Supplementary Information for

Disturbance-Introduced Interferometry: Surface Topography Metrology Beyond Vibration Isolation and Phase-Shifting Control

Lin Li[1,2,†], Rui Xu[3,4,†], Yawen He[1,2], Lihua Lei[5,6], Xinzhe Li[1,2], Lintong Du[1,2], Yue Pan[1,2], Yumai Ji[1,2], Yang Shen[1,2], Shuxin Liu[1,2,7], Sen Han[8,9] and Jiamiao Yang[1,2,7,*]

[1] School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, 200240, China

[2] State Key Laboratory of Submarine Geoscience, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, 200240 Shanghai, China

[3] State Key Laboratory of Dynamic Optical Imaging and Measurement, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

[4] University of Chinese Academy of Sciences, Beijing 100049, China

[5] Shanghai Institute of Measurement and testing Technology Co., Ltd., Shanghai 201203, China

[6] Shanghai Key Laboratory of Online Test and Control Technology, Shanghai 201203, China

[7] Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai 200240, China

[8] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

[9] Suzhou H&L Instruments LLC, Suzhou 215123, China.

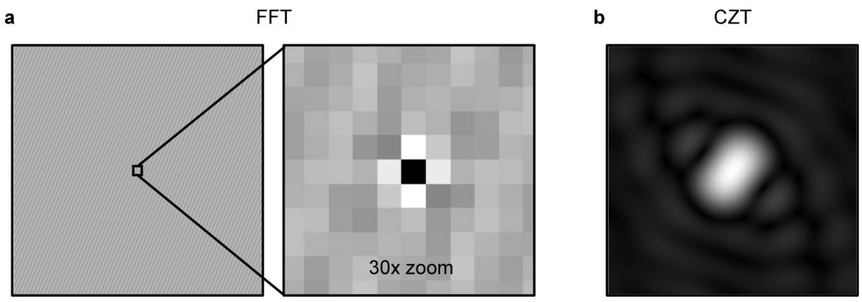
* jiamiaoyang@sjtu.edu.cn

† These authors contributed equally to this work.

Supplementary Note 1: High-resolution parameter estimation based on Chirp-Z transform

In contrast to standard Discrete Fourier Transform (DFT), which imposes uniform resolution across the spectrum, Chirp-Z Transform (CZT) samples spirally in the Z-plane, enabling super-resolution analysis within targeted frequency bands. For a one-dimensional sequence $x[n]\left(n=0,1,\ldots,N-1\right)$, the CZT is expressed as $X_{\omega}=\sum_{n=0}^{N-1}x[n]\cdot A^{-n}\cdot W^{nk}$,where $A=A_0e^{j\theta_0}$ denotes the initial sampling point, $W=W_0e^{j\phi_0}$ is the spiral step factor, and ω indexes the output spectral components. For two-dimensional signals $x[m,n]\left(m=0,1,\ldots,M-1,n=0,1,\ldots,N-1\right)$, CZT is computed sequentially: $X[\omega_x,\omega_y]=\mathrm{CZT}_x\{\mathrm{CZT}_y\{x[m,n]\}\}$.

By tuning A and W, CZT flexibly adjusts the sampling range and density in the frequency domain, resolving closely spaced spectral peaks that standard DFT cannot distinguish, while minimizing computational overhead. In the coaxial interferometric measurements described here (Fig. S1), the spectral peaks of $\cos\left(\psi_n\right)$ cluster near zero frequency. CZT achieves sub-pixel resolution estimation of $\left(k_{xn},k_{yn}\right)$ without requiring additional spatial sampling points, enhancing parameter resolution. Efficient computation is enabled by the Bluestein algorithmp[1], with a complexity of $O((M+K)(N+L)\log(M+K+N+L))$.

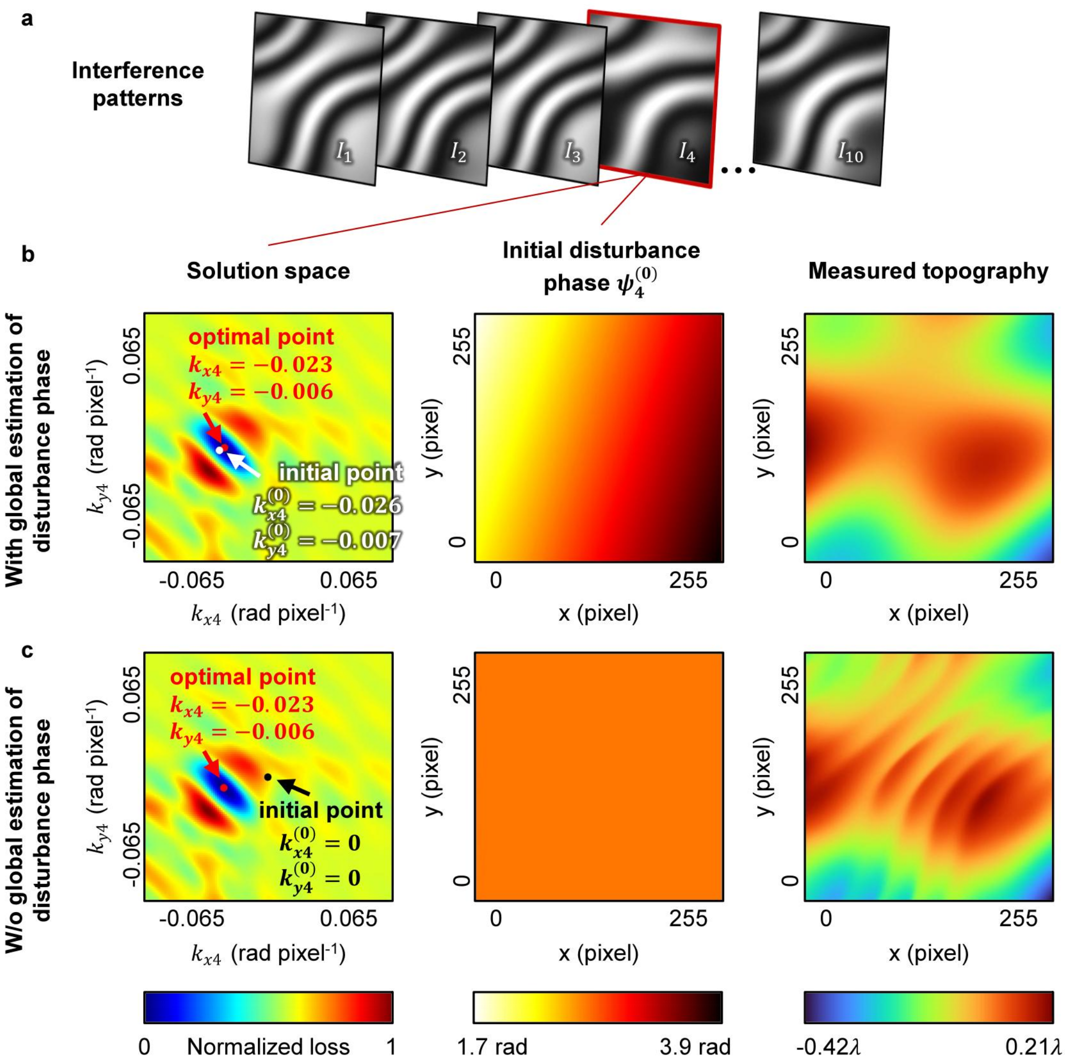


Supplementary Figure S1. Comparison of FFT (a) and CZT (b).

Supplementary Note 2: Role of global coarse estimation of disturbance phases

Random mechanical disturbances (RMD) introduce numerous unknowns in phase-shifting interferometry (PSI), requiring both traditional anti-vibration algorithms and Natural Phase Decoding Algorithm (NPDA) developed here to initiate iterative solutions from suitable parameter estimates. Traditional PSI algorithms effectively estimate linear disturbance parameters $d_n^{(0)}$, but struggle with nonlinear parameters k_{xn}, k_{yn}. Typically, these methods set initial values $k_{xn}^{(0)}, k_{yn}^{(0)} = 0$, which, under significant disturbances, leads to convergence at local optima and erroneous outcomes as true values diverge from zero.

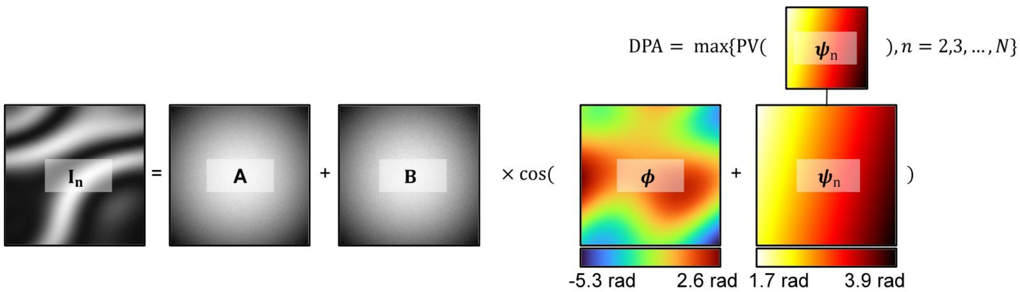
NPDA mitigates this by performing a coarse global estimation of the disturbance phase. For surface profile reconstruction from a 10-frame 256×256 interferogram sequence (Fig. S2a), NPDA estimates k_{xn}, k_{yn}. Fig. S2b demonstrates the estimation of the fourth frame's parameters $k_{x4}^{(0)}, k_{y4}^{(0)}$ (left) and the corresponding reconstructed disturbance phase $\psi^{(0)} = k_{xn}^{(0)} x + k_{yn}^{(0)} y + d_n^{(0)}$ (middle) yields values near the global optimum, ensuring accurate subsequent reconstruction (right). Without this step (Fig. S2c), initializing $k_{xn}^{(0)}, k_{yn}^{(0)} = 0$ as in traditional PSI introduces substantial errors in the final result.



Supplementary Figure S2. Effects of global disturbance phase estimation. For the interferogram sequence **(a)**, estimation **(b)** initiates optimization near the global optimum (left, middle), yielding accurate reconstruction (right). Without estimation **(c)**, starting from zero (left, middle) results in poor reconstruction (right).

Supplementary Note 3: Quantification of nonlinear disturbance level

Vibration tolerance in laser interferometric surface topography measurements is often assessed using Vibration Criterion (VC) curves, which characterize structural vibration levels. While effective for equipment evaluation, this metric is less suited for methodological comparisons. In such measurements, the phase difference between test and reference beams responds anisotropically to six-degree-of-freedom pose shifts induced by random mechanical disturbances (RMD). With linear disturbance phases well-managed, nonlinear components demand greater scrutiny. We define the Disturbance Phase Amplitude (DPA) as the maximum peak-to-valley (PV) value of the nonlinear disturbance phase within the measurement cavity under RMD (Fig. S3), providing a targeted metric for nonlinear disturbance levels.

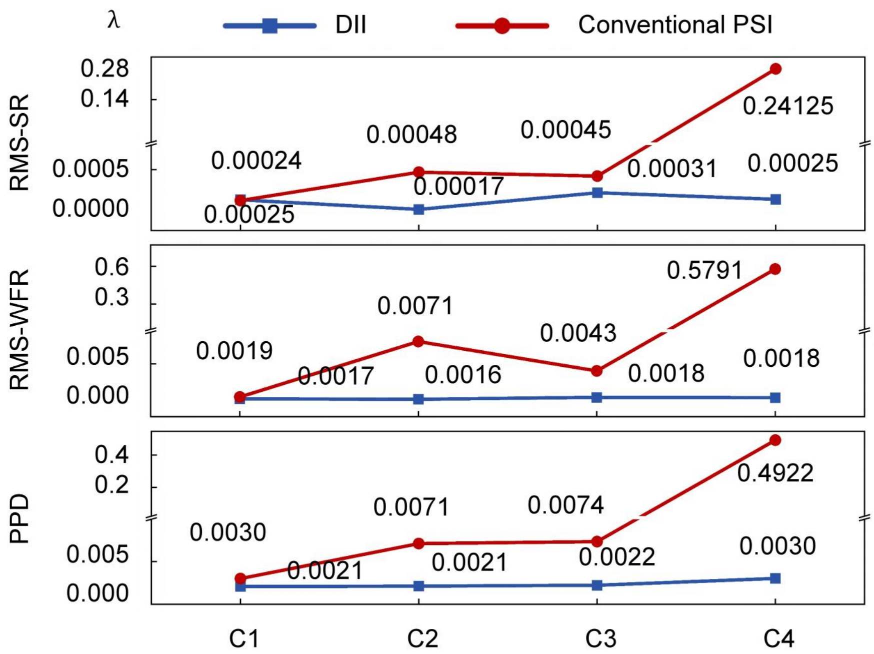


Supplementary Figure S3. Diagram illustrating the DPA concept for quantifying nonlinear disturbances in laser interferometric measurements.

Supplementary Note 4: Measurement repeatability calibration results

To assess measurement stability and compare the proposed method (DII) with traditional PSI under varying disturbance conditions (C1–C4, see Methods), we employed three metrics: RMS Simple Repeatability (RMS-SR), RMS Wavefront Repeatability (RMS-WFR), and Peak Pixel Deviation (PPD). RMS-SR, gauging macroscopic stability, is twice the standard deviation of RMS values from 50 consecutive measurements. RMS-WFR, assessing wavefront stability, is the mean RMSE of even-numbered measurements against the odd-numbered average, plus twice the standard deviation, over 50 runs. PPD, reflecting pixel-level stability, is the 99th percentile of pixel-wise standard deviation across 50 measurements.

As shown in Fig. S4, under condition C1, SI achieves RMS-SR of 0.00025λ, RMS-WFR of 0.0017λ, and PPD of 0.0021λ (99th percentile). Across RMD conditions (C2–C4), averages are RMS-SR 0.00025λ, RMS-WFR 0.00176λ, and PPD 0.00246λ (99th percentile), demonstrating consistency with controlled settings and exceptional precision. Conversely, traditional PSI errors escalate with increasing DPA, reaching RMS-SR 0.24125λ, RMS-WFR 0.5791λ, and PPD 0.4922λ (99th percentile) under C4.



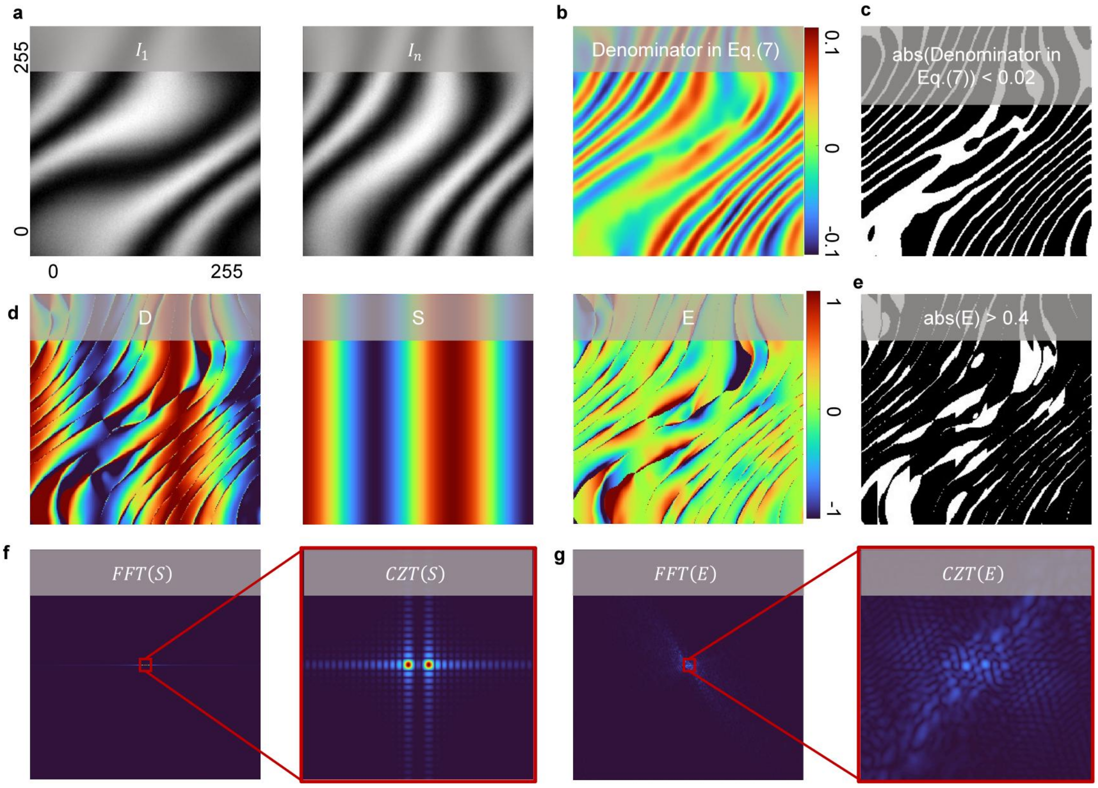
Supplementary Figure S4. Measurement repeatability calibration results.

Supplementary Note 5: Analysis of ill-conditioned scenarios in Eq. (7)

In instances where the denominator in Eq. (7) approaches zero for certain pixels, the calculation becomes ill-conditioned. We conducted some simulations to analyze the impact of it.

As illustrated in Fig. S5, we simulated two interferograms with a relative tilt of 2π rad (Fig. S5a). The denominator calculated using Eq. (7) contains regions where its value is close to zero (Fig. S5b). We decomposed the output of Eq. (7) into an ideal signal (S) and an error component (E). A strong spatial correlation is evident between the pixels with a near-zero denominator (where abs(Denominator) < 0.02, as shown in Fig. S5c) and the pixels with high error (where abs(E) > 0.4, as shown in Fig. S5e). This confirms that these ill-conditioned points are a source of significant error.

However, the crucial insight comes from the frequency domain analysis. The frequency spectrum of the error E (Fig. S5f) shows that this noise is scattered across high-frequency components. In stark contrast, the spectrum of the ideal signal S (Fig. S5g) is highly concentrated in the low-frequency region. Therefore, when we apply spectral peak analysis (via Chirp Z-Transform in Eqs. (8)-(9)) in the subsequent step, the algorithm naturally isolates the concentrated, low-frequency signal from the scattered, high-frequency noise. This ensures that the ill-conditioned points have a negligible impact on the accuracy of the estimated disturbance parameters.



Supplementary Figure S5. Analysis of signal and noise in the presence of ill-conditioned points in Eq. (7). (a) Two simulated interferograms with a relative tilt of 2π rad. **(b)** The corresponding denominator map calculated using Eq. (7). **(c)** A binary mask highlighting regions where the absolute value of the denominator is less than 0.02. **(d)** The result of Eq. (7), denoted as D, is decomposed into an ideal signal S and an error component E. **(e)** A binary mask showing regions where the absolute value of the error E is greater than 0.4. A strong spatial correlation is observed between the ill-conditioned points in (c) and the high-error regions in (e). **(f, g)** Frequency spectra of the signal S and the error E, respectively. The spectra reveal that the noise introduced by these ill-conditioned points manifests primarily as high-frequency, scattered components, while the true disturbance signal is concentrated in the low-frequency region.

Supplementary Note 6: Robustness to Intensity Level Noise

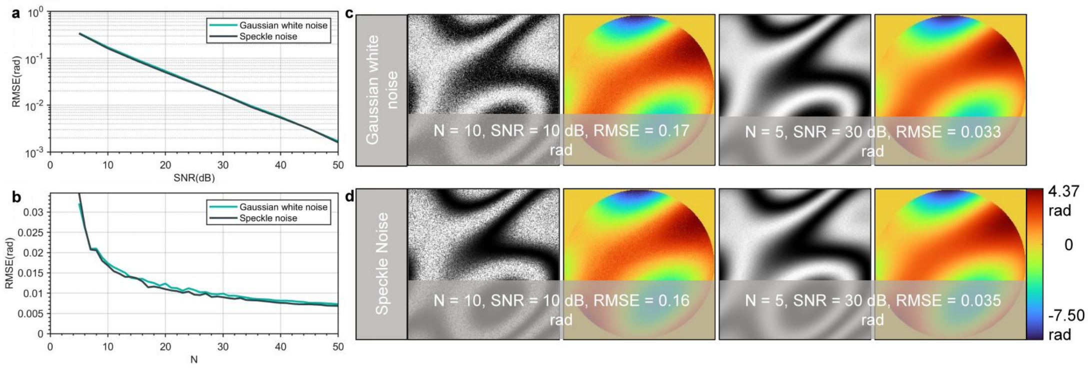
Intensity level noise, originating from sources like speckle and sensors, degrades signal quality by changing to the pixel intensity values and thus lowering the Signal-to-Noise Ratio (SNR). The robustness against noise is critical. NPDA addresses this issue by leveraging the random and zero-mean characteristics of noise through its intrinsic modeling and processing mechanisms.

In the Coarse Estimation stage, an algebraic operation, as defined in Eq. (7), is employed to isolate the signal into a single frequency component. Subsequently, Eq. (8) separates this signal from the spectrally-dispersed noise, thereby mitigating its influence (see Supplementary Note 5 for details).

During the Fine Refinement stage, the algorithm minimizes the sum of squared residuals between the theoretical and the measured interferogram intensities via a least-squares approach. This optimization effectively smooths and rejects the intensity noise by treating it as part of the residual error. Moreover, the algorithm's pointwise update scheme leverages the entire set of captured frames, creating an averaging effect that significantly diminishes the influence of random intensity noise on the final reconstructed phase.

To quantitatively evaluate the performance of NPDA under noisy conditions, we conducted a series of numerical simulations. We tested the algorithm against both additive white Gaussian noise and multiplicative speckle noise. The speckle noise was simulated using the model $I_n(\vec{p}) = I(\vec{p}) + v(\vec{p})I(\vec{p})$, where v is a zero-mean, uniformly distributed random variable. The noise level was quantified by the SNR of the generated images.

The simulation results are presented in Fig. S6. Fig. S6a plots the root-mean-square error (RMSE) of the reconstructed phase as a function of image SNR, which ranged from 5 dB to 50 dB, using a fixed set of 10 interferograms. For both Gaussian and speckle noise, the logarithm of the phase RMSE exhibits a clear inverse relationship with the image SNR. Fig. S6b illustrates the relationship between the phase RMSE and the number of interferograms (N), varying from 5 to 50, at a constant SNR of 30 dB. The result shows that the reconstruction error is approximately inversely proportional to the number of frames used. Fig. S6c and Fig. S6d provide qualitative examples of the input interferograms and the corresponding reconstruction results under challenging conditions with significant Gaussian and speckle noise, respectively. Collectively, these results demonstrate the exceptional robustness of our method. Even under severe noise conditions, such as an SNR of only 10 dB or with as few as 5 interferograms, NPDA can accurately retrieve the underlying phase. This confirms the algorithm's high resilience to noise and its suitability for real-world measurement scenarios.



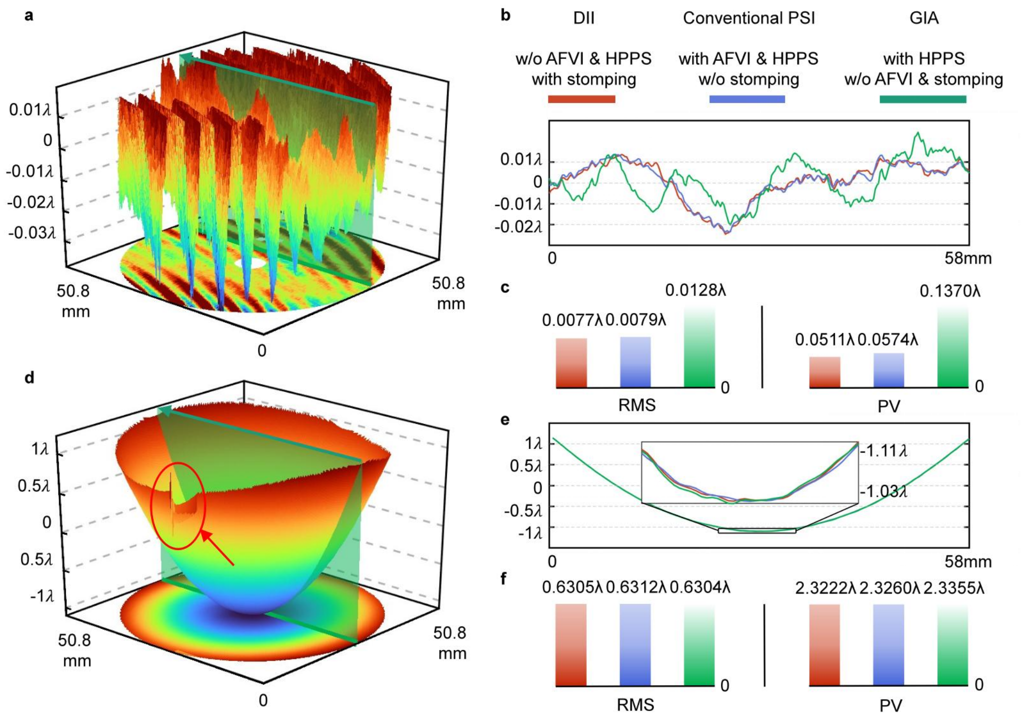
Supplementary Figure S6. Robustness and performance of the NPDA method in numerical simulations. The plots show the phase reconstruction RMSE as a function of **(a)** the image SNR with a fixed number of frames (N=10) and **(b)** the number of frames (N) with a fixed SNR of 30 dB. Both additive Gaussian and multiplicative speckle noise scenarios were tested. Panels **(c)** and **(d)** show representative examples of the method's performance on a single interferogram corrupted by significant Gaussian and speckle noise, respectively, demonstrating successful phase retrieval even under severe noise conditions.

Supplementary Note 7: Additional comparison using a conventional PZT-based configuration without AFVI

To further complement the comparison in Fig. 6, we additionally considered a conventional PZT-based configuration without an air-floating vibration isolation platform (AFVI). Unlike the DII experiment in Fig. 6b, no intentional floor stomping was introduced in this additional test; the purpose here was to isolate the effect of removing AFVI while retaining conventional phase shifting. The corresponding results are shown in Fig. S7.

For the spherical sample, the reconstruction breaks down completely, as shown in Fig. S7a. Its RMS and PV reach 0.0128λ and 0.1370λ, respectively (Fig. S7c), which are clearly inconsistent with the results obtained by DII and by conventional PSI under controlled conditions, whose RMS values are 0.0077-0.0079λ and PV values are 0.0511-0.0554λ. This discrepancy is also clearly visible in the profile comparison in Fig. S7b. For the planar sample, the reconstructed surface does not collapse globally, but still exhibits a clear local reconstruction error, as shown in Fig. S7d, where the abnormal region is highlighted by the red circle. The corresponding profile comparison in Fig. S7e and the RMS/PV comparison in Fig. S7f further confirm that this conventional configuration remains less reliable than both DII and conventional PSI under controlled conditions.

These results further support that, although a conventional phase shifter can provide controlled phase steps, the lack of effective vibration isolation still degrades the measurement reliability of traditional approaches even without intentional stomping disturbance.



Supplementary Figure S7. Additional comparison using a conventional PZT-based configuration without AFVI. No intentional floor stomping was introduced in this test. (a) Reconstructed spherical surface obtained using the conventional PZT-based configuration without AFVI. **(b)** Profile comparison for the spherical sample among DII, conventional PSI under controlled conditions, and the conventional PZT-based configuration without AFVI. **(c)** Corresponding RMS and PV comparison for the spherical sample. **(d)** Reconstructed planar surface obtained using the conventional PZT-based configuration without AFVI, showing a clear local reconstruction error highlighted by the red circle. **(e)** Profile comparison for the planar sample among DII, conventional PSI under controlled conditions, and the conventional PZT-based configuration without AFVI. **(f)** Corresponding RMS and PV comparison for the planar sample.

Supplementary Note 8. Computational scaling and optimization considerations for high-resolution implementation

Without implementation-level optimization, the computational cost of the current method increases approximately in proportion to the number of pixels. Therefore, for the same number of frames, extending from 512 × 512 to 4K × 4K images would increase the runtime by about 64×.

In practice, however, substantial optimization space remains. The coarse estimation stage does not require full-resolution data, because it only estimates low-order disturbance parameters; for high-resolution images, downsampling to 512 × 512 or even 256 × 256 is generally sufficient for this purpose. Likewise, in the temporal refinement stage, only three disturbance parameters are updated for each frame, so full-frame data are not strictly necessary. For example, a subset of 65,536 pixels (equivalent to 256 × 256 samples) is already sufficient in practice to update these three parameters for each frame.

The iterative structure is also naturally parallelizable. In particular, the framewise disturbance updates can be parallelized across frames, and the pixelwise spatial updates can be parallelized across pixels. Therefore, the computational burden at high resolution is primarily an implementation issue rather than a fundamental limitation of the algorithmic framework.

For memory usage, the dominant cost comes from storing full-resolution image stacks and per-pixel intermediate variables, rather than from the disturbance parameters themselves. Taking 4K × 4K to mean 4096 × 4096 pixels, a 10-frame double-precision image stack already requires about 4096 × 4096 × 10 × 8 bytes $\approx$ 1.25 GiB. If the normalized interferograms are stored simultaneously, this increases to about 2.50 GiB. In addition, the full-resolution spatial maps A, B, and φ require another 3 × 4096 × 4096 × 8 bytes $\approx$ 384 MiB. Thus, even before accounting for temporary arrays, a straightforward MATLAB implementation can already approach about 3 GiB of memory usage. If the temporal refinement for one frame is implemented by explicitly forming the full-frame residual vector and its 3-column Jacobian, this would introduce about 128 MiB and 384 MiB more, respectively, and the actual peak memory can be higher due to intermediate array copies in MATLAB. These estimates indicate that the memory demand is manageable on standard computers with sufficient RAM.

Supplementary Reference

1 Rabiner, L. R., Schafer, R. W. & Rader, C. M. The chirp z-transform algorithm and its application. *The Bell System Technical Journal* **48**, 1249-1292, doi:10.1002/j.1538-7305.1969.tb04268.x (1969).