*
[7]*Department of Biomedical Engineering, College of Engineering, Peking University, Beijing 100871, China*
[8]*School of Physics, University of the Witwatersrand, Private Bag 3, Johannesburg 2050, South Africa*
[9]*chao.he@eng.ox.ac.uk*
[10]*martin.booth@eng.ox.ac.uk*

Supplementary Note 1: Universal full Poincaré generator and full Poincaré beam polarimetry

There are two types of systems that can generate a full Poincaré beam (FPB) [1–3]. The first type – shown in Supplementary Figure 1a – has the functionality of transferring fixed (or limited state) state of polarisation (SOP) into a FPB. A typical system configuration is based on two liquid crystal spatial light modulators (SLMs) or a system using multiple passes from a single SLM [4]. Under such a geometry, due to the SLM having a uniformly distributed slow/fast axis orientation, it is strongly polarisation dependent. It can be used to generate the FPB to some extent but cannot meet the requirements of generating a FPB with an arbitrary incident SOP. For example, if we use a uniform linear SOP that is aligned in the same direction as the fast axis orientation of the first SLM, then the modulation of such a pass would lose all functionality. Hence with only one degree of freedom introduced by the second SLM (or the second pass) an arbitrary SOP cannot be generated [4,5]. Under our definition in the main article, such a system cannot be referred to as a universal full Poincare generator (UFPG).

The second type of system is based on the functionality of transferring each arbitrary SOP into a different FPB (see Supplementary Figure 1b), which is the UFPG we defined above. It can be a linear retarder array – such as the GRIN lens we used in the main article; or a mixed diattenuator array (see Supplementary Figure 1b). The former is a combination of all effective fast axis orientations (θ from 0° to 180°) and retardance values (δ from 0° to 180°). The latter contains all possibilities of the eigenvector (determined via transmission axis orientation θ' and eigenvector elliptical ratio b/a). The UFPG (GRIN lens) that we used in the main article, has spatially variant birefringence that provides, in effect, the full range of retardances required with order number two (see Supplementary Figure 1c). Note that the UFPG system is not limited to these two types but could also in principle be generated by other mechanisms.

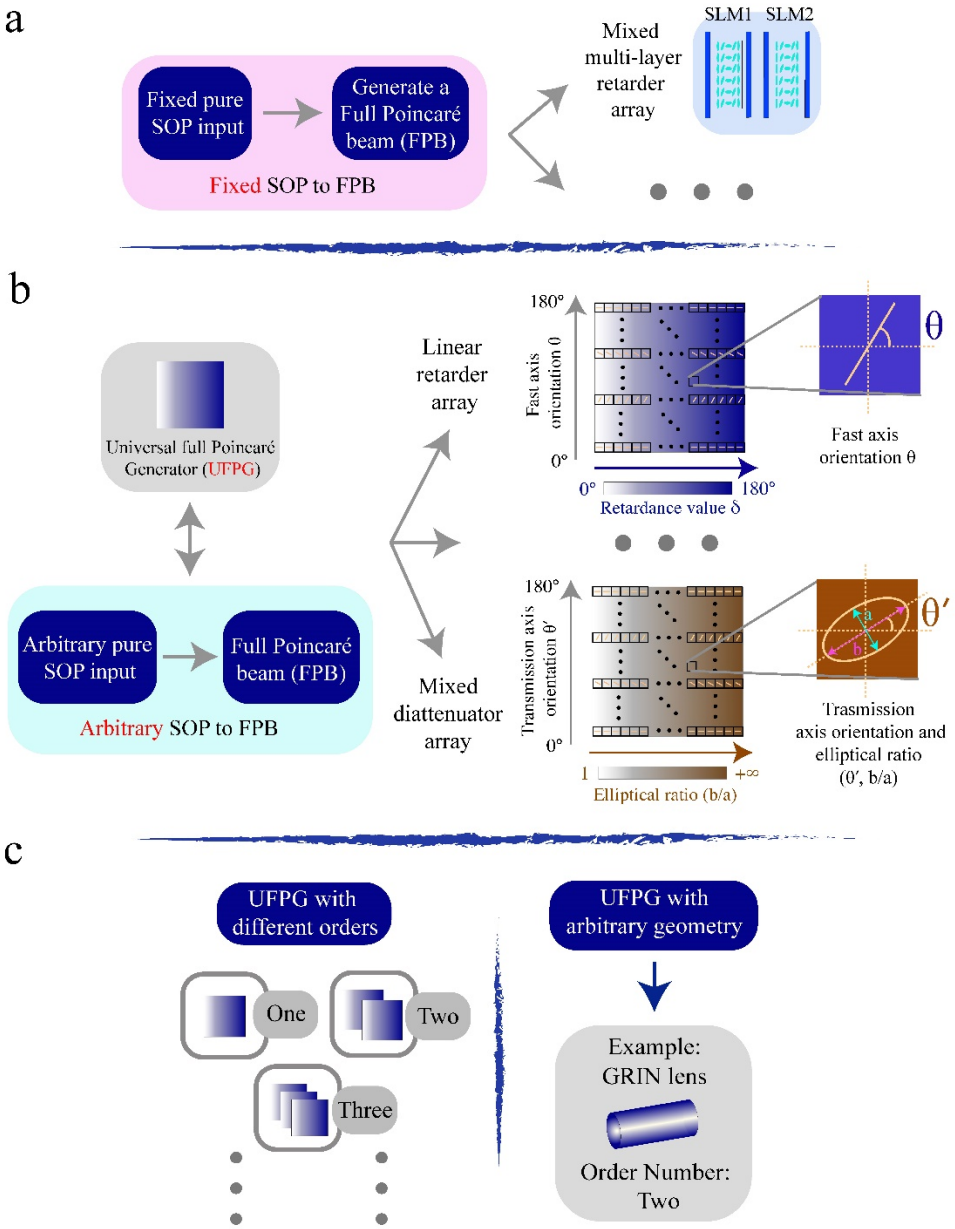

Fig. S1. Two types of systems that can generate a FPB. (a) Traditional FPB generation methods – restricted by the ability of generating a FPB from only a fixed incident SOP. The device commonly consists of a mixed-layered retarder array, such as multiple SLMs. (b) The UFPG as defined in this paper – any pure incident SOP can be used to generate a different FPB. The detailed mathematical proof of such property can be found in Method. Two typical examples are given, one is a linear retarder array (with both the value of the eigenvector and eigenvalue ranging from 0° to 180°) and the second is a mixed diattenuator array (with all possible eigenvectors, θ' ranging from 0° to 180°, and b/a from 1 to +∞). (c) The order number of the UFPG system corresponding to the order of the underlying FPB. A practical implementation of a UFPG through a GRIN lens, with order number two, consisting of linear retarders. Here, the full range of retardances and angles is provided through the spatial variation of birefringence in the GRIN medium [6–8].

There exist various methods of FPB polarimetry ranging from Stokes vector polarimeter to Mueller matrix polarimeter [9–14], utilizing devices include stress engineered optics (SEO), GRIN lenses, multi-core fibres (MCF), and uniaxial crystals. However, their processes of Stokes vector retrieval rely upon matrix calculation (domain 1 in Supplementary Figure 2), the same as other non-FPB polarimetry [9]. Our new paradigm is in a separate domain (domain 2 in Supplementary Figure 2) that parallels the traditional one. As shown in the figure, existing works (within the yellow region) are based on matrix calculation; however, a UFPG based system can enable a new paradigm (within the purple region) via image processing (such as machine learning (ML)), which uses inference from a physical model of locating the brightest points to retrieve Stokes vector. To the best of our knowledge, such a UFPG-based paradigm has not been shown, even though similar devices for domain 2 (UFPG) have been harnessed for methods in domain 1. Domain 2 offers two important features – 1) circumventing the error amplification factor (see details in Supplementary Note 3) and 2) an 'end to end' solution for SOP retrieval (see details in Supplementary Note 4).

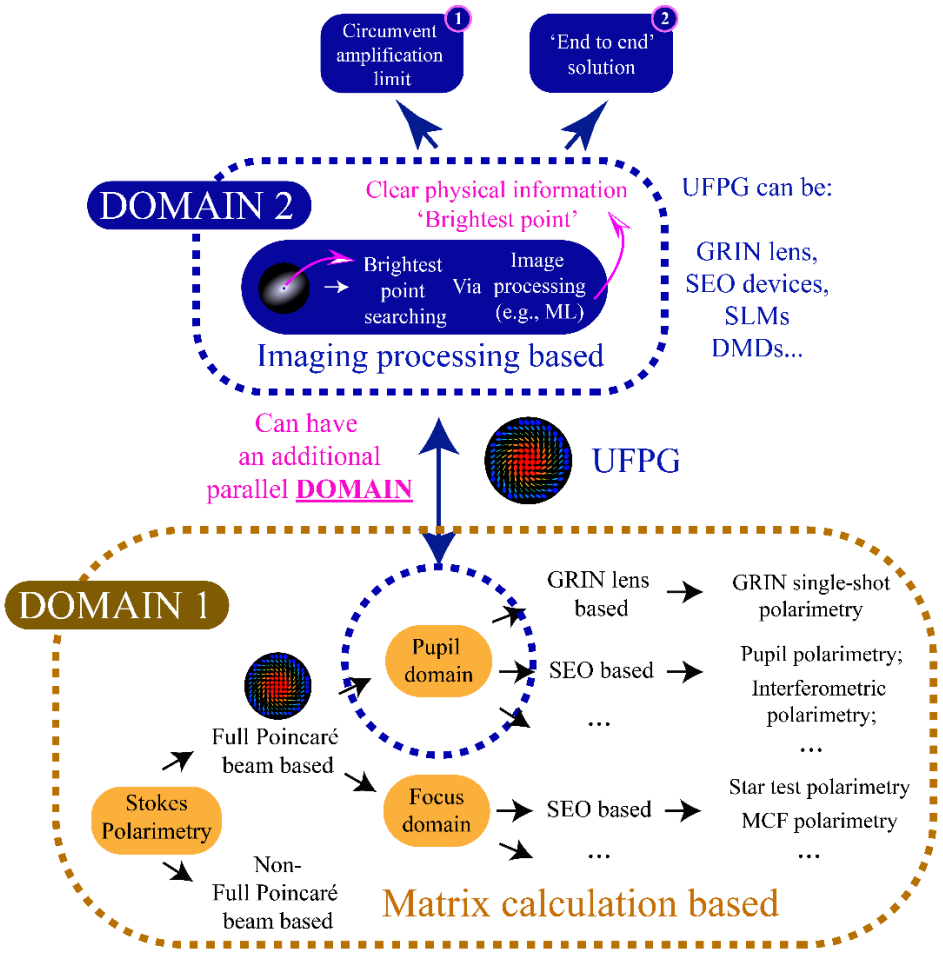

Fig. S2. Two domains of current Stokes polarimetry. Domain 1 consists of matrix calculation-based Stokes retrieval approaches, while domain 2 utilizes imaging processing to retrieve Stokes vector, with an emphasis on inference from clear physical information.

Supplementary Note 2: Experimental set-up and operating principles of universal full Poincaré generator polarimetry

The experiments to validate the feasibility of the PFG based paradigm were performed with the set-up that consists of a spatial light modulator (SLM, Hamamatsu, X10468−02) based polarisation state generator (PSG) and a GRIN lens-based polarisation state analyser (PSA) (Supplementary Figure 3a). We used a LED (3 W, 633 nm, $\Delta\lambda$ = 20 nm) as the illumination source, then the light beam passed through a polariser (Thorlabs, LPVIS050) to generate linear polarised light at 45° with respect to the slow (modulating) axis orientation of the SLM. Here we adopted the double-path geometry using a single SLM according to the method in Ref[5], in order to generate an arbitrary SOP. This PSG (yellow dashed box) was used to generate training data for the convolutional neural network to reconstruct the polarisation filter (PF) determined vector map (see later sections). On the analysis side (blue dashed box), the UFPG system we used was a GRIN lens (Femto Technology Co. Ltd., G-B161157-S1484, NA=0.1, Pitch=2, the same as that used in Ref[8]) followed by a fixed circular polariser (CP) (Thorlabs, CP1L633), which formed the PSA. The intensity images were obtained by a camera (Thorlabs, DCC3240N). After the network was trained, the PSG was removed from the system before conducting the actual sensing experiments. As we have discussed before, the order number of the UFPG is extensible and the distribution of retardance (or diattentuation) is not confined to a particular physical arrangement. Supplementary Figure 3b shows the link between an UFPG with order number two and a GRIN lens, which was used throughout this paper. In the GRIN lens, the retardances are arranged with circular symmetry, rather than the conceptual Cartesian arrangement shown in Fig. 1 in the main article. Note a pinhole is utilized at the focal plane of the UFPG system, based on which we can assume the incident SOP is uniform before being analysed by the polarimeter. In this work, we focus on 2D polarimetry.

The UFPG paradigm is based on the concept of 'full generation/analysis'. As a UFPG is a complete system that contains all possible SOP generating channels, conceptually through a continuous array of polarisation elements, it can create a FPB in a single shot. Those channels can also perform 'full analysis' of any incident SOP, in effect through the array of polarisation elements. A GRIN lens with appropriately chosen properties can provide a practical UFPG, as it behaves as a spatially variant waveplate, due to its intrinsic birefringence. For any uniform SOP at the GRIN lens input, a FPB will be created at the GRIN lens output. The PF following the GRIN lens post-selects the chosen PF eigenstate, effectively by projecting the output state at each field point onto the eigenstate. This creates an intensity distribution that is characteristic of the state present at the GRIN lens input. The brightest points of this distribution correspond to the points at which the GRIN lens outputs the SOP equivalent to the PF eigenstate, while the same position of the brightest point on the PF-determined map (which is a certain vector field generated via the PF eigenstate incident GRIN lens under the local coordinate system) is the incident SOP. If the input state changes, then the positions of the brightest points change. There exists therefore a mapping between the input states and the output intensity distribution. For instance, when a 45° linear SOP illuminates the GRIN lens before being filtered by a PF that contains the eigenstate of right hand circular (RC) SOP, the final intensity would have two brightest points at the locations with exact RC-SOP (the number of points is determined by the order of UFPG; for the GRIN lens there are two points). Note that the PF-determined map can only be changed through using a different PF eigenstate, which would lead to a change in the locations of the brightest points and the number of the points on the intensity distribution for a given input SOP. It should be emphasised again that this new paradigm directly determines the SOP from the intensity pattern without recourse to conventional instrument matrix calculations.

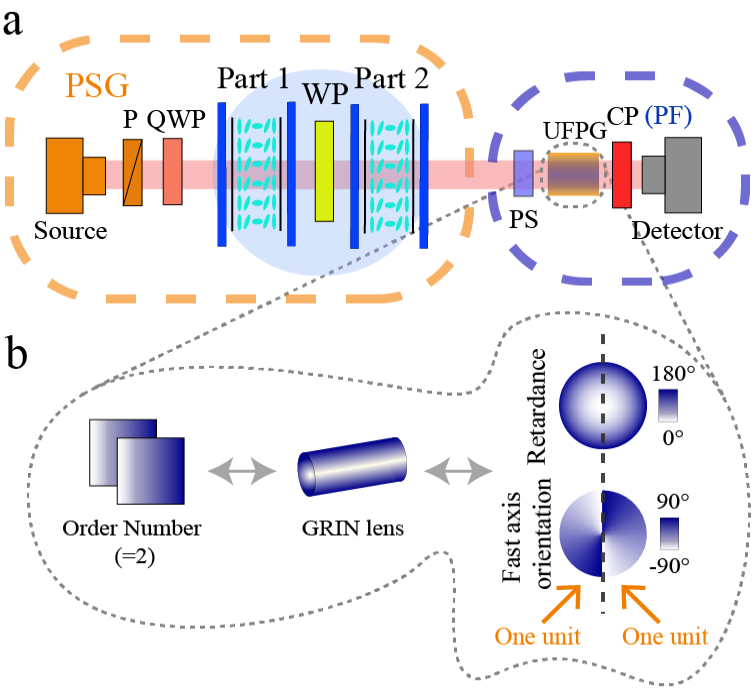

Fig. S3. Experimental setup of the UFPG polarimeter. (a) Sketch of experiment: P, CP: fixed polariser; QWP: rotating quarter waveplate; WP: quarter waveplate (note here we illustrate using a transmission geometry the both SLM parts were actually two adjacent sections of a single SLM; in practice, the WP was transited twice in a reflection geometry to act as a half wave plate[5]); UFPG: universal full Poincare generator, which is in this case a GRIN lens; PSG and PFG based polarimetry are shown in the dashed yellow/blue circles. (b) Sketch for the illustration of two UFPG units within a GRIN lens. In this manifestation of the UFPG, the retardance is arranged in a polar coordinate system[8]. Before the UFPG there exists a pinhole filtering system (PS). The black dotted line shows the line of symmetry separating two UFPG units.

Supplementary Note 3: Stokes vector measurement

In traditional approaches to Stokes vector measurement, intensities are recorded on the detector for calculation of the target SOPs [15–19]. The core equations for calculation of the Stokes vector are shown in Eq. (3-1), where S_{in} is the Stokes vector of the incident light field, consisting of 4 components (S_0, S_1, S_2, S_3). A is an $n\times4$ matrix known as the instrument matrix, which is determined by the optical properties of the measurement system [15–19]. I is the intensity information recorded by the detector.

$$I = A \cdot S_{in}, \qquad S_{in} = inv(A) \cdot I. \tag{S1}$$

In practice, we also need to consider the occurrence of errors that would affect the measurement precision and accuracy [15–23]. These two parameters will both, however, be affected via systematic errors ΔA, ΔI or random errors δA, δI. Hence three key processes – denoising, optimisation, and calibration – have been investigated and explored by numerous researchers

in the quest to enhance the precision and accuracy of a polarisation sensing system by eliminating these errors [15–23].

Amongst the three processes, optimisation is used to deal with the instrument matrix A – which determines the amplitude of the error transformation of a designed polarisation sensing system [15–18]. Several criteria such as the condition number (CN) have been put forward to evaluate such an optimisation [15–18]. In the context of polarisation sensing, this is defined as

$$\begin{aligned} CN\ (A) &= \|A\| \cdot \|inv(A)\|, \\ \max(\mu_i) &= \|A\|, \\ CN\ (A) &= \frac{\mu_{max}}{\mu_{min}}. \end{aligned} \tag{S2}$$

Here, $\|.\|$ represents the matrix L-2 norm, A is the instrument matrix during Stokes measurements; μ_i $(i = 1,2,3 \ldots)$ are the singular values of matrix A. μ_{min} and μ_{max} are the minimum and maximum values of μ_i. For a complete calculation of S_{in}, at least four intensity parameters are required, which is determined by the component number of the Stokes vector. To calculate four unknown parameters, there needs to be at least four equations, hence the instrument matrix A cannot be a single column/row [15–18]. Additionally, as each row of A is related to a different analysis state, any combination of the related vectors cannot be mutually orthogonal. Hence for polarisation sensing, the minimum CN value is $\sqrt{3}$, which is the theoretical limit with considerations of a systematic error amplification [15–18], as opposed to the minimum CN value ($\mathrm{CN} = 1$) without matrix inversion induced systematic amplification. However, in the new paradigm, we can avoid any such issues related to this matrix inversion though the use of a different computational process that is based upon the physical phenomenon of the mapping of polarisation state to an intensity pattern by the UFPG. As mentioned in main article, to deal with the additional errors such as temporal and spatial errors and noise, we introduce machine learning (ML) based 'end to end' method to deal all factors in one step. Details can be found in Supplementary Note 4.

Supplementary Note 4: Machine learning based image processing approach for polarisation estimation

For a chosen combination of UFPG and PF, there is a unique correspondence between the input SOP and the intensity pattern at the detector. Conceptually, we can explain intuitively that the positions of the brightest points encode the input SOP. However, practically, we wish to use the whole of the intensity pattern in order to estimate the input state, to ensure highest accuracy, particularly in the presence of noise. Furthermore, the exact mapping between the input state and intensity pattern is not perfectly known in advance, due to variations in system configuration as well as spatial and temporal noise. Rather than performing denoising or calibration step, we adopted a convolutional neural network (CNN) to approach the estimation holistically [24,25]. A U-Net like network has been chosen [26–28], due to its robustness to image noise, stability towards hyperparameter changing, and promising detection performance. U-Net is a non-polarisation-specific network, which has been widely used in recent ML research, and has hence been widely validated and applied either in scientific research or applications in industry [26]. It is therefore well-suited for further duplication, dissemination, and enhancement if compared with the specifically designed polarisation-oriented networks

that are currently being applied to polarisation imaging [23]. For ML based image processing, we are dealing with clear physical information – brightest point locations – to infer directly the Stokes vector. Here the different locations of the brightest points represent different SOPs. Such a well-defined physical foundation contributes to a clear target for a physical inference ML process, which goes beyond a pure black box approach.

The main process of the approach (when conducting real sensing) is depicted in Supplementary Figure 4a and 4b with a simple flow chart. The imaging processing pipeline comprises three steps: 1) the obtained input intensity images (with unknown SOP) are fed into the CNN and the areas around the brightest points are highlighted in the output heatmap (which represent probable locations of the peak); 2) refined estimates of the brightest points are localised from the heatmap using the centrosymmetric constraint, due to the symmetry of the GRIN lens; 3) the positions are converted via a look-up table (LUT) to the SOPs. The target SOP can then be obtained as the output (see Supplementary Figure 4c for the three steps). The LUT can be built offline prior to the processing procedure and is generated from simulation images with known SOPs. This enables fast and efficient SOP prediction from the brightest points during inference. Note the first step belongs to 'neural network', second and third steps belongs to 'image post-processing' (in the main article).

This network follows an encoder-decoder structure (see Supplementary Figure 4b), where the encoder down-samples the input to extract deeper features, and the decoder up-samples the feature map to integrate information from the encoder at different scales. After the last convolutional layer, a heatmap with pixel values ranging from 0.0-1.0 is generated, which can be seen as a probability map for the brightest points. As the number of possible brightest point locations increases along with the resolution increase, the input image resolution determines the upper bound of SOP prediction precision. 384×384 resolution was adopted for the input intensity image to reach a trade-off between prediction precision, training convergence, and inference speed. Other resolutions including 256×256 and 512×512 have also been tested, as shown in Table 3. 57877 pairs of simulated/experimental images were used to generate the training set. The simulated images were calculated via a GRIN lens retardance model using the approaches described in Ref [8] as the ground truth. The experimental images were acquired with known SOP input using the system shown in Supplementary Figure 4a. We generated the data by quasi-uniformly sampling on the Poincaré sphere, in order to cover as large a parameter range as possible. The locations of the brightest points were directly found from the simulated images, followed by a Gaussian distribution to indicate the local area around them to model the expected network output heatmap. Then pairs of noisy intensity images and heatmaps that have a one-to-one correspondence, were used to train the network. To increase the training set size as well as to simulate data conditions in real-world applications, data augmentation [29] including contrast and brightness changes have been carried out to improve the robustness of the network.

During the training, the network was trained with a stochastic gradient descent (SGD) optimizer [30] using gradients computed with backpropagation [31], with batch size set to 4, learning rate 0.001, momentum 0.9. A weighted L-2 loss function (Eq. (4-1)) was adopted to deal with the "imbalanced classification" problem, since the bright area only takes up a small part of the image:

$$Loss = \frac{1}{2N}\sum_{i=0}^{N-1} w_i(v_i - {v_i}^*)^2 + \lambda\|A\|^2\,, \qquad \text{(S3)}$$

where N is the total number of pixels, v_i the predicted value of the i th pixel, and ${v_i}^*$ the ground truth value of the i th pixel. w_i is the weight of the i th pixel, which was set to 50 if ${v_i}^* > 0$, otherwise 1. $\lambda = 0.0005$ is the coefficient of the regulariser A, where $A = [a_0, a_1, a_2, \dots, a_k]$ is the set of all parameters in the network. The network was trained over 5 epochs and converged in one hour on a PC (OS: Ubuntu 16.04; CPU: i7-4770; GPU: NVIDA GTX 1080 Ti).

As it is shown in the main article, we demonstrate the feasibility of using the proposed ML based method to estimate the SOP. This is a highly efficient image retrieval system: the brightest points function as image signatures that are unique to SOPs in certain ranges, so the problem of predicting SOPs is simplified and converted to a search task. There are several advantages of this approach: 1) preparation of the training set is straightforward and it is easy to cover an adequate domain; 2) finding the SOP takes only 30 ms on a normal desktop GPU, enabling real-time online SOP detection; 3) the network is invulnerable to temporal/spatial noise from the image acquisition system. Future work will centre around expanding the capability of the proposed ML based approach to enable robust and accurate real-world applications.

As the new paradigm is based on estimation the brightest point on the image, there is an intrinsic link between the image resolution (with pixel number $n \times n$) and the polarisation resolution (sensitivity S_p) of the system. This hardware parameter could be used to indicate the maximum sensitivity that the system can perform, assuming other noise sources are minimised, which can guide the training process of the CNN with respect to the best effective dataset. This sensitivity can be calculated as

$$S_p = K \cdot \sqrt[D_s]{\frac{\eta}{R}}\,, \qquad \text{(S4)}$$

where D_s is the dimension of the Stokes vector, $R = \frac{\pi \cdot n^2}{4}$ represents the effective pixel number (in the GRIN lens based UFPG, we have a circular area). K is a constant parameter. As the topological order η of the GRIN lens is 2, there would in effect be half the number of pixels to determine S_p. Note here we assume the sampling depth is sufficient and the non-linearity of the system is low. Following the above equation, we could calculate and plot the theoretical relationship between S_p and intensity image with resolution $n \times n$, if systematic and random errors are minimised (see Supplementary Figure 4d). S_p can be boosted if we use a higher camera pixel resolution.

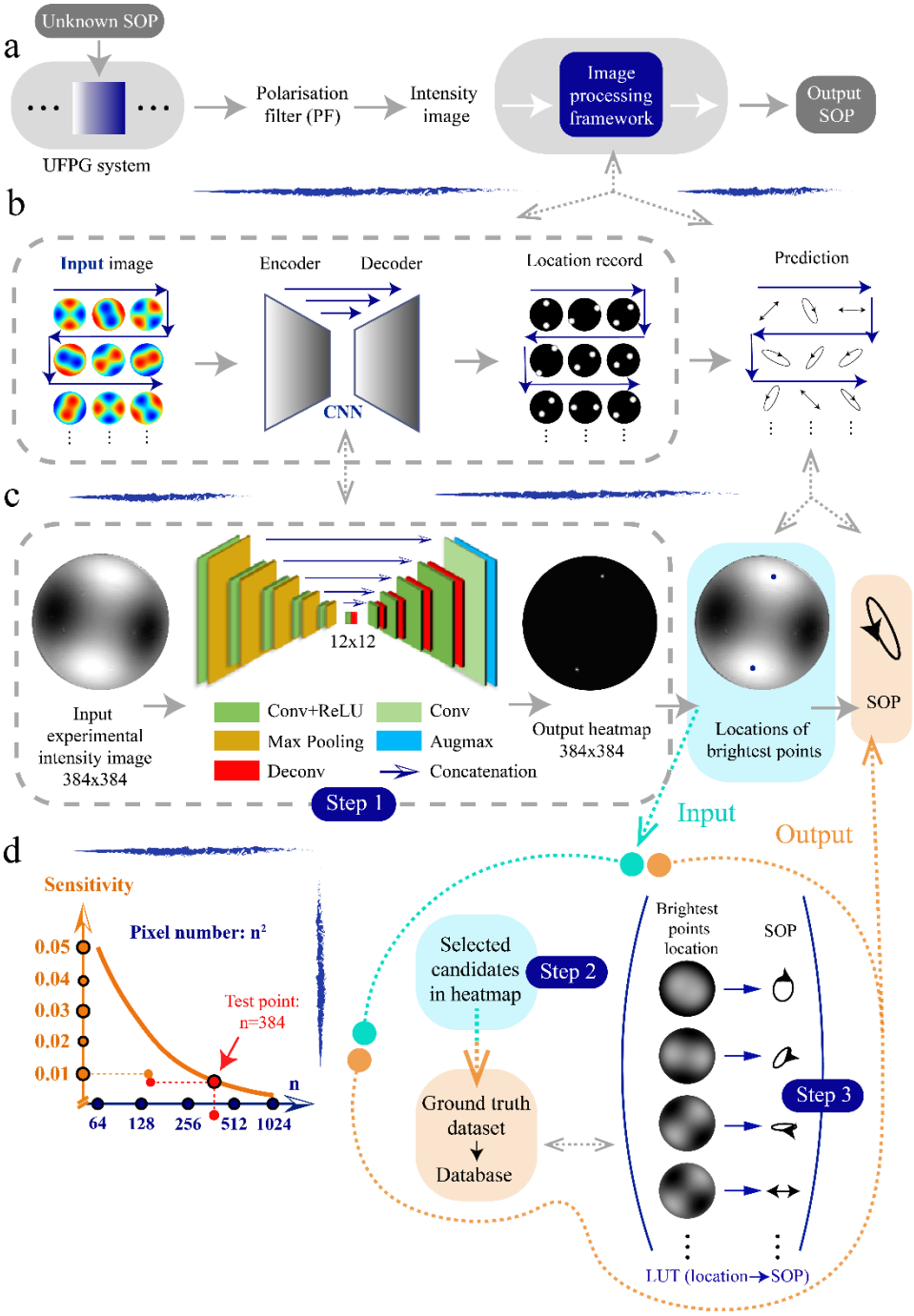

Fig. S4. Schematic of the machine learning based image processing approach to identify the unknown SOP. (a) Simple flow chart of the UFPG based paradigm enabled by ML. (b) ML implementation process: intensity distribution images are used as the network input, which then pass through the CNN (encoder plus decoder) then to find the brightest point locations and predict the unknown SOP. (c) One practical example of using the CNN based imaging process approach to acquire the unknown SOP from an experimental intensity distribution image. (d)

The theoretical relationship between the sensitivity of the UFPG and the pixel number of the acquired image (an image with pixel number $n \times n$).

Compared with traditional Stokes retrieval methods, this new imaging processing framework provides advances. The traditional approach treats denoising, optimisation and calibration as separately optimised processes. However, this traditional approach does not consider the interaction between such three processes, even though their effects on the overall performance of the polarimeter are not independent. Calibration methods can also be laborious, requiring different capabilities to conduct and complete. Those points strongly affect the performance and robustness of the polarimeter, limiting their broader usage. By remodelling these three processes into a single step, as well as considering other residual errors sources and their interactions in an 'end-to-end' manner, our new paradigm provides polarimetry with robust performance. The future challenge for our approach might be how to bridge the gap between a trained model and real-world situations. We have already partly addressed this by remodeling the problem into a simple brightest-point search, which greatly reduces the training requirements. Furthermore, we may harness simulation approaches like Monte Carlo simulation [9] which can be used to generate synthetic data to strongly aid the training process of the model.

Supplementary Note 5: Statistical analysis, validation and comparison of the measurement precision/systematic sensitivity

In this section, we examine the numerical performance of the proposed UFPG method and the established method of the point Stokes polarimeter (PSP). In particular, we compare their performance based on mean square error (MSE) and mean absolute error (MAE) [32–34]. We denote $(Y_1, ..., Y_n)$ as the vector of n observed ground truth samples, with $(\hat{Y}_1, ..., \hat{Y}_n)$ being the predicted values. MSE and MAE are computed as,

$$MSE = \frac{1}{n}\sum_{i=1}^{n} e_i^2, \;\; MAE = \frac{1}{n}\sum_{i=1}^{n} |e_i|, \tag{S5}$$

where $e_i = \hat{Y}_i - Y_i$. We first conduct an induction process via measuring a standard light field that compares to the ground truth. In terms of robustness, error plots for both methods comparing to the ground truth are demonstrated in Supplementary Figure 5 and Table 1, where a conditional mean (CM) for each point with the conditional confidence interval (CCI) related parameters are included. The definition for 90%-CI (P) throughout this work follows $P\big(Y \in \big[F_{lower}(X), F_{upper}(X)\big] | X = x\big) = 90\%$. Throughout this work, we used P to characterize the system sensitivity of the UFPG polarimeter, as it reflects the robustness of the data variation compared with the ground truth. The CM is estimated via least squares regression, as shown in Eq. (5-2), and CCI is estimated via quantile regression [35], as shown in Eq. (5-3).

$$\hat{F} = \underset{F}{\operatorname{argmin}} \frac{1}{2}\sum_{i=1}^{n} \big(Y_i - F(X_i)\big)^2, \tag{S6}$$

$$\hat{F} = \underset{F}{\operatorname{argmin}} \frac{1}{2}\sum_{i=1}^{n} \rho_\tau\big(Y_i - F(X_i)\big), \tag{S7}$$

where $\rho_\tau(\mu) = \mu(\tau - \mathcal{H}(\mu < 0))$; $\mathcal{H}(\cdot)$ is an indicator function, τ= (0.05, 0.95) corresponds to (5%, 95%) in quantile. F in both Eq. (5-2) and Eq. (5-3) is formulated via a Gradient Boosting algorithm [36], which is also a ML method for regression and classification, presented as

$$F(X) = \sum_{j=1}^{J_m} \gamma_{jm} \mathcal{H}(x \in R_{jm}). \tag{S8}$$

where J_m is the number of its leaves, γ_{jm} is the value predicted in region R_{jm}. Here the tuning parameters are set as follows: number of trees = 2, and max depth = 2 for estimating the CM for the validation case. The numerical results are summarized in Supplementary Table 1, where the upper, middle, lower lines are the 95% quantile, the CM, and the 5% quantile. As suggested by the results in Supplementary Figure 5 and Supplementary Table 1, the biases for the UFPG polarimeter are much smaller, and the minimal and maximal width of 90%-CI are both much narrower than that from the PSP in different cases, hence the UFPG polarimeter shows better sensitivity. The corresponding quantitative results also indicate that the UFPG polarimeter consistently provides more robust and precise measurement.

The sensitivity is vital for applications where weak polarisation variations play important roles. This includes, for example, optical super-resolution microscopes, where polarisation aberrations would detrimentally affect the system resolution or measurements of weak vector information [4,5,7]; as the quality of a vectorially structured light beam is destroyed. In order to compensate for such errors, having a highly sensitive sensor is crucial, since the higher the sensitivity, the more precisely the compensation system can then conduct the correction [4,5,7]. Here, we demonstrate the sensitivity of the UFPG polarimeter (see Supplementary Figure 5) by 1) measuring a standard horizontal polarised light field from the light source (generated via a polariser, Supplementary Figure 5b), which we use as a standard sensitivity evaluation process throughout this work; 2) testing the output of the horizontal middle line of a slightly tilted spatially variant half-wave plate array (see Supplementary Figure 5c); which we use as a real weak polarisation errors case in this work. Such an array can be used, for example, to generate the vectorial light field for the depletion beam of a STED microscope system, where a fine polarisation error could be catastrophic [37]. Both results are given with comparisons to the results from a PSP.

For the demonstration, we sampled 200 points along the y axis of the target shown in Supplementary Figure 5b and 5c (white dotted boxes) by using our UFPG polarimeter and the PSP. Then we calculated the bias of second component of the output Stokes vectors recorded using two approaches and plotted them in Supplementary Figure 5b (first validation case) and Supplementary Figure 5c (second real sensing case). It can be found that the performance of the UFPG polarimeter (pink data) showed higher consistency and higher stability than the traditional polarimeter (blue data) with respect to the reference line (grey). For second case, a time average to reduce uncertainty in the traditional approach is given, producing the grey curve as a reference. The measurements are all from small regions (enclosed in the white dotted boxes in Supplementary Figure 5). The MSE, MAE as well as sensitivity are calculated for former case (here we focus on relative error characterisation); further results can be found in Supplementary Table 1. We can see that the UFPG polarimeter shows sensitivity ranging from a minimum 0.79 ($/10^2$) to a maximum 0.81 ($/10^2$), validating that the UFPG method provides robust performance. The results also show successful sensing of the trends of the weak polarisation errors which cannot be detected via a traditional PSP. We used here 384×384 pixels; the performance of the UFPG with different numbers of pixels can be found in Supplementary Table 3, which validates the trend in Supplementary Figure 4d, given the fact

that the pixel number affects the sensitivity of the system. A detailed analysis of the effects of various parameters on sensitivity for higher pixel numbers would be required to determine the ultimate limits of performance. Furthermore, Supplementary Table 2 gives a quantitative comparison of the measurement precision between the UFPG polarimeter and the PSP as well as other polarimetry counterparts (see more details in Ref [38]). The UFPG polarimeter shows superior performance.

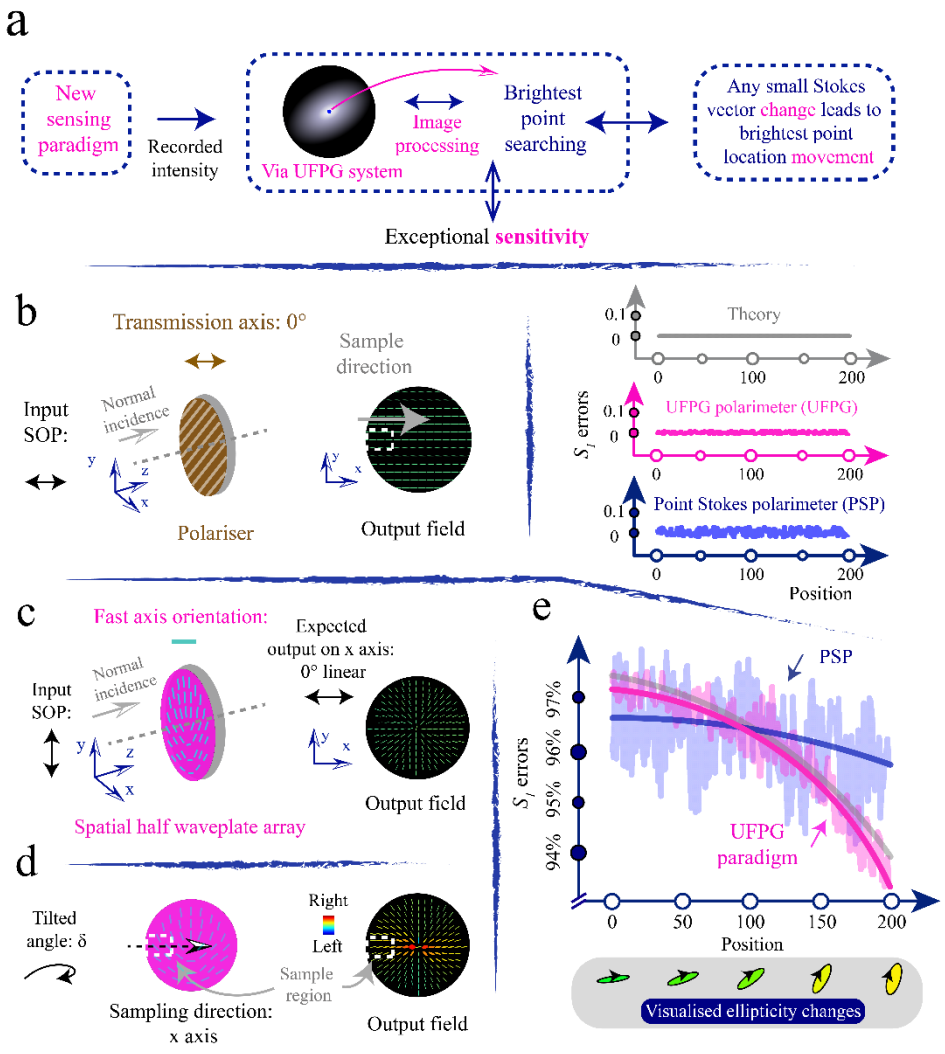

Fig. S5. Performance of UFPG and PSP. (a) A test region of the horizontal linear SOP light field generated by a high precision polariser. Theoretical results together with experiment results using UFPG and PSP methods are given. In the plots of the data in the sub-figures, the horizontal axis represents 200 points measured along the zero degree linear polarised light field (the direction within the white-dotted rectangle is shown by the "sample direction" arrow); the vertical axis shows the percentage of the vertical polarised components, which represents the error. (b) A conceptual flow chart of the mechanism of UFPG sensing paradigm. The paradigm consists of two core parts: a UFPG system, and an imaging processing method (a ML technique in this case) to determine the locations of the brightest points and thereby determination of the unknown SOP. Any small input Stokes vector change can lead to the movement of the brightest point location. Note here we focus on the relative SOP changes rather than measuring the

absolute value of the SOP. (c) Schematic of the spatially variant half wave plate array. When a vertically linear SOP is normally incident upon a plate with fixed fast axis orientations shown in the figure, the output should be a horizontally linear SOP. The output vector field is given. (d) Plate with a fine tilt (δ=3°) angle. We measured the Stokes vector output in the white dotted box region along the x axis, via the UFPG method and the traditional single-shot point Stokes polarimeter. The output vector field is demonstrated as well. (e) The obtained horizontal linear component of 200 measured points using both approaches. Pink and blue dots represent the measurement points using the UFPG polarimeter and the point Stokes polarimeter. The grey line is a reference line. The additional variations are attributed to manufacturing imperfections. More details of their initial comparisons and related analysis are provided above. A visualized ellipticity change (pseudo-magnitude) is provided as illustration of the effect of the error in the horizontal component.

Table 1 Results of UFPG and PSP methods

Method	MSE ($*10^{-5}$)	MAE ($*10^{-3}$)	Sensitivity ($*10^{-2}$) (min, max)
UFPG	0.72±0.65	2.33±1.33	0.79, 0.81
PSP	25.31±32.54	13.04±9.14	4.65, 4.88

Table 2 Quantitative comparision data

	DOP error (max)	DOP error (mean)	Sensitivity ($*10^{-2}$) (max)
Product 1 (UFPG)	±0.84%	±0.02%	0.81
Product 2 (PSP)	N/A	±1%	4.88

Error of Stokes parameters (mean)	
Product 1 (UFPG)	<1% (S_1), <1% (S_2), <1% (S_3)
Product 3 (X-shaped aperture array)$^{\Delta}$	>7.3% (S_1), >7.2% (S_2), >5.2% (S_3)
Product 4 (Integrated plasmonic)$^{\Delta}$	~45% (S_1, S_2, S_3)
Product 5 (Dielectric metasurface)$^{\Delta}$	7.5~15% (S_1, S_2, S_3)
• • •	• • •

Table 3 Comparasion of different pixel numbers

	256x256	384x384	512x512
DOP error ($*10^{-2}$) (max)	±2.35%	±0.84%	±0.69%
DOP error ($*10^{-2}$) (mean)	±0.31%	±0.02%	±0.05%
Sensitivity ($*10^{-2}$) (max)	1.76	0.81	0.63

Table. S1, S2 and S3. *Degrees of polarisation (DOP) error is what the product claimed; $^{\Delta}$Refer to Ref [38] for data of other comparisons.

Supplementary Note 6: Low intensity and depolarised scenarios.

To demonstrate the wider feasibility of the polarimeter, we consider measurement using our UFPG polarimeter in low intensity and depolarised scenarios, as these commonly occur in real world sensing. Supplementary Figure 6a shows retrieval errors from 200 random SOPs (methods in the main article) with respect to different normalized intensity levels. Supplementary Figure 6b shows the errors of different SOPs associated with different DOPs (the DOP level here is normalized and controlled following the method explained in Ref [39]). The S_1, S_2, S_3 errors are represented via mean errors and standard deviation. Based on the results from Supplementary Figure 6, we can conclude that the UFPG polarimeter has robust performance both at low intensity and with depolarisation, which are comparable to its performance with normalized intensity level and DOP both equal to 1 (main article). This robust performance is a consequence of the ML based imaging processing procedure. The CNN model provides many parameters to fit to a wide range of real scenarios. This is helped by the data augmentation [40] that included models of the various scenarios for brightness changes (including low intensity) and different contrast (related to low DOP) (see Supplementary Note 4).

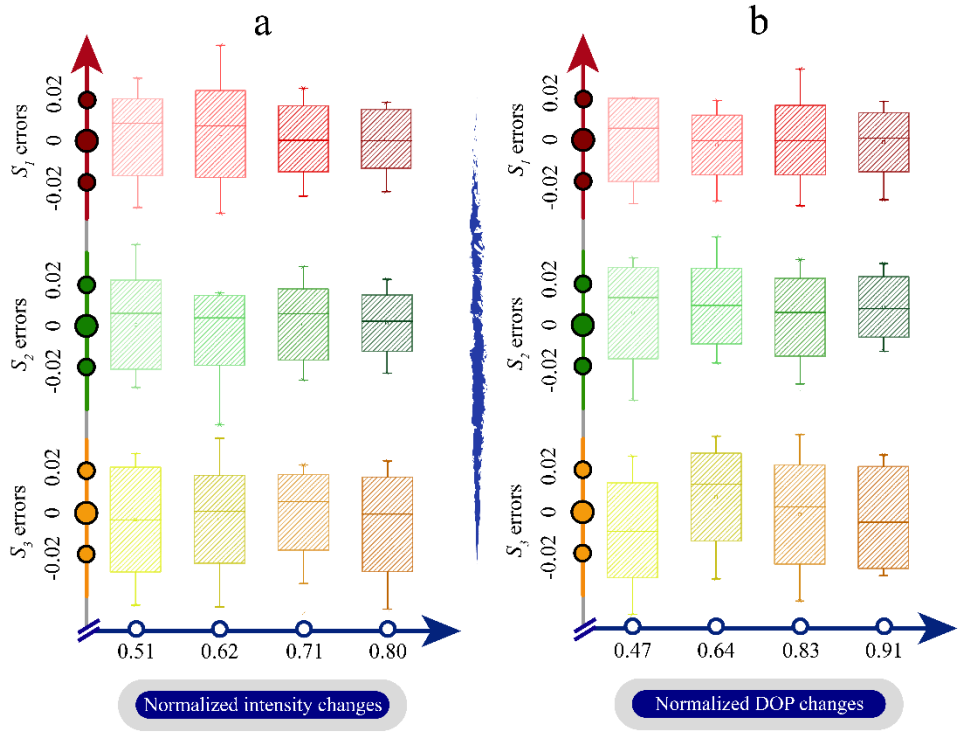

Fig. S6. Experimental results for low intensity and depolarised occasions. (a) The errors of Stokes components S_1, S_2, S_3 with respect to different intensity conditions: 0.51, 0.62, 0.71, 0.80 (normalized intensity). (b) The errors of Stokes components S_1, S_2, S_3 with respect to different DOP conditions: 0.47, 0.64, 0.83, 0.91 (normalized DOP).

The depolarisation parameter is also an important vectorial parameter in numerous techniques and applications [41,42] and can be additionally extracted via the obtained image. Supplementary Figure 7 shows the typical intensity distributions of several random SOPs selected from the Poincaré sphere under different levels of depolarisation (0% for 7a, 20% for 7b, 60% for 7c, respectively). The SOP of their polarised parts, however, remain the same. The level of contrast within the image directly represents the level of depolarisation of the target beam. Its DOP – can be theoretically calculated via the normalised intensity value of the brightest and darkest points on the intensity image according to a simple calculation, $(I_{max} - I_{min})/(I_{max} + I_{min})$, This is another unique feature of the UFPG polarimeter that enables the calculation of depolarisation in a simple way.

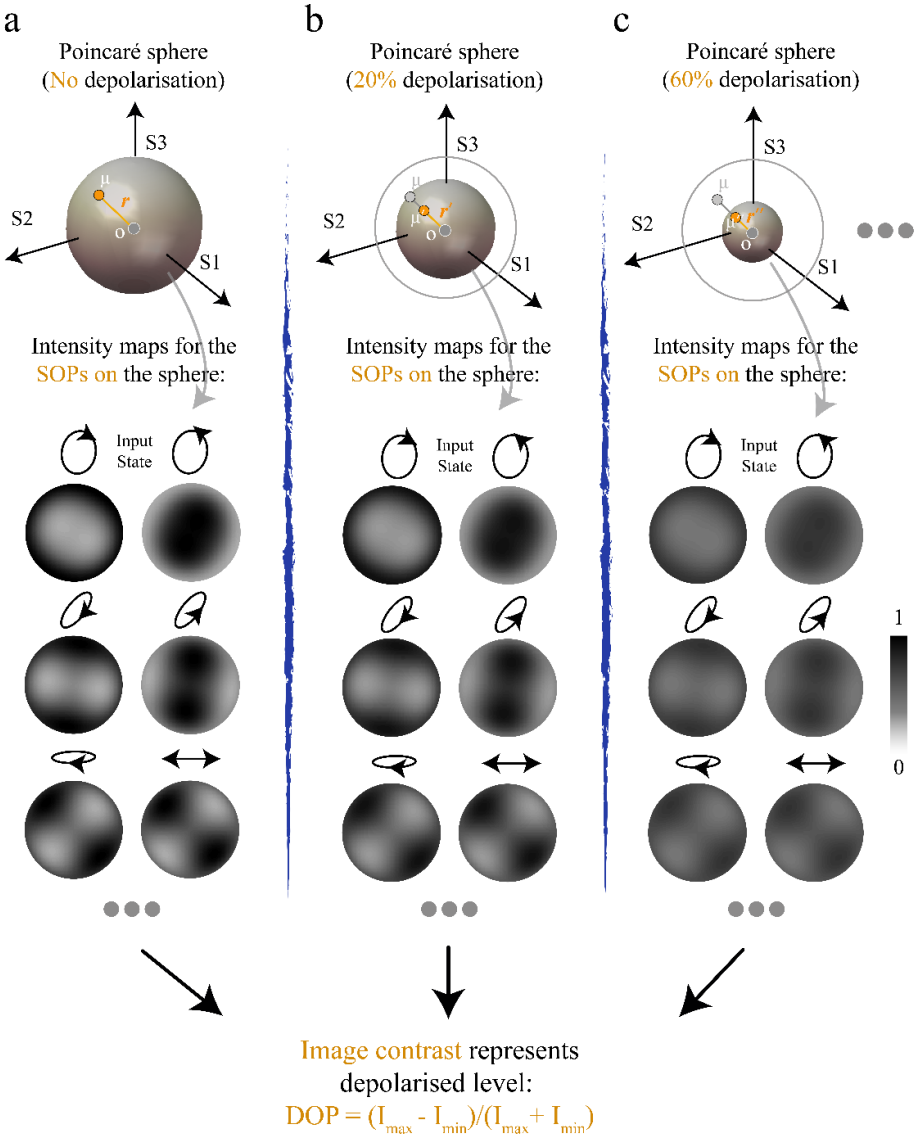

Fig. S7. The relationship between depolarised beams and their corresponding intensity images. (a) Before depolarisation – point μ on the surface of the Poincare sphere, with radius r from the centre o. Recorded intensity images of the UFPG polarimeter for several randomly chosen SOPs on the surface. (b) and (c) After depolarisation – points μ' and μ'' *in* the Poincare sphere with radius r' and r'' from the centre o (the fully polarised part representing the same SOP as μ). The final obtained intensity images under such depolarisation levels with previously chosen incident SOPs are also given below. The DOP of the incident beam can be calculated from the contrast level of their corresponding intensity images.

Supplementary Note 7: Single-shot full Stokes multi-point sensing using the UFPG paradigm

While we can implement single-shot sensing through our UFPG-based paradigm, it is also possible to expand the concept into multi-point measurement as well as into the imaging

domain. An $n \times n$ UFPG array consisting of $n \times n$ UFPGs is designed (see Supplementary Figure 8a, 10×10 UFPGs are used for demonstration). Next it is integrated into the image plane of an imaging polarimeter (see Supplementary Figure 8b) – on the surface of the camera sensor combined with a polariser. The assembly can then directly perform multi-point polarisation sensing by determining the brightest point on each super-pixel (10×10) via the same procedure as for point sensing. Future engineering solutions should be developed to conquer certain problems if we wanted to perform imaging, such as finding a balance between the number of photo-sensors and the image resolution – there is always a trade-off when considering the focus plane division technique [9] and this issue can be notable in the case of our UFPG.

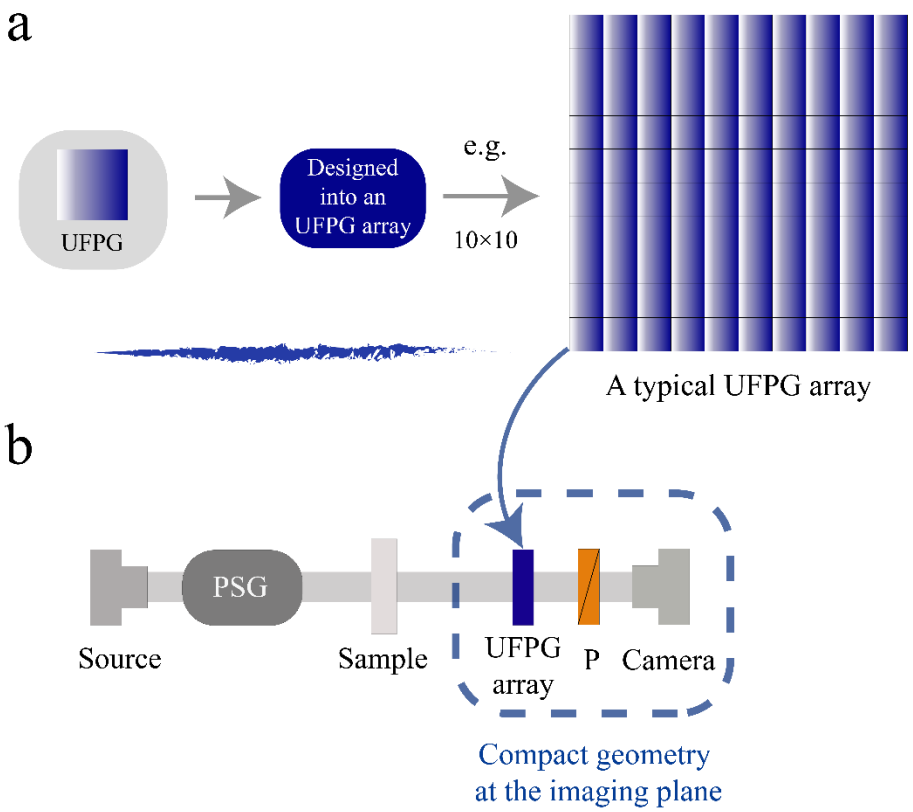

Fig. S8. Extended application for the UFPG system. (a) A typical UFPG array with super-pixel number 10×10. Note that here one UFPG acts as one pixel. (b) A schematic for a new compact snapshot Stokes multi-point sensing/imaging device. It possesses the same advantages as a single UFPG, such as direct Stokes vector detection without any additional matrix calculations. PSG: polarisation state generator; P: polariser.

References

1. A. M. Beckley, M. A. Alonso, and T. G. Brown, "Full Poincaré beams," Optics Express **18**, 10777–10785 (2010).
2. Q. Zhan, W. Han, and W. Cheng, "Flattop focusing with full Poincaré beams under low numerical aperture illumination," Optics Letters **36**, 1605–1607 (2011).

3. G. Piquero, L. Monroy, M. Santarsiero, M. Alonzo, and J. C. G. D. Sande, "Synthesis of full Poincaré beams by means of uniaxial crystals," Journal of Optics **20**, 065602 (2018).
4. Q. Hu, Y. Dai, C. He, and M. J. Booth, "Arbitrary vectorial state conversion using liquid crystal spatial light modulators," Optics Communications **459**, 125028 (2020).
5. Y. Dai, C. He, J. Wang, R. Turcotte, L. Fish, M. Wincott, Q. I. Hu, and M. J. Booth, "Active compensation of extrinsic polarization errors using adaptive optics," Optics Express **27**, (2019).
6. C. He, M. Booth, and M. Booth, "Extraordinary Beam Modulation with Ordinary GRIN Lenses," Optics and Photonics News, Vol. 31, Issue 12, p. 47 (2020).
7. C. He, J. Antonello, and M. J. Booth, "Vectorial adaptive optics," (2021).
8. C. He, J. Chang, Q. Hu, J. Wang, J. Antonello, H. He, S. Liu, J. Lin, B. Dai, D. S. Elson, P. Xi, H. Ma, and M. J. Booth, "Complex vectorial optics through gradient index lens cascades," Nature Communications **10**, 1–8 (2019).
9. C. He, H. He, J. Chang, B. Chen, H. Ma, and M. J. Booth, "Polarisation optics for biomedical and clinical applications: a review," Light: Science & Applications **10**, 1–20 (2021).
10. R. Ramkhalawon, A. M. Beckley, and T. G. Brown, "Star test polarimetry using stress-engineered optical elements," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XIX* (SPIE, 2012), Vol. 8227, p. 82270Q.
11. J. Chang, H. He, C. He, Y. Wang, N. Zeng, R. Liao, and H. Ma, "Optimization of GRIN lens Stokes polarimeter," Applied Optics **54**, 7424 (2015).
12. J. Chang, N. Zeng, H. He, Y. He, and H. Ma, "Single-shot spatially modulated Stokes polarimeter based on a GRIN lens," Optics Letters **39**, 2656 (2014).
13. S. Sivankutty, E. R. Andresen, G. Bouwmans, T. G. Brown, M. A. Alonso, and H. Rigneault, "Single-shot polarimetry imaging of multicore fiber," Optics Letters **41**, 2105 (2016).
14. B. G. Zimmerman, T. G. Brown, P. Yang, H. Wei, G. W. Kattawar, Y. X. Hu, D. M. Winker, C. A. Hostetler, and B. A. Baum, "Star test image-sampling polarimeter," Optics Express **24**, 23154–23161 (2016).
15. D. S. Sabatke, E. L. Dereniak, G. S. Phipps, M. R. Descour, S. A. Kemme, and W. C. Sweatt, "Optimization of retardance for a complete Stokes polarimeter," Optics Letters **25**, 802–804 (2000).
16. J. S. Tyo, "Design of optimal polarimeters: maximization of signal-to-noise ratio and minimization of systematic error," Applied Optics **41**, 619–630 (2002).
17. C. He, H. He, H. Ma, J. Chang, N. Zeng, R. Liao, and Y. Wang, "Linear polarization optimized Stokes polarimeter based on four-quadrant detector," Applied Optics **54**, 4458–4463 (2015).
18. J. Chang, H. He, Y. Wang, Y. Huang, X. Li, C. He, R. Liao, N. Zeng, S. Liu, and H. Ma, "Division of focal plane polarimeter-based 3 × 4 Mueller matrix microscope: a potential tool for quick diagnosis of human carcinoma tissues," Journal of Biomedical Optics **21**, 056002 (2016).
19. J. Qi and D. S. Elson, "Mueller polarimetric imaging for surgical and diagnostic applications: a review," Journal of Biophotonics **10**, 950–982 (2017).
20. A. Skumanich, B. W. Lites, V. M. Pillet, and P. Seagraves, "The Calibration of the Advanced Stokes Polarimeter," The Astrophysical Journal Supplement Series **110**, 357–380 (1997).
21. H. Gu, X. Chen, H. Jiang, al -, W. Li, and C. Zhang, *Optimal Broadband Mueller Matrix Ellipsometer Using Multi-Waveplates with Flexibly Oriented Axes Depolarization Artifacts in Dual Rotating-Compensator Mueller Matrix Ellipsometry* (2000).
22. A. Abubakar, X. Zhao, M. Takruri, E. Bastaki, and A. Bermak, "A Hybrid Denoising Algorithm of BM3D and KSVD for Gaussian Noise in DoFP Polarization Images," IEEE Access **8**, 57451–57459 (2020).
23. X. Li, H. Li, Y. Lin, J. Guo, J. Yang, H. Yue, K. Li, C. Li, Z. Cheng, H. Hu, and T. Liu, "Learning-based denoising for polarimetric images," Optics Express **28**, 16309–16321 (2020).
24. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM **60**, 84–90 (2017).

25. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," 770–778 (2016).
26. T. von Eicken, A. Basu, V. Buch, and W. Vogels, "U-Net: A User-Level Network Interface for Parallel and Distributed Computing," ACM SIGOPS Operating Systems Review 29, no. 5, 40-53 (1995).
27. N. Navab, J. Hornegger, W. Wells, and A. Frangi, *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015* (2015).
28. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **9351**, 234–241 (2015).
29. C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data **6**, (2019).
30. L. B.-O., "Online algorithms and stochastic approximations," *Online learning and neural networks* (1998).
31. D. Touretzky and G. Hinton, *Connectionist Models Summer School* (1988).
32. R. Tibshiranit, "Regression Shrinkage and Selection Via the Lasso," Journal of the Royal Statistical Society: Series B (Methodological) **58**, 267–288 (1996).
33. T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning. Springer series in statistics," (2001).
34. B. Dai, J. Wang, X. Shen, and A. Qu, "Smooth neighborhood recommender systems," Journal of Machine Learning Research **20**, 1–24 (2019).
35. R. Koenker and K. F. Hallock, "Quantile Regression," Journal of Economic Perspectives **15**, 143–156 (2001).
36. J. H. Friedman, "Greedy function approximation: a gradient boosting machine," Annals of statistics, 1189-1232 (2001).
37. Z. Rong, C. Kuang, Y. Fang, G. Zhao, Y. Xu, and X. Liu, "Super-resolution microscopy based on fluorescence emission difference of cylindrical vector beams," Optics Communications **354**, 71–78 (2015).
38. A. Basiri, X. Chen, J. Bai, P. Amrollahi, J. Carpenter, Z. Holman, C. Wang, and Y. Yao, "Nature-inspired chiral metasurfaces for circular polarization detection and full-Stokes polarimetric measurements," Light: Science & Applications **8**, 1–11 (2019).
39. H. Shao, Y. He, W. Li, and H. Ma, "Polarization-degree imaging contrast in turbid media: a quantitative study," Applied Optics **45**, 4491–4496 (2006).
40. C. Steger, M. Ulrich, and C. Wiedemann, *Machine Vision Algorithms and Applications* (2018).
41. N. Ghosh, "Tissue polarimetry: concepts, challenges, applications, and outlook," Journal of Biomedical Optics **16**, 110801 (2011).
42. T. Novikova, I. Meglinski, J. C. Ramella-Roman, and V. V. Tuchin, "Special Section Guest Editorial: Polarized Light for Biomedical Applications," Journal of Biomedical Optics **21**, 071001 (2016).

Given an input light S with $\chi \in (-\frac{\pi}{2}, \frac{\pi}{2})$, and $\psi \in [-\pi, \pi]$:

$$S = \begin{bmatrix} 1 \\ \cos(\chi)\cos(\psi) \\ \cos(\chi)\sin(\psi) \\ \sin(\chi) \end{bmatrix}, \tag{1}$$

and the $\mathbf{M}_{\theta,\delta}$ is the Mueller matrix of the half of a GRINLens with $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and $\delta \in (0, \pi)$:

$$\begin{aligned} \mathbf{M}_{\theta,\delta} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -\sin 2\theta & 0 \\ 0 & \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\delta & \sin\delta \\ 0 & 0 & -\sin\delta & \cos\delta \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos\delta & \sin 2\theta \cos 2\theta(1-\cos\delta) & -\sin 2\theta \sin\delta \\ 0 & \sin 2\theta \cos 2\theta(1-\cos\delta) & \sin^2 2\theta + \cos^2 2\theta \cos\delta & \cos 2\theta \sin\delta \\ 0 & \sin 2\theta \sin\delta & -\cos 2\theta \sin\delta & \cos\delta \end{pmatrix}. \end{aligned}$$

Then, the outcome of interest is computed as

$$O = \mathbf{P}\mathbf{M}_{\theta,\delta}S = \begin{bmatrix} \frac{1}{2}\big(1 + \cos(\delta)\sin(\chi) + \cos(\chi)\sin(\delta)\sin(2\theta - \psi)\big) \\ 0 \\ 0 \\ \frac{1}{2}\big(1 + \cos(\delta)\sin(\chi) + \cos(\chi)\sin(\delta)\sin(2\theta - \psi)\big) \end{bmatrix}, \tag{2}$$

where

$$\mathbf{P} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix},$$

and the intensity of the outcome is $\mathcal{I}(\theta, \delta) := O_1 = \frac{1}{2}\big(1 + \cos(\delta)\sin(\chi) + \cos(\chi)\sin(\delta)\sin(2\theta - \psi)\big)$. The following theorem indicates the correspondence between the input light and the optimal GRINLens parameters.

Theorem 1. *Define the intensity function as:*

$$\mathcal{I}(\theta, \delta) = \frac{1}{2}\big(1 + \cos(\delta)\sin(\chi) + \cos(\chi)\sin(\delta)\sin(2\theta - \psi)\big),$$

where $\chi \in (-\frac{\pi}{2}, \frac{\pi}{2})$ and $\psi \in [-\pi, \pi]$ are the parameters of an input light. Let

$$\widehat{\theta}(\chi, \psi), \widehat{\delta}(\chi, \psi) = \underset{\theta \in [-\pi/2, \pi/2]; \delta \in (0,\pi)}{\operatorname{argmax}} \mathcal{I}(\theta, \delta).$$

Then, $(\widehat{\theta}(\chi,\psi),\widehat{\delta}(\chi,\psi))$ *is an injective function (one-to-one function) of* (χ,ψ)*:*

$$\widehat{\theta}(\chi,\psi)=\widehat{\theta}(\psi)=\psi/2+\frac{1}{4}\pi,\ \textit{if}\ \psi<\pi/2,\quad \psi/2-\frac{3}{4}\pi,\ \textit{if}\ \psi\geq\pi/2,$$
$$\widehat{\delta}(\chi,\psi)=\widehat{\delta}(\chi)=\pi/2-\chi.$$

Proof of Theorem 1. To proceed, we first compute the gradients of $\mathcal{I}(\theta,\delta)$ with respect to θ and δ, respectively.

$$\begin{aligned}\frac{\partial\mathcal{I}(\theta,\delta)}{\partial\theta}&=\cos(\chi)\cos(2\theta-\psi)\sin(\delta)\\ \frac{\partial\mathcal{I}(\theta,\delta)}{\partial\delta}&=\frac{1}{2}\big(\cos(\delta)\cos(\chi)\sin(2\theta-\psi)-\sin(\delta)\sin(\chi)\big).\end{aligned}\tag{3}$$

According to the extreme value theorem, the maximizer of $\mathcal{I}(\theta,\delta)$ are obtained, either on the local maximizer within the domain, or in the boundary. We consider those two cases separately.

Boundary. There are four boundaries in the (closure of) domain: $\mathcal{B}_1=\{\theta\in[-\pi/2,\pi/2],\delta=0\}$, $\mathcal{B}_2=\{\theta\in[-\pi/2,\pi/2],\delta=\pi\}$, $\mathcal{B}_3=\{\theta=-\pi/2,\delta\in[0,\pi]\}$, and $\mathcal{B}_4=\{\theta=\pi/2,\delta\in[0,\pi]\}$.

The evaluation of $\mathcal{I}(\theta,\delta)$ is computed as follows.

- $\mathcal{B}_1:\mathcal{I}(\theta,\delta)=\frac{1}{2}\big(1+\sin(\chi)\big)$,
- $\mathcal{B}_2:\mathcal{I}(\theta,\delta)=\frac{1}{2}\big(1-\sin(\chi)\big)$,
- $\mathcal{B}_3:\mathcal{I}(\theta,\delta)=\frac{1}{2}\big(1+\cos(\delta)\sin(\chi)+\cos(\chi)\sin(\delta)\sin(\psi)\big)$,
- $\mathcal{B}_4:\mathcal{I}(\theta,\delta)=\frac{1}{2}\big(1+\cos(\delta)\sin(\chi)+\cos(\chi)\sin(\delta)\sin(\psi)\big)$,

Therefore, the maximizer in the boundary is obtained at $(\theta=-\pi/2,\delta=\delta_b)$ or $(\theta=\pi/2,\delta=\delta_b)$, since the evaluation of $\mathcal{B}_3,\mathcal{B}_4$ incorporates the cases in $\mathcal{B}_1,\mathcal{B}_2$. If $\delta_b\neq 0,\pi$, then $\frac{\partial\mathcal{I}(\theta,\delta)}{\partial\theta}|_{\theta=-\pi/2}=-\cos(\chi)\cos(\psi)\sin(\delta_b)=\frac{\partial\mathcal{I}(\theta,\delta)}{\partial\theta}|_{\theta=\pi/2}$, which yields that the maximizer must be obtained at a local maximizer.

Local maximizer. Next, we consider the local maximizer. Letting the equations in (3) equal to zeros, we have

$$\cos(2\theta-\psi)=0\implies 2\theta-\psi=-\frac{3}{2}\pi,-\frac{1}{2}\pi,\frac{1}{2}\pi,\frac{3}{2}\pi$$
$$\cos(\delta)\cos(\chi)\sin(2\theta-\psi)-\sin(\delta)\sin(\chi)=0\implies\cos(\delta\pm\chi)=0.$$

This yields the following cases:

(i) if $\psi\geq\pi/2$, there are two local maximizers/minimizers $\theta=\psi/2-\frac{3}{4}\pi,\delta=\pi/2-\chi$, or $\theta=\psi/2-\pi/4,\delta=\chi+\pi/2$. In this case, we have:

θ	$-\pi/2$	$\psi/2-3\pi/4$	0	$\psi/2-\pi/4$	$\pi/2$
$\frac{\partial\mathcal{I}(\theta,\delta)}{\partial\theta}$	$-\cos^2(\chi)\cos(\psi)>0$	0	$\cos^2(\chi)\cos(\psi)<0$	0	$-\cos^2(\chi)\cos(\psi)>0$

Therefore, the local maximizer is $(\theta = \psi/2 - \frac{3}{4}\pi, \delta = \pi/2 - \chi)$, and its evaluation is $\mathcal{I}(\theta, \delta) = 1$.

(ii) if $-\pi/2 \leq \psi < \pi/2$, there are two local maximizers/minimizers, $\theta = \psi/2 - \pi/4, \delta = \chi + \pi/2$ or $\theta = \psi/2 + \pi/4, \delta = \pi/2 - \chi$. In this case, we have

θ	$-\pi/2$	$\psi/2 - \pi/4$	0	$\psi/2 + \pi/4$	$\pi/2$
$\frac{\partial \mathcal{I}(\theta,\delta)}{\partial \theta}$	$-\cos^2(\chi)\cos(\psi) < 0$	0	$\cos^2(\chi)\cos(\psi) > 0$	0	$-\cos^2(\chi)\cos(\psi) < 0$

Therefore, the local maximizer is $(\theta = \psi/2 + \frac{1}{4}\pi, \delta = \pi/2 - \chi)$, and its evaluation is $\mathcal{I}(\theta, \delta) = 1$.

(iii) if $-\pi \leq \psi < -\pi/2$, there are two local maximizers/minimizers, $\theta = \psi/2 + \pi/4, \delta = \pi/2 - \chi$ or $\theta = \psi/2 + 3/4\pi, \delta = \pi/2 + \chi$. In this case, we have

θ	$-\pi/2$	$\psi/2 + \pi/4$	0	$\psi/2 + 3/4\pi$	$\pi/2$
$\frac{\partial \mathcal{I}(\theta,\delta)}{\partial \theta}$	$-\cos^2(\chi)\cos(\psi) > 0$	0	$\cos^2(\chi)\cos(\psi) < 0$	0	$-\cos^2(\chi)\cos(\psi) > 0$

Therefore, the local maximizer is $(\theta = \psi/2 + \frac{1}{4}\pi, \delta = \pi/2 - \chi)$, and its evaluation is $\mathcal{I}(\theta, \delta) = 1$.

Since the evaluation in local maximizers are larger than the one in the boundary. Taken together, we have the following conclusions:

CASE 1. When $\psi \geq \pi/2$, the local maximizer is $(\theta = \psi/2 - \frac{3}{4}\pi, \delta = \pi/2 - \chi)$, and its evaluation is $\mathcal{I}(\theta, \delta) = 1$, and it is larger than $\mathcal{I}(\pm\pi/2, 0) = \mathcal{I}(\pm\pi/2, \pi) = \frac{1}{2}(1 + \sin(\chi))$. Consequently, $(\theta = \psi/2 - \frac{3}{4}\pi, \delta = \pi/2 - \chi)$ is the global maximizer.

CASE 2. When $\psi < \pi/2$, with the same argument, $(\theta = \psi/2 + \frac{1}{4}\pi, \delta = \pi/2 - \chi)$ is the global maximizer.

Thus, both $\widehat{\theta}(\chi, \psi)$ and $\widehat{\delta}(\chi, \psi)$ are one-to-one functions. This completes the proof. □