Accelerating ptychographic reconstructions using spectral initializations

LORENZO VALZANIA,1,2,*,† JONATHAN DONG,1,3,† AND SYLVAIN GIGAN1

1Laboratoire Kastler Brossel, École Normale Supérieure - Paris Sciences et Lettres (PSL) Research University, Sorbonne Université, Centre National de la Recherche Scientifique (CNRS) UMR 8552, Collège de France, 24 rue Lhomond, 75005 Paris, France
2Laboratory for Transport at Nanoscale Interfaces, Empa, Swiss Federal Laboratories for Materials Science and Technology, 129 Uberlandstrasse, Dubendorf 8600, Switzerland
3Laboratoire de Physique de l’École Normale Supérieure - Paris Sciences et Lettres (PSL) Research University, Sorbonne Université, CNRS, Université de Paris, 24 rue Lhomond, 75005 Paris, France

*lorenzo.valzania@lkb.ens.fr
†These authors equally contributed to this work.

Abstract: Ptychography is a promising phase retrieval technique for label-free quantitative phase imaging. In particular, the combination of a synthetic aperture approach with a reference-less setup makes it appealing to the biomedical community. Recent advancements in phase retrieval witnessed the development of spectral methods, in order to accelerate gradient descent algorithms. Using spectral initializations on both simulated and experimental data, for the first time we report 3 times faster ptychographic reconstructions than with a standard gradient descent algorithm. Spectral methods represent a paradigmatic change to develop new theoretical insights as well as experimental implementations of ptychography.

© 2020 Optical Society of America

Ptychography is a computational imaging technique that enables label-free, quantitative phase imaging [1]. It is based on a simple principle: scan a probe across a sample, collect the corresponding intensity diffraction patterns (also known as ‘ptychograms’), and reconstruct an image of the object of interest. Because it does not require complex optical elements, it has been adapted to a variety of spectral ranges, from X-rays [2] and extreme ultra-violet light [3] all the way down to the terahertz (THz) range [4], and to electron microscopy [5] for which it was originally conceived in the late 60s [6]. Fourier ptychography, a closely related approach, has been developed mostly in the visible range, as a way to increase the resolution and field-of-view of optical microscopes [7, 8].

The computational reconstruction in ptychography requires to solve a phase retrieval problem. From intensity-only measurements, one needs to retrieve the phase of the electric field at the object plane. Several algorithms have been proposed to solve this problem. Alternating projections algorithms, delivering a new estimate of the solution every time a ptychogram is used, are the most common ones. For example, all the various implementations of the ptychographic iterative engine (PIE) [9] belong to this family. Gradient-descent-based algorithms, using the whole dataset at each iteration, have recently been demonstrated to be more robust against noise [10–12]. Because two subsequent probe positions present an overlap (typically larger than 60%), this redundancy helps make the phase retrieval tractable [13]. Despite all the advancements, such a non-linear optimization problem is still not completely understood and convergence of these algorithms towards the solution is not guaranteed.

To avoid local minima in a non-convex optimization problem, one solution is to provide an initial estimate already close to the solution. Spectral methods have been proposed for the general phase retrieval problem [14]: the initial estimate is defined as the leading eigenvector of a particular matrix, and this spectral initialization will then be followed by iterative algorithms. Important
theoretical breakthroughs by [15, 16] show that for random and independent measurements one particular spectral method is optimal. Interestingly, Marchesini and colleagues also proposed a similar spectral initialization for ptychography [17], supporting their conclusions with simulated data. However, when spectral methods were tried on experimental ptychography [18] and Fourier ptychography [19, 20] datasets, no gain was observed.

In this work, we prove with both numerical and experimental results that spectral initializations can be useful for ptychography. We show that the spectral method of [17] is a particular case of the general spectral methods studied in [15, 16]. This allows us to propose a novel spectral method that provides a better initial estimate, that will then be used for gradient descent. We show that although spectral methods are not perfect and present reconstruction artefacts, especially when measurements are neither random nor independent, they are valuable tools to make the ptychographic reconstruction faster and more robust to noise. It is indeed the first time that the theoretical works from [15, 16] have been applied to a completely non-random setting.

Continuous-domain model of ptychography

A ptychogram $y^{(l)}(x)$ is typically described by the following forward model

$$y^{(l)}(x) = |P_d(a(x - x^{(l)})\psi(x))|^2,$$

where $x$ is a two-dimensional spatial coordinate, $\psi(x)$ is the complex transmission function of the object, scanned with the illumination denoted by $a(x)$ at the positions $x^{(l)}$ for $l = 1, \ldots, L$, with $L$ being the total number of images. $P_d$ is a known linear operator describing the transmission through the optical system of optical length $d$ comprised between the object plane and the detector plane. We hereby point out that the model equally applies to Fourier ptychography [8], if one regards $\psi(x)$ and $a(x)$ as the Fourier transform of the object transmission and the illumination functions, respectively, and $P_d$ is an inverse Fourier transform.

Therefore, without loss of generality, in what follows the case of ptychography will be considered, with a setup shown in Fig. 1. Usually, one defines $P_d$ as a Fourier transform since the camera is placed in the far-field of the object, but in our setting the Rayleigh-Sommerfeld convolution integral [21] is more suitable, i.e. $P_d(f(x)) \equiv f(x) * h_d(x)$, with $h_d(x) \equiv (|d|/i\lambda) \exp[2\pi i \text{sgn}(d) r/\lambda r^2]$, $r \equiv (||x||^2 + d^2)^{1/2}$ and where $\lambda$ is the wavelength, $\text{sgn}(\cdot)$ is the sign function and $d > (<) 0$ is the propagation distance in the forward (resp. backward) direction.

![Fig. 1. Sketch of a ptychography setup. An aperture generates a simple Airy-disk probe $a(x)$, that is scanned with shifts $x^{(l)}$ across an object of interest $\psi(x)$, for $l = 1, \ldots, L$. For each probe position, a camera records an intensity image $y^{(l)}(x)$ after free space propagation of a distance $d$.](image)

Spectral methods
To recover the object, ptychography requires to solve a phase retrieval problem, that can be more generally formulated as:

\[ y = |S\psi|^2, \]

(2)

where \( \psi(x) \in \mathbb{C}^n \) and \( S \in \mathbb{C}^{p \times n} \) is the sampling matrix, with \( n \) the number of parameters to estimate and \( p \) the number of intensity measurements. Spectral methods have been proposed as a computationally efficient technique to find an approximate solution to the phase retrieval problem. They consider the following matrix (where \( S^\dagger \) denotes the Hermitian conjugate of the matrix \( S \)):

\[ Z = \frac{1}{L} S^\dagger \text{diag}(T(y)) S, \]

(3)

where \( T(\cdot) \) is a preprocessing function acting on the measurements. The leading eigenvector of \( Z \) will converge towards the solution \( \psi \) provided the oversampling (or overlap in ptychography) is high enough. Theoretical works have provided a considerable understanding of these spectral methods for the random setting, i.e. when the sampling matrix \( S \) is random. For example, rigorous convergence proofs can be written for any increasing preprocessing function bounded above, as proven in [22]. Indeed over the years, spectral methods based on different preprocessing functions have been proposed, \( T_1 \) in [14], \( T_2 \) in [17], and the optimal preprocessing function in the noiseless setting \( T^* \) in [16], defined as:

\[
T_1(y) = y \\
T_2(y) = \begin{cases} 
0 & \text{if } y \leq T \\
1 & \text{if } y > T 
\end{cases} \\
T^*(y) = 1 - \frac{1}{y},
\]

(4)

with \( T \) being a threshold intensity value. To compute the leading eigenvector efficiently, we use power iterations, that converge exponentially towards the largest eigenvalue in absolute value.

Although spectral methods provide an efficient way to have an approximate solution to a phase retrieval problem, this solution can further be refined by iterative techniques. We chose here to use gradient descent with the quadratic loss:

\[ l(\psi) = \|y - |S\psi|^2\|^2 \]

(5)

Each iteration uses the whole set of images, which makes it more robust against noise than alternating projections using images one at a time [10, 12].

**Discrete and vectorized model of ptychography**

In order to apply spectral methods to ptychography, the forward model in the continuous domain of Eq. (1) needs to be brought into the discrete vectorized form of Eq. (2). Vectorized quantities are indicated with bold symbols of the corresponding quantities in Eq. (1). Therefore, let \( \psi \in \mathbb{C}^n \) be a column vector representing the object complex transmission function, \( y \equiv (y_k) \in \mathbb{R}^{L_m} \) be the stack of measured intensities, where \( m \) is the number of camera pixels. Let also \( a^{(l)} \in \mathbb{C}^{m \times n} \) be the discrete and vectorized version of \( a(x - x^{(l)}) \) and \( P_d \in \mathbb{C}^{m \times m} \) be the matrix describing the linear transform of the Rayleigh-Sommerfeld integral.

Now we can adopt the matrix factorization already used in [17, 23] and finally write \( y = |S\psi|^2 \), with \( S \equiv PA \in \mathbb{C}^{L_m \times n} \) and

\[
P \equiv \begin{bmatrix} P_d & \cdots & 0 \\
0 & \ddots & 0 \\
0 & \cdots & P_d 
\end{bmatrix} \in \mathbb{C}^{L_m \times L_m}, \quad A \equiv \begin{bmatrix} a^{(1)} \\
\vdots \\
a^{(L)} 
\end{bmatrix} \in \mathbb{C}^{L_m \times n},
\]

(6)

from which, using Eq. (3) and the unitarity of \( P \) (i.e., \( P^\dagger = P^{-1} \)), one obtains

\[ Z = A^\dagger P^{-1} \text{diag}(T(y)) PA. \]

(7)
The overall algorithm, consisting of the combination of a spectral initialization and a gradient
descent, is summarized in the box Algorithm 1. Code is available at the following address:
https://github.com/laboGigan/ptychography-spectral.

The previous matrices are very large, for example $P$ is of size $Lm \times Lm$, with $Lm$ the total
number of measured pixels. As typically $Lm \gg 10^4$, we want to avoid the explicit computation
of $Z$ to retrieve its leading eigenvector. Instead, we perform power iterations using $L$ partial
estimates $\psi^{(t)}_{t+1}(x)$:

$$
\psi^{(t)}_{t+1}(x) = \tilde{a}(x-x^{(t)})P_{-d}\{T(v^{(t)}(x))P_d\{a(x-x^{(t)})(x))\},
$$

where $\tilde{a}(x-x^{(t)})$ denotes the complex conjugate of $a(x-x^{(t)})$. The estimate $\psi_{t+1}(x)$ at the
$(t+1)$-th iteration is obtained by stitching the partial estimates, shifted at the corresponding scan
positions.

**Algorithm 1** Solve ptychography with a spectral initialization

**Require:** Measurements $y$, scan parameters $A$, block diagonal matrix $P$ of linear operators,
preprocessing function $T(\cdot)$, number of diffraction patterns $L$, number of power iterations
$M$, number of gradient descent iterations $N$, step size $\gamma$

1: function PTYCHOGRAPHY WITH SPECTRAL INITIALIZATION($y$, $A$, $P$, $T$, $L$, $M$, $N$, $\gamma$)
2: $S \leftarrow PA$
3: $Z \leftarrow L^{-1}S^* \text{diag}\{T(y)\}S$
4: $\psi_0 \leftarrow $ random
5: for $t \leftarrow 1$ to $M$ do $\triangleright$ Spectral initialization through power iterations
6: $\psi_t \leftarrow Z\psi_{t-1}$
7: $\psi_t \leftarrow \psi_t/||\psi_t||$
8: end for
9: for $t \leftarrow M + 1$ to $M + N$ do $\triangleright$ Gradient descent
10: $\nabla l(\psi) \leftarrow S^T[[|S\psi_{t-1}|^2 - y] \odot (S\psi_{t-1})]
11: $\psi_t \leftarrow \psi_{t-1} - \gamma \nabla l(\psi)$
12: end for
13: end function

Experiments were performed using the THz imaging setup at Empa, the Swiss Federal
Laboratories for Materials Science and Technology, equipped with a far-infrared gas laser (FIRL
100, Edinburgh Instruments, Livingston, Scotland) emitting several tens of mW of continuous
wave power at $\lambda = 96.5$ μm. An uncooled microbolometer array detector featuring $m = 480 \times
640$ pixels on a pitch of 17 μm (Gobi-640-GigE, Xenics, Leuven, Belgium), originally designed
for infrared radiation but well performing at longer wavelengths too [24], was used as a THz
camera. The object was a metallic resolution test target in the shape of a nine-spoked Siemens
star (Fig. 1), acting as an amplitude binary object and whose transmission function is shown in
Fig. 2(a). We let a plane wave diffract through a circular aperture with a diameter of 3 mm and
propagate by 7 mm before impinging on the object, yielding an Airy disk with NA = 0.21. The
scan of the object was performed across a square grid of $11 \times 11$ points at a relative overlap of
77%, according to the definition given by Bunk et al. [13]. The ptychograms were recorded after
a free space propagation of 6 mm, corresponding to a numerical aperture of about 0.6.

Results from simulated as well as experimental data are plotted in Fig. 2. We initiate our
reconstructions with values drawn from a standard complex Gaussian random distribution (whose
modulus is shown in Fig. 2(b)). This way, we ensure a negligible correlation between the ground
truth and the initial estimate. The most striking effect of the spectral initialization is apparent
at the earliest iterations. The power method yields the reconstructions shown in Figs. 2(c, e)
after 10 and 3 iterations on the simulated and experimental dataset, respectively. The same number of iterations, used in a gradient descent scheme, delivers much less informative results (Figs. 2(d, f)), especially when using experimental data. The reconstructions in Figs. 2(c-f) are further refined with a gradient descent algorithm until the solutions obtained with and without spectral initialization reach a comparable quality (Figs. 2(g-l)). We realized that starting the reconstruction with a spectral method has a two-fold relevance. First, the final reconstruction is obtained with one third of the iterations needed without spectral initialization. Second, we achieve results from experimental data that are more resilient to noise. The artifacts observed in Figs. 2(j, l) are introduced by the gradient descent algorithm which, when run for too many iterations, is known to deteriorate the quality of the reconstruction [10]. We indeed observed the same drift to noisy reconstructions starting with a spectral initialization, too, however at a number of iterations $\sim 2.5$ times larger than for standard gradient descent.

In Fig. 3 we study in more detail the different spectral methods, showing the spectral initializations for various choices of the preprocessing function $T(\cdot)$, the number of power iterations $M$, and the spatial distribution of the illumination function $a(x)$. We have seen

Fig. 2. Comparison between a ptychographic reconstruction with and without a spectral initialization. (a) Simulated object; (b) Modulus of the random initial estimate used as an input to the power iteration method (top row) and the gradient descent (bottom row); (c) [(e)] spectral initialization from simulated [experimental] data; (d) [(f)] solution via gradient descent run for the same number of iterations as the power method; (g) [(i, k)] final reconstruction via spectral initialization and gradient descent; (h) [(j, l)] final reconstruction via gradient descent only. Scale bars: 1 mm ($\sim 10\lambda$).

Fig. 3. Spectral initializations at different numbers of power iterations and different preprocessing and illumination functions (simulated data from the object shown in Fig. 2(a)). Note that, in order for all the images to share the same colorbar, shown on the right, (a) and (b) have been multiplied by 5 and 200, respectively. Scale bar: 1 mm ($\sim 10\lambda$).
empirically that these three parameters are impacting the most the spectral method performance.

Note that when the illumination function is the Airy disk used in Fig. 2 and Figs. 3(a-f), the hypotheses of random sampling required by the spectral method [14] break down. For this reason, even in a noiseless setting like the one presented in Fig. 3, we cannot expect to indefinitely approach the solution using power iterations, as confirmed by Figs. 3(a, b). However, the preprocessing functions $T_2$ and $T^*$ make the estimates more robust to the power iterations, while delivering informative estimates already after 10 power iterations. Therefore, the estimate in Fig. 3(e) was used to run the simulations of Fig. 2, and the corresponding preprocessing function $T^*$ was also employed on the experimental dataset.

In Figs. 3(g, h), we simulated a speckle beam with a grain size about 10 times smaller than the illumination shifts, so to boost the diversity of the ptychograms upon translation. This makes the acquisition closer to that of coded diffraction imaging, for which spectral methods were originally developed [14,15]. As a result, a much more effective spectral initialization is obtained, with no need to preprocess the measurement data. Besides an increase of the spatial resolution, we notice a more reliable quantification of the amplitude. These results can also be seen in the context of ptychography with randomized and structured illuminations, which has already been implemented with three main advantages: adding diversity in the captured images [25,26], accessing higher spatial frequencies [27] as in coherent structured illumination microscopy [28], and reducing dynamic range requirements in X-ray imaging [29,30].

Our study can be regarded as an additional evidence that random probes are beneficial in ptychography. This paves the way for improved algorithms and randomized optics for ptychography. Moreover, we foresee that smart spectral initializations will be applied in this context, both to increase the reconstruction speed and escape local minima. This may especially arise when complicated objects have to be imaged, for example with a large field-of-view or when a reliable initial estimate is not available a priori.

In conclusion, we accelerated experimental ptychographic reconstructions through spectral initialization. Compared to a standard gradient descent implementation, three times faster reconstructions were obtained. Although spectral methods have been developed for random sampling, where they have been shown to be optimal, we investigated the most relevant reconstruction parameters to make our methods more general, and applicable even in the absence of randomness. Besides providing informative estimates of the solution from less a priori information, we believe that spectral methods represent a step forward in the development of ptychography with structured and randomized illuminations. Ours results should be applicable whatever the wave domain of interest (from electron microscopy to THz) and provide a simple and robust way to accelerate ptychographic methods.

**Funding**

LV acknowledges the financial support from the Swiss National Science Foundation (grants no. 200021_160078/1 and P2BEP2_188152). SG is funded by the H2020 European Research Council (ERC, H2020, SMARTIES-724473).

**Acknowledgments**

LV wishes to thank E. Hack and P. Zolliker (Empa) for contributing constructive comments during the data acquisition.

**Disclosures**

The authors declare no conflicts of interest.
