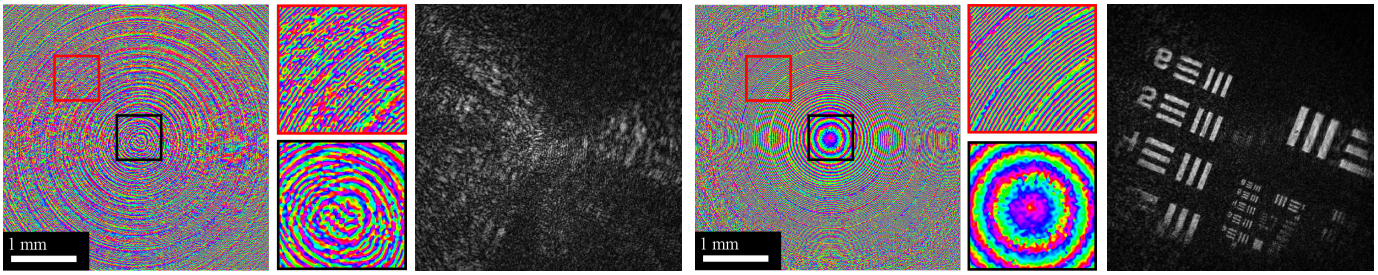
Reference Wave Design for Wavefront Sensing

Wei-Yu Chen, Anat Levin, Matthew O'Toole, and Aswin C. Sankaranarayanan

Abstract—One of the classical results in wavefront sensing is phase-shifting point diffraction interferometry (PS-PDI), where the phase of a wavefront is measured by interfering it with a planar reference created from the incident wave itself. The limiting drawback of this approach is that the planar reference, often created by passing light through a narrow pinhole, is dim and noise sensitive. We address this limitation with a novel approach called ReWave that uses a non-planar reference that is designed to be brighter. The reference wave is designed in a specific way that would still allow for analytic phase recovery, exploiting ideas of sparse phase retrieval algorithms. ReWave requires only four image intensity measurements and is significantly more robust to noise compared to PS-PDI. We validate the robustness and applicability of our approach using a suite of simulated and real results.

Index Terms—Wavefront sensing, Phase retrieval, Interferometry



(a) PS-PDI [1] (b) ReWave (ours)

Fig. 1: **PS-PDI and ReWave results.** Wavefront sensing techniques can capture the phase mask pattern of a Fresnel lens, and refocus the wave to form an image of a target (e.g., a resolution chart) beyond it. (a) PS-PDI recovers a noisy phase pattern of the Fresnel lens, and cannot be used to form an image of the resolution chart. (b) Our reference wave design technique, ReWave, clearly reproduces the quadratic phase pattern of the Fresnel lens. By correcting for the phase of the Fresnel lens, we can also successfully synthesize a sharp image of the resolution chart.

1 INTRODUCTION

Image sensors only measure the intensity of the incident illumination. However, the field incident on them has an associated phase which is desired in many applications. Wavefront sensing systems try to address this challenge, attempting to measure both the amplitude and phase of a light wave. While there are numerous techniques for wavefront sensing based on phase retrieval [3], [6], [7] and interferometry [8], [9], [10], it remains a challenge to build a fast and high-resolution wavefront sensor that is robust to noise [6].

Among the numerous wavefront sensing approaches, phase-shifting point diffraction interferometry (PS-PDI) [1], [11] is perhaps one of the most elegant. PS-PDI creates a planar self-reference wave from the incident wave that is being sensed. Interference patterns observed between the wave and the reference, under multiple phase shifts, provide analytic extraction of the phase of the incident wave.

PS-PDI has many desirable properties: it creates a self-reference wave that bypasses the need for an external reference wave, requires only three intensity measurements, provides a simple analytical solution to the wavefront sensing problem, and can measure wavefronts at the spatial resolution of the image sensor. However, its key limitation is the need for a planar reference wave, produced by collimating light that has been passed through a narrow pinhole mask. The self-reference wave is *extremely* dim, and the resulting technique is therefore highly sensitive to noise.

We propose *ReWave*, an adaptive wavefront sensing technique that overcomes the limitations associated with PS-PDI, by designing a mask with higher light throughput to produce a brighter reference wave. ReWave uses a second image sensor to observe the light intensity distribution in the Fourier plane and identify candidate pinholes where the light intensity is brightest. As the resulting reference would no longer be planar, we need to recover an unknown reference wave as part of our algorithm. For that, the reference mask is strategically chosen such that the resulting reference wavefront can be uniquely and analytically recovered using a key result from sparse phase retrieval [12]. This result relies on the design of a so-called "collision-free mask", where any two pairs of pinholes have a unique displacement between them. When put together, ReWave creates a significantly brighter reference wave for phase-shifting interferometry, while also retaining the many advantages of PS-PDI. These advantages include a reference wave that

- *Chen and Sankaranarayanan are with the ECE Department, Carnegie Mellon University, Pittsburgh.*
- *Levin is with the EE Department, Technion, Israel.*
- *O'Toole is with the Robotics Institute, Carnegie Mellon University. Corresponding Author E-mail: saswin@andrew.cmu.edu*

TABLE 1: **Different phase acquisition approaches:** A summary of the advantages and drawbacks of different approaches.

	Shack-Hartmann [2]	Gerchberg-Saxton [3]	Multi-planes [4]	WISH [5]	PS-PDI [1]	**ReWave (ours)**
limiting resolution	lens array	camera	camera	camera	camera	camera
analytical solution	yes	no	no	no	yes	yes
number of measurement	1	2	∼20	>8	3	3 +1
reference brightness					low	high

is created from the incident signal, and a small number of measurements that are required to recover the wave through a simple analytical expression.

Contributions. This paper proposes *ReWave*, an analytical wavefront sensing method that advances PS-PDI with an adaptively-constructed reference wave to enable robustness to noise. Our main contributions are as follows:

- *Adaptive reference wave construction.* Our primary contribution is the adaptive design of a non-planar reference wave that is significantly brighter than the traditional planar reference that is typically used in PS-PDI, which makes the resulting technique robust to noise. Further, following work in sparse phase retrieval, the reference wave is created so as to permit an analytical reconstruction.
- *Analytical solution.* In the context of phase retrieval, our method provides an analytical solution with only four measurements. In contrast, state of the art techniques based on phase modulation [5] require more measurements, and are computationally intensive requiring iterative optimization while lacking convergence guarantees.
- *Experimental prototype.* We have implemented our algorithm on a lab prototype and successfully demonstrated various tasks like refocusing and imaging through scattering media.

In Fig. 1, we recover the phase profile of a Fresnel lens using PS-PDI and ReWave. The reconstruction from ReWave shows a clear quadratic phase profile, while that of PS-PDI is riddled with errors. As a consequence, when we computationally correct the ghosting introduced by the Fresnel lens, ReWave produces a higher quality result. Note that the code and data associated with the paper can be found at https://github.com/Image-Science-Lab-cmu/ReWave_ICCP2021.

2 RELATED WORK

Prior work in wavefront sensing can be categorized in three groups: classical, phase retrieval, and interferometry-based. We summarize them here. A list of representative works and their relative merits is provided in Table 1.

2.1 Classical approaches for wavefront sensing

One of the basic wavefront sensor designs relies on the Shack-Hartmann sensor [2]. This sensor uses a lenslet array to capture a wavefront in a single snapshot. Assuming the wavefront is locally planar over each lenslet, the location of the brightest spot beneath each lenslet corresponds to the tilt of the local wavefront. Despite its simplicity, the main disadvantage of the Shack-Hartmann sensor is its poor spatial and angular resolutions.

There have been many attempts to improve on the standard Shack-Hartmann design. Pyramid wavefront sensor [13] uses a pyramidal-shaped prism to encode the gradient of the incident wave. However, adapting the working range of the approach relies on moving components [14]. A curvature wavefront sensor [15] captures the Laplacian of the phase, recovered by using the Transport of Intensity Equation (TIE) [16]. The method cannot reconstruct smooth waves where the Laplacian is small. A coded wavefront sensor [17] achieves higher resolutions by replacing the lenslet array with a coded mask, and measures the wavefront by observing local shifts in a diffraction pattern. However, this method assumes the wave has uniform amplitude and smooth phase.

2.2 Phase retrieval

2.2.1 *Phase retrieval using iterative methods*

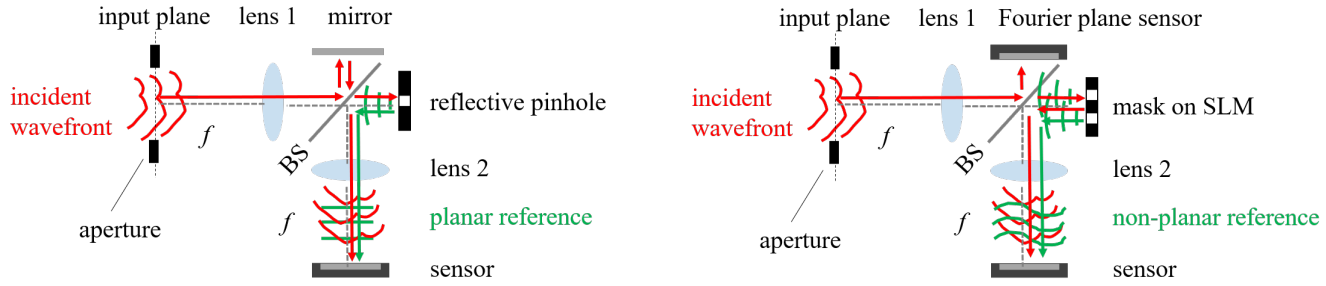
Phase estimation is challenging as cameras can only measure the intensity of the wave, which is nonlinearly related to the complex wave of interest. One class of techniques acquires multiple intensity images of the incident wave, under a set of transformations. The unknown wavefront is estimated using an iterative scheme, where in each iteration, the intensity of the transformed wave is simply replaced by the measured one. This iterative scheme has a fixed point solution that produces the measured intensities under the approriate set of transformations.

The classic technique in this class is the Gerchberg-Saxton (GS) algorithm [3], where the intensity of the wave and its Fourier transform are measured. Although simple to implement, convergence guarantees are limited as the solution space is riddled with local minima [18].

Nonlinear optimization has been successfully applied with other type of measurements. For example, multi-plane propagation [4] measures the intensity of the incident waves at multiple depth planes. Other techniques provide more diverse measurements using an amplitude [19] or phase based [5] masks implemented using a spatial light modulator (SLM). Compared with the original GS algorithm, these intensity observations capture more information on the phase and usually converge to a better solution, but still have little theoretical guarantees. It is also noted empirically that they require at least eight input images for robust performance.

2.2.2 *Convex optimization in phase retrieval*

While phase retrieval itself is a non-convex optimization problem, it is possible to relax it as a convex problem. For example, PhaseLift [20] and PhaseCut [21] map the problem



(a) PDI (b) ReWave

Fig. 2: **PDI vs. ReWave setup.** A $4f$ interferometric setup used to implement classical PDI and our ReWave approach. (a) Point diffraction interferometry — the incident wave is split by a beamsplitter (BS) into two arms. One arm reflects the incident wave as is using a mirror, while the other includes a pinhole, which when collimated, forms a planar reference field which corresponds to the DC component of the target incident wave. (b) ReWave — As with PDI, the incident wave is split into two arms using a BS. On one arm, we replace the pinhole with a phase only spatial light modulator which implements a mask $\mathbf{M}$ for the reference field; as explained in Sec. 4.1, this arm reflects both the target and reference fields. On the other arm, we include an image sensor which measures the intensity on the Fourier plane of the incident field; this intensity image is used to select pixels on the phase SLM, or equivalently, design the mask $\mathbf{M}$ that is used to build the non-planar reference wave.

to a semi-definite programming or a relaxed MaxCut problem. Jaganathan et al. [22] design specific masks to ensure a unique solution. However, all of these techniques rely on lifting, mapping a length-n vector into a matrix with n^2 elements. This lifting operation is impractical when dealing with high resolution wavefronts, where n is large.

2.2.3 *Sparse phase retrieval*

The phase retrieval problem can be solved more efficiently when the underlying wave is sparse [12], [23], [24], [25]. The work of Ranieri et al. [12] is central to this paper, as it provides the conditions under which a unique analytic solution is feasible. We will discuss this approach in detail in Sec. 4.4.

2.3 Interferometry-based wavefront sensing

2.3.1 *Phase-shifting interferometry*

Phase-shifting interferometry (PSI) [9] is a technique for recovering the wavefront by interfering it with a reference wave under three or more phase shifts. In a typical implementation, the illumination source is split into two arms. One of them interacts with the target sample, while the other arm serves as a known reference. The two-path design make the set-up bulky and sensitive to vibrations. Moreover, in applications like adaptive optics [26] in the context of astronomy, we do not have access to the coherent source and, consequently, the reference.

2.3.2 *Common-path interferometry*

The proposed work falls under the category of common-path interferometer, where the reference and target waves travel along the same optical path. For example, lateral shearing interferometry interferes the wave with a shifted copy of itself, resulting in the wavefront gradient. Different implementation includes quadri-wave lateral shearing interferometry [27] or optical differentiation wavefront sensor [28]. However, these methods generally require the wavefront to be smooth for successful recovery.

2.3.3 *Point diffraction interferometry*

Point diffraction interferometry [11] creates a reference plane wave by passing the incident wave through a pinhole. Phase-shifting point diffraction interferometry (PS-PDI) [1], [29] captures 3 shots while changing the phase of this reference wave. Our work builds upon PS-PDI, but allows for a brighter reference and thus has a higher SNR [30]. While most PDI is a common-path setup, it can also be implemented using a two-path design [31], which we use in Sec. 3 to introduce the ReWave setup.

2.4 Combination of phase retrieval and interferometry

Interferometry-based wavefront sensing can be used in conjunction with phase retrieval techniques. For example, in Fourier holography [32], a single interference pattern (a hologram) between the far-field diffraction of an object and a known reference wave is measured. Recently, Barmherzig et al. [33] applied phase retrieval technique to estimate phase from the recorded hologram.

3 Background on Point Diffraction Interferometry

We start with a review of the point diffraction interferometry (PDI) technique [11] for measuring the phase of a wavefront, which forms the basis for the proposed approach.

In PDI, the incident wave, whose phase we aim to measure, is split into two arms, as illustrated in Fig. 2(a). One arm transfers a copy of the wave while the other arm creates a self-reference wave. The self-reference wave is created by placing a pinhole in a Fourier plane, letting into the system only a plane wave corresponding to the DC component of the incident wave. There are many ways to realize this idea via an optical setup [1], [29], and we illustrate a setup in Fig. 2(a) which is analogous to what we will use in our approach below. Specifically, Fig. 2(a) assumes a beamsplitter and two mirrors in its two arms, reflecting the target

and reference signals. The reflected signals are interfered via the same beam-splitter and captured by a camera. This camera is focused such that it images a copy of the input plane before the beam splitter.

In phase-shifting point diffraction interferometry (PS-PDI) [1], [29], the reference wave is subject to $K \geq 3$ equally-spaced phase-shifts $\phi_k = 2\pi(k-1)/K$, leading to intensity images of the form:

$$\begin{aligned} I_k(x) &= |\mathbf{u}(x) + e^{j\phi_k}\mathbf{r}(x)|^2 \\ &= |\mathbf{u}(x)|^2 + |\mathbf{r}(x)|^2 + 2Re\left(\mathbf{u}(x)\mathbf{r}(x)^* e^{-j\phi_k}\right), \end{aligned} \tag{1}$$

where $\mathbf{u}$ is the incident wave, $\mathbf{r}$ is the reference wave, and x denotes a 2D position on the image plane sensor. From these measurements, we can extract the interference signal as in [8]

$$\mathbf{u}(x)\mathbf{r}(x)^* = \frac{1}{K}\sum_k I_k(x)e^{j\phi_k}. \tag{2}$$

As few as $K = 3$ phase shifts ϕ_k are sufficient to recover the wavefront. In the setup above, the reference wave is basically a constant plane wave equivalent to the DC of the target signal. Thus, Eq. (2) provides the desired incident wave up to a global scale factor.

While this algorithm is extremely simple to derive, in practice it suffers from a serious disadvantage: the reference wave is created from a single pinhole and hence, it is extremely dim and noise sensitive. A well-known result in interferometry [30], [34] states that the performance of phase-shifting interferometry improves with increasing brightness of the reference wave, as long as we avoid saturation at the sensor. Unfortunately, there is no straightforward way to create a reference wave which is both fronto-parallel and of high-intensity. For example, if we naively increase the size of the pinhole, the reference wave would no longer be planar [11].

4 ReWave approach

Our proposed ReWave approach is designed to increase the brightness of the reference wave. We do so by placing a mask $\mathbf{M}$ that lets light through multiple pinholes, rather than a single one. This brighter reference wave also has a more complex form, requiring us to derive proper estimation techniques.

4.1 Setup

In the following, we will use $\mathcal{F}_\mathbf{u}, \mathcal{F}_\mathbf{r}$ to denote the Fourier transforms of the target and reference waves, respectively, and use ω to denote a 2D position on the Fourier plane of the system.

Our setup is illustrated in Fig. 2(b). We replace the pinhole in the PDI setup of Fig. 2(a) with a spatial light modulator (SLM) that will allow us to programmably implement any desired mask. One approach is to use an amplitude SLM in the reference arm of the interferometer and a clear mirror in the second arm, as in the PDI setup of Fig. 2(a). We instead use a common-path interferometric scheme, where we instead place a phase SLM on the reference arm; since the phase SLM does not block light, this generates a reference wave $\mathbf{r}$ which is interfered with an adjusted target $\mathbf{u} - \mathbf{r}$ on the *same arm* of the interferometer in Fig. 2(b), simplifying both the alignment and reducing vibration issues. This also frees up the second arm of the beamsplitter as we already have a copy of the incident wave, sans the reference, to interfere with the reference.

Given a binary mask $\mathbf{M}$, we define the reference and target signals as

$$\mathcal{F}_\mathbf{r}(\omega) = \mathcal{F}_\mathbf{u}(\omega)\mathbf{M}(\omega), \quad \mathcal{F}_{\mathbf{u}-\mathbf{r}}(\omega) = \mathcal{F}_\mathbf{u}(\omega)(1 - \mathbf{M}(\omega)). \tag{3}$$

To phase shift the reference, the SLM displays a phase pattern equal to $e^{j\phi}\mathbf{M} + (1 - \mathbf{M})$, i.e., a phase of ϕ at the pixels in the mask $\mathbf{M}$ and zero elsewhere. Thus we generate a phase shift at the mask pixels, and reflect light without any change at the pixels belonging to $\mathcal{F}_{\mathbf{u}-\mathbf{r}}$. At the sensor, we measure the intensity resulting from the interference of the reference $\mathbf{r}$ with the wave $\mathbf{u} - \mathbf{r}$.

The common-arm configuration in Fig. 2(b) is important in another way. In the second arm of the beam-splitter, we place a sensor measuring the intensity of the incoming wave at the Fourier plane. We use this information to adapt the mask layout to the signal being captured.

4.2 Approach

With the setup as described above, we can capture $K = 3$ phase-shifted interference patterns of the form

$$I_k(x) = |\mathbf{u}(x) - \mathbf{r}(x) + e^{j\phi_k}\mathbf{r}(x)|^2, \tag{4}$$

with phase shifts that are equally-spaced, i.e., $\phi_k = 2\pi(k-1)/K$. By interpreting $(\mathbf{u}(x) - \mathbf{r}(x))$ as the "incident wave", this expression becomes identical to Eq. (1) and hence we can measure $(\mathbf{u}(x) - \mathbf{r}(x))\mathbf{r}^*(x)$ using Eq. (2). However, unlike PS-PDI, $\mathbf{r}(x)$ is not constant and unknown to us; hence, we need to estimate it as well.

Our approach to recovering the reference wave has two steps: first, estimate the amplitude $|\mathbf{r}(x)|$, and second, estimate its phase $\angle\mathbf{r}(x)$. The latter is a phase retrieval problem in its own right, and it is here where the design of the reference wave, and the selection of the mask $\mathbf{M}$, comes in play; we invoke prior work in sparse phase retrieval to obtain an analytical solution. Once we have both magnitude and phase, we can form the complex-valued reference $\mathbf{r}(x) = |\mathbf{r}(x)|e^{j\angle\mathbf{r}(x)}$, and subsequently estimate the incident wave from the measurement of $(\mathbf{u}(x) - \mathbf{r}(x))\mathbf{r}^*(x)$.

4.3 Estimating the amplitude of the reference wave

As mentioned above, we first want to estimate $|\mathbf{r}(x)|^2$. From Eq. (4) we can compute the DC component and the amplitude of the captured sinusoid, which provide two constraints on the reference and target waves:

$$\alpha(x) = \frac{1}{K}\sum_k I_k(x) = |\mathbf{u}(x) - \mathbf{r}(x)|^2 + |\mathbf{r}(x)|^2 \tag{5}$$

$$\beta(x) = \frac{1}{K}\sum_k I_k(x)e^{j\phi_k} = (\mathbf{u}(x) - \mathbf{r}(x))\,\mathbf{r}(x)^* \tag{6}$$

Given $\alpha(x), \beta(x)$, we can eliminate $\mathbf{u}(x) - \mathbf{r}(x)$ to get the following quadratic expression on $|\mathbf{r}(x)|^2$:

$$|\mathbf{r}(x)|^4 - \alpha(x)|\mathbf{r}(x)|^2 + |\beta(x)|^2 = 0 \tag{7}$$

Solving for the roots of a quadratic equation leads to two candidate solutions for $|\mathbf{r}(x)|^2$:

$$\frac{1}{2}\left(\alpha(x) \pm \sqrt{\alpha(x)^2 - 4|\beta(x)|^2}\right) \tag{8}$$

In a similar way, we can get a quadratic expression for $|\mathbf{u}(x) - \mathbf{r}(x)|^2$. In fact, it can be shown that one of the roots in Eq. (8) corresponds to $|\mathbf{r}(x)|^2$, and the other to $|\mathbf{u}(x) - \mathbf{r}(x)|^2$. While Eq. (8) provides the values of $|\mathbf{r}(x)|^2$ and $|\mathbf{u}(x) - \mathbf{r}(x)|^2$, it does not tell them apart. However, as the magnitude of the reference wave is usually lower than that of the actual signal, we estimate $|\mathbf{r}(x)|^2$ as the minimum of the two values in Eq. (8). A similar assumption has been used successfully for transmission matrix estimation [35].

4.4 Phase estimation from collision-free masks

The above algorithm estimates $|\mathbf{r}(x)|^2$ from the phase shifted interference patterns. However, to actually estimate the field of interest $\mathbf{u}$ we need to know the phase of $\mathbf{r}$, not only its magnitude. To get from $|\mathbf{r}|$ to $\mathbf{r}$, we essentially need to solve the original problem of this paper: phase retrieval. However, if the mask $\mathbf{M}$ corresponds to a sparse subset of holes which are carefully selected to satisfy a certain collision-free arrangement that we define below, the phase estimation simplifies to singular value decomposition.

Recalling that the auto-correlation and power spectral density of a signal are Fourier pairs, we can recover the auto-correlation of $\mathcal{F}_\mathbf{r}$ by computing the Fourier transform on $|\mathbf{r}|^2$. From Eq. (3), we observe that $\mathcal{F}_\mathbf{r}$ is a sparse signal since it is constructed by choosing a sparse set of phase SLM pixels; further, the support of this signal, i.e., the location of its non-zero entries, is known. We now show that, if the mask is designed carefully, $\mathcal{F}_\mathbf{r}$ can be estimated from the SVD of a matrix extracted from its auto-correlation.

The auto-correlation of $\mathcal{F}_\mathbf{r}$ can be expressed as a sum of products of pairwise elements

$$\mathcal{F}^{-1}_{|\mathbf{r}|^2}(\omega) = (\mathcal{F}_\mathbf{r} \star \mathcal{F}_\mathbf{r})(\omega) = \sum_{(\omega_p,\omega_q)|\omega_p-\omega_q=\omega} \mathcal{F}_\mathbf{r}(\omega_p)\mathcal{F}_\mathbf{r}(\omega_q)^* \tag{9}$$

where the sum is taken over all pairs of positions such that $\omega_p - \omega_q = \omega$. Given Eq. (9), if we design a mask such that the set $\{(\omega_p, \omega_q) \mid \omega_p - \omega_q = \omega\}$ consists *at most* of one member, for any non-zero value of ω, we can associate the value of $\mathcal{F}^{-1}_{|\mathbf{r}|^2}(\omega)$ to the product $\mathcal{F}_\mathbf{r}(\omega_p)\mathcal{F}_\mathbf{r}(\omega_q)^*$ for the appropriate ω_p and ω_q.

Definition 1. A mask $\mathbf{M}$ is *collision free* if, for every $\omega \neq 0$, there is at most one pair (ω_p, ω_q) with $p \neq q$, satisfying $\omega_p - \omega_q = \omega$.

For a collision-free mask, we can extract $\mathcal{F}_\mathbf{r}(\omega)$ from $\mathcal{F}_\mathbf{r}(\omega_p)\mathcal{F}_\mathbf{r}(\omega_q)^*$ using a simple SVD. Let the support of $\mathcal{F}_\mathbf{r}$ be $\{\widetilde{\omega}_p, 1 \leq p \leq P\}$; since we choose the SLM pixels on the reference, this support set is known. We now create a $P \times P$ matrix $Z = [z_{p,q}]$, whose (p,q)-th entry is $\mathcal{F}_\mathbf{r}(\widetilde{\omega}_p)\mathcal{F}_\mathbf{r}(\widetilde{\omega}_q)^*$. Using Eq. (9), we can extract this value from the auto-correlation function $\mathcal{F}^{-1}_{|\mathbf{r}|^2}(\omega)$ at $\omega = \widetilde{\omega}_p - \widetilde{\omega}_q$. The diagonal elements of Z corresponding to $p = q$ are not included in the auto-correlation, but since in our setup we have a Fourier plane sensor observing the Fourier plane of the field we get all measurements of the form $|\mathcal{F}_\mathbf{r}(\omega)|^2$ directly. With this, we have a rank-1 matrix Z, whose left/right singular vectors provide $\mathcal{F}_\mathbf{r}$, the Fourier transform of the reference wave, up to an unknown scalar. We can now estimate the reference wave $\mathbf{r}$ simply by computing the inverse Fourier transform of $\mathcal{F}_\mathbf{r}$.

Once we solved for the reference wave, we can extract the target field $\mathbf{u}$ using

$$\widehat{\mathbf{u}}(x) = \frac{\beta(x) + |\mathbf{r}(x)|^2}{\mathbf{r}(x)^*} \tag{10}$$

with $\beta(x)$ defined in Eq. (6). The resulting algorithm is summarized in Algorithm 1.

Connection to Ranieri et al. [12]. The construction of the collision-free mask is based on [12]. However, our implementation here offers three main simplifications. First, in [12], the sparse support is unknown and needs to be estimated, while in our approach we control the mask $\mathbf{M}$ and its precise support as in an endoscopy application [24]. Second, in [12], the diagonal components of the product matrix, $|\mathcal{F}_\mathbf{r}(\omega_p)|^2$, are unknown and they calculate them using matrix inversion step; we however directly measure $|\mathcal{F}_\mathbf{r}(\omega_p)|^2$. Finally, [12] only assumes the support is sparse and collision free, while we design our mask and guarantee that this condition holds.

Algorithm 1: ReWave pipeline

1. Capture $|\mathcal{F}_\mathbf{u}|^2$ using the Fourier plane sensor.
2. Design a collison-free mask $\mathbf{M}$ using Algorithm 2.
3. Place $K \geq 3$ phase masks of the form $1 - \mathbf{M} + e^{j\phi_k}\mathbf{M}$ in the Fourier plane and use the image plane sensor to capture $I_k = |\mathbf{u}(x) - \mathbf{r}(x) + e^{j\phi_k}\mathbf{r}(x)|^2$.
4. Evaluate $\alpha(x), \beta(x)$ using Eqs. (5) and (6).
5. Set
$$|\mathbf{r}(x)|^2 = \min\left\{\frac{1}{2}\left(\alpha(x) \pm \sqrt{\alpha(x)^2 - 4|\beta(x)|^2}\right)\right\}$$
6. Compute $\mathcal{F}^{-1}_{|\mathbf{r}|^2}$ and estimate $\mathcal{F}_\mathbf{r}$ using SVD.
7. Use the estimated reference wave to extract $\mathbf{u}$ from the interference signal
$$\widehat{\mathbf{u}}(x) = \frac{\beta(x) + |\mathbf{r}(x)|^2}{\mathbf{r}(x)^*}$$

Algorithm 2: Collision-free mask design

Input: Brightness map $\mathbf{y} = |\mathcal{F}_\mathbf{u}|^2$
Output: Support set $\mathbf{S}_p$
$\mathbf{S}_0 = \emptyset$;
for *p = 1 to N* **do**
  $\omega_p = \arg\max_\omega \mathbf{y}(\omega)$
  forall $\omega_m, \omega_n \in \mathbf{S}_p$ **do**
    $\mathbf{y}((\omega_p + \omega_n)/2) = 0$
    $\mathbf{y}(\omega_p + \omega_m - \omega_n) = 0$
  end
  $\mathbf{S}_p = \mathbf{S}_{p-1} \cup \{\omega_p\}$;
end

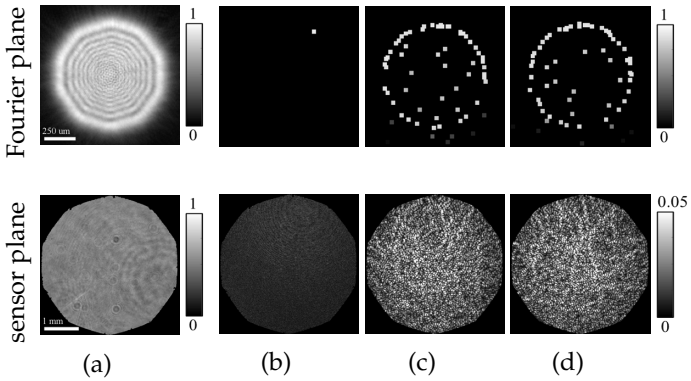


Fig. 3: **Visualizing the reference wave.** We show the amplitude of the waves in the Fourier (top row) and image (lower row) planes, demonstrating the brightness increase achieved by our ReWave approach. (a) Incident wave. (b) The single hole mask of the PDI setup leads to a very dim planar reference wave. (c-d) The reference waves formed by using different collision-free masks.

4.5 Collision-free mask design

To design the mask $\mathbf{M}$, we want to let as much light through the pinholes, while still respecting the collision free conditions. As our setup includes a Fourier plane sensor capturing $\mathcal{F}_{\mathbf{u}}$, we can adapt the mask holes to the content of the signal, and position them at points where the Fourier intensity of the target wave $\mathcal{F}_{\mathbf{u}}$ is high. We use a greedy selection strategy summarized in Algorithm 2. Specifically, we sequentially add the brightest point to the mask such that it does not create a collision with any of the previously selected points. Given an $N \times N$ image, the mask $\mathbf{M}$ can include $O(N)$ holes thereby providing significant improvements on the brightness of the reference wave which, as we demonstrate later, is quite effective in producing high quality results. With this greedy approach, we can also easily create multiple possible masks by picking randomly among top choices; this provides additional diversity in our measurements, which is desirable when we have the luxury of capturing more than three images. We illustrate these benefits in Fig. 3(a), where we observe the intensity of an incident wave both in the sensor and the Fourier plane. If we implement PDI by selecting only one point in the Fourier plane, the reference wave would still be very dim even if we already pick the brightest point, as illustrated in Fig. 3(b). On the other hand, the ReWave technique selects several points to create a much brighter reference.

4.6 Range and resolution

We now discuss the factors that limit the spatial and angular range and resolution of the wavefront sensor.

4.6.1 Spatial range and angular resolution

The spatial range of the wavefront sensor Ω_x cannot be larger than the width of the image sensor. However, to use the entire sensor area, we need to ensure that the wavefront is accurately propagated from the input aperture to the image sensor by the 4f system without any significant aberations. This, in turn, requires that the lens diameter is larger than the sensor width and the pupil plane aperture, which in turn is determined by the SLM area.

In our implementation, a second limiting factor emerges from the need to separate between the signal of interest and its ghosting copy due to the unmodulated DC term reflecting off the SLM. As we discuss later in Sec. 6, we reduce this ghosting by placing an aperture in the input plane and using a phase ramp on the SLM to isolate the wave of interest from the DC component. The size of the aperture is smaller than the sensor area and hence, determines the spatial range Ω_x of the wavefront sensor. This restriction can be avoided by placing a lithographic made phase ramp in the Fourier plane rather than implementing it with an SLM whose pitch is limited. In that case, we can set Ω_x to be equivalent to the support of the imaging sensor.

The spatial range is also a key determining factor in the angular resolution Δ_ω of the wavefront sensor, since it determines the spread of the airy disk in the Fourier plane.

$$\Delta_\omega = \frac{\lambda f}{2\Omega_x} \tag{11}$$

This airy disk is invariably larger than the SLM pitch, which would be the natural limiting factor for larger spatial ranges.

Note that, given the sensor's limited spatial range, we will provides discussion in Sec. 4.7.2 on how to choose the size of pinhole features in the Fourier plane mask.

4.6.2 Spatial resolution and angular range

The spatial resolution Δ_x of the wavefront sensor is determined by the smaller of the sensor pixel pitch, as well as size of the pupil plane aperture, which in our case is the SLM width. In practice, a signal at spatial resolution Δ_x will span in the Fourier plane a range $[-\Omega_\omega, \Omega_\omega]$ with

$$\Omega_\omega = \frac{\lambda f}{2\Delta_x}, \tag{12}$$

and in our implementation in Sec. 6, the SLM width is roughly equivalent to this range. The angular range of the sensor is determined by the pupil plane aperture, which, as mentioned earlier, is the SLM width.

4.7 Additional improvements

The discussion in the previous subsections outlines the basic ReWave approach. Below, we describe two additional components which can further improve its performance.

4.7.1 Multiple reference waves

One drawback of the above approach is that our reference wave $\mathbf{r}(x)$ will have spatially-varying intensity. In pixels where the reference has a very low magnitude, our estimate of the incident wave will be poor, as the estimate of $\mathbf{u}(x)$ in Eq. (10) involves a division by the reference wave. To overcome this, we can capture phase-shifted interference measurements with multiple collision free masks $\mathbf{M}_t$. These masks can produce different reference waves $\mathbf{r}_t$, varying the location of low intensity $|\mathbf{r}_t(x)|$ values. The incident waves $\widehat{\mathbf{u}}_t$ obtained from the masks $\mathbf{M}_t$ (or equivalently, the reference waves $\mathbf{r}_t$) can be robustly fused using:

$$\bar{\mathbf{u}}(x) = \frac{\sum_t \widehat{\mathbf{u}}_t(x)|\mathbf{r}_t(x)|^2}{\sum_t |\mathbf{r}_t(x)|^2} = \frac{\sum_t (\beta_t(x) + |\mathbf{r}_t(x)|^2)\mathbf{r}_t(x)}{\sum_t |\mathbf{r}_t(x)|^2} \tag{13}$$

We provide a detailed justification for this weighted averaging of the individual estimate in our App. A. It is also worth noting that all masks $\mathbf{M}_t$ share one phase pattern, corresponding to a phase shift $\phi_1 = 0$. As a result, we only need to capture two images per each additional mask $\mathbf{M}_t$, and hence, we need $K = 1 + 2L$ measurements when seeking to sense with L reference waves, plus one more measurement by the Fourier plane sensor.

4.7.2 *Overcoming sensor range limits*

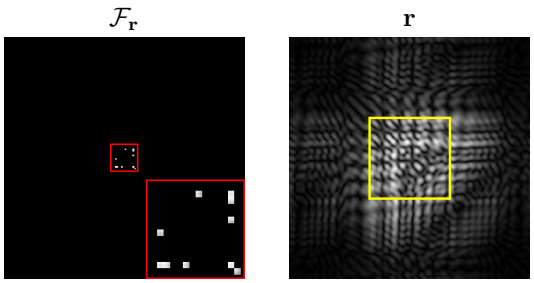
An important parameter for our system is the size of the individual pinholes composing our masks. Ideally, we want to make them as wide as possible to allow more light, but increasing their size beyond the Nyquist limit introduces dependencies in the mask which violate the collision free requirement. Hence, we set the size of the pinholes forming the reference wave to be equal to Δ_ω, as given in (11).

However, the reference wave $r(x)$ emerging from such pinholes can exceed the sensor range. For example, given a piecewise constant pinhole of size Δ_ω in the Fourier plane, one gets in the image plane a sinc whose main lobe matches the spatial range of the sensor Ω_x. The secondary lobes of this sinc lie outside the range Ω_x, and are not captured by the sensor.

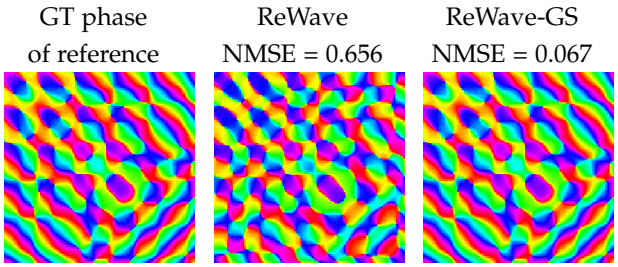
This affects the reference wave estimation technique described in Section 4.4; specifically, when we compute the auto-correlation of $\mathcal{F}_\mathbf{r}$ by taking the Fourier transform of $|\mathbf{r}(x)|^2$, an implicit assumption is that $\mathbf{r}(x) = 0$ outside the image sensor. As a consequence, the estimate of the reference wave can be off by a small amount (see Fig. 4.)

ReWave-GS. One way to fix this is to use an iterative approach to fine tune the estimate of the reference wave, relaxing the assumption that the reference wave is zero outside of the sensor range. This approach is motivated by the popular Gerchberg–Saxton phase retrieval algorithm [3]. In this approach, we start with the estimate of $|\mathbf{r}|^2$ using Eq. (8) and subsequently, we obtain an initial expression for $\mathcal{F}_\mathbf{r}$ using SVD as described in Sec. 4.4. This field is transformed back into the image plane to get $\mathbf{r}$, but crucially over a spatial range that is wider than the actual sensor size. For the part of $\mathbf{r}$ that falls inside the sensor range, we set the magnitude to match the measurements, while the values outside the sensor range are left unchanged. We then recompute $\mathcal{F}_\mathbf{r}$ as before with this new estimate of $|\mathbf{r}|^2$, and enforce the sparsity and amplitude constraint on the Fourier plane, and iterate between the Fourier and image planes until convergence. In our evaluation below, we refer to this scheme as *ReWave-GS*. Being careful about these out of boundary values improves the estimate of the reference wave, and consequently, the incidence wavefront.

Fig. 4 illustrates this issue. Here, $\mathcal{F}_\mathbf{r}$ is selected to include a few pinholes whose width is set as the Fourier resolution in Eq. (11). Transforming $\mathcal{F}_\mathbf{r}$ back to the image plane, we see its support somewhat exceeds the binary support of $\mathbf{u}$. In the lower part of the figure we see the ground truth phase of $\mathbf{r}$, compared with a reconstruction from the original ReWave approach (which assumes that the reference is zero-valued outside of the sensor's spatial range) and the iterative ReWave-GS. This fine tuning improves the accuracy of the reconstructed phase.



(a) Reference wave example



(b) Reconstructed phase of reference wave

Fig. 4: **Reconstructing the reference wave with a limited sensor.** (a) Visualize the reference mask $\mathcal{F}_\mathbf{r}$ in the Fourier plane and the resulting reference wave $\mathbf{r}$ in the image plane. The support in the image plane is wider than the sensor support (marked with a yellow frame). (b) Ignoring this wider support and assuming the reference wave is zero outside the sensor leads to erroneous result. Applying a simple optimization on the ReWave output, ReWave-GS can significantly improve the results.

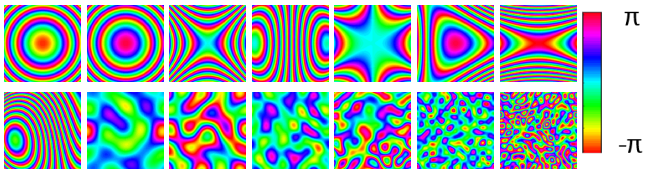
5 Synthetic Evaluation

We start with a synthetic comparison of our approach against prior work in wavefront sensing. This allows us to study the convergence and noise sensitivity of the approaches, independently of optics and implementation challenges. We used the 14 wavefronts in Fig. 5(a), all involving a uniform amplitude and a spatially-varying phase. This set includes eight linear combinations of Zernike wavefronts up to third orders and six random wavefronts at different resolutions. Each wavefront has a resolution of 150×150 pixels, zero padded inside a 450×450 pixels image. Propagation of wavefronts, when needed, is simulated by the angular spectrum method as described in [36]. Similarly, we account for the limitations induced by SLM pixelation for techniques that require phase modulation.

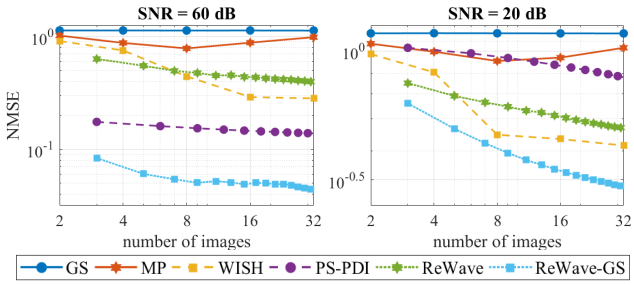
We use normalized mean squared error (NMSE) as the evaluation metric, following [23]; given an estimate, $\widehat{\mathbf{u}}$, for ground truth wavefront, $\mathbf{u}$, NMSE is defined as

$$\frac{\|\mathbf{u} - \widehat{\mathbf{u}}\|_2}{\|\mathbf{u}\|_2}.$$

Evaluation results are provided in Fig. 5(b). In all our simulations, we added noise to the measurements, leading to two different SNR values of 20 dB and 60 dB. We performed the evaluation as a function of the number of image intensity measurements provided to each algorithm. We compared the following approaches:



(a) Input phase patterns



(b) Reconstruction error comparison

Fig. 5: **Synthetic evaluation:** We compare several phase acquisition approaches using the phase masks in (a) as ground truth. The results of this evaluation, demonstrated in (b) consider low (left graph, SNR = 60dB) and high (right, SNR = 20dB) noise scenarios, as a function of the number of images provided to the algorithm. Each line represents one method, and each point is an averaged reconstruction error for the 14 different phase patterns in (a). Our ReWave-GS steadily outperforms all alternatives.

1) *Gerchberg–Saxton (GS).* This approach captures the intensity of the target field in the Fourier and image planes $|\mathcal{F}_{\mathbf{u}}|^2, |\mathbf{u}|^2$ and uses alternate optimization to solve for the phase [3]. The optimization is usually noise sensitive and often converges to local minima. As this approach only requires two input images, in the simulation of Fig. 5(b), we only use additional measurements to reduce noise.
2) *Multi-plane propagation (MP) [4].* In this approach, one places the target at multiple planes away from the imaging system, which is equivalent to adding a quadratic phase in the Fourier plane. Multiple intensity images are captured at different defocus settings. A GS-style iterative technique is used to solve for the phase.
3) *WISH [5].* This approach places different random phase masks a finite distance before the image plane sensor, capturing multiple intensity images that have been scrambled at different ways at the image plane. It solves for the phase as before, using iterative techniques [3].
4) *PS-PDI*, as described in Sec. 3, requires three input images; we use additional measurements to reduce noise by repeating the measurements and averaging.
5) *ReWave.* When more than three input images are allowed, we use different collision-free masks $\mathbf{M}_t$ and average their wavefront estimates obtained from each using Eq. (13).
6) *ReWave-GS*, as described in Sec. 4.7.2.

From the comparison in Fig. 5, the best results are achieved by the ReWave-GS approach. The WISH approach is the strongest competitor among all these algorithms, and the basic (unoptimized) ReWave approach overcomes it only when a low number of shots is provided. PS-PDI achieves good results in low noise scenarios but fails drastically in higher noise levels.

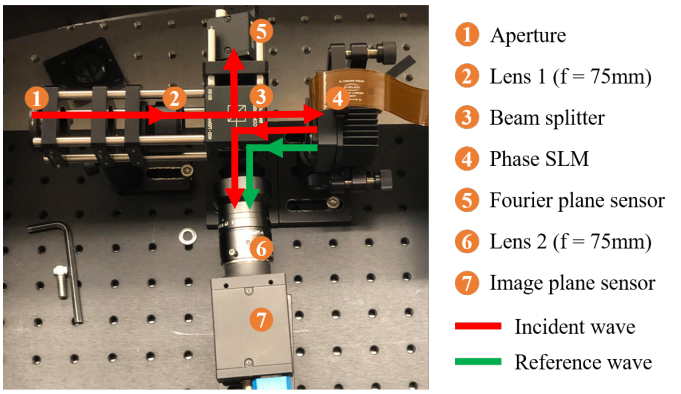


Fig. 6: **Hardware prototype:** Image of our prototype, analogous to the schematic in Fig. 2(b). The incident wave (not shown in this figure) enters via the aperture.

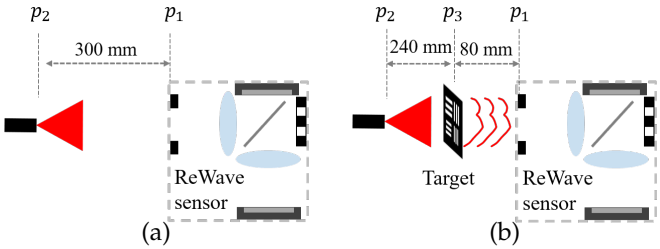


Fig. 7: **Setup for phase acquisition experiments:** (a) Setup used for Fig. 8, where a diverging spherical wave is measured at the input plane of the system. (b) Setup used for Fig. 9, where a resolution chart is placed $80mm$ before the input plane. The phase measured at the input plane is back-propagated to produce a sharp image at the chart plane.

6 Hardware Prototype and Results

Fig. 6 illustrates our hardware prototype, which follows the schematic in Fig. 2(b). We illuminate the setup using a $520nm$ laser source. We use a phase SLM (Holoeye GAEA-2, 4160×2464 pixels, $3.74\mu m$ pitch). Both relay lens and camera lens in the $4f$ relay have focal length $f = 75mm$, where the relay lens is calibrated to be at distance f from the SLM, and the camera lens is focus at infinity. The image plane sensor has 1920×1440 resolution with a pixel pitch of $4.54\mu m$. A typical challenge with LCoS-based phase modulation is that some DC component is added to every pattern displayed. This DC component adds ghosting to the pattern we observe on the sensor plane (see supplementary Fig. 14), which degrades the result. To avoid this ghosting, we add on the SLM a phase ramp which shifts the interference image on the sensor plane. The shifted image only occupied an area of 900×900 pixels on the sensor plane. Finally, we align the Fourier plane camera (1440×1080 resolution with $3.45\mu m$ pitch) with the SLM and calculate the exact transformation between them. Details about the system calibration are provided in App. B.

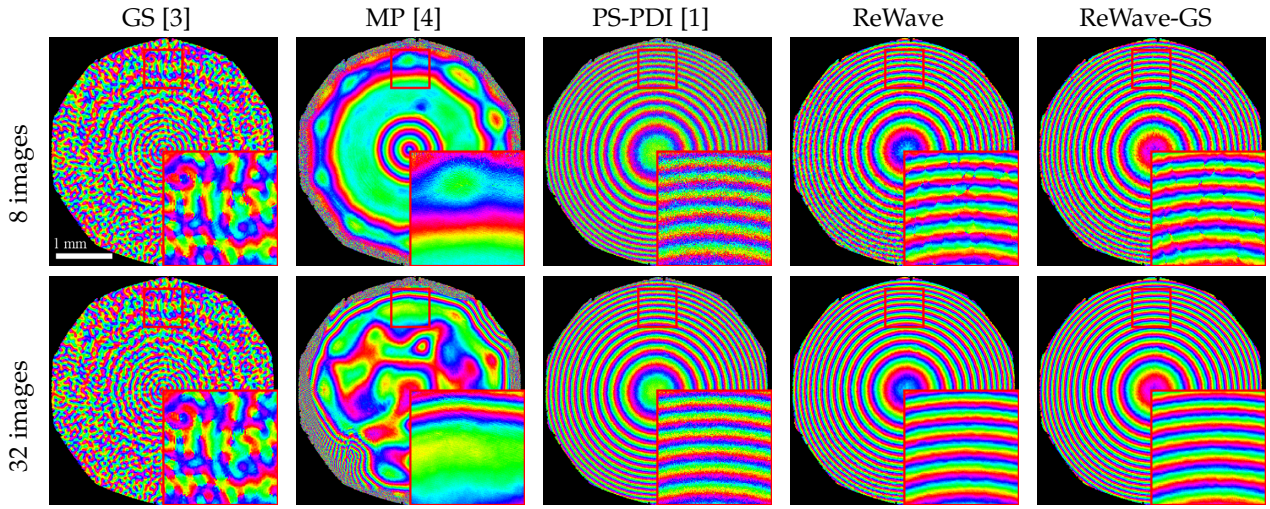


Fig. 8: **Diverging wave reconstruction:** We reconstruct the phase of a diverging wave with the different methods using 8 (top) and 32 (bottom) input images. From left to right, GS and MP fail to converge. Given 8 images PDI recovers the wave subject to some high frequency noise fluctuations, which improve given 32 images. Our ReWave approach achieves a clear phase pattern with only 8 images, and further improves given 32 images.

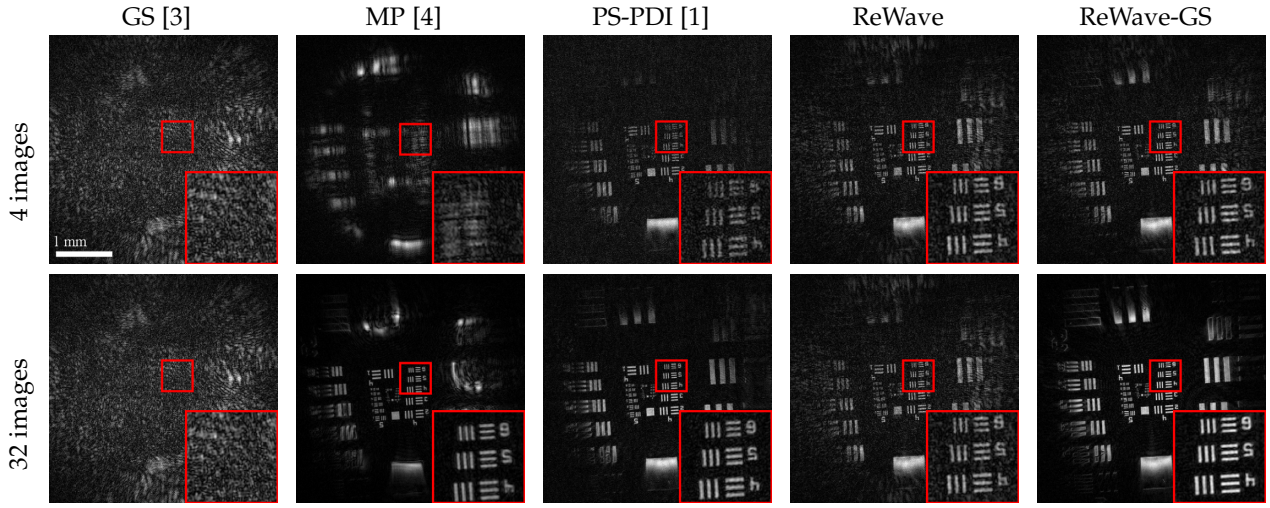


Fig. 9: **Resolution chart reconstruction:** We capture a resolution chart $80mm$ away from the input plane of the system (see setup in Fig. 7) and use the phase of the captured wave to refocus the camera at the chart plane.

6.1 Comparison on real data

We compare our algorithm against the competing alternatives mentioned in Sec. 5, excluding WISH [5] that cannot be realized with our setup. We will address more alternatives in App. D. In Fig. 8, we show reconstruction from data acquired in our setup, measuring the phase of a wave diverging from a point light source $300mm$ before the input plane of the system, as illustrated in Fig. 7(a). When only 8 input images are provided, GS and MP fail to converge to the global optimum, and PS-PDI suffers from high frequency noise which reduces its contrast. On the other hand, the proposed methods (ReWave and ReWave-GS) can both reconstruct a reasonably clear image despite small artifacts. Given 32 images, PDI and ReWave converge to accurate results.

Now we put a resolution chart a distance $80mm$ from the input plane, as illustrated in Fig. 7(b). Assuming we recover the phase at the input plane correctly, we can back-propagate it and simulate an image focused at the plane of the resolution chart. The result is illustrated in Fig. 9. Our ReWave and ReWave-GS approaches lead to the best results. PDI is also successful, though from only 4 images our approach leads to a better contrast. In this case, the MP converged to the desired result from 32 images.

In the following subsections, we demonstrate additional applications of our method. All the images are reconstructed by ReWave-GS using 32 images. *Please see animations in the supplementary video.*

6.2 Application 1: Imaging phase masks

In Fig. 10, we accurately capture the phase of a few target phase masks at the input plane of our ReWave setup. In the first example of Fig. 10, we display a phase-only text on another transmissive SLM (Holoeye LC2012) with increasing phase contrast from left to right. In the second example, we measure the phase of an holographic diffuser

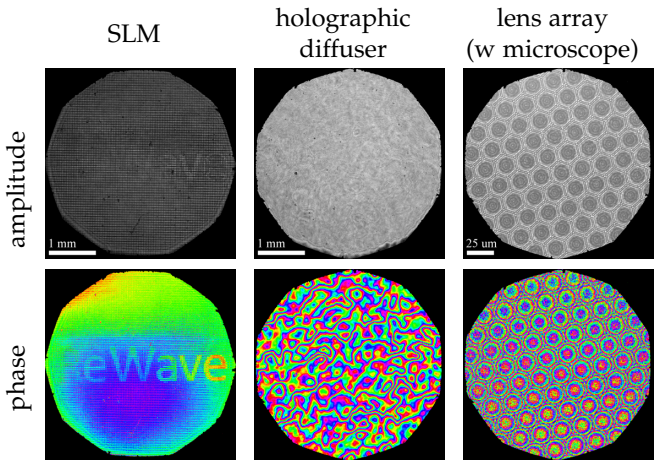


Fig. 10: **Phase masks imaging:** Left: text displayed on a phase SLM. Center: random phase patterns on a holographic diffuser. Right: quadratic phase pattern of a lenslet array. The patterns are often invisible from the corresponding intensity images.

(Edmund Optics 47-988) as in [5]. Unlike the capture in [5], we did not use a lens to shrink the pattern, so we observe a smaller region of the diffuser with a finer pixel pitch. In the last example of Fig. 10, we reconstruct the phase of a lens array. Since each lenslet has a very small diameter of around $20um$, we magnified it by a factor $\times 23$, adding a microscope to the system. This included a camera lens with a $100mm$ focal length as a tube lens and a $\times 40$ objective lens with a $4.33mm$ focal length. We can clearly observe the hexagonal structure array as well as the quadratic pattern within each lenslet.

6.3 Application 2: Refocusing

As we capture both the phase and amplitude of the incident wave, we can synthesize images focused at different planes, using basic wave propagation. This synthetic refocusing is particularly useful in microscopy, where the depth of field is extremely shallow. In the top two rows of Fig. 11, we captured two microscope slides at two different distances ($d_1 = 23mm$, $d_2 = 51mm$). By refocusing at these two planes, we see the upper left and bottom right parts of the image coming in and out of focus. In the bottom two rows, we used the microscope system mentioned in Sec. 6.2 to image a honey bee leg. The depth of field now is less then $0.1mm$. When we computationally propagate the wave to $d_1 = -40\mu m$ and $d_2 = 118\mu m$, we can refocus the wave and observe spots on the front or the hair at the back. *Please see the supplementary video for a continuous refocusing sequence of this scene.*

6.4 Application 3: Correcting Fresnel lens artifacts

We re-implement the Fresnel lens correction experiment of WISH [5]. A Fresnel lens is a thin mask with a lens-like quadratic phase pattern. However, the warping of the lens phase profile as well as imperfections in fabrication prevent it from creating a high-quality image as an ideal lens. In the setup of Fig. 12(b), we place a Fresnel lens

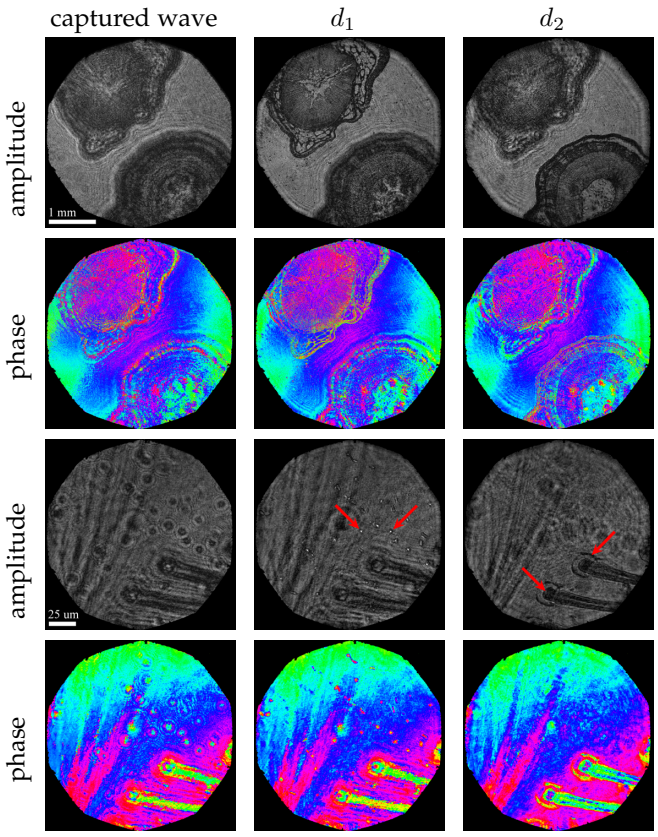


Fig. 11: **Refocusing:** Upper panel: patterns on two adjacent microscope slides, we can back-propagate the captured wave to focus on the front and rear slides. Lower panel: A bee leg tissue observed under a shallow depth of field microscope. While we cannot see an all-focus image, we can refocus the wave on the dust particles in the front or on the hair in the back. Complete refocusing results can be found in the supplementary video.

at plane p_2 $107mm$ before the plane p_1 imaged by our system. We place a resolution chart at plane p_3 at a distance $100mm$ beyond p_2. The focal length of the fresnel lens is $1/(1/107 + 1/100) \sim 51.7mm$, so that ideally we should get a sharp image of the chart. However, the poor lens quality leads to ghosting artifacts visualized in Fig. 12(d). To overcome this, we first used the setup of Fig. 12(a) to capture the phase of a Fresnel lens alone, illuminated by a simple plane wave. The recovered phase is illustrated in Fig. 12(c). We back-propagate the wave to get the phase of the Fresnel lens at plane p_2. In Fig. 12(e), we use this phase to correct the aberrations and display a sharp image of the calibration chart. For that, we first back-propagate the wave in p_1 to p_2, divide by the phase of Fresnel lens, and back-propagate it further to plane p_3.

Fig. 1 demonstrates a similar experiment using only 4 input images when reconstructing the Fresnel lens and compare the ReWave result with simple PS-PDI. The amount of noise involved significantly scrambles the result.

6.5 Application 4: Seeing through a diffuser

In Fig. 12(f-h), we use the same procedure as with the Fresnel lens to image through a diffuser. In the setup, a diffuser

is placed at plane p_2, $235mm$ before the input plane p_1, and a resolution chart is placed at plane p_3, $75mm$ beyond it. In Fig. 12(f), we capture the phase of the diffuser alone illuminated by a pure plane wave. This diffuser scrambles the light significantly. In Fig. 12(g), we attempt to back-propagate the wave captured on p_1 (the input plane of the setup) to plane p_3 where the chart is located, ignoring the presence of the diffuser, the chart is completely unrecognizable. However, using the calibrated phase from Fig. 12(f), we can correct some of the aberrations and obtain an image of the chart in Fig. 12(h). For that, we follow the same procedure as before: we propagate from p_1 to p_2, cancel the phase of the diffuser and continue to propagate to p_3. Here, we use a rather coarse diffuser (Thorlab DG-120) with a spread angle of up to $\pm 30°$. What makes the reconstruction quite challenging is the fact that the changes in the diffuser phase pattern is of a higher frequency than that of our imaging system, so we can not reconstruct the complete phase pattern. However, we can still reconstruct a smoothed version of the pattern by capturing the diffracting wave in a finite aperture, and thus back-propagate a band-limited wave.

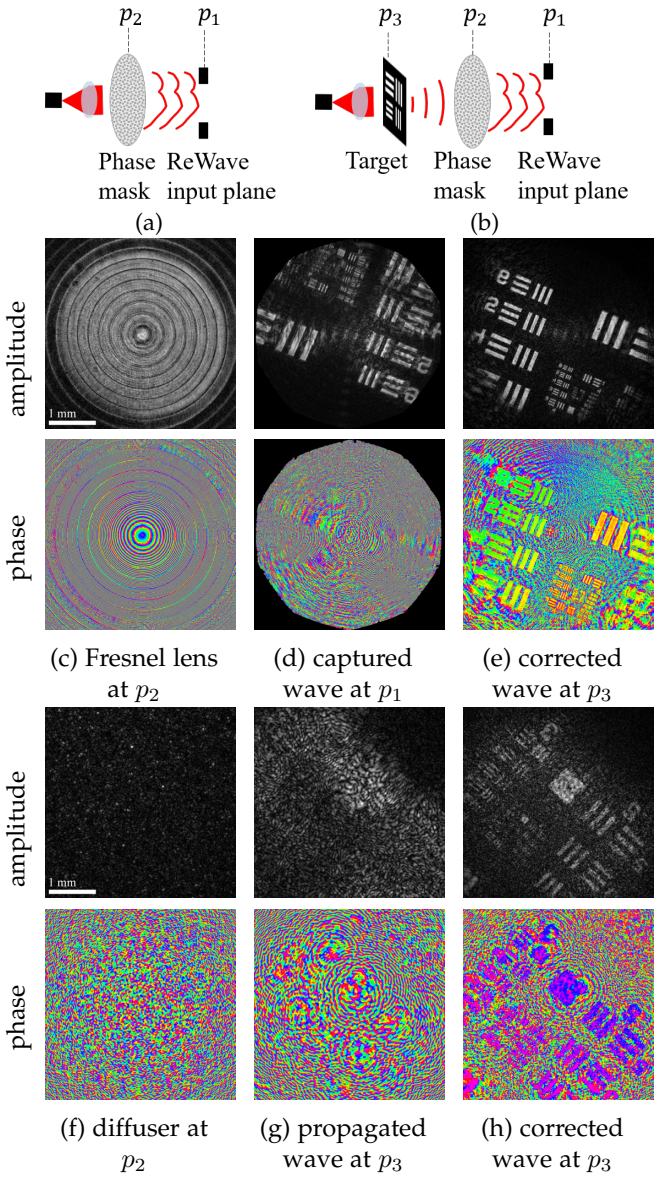


Fig. 12: **Correcting phase mask aberrations:** Both Fresnel lens and diffusers can be modeled as phase masks. (a) A setup for calibrating a phase mask (Fresnel lens or a diffuser): a mask placed at plane p_2 is illuminated by a plane wave. The resulting wave is captured at plane p_1, the input plane of our setup. (b) A setup imaging through the phase mask, where a resolution chart target is placed at plane p_3 behind the mask. The captured wave at p_1 is aberrated. Yet, applying the inverse of the calibrated mask, we can correct the aberrations and see through the mask. (c-e) Correcting the aberrations of a cheap Fresnel lens to get a sharp image of the target. (f-h) Seeing through a wide-angle diffuser. More details can be found in the supplementary video.

7 Conclusion

This paper presents a phase acquisition approach that builds on the self-interference idea of point diffraction interferometry (PDI). While PDI is highly noise sensitive due to the narrow pinhole used to create the reference wave, we introduce a reference mask that can admit much more energy. To recover the reference wave, the mask is designed to respect collision-free conditions which allow us to use simple techniques from sparse phase retrieval. We demonstrate the robustness of this new approach through a careful comparison against previous algorithms using both synthetic and real camera data. We also validate our approach with numerous applications for phase imaging including refocusing and imaging through diffusers.

Appendices

In this appendix, we detail supplementary material on three topics: the multi-reference fusion technique, details of calibration, and finally, some additional comparisons. Equation references here point at corresponding equations in the main manuscript.

Appendix A
Fusion of multiple phase estimates

To derive the fusion scheme in Eq. (13) of the main paper, we first need derive the statistical properties of the estimates derived from each reference wave. This involves starting with a realistic noise model of the measurements, including photon and read noise, and propagating the distributions, and, in particular, the means and covariances, through the relevant expressions.

Noise model. The complete probabilistic model for a camera signal is:

$$\widetilde{b}(x) \sim \frac{G}{F}\mathcal{P}ois\left(\frac{F}{G}b(x)\right) + \mathcal{N}(0, F^2 10^{-\frac{R}{10}}),$$

where G is the gain, F is the full well capacity, and R is the dynamic range. Multiply both sides by $\frac{F}{G}$, we have:

$$\frac{F}{G}\widetilde{b}(x) \sim \mathcal{P}ois\left(\frac{F}{G}b(x)\right) + \mathcal{N}(0, G^2 10^{-\frac{R}{10}})$$

Defining the image is on the scale that $I(x) = \frac{F}{G}b(x)$, and $\sigma = G10^{-\frac{R}{20}}$,

$$\widetilde{I}(x) \sim \mathcal{P}ois(I(x)) + \mathcal{N}(0, \sigma^2).$$

Lets consider the variable $\beta(x) = (\mathbf{u}(x) - \mathbf{r}(x))\mathbf{r}(x)^*$ as defined in Eq. (6) of the main paper. Suppose we reconstruct $\beta(x)$ from $K = 3$ phase shifted measurements. The expression for the imaginary component of $\beta(x)$ is given as

$$\begin{aligned}\mathrm{Im}(\widetilde{\beta}(x)) &= \frac{1}{3}\sum_{k=1}^{3}\widetilde{I}_k(x)\left(\frac{-1+\sqrt{3}i}{2}\right)^{k-1} \\ &= \frac{1}{2\sqrt{3}}(\widetilde{I}_2(x) - \widetilde{I}_3(x)).\end{aligned}$$

Lets first assume that read noise is minimal, and that the measurements $\widetilde{I}_k$ are corrupted by Poisson noise. With this, $\mathrm{Im}(\widetilde{\beta}(x))$ is a scaling of $\widetilde{Z} \sim \mathcal{P}ois(I_2(x)) - \mathcal{P}ois(I_3(x))$, which follows a Skellam distribution. While Skellam has a complicated probability density function, it can be approximated by a normal distribution [37] when its variance is much larger than it mean.

We can now derive the mean and variance of $\widetilde{Z}$.

$$\begin{aligned}\mathrm{E}[\widetilde{Z}] &= \mathrm{E}[(\widetilde{I}_2(x) - \widetilde{I}_3(x)] \\ &= I_2(x) - I_3(x) = 2\sqrt{3}\,\mathrm{Im}(\beta(x))\end{aligned}$$

Since variance of a Poisson random variable is just its mean, we have

$$\mathrm{var}[\widetilde{Z}] = I_2(x) + I_3(x) \approx 2|\mathbf{u}(x)|^2$$

The approximation made above assumes that the intensity measured at a pixel, which is the interference between the wave and reference, is approximately equal to $|\mathbf{u}(x)|^2$. Since $\mathrm{Im}(\beta(x)) = \mathrm{Im}(\mathbf{u}(x)\mathbf{r}(x)^*)$, we operate under the assumption that $|\mathbf{r}| \ll |\mathbf{u}|$, and it immediately follows that $\mathrm{E}[\widetilde{Z}] \ll \mathrm{var}[\widetilde{Z}]$. Therefore, we can invoke the Skellam - Gaussian approximation and obtain

$$\widetilde{Z} \sim \mathcal{N}\left(2\sqrt{3}\,\mathrm{Im}(\beta(x)), 2|\mathbf{u}(x)|^2\right)$$

Now taking the read noise into consideration,

$$2\sqrt{3}\mathrm{Im}(\widetilde{\beta}(x)) \sim \mathcal{N}\left(2\sqrt{3}\,\mathrm{Im}(\beta(x)), 2|\mathbf{u}(x)|^2 + 2\sigma^2\right)$$

Finally,

$$\mathrm{Im}(\widetilde{\beta}(x)) \sim \mathcal{N}\left(\mathrm{Im}(\beta(x)), \frac{(|\mathbf{u}(x)|^2 + \sigma^2)}{6}\right)$$

The real part can be derived in the same way.

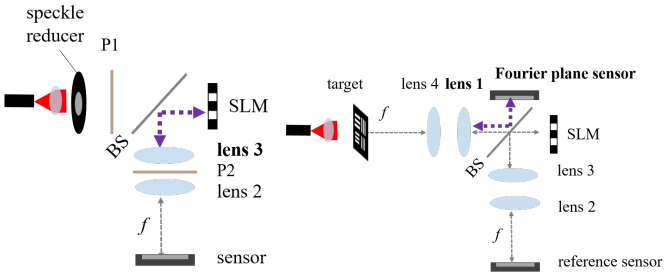
$$Re(\widetilde{\beta}(x)) \sim \mathcal{N}\left(Re(\beta(x)), \frac{(|\mathbf{u}(x)|^2 + \sigma^2)}{6}\right)$$

An important conclusion is that the variance of the estimates of $\widetilde{\beta}(x)$ is not dependent on the actual reference wave; this motivates the fusion scheme when we measure with multiple reference waves. Note that the above result *does not* imply that the phase estimates are independent of the reference wave; this is easily seen when we try to recover $\mathbf{u}(x)$ from $\widetilde{\beta}(x)$ which requires dividing by $|\mathbf{r}(x)|^2$.

Fusing measurements from multiple reference waves. Now we want to merge several measurements under different $\mathbf{r}_t$. As we have shown above, the estimates $\widetilde{\beta}(x)$ follow a normal distributions, and each distribution has the same variance, regardless of $\mathbf{r}_t$. To proceed further, we need a statistical model on $\mathbf{r}(x)$ which we assume has very low variance, since it is estimated from an overdetermined problem; specifically, the sparse Fourier transform of length K is measured from K^2 measurements of its autocorrelation. Hence, we can assume $\mathbf{u}(x)\mathbf{r}_t(x)^* = \beta(x) + |\mathbf{r}_t(x)|^2$ would still be a normal distribution with the same variance. Therefore, we can solve for $\mathbf{u}$ by solving the linear system below:

$$\begin{bmatrix}\mathbf{r}_1(x)^* \\ \vdots \\ \mathbf{r}_T(x)^*\end{bmatrix}\bar{\mathbf{u}}(x) = \begin{bmatrix}\beta_1(x) + |\mathbf{r}_1(x)|^2 \\ \vdots \\ \beta_T(x) + |\mathbf{r}_T(x)|^2\end{bmatrix}$$

The solution reduces to $\bar{\mathbf{u}}(x) = \frac{\sum_t(\beta_t(x)+|\mathbf{r}_t(x)|^2)\mathbf{r}_t(x)}{\sum_t|\mathbf{r}_t(x)|^2}$, as stated in Eq. (13).



(a) align sensor with SLM (b) align two sensors

Fig. 13: **Calibration setup:** To calibrate the transformation between the Fourier plane sensor and the SLM, we use a two step process, solving for a transformation between the SLM and the image plane sensor and the transformation between the image plane and Fourier plane sensors. For each step, we use the two setups in (a) and (b), where the purple arrow and the bold text highlight the key parts for each step. (a) Given lens 2 is already f-away from the sensor, we align lens 3 so the sensor can observe the focused intensity pattern on the SLM. P1 and P2 are rotated polarizers to operate the SLM in amplitude modes. We use this to compute the transformation between the SLM and the sensor. (b) We use lens 1 and 4 to form a relay between the SLM and the target. We align the Fourier plane sensor to be f away from lens 1 by adjusting it to see a focused image of the target. Using the SLM as a mirror both sensors observe the same target and the transformation can be calculated.

APPENDIX B
CALIBRATION DETAILS

Here we provide more details on the calibration of our setup. The first challenge is to align the Fourier plane sensor with the SLM and calculate the exact transformation between them. To do that, we introduce an additional lens that make the image plane sensor observe the Fourier plane as well, and use it to align the actual Fourier plane sensor with the SLM. For that, we use Fig. 13(a), where we added lens 3 to the actual setup. This creates a relay system between

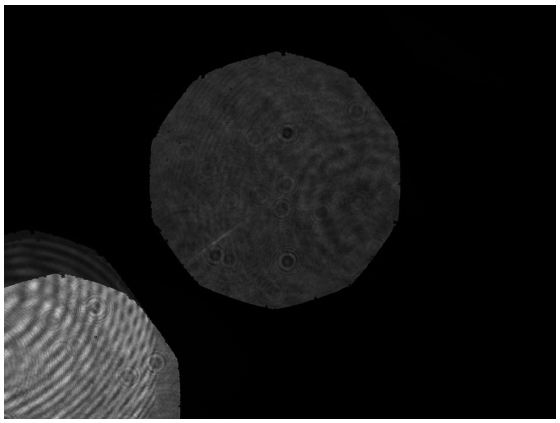


Fig. 14: **SLM ghosting:** Due to SLM imperfections, the phase pattern we display involves a DC component. In the image plane, this creates a constant part of the field (lower left disc) which is added to the measurement. To overcome this, we added a phase ramp to the SLM pattern, which shifts the desired part of the wave on the image sensor (central disc). This has the disadvantage of limiting the sensor area we can use, but allows the measurement of a clean signal at the usable pixels.

the lower sensor, switching it to the the Fourier plane rather than the image plane. In this configuration the sensor can directly image the SLM. We operate the SLM in amplitude mode rather than phase mode by introducing 2 orthogonal polarizers. We also add a speckle reducer to make laser light source incoherent. Given the distance between lens 2 and the camera have been calibrated to be f, we only need to adjust the position of lens 3 to observe a clear image on the SLM. We display a known pattern on the SLM and use it to compute an affine 2D transformation between the sensor and the SLM (we solve for scale and translation).

In the next step, we build another relay system between the target and the SLM. For that, we add to the setup lens 4, as illustrated in Fig. 13(b). We adjust the position of the Fourier plane sensor to observe a sharp image of the target. If we now use the SLM as a planar mirror, both sensors observe the same target, and we can calculate the transformation between both sensors.

Combining the two affine transformation estimated above, we now have the alignment between the SLM and the Fourier plane sensor. We can now remove lens 3 and 4 to get our final setup. Note that to avoid the ghosting problem as in Sec. 6 and in Fig. 14, we display a phase ramp on the SLM to shift the image, and adjust the sensor location accordingly to observe the image.

APPENDIX C
ADDITIONAL PHASE IMAGING RECONSTRUCTION

In Fig. 15, we show reconstruction results for the diverging wavefront as in Fig. 8 but given only 4 images. While the diverging wavefront needs 8 input images for a clear reconstruction, our ReWave method can still give a reasonably good result given 4 images.

APPENDIX D
ADDITIONAL PHASE IMAGING ALTERNATIVES

Here we address two additional competing alternatives for Fig. 8 and Fig. 9. First, we implement another version of the multi-plane (MP) approach where the quadratic function on the SLM corresponds to a different propagation range. In the implementation in the main paper, we use quadratic patterns corresponding to a $[20, 50]mm$ propagation range. Here, the propagation range corresponding to $[240, 270]mm$, which we denote as MP(far) below. This experiment demonstrates that the performance of the MP algorithm highly depends on the propagation range where it is used. The far MP range better matches the phase of the diverging point source, which was placed at distance $300mm$. Unsurprisingly, the reconstruction of the quadratic wave in Fig. 16 is now significantly better. The original MP in Fig. 8 failed in this experiment. However, for the resolution chart experiment in Fig. 17, MP(far) fails completely on 4 images and give degraded ones on 32 images, while the original MP range in Fig. 9 gave much better results.

We also attempt to place some equivalence of WISH [5], where we placed a random mask at the Fourier plane where our SLM is positioned, rather than a finite distance before the sensor as in the original design. We refer to this as Fourier-WISH. In Fig. 16, this Fourier-WISH fails to reconstruct the diverging wave, but succeeds using 32 images. In Fig. 17, Fourier-WISH fail to converges given 4 images, and still has artifacts with 32 images.

ACKNOWLEDGMENTS

This work was supported in part by the NSF CAREER grant CCF-1652569 and the NSF Expeditions in Computing Grant # 1730147. WYC also gratefully acknowledges support via the Wei Shen and Xuehong Zhang Presidential Fellowship. AL acknowledges ERC 635537, ISF 1947, NSF-BSF 2019758.

REFERENCES

[1] V. Akondi, A. Jewel, and B. Vohnsen, "Digital phase-shifting point diffraction interferometer," *Optics Letters*, vol. 39, no. 6, pp. 1641–1644, 2014.

[2] B. C. Platt and R. Shack, "History and principles of shack-hartmann wavefront sensing," *Journal of Refractive Surgery*, vol. 17, no. 5, pp. S573–S577, 2001.

[3] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied Optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[4] G. Pedrini, W. Osten, and Y. Zhang, "Wave-front reconstruction from a sequence of interferograms recorded at different planes," *Optics Letters*, vol. 30, no. 8, pp. 833–835, 2005.

[5] Y. Wu, M. K. Sharma, and A. Veeraraghavan, "Wish: wavefront imaging sensor with high resolution," *Light: Science & Applications*, vol. 8, no. 1, pp. 1–10, 2019.

[6] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[7] T. Bendory, R. Beinert, and Y. C. Eldar, "Fourier phase retrieval: Uniqueness and algorithms," in *Compressed Sensing and its Applications*. Springer, 2017, pp. 55–91.

[8] J. H. Bruning, D. R. Herriott, J. Gallagher, D. Rosenfeld, A. White, and D. Brangaccio, "Digital wavefront measuring interferometer for testing optical surfaces and lenses," *Applied Optics*, vol. 13, no. 11, pp. 2693–2703, 1974.

[9] P. Hariharan, B. Oreb, and T. Eiju, "Digital phase-shifting interferometry: a simple error-compensating phase calculation algorithm," *Applied Optics*, vol. 26, no. 13, pp. 2504–2506, 1987.

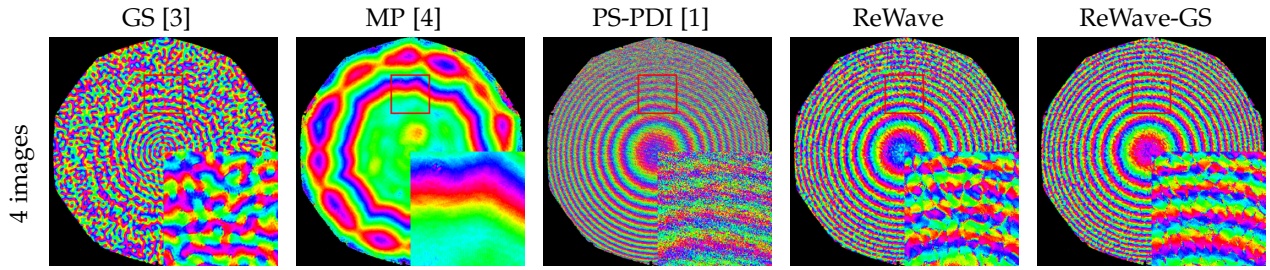


Fig. 15: **Diverging wave reconstruction under 4 images:** We compare the reconstructed phase of a diverging wave with different methods as in Fig. 8 but using 4 input images. From left to right, GS and MP fail to converge, and PS-PDI give a very noisy reconstruction. On the other hand, our ReWave approach, while not perfect, shows a less noisy reconstruction.

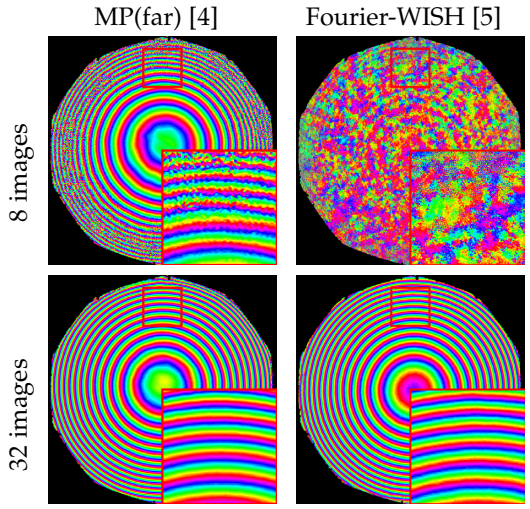


Fig. 16: **Additional diverging wave reconstructions:** We follow on Fig. 8, demonstrating the reconstruction of a diverging wave using MP(far) at a different propagation range and a Fourier-WISH setup which displays a phase mask at the Fourier plane sensor.

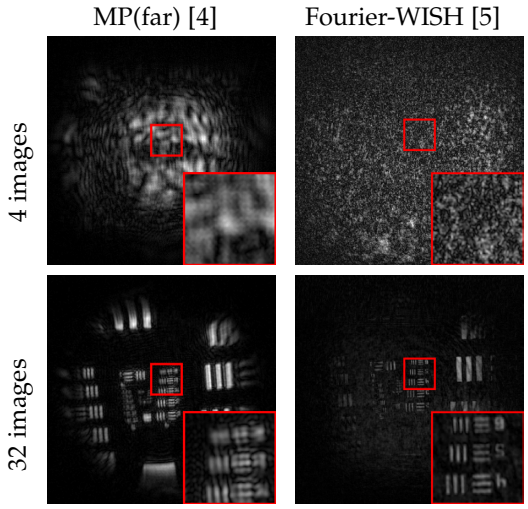


Fig. 17: **Additional resolution chart reconstructions:** We follow on Fig. 9 reconstructing a resolution chart $80mm$ away from the input plane of the system, and use the phase of the captured wave to refocus the camera at the chart plane. We demonstrate results with the MP(far) phase patterns and the Fourier-WISH masks.

[10] U. Schnars, C. Falldorf, J. Watson, and W. Jüptner, *Digital Holography and Wavefront Sensing*. Springer, 2016.

[11] R. Smartt and W. Steel, "Theory and application of point-diffraction interferometers," *Japanese Journal of Applied Physics*, vol. 14, no. S1, p. 351, 1975.

[12] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli, "Phase retrieval for sparse signals: Uniqueness conditions," *submitted to IEEE Transactions on Information Theory*, 2013.

[13] R. Ragazzoni, "Pupil plane wavefront sensing with an oscillating prism," *Journal of Modern Optics*, vol. 43, no. 2, pp. 289–293, 1996.

[14] A. Burvall, E. Daly, S. R. Chamot, and C. Dainty, "Linearity of the pyramid wavefront sensor," *Optics Express*, vol. 14, no. 25, pp. 11 925–11 934, 2006.

[15] F. Roddier, "Curvature sensing and compensation: a new concept in adaptive optics," *Applied Optics*, vol. 27, no. 7, pp. 1223–1225, 1988.

[16] M. R. Teague, "Deterministic phase retrieval: a green's function solution," *JOSA*, vol. 73, no. 11, pp. 1434–1441, 1983.

[17] C. Wang, X. Dun, Q. Fu, and W. Heidrich, "Ultra-high resolution coded wavefront sensor," *Optics Express*, vol. 25, no. 12, pp. 13 736–13 746, 2017.

[18] D. Noll and A. Rondepierre, "On local convergence of the method of alternating projections," *Foundations of Computational Mathematics*, vol. 16, no. 2, pp. 425–455, 2016.

[19] B.-Y. Wang, L. Han, Y. Yang, Q.-Y. Yue, and C.-S. Guo, "Wavefront sensing based on a spatial light modulator and incremental binary random sampling," *Optics Letters*, vol. 42, no. 3, pp. 603–606, 2017.

[20] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.

[21] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[22] K. Jaganathan, Y. Eldar, and B. Hassibi, "Phase retrieval with masks using convex optimization," in *IEEE International Symposium on Information Theory (ISIT)*, 2015.

[23] Y. Shechtman, A. Beck, and Y. C. Eldar, "Gespar: Efficient phase retrieval of sparse signals," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 928–938, 2014.

[24] D. Kogan, S. Sivankutty, V. Tsvirkun, G. Bouwmans, E. R. Andresen, H. Rigneault, and D. Oron, "Phase retrieval in multicore fiber bundles," *Optics Letters*, vol. 42, no. 3, pp. 647–650, 2017.

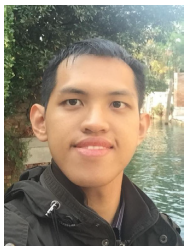
[25] S. Sivankutty, V. Tsvirkun, O. Vanvincq, G. Bouwmans, E. R. Andresen, and H. Rigneault, "Nonlinear imaging through a fermat's golden spiral multicore fiber," *Optics Letters*, vol. 43, no. 15, pp. 3638–3641, 2018.

[26] J. M. Beckers, "Adaptive optics for astronomy: principles, performance, and applications," *Annual Review of Astronomy and Astrophysics*, vol. 31, no. 1, pp. 13–62, 1993.

[27] S. Velghe, J. Primot, N. Guérineau, R. Haïdar, S. Demoustier, M. Cohen, and B. Wattellier, "Advanced wave-front sensing by quadri-wave lateral shearing interferometry," in *Interferometry XIII: Techniques and Analysis*, vol. 6292. International Society for Optics and Photonics, 2006, p. 62920E.

[28] J. E. Oti, V. F. Canales, and M. P. Cagigal, "Analysis of the signal-to-noise ratio in the optical differentiation wavefront sensor," *Optics*

Express, vol. 11, no. 21, pp. 2783–2790, 2003.
[29] H. Medecki, E. Tejnil, K. Goldberg, and J. Bokor, "Phase-shifting point diffraction interferometer," *Optics Letters*, vol. 21, no. 19, pp. 1526–1528, 1996.
[30] C. Wykes, "A theoretical approach to the optimization of electronic speckle interferometry fringes with limited laser power," *Journal of Modern Optics*, vol. 34, no. 4, pp. 539–554, 1987.
[31] J. Notaras and C. Paterson, "Point-diffraction interferometer for atmospheric adaptive optics in strong scintillation," *Optics Communications*, vol. 281, no. 3, pp. 360–367, 2008.
[32] I. McNulty, J. Kirz, C. Jacobsen, E. H. Anderson, M. R. Howells, and D. P. Kern, "High-resolution imaging by fourier transform x-ray holography," *Science*, vol. 256, no. 5059, pp. 1009–1012, 1992.
[33] D. A. Barmherzig, J. Sun, P.-N. Li, T. J. Lane, and E. J. Candès, "Holographic phase retrieval and reference design," *Inverse Problems*, vol. 35, no. 9, p. 094001, 2019.
[34] C. S. Anderson, "Fringe visibility, irradiance, and accuracy in common path interferometers for visualization of phase disturbances," *Applied Optics*, vol. 34, no. 32, pp. 7474–7485, 1995.
[35] Y. Chen, M. K. Sharma, A. Sabharwal, A. Veeraraghavan, and A. C. Sankaranarayanan, "3PointTM: Faster measurement of high-dimensional transmission matrices," in *Euro. Conf. Computer Vision (ECCV)*, 2020.
[36] D. Voelz, "Computational fourier optics: a matlab tutorial." Society of Photo-Optical Instrumentation Engineers, 2011.
[37] D. Karlis and I. Ntzoufras, "Bayesian analysis of the differences of count data," *Statistics in Medicine*, vol. 25, no. 11, pp. 1885–1905, 2006.



Wei-Yu Chen is a Ph.D. student in ECE Department at Carnegie Mellon University. He received his B.S. in EE in 2015 and M.S. in 2017 from National Taiwan University, Tapiei, Taiwan. He is the recipient of Wei Shen and Xuehong Zhang Presidential Fellowship at Carnegie Mellon University in 2021, and his research interests are optics, computational imaging and computer vision.

Anat Levin is an Associate Professor at the department of Electrical Engineering, Technion, Israel, doing research in the field of computational imaging. She received her Ph.D. from the Hebrew University at 2006. During the years 2007-2009 she was a postdoc at MIT CSAIL, and during 2009-2016 she was an Assistant and Associate Professor at the department of Computer Science and Applied Math, the Weizmann Inst. of Science.

Matthew O'Toole is an Assistant Professor with the Robotics Institute and Computer Science Department at Carnegie Mellon University (CMU). He earned his Ph.D. from the University of Toronto and was previously a Banting postdoctoral fellow with the Computational Imaging group at Stanford University. He is the recipient of the ACM SIGGRAPH 2017 Outstanding Thesis Honorable Mention award, and runner-up best paper awards at CVPR 2014 and ICCV 2007.

Aswin C. Sankaranarayanan is an Associate Professor in the ECE Department at Carnegie Mellon University (CMU). He earned his Ph.D. from University of Maryland, College Park and was a post-doctoral researcher at the DSP group at Rice University before joining CMU. Aswin's research spans topics in imaging, vision, and image processing. He is the recipient of the CVPR 2019 best paper award, the CIT Dean's Early Career Fellowship in 2018, the NSF CAREER award in 2017, the Eta Kappa Nu (Sigma chapter) Excellence in Teaching award in 2017, the 2016 Herschel M. Rich invention award, and the distinguished dissertation fellowship from the ECE Department at University of Maryland in 2009.