LOW RANK FOURIER PTYCHOGRAPHY

Zhengyu Chen, Gauri Jagatap, Seyedehsara Nayer, Chinmay Hegde, Namrata Vaswani

Iowa State University

ABSTRACT

In this paper, we introduce a principled algorithmic approach for Fourier ptychographic imaging of dynamic, time-varying targets. To the best of our knowledge, this setting has not been explicitly addressed in the ptychography literature. We argue that such a setting is very natural, and that our methods provide an important first step towards helping reduce the sample complexity (and hence acquisition time) of imaging dynamic scenes to manageable levels. With significantly reduced acquisition times per image, our approach provides an path towards practical dynamic ptychographic imaging of time varying scenes.

Index Terms— Phase retrieval, ptychography, low rank

1. INTRODUCTION

1.1. Motivation

In recent years, the classical problem of *phase retrieval* has attracted renewed interest in the signal and image processing community. The phase retrieval problem involves reconstructing a length-n discrete-time signal (or image) given noisy observations of the magnitudes of its discrete Fourier transform (DFT) coefficients. A generalized version of phase retrieval studies a similar reconstruction problem where the DFT is replaced by a generic linear measurement operator. A series of recent breakthrough results [1, 2, 3, 4, 5] have introduced principled and provably accurate algorithms for generalized phase retrieval, provided the measurement operator is constructed by sampling vectors from certain families of multivariate probability distributions.

Phase retrieval algorithms enable a variety of imaging applications ranging from X-ray crystallography and biomedical imaging [6, 7, 8]. A related imaging technique is known as *Fourier ptychography*, which can be used for super-resolving images obtained in microscopic imaging systems. The high level approach is to capture *multiple* snapshots of a target scene using a programmable coherent illumination source coupled with a system involving two lenses, and reconstruct a high-resolution image of the target scene via (generalized) phase retrieval. One way to engineer multiple snapshots of a scene is to fix the position of the illumination source, and either let the camera aperture undergo spatial translations [9], or construct an array of fixed cameras, each of which captures a specific portion of the Fourier spectrum of the desired high-resolution image. Zheng et al. [10] have demonstrated that using such a system, one can image beyond the diffraction-limit of the objective lens in a microscope. Recently, Holloway et al. have demonstrated similar, very promising results in the context of long-distance sub-diffraction imaging [11].

While the above results indicate the considerable promise of Fourier ptychography, an *algorithmic* understanding of the image reconstruction problem is nascent, and several important questions remain unanswered. Rigorous guarantees of correctness of the methods proposed in [9, 11] are not yet available. More worrisome is the issue of *sample complexity* of image reconstruction, since a basic requirement in all existing methods is that the number of measurements must exceed the resolution of the image. However, this requirement can be particularly challenging when imaging a *dynamic* scene involving a moving target; for a video sequence with q images each with resolution n, without using any structural prior assumptions, the number of observations must be at least $\Omega(nq)$, which can quickly become prohibitive.

1.2. Our Contributions

In this paper, we introduce a principled algorithmic approach for Fourier ptychographic imaging of dynamic, time-varying targets. To the best of our knowledge, this setting has not been explicitly addressed in the ptychography literature. However, we argue that such a setting is very natural, and that our methods provide an important first step towards alleviating the aforementioned issues of sample complexity that can arise in Fourier ptychography.

The high level idea is that if the dynamics of the scene are sufficiently slow, then the underlying video signal can be well-modeled by a *low-rank matrix*. This modeling assumption has been successfully employed in a variety of video acquisition, compression, and enhancement applications [12, 13, 14, 15, 16]. Specifically, if we reshape each image in the video sequence as a vector $\boldsymbol{x}_k \in \mathbb{R}^n$ and stack up q consecutive frames into a matrix $\boldsymbol{X} \in \mathbb{R}^{n \times q}$, then $\boldsymbol{X}$ is (approximately) rank-r, with $r \ll \min(n, q)$.

The crux of this paper is to demonstrate how we can effectively leverage such low-rank structure in order to enable better image reconstruction (specifically, one that surpasses the naïve approach of reconstructing each image frame by frame using an existing method). To this end, we advocate a new Fourier ptychographic approach that consists of two key ingredients:

1. We develop two novel "under-sampling" strategies that can considerably reduce the sample complexity of video Fourier ptychography. Moreover, these strategies can be implemented in common Fourier ptychographic imaging setups (such as [10, 11]) in a straightforward fashion.
2. We couple these under-sampling strategies with a new image reconstruction algorithm which fully exploits the underlying low-rank structure of the target video sequence. Moreover, we confirm the advantages of this algorithm via a number of simulation experiments.

Our algorithm builds upon those introduced in our recent previous work on low rank phase retrieval [17]. While that paper analyzes a similar setup for leveraging low-rank structure for (generalized) phase retrieval, we only considered special families of measurement operators (i.i.d. Gaussian measurements [1] and coded-diffraction patterns (CDP) [18]). We expand the utility of our previous work

to the Fourier ptychographic measurement setup, hence demonstrating a real-world application of our approach for the first time.

Our reconstruction algorithm involves a non-convex, iterative estimation procedure; hence, initializing the algorithm properly is crucial. It turns out that the initialization procedure in [17] is not particularly effective (since they are specially tailored to the Gaussian or the CDP case). Instead, we devise a novel initialization mechanism for our reconstruction algorithm, and justify it conceptually as well as in experiments. We experimentally demonstrate that our new modified reconstruction algorithm (that we call Low-Rank Ptychography, or *LRPtych*), compares very favorably in terms of sample complexity when contrasted with existing "single-frame" methods, such as the Iterative Error Reduction Algorithm (IERA) of [11].

This paper focuses on Fourier ptychographic acquisition of dynamic scenes that are well approximated as forming a low-rank matrix. In a companion paper [19], we develop algorithms for Fourier ptychography for *static* scenes that obey intra-frame modeling assumptions such as sparsity and/or structured sparsity. While the undersampling strategies in both cases are similar, the associated reconstruction algorithms are very different. See [19] for details.

1.3. Prior Work

There exists a rich body of literature on phase retrieval, dating back to the optical imaging work of [20, 21, 22, 23].

There is also a large (and newer) body of work on low-rank models for video acquisition and analysis [12, 13, 14]. Our recent work [17] is the first to leverage low-rank models for (generalized) phase retrieval. However, the algorithms introduced in that work are designed to succeed in the case of i.i.d. Gaussian measurements, and the applicability of that technique to other families of measurements has not been explored.

On the other hand, the literature on Fourier ptychography has mainly focused on achieving improved resolution of the reconstructed images [9, 10, 24], and relatively little attention has been given to the issue of (potentially) large measurement complexity. As discussed above, this issue is particularly acute in the case of dynamic (video) imaging. Below, we discuss a ptychographic framework that integrates low-rank models in the reconstruction step to achieve considerably reduced measurement rates, at little or no reduction of image reconstruction quality.

2. DATA ACQUISITION SETUP

2.1. Optical setup

We describe the standard Fourier ptychographic measurement procedure, following the setup of [11]. The illumination field emerging from the object can be treated as a 2D signal (or image). Using a Fraunhofer approximation, the field at the aperture plane can be expressed as the Fourier transform of the spatial signal. The light passes through a lens (having a limited-size aperture) and propagates to the plane which is occupied by the image sensor. This propagation procedure can be modeled as an inverse Fourier transform. The image sensor can only capture the magnitude of the light field at the sensor plane, and discards its phase. To overcome the diffraction limit, the authors of [11] advocate using an array of cameras placed at different locations in the sensor plane. Optically, this corresponds to measuring different (overlapping) regions of the Fourier domain signal; typically, the camera locations are chosen as a square grid, and their combined effect is to simulate a large *synthetic aperture* which can potentially yield a high resolution image.

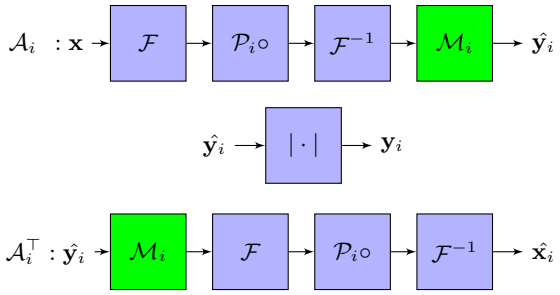

Fig. 1. Sequence of operations defined by $\mathcal{A}_i$. Here the green box indicates the extra sub-sampling step and $i = [N]$ denotes the camera index.

2.2. Mathematical model

We now represent the optical setup mathematically. Consider a (high resolution) spatial light field at time instant k, denoted by $\boldsymbol{x}_k$. We define the video matrix $\boldsymbol{X}$, composed of frames $\boldsymbol{x}_k$ for $k \in [1, \ldots, q]$:

$$\boldsymbol{X} := [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_q],\ \boldsymbol{X} \in \mathbb{R}^{n \times q}.$$

We assume that the dynamics of the video are sufficiently slow, and hence the rank of matrix is $\boldsymbol{X}$ is no greater than r, with $r \ll \min(n, q)$. For each video frame $\boldsymbol{x}_k$, the ptychographic measurements corresponding to the i^{th} camera location takes the following form:

$$\boldsymbol{y}_{i,k} = |\mathcal{A}_{i,k}(\boldsymbol{x}_k)|.$$

Here, we introduce the operator $\mathcal{A}_{i,k}$, where the index $i \in [1, \ldots, N]$ corresponds to different camera positions, and the index $k = 1, 2, \ldots, q$ indicates the time stamp. Collectively, all measurements of a single frame $\boldsymbol{x}_k$ can be expressed in terms of the cumulative measurement vector $\boldsymbol{y}_k$, defined as follows:

$$\boldsymbol{y}_k = \begin{bmatrix} \mathcal{A}_{1,k}(\boldsymbol{x}_k) \\ \mathcal{A}_{2,k}(\boldsymbol{x}_k) \\ \vdots \\ \mathcal{A}_{N,k}((\boldsymbol{x}_k) \end{bmatrix}.$$

In terms of various stages of the data acquisition process, the operator $\mathcal{A}_{i,k}$ can be expressed as:

$$\mathcal{A}_{i,k}(\cdot) = \mathcal{M}_{i,k}\mathcal{F}^{-1}\mathcal{P}_i \circ \mathcal{F}(\cdot). \tag{1}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier and inverse Fourier transforms, and $\mathcal{P}_i$ is a pupil mask corresponding to the i^{th} camera. A stage-wise illustration of this setup is shown in Fig. 1.

In addition, since our ultimate goal is to reduce the measurement complexity of the overall procedure, we suppose that not all observations are recorded at the sensor plane. This can be modeled via an "under-sampling" operator $\mathcal{M}_{i,k}$ which is applied subsequent to the inverse Fourier transform. In the standard ptychographic setup, this corresponds to an identity operation; however, we allow for more general operators $\mathcal{M}_{i,k}$.

Under this premise, we can modify the conventional measurement setup to benefit from the low-rank structure of the video to be acquired. This can be in terms of utilizing a subset of all measurements (hence, reducing the number of samples required for successful recovery). This subset can be chosen in different ways, which we will discuss in detail in Section 4.

3. RECONSTRUCTION ALGORITHM: LR-PTYCH

We develop a reconstruction method that exploits the assumption that a sequence of slowly changing images is often well approximated by a low rank matrix (with each column of the matrix being one image arranged as a 1D vector). For real videos, this means that the first few singular values of $\boldsymbol{X}$ contain most of the energy. Under this assumption, the desired $\boldsymbol{X}$ will be the solution to the non-convex optimization problem:

$$\arg\min_{\boldsymbol{X}} \sum_{k=1}^{q}\sum_{i=1}^{N} \|\boldsymbol{y}_{i,k} - |\mathcal{A}_{i,k}(\boldsymbol{x}_k)|\|_2^2, \quad \text{s.t.} \quad \text{rank}(\boldsymbol{X}) = r. \tag{2}$$

To solve (2), we adapt the *low-rank phase retrieval* (LRPR) algorithm in [17]. Our algorithm relies on the fact that a rank-r matrix $\boldsymbol{X}^*$ can be written as $\boldsymbol{X}^* = \boldsymbol{U}\boldsymbol{B}$, where $\boldsymbol{U}$ is a matrix of size $n \times r$ with mutually orthonormal columns, and $\boldsymbol{B}$ is a matrix of size $r \times q$.

The original LRPR algorithm used a spectral initialization approach that was a modification of the ideas in [25, 2]. However, directly borrowing the approach of LRPR does not work for the ptychographic application. LRPR suggests using the top eigenvector of:

$$\boldsymbol{M} = \frac{1}{nNq}\sum_{k=1}^{q}\sum_{i=1}^{N} \boldsymbol{A}_{i,k}^{\top}\text{diag}(\boldsymbol{y}_{i,k}^2)\boldsymbol{A}_{i,k}$$

(with $\mathbf{A}_{i,k} = \mathcal{A}_{i,k}(\mathbb{I})$) defined as suggested in [17] However, in our present setup, $\boldsymbol{M}$ does not have a clear separation between the r-th and the $(r+1)$-th eigenvalues.

Instead, we use a modification of the initialization idea suggested in [11]. We use $\sqrt{\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_{i,k}^2}$ as the initial guess of image frame $\boldsymbol{x}_k$, followed by computing a rank r approximation of the resulting matrix and using its components to initialize $\boldsymbol{U}$ and $\boldsymbol{B}$. Please see the first five lines of Algorithm 1. This approach relies on the following intuition. If the measurements were not phaseless, then $\boldsymbol{y}_{i,k}$ would contain random samples of a bandpass filtered version of the signal (with different i's corresponding to different random samples of different bands). Hence summing (or averaging) all the $\boldsymbol{y}_{i,k}$'s, would provide a good initial estimate of the $\boldsymbol{x}_k$.

The same would also be true if the operation before the step of taking phaseless measurements returned a vector with all non-negative entries. In our setting, neither is exactly true, however the same idea still returns a good enough initial estimate. We believe the reason is that the image itself is all non-negative and hence its low-pass filtered measurements are definitely all non-negative as well. These likely dominate in the summation, and because of this, the same approach works even though we are often removing the sign of negative entries as well (the higher frequency entries can be negative). Also, experimentally we observe that, instead of averaging, taking the root mean squared estimate yields a better initial estimate. This is likely because the large (low pass) entries dominate even more in this estimate than in a simple average. In ongoing work, we are working on developing a formal proof of correctness that relies on this intuition.

Once we obtain an initial estimate, we then refine it using a procedure similar to the LRPR2 algorithm of [17], which is an alternating-minimization algorithm that alternates between three steps: estimating the phase of the measurements, and the components $\boldsymbol{U}$ and $\boldsymbol{B}$ of the low rank matrix $\boldsymbol{X}$. The complete algorithm is summarized in Algorithm 1. For the least-squares step (step (b) of Alg. 1), we use the conjugate gradient (CG) method to obtain a fast, approximate solution, and thus avoid any need for explicit matrix inversions.

Algorithm 1 Low Rank Ptychography (LR-Ptych)

1: **for** $k = 1, 2, \ldots, q$ **do**
2: $\quad \boldsymbol{x}_k^0 \leftarrow \frac{1}{L}\sum_{i=1}^{L}\boldsymbol{y}_{i,k}$
3: **end for**
4: $[\boldsymbol{U}^0, \boldsymbol{S}^0, \boldsymbol{V}^0] \leftarrow SVD((\boldsymbol{X}^0))$
5: **for** $k = 1, 2, \ldots, q$ **do**
6: $\quad \boldsymbol{b}_k^0 \leftarrow (\boldsymbol{S}^0\boldsymbol{V}^{0\top})_k$
7: **end for**
8: **for** $t = 1, 2, \ldots, T$ **do**
9: $\quad$ a) $\boldsymbol{C}_k^t \leftarrow \text{diag}(phase(\mathcal{A}_k(\boldsymbol{U}^{t-1}\boldsymbol{b}_k^{t-1}))), k = 1, 2, \ldots, q$
10: $\quad$ b) $\boldsymbol{U}^{tmp} \leftarrow \arg\min_{\tilde{\boldsymbol{U}}} \sum_k \left\|\boldsymbol{C}_k^t\boldsymbol{y}_k - \mathcal{A}_k(\tilde{\boldsymbol{U}}\boldsymbol{b}_k^{t-1})\right\|^2$
11: $\quad$ c) $\boldsymbol{U}^t \leftarrow QR(\boldsymbol{U}^{tmp})$
12: $\quad$ d) $\boldsymbol{b}_k^t \leftarrow \arg\min_{\tilde{\boldsymbol{b}}_k} \left\|\boldsymbol{C}_k^t\boldsymbol{y}_k - \mathcal{A}_k(\boldsymbol{U}^t\tilde{\boldsymbol{b}}_k)\right\|^2, \quad k = 1, 2, \ldots, q$
13: **end for**
14: **Output** $\boldsymbol{x}_k^* = \boldsymbol{U}^T\boldsymbol{b}_k^T$

4. EXPERIMENTAL RESULTS

We study the performance of our LR-Ptych approach using two different under-sampling scenarios, which we explain below.

4.1. Uniform random undersampling

In our first scenario, we apply an under-sampling mask $\mathcal{M}_i$ on the measurements. Here the under-sampling mask consists of $1s$ and $0s$ sampled according to a Bernoulli distribution with probability of 1 equal to f. This operation is equivalent to zeroing out a large random fraction of pixels observed by each camera.

Thus for an input $v \in \mathbb{C}^n$, the sub-sampling mask operates as

$$\mathcal{M}_{i,k}(\mathbf{v})_j = \begin{cases} 0 & u_j^i > f, \\ v_j & u_j^i < f. \end{cases} \tag{3}$$

where each u_j^i is independently drawn from the uniform distribution on $[0, 1]$. We compare the performance of LR-Ptych algorithm with the Iterative Error Reduction Algorithm (IERA) [11] on several real videos which are assumed to be approximately low rank at different under-sampling ratios.

In this experiment, eight different real videos, each resized to$(180 \times 180 \times q)$ were used, q being number of frames. Other parameters of the ptychographic measurement setup are as following: diameter of aperture is 40, fraction of overlap between adjacent cameras is 0.48, and the number of cameras in an array is 81 (9×9). Thus, if all measurements are considered $(f = 1)$, then $nN = 180 \times 180 \times 81 = 2624400$. We analyze the performance of all algorithms at different sampling rates of $f = 0.75, 0.5, 0.25, 0.05$.

To compare the quality of recovery, we need a metric to describe the difference between recovered video $\boldsymbol{X}^*$ and the ground truth $\boldsymbol{X}$. We use the Structural SIMilarity index (SSIM) for all comparisons as it relies on calculating cross-correlations instead of just a distance measure [26]. However, SSIM can be only used on images (in this case, individual frames of the video). For the whole video, we use the average of the SSIMs corresponding to all the frames of $\boldsymbol{X}^*$.

A total of 8 videos were used for this experiment which we call "Fish", "Fish1", "Dog", "Dog1", and "Mouse". For all videos we assumed that the rank $r = 20$. We verified that this was indeed a valid approximation for most vidoes; also it results in manageable

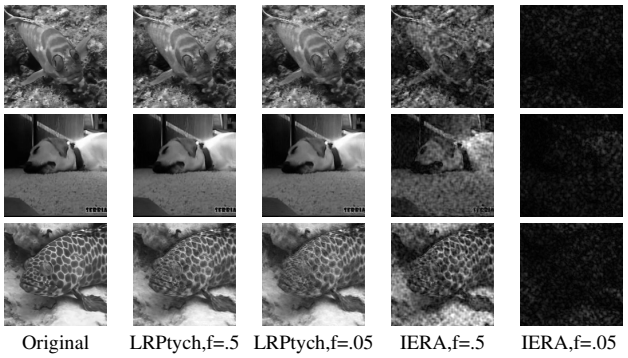

Fig. 2. Comparison of frame 66 of different videos under different algorithms and different sampling rates. Time taken in seconds for "Dog1" video is '8543', '8607' for LRPRtych, and '181', '188', for IERA, respectively with $f = 0.5$, and $f = 0.25$.

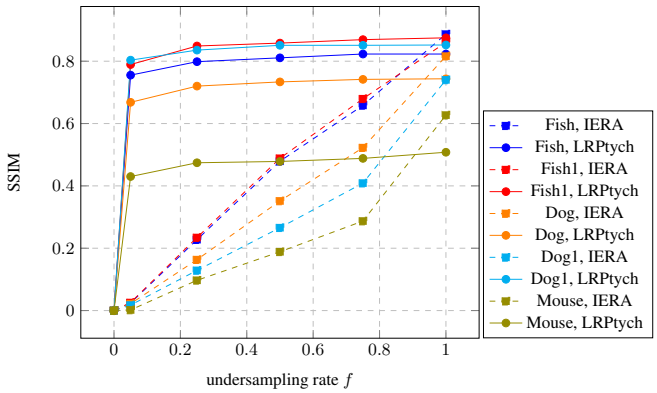

Fig. 3. SSIM of four different *slow changing* videos and one without *slow changing*, at different sampling rates.

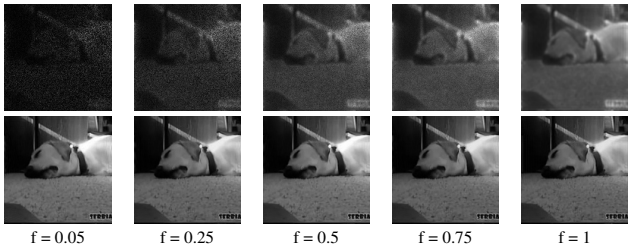

Fig. 4. Visual comparison for random pixel undersampling of frame number 66. First row shows the results for LRPR1, and the second row shows results for the proposed method.

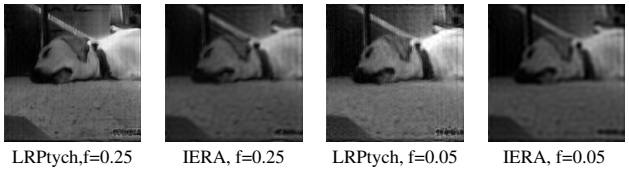

Fig. 5. Visual comparison for random camera undersampling for frame number 66.

speed of the resulting algorithm. In Figure 3, we plot the SSIM computed for all these videos as a function of the under-sampling rate f. We also show the visual comparisons for one frame of "Fish1",

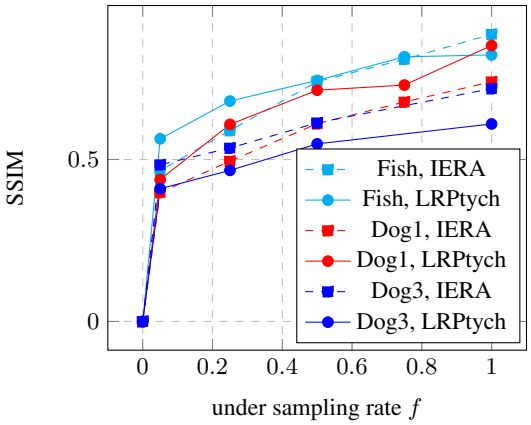

Fig. 6. Results of using random camera under-sampling pattern.

"Dog1", and "Fish" videos in Fig. 2, respectively from top. As can be seen from both figures, when fewer samples are used (f is smaller), the proposed algorithm (LR-Ptych) significantly outperforms IERA. The reason is of course that LR-Ptych exploits the approximately low rank structure inherent in all these videos.

In another experiment we compared the results of our refinement process with that of LRPR1 in [17]. As it can be seen from the Fig. 4, low rank alternating minimization works better. Both algorithms are using the initialization step of 1, under different rates of undersampling for "Dog1" video.

4.2. Under sampling using uniform random camera pattern

In this version, we select some cameras of the camera array at random while not using the others. The sub-sampling matrix in this approach can be written as

$$\mathcal{M}_{i,k}(\mathbf{v}) = \begin{cases} \mathbf{0} & u_i > f, \\ \mathbf{v} & u_i < f. \end{cases}$$

where u_i is independent standard uniform random variables.

Figure 6 shows results for this under-sampling strategy for videos of "Fish", "Dog1". As can be seen, our algorithm again outperforms IERA, although the gain over IERA is not as much in this setting. The reason is that we have less diversity of measurements here. A visual example for this experiment is shown in Figure 5, which is frame number 66 of "Dog1" video. It can be observed that our algorithm is capable of recovering more details, in comparison to IERA.

5. CONCLUSIONS AND FUTURE WORK

In conclusion, we demonstrate that we can acquire slowly changing low-rank videos from a pytchographic setup, using much fewer samples, as compared to conventional imaging techniques. The quality of such reconstruction has been shown to be much superior and also translates to lower operational costs and imaging time.

As future work, we aim to (i) design a more principled initialization strategy, (ii) establish a formal bound on the sample complexity of low-rank phase retrieval algorithms for pytchography and coherent camera array imaging, and (iii) accelerate the algorithm to be applicable to higher resolution video sequences.

6. REFERENCES

[1] E. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," vol. 66, no. 8, pp. 1241–1274, 2013.

[2] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[3] Yuxin Chen and Emmanuel Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Advances in Neural Information Processing Systems*, 2015, pp. 739–747.

[4] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," 2013, pp. 2796–2804.

[5] Gang Wang, Georgios B Giannakis, and Yonina C Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Transactions on Information Theory*, 2017.

[6] Yoav Shechtman, Yonina C Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[7] R. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.

[8] A. Maiden and J. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009.

[9] Siyuan Dong, Roarke Horstmeyer, Radhika Shiradkar, Kaikai Guo, Xiaoze Ou, Zichao Bian, Huolin Xin, and Guoan Zheng, "Aperture-scanning fourier ptychography for 3d refocusing and super-resolution macroscopic imaging," *Optics express*, vol. 22, no. 11, pp. 13586–13599, 2014.

[10] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang, "Wide-field, high-resolution fourier ptychographic microscopy," *Nature photonics*, vol. 7, no. 9, pp. 739–745, 2013.

[11] Jason Holloway, M Salman Asif, Manoj Kumar Sharma, Nathan Matsuda, Roarke Horstmeyer, Oliver Cossairt, and Ashok Veeraraghavan, "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 251–265, 2016.

[12] Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi, "Low-rank matrix completion using alternating minimization," in *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*. ACM, 2013, pp. 665–674.

[13] Emmanuel J Candès and Benjamin Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, pp. 717, 2009.

[14] John Wright, Arvind Ganesh, Shankar Rao, Yigang Peng, and Yi Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in *Advances in neural information processing systems*, 2009, pp. 2080–2088.

[15] Samet Oymak, Amin Jalali, Maryam Fazel, Yonina C Eldar, and Babak Hassibi, "Simultaneously structured models with application to sparse and low-rank matrices," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2886–2908, 2015.

[16] Yonina C Eldar, Deanna Needell, and Yaniv Plan, "Uniqueness conditions for low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 33, no. 2, pp. 309–314, 2012.

[17] Namrata Vaswani, Seyedehsara Nayer, and Yonina C Eldar, "Low-rank phase retrieval," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 4059–4074, 2016.

[18] E. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," vol. 39, no. 2, pp. 277–299, 2015.

[19] G. Jagatap, Z. Chen, C. Hegde, and N. Vaswani, "Sub-diffraction imaging using fourier ptychography and structured sparsity," 2017, Project website: https://gaurijagatap.github.io/Sparse-image-super-resolution/.

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

[21] R. Gerchberg and W. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures.," *Optik*, vol. 35, no. 237, 1972.

[22] S. Marchesini, "Phase retrieval and saddle-point optimization," *JOSA A*, vol. 24, no. 10, pp. 3289–3296, 2007.

[23] K. Nugent, A. Peele, H. Chapman, and A. Mancuso, "Unique phase recovery for nonperiodic objects," *Physical review letters*, vol. 91, no. 20, pp. 203902, 2003.

[24] A. Maiden, M. Humphry, F. Zhang, and J. Rodenburg, "Superresolution imaging via ptychography," *J. Opt. Soc. Am. A*, vol. 28, no. 4, pp. 604–612, Apr 2011.

[25] Praneeth Netrapalli, Prateek Jain, and Sujay Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 2796–2804.

[26] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.