

Provable and Robust Wavefront Sensing via Self-Reference Interferometry Supplementary Material

Nebiyou Yismaw¹, Vishwanath Saragadam¹,
Aswin C. Sankaranarayanan² and M. Salman Asif¹ (✉)

¹ University of California, Riverside, CA, USA
{nyismaw, vishwans, sasif}@ucr.edu

² Carnegie Mellon University, Pittsburgh, PA, USA
saswin@andrew.cmu.edu

This supplementary material provides theoretical proofs, implementation details, experimental settings, and additional results supporting the main paper. Complete proofs of all lemmas and theorems are presented in Sec. [A1](#) along with additional noise analysis in Sec. [A2](#). We discuss additional algorithm implementation details in Sec. [A3](#). We provide hardware specifications, experimental settings, and additional real and simulated experiments in Secs. [A4](#) to [A6](#).

A1 Proofs of Theoretical Results

In this section, we provide complete proofs of our lemmas and theorems stated in the main paper. For completeness, we first restate the relevant notation. Each pixel or measurement location $i \in \{-\lfloor N/2 \rfloor, \dots, \lfloor N/2 \rfloor - 1\}$ corresponds to a *node* in the graph. The set of graph nodes is

$$V = \left\{ -\left\lfloor \frac{N}{2} \right\rfloor, \dots, \left\lfloor \frac{N}{2} \right\rfloor - 1 \right\}, \quad (1)$$

where $|V| = N$. An *edge* exists between node i and j if we have direct phase difference measurements between them as $\nabla_k \phi$. For a set of shifts $\mathcal{S} = \{s_1, s_2, \dots, s_K\}$, the edge set is

$$E(\mathcal{S}) = \{(i, i + s_k) : i, i + s_k \in V, s_k \in \mathcal{S}\}. \quad (2)$$

Each shift contributes an edge between i and $i + s_k$ for every valid i . We call one edge traversal a step or hop and the total number of hops/steps required to go from a reference node to any other node as the hop distance.

A1.1 Connectivity guarantees

We now provide a complete proof of our Theorem [1](#). Our main objective is to show that any node in the graph can be accessed from the reference node through a sequence of hops/steps along the connected nodes. In other words, given two

co-prime integers s, t , we can represent any node as $n = \alpha s + \beta t$, using integers α, β that keep the entire path within the graph (V, E) .

Our required condition follows directly from Bezout's identity [2], which for co-prime s and t (*i.e.* $\gcd(s, t) = 1$) guarantees that any integer can be written as a linear combination $\alpha s + \beta t$. This identity alone does not prove connectivity in our case because the intermediate nodes visited by such combinations are not guaranteed to remain within the valid range of V . In other words, the identity only states the existence of integers α, β such that $\alpha s + \beta t = n$, but the corresponding sequence of hops (*i.e.* αs or βt) may leave the valid domain before reaching the desired node. Therefore, we need some additional constraints on s and t to prove that valid propagation paths exist entirely inside V that connect all nodes.

Our proof proceeds in two steps. First, Lemma 1 shows that when s and t are co-prime, repeated hops of size s generate all possible residues modulo $s + t$ (*i.e.* $0, 1, \dots, s + t - 1$). Note that a $-t$ step is equivalent to a $+s$ step modulo $s + t$ since $-t \equiv s \pmod{s + t}$. Thus, propagation using $+s$ hops, with $-t$ hops when the traversal reaches the boundary of the interval, visits all residues modulo $s + t$. As illustrated in Fig. A1, these hops cover every index within any interval of length $s + t$. Next, Lemma 2 shows that such intervals can be translated across the domain while remaining inside V as long as $s + t \leq N$. This allows connectivity to propagate from the reference node to the entire graph. Combining these results establishes the global reachability and connectivity stated in Theorem 1.

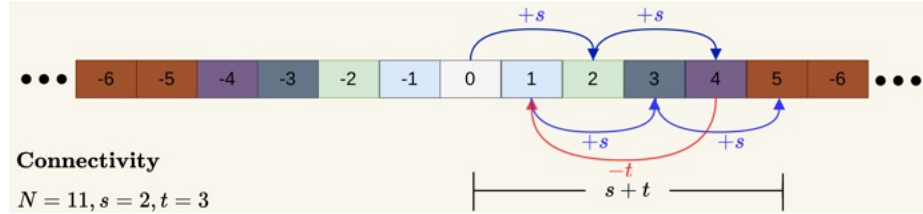


Fig. A1: Illustration of the connectivity using two co-prime shifts for $N = 11$, $s = 2$, and $t = 3$ using hops $+s$ and $-t$. Blue arrows represent forward propagation using $+s$, while the red arrow represents taking a step of size $-t$ when the boundary is reached. Over an interval of length $s + t$ since s and t are co-prime, the hops will visit all nodes without repetition.

Lemma 1 (Co-prime Residue Coverage). *Let $k \geq 1$, $s \in \mathbb{Z}$ and $\gcd(s, k) = 1$. Then for any $p \in \mathbb{Z}$,*

$$\{(p + is) \bmod k : i = 0, 1, \dots, k - 1\} = \{0, 1, \dots, k - 1\}.$$

Proof. Consider the k residues $\{(p + is) \bmod k : i = 0, 1, \dots, k - 1\}$. Showing that each of the k residues is unique will complete the proof, because we know there are k possible non-negative remainders that are less than k . If they are distinct, it means that the set is exactly $\{0, 1, \dots, k - 1\}$.

Let us assume that they are not distinct and for some i_1, i_2 where $0 \leq i_1 < i_2 \leq k-1$ and $i_1 \neq i_2$ we have

$$p + i_1 s \equiv p + i_2 s \pmod{k}, \quad (3)$$

which implies $(i_2 - i_1)s \equiv 0 \pmod{k}$. This is true because Eq. (3) would imply that $\exists m, n \in \mathbb{Z}$ such that

$$(p + i_1 s) - nk = (p + i_2 s) - mk \implies (i_2 - i_1)s = (m - n)k.$$

We also know that $\gcd(s, k) = 1$, $(i_2 - i_1)s \equiv 0 \pmod{k}$, which leads to $i_2 - i_1 \equiv 0 \pmod{k}$. This is a contradiction because we know $1 \leq i_2 - i_1 \leq k-1$ and $i_2 - i_1 \equiv 0 \pmod{k}$ is valid if and only if $i_2 = i_1$, which contradicts our assumption.

The argument above proves that all k residues are distinct and the set must exactly be $\{0, \dots, k-1\}$. $\square$

Lemma 2 (Sliding Window Coverage). *Let $s, t \in \mathbb{N}$ with $s+t \leq N$, and let V be defined as Eq. (1). For any $x \in V$, there exists an interval I_1 of size $s+t$ containing x , and a finite sequence of intervals $I_1, I_2, \dots, I_K$ (for some $K \in \mathbb{N}$) of size $s+t$, each contained in V , such that consecutive intervals share at least one common element and I_K contains 0.*

Proof. Since $s+t \leq N = |V|$, any interval of size $s+t$ fits within V . For any $x \in V$, set the starting point of I_1 as

$$I_1 = [i_0, i_0 + s + t - 1] \quad \text{where} \quad i_0 = \min(x, \left\lfloor \frac{N}{2} \right\rfloor - s - t + 1), \quad (4)$$

which contains x and lies within V . If $0 \in I_1$, then $K = 1$ and the sequence is complete. Otherwise, we construct $I_2, \dots, I_K$ by sliding each interval by one position towards 0, so the consecutive intervals overlap in $s+t-1$ elements. This sequence ends in at most $|i_0|$ slides, i.e. $K \leq |i_0| + 1$, as each slide reduces the distance to 0 by exactly one. Thus, we guarantee I_K contains 0. $\square$

Theorem 1 (Connectivity via Co-Prime Shifts). *Let $s, t \in \mathbb{N}$ be co-prime, i.e., $\gcd(s, t) = 1$, and let the graph V be defined as Eq. (1) and edges $E(\{s, t\})$ defined following Eq. (2) using shifts s and t . Then every node in V is reachable from the reference node 0 through a sequence of hops along edges, provided $s+t \leq N$.*

Proof. By Bezout's identity (2) since s and t are co-prime, for any integer n , there exist integers α, β such that

$$\alpha s + \beta t = n.$$

Therefore, any integer n can be expressed as a linear combination of s and t . In particular, we can take α hops of size s and β hops of size t to reach n . However, this identity alone does not establish connectivity, as it does not guarantee that the hops land on valid nodes inside V .

Next, we show that under the condition $s + t \leq N$, there exists a valid sequence of paths that remain entirely within V and connect any node in V to the reference node. Consider the propagation sequences n_k , where we start from node $n_0 = 0$ take $+s$ steps to the right until we reach the boundary $\lfloor N/2 \rfloor$. If we reach the boundary, we take the step $-t$. We can write these sequences as

$$n_k = \begin{cases} n_{k-1} + s & \text{if } n_{k-1} + s \leq \lfloor N/2 \rfloor \\ n_{k-1} - t & \text{otherwise .} \end{cases} \quad (5)$$

Notice that the iteration can proceed for arbitrarily many k without leaving V , i.e., $n_k \in V$ for all k . This is because if $n_{k-1} + s > \frac{N}{2}$, then we must have $n_{k-1} - t > -\frac{N}{2}$. This follows from

$$\text{if } n_{k-1} + s > \frac{N}{2} \implies n_{k-1} - t > \frac{N}{2} - s - t \geq \frac{N}{2} - N \geq -\frac{N}{2}.$$

We can see that n_k and $n_0 + ks$ have the same remainder when divided by $s + t$. Formally, $n_k \equiv n_0 + ks \pmod{s+t}$. At each step, the update is either

$$n_{k+1} = n_k + s \quad \text{or} \quad n_{k+1} = n_k - t.$$

In both cases, $n_{k+1} \equiv n_k + s \pmod{s+t}$, since $-t \equiv s \pmod{s+t}$. Therefore, $n_k \equiv n_0 + ks \pmod{s+t}$.

Let us now consider an interval of length $s + t$ starting from node 0. Since $\gcd(s, t) = 1$, it follows that $\gcd(s, s+t) = 1$. By Lemma 1, the sequence $n_0 + ks$ for $k = 0, \dots, s+t-1$ attains all residues $\{0, 1, \dots, s+t-1\} \pmod{s+t}$. Since the hop sequence n_k is congruent with $n_0 + ks \pmod{s+t}$, the sequence n_k connects all nodes starting from n_0 up to n_{s+t-1} (i.e. the sequence forms a contiguous interval of length $s + t$).

Finally, it remains to show that for any $n \in V$, there exists a valid path from the reference node to n using $+s$ and $-t$ hops. By Lemma 2, for any $n \in V$ there exists a sequence of intervals $I_1, \dots, I_K$ each of size $s + t$ and contained in V , with $n \in I_1$, consecutive intervals sharing a common node, and $0 \in I_K$. Since every node within each interval is mutually reachable via $+s$ and $-t$ steps, the shared nodes between consecutive intervals provide a valid path from n to 0. $\square$

A1.2 Optimal co-prime shift design

In this section, we provide the complete proof of our theorem for the optimal shift pairs. We first establish a lower bound on the maximum hop distance h^* that is necessary for full connectivity by any pair of shifts (see Lemma 3). This lemma counts the maximum number of nodes that can be reached at each hop distance. Note that each node has four possible neighbors with two shifts (i.e., separated by $\pm s$ and $\pm t$). By counting the number of nodes reachable within h hops, we obtain an upper bound on the number of distinct nodes that can be visited after h hops (or steps). Summing these counts determines how many hops are required to cover all N nodes. Since this count ignores the fact that multiple

paths may lead to the same node, the count gives us a lower bound on the number of hops required for full connectivity. In summary, the count represents the maximum possible number of reachable nodes and provides a universal lower bound on the required hop length.

Lemma 3 (Lower Bound on the Maximum Hop Distance). *Consider the graph V with N nodes defined in Eq. (1) and edges $E(\{s, t\})$ as in Eq. (2). Let h^* denote the minimum number of hops required to reach all nodes from a reference node. Then, for any choice of two shifts,*

$$h^* \geq \left\lceil \frac{-1 + \sqrt{2N-1}}{2} \right\rceil. \quad (6)$$

Proof. If we fix a hop budget h , any node reachable from the reference within h hops must be written as

$$x = \alpha s + \beta t, \quad |\alpha| + |\beta| \leq h, \quad (7)$$

which means we take α hops using shift s and β hops using shift t . Therefore, the set of nodes reachable within h hops is

$$\mathcal{S}_h = \{\alpha s + \beta t : |\alpha| + |\beta| \leq h\}.$$

The size of the set is

$$1 + 4 \sum_{i=1}^h i = 2h^2 + 2h + 1,$$

which is found by counting the integer points in the diamond $|\alpha| + |\beta| = i$, for $i = 1, \dots, h$. Since different pairs (α, β) may produce the same node n , the number of distinct nodes reachable within h hops satisfies $|\mathcal{S}_h| \leq 2h^2 + 2h + 1$. In order to reach all N nodes in the interval, it is therefore necessary that

$$2h^2 + 2h + 1 \geq N.$$

Solving this quadratic inequality gives

$$h \geq \frac{-1 + \sqrt{2N-1}}{2}.$$

Since h is an integer for any $N > 1$, we obtain the universal lower bound

$$h^* = \left\lceil \frac{-1 + \sqrt{2N-1}}{2} \right\rceil.$$

This result implies that, for any pair of two shifts s and t , we will not be able to connect all the nodes in V using less than h^* hops (or steps). $\square$

Next, we prove Theorem 2, which shows that the co-prime shifts $s^* = \lfloor \sqrt{N/2} \rfloor$ and $t^* = s^* + 1$ achieve this bound. We first show that using the lower bound h^* hops, the nodes visited by the optimal shifts form a continuous

interval. Finally, we prove that the endpoints of this interval exceed $\lfloor N/2 \rfloor$ in magnitude, which implies that the entire node set V lies within the reachable region. Note that this theorem does not claim that the provided optimal pair is unique. Experimentally, we observe that some other shift pairs with values close to $\sqrt{N/2}$ also achieve the same bound.

Theorem 2. *For a graph with N nodes as in Eq. (1) and edges defined by the consecutive **co-prime** pair $s, t \in \mathbb{N}$, $s = \lfloor \sqrt{\frac{N}{2}} \rfloor$ and $t = \lfloor \sqrt{\frac{N}{2}} \rfloor + 1$, achieves the minimum possible hop distance h^* given in Lemma 3.*

Proof. Let $k \in \mathbb{N}$ and define consecutive co-prime shifts as $s = k$ and $t = k + 1$. Since consecutive integers satisfy $\gcd(s, t) = 1$, this choice of s and t will give us a co-prime set and any integer reachable within hop budget h can be written as

$$\begin{aligned} n &= \alpha k + \beta(k + 1), & |\alpha| + |\beta| &\leq h \\ &= (\alpha + \beta)k + \beta. \end{aligned} \quad (8)$$

Let us define an anchor point $u = \alpha + \beta$ and offset v and rewrite Eq. (8) as

$$n = uk + v. \quad (9)$$

We can view the propagation as making u hops of k units and visiting v consecutive nodes. For a fixed budget h , the constraint on u and v can be written as

$$|u - v| + |v| \leq h.$$

By the triangle inequality, we have $|u| = |\alpha + \beta| \leq |\alpha| + |\beta| \leq h$. Similarly for a fixed u , the corresponding values of v must satisfy $u - h \leq 2v \leq u + h$, which can be written as

$$\left\lceil \frac{u - h}{2} \right\rceil \leq v \leq \left\lfloor \frac{u + h}{2} \right\rfloor.$$

Therefore, the integers reachable within h hops can be written as

$$n = uk + v, \quad |u| \leq h, \quad \left\lceil \frac{u - h}{2} \right\rceil \leq v \leq \left\lfloor \frac{u + h}{2} \right\rfloor. \quad (10)$$

For each fixed u , the range of possible values of v consists of consecutive integers between $\lceil \frac{u-h}{2} \rceil$ and $\lfloor \frac{u+h}{2} \rfloor$. Thus, we can define the interval of nodes visited by a given u and k as

$$I_{u,k} = \left[uk + \left\lceil \frac{u - h}{2} \right\rceil, uk + \left\lfloor \frac{u + h}{2} \right\rfloor \right].$$

We can find that the difference between the end of the interval $I_{u,k}$ and the start of the interval $I_{u+1,k}$ (*i.e.*, the gap between the intervals) is

$$\min(I_{u+1,k}) - \max(I_{u,k}) \leq k - h. \quad (11)$$

Therefore, if $h \geq k$, the intervals $I_{u,k}$ and $I_{u+1,k}$ are either overlapping or contiguous.

The interval $I_{u,k}$ will have a maximum value of $h(k+1)$ for $u = h$ and a minimum value of $-h(k+1)$ for $u = -h$. Therefore, following Eq. (11) if $k \leq h$, the union of intervals for $|u| \leq h$ is

$$\bigcup_{|u| \leq h} I_{u,k} = [-h(k+1), h(k+1)].$$

Let us now consider the choices: $k^* = \lfloor \sqrt{\frac{N}{2}} \rfloor$, $s = k^*$, $t = k^* + 1$, and $h = h^*$ from Lemma 3. First, notice that $h^* \geq k^*$ because

$$h^* \geq \frac{-1 + \sqrt{2N-1}}{2} \underset{(i)}{>} \frac{-1 + \sqrt{2N} - 1}{2} = \sqrt{\frac{N}{2}} - 1,$$

where (i) comes from the fact that $\sqrt{2N-1} > \sqrt{2N} - 1$. Finally, we have

$$h^* > \sqrt{\frac{N}{2}} - 1 \implies h^* \geq \left\lfloor \sqrt{\frac{N}{2}} \right\rfloor = k^*. \quad (12)$$

Second, we need to show that $h^*(k^* + 1) \geq \lfloor N/2 \rfloor$. The choice of $k = k^*$ can map to any possible values of N that are in the range $2(k^*)^2 \leq N < 2(k^* + 1)^2$. For these values, note that from Eq. (12) h^* can take the values of $h^* = k^*$ or $h^* = k^* + 1$. In particular, we have the two cases written as

$$h^* = \begin{cases} k^* & \text{if } 2(k^*)^2 \leq N \leq 2(k^*)^2 + 2k^* + 1 \\ k^* + 1 & \text{if } 2(k^*)^2 + 2k^* + 2 \leq N < 2(k^* + 1)^2. \end{cases}$$

In both cases, we can verify that

$$h^*(k^* + 1) \geq \left\lfloor \frac{N}{2} \right\rfloor. \quad (13)$$

We have shown that there are no gaps between the intervals since $h^* \geq k^*$, and that we can reach all N nodes of V . Consequently, the union of intervals for $|u| \leq h^*$ covers all nodes of V . In other words,

$$V = \left\{ -\left\lfloor \frac{N}{2} \right\rfloor, \dots, \left\lfloor \frac{N}{2} \right\rfloor - 1 \right\} \subset \bigcup_{|u| \leq h^*} I_{u,k^*}.$$

Therefore, the shifts $s = k^*$ and $t = k^* + 1$ achieve the minimum hop bound h^* and connect all nodes in V . $\square$

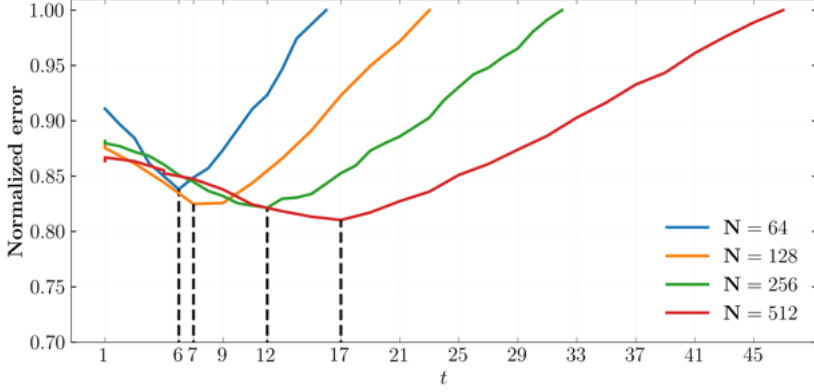


Fig. A2: Normalized reconstruction error as a function of the shift t for different sizes N . For each curve, the first co-prime pair is fixed to $s = k^* = \lfloor \sqrt{N}/2 \rfloor$ and we sweep through all possible choices of t . Across all tested sizes, the minimum error occurs near $t = s + 1$ (indicated using the dashed lines). This confirms the consecutive co-prime pairs $(s, t) = (k^*, k^* + 1)$ yields the optimal propagation structure and minimizes noise accumulation.

A2 Noise Analysis and Optimality

In this section, we provide additional empirical results that support the theoretical analysis. In particular, we investigate how the choice of shifts affects noise propagation and reconstruction error. The **hop distance** denotes the number of propagation steps required to reach a given node from the reference node.

Optimal shift selection. To validate Theorem 2, we fix the first shift to $s = k^*$ and sweep the second shift t across a range of values while measuring the reconstruction error. The results for several signal sizes are shown in Fig. A2. The minimum error consistently occurs near $t = s + 1$, supporting the theoretical result that the consecutive co-prime pair $(s, t) = (k^*, k^* + 1)$ achieves the lowest error.

For the number of nodes N shown in Fig. A2 we obtain the exact optimal pairs predicted by the theory for most cases: $(s, t) = (5, 6)$ for $N = 64$, $(11, 12)$ for $N = 256$, and $(16, 17)$ for $N = 512$, which follow $k^* = \lfloor \sqrt{N}/2 \rfloor$. For $N = 128$, the minimum occurs at $(7, 8)$, which is very close to the predicted pair $(8, 9)$. The recovery error observed in Fig. A2 is very close, as both shifts achieve similar hop lengths.

Noise reduction with path averaging. We can reduce noise accumulation by averaging the phasor estimates obtained from multiple propagation paths. When several paths reach the same node, their estimates can be combined which reduces estimator variance and improves robustness to measurement noise.

Figure A3 compares the cumulative phase error with and without path averaging for several shift configurations. In all multi-shift cases, averaging leads to

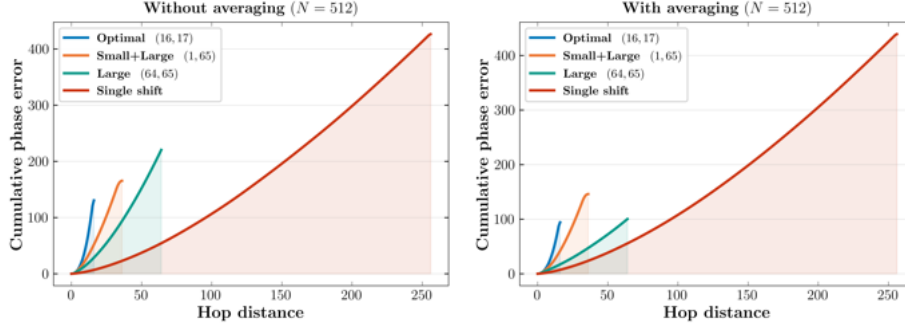


Fig. A3: Cumulative phase error as a function of hop distance with and without path averaging for several shift configurations at $N = 512$. Averaging reduces noise accumulation when multiple propagation paths exist because estimates from different paths can be combined.

a clear reduction in reconstruction error across hop distances. This behavior is expected since averaging combines multiple independent estimates of the phase difference and therefore suppresses noise. For the single-shift case there is no improvement since the phase at each node is obtained from only one propagation path.

A3 Additional Algorithm Implementation Details

Our main algorithm (Algorithm 1 of the main paper) propagates phase estimates using a breadth-first search (BFS) starting from a reference pixel. In this formulation, a queue stores pixels whose phase has already been estimated. Each time a pixel is dequeued, its neighbors defined by the shifts $\{\Delta_k\}$ are visited. The phases of these neighbors are computed by multiplying the current phasor estimate with the measured phase differences.

This queue-based implementation is inherently sequential and therefore inefficient on GPUs. To address this, we reformulate the propagation as a parallel wavefront update over the entire signal. Instead of processing one pixel at a time, we maintain a binary mask that identifies the currently active pixels whose phases have already been estimated. At each iteration, all active pixels simultaneously propagate their phase estimates to neighboring pixels defined by the shifts $\{\Delta_k\}$. The propagation is implemented using tensor slices that align source and destination pixels. For each shift s_k , we update

$$\hat{\mathbf{p}}_{i+s_k, j}^{(t+1)} += \hat{\mathbf{p}}_{i, j}^{(t)} \mathbf{p}_k^{(v)}(i + s_k, j), \quad (14)$$

for all active pixels (i, j) simultaneously. Despite propagating from all visited pixels, hop distances are assigned only on first visit. This ensures that each pixel records its shortest-path hop count. After propagation, pixels reached for the first time are assigned their hop distance, and phasors arriving with the same

Table A1: Optical hardware used in the real experiment setup.

Component	Model	Quantity	Notes
Laser diode (532 nm)	CPS532-C2	1	Collimated illumination source
Spatial Light Modulator (SLM)	Holoeye LUNA VIS 111	1	1920×1080 , $4.5 \mu\text{m}$ pitch
FLIR Camera	BFS-U3-23S3M-C	1	1920×1200 , $3.45 \mu\text{m}$ pitch
Non-polarizing Beam splitter	Thorlabs CCM1-BS013	1	Routes reference and modulated beams
Lens	Thorlabs AC254-100-A	2	$f = 100 \text{ mm}$, $\varnothing 25.4 \text{ mm}$
Diffuser	Thorlabs DG10-120-A	1	Glass diffuser
Aperture	ID12/M	1	Iris, 12 mm Max Aperture
Linear polarizers	-	2	Input/output polarization control

hop length are averaged. The algorithm updates the active pixels and repeats the same process until all nodes are visited. This parallel formulation preserves the behavior of the BFS propagation while allowing all updates to be computed efficiently using batched tensor operations.

A4 Additional Real Experiments and Hardware Details

We designed a hardware prototype to realize a self-referenced PSI model and to validate our method on real experiments. The optical components used in our setup are summarized in Tab. A1. We use a collimated 532 nm laser diode to illuminate the target object. We built a $4f$ system using two lenses each with focal length $f = 100 \text{ mm}$. A Holoeye LUNA phase-only SLM (1920×1080 , $4.5 \mu\text{m}$ pitch) is placed at the Fourier plane of the input. A non-polarizing beam splitter between the first lens and the SLM routes both the unmodulated and modulated reflected wavefronts through the second lens to a camera sensor (1920×1200 , $3.45 \mu\text{m}$ pitch) located one focal length away from the second lens. Linear polarizers placed before and after the beam splitter ensure the correct polarization state for phase modulation at the SLM. Phase ramps and global quadrature phase shifts displayed on the SLM produce calibrated spatially shifted interferometric measurements.

Additional refocusing experiments. In Fig. A4, we show the computational refocusing results by propagating the wavefront reconstructed from two sets of measurements: 16 measurements using shift pairs (24, 25), and 24 measurements using shifts (21, 24, 25). The wavefront reconstructed at $z = 0 \text{ mm}$ is numerically propagated from $z = 0 \text{ mm}$ to $z = 100 \text{ mm}$. We obtain a sharp focus around the correct target plane at $z = 75 \text{ mm}$ or $z = 67 \text{ mm}$ for three different scenes. The zoomed-in views highlight that fine structures and edges in the scene are clearly resolved at the correct focus plane.

Fresnel versus angular-spectrum propagation. To verify that the refocusing results are not substantially affected by the choice of numerical propagation model, we propagated the same recovered complex field using both Fresnel propagation and the angular spectrum method (ASM). Figure A5 shows a representative comparison at $z = 65 \text{ mm}$. The focused intensity images obtained with the two methods are visually similar.

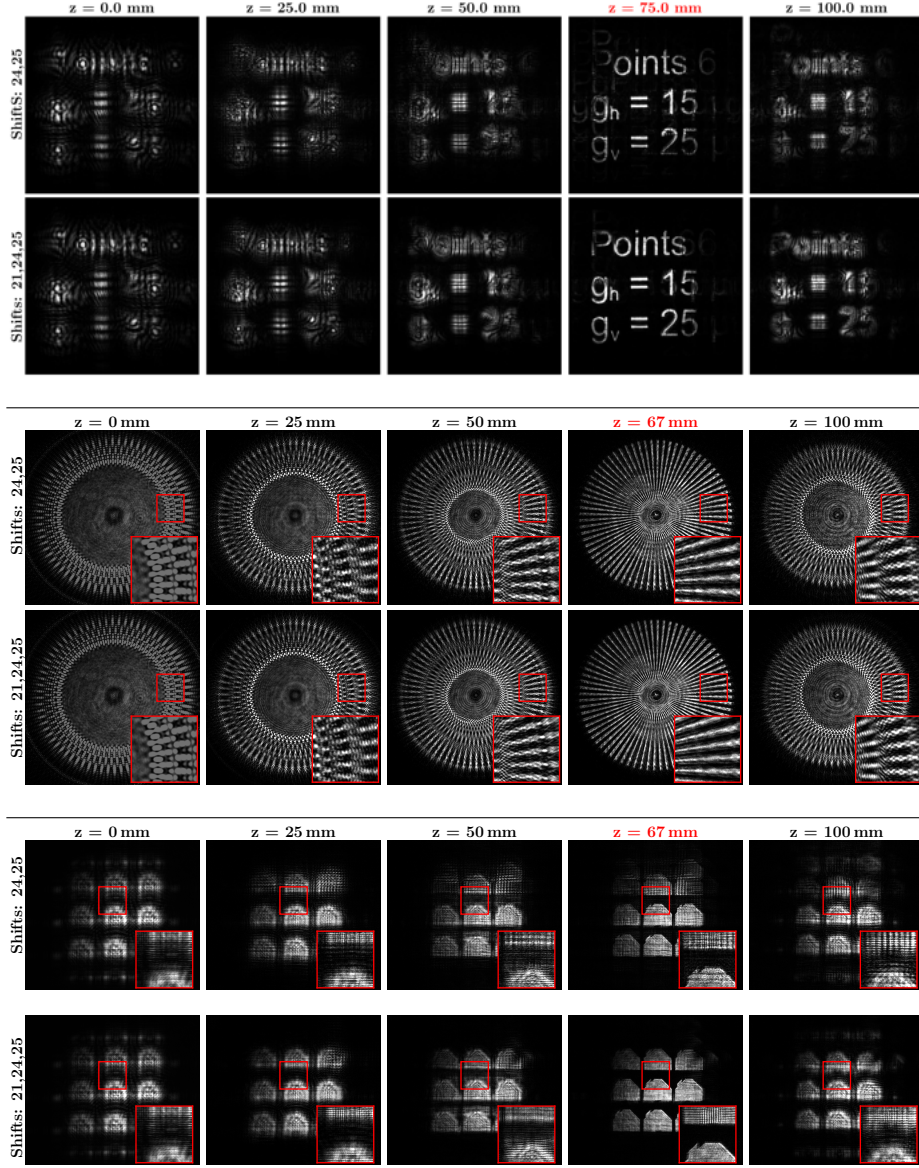


Fig. A4: Computational refocusing of the recovered wavefront for three different scenes using free-space propagation. The reconstructed complex field is propagated from $z = 0$ mm to $z = 100$ mm away from the object plane. A sharp focused image appears in the fourth column for each image.

Comparison experiments. We next present visual comparisons of refocusing results for different reconstruction methods in Fig. A9. Details about the comparison methods are presented in Sec. A5. We consider the same two measurement

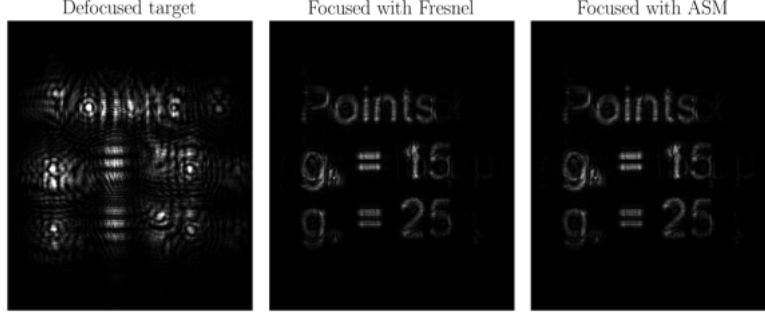


Fig. A5: Comparison of Fresnel and ASM propagation for computational refocusing.

Table A2: Lens phasor MAE. (Best, second best).

Configuration	Ours	DIP	GD-Rand	PnP-FISTA	WISH
$f = 75$ mm, $\#meas=16$	1.445	1.539	<u>1.527</u>	1.566	1.571
$f = 75$ mm, $\#meas=32$	1.456	1.495	<u>1.421</u>	1.551	1.378
$f = 100$ mm, $\#meas=16$	1.267	1.500	1.509	<u>1.493</u>	1.571
$f = 100$ mm, $\#meas=32$	1.271	<u>1.439</u>	1.465	1.453	1.460

settings: 16 measurements using shift pairs (24, 25), and 24 measurements using shifts (21, 24, 25). In all cases, the targets are refocused using the phases recovered by each method. Qualitatively, our method yields the best reconstructions across all targets. In addition, applying least-squares refinement further improves the sharpness and quality of the recovered structures.

Quantitative hardware metrics. We compare the recovered phase profiles of two biconvex lenses ($f = 75$ and 100 mm) against their ideal thin-lens quadratic profiles $\phi(x, y) = \frac{\pi}{\lambda f}(x^2 + y^2)$ with $\lambda = 532$ nm. Figures A6 and A7 show the recovered phase maps for the $f = 75$ and 100 mm lenses using 16 and 32 measurements. Table A2 reports the Phasor MAE across configurations, where our method achieves the best results in most settings.

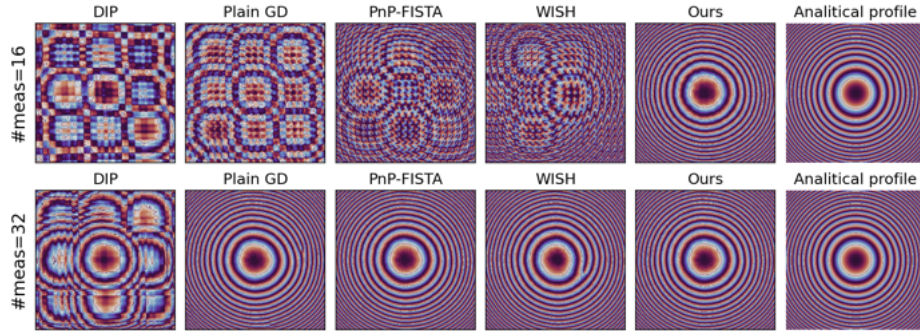


Fig. A6: Lens ($f = 75$ mm) phase recovery with 16, 32 measurements.

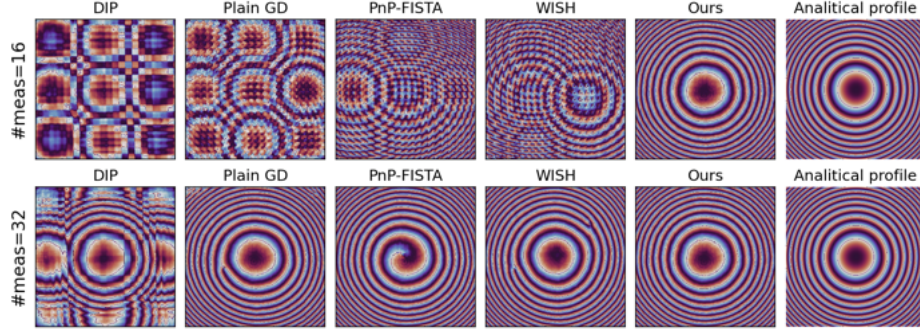


Fig. A7: Lens ($f = 100$ mm) phase recovery with 16, 32 measurements.

Lateral resolution analysis. We characterize the lateral resolution of a 3 mm Siemens star with 72 line-pairs (Fig. 8 in the paper), imaged with $1\times$ magnification. We sampled the intensity profiles at different radii and computed contrast for those line pairs. Figure A8 shows the mean and ± 1 std across the 9 sectors. The system reaches MTF25 at 0.329 lp/px (95 lp/mm). We note that while Fig. A8 analyzes the resolution after wavefront correction, the inherent lateral resolution of the system is limited by the sensor pitch, SLM pitch, and the physical width of the SLM.

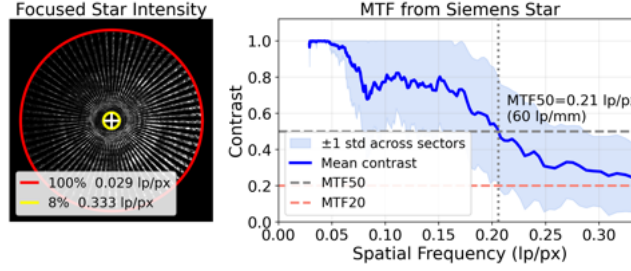


Fig. A8: Focused star target (left) and MTF plot (right).

A5 Comparison Methods and Experimental Details

We conducted comparison experiments against the following methods: gradient descent-based complex field recovery with random initialization (**GD-Rand**) and spectral initialization [5] (**GD-Spec**), a **PnP-FISTA** plug-and-play recovery method using a pretrained **DRUNet** denoiser [9] as a prior, **WISH** [8], and a Deep Image Prior (**DIP**)-based approach [7]. All experiments use multiple shifted measurements, where the field is interfered with spatially shifted copies along different directions and captured at multiple shift amounts. In all exper-

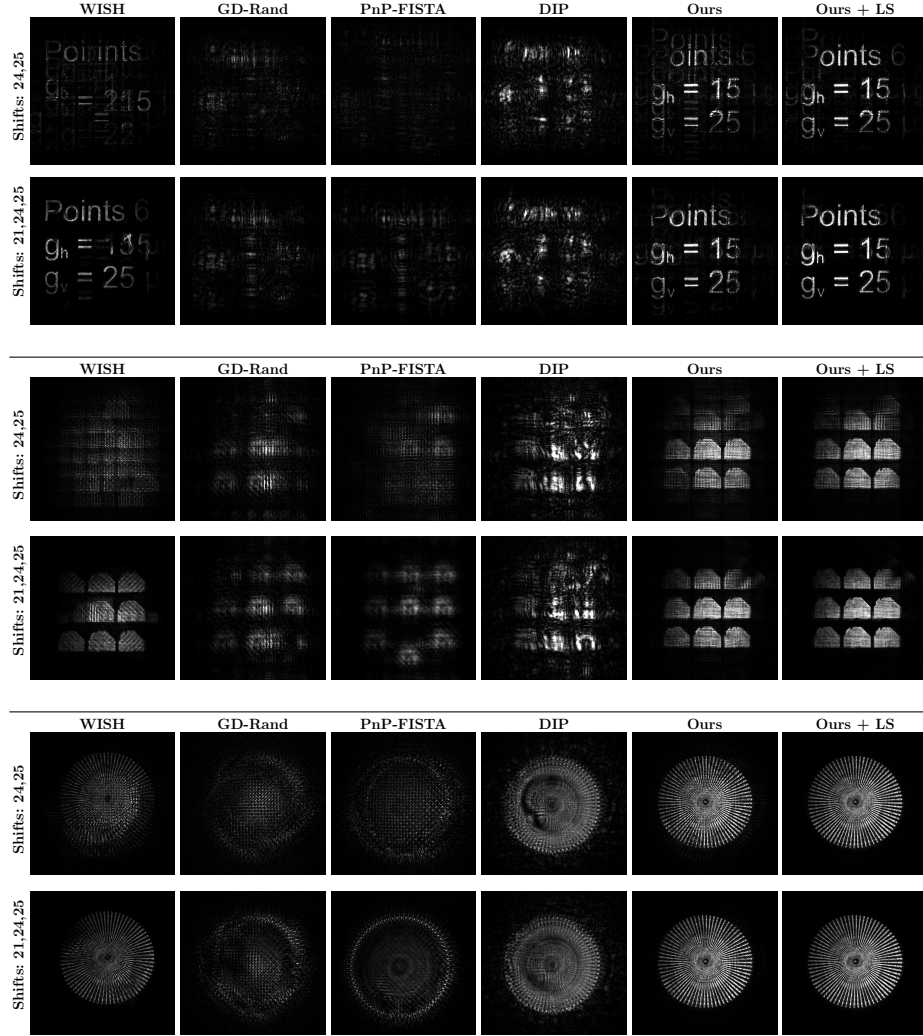


Fig. A9: Comparison of refocused reconstructions for three targets using different methods.

iments, the measured amplitude that is captured with no shifts is used during initialization.

- **Gradient-based methods.** We implement complex field recovery using gradient descent (GD) with stochastic gradient descent (SGD) in PyTorch, using a momentum of 0.9 and running for 1000 iterations. The learning rate is set to 5×10^{-3} for all experiments. Two initialization strategies are used: random phase initialization (**GD-Rand**) and spectral initialization (**GD-Spec**) [5], where the initial field is computed using the adjoint operator

from the DeepInverse library [6]. The loss minimized is the mean-squared error between predicted and measured intensities.

- **PnP-FISTA.** We use the plug-and-play (PnP) framework for phase recovery using the DeepInverse library [6]. The amplitude is initialized from the measured magnitude, and the phase is initialized randomly. A pretrained **DRUNet** denoiser [9] is used as the prior for the PnP framework. FISTA iterations run for 500 steps, with a stepsize of 8×10^{-2} , and the denoiser strength σ is set according to the measurement SNR.
- **WISH.** The complex field is recovered using the WISH algorithm [8]. Each shifted interference measurement is modeled via a two-step propagation: a Fourier transform to the Fourier plane, phase modulation to simulate the shift, and an inverse Fourier transform, corresponding to a $4f$ optical system. The field at the Fourier plane is initialized by back-propagating the measured amplitudes with a random phase for each shift. The recovery algorithm follows a Gerchberg–Saxton-style alternating minimization: the algorithm iteratively enforces amplitude constraints at the measurement plane and updates the Fourier-plane estimate using a least-squares aggregation across all measurements. The process is repeated for 1000 iterations.
- **DIP.** We use the Deep Image Prior (DIP) framework [7] to recover the phase of the complex field. The phase is parameterized using a convolutional decoder network [3] that maps a fixed random latent input of size 16×16 to the output phase, while the amplitude is fixed to the measured amplitude. The network is optimized using Adam [4] with an initial learning rate of 10^{-3} for 1000 iterations, minimizing the mean-squared error between the predicted and measured intensities over all shifted measurements.

A6 Additional Simulation Experiments

We report additional simulations to compare methods with different noise levels and numbers of measurements. To generate simulated data, we use 100 images from the DIV2K validation dataset [1] as ground truth intensity patterns. All images are center cropped to size 512×512 . For each image, we synthesize complex wavefronts by combining the intensity with three types of phase profiles, including random phase, quadratic phase, and smooth phase profiles³. We simulate shifted interferometric measurements with Poisson and Gaussian noise at multiple SNR levels and multiple shifts. We use the following shift configurations in the simulated experiments:

- 8 measurements: $\{\Delta_1\} = \{1\}$
- 16 measurements: $\{\Delta_1, \Delta_2\} = \{16, 17\}$
- 32 measurements: $\{\Delta_i\}_{i=1}^4 = \{16, 17, 23, 31\}$
- 64 measurements: $\{\Delta_i\}_{i=1}^8 = \{16, 17, 21, 33, 11, 22, 13, 63\}$

³ The MATLAB `peaks` function is used to generate smooth synthetic profiles.

Note that (16, 17) correspond to optimal shifts, whereas other shifts are selected at random.

Phase errors for quadratic phase profiles using 8, 16, 32, 64 measurements are reported in Tabs. A3 and A4, with visual reconstructions shown in Figs. A10 and A11. Across all measurement and noise settings our method performs the best. When using only 8 measurements, our method significantly outperforms the competing approaches, especially at 22 dB SNR. In addition, the least-squares refinement further improves the results in the 13 dB SNR setting.

A similar pattern is observed for peak phase profiles (Tabs. A5 and A6; Figs. A14 and A15). For random phase profiles (Tabs. A7 and A8; Figs. A12 and A13), gradient descent achieves the best performance, while our method with least-squares refinement provides the second best results.

Table A3: Phase errors for **Quadratic** phase recovery (mean $\pm$ std) at **22 dB SNR**. Our method achieves the best performance. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	1.536 $\pm$ 0.024	0.695 $\pm$ 0.365	0.470 $\pm$ 0.434	0.253 $\pm$ 0.385
GD-Spec	1.479 $\pm$ 0.063	0.737 $\pm$ 0.400	0.400 $\pm$ 0.432	0.358 $\pm$ 0.412
PnP-FISTA	1.505 $\pm$ 0.036	0.669 $\pm$ 0.341	0.456 $\pm$ 0.304	0.342 $\pm$ 0.225
WISH	1.524 $\pm$ 0.022	0.852 $\pm$ 0.299	0.423 $\pm$ 0.366	0.302 $\pm$ 0.401
DIP	1.541 $\pm$ 0.024	1.486 $\pm$ 0.031	1.491 $\pm$ 0.051	1.425 $\pm$ 0.071
Ours	<u>0.699 $\pm$ 0.247</u>	0.158 $\pm$ 0.104	0.099 $\pm$ 0.083	0.076 $\pm$ 0.079
Ours + LS	0.623 $\pm$ 0.277	<u>0.183 $\pm$ 0.082</u>	<u>0.126 $\pm$ 0.065</u>	<u>0.112 $\pm$ 0.071</u>

Table A4: Phase errors for **Quadratic** phase recovery (mean $\pm$ std) at **13 dB SNR**. Our method achieves the best performance and least-square refinements improve performance. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	1.539 $\pm$ 0.010	0.729 $\pm$ 0.361	0.471 $\pm$ 0.422	0.266 $\pm$ 0.369
GD-Spec	1.486 $\pm$ 0.051	0.859 $\pm$ 0.417	0.384 $\pm$ 0.285	0.370 $\pm$ 0.312
PnP-FISTA	1.522 $\pm$ 0.027	0.936 $\pm$ 0.304	0.518 $\pm$ 0.305	0.398 $\pm$ 0.267
WISH	1.536 $\pm$ 0.017	1.012 $\pm$ 0.276	0.500 $\pm$ 0.360	0.377 $\pm$ 0.405
DIP	1.541 $\pm$ 0.022	1.485 $\pm$ 0.030	1.485 $\pm$ 0.053	1.415 $\pm$ 0.082
Ours	<u>1.342 $\pm$ 0.161</u>	<u>0.577 $\pm$ 0.311</u>	<u>0.350 $\pm$ 0.240</u>	<u>0.254 $\pm$ 0.184</u>
Ours + LS	1.251 $\pm$ 0.225	0.445 $\pm$ 0.287	0.274 $\pm$ 0.194	0.200 $\pm$ 0.124

Table A5: Phase errors for **Peak** phase recovery (mean $\pm$ std) at **22 dB SNR**. Our method achieves the best results across all settings. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	1.478 $\pm$ 0.021	0.791 $\pm$ 0.361	0.570 $\pm$ 0.404	0.286 $\pm$ 0.363
GD-Spec	1.468 $\pm$ 0.075	0.686 $\pm$ 0.397	0.371 $\pm$ 0.374	0.209 $\pm$ 0.367
PnP-FISTA	1.506 $\pm$ 0.035	0.663 $\pm$ 0.341	0.479 $\pm$ 0.338	0.373 $\pm$ 0.296
WISH	1.517 $\pm$ 0.024	0.898 $\pm$ 0.330	0.425 $\pm$ 0.367	0.309 $\pm$ 0.391
DIP	1.471 $\pm$ 0.022	1.442 $\pm$ 0.063	1.423 $\pm$ 0.081	1.352 $\pm$ 0.112
Ours	<u>0.665 $\pm$ 0.238</u>	0.155 $\pm$ 0.096	0.097 $\pm$ 0.071	0.071 $\pm$ 0.063
Ours + LS	0.559 $\pm$ 0.247	<u>0.205 $\pm$ 0.063</u>	<u>0.159 $\pm$ 0.050</u>	<u>0.124 $\pm$ 0.050</u>

Table A6: Phase errors for **Peak** phase recovery (mean $\pm$ std) at **13 dB SNR**. Our method achieves the best results across all settings. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	1.485 $\pm$ 0.012	0.830 $\pm$ 0.338	0.593 $\pm$ 0.401	0.314 $\pm$ 0.358
GD-Spec	1.482 $\pm$ 0.054	0.800 $\pm$ 0.381	0.415 $\pm$ 0.361	<u>0.216 $\pm$ 0.195</u>
PnP-FISTA	1.525 $\pm$ 0.022	0.948 $\pm$ 0.316	0.535 $\pm$ 0.327	<u>0.377 $\pm$ 0.279</u>
WISH	1.534 $\pm$ 0.018	1.019 $\pm$ 0.278	0.552 $\pm$ 0.375	0.408 $\pm$ 0.404
DIP	1.470 $\pm$ 0.022	1.452 $\pm$ 0.054	1.431 $\pm$ 0.074	<u>1.360 $\pm$ 0.103</u>
Ours	<u>1.307 $\pm$ 0.194</u>	<u>0.532 $\pm$ 0.277</u>	<u>0.342 $\pm$ 0.227</u>	0.255 $\pm$ 0.200
Ours + LS	1.181 $\pm$ 0.276	0.394 $\pm$ 0.234	0.273 $\pm$ 0.175	0.206 $\pm$ 0.152

Table A7: Phase errors for **Random** phase recovery (mean $\pm$ std) at **22 dB SNR**. Our method consistently achieves the best or second-best results. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	<u>0.505 $\pm$ 0.142</u>	0.138 $\pm$ 0.148	0.079 $\pm$ 0.152	0.043 $\pm$ 0.126
GD-Spec	1.481 $\pm$ 0.064	0.691 $\pm$ 0.367	0.449 $\pm$ 0.440	0.148 $\pm$ 0.203
PnP-FISTA	1.511 $\pm$ 0.030	0.735 $\pm$ 0.329	0.425 $\pm$ 0.264	0.364 $\pm$ 0.233
WISH	1.519 $\pm$ 0.027	0.854 $\pm$ 0.320	0.453 $\pm$ 0.377	0.285 $\pm$ 0.377
DIP	1.423 $\pm$ 0.063	1.495 $\pm$ 0.037	1.447 $\pm$ 0.065	1.384 $\pm$ 0.104
Ours	0.659 $\pm$ 0.236	<u>0.156 $\pm$ 0.099</u>	<u>0.094 $\pm$ 0.075</u>	<u>0.073 $\pm$ 0.066</u>
Ours + LS	0.492 $\pm$ 0.273	0.158 $\pm$ 0.075	0.121 $\pm$ 0.055	0.107 $\pm$ 0.054

Table A8: Phase errors for **Random** phase recovery (mean $\pm$ std) at **13 dB SNR**. Our method consistently achieves the best or second-best results. (Best bold, second-best underlined).

Method \ #Meas.	8	16	32	64
GD-Rand	0.642 $\pm$ 0.210	0.215 $\pm$ 0.205	0.126 $\pm$ 0.159	0.092 $\pm$ 0.173
GD-Spec	1.485 $\pm$ 0.055	0.842 $\pm$ 0.383	0.495 $\pm$ 0.377	0.360 $\pm$ 0.331
PnP-FISTA	1.522 $\pm$ 0.027	0.951 $\pm$ 0.340	0.595 $\pm$ 0.339	0.390 $\pm$ 0.275
WISH	1.533 $\pm$ 0.021	0.995 $\pm$ 0.277	0.534 $\pm$ 0.362	0.370 $\pm$ 0.385
DIP	1.426 $\pm$ 0.061	1.499 $\pm$ 0.036	1.446 $\pm$ 0.068	1.368 $\pm$ 0.086
Ours	1.283 $\pm$ 0.173	0.555 $\pm$ 0.299	0.360 $\pm$ 0.246	0.254 $\pm$ 0.176
Ours + LS	<u>1.173 $\pm$ 0.270</u>	<u>0.407 $\pm$ 0.266</u>	<u>0.283 $\pm$ 0.208</u>	<u>0.195 $\pm$ 0.125</u>

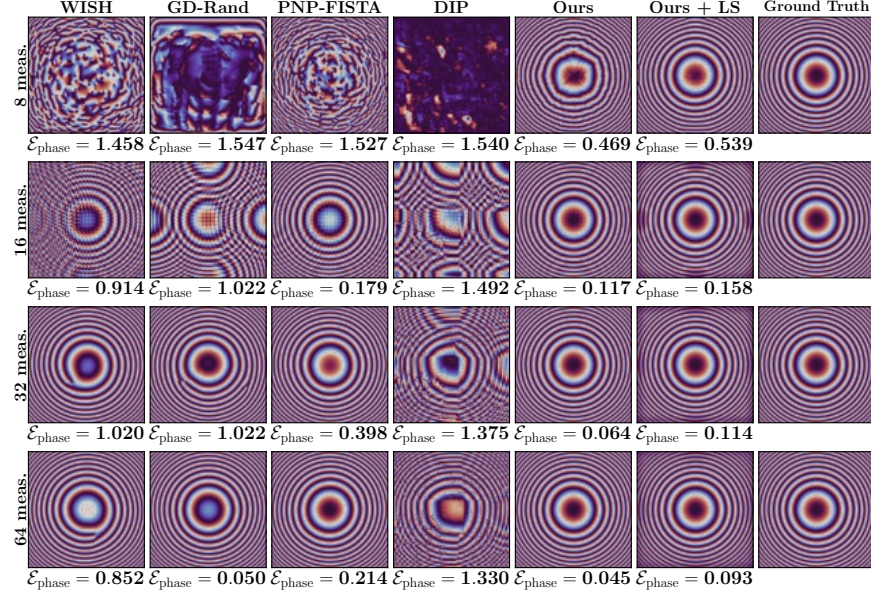


Fig. A10: Quadratic phase recovery at different number of measurements at 22dB SNR. Our method consistently yields the best performance both quantitatively and qualitatively.

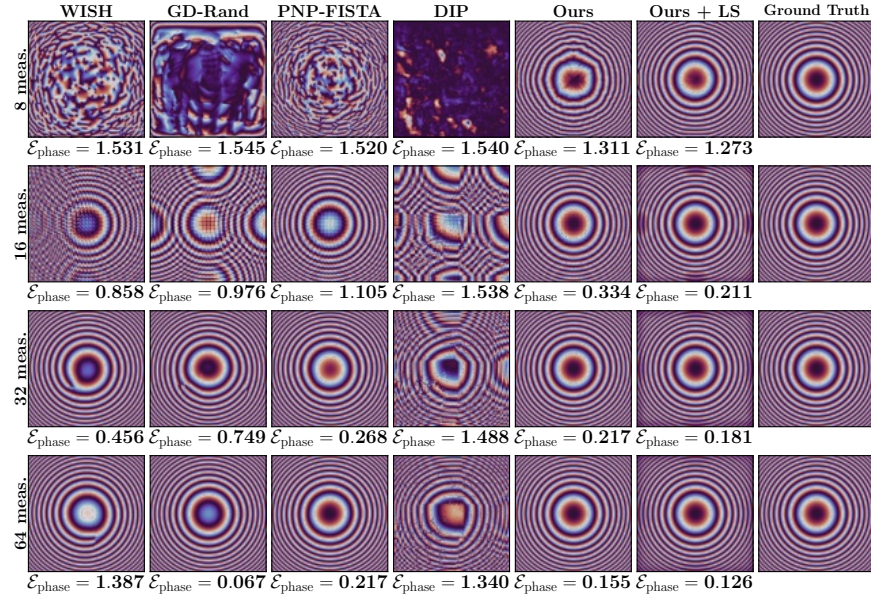


Fig. A11: Quadratic phase recovery at 13 dB SNR. Our method is the only approach that reliably reconstructs the phase profile using only 8 measurements.

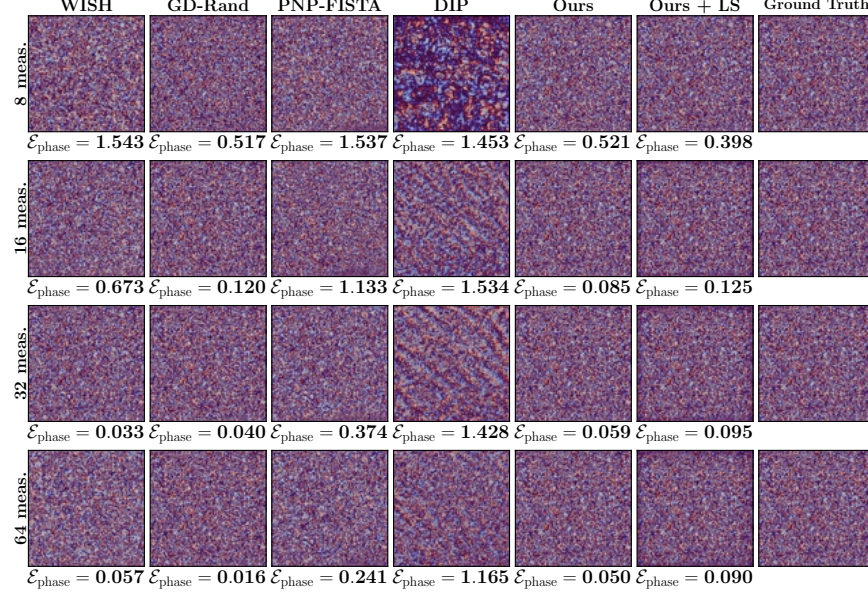


Fig. A12: Random phase recovery at 22 dB SNR with different numbers of measurements. Our method with least-squares refinement achieves the second-best performance across most measurement settings.

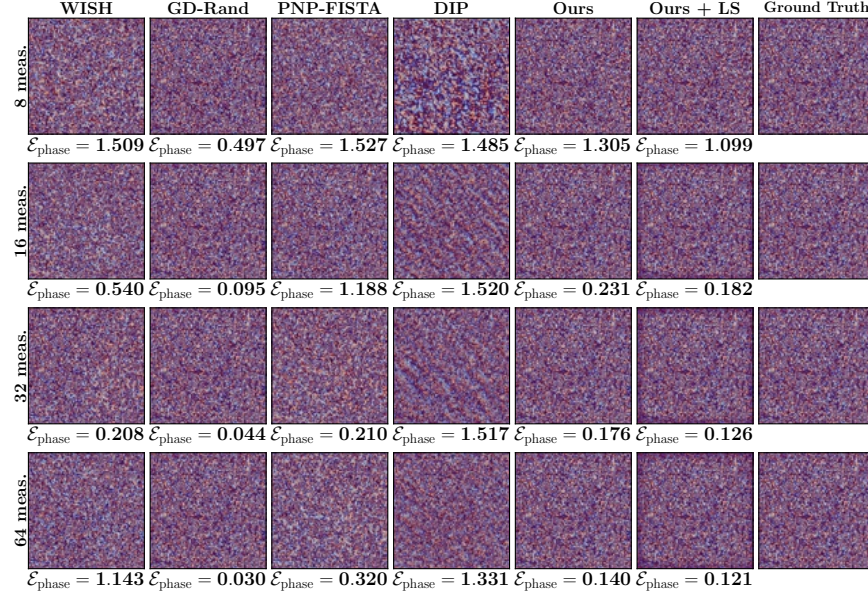


Fig. A13: Random phase recovery at 13 dB SNR with different numbers of measurements. Our method with least-squares refinement achieves the second-best performance across most measurement settings.

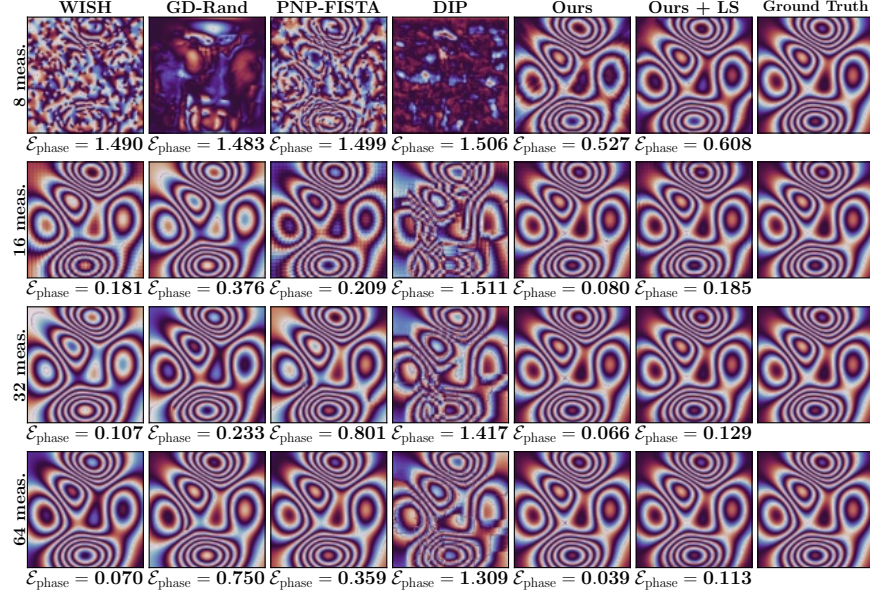


Fig. A14: Peak phase recovery at 22 dB SNR with different numbers of measurements. Our method achieved the best performance across all settings.

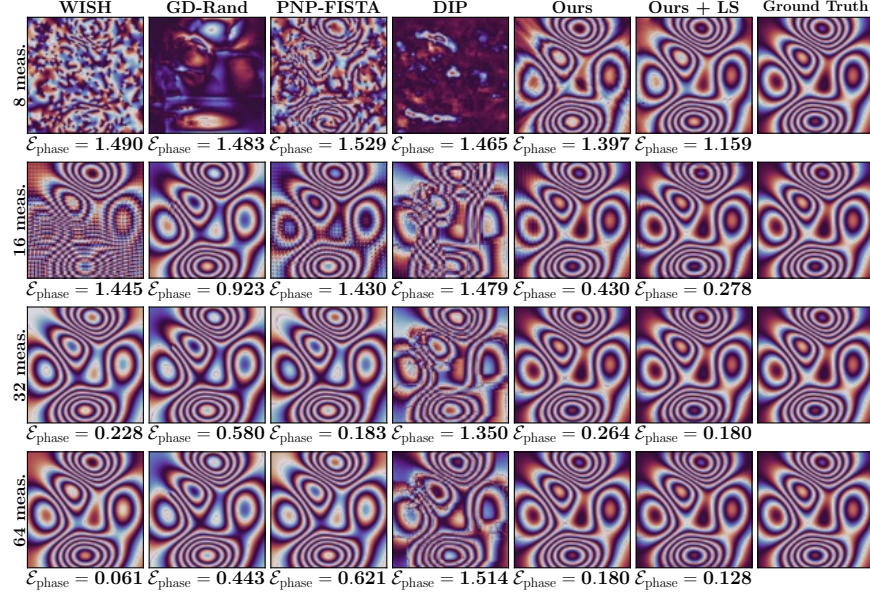


Fig. A15: Peak phase recovery at 13 dB SNR with different numbers of measurements. Our method achieves the best performance in most settings and is the only method that produces a consistent phase profile with very few measurements.

References

1. Agustsson, E., Timofte, R.: NTIRE 2017 challenge on single image super-resolution: Dataset and study. In: IEEE Conf. Comput. Vis. Pattern Recog. Worksh. (2017)
2. Bézout, E.: Théorie générale des équations algébriques. Ph.-D. Pierres (1779)
3. Darestani, M.Z., Heckel, R.: Accelerated MRI with un-trained neural networks. IEEE Transactions on Computational Imaging **7**, 724–733 (2021)
4. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
5. Netrapalli, P., Jain, P., Sanghavi, S.: Phase retrieval using alternating minimization. In: Adv. Neural Inform. Process. Syst. (2013)
6. Tachella, J., Terris, M., Hurault, S., Wang, A., Davy, L., Scanvic, J., Sechaud, V., Vo, R., Moreau, T., Davies, T., Chen, D., Laurent, N., Monroy, B., Dong, J., Hu, Z., Nguyen, M.H., Sarron, F., Weiss, P., Escande, P., Massias, M., Modrzyk, T., Levac, B., Liaudat, T.I., Song, M., Hertrich, J., Neumayer, S., Schramm, G.: DeepInverse: A Python package for solving imaging inverse problems with deep learning. Journal of Open Source Software **10**(115), 8923 (2025)
7. Ulyanov, D., Vedaldi, A., Lempitsky, V.: Deep image prior. In: IEEE Conf. Comput. Vis. Pattern Recog. (2018)
8. Wu, Y., Sharma, M.K., Veeraraghavan, A.: WISH: Wavefront imaging sensor with high resolution. Light: Science & Applications **8**(1), 44 (2019)
9. Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R.: Plug-and-play image restoration with deep denoiser prior. IEEE Trans. Pattern Anal. Mach. Intell. **44**(10), 6360–6376 (2021)