Official Journal of CIOMP
elight.springeropen.com

RESEARCH ARTICLE

Open Access

Large-scale phase retrieval

Xuyang Chang, Liheng Bian* and Jun Zhang

Abstract

High-throughput computational imaging requires efficient processing algorithms to retrieve multi-dimensional and multi-scale information. In computational phase imaging, phase retrieval (PR) is required to reconstruct both amplitude and phase in complex space from intensity-only measurements. The existing PR algorithms suffer from the tradeoff among low computational complexity, robustness to measurement noise and strong generalization on different modalities. In this work, we report an efficient large-scale phase retrieval technique termed as *LPR*. It extends the plug-and-play generalized-alternating-projection framework from real space to nonlinear complex space. The alternating projection solver and enhancing neural network are respectively derived to tackle the measurement formation and statistical prior regularization. This framework compensates the shortcomings of each operator, so as to realize high-fidelity phase retrieval with low computational complexity and strong generalization. We applied the technique for a series of computational phase imaging modalities including coherent diffraction imaging, coded diffraction pattern imaging, and Fourier ptychographic microscopy. Extensive simulations and experiments validate that the technique outperforms the existing PR algorithms with as much as 17dB enhancement on signal-to-noise ratio, and more than one order-of-magnitude increased running efficiency. Besides, we for the first time demonstrate ultra-large-scale phase retrieval at the 8K level (7680 × 4320 pixels) in minute-level time.

Keywords: Phase retrieval, Computational imaging, Phase imaging, Large scale

1 Introduction

Wide field of view and high resolution are both desirable for various imaging applications, such as medical imaging [1–4] and remote sensing [5], providing multi-dimensional and multi-scale target information. As the recent development of computational imaging, large-scale detection has been widely employed in a variety of computational imaging modalities [3, 4, 6, 7]. These computational imaging techniques largely extend the spatial-bandwidth product (SBP) [8] of optical systems from million scale to billion scale. As an example, the SBP of the real-time, ultra-large-scale, high-resolution (RUSH) platform [4] and the Fourier ptychographic microscopy (FPM) [3] have reached to as high as 10^8–10^9. Such a large amount of data poses a great challenge for post software processing. Therefore, large-scale processing algorithms with low computational complexity and high fidelity are of great significance for those imaging and perception applications in various dimensions [9].

In computational phase imaging, phase retrieval (PR) is required to reconstruct both amplitude and phase in complex space from intensity-only measurements. This problem originates from the limitation of the low response speed of photodetectors that impedes direct acquisition of light wavefront. Mathematically, the underlying goal of PR is to estimate an unknown complex-field signal from the intensity-only measurements of its complex-valued transformation, which is described as

$$I = |\boldsymbol{A}u|^2 + \omega, \tag{1}$$

where u is the underlying signal to be recovered $\left(u \in \mathbb{C}^{n\times 1}\right)$, I contains the intensity-only measurements $\left(I \in \mathbb{R}^{m\times 1}\right)$, $\boldsymbol{A}$ represents measurement matrix $\left(\boldsymbol{A} \in \mathbb{R}^{m\times n} \text{ or } \mathbb{C}^{m\times n}\right)$, and ω stands for measurement noise. Phase retrieval has been widely applied in plenty fields such as astronomy, crystallography, electron

*Correspondence: bian@bit.edu.cn
School of Information and Electronics & Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing, China

microscopy and optics [10]. It solves various nonlinear inverse problems in optical imaging, such as coherent diffraction imaging [11] (CDI), coded diffraction pattern imaging [12] (CDP), Fourier ptychographic microscopy [3] (FPM) and imaging through scattering medium [13].

In the past few decades, different phase retrieval algorithms have been developed. Gerchberg and Saxton pioneered the earliest alternating projection (AP) algorithm in the 1970s [14], which was then extended by Fienup et al. with several variants [15]. Due to its strong generalization ability, AP has been widely employed in multiple phase imaging models. Nevertheless, it is sensitive to measurement noise, suffering from poor noise robustness. Afterwards, researchers introduced optimization into PR, deriving a series of semi-definite programming (SDP) based algorithms [16, 17] and Wirtinger flow (WF) based algorithms [18–20]. These techniques enhance robustness to measurement noise, but they require high computational complexity and high sampling rate, making them inapplicable for large-scale phase retrieval. Although the sparsity prior of natural images in transformed domains can be incorporated as an additional constraint to lower sampling rate [21, 22], it further increases computational complexity. Although these algorithms can theoretically employ patch-wise [23] and parallel strategies to deal with large-scale data, such a manner leads to a heavier load of memory requirement.

In the last few years, the booming deep learning (DL) technique has also been introduced for phase retrieval [24]. Following the large-scale training framework, the DL strategy outperforms the above traditional PR techniques with higher fidelity. However, it provides poor generalization that each suits only for specific models, such as holography [24] and FPM [25]. For different models and even different system parameters, the deep neural network requires to be retrained with new large-scale data sets. Recently, the prDeep technique [26] integrates iterative optimization and deep learning together, enabling to benefit from respective advantages. However, prDeep cannot recover complex-domain signals, leading to limited applications in practice. To sum, despite of different workflows, the above existing PR algorithms suffer from the tradeoff among low computational complexity, robustness to measurement noise and strong generalization, making them inapplicable for general large-scale phase retrieval.

In this work, we report an efficient large-scale phase retrieval technique termed as *LPR*, as sketched in Fig. 1. It builds on the plug-and-play (PNP) [27] optimization framework, and extends the efficient generalized-alternating-projection (GAP) [9, 28, 29] strategy from real space to nonlinear. The complex-field PNP-GAP scheme ensures strong generalization of *LPR* on various imaging modalities, and outperforms the conventional first-order PNP techniques (such as ADMM [27], ISTA [30] and FISTA [31] used in prDeep) with fewer auxiliary variables, lower computational complexity and faster convergence. As PNP-GAP decomposes reconstruction into separate sub-problems including measurement formation and statistical prior regularization [9, 32], we further introduce an alternating projection solver and an enhancing neural network respectively to solve the two sub-problems. These two solvers compensate the shortcomings of each other, allowing the optimization to bypass the poor generalization of deep learning and poor noise robustness of AP. As a result, *LPR* enables generalized large-scale phase retrieval with high fidelity and low computational complexity, making it a state-of-the-art method for various computational phase imaging applications.

We compared *LPR* with the existing PR algorithms on extensive simulation and experiment of different imaging modalities. The results validate that compared to the AP based PR algorithms, *LPR* is robust to measurement noise with as much as 17dB enhancement on signal-to-noise ratio. Compared with the optimization based PR algorithms, the running time is significantly reduced by more than one order of magnitude. Finally, we for the first time demonstrated ultra-large-scale phase retrieval at the 8K level (7680 × 4320 pixels) in minute-level time, where most of the other PR algorithms failed due to unacceptable high computational complexity.

2 Results

We applied *LPR* and the existing PR algorithms on both simulation and experiment data of three computational phase imaging modalities including CDI, CDP and FPM, to investigate respective pros and cons. The competing algorithms for comparison includes the alternating projection technique (AP) [14, 15], the SDP based techniques (PhaseMax (PMAX) [33], PhaseLift (PLIFT) [16], PhaseLamp (PLAMP) [34]), the Wirtinger flow based techniques (Wirtinger Flow (WF) [18], Reweighted Wirtinger Flow (RWF) [35]), the amplitude flow based techniques [36, 37] (AmpFlow (AF), Truncated AmpFlow (TAF), Reweighted AmpFlow (RAF)), Coordinate Descent (CD) [38], KACzmarz (KAC) [39], prDeep [26] and the deep learning technique (DL) [24]. Most of these algorithms parameters were tuned based on the Phasepack [40] to achieve best performance. The convergence is determined when the intensity difference of reconstructed image between two successive iterations is smaller than a preset threshold. We employed the peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [41] and root mean squared error (RMSE) to quantify reconstruction quality. All the calculation was

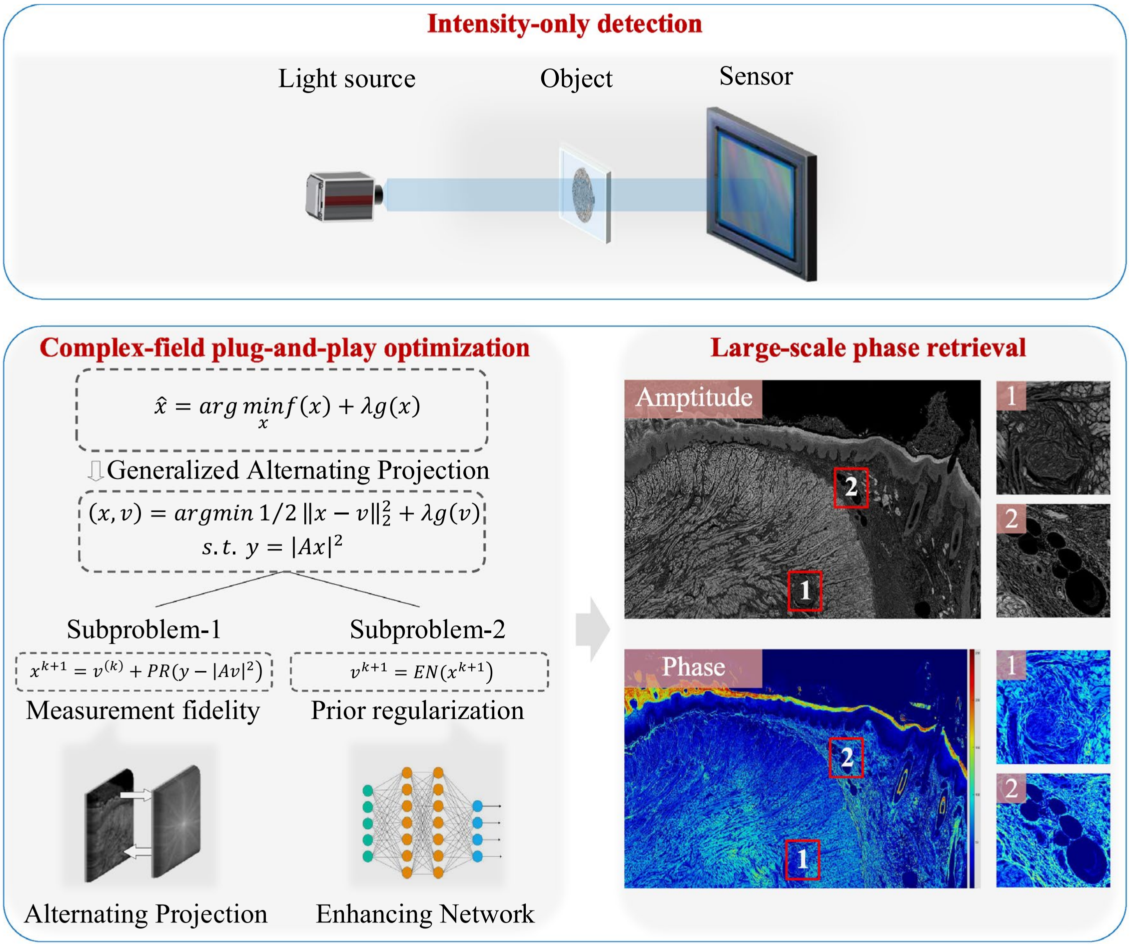

Fig. 1 The schematic of the reported *LPR* technique for large-scale phase retrieval. *LPR* decomposes the large-scale phase retrieval problem into two subproblems under the PNP-GAP framework, and introduces the efficient alternating projection (AP) and enhancing network solvers for alternating optimization. The workflow realizes robust phase retrieval with low computational complexity and strong generalization on different imaging modalities

tested on a desktop PC with an Intel i7-9700 CPU, 16G RAM and an Nvidia GTX 1660s GPU.

2.1 Coherent diffraction imaging

CDI is a representative non-interferometric phase imaging technique, and has been widely applied in physics, chemistry and biology due to its simple setup [10]. It illuminates a target using coherent plane waves, and records the intensity of the far-field diffraction pattern. By oversampling the diffracted light field and applying phase retrieval, both the target's amplitude and phase information can be reconstructed. Mathematically, the measurement formation of CDI is

$$I = |\mathcal{F}(u)|^2, \tag{2}$$

where u denotes the target information, and $\mathcal{F}$ represents the Fourier transformation that approximates the far-field diffraction.

Following the above formation model, we employed a high-resolution image (1356 × 2040 pixels) from the DIV2K [42] dataset and an onion cell image [43] as the latent real-domain signals to synthesize two groups of CDI measurements. Because the prDeep technique for comparison is only applicable in real domain [26], we did not introduce phase into the latent signals. Due to the uniqueness guarantee of the solution, CDI requires at least 4 times oversampling in the Fourier domain [44]. Correspondingly, we padded zeros around the image matrix to generate a 2712 × 4080 image. We implemented Fourier transform to the image and retained only its intensity as measurements. Additionally, to investigate the techniques' robustness to measurement noise, we further added different levels of white Gaussian noise (WGN) to the measurements.

Table 1 presents the quantitative reconstruction evaluation of different techniques. The results show that the CD and KAC methods failed with no convergence. This is because these techniques require higher sampling ratio. The PLIFT and PLAMP methods do not work as well, because they require matrix lifting and involve a higher dimensional matrix that is out of memory in large-scale reconstruction (Additional file 1: Fig. S1 shows the memory requirements of different algorithms under different image sizes). The other methods except for prDeep obtain little improvement compared to the AP algorithm. Specifically, the WF, AF and PMAX methods even degrade due to limited sampling ratio and noise corruption. The reconstruction of prDeep is better than the conventional algorithms, but with only 2dB enhancement on PSNR, and almost no SSIM improvement compared to AP. In contrast, *LPR* produces significant enhancement on reconstruction quality, with as much as 6dB and 0.29 improvement on PSNR and SSIM, respectively. Due to limited space, the results of another set of simulation is presented in Additional file 1: Table S1 and Figs. S2 and S3, which coincides with the above quantitative results.

Table 1 also presents the running time of these techniques. Because all the other algorithms used the result of AP as initialization, we recorded the excess time as the running time of these algorithms. From the results, we can see that prDeep consumes the most running time. *LPR* takes the same level of running time compared to the conventional algorithms, but with significantly improved reconstruction quality.

We further compared these algorithms on experiment CDI data [45], to validate their effectiveness in practical applications. The imaging sample is live glioblastoma cell line U-87 MG. The setup includes a HeNe laser (543nm 5mW), a dual pinhole aperture that consists of two 100 m pinholes spaced 100 m apart from edge to edge, a 35 mm objective lens and a CCD camera (1340 × 1300, 16 bits). Although the situ CDI modality used dual-pinhole illumination that is slightly different from the standard CDI, its reconstruction is still a phase retrieval task in essence. The sequential measurements contain far-field diffraction patterns of several moments in the cell fusion process. Because the conventional algorithms obtain little improvement compared to AP and prDeep is not applicable for complex-field samples [26], we only present the reconstruction results of AP and *LPR* in Fig. 2. The results show that there exist serious noise artifacts in AP reconstruction, especially in the amplitude images. The cells are almost submerged by background noise at 0 and 135 min, and the contours and edges of cells can not be clearly observed. In comparison, *LPR* produces high-fidelity results that effectively preserve fine details while attenuating measurement noise. The complete results of all the 48 moments are shown in Additional file 1: Figs. S4, S5, S6 and S7.

Table 1 Quantitative comparison under the CDI modality

Algorithm	SNR = 20dB			SNR = 25dB			SNR = 30dB		
	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
AP	18.46	0.50	819.67	21.75	0.58	854.37	22.29	0.65	863.14
WF	19.05	0.52	+ 27.15	20.84	0.62	+ 31.98	21.27	0.70	+ 32.41
RWF	18.52	0.50	+ 25.69	21.98	0.61	+ 27.53	22.41	0.71	+ 27.98
AF	16.55	0.42	+ 28.61	19.63	0.49	+ 29.74	19.83	0.54	+ 27.29
TAF	18.57	0.53	+ 26.04	21.81	0.59	+ 25.99	22.30	0.65	+ 26.49
RAF	18.52	0.53	+ 22.55	21.79	0.58	+ 21.80	22.27	0.65	+ 22.19
PLIFT	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PLAMP	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PMAX	16.64	0.42	+ 38.48	19.73	0.49	+ 39.04	19.97	0.54	+ 38.11
CD	✘-no convergence			✘-no convergence			✘-no convergence		
KAC	✘-no convergence			✘-no convergence			✘-no convergence		
prDeep	20.60	0.52	+ 49.01	21.83	0.58	+ 43.36	23.33	0.65	+ 35.46
LPR	**23.30**	**0.79**	+28.52	**25.52**	**0.83**	+ 29.97	**28.11**	**0.86**	+ 27.19

CD and KAC fail with no convergence. PLIFT and PLAMP are out of computer memory. Most of the conventional algorithms produce little improvement than AP. *LPR* outperforms the other algorithms, with as much as 6dB (SNR = 30) and 0.29 (SNR = 20) improvement on PSNR and SSIM, respectively. We use the excess time beyond AP as the other algorithms' running time, which shows that prDeep consumes the most running time. In comparison, *LPR* takes the same level of running time as the conventional methods

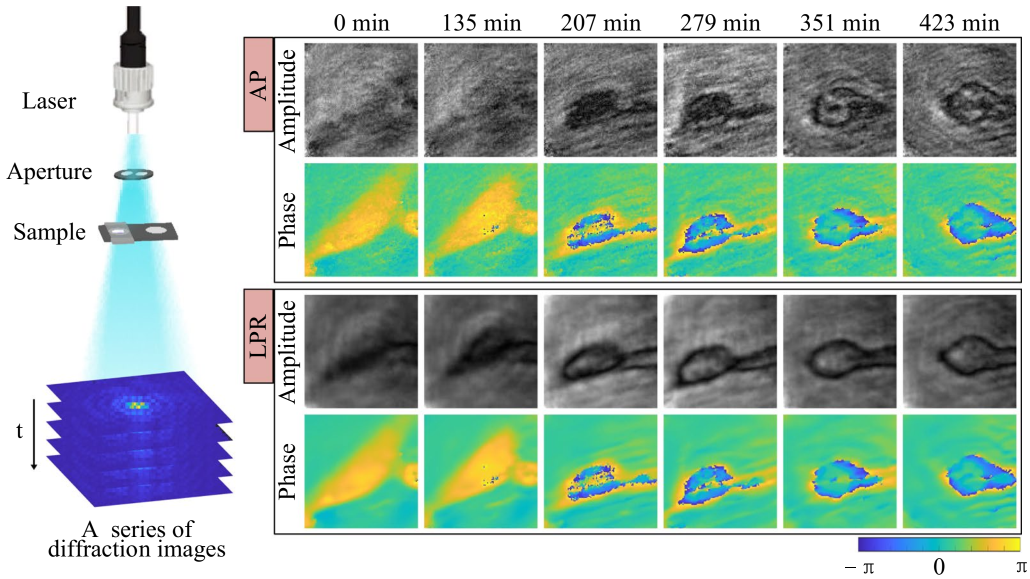

Fig. 2 Comparison of experiment results under the CDI modality [45]. A dual-pinhole aperture is illuminated by a coherent light. A live glioblastoma cell sample is imaged in a time series of diffraction patterns. The reconstructed results describe the fusion process of two glioblastoma cells and form a high-density area. The AP technique is sensitive to measurement noise, and produces unsatisfying results. The reported *LPR* technique enables to remove noise artifacts and preserve fine details with high fidelity

2.2 Coded diffraction pattern imaging

CDP [12] is a coded version of CDI, which introduces wavefront modulation to increase observation diversity. The strategy of multiple modulations and acquisitions enables to effectively bypass the oversampling limitation of the conventional CDI. Generally, the target light field is additionally modulated by a spatial light modulator (SLM), and the measurements after far-field Fraunhofer diffraction can be modeled as

$$I = |\mathcal{F}(u \odot d)|^2, \tag{3}$$

where d represents the modulation pattern, and $\odot$ denotes the Hadamard product.

We simulated CDP measurements with five and single phase modulations, respectively. The modulation patterns d are subject to Gaussian distribution [12]. We employed the same image as CDI to be the ground-truth signal (real domain), and added various levels of WGN to the measurements. Table 2 presents the quantitative evaluation of different techniques under the CDP modality (5 modulations). The results show that the Wirtinger flow based techniques (WF and RWF) failed because of insufficient measurements [18]. The PLIFT and PLAMP methods are still out of memory. The other conventional methods produce either little improvement or even worse reconstruction compared to AP. Although prDeep outperforms AP, it consumes around triple running time with high computational complexity. In comparison, the reported *LPR* obtains the best reconstruction performance, with as much as 8.3dB on PSNR and 0.61 on SSIM. Besides, it also shares the same level of running time as AP, which maintains the highest efficiency among all the algorithms. The detailed visual comparison of different methods is presented in Additional file 1: Fig. S8.

To further demonstrate the strong reconstruction performance of *LPR*, we also compared these algorithms in the case of a limited sampling ratio with only single modulation, as shown in Table 3 and Fig. 3. Due to extremely insufficient measurements, most of the methods failed with either no convergence or poor reconstruction quality. Under heavy measurement noise, the target information is either buried or smoothed. In contrast, the reported *LPR* technique enables as much as 17dB enhancement on PSNR and 0.8 improvement on SSIM. As validated by the close-ups in Fig. 3, *LPR* is able to retrieve fine details, even in the case of heavy measurement noise. Meantime, it is effective to attenuate noise and artifacts, producing smooth background.

In order to further illustrate the computational complexity of different techniques, we show the computation time as a function of image size in Additional file 1: Fig. S9 . We can see that as the image size increases, *LPR* obtains a lower computational complexity than prDeep.

Table 2 Quantitative comparison under the CDP modality (5 modulations)

Algorithm	SNR = 10dB			SNR = 15dB			SNR = 20dB		
	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
AP	15.60	0.21	105.76	18.61	0.33	110.73	23.22	0.55	174.98
WF	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
RWF	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
AF	13.93	0.19	247.07	17.84	0.33	231.38	23.13	0.60	211.39
TAF	13.40	0.16	257.57	18.14	0.34	225.67	22.71	0.59	213.65
RAF	13.88	0.19	261.59	17.86	0.38	222.38	23.10	0.59	212.09
PLIFT	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PLAMP	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PMAX	11.08	0.13	295.84	11.36	0.14	300.21	11.66	0.15	296.28
CD	8.69	0.22	357.52	9.47	0.20	321.81	9.78	0.20	264.89
KAC	10.83	0.13	192.44	10.97	0.15	161.48	11.01	0.16	114.75
prDeep	22.67	0.61	301.41	24.42	0.72	282.14	26.85	0.76	380.60
LPR	**22.73**	**0.82**	124.80	**26.92**	**0.88**	137.33	**31.89**	**0.94**	228.42

The Wirtinger flow based (WF, RWF) techniques fail because of insufficient measurements. PLIFT and PLAMP are out of memory. The other methods produce little improvement or consume extremely long running time compared to AP. In comparison, *LPR* consumes the same level of running time as AP, and obtains the best performance with as much as 8.3dB on PSNR (SNR = 15) and 0.61 on SSIM (SNR = 10)

Table 3 Quantitative comparison under the CDP modality (single modulation)

Algorithm	SNR = 10dB			SNR = 15dB			SNR = 20dB		
	PSNR	SSIM	TIME	PSNR	SSIM	TIME	PSNR	SSIM	TIME
AP	11.71	0.08	13.96	12.82	0.09	13.55	13.02	0.10	13.34
WF	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
RWF	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
AF	10.47	0.08	24.61	10.53	0.08	23.73	10.82	0.09	23.36
TAF	10.52	0.08	24.05	10.93	0.07	24.21	11.02	0.08	23.09
RAF	10.38	0.06	26.17	10.43	0.07	25.83	10.78	0.08	25.82
PLIFT	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PLAMP	✘-memory limitation			✘-memory limitation			✘-memory limitation		
PMAX	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
CD	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
KAC	✘-insufficient measurements			✘-insufficient measurements			✘-insufficient measurements		
prDeep	18.29	0.39	153.41	19.21	0.54	142.34	23.92	0.68	104.84
LPR	**21.11**	**0.81**	77.80	**25.64**	**0.87**	81.51	**30.10**	**0.89**	62.89

Most of the conventional algorithms fail with either no convergence or poor reconstruction quality because of extremely insufficient measurements. In comparison, *LPR* still obtains the best reconstruction quality, with more than 17dB improvement on PSNR and nearly 0.8 on SSIM (SNR=20)

2.3 Fourier ptychographic microscopy

FPM is a novel technique to increase optical system's bandwidth for wide-field and high-resolution imaging. It illuminates the target with coherent light at different incident angles, and acquires corresponding images that contain information of different sub-regions of the target's spatial spectrum. Mathematically, the measurement formation model of FPM is

$$I = \left| \mathcal{F}^{-1}[P \odot \mathcal{F}\{u \odot \mathcal{S}\}] \right|^2, \tag{4}$$

where $\mathcal{F}^{-1}$ is inverse Fourier transform, P denotes system's pupil function, and $\mathcal{S}$ represents the wave function of incident light.

Following the formation model, we first implemented a simulation comparison with the following setup

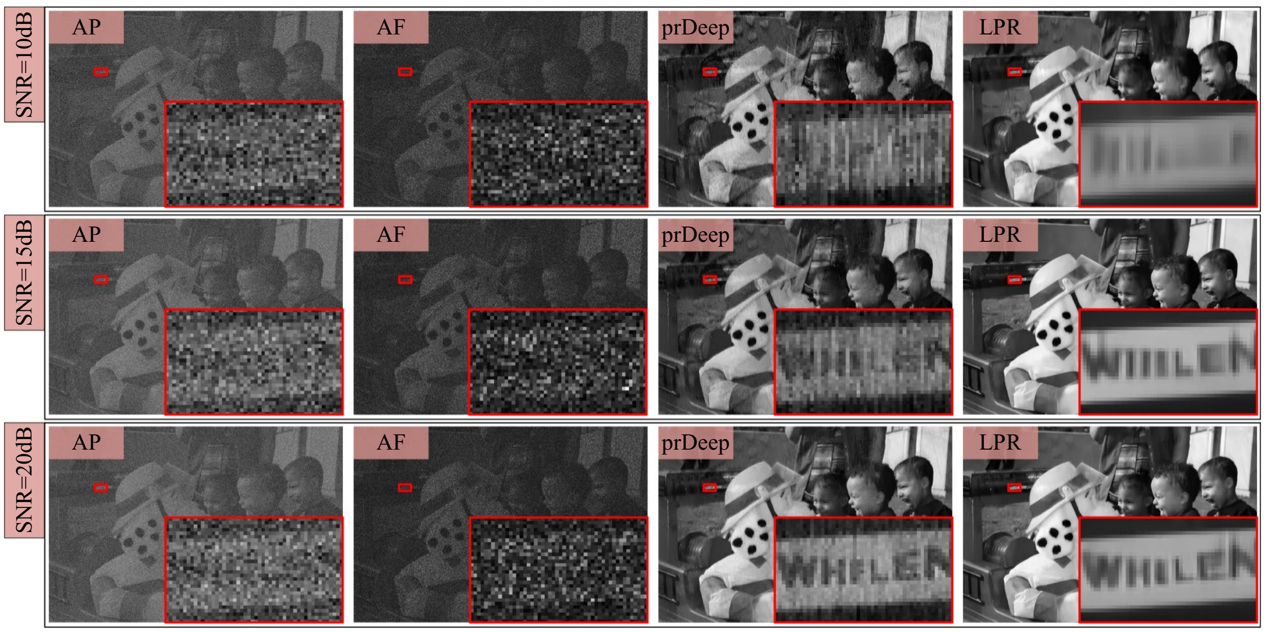

Fig. 3 Visual comparison under the CDP imaging modality (single modulation). In such a low sampling ratio with measurement noise, all the conventional algorithms produce low-contrast resolution. The prDeep technique also produces serious reconstruction artifacts. The reported *LPR* technique outperforms the other methods with much higher fidelity

parameters: the wavelength is 625nm, the numerical aperture (NA) of objective lens is 0.08, the height from the light source to the target is 84.8mm, and the distance between adjacent light sources is 4mm. The pixel size of camera is 3.4μm.Two microscopy images of blood cells [46] (2048 $\times$ 2048 pixels) were employed as the latent high-resolution (HR) amplitude and phase, respectively. The size of captured low-resolution images (LR) was one fourth of the HR images.

Figure 4 presents the reconstruction results of AP [3], WF [47], deep learning (DL) [24] and *LPR*. For the DL technique, we used the result of the AP algorithm as the network's input, and the network outputted the enhanced reconstruction results. In the training process, we used 20,000 images (10,000 each for amplitude and phase) from the PASCAL Visual Object Classes dataset [48] and DIV2K dataset [42], and trained the network individually for different noise levels. From the results, we can see that AP is sensitive to measurement noise. WF can better handle noise, but it requires high computational complexity and long running time (more than one order of magnitude). Although DL consumes the least inferring time and outperforms the AP and WF methods, its reconstruction quality is still worse than LPR in the presence of measurement noise. Compared with AP, *LPR* obtains as much as nearly 10dB enhancement on PSNR (SNR $=$ 10). Besides, it consumes the same order of running time as AP. The visual comparison also validates that *LPR* enables high-fidelity reconstruction of both amplitude and phase. Due to space limitation, we present the other two sets of simulation results in Additional file 1: Figs. S10 and S11.

We also implemented the algorithms on experiment FPM measurements. The imaging sample is a blood smear stained by HEMA 3 Wright-Giemsa. The setup consists of a 15 $\times$ 15 LED array, a 2$\times$ 0.1 NA objective lens (Olympus), and a camera with 1.85μm pixel size. The central wavelength of the LEDs is 632nm, and the lateral distance between adjacent LEDs is 4mm. The LED array is placed 80mm from the sample. We captured two sets of 225 LR images that correspond to the 15 $\times$ 15 LEDs, respectively under 1ms and 0.25ms exposure time. The reconstructed results are presented in Fig. 5, which shows that AP is seriously degraded under limited exposure. Only the cell nucleus can be observed in amplitude, and other details are lost. *LPR* produces state-of-the-art reconstruction performance. The measurement noise is effectively removed, and the cell structure and morphology details are clearly retrieved.

2.4 Ultra-large-scale phase retrieval

In ultra-large-scale imaging applications such as 4K (4096 $\times$ 2160 pixels) or 8K (7680 $\times$ 4320 pixels), most reconstruction algorithms are not applicable due to either highly large memory requirement or extremely long running time. Nevertheless, the reported *LPR* technique

		Amplitude		Phase	
		PSNR	SSIM	RMSE	TIME
SNR 5dB	AP	14.8	0.38	2.57	27.1
	WF	17.1	0.46	2.56	1630
	DL	20.1	0.78	2.38	5.4
	LPR	22.3	0.86	1.45	42.5
SNR 10dB	AP	19.1	0.41	2.46	25.2
	WF	22.5	0.48	2.47	1604
	DL	28.9	0.78	2.04	5.2
	LPR	29.0	0.89	1.33	41
SNR 15dB	AP	25.4	0.78	2.41	24.9
	WF	27.9	0.85	2.42	1490
	DL	29.9	0.82	1.48	5.2
	LPR	30.6	0.92	1.32	41.7

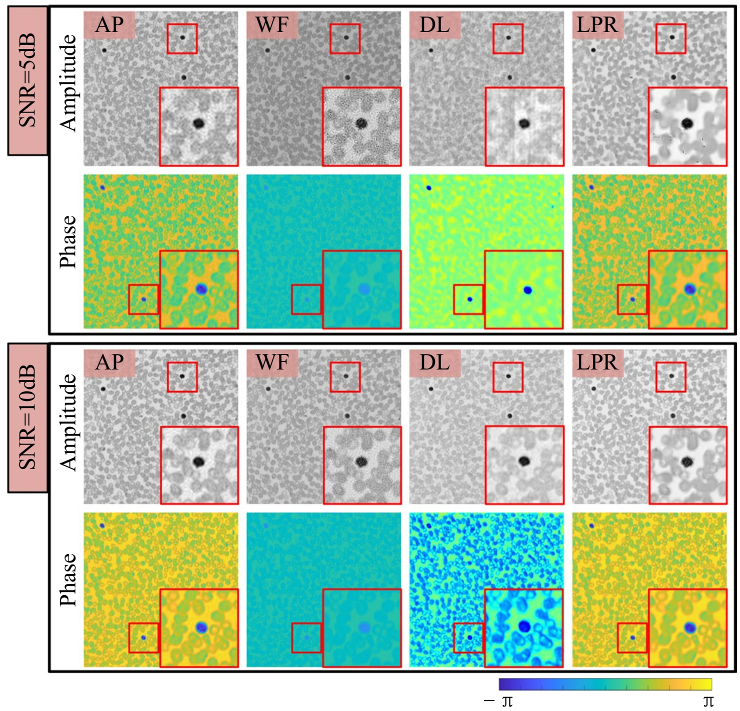

Fig. 4 Comparison of simulation results under the FPM modality. The left table presents quantitative comparison, while the right images show visual comparison. AP suffers from poor noise robustness. WF requires high computational complexity with longer running time (more than one order of magnitude). Although the deep learning technique consumes the least running time and outperforms the AP and WF methods, its reconstruction quality is still worse than *LPR* in the presence of measurement noise. In contrast, *LPR* produces the highest reconstruction quality with as much as nearly 10dB enhancement on PSNR (SNR = 10) and consumes the same order of running time as AP

still works well in such applications. As a demonstration, we implemented a simulation of 8K-level CDP (5 modulations), using an 8K outer space color image as the real-domain ground truth (released by NASA using the Hubble Telescope). Its spatial resolution is 7680×4320 (each color channel) with in total 33.1 million pixels. We simulated intensity-only measurements individually for different RGB channels, and the reconstruction was also implemented separately for different channels. Figure 6 presents the reconstruction results of AP and *LPR*, with the input SNR being 5dB. The close-ups show that the result of AP is drowned out by measurement noise, leading to dimness and loss of target details. In comparison, *LPR* outperforms a lot with strong robustness. Both their running times lie in the minute level. Another set of 8K reconstruction results is shown in Additional file 1: Fig. S12).

3 Methods

Following optimization theory, the phase retrieval task can be modeled as

$$\hat{u} = \arg\min_{u} f(u) + \lambda g(u), \tag{5}$$

where u denotes the target complex field to be recovered, $f(u)$ is a data-fidelity term that ensures consistency between the reconstructed result and measurements, and $g(u)$ is a regularizer that imposes certain statistical prior knowledge. Conventionally, Eq. (5) is solved following the first-order proximal gradient methods, such as ISTA and ADMM that are time-consuming to calculate gradients in large-scale nonlinear tasks [32]. In this work, instead, we employ the efficient generalized-alternating-projection (GAP) strategy [32] to transform Eq. (5) with fewer variables to

$$\begin{aligned}(u, v) = \operatorname{argmin} 1/2\|u - v\|_2^2 + \lambda g(v) \\ \text{s.t. } I = |Au|^2,\end{aligned} \tag{6}$$

where v is an auxiliary variable balancing the data fidelity term and prior regularization, A denotes measurement matrix, and I represents measurement. The difference between the conventional ADMM and GAP optimization is the constraint on the measurement [32]. ADMM minimizes $\left\|I - |Au|^2\right\|$, while GAP imposes the constraint $I = |Au|^2$.

To tackle the large-scale phase retrieval task, we extend the efficient plug-and-play (PNP) optimization framework

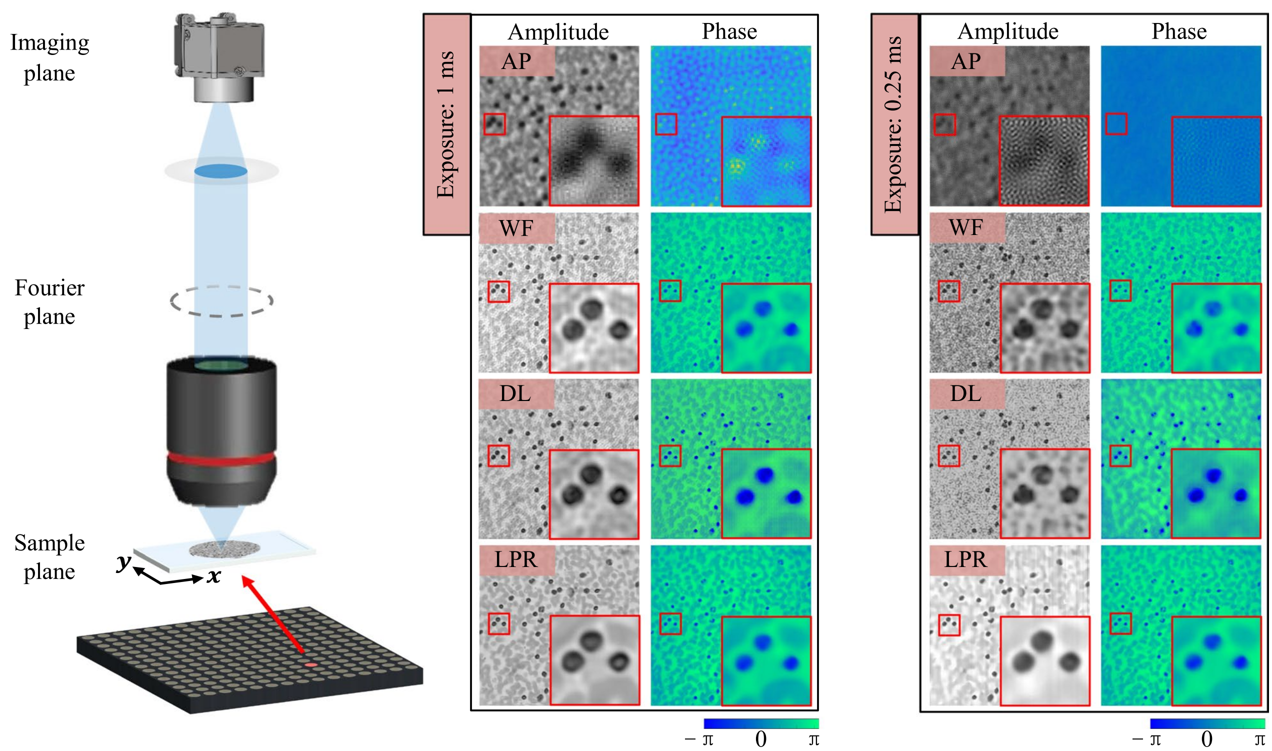

Fig. 5 Comparison of experiment results under the FPM modality. The target is a red blood cell sample that is prepared on a microscope slide stained with Hema 3 stain set (Wright-Giemsa). The limited exposure results in serious measurement noise, which directly flows into the reconstruction results of AP. The WF technique outperforms AP, but it still degrades a lot under a short exposure time (0.25ms). The reported *LPR* technique maintains strong robustness to measurement noise, and enables to retrieve clear cell structure and morphology details

[27] from real space to nonlinear complex space. Fundamentally, PNP decomposes optimization into two separate sub-problems including measurement formation and prior regularization, so as to incorporating inverse recovery solvers together with various image enhancing solvers to improve reconstruction accuracy, providing high flexibility for different applications. Mathematically, Eq. (6) is decomposed into the following two sub-problems, to alternatively update the two variables u and v.

- Updating u: given $v^{(k)}$, $u^{(k+1)}$ is updated via a Euclidean projection of $v^{(k)}$ on the manifold $I = |Au|^2$ as

$$u^{k+1} = v^{(k)} + \lambda \cdot PR\left(I - |Av|^2\right), \tag{7}$$

  where PR is phase retrieval solver. Considering its great generalization ability on various imaging modalities and low computational complexity, we employ the AP method as the PR solver. It alternates between the target and observation planes allowing to incorporate any information available for the variables, providing low sampling rate requirement.
- Updating v: given $u^{(k+1)}$, $v^{(k+1)}$ is updated by an image enhancing solver EN as

$$v^{k+1} = EN\left(u^{k+1}\right). \tag{8}$$

Although the iterative image enhancing research has made great progress in recent years with such as non-local optimization and dictionary learning [49], they maintain high computational complexity for large-scale reconstruction [50]. In this work, considering its state-of-the-art enhancement performance and flexibility to tackle different noise levels, we employed the deep learning based FFDNET [51] to deal with the sub-problem with high fidelity and self-adaptation. The neural network consists of a series of 3×3 convolution layers. Each layer is composed of a specific combination of three types of operations including convolution, rectified linear units and batch normalization. The architecture provides a balanced tradeoff between noise suppression and detail fidelity. While an image is input into the network, it is first down sampled into several sub-blocks, which then flow through the network for quality enhancement. Finally, these optimized blocks are stitched together to the original size. Such a workflow enables

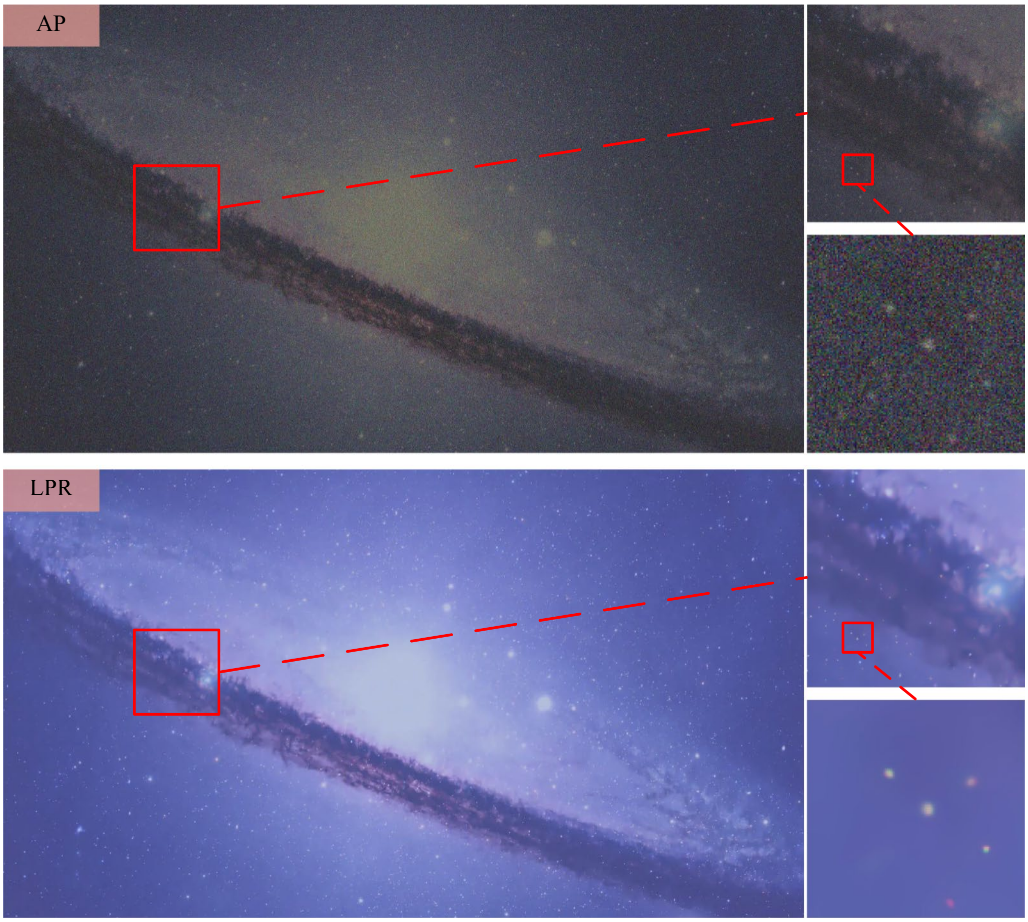

Fig. 6 The first demonstration of ultra-large-scale phase retrieval at the 8K level (7680 × 4320 ×3 pixels). The imaging modality is CDP with 5 modulations. At such a large scale, only the AP and the reported *LPR* techniques still work, while the other ones fail due to high computational complexity. The results validate that *LPR* significantly outperforms AP with effective noise removal and detail reservation

its great generalization ability on different image sizes.

After initialization, the variables are updated alternatively following Eqs. (7) and (8). When the intensity difference of the reconstructed image between two successive iterations is smaller than a given threshold, the iteration stops with convergence. Since both the two solvers *PR* and *EN* are highly efficient and flexible, the entire reconstruction maintains low computational complexity and great generalization. The complete *LPR* algorithm is summarized in Algorithm 1 (Additional file 1), and the demo code has been released at https://github.com/bianlab/bianlab.github.io.

4 Conclusion and discussion

In this work, we engaged to tackle the large-scale phase retrieval problem, and reported a generalized *LPR* optimization technique with low computational complexity and strong robustness. It extends the efficient PNP-GAP framework from real space to nonlinear complex space, and incorporates the alternating projection solver and enhancing neural network. As validated by extensive simulations and experiments on three different computational phase imaging modalities (CDI, CDP and FPM), *LPR* exhibits unique advantages in large-scale phase retrieval tasks with high fidelity and efficiency.

The PNP framework has a theoretical guarantee of convergence for most real-domain tasks, such as denoising, deblurring [52, 53], etc. However, to the best of our knowledge, there is no theoretical proof of PNP's convergence in the complex domain. Further, there is also no theoretical guarantee of convergence for the alternating projection solver that has been widely used for ~50 years [10]. Even though, the extensive experimental results of various imaging modalities in this work and other studies (e.g. Fourier ptychographic microscopy [3], coherent diffraction imaging [11], ptychography [54], and coded diffraction patterns [12]) have validated that the PNP framework and the alternating-projection solver can successfully converge to a global minimum.

The *LPR* technique can be further extended. First, it involves multiple algorithm parameters that are currently adjusted manually. We can introduce the reinforcement learning technique [55] in our future work to automatically adjust these parameters for best performance. Second, *LPR* is sensitive to initialization, especially under low sampling rate. The optimal spectral initialization [56] technique can be incorporated for stronger robustness. Third, the stagnation problem in blind ptychographic reconstruction [54] deserves further study under the reported framework. This enables to simultaneously recover both object and system parameters. Fourth, it is interesting to investigate the influence of employing other image enhancing solvers such as super-resolution neural network, deblurring network and distortion removal network. This may open new insights for phase retrieval with further boosted quality.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s43593-021-00004-w.

Additional file 1: Figure S1. The relationship between memory requirement and image size under the CDI modality. **Figure S2.** Visual comparison under the CDI modality. **Table S1.** Quantitative comparison under the CDI modality (onion cell). **Figure S3.** Visual comparison under the CDI modality (onion cell). **Figure S4.** Experiment amplitude results of AP under the CDI modality. **Figure S5.** Experiment phase results of AP under the CDI modality. **Figure S6.** Experiment amplitude results of LPR under the CDI modality. **Figure S7.** Experiment phase results of LPR under the CDI modality. **Figure S8.** Visual comparison of simulation results under the CDP modality (5 modulations). **Figure S9.** The relationship between running time and image size under the CDP modality. **Figure S10.** Comparison of simulation results under the FPM modality. **Figure S11.** Comparison of large-phase-range phase retrieval results under the FPM modality. **Figure S12.** Ultra-large-scale phase retrieval at the 8K level (7680×4320 pixels).

Acknowledgements
The authors would like to thank anonymous reviewers for helpful and stimulating comments.

Authors' contributions
LB and XC conceived the idea and designed the experiments. XC conducted the simulations and experiments. All the authors contributed to writing and revising the manuscript, and convolved in discussions during the project. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (Nos. 61971045, 61827901, 61991451), National Key R&D Program (Grant No. 2020YFB0505601), Fundamental Research Funds for the Central Universities (Grant No. 3052019024).

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare no competing financial interests.

Received: 28 April 2021 Revised: 24 June 2021 Accepted: 13 July 2021
Published online: 03 September 2021

References

1. H. Pahlevaninezhad, M. Khorasaninejad, Y.-W. Huang, Z. Shi, L.P. Hariri, D.C. Adams, V. Ding, A. Zhu, C.-W. Qiu, F. Capasso et al., Nano-optic endoscope for high-resolution optical coherence tomography in vivo. Nat. Photonics **12**(9), 540–547 (2018)
2. A. Lombardini, V. Mytskaniuk, S. Sivankutty, E.R. Andresen, X. Chen, J. Wenger, M. Fabert, N. Joly, F. Louradour, A. Kudlinski et al., High-resolution multimodal flexible coherent Raman endoscope. Light. Sci. Appl. **7**(1), 1–8 (2018)
3. G. Zheng, R. Horstmeyer, C. Yang, Wide-field, high-resolution Fourier ptychographic microscopy. Nat. Photonics **7**(9), 739–745 (2013)
4. J. Fan, J. Suo, J. Wu, H. Xie, Y. Shen, F. Chen, G. Wang, L. Cao, G. Jin, Q. He et al., Video-rate imaging of biological dynamics at centimetre scale and micrometre resolution. Nat. Photonics **13**(11), 809–816 (2019)
5. W.-Q. Wang, Space-time coding MIMO-OFDM SAR for high-resolution imaging. IEEE T. Geosci. Remote **49**(8), 3094–3104 (2011)
6. D.J. Brady, M.E. Gehm, R.A. Stack, D.L. Marks, D.S. Kittle, D.R. Golish, E. Vera, S.D. Feller, Multiscale gigapixel photography. Nature **486**(7403), 386–389 (2012)
7. H. Wang, Z. Göröcs, W. Luo, Y. Zhang, Y. Rivenson, L.A. Bentolila, A. Ozcan, Computational out-of-focus imaging increases the space-bandwidth product in lens-based coherent microscopy. Optica **3**(12), 1422–1429 (2016)
8. A.W. Lohmann, R.G. Dorsch, D. Mendlovic, Z. Zalevsky, C. Ferreira, Space-bandwidth product of optical signals and systems. JOSA A **13**(3), 470–473 (1996)
9. X. Yuan, Y. Liu, J. Suo, Q. Dai, Plug-and-play algorithms for large-scale snapshot compressive imaging. In: Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1447–1457 (2020)
10. Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, M. Segev, Phase retrieval with application to optical imaging: a contemporary overview. IEEE Signal Proc. Mag. **32**(3), 87–109 (2015)
11. J. Miao, P. Charalambous, J. Kirz, D. Sayre, Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens. Nature **400**(6742), 342–344 (1999)
12. E.J. Candes, X. Li, M. Soltanolkotabi, Phase retrieval from coded diffraction patterns. Appl. Comput. Harmon. A. **39**(2), 277–299 (2015)
13. O. Katz, P. Heidmann, M. Fink, S. Gigan, Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. Nat. Photonics **8**(10), 784–790 (2014)
14. R.W. Gerchberg, A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik **35**, 237–246 (1972)
15. J.R. Fienup, Phase retrieval algorithms: a comparison. Appl. Optics **21**(15), 2758–2769 (1982)
16. E.J. Candes, T. Strohmer, V. Voroninski, Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. Commun. Pur. Appl. Math. **66**(8), 1241–1274 (2013)
17. L. Vandenberghe, S. Boyd, Semidefinite programming. SIAM Rev. **38**(1), 49–95 (1996)
18. E.J. Candes, X. Li, M. Soltanolkotabi, Phase retrieval via Wirtinger flow: Theory and algorithms. IEEE T. Inform. Theory **61**(4), 1985–2007 (2015)

19. Y. Chen, E. Candes, Solving random quadratic systems of equations is nearly as easy as solving linear systems. In: International Conference on Neural Information Processing Systems (NIPS), pp. 739–747 (2015)
20. W.-J. Zeng, H.-C. So, Coordinate descent algorithms for phase retrieval. Signal Process. **169**, 107418 (2020)
21. V. Katkovnik, Phase retrieval from noisy data based on sparse approximation of object phase and amplitude. arXiv preprint arXiv:1709.01071 (2017)
22. C.A. Metzler, A. Maleki, R.G. Baraniuk, BM3D-PRGAMP: Compressive phase retrieval based on BM3D denoising. In: International Conference on Image Processing (ICIP), pp. 2504–2508 (2016). IEEE
23. S. Chowdhury, M. Chen, R. Eckert, D. Ren, F. Wu, N. Repina, L. Waller, High-resolution 3D refractive index microscopy of multiple-scattering samples from intensity images. Optica **6**(9), 1211–1219 (2019)
24. Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, A. Ozcan, Phase recovery and holographic image reconstruction using deep learning in neural networks. Light Sci. Appl. **7**(2), 17141–17141 (2018)
25. A. Kappeler, S. Ghosh, J. Holloway, O. Cossairt, A. Katsaggelos, Ptychnet: CNN based Fourier ptychography. In: International Conference on Image Processing (ICIP), pp. 1712–1716 (2017). IEEE
26. C. Metzler, P. Schniter, A. Veeraraghavan, *et al*: prDeep: robust phase retrieval with a flexible deep network. In: International Conference on Machine Learning (ICML), pp. 3501–3510 (2018). PMLR
27. S.V. Venkatakrishnan, C.A. Bouman, B. Wohlberg, Plug-and-play priors for model based reconstruction. In: Global Conference on Signal and Information Processing (GlobalSIP), pp. 945–948 (2013). IEEE
28. X. Liao, H. Li, L. Carin, Generalized alternating projection for weighted-2,1 minimization with applications to model-based compressive sensing. SIAM J. Imaging Sci. **7**(2), 797–823 (2014)
29. X. Yuan, Generalized alternating projection based total variation minimization for compressive sensing. In: International Conference on Image Processing (ICIP), pp. 2539–2543 (2016). IEEE
30. J.M. Bioucas-Dias, M.A. Figueiredo, A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. IEEE T. Image Process. **16**(12), 2992–3004 (2007)
31. A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
32. Y. Liu, X. Yuan, J. Suo, D.J. Brady, Q. Dai, Rank minimization for snapshot compressive imaging. IEEE T. Pattern Anal. **41**(12), 2990–3006 (2018)
33. T. Goldstein, C. Studer, Phasemax: Convex phase retrieval via basis pursuit. IEEE T. Inform. Theory **64**(4), 2675–2689 (2018)
34. O. Dhifallah, C. Thrampoulidis, Y.M. Lu, Phase retrieval via linear programming: Fundamental limits and algorithmic improvements. In: Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 1071–1077 (2017). IEEE
35. Z. Yuan, H. Wang, Phase retrieval via reweighted Wirtinger flow. Appl. Optics **56**(9), 2418–2427 (2017)
36. G. Wang, G.B. Giannakis, Y.C. Eldar, Solving systems of random quadratic equations via truncated amplitude flow. IEEE T. Inform. Theory **64**(2), 773–794 (2017)
37. G. Wang, G.B. Giannakis, Y. Saad, J. Chen, Phase retrieval via reweighted amplitude flow. IEEE T. Signal Proces. **66**(11), 2818–2833 (2018)
38. W.-J. Zeng, H.-C. So, Coordinate descent algorithms for phase retrieval. arXiv preprint arXiv:1706.03474 (2017)
39. K. Wei, Solving systems of phaseless equations via Kaczmarz methods: A proof of concept study. Inverse Probl. **31**(12), 125008 (2015)
40. R. Chandra, T. Goldstein, C. Studer, Phasepack: A phase retrieval library. In: International Conference on Sampling Theory and Applications (SampTA), pp. 1–5 (2019). IEEE
41. Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE T. Image Process. **13**(4), 600–612 (2004)
42. E. Agustsson, R. Timofte, Ntire 2017 challenge on single image super-resolution: Dataset and study. In: Conference on Computer Vision and Pattern Recognition (CVPR), pp. 126–135 (2017)
43. Choksawatdikorn: Onion cells under microscope view. https://www.shutterstock.com/zh/image-photo/onion-cells-microscope-1037260501. [Online; accessed 20-June-2021] (2021)
44. J. Miao, T. Ishikawa, I.K. Robinson, M.M. Murnane, Beyond crystallography: Diffractive imaging using coherent X-ray light sources. Science **348**(6234), 530–535 (2015)
45. Y.H. Lo, L. Zhao, M. Gallagher-Jones, A. Rana, J.J. Lodico, W. Xiao, B. Regan, J. Miao, In situ coherent diffractive imaging. Nat. Commun. **9**(1), 1–10 (2018)
46. Choksawatdikorn: Blood cells under microscope view for histology education. https://www.shutterstock.com/zh/image-photo/blood-cells-under-microscope-view-histology-1102617128. [Online; accessed 5-November-2020] (2020)
47. L. Bian, J. Suo, G. Zheng, K. Guo, F. Chen, Q. Dai, Fourier ptychographic reconstruction using Wirtinger flow optimization. Opt. Express **23**(4), 4856–4866 (2015)
48. M. Everingham, L. Van Gool, C.K.I. Williams, J. Winn, A. Zisserman, The PASCAL Visual Object Classes Challenge 2012 (VOC2012) Results. http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html
49. M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries. IEEE T. Image Process. **15**(12), 3736–3745 (2006)
50. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE T. Image Process. **26**(7), 3142–3155 (2017)
51. K. Zhang, W. Zuo, L. Zhang, FFDNet: Toward a fast and flexible solution for CNN-based image denoising. IEEE T. Image Process. **27**(9), 4608–4622 (2018)
52. S.H. Chan, X. Wang, O.A. Elgendy, Plug-and-play admm for image restoration: Fixed-point convergence and applications. IEEE Transact. Comput Imaging **3**(1), 84–98 (2016)
53. P. Nair, R.G. Gavaskar, K.N. Chaudhury, Fixed-point and objective convergence of plug-and-play algorithms. IEEE Transactions on Computational Imaging **7**, 337–348 (2021)
54. S. Jiang, J. Zhu, P. Song, C. Guo, Z. Bian, R. Wang, Y. Huang, S. Wang, H. Zhang, G. Zheng, Wide-field, high-resolution lensless on-chip microscopy via near-field blind ptychographic modulation. Lab Chip **20**(6), 1058–1065 (2020)
55. K. Wei, A. Aviles-Rivero, J. Liang, Y. Fu, C.-B. Schönlieb, H. Huang, Tuning-free plug-and-play proximal algorithm for inverse imaging problems. In: International Conference on Machine Learning (ICML), pp. 10158–10169 (2020). PMLR
56. W. Luo, W. Alghamdi, Y.M. Lu, Optimal spectral initialization for signal recovery with applications to phase retrieval. IEEE T. Signal Proces. **67**(9), 2347–2356 (2019)