



**HAL**  
open science

## Polarimetric Fourier phase retrieval

Julien Flamant, Konstantin Usevich, Marianne Clausel, David Brie

► **To cite this version:**

Julien Flamant, Konstantin Usevich, Marianne Clausel, David Brie. Polarimetric Fourier phase retrieval. 2022. hal-03613352v3

**HAL Id: hal-03613352**

**<https://hal.science/hal-03613352v3>**

Preprint submitted on 7 Apr 2023 (v3), last revised 11 Mar 2024 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Polarimetric Fourier phase retrieval\*

Julien Flamant<sup>†</sup>, Konstantin Usevich<sup>†</sup>, Marianne Clausel<sup>‡</sup>, and David Brie<sup>†</sup>

**Abstract.** This work introduces *polarimetric Fourier phase retrieval* (PPR), a physically-inspired model to leverage polarization of light information in Fourier phase retrieval problems. We provide a complete characterization of its uniqueness properties by unraveling equivalencies with two related problems, namely bivariate phase retrieval and a polynomial autocorrelation factorization problem. In particular, we show that the problem admits a unique solution, which can be formulated as a greatest common divisor (GCD) of measurements polynomials. As a result, we propose algebraic solutions for PPR based on approximate GCD computations using the null-space properties Sylvester matrices. Alternatively, existing iterative algorithms for phase retrieval, semidefinite positive relaxation and Wirtinger-Flow, are carefully adapted to solve the PPR problem. Finally, a set of numerical experiments permits a detailed assessment of the numerical behavior and relative performances of each proposed reconstruction strategy. They further demonstrate the fruitful combination of algebraic and iterative approaches towards a scalable, computationally efficient and robust to noise reconstruction strategy for PPR.

**Key words.** Fourier phase retrieval, polarization, approximate greatest common divisor, semidefinite positive relaxation, Wirtinger Flow

**MSC codes.** ?

**1. Introduction.** The problem of Fourier phase retrieval, i.e., the recovery of a signal given the magnitude of its Fourier transform, has a long and rich history dating back from the 1950s [56]. It has been – and continues to be – of tremendous importance for many applications areas involving optics, such as crystallography [20, 21, 47], astronomy [25, 26], coherent diffraction imaging (also known as lensless imaging) [46, 44], among others. Such problem arises in optics since *phase information* of light cannot be measured directly due to the high oscillating frequency of the electromagnetic field: indeed there is no conventional detector that can sample at a rate of  $\sim 10^{12}$  Hz (infrared) up to  $\sim 10^{18}$  Hz (hard x-rays). In addition, many imaging applications rely on diffraction measurements in the far-field, where light propagation essentially acts as a Fourier transform operator of the field near the imaged object [30]. Examples include one-dimensional (1D) temporal Fourier transforms performed by spectrometers in ultra-short laser pulse characterization [68] or two-dimensional (2D) spatial Fourier transforms recorded on far-field pixelated detectors in X-ray coherent diffraction imaging [18]. These Fourier-domain detectors, together with the impossibility to measure phase information, yield phaseless Fourier intensity measurements. Therefore, reconstruction of the imaged object requires solving a Fourier phase retrieval problem. See [58] for a comprehensive overview of such problems in optical imaging.

Just like color (wavelength), *polarization* is a fundamental property of light. It encodes

---

\*Corresponding author: Julien Flamant.

**Funding:** This work was funded by CNRS and GdR ISIS OPENING exploratory research project grant.

<sup>†</sup>CNRS, Université de Lorraine, CRAN, F-54000 Nancy France ([julien.flamant@cnrs.fr](mailto:julien.flamant@cnrs.fr), [konstantin.usevich@univ-lorraine.fr](mailto:konstantin.usevich@univ-lorraine.fr), [david.brie@univ-lorraine.fr](mailto:david.brie@univ-lorraine.fr)).

<sup>‡</sup>Université de Lorraine, CNRS, IECL, F-54000 Nancy France ([marianne.clausel@univ-lorraine.fr](mailto:marianne.clausel@univ-lorraine.fr))

the geometry of oscillations of the electromagnetic field, which describes an ellipse in the 2D plane perpendicular to the propagation direction for vacuum-like media [19]. As polarized light propagates in media, its polarization can change, thus revealing key properties, such as medium anisotropy or structural properties that are inaccessible to conventional, non-polarized light [27]. As a result, polarized light imaging has found many applications such as material characterization [31], remote sensing [62] or bio-imaging [34]. Despite the important practical interests of polarization, only a few authors have considered leveraging this fundamental attribute of light in phase retrieval problems. The authors in [59, 55] pioneered the use of polarization in Fourier phase retrieval for ultrashort attosecond ( $10^{-18}$  s) laser pulse characterization. The motivation for polarimetric measurements arises from a fundamental physical limitation, which prevents the direct use of standard pulse characterization strategies based on nonlinear light-matter interaction such as Frequency-Resolved Optical Gating (FROG) [61] and its variants. Another line of work regards the extension of a scanning coherent diffraction imaging technique, known as *ptychography*, to take into account the polarization of light. This novel imaging modality, called *vectorial ptychography* [23, 24] combines spatially redundant measurements with polarimetric measurements. This allows quantitative imaging of complex anisotropic media, such as biominerals [5, 6].

*Related work.* Fourier phase retrieval is a long standing problem and therefore has generated a continuous interest from researchers of various horizons, leading to a vast literature ranging from theoretical results to practical imaging algorithms, see [12] for an overview. A recent survey of uniqueness and stability of Fourier phase retrieval can be found in [33]; see also [13] for a discussion of its algebraic properties. A comprehensive tour of existing algorithms is given in [22]; see also [4] for an extensive discussion of related geometric aspects.

One-dimensional Fourier phase retrieval does not admit a unique solution in general [10]. Therefore, many strategies to enforce uniqueness have been devised. These include additional information on the signal, such as knowledge of some entries [11], non-negativity [8], sparsity [39, 52] or minimum phase [35]. Another approach consists in generating additional measurements, e.g., using deterministic masks [36, 15], (randomly) coded diffraction patterns [16] or using redundant, overlapping measurements inspired by ptychography [14, 37].

More closely related to the present work is the use of additional, interference-like measurements in Fourier phase retrieval. The main idea roots in a imaging technique known as *holography*, which involves the coherent interference of the object of interest  $\mathbf{x}$  with some reference signal  $\mathbf{y}$ . Pushing this idea further, authors have developed a strategy ensuring uniqueness in Fourier phase retrieval, called *vectorial phase retrieval* [53] or double-blind holography [42, 54, 50]. More precisely, they show (and exploit) that almost all signals  $\mathbf{x}$  and  $\mathbf{y}$  can be recovered from four Fourier magnitudes measurements, of  $\mathbf{x}$ ,  $\mathbf{y}$ ,  $\mathbf{x} + j\mathbf{y}$  (with  $j^2 = -1$ ) and  $\mathbf{x} + \mathbf{y}$ , respectively. Similar ideas appear in [38], where the reconstruction problem is formulated using correlations functions instead of Fourier transforms.

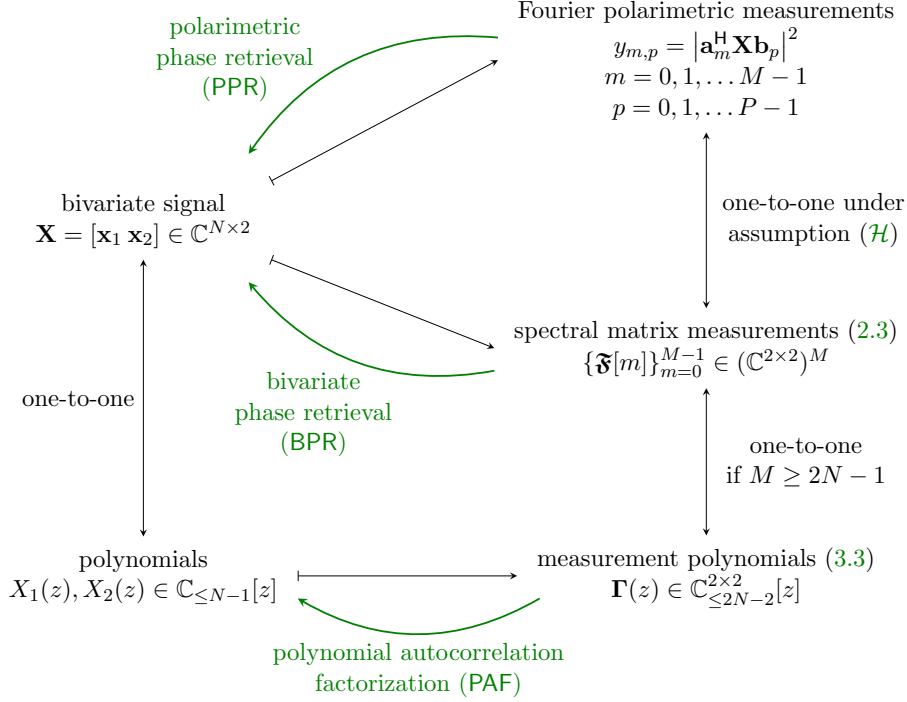
While these works share several features with the present paper, they also differ on a number of important points. First, they do not exploit a polarimetric acquisition scheme, which limits their use in contexts where one is interested in reconstructing the polarized (or bivariate) electromagnetic field (such as in polarized coherent diffraction imaging techniques [60]). In particular, we will show that the proposed polarimetric Fourier phase retrieval model encompasses vectorial phase retrieval as a special case, for a specific choice of *four*

82 *polarimetric projections*. In addition, while the connection between vectorial phase retrieval  
83 and greatest common divisor of polynomials was observed in [38], it was not investigated in  
84 detail as the authors focused on a semidefinite programming relaxation. In contrast, algebraic  
85 approaches based on greatest common divisor computations are a cornerstone of the proposed  
86 methodology for the polarimetric Fourier phase retrieval model.

87 **Contributions.** This work introduces a novel Fourier phase retrieval model, called *polarimetric Fourier phase retrieval (PPR)*, which takes advantage of the physical measurement  
88 of polarization properties in optics. In particular, measurements are readily interpreted in  
89 terms of polarimetric Fourier projections of the bivariate electromagnetic field. As such, the  
90 proposed model can be implemented using standard optical components, such as polarizers or  
91 waveplates. It is flexible: more polarimetric measurements can be performed if desired. We  
92 focus on the 1D Fourier case in this paper, as a first step to demonstrate the potential of polar-  
93 ization information in Fourier phase retrieval problems. First, we characterize its uniqueness  
94 properties by carefully establishing equivalences with two other problems, namely bivariate  
95 Fourier phase retrieval (**BPR**) and polynomial autocorrelation factorization (**PAF**). In par-  
96 ticular, we show that the **PPR** problem can be solved through algebraic methods based on  
97 approximate greatest common divisor computations. We compare in detail these approaches  
98 with tailored adaptations of standard iterative algorithms for Fourier phase retrieval, namely  
99 semidefinite positive relaxation and Wirtinger-Flow, to the case of **PPR**. Finally, numerical  
100 experiments demonstrate that combining algebraic and iterative approaches yields a scalable,  
101 computationally efficient and robust to noise reconstruction strategy for **PPR**.

103 **Organization of the paper.** A crucial feature of the present paper is the extensive use of  
104 equivalences between the polarimetric Fourier phase retrieval (**PPR**) problem and two other  
105 problems, namely bivariate Fourier phase retrieval (**BPR**) and polynomial autocorrelation fac-  
106 torization (**PAF**). For reference, these equivalences are stated in **Figure 1**, with pointers to  
107 relevant definitions and equations. **Section 2** introduces the **PPR** model and discusses its physi-  
108 cal interpretations in terms of polarimetric measurement. Under some very general conditions,  
109 the equivalence with **BPR** is then established, which permits the study of trivial ambiguities.  
110 The relation of **PPR** with a standard 1D Fourier phase retrieval problem is also discussed.  
111 **Section 3** starts by reformulating the **BPR** problem using a polynomial representation, leading  
112 to **PAF**. Then, we leverage uniqueness results on multivariate spectral representations [63]  
113 to establish a necessary and sufficient characterization of uniqueness in **PAF** (**Theorem 3.3**).  
114 **Corollary 3.4** states that **PAF** is almost everywhere unique, and as a result, an algebraic  
115 solution can be found using greatest common divisors of measurement polynomials (**Propo-**  
116 **sition 3.5**). **Section 4** goes back to **PPR** and exploits uniqueness results to propose a fully  
117 algebraic reconstruction method for **PPR** (**Algorithm 1**) based on two variations of approxi-  
118 mate greatest common divisor computations. **Section 5** focus instead on iteratives algorithms  
119 for **PPR**, by tailoring semidefinite relaxation (**Algorithm 4**) and Wirtinger Flow (**Algorithm**  
120 **5**). **Section 6** presents several numerical experiments to illustrate and assess the practical  
121 performances of the proposed reconstruction strategies. **Section 7** collects concluding remarks  
122 and Appendices gather technical details and proofs.

123 **Notations.** In this paper, we denote by  $\mathbb{R}$  the set of real numbers and by  $\mathbb{C}$  the set of  
124 complex numbers with imaginary unit  $j$  such that  $j^2 = -1$ . Vectors and matrices are denoted  
125 in bold lowercase letters and bold capital letters, respectively. Dependence of quantities in



**Figure 1.** Equivalences of data and solutions in problems PPR, BPR and PAF.

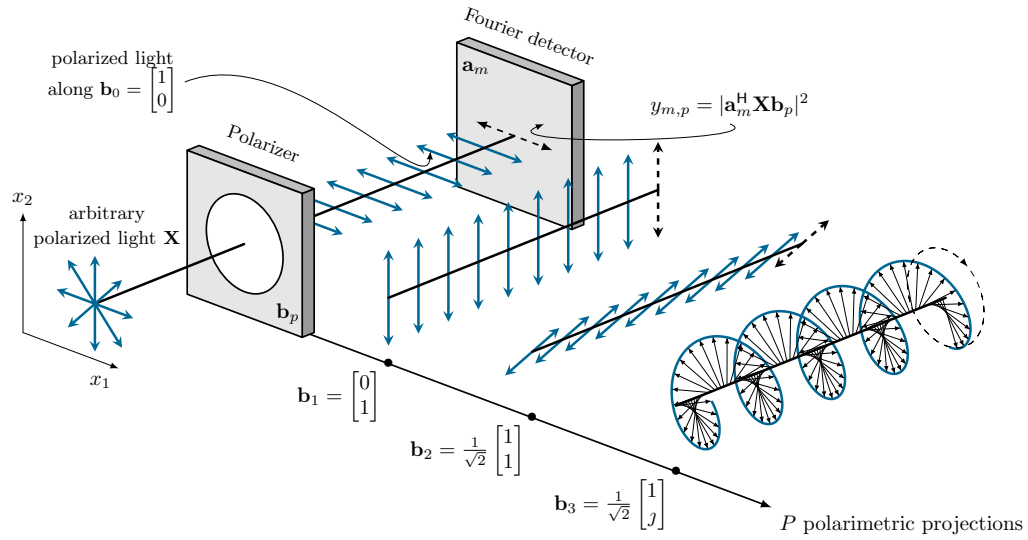
126 terms of a discrete index are indicated by brackets, i.e.,  $\mathbf{x}[n]$  denotes the  $n$ -th entry of the set  
 127 of vectors  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$ . Notation  $\mathbf{a}^*$ ,  $\mathbf{A}^*$  indicate the complex conjugate of vector  $\mathbf{a}$  and matrix  
 128  $\mathbf{A}$ , respectively. The transpose of a matrix  $\mathbf{A}$  is  $\mathbf{A}^\top$  and its conjugate transpose is given  
 129 by  $\mathbf{A}^H$ . Fourier domain quantities are denoted using capital gothic letters, i.e., the vector  
 130  $\mathfrak{X}[m] \in \mathbb{C}^2$  denotes the  $m$ -th entry of the (one-dimensional) discrete Fourier transform of the  
 131 vector signal  $\{\mathbf{x}[n] \in \mathbb{C}^2\}_{n=0}^{N-1}$ , evaluated at a frequency indexed by integer  $m$ .

132 **2. Polarimetric Fourier phase retrieval model.** For conciseness, we use from now on the  
 133 term *phase retrieval* as a synonym for Fourier phase retrieval.

134 **2.1. General formulation.** Consider a discrete bivariate signal  $\mathbf{x}[n] = (x_1[n], x_2[n])^\top \in \mathbb{C}^2$   
 135 defined for  $n = 0, 1, \dots, N-1$ . Let  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  be the matrix representation of  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$   
 136 obtained by stacking samples row-wise such that

$$137 \quad (2.1) \quad \mathbf{X} = \begin{bmatrix} x_1[0] & x_2[0] \\ x_1[1] & x_2[1] \\ \vdots & \vdots \\ x_1[N-1] & x_2[N-1] \end{bmatrix} = [\mathbf{x}_1 \ \mathbf{x}_2],$$

138 where  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^N$  collect the two vector components of the signal. We define the *polarimetric*  
 139 (*Fourier*) *phase retrieval (PPR)* problem as the recovery of  $\mathbf{X}$  given  $MP$  Fourier polarimetric



**Figure 2.** Physical interpretation of the polarimetric phase retrieval model (PPR) in terms of polarization optics. The four polarimetric projections shown correspond to the standard measurement scheme described by (2.4) and (2.5), see Example 1.

140 projections. Formally,

141 (PPR) 
$$\text{find } \mathbf{X} \in \mathbb{C}^{N \times 2} \text{ given measurements } y_{m,p} = \left| \mathbf{a}_m^H \mathbf{X} \mathbf{b}_p \right|^2,$$

$$m = 0, 1, \dots, M-1, \quad p = 0, 1, \dots, P-1$$

142 where  $\mathbf{a}_m \in \mathbb{C}^N$  is the discrete Fourier vector corresponding to frequency  $f_m = (2\pi m)/M$ ,  
 143 such that  $a_m[n] = \exp[jn f_m]$  for  $n = 0, 1, \dots, N-1$ . The vector  $\mathbf{b}_p \in \mathbb{C}^2$ , normalized such that  
 144  $\|\mathbf{b}_p\|_2^2 = 1$ , denotes an arbitrary projection acting on the two vector components of  $\mathbf{X}$ .

145 Figure 2 permits to attach precise physical interpretations of PPR measurements in terms  
 146 of polarization optics. The matrix  $\mathbf{X}$  represents the one-dimensional bivariate electromagnetic  
 147 field, where each row is a vector of  $\mathbb{C}^2$  describing an arbitrary polarization state (the so-called  
 148 Jones vector [27]). This states passes through a polarizer defined by  $\mathbf{b}_p \in \mathbb{C}^2$ , evaluating the  
 149 projection of polarization states of  $\mathbf{X}$  onto  $\mathbf{b}_p$ . Finally, light impinges on a Fourier detector  
 150 described by  $\mathbf{a}_m \in \mathbb{C}^N$ , leading to squared magnitude PPR measurements  $y_{m,p}$ .

151 The measurement model PPR can be easily implemented experimentally. Indeed, Fourier  
 152 vectors  $\{\mathbf{a}_m\}_{m=0}^{M-1}$  correspond to far-field measurements in optics, as encountered in coherent  
 153 diffraction imaging techniques (for the case of 2D/3D images) or in spectrometry (for the  
 154 1D case of ultra-short pulses). On the other hand, the set  $\{\mathbf{b}_p\}_{p=0}^{P-1}$  describes the different  
 155 polarizers (or polarization analysers) required to measure polarization of light. Any arbitrary  
 156 polarizer (in mathematical terms, any unit-norm vector  $\mathbf{b}_p \in \mathbb{C}^2$ ) can be constructed as  
 157 combination of standard optical components, such as linear polarizers or waveplates [27].  
 158 Therefore, polarimetric measurements are very flexible: their number, as well as the reference  
 159 polarization states  $\{\mathbf{b}_p\}_{p=0}^{P-1}$  can be tailored at will depending on the context.

160 **2.2. Relation with Fourier matrix measurements.** A closely related problem to PPR is  
 161 the *bivariate phase retrieval* (BPR) problem. Let us introduce the discrete Fourier transform  
 162 of the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  as

$$163 \quad (2.2) \quad \mathfrak{X}[m] = \sum_{n=0}^{N-1} \mathbf{x}[n] \exp\left(-2\pi j \frac{mn}{M}\right) = \begin{bmatrix} \mathfrak{X}_1[m] \\ \mathfrak{X}_2[m] \end{bmatrix} = (\mathbf{a}_m^H \mathbf{X})^\top \in \mathbb{C}^2$$

164 for  $m = 0, 1, \dots, M-1$ . Then let  $\mathfrak{F}[m]$  denote the rank-1 complex spectral matrix such that

$$165 \quad (2.3) \quad \mathfrak{F}[m] = \mathfrak{X}[m] \mathfrak{X}[m]^H = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m] \mathfrak{X}_2[m]^* \\ \mathfrak{X}_2[m] \mathfrak{X}_1[m]^* & |\mathfrak{X}_2[m]|^2 \end{bmatrix} \in \mathbb{C}^{2 \times 2}.$$

166 At a given frequency indexed by  $m$ , the spectral matrix  $\mathfrak{F}[m]$  collects the squared Fourier  
 167 amplitudes of the two components  $\mathbf{x}_1$  and  $\mathbf{x}_2$  of the bivariate signal as well as their relative  
 168 Fourier phase. The recovery of the original bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  (or equivalently its  
 169 matrix representation  $\mathbf{X}$ ) from its spectral matrices defines the BPR problem:

$$170 \quad (\text{BPR}) \quad \text{find } \mathbf{X} \in \mathbb{C}^{N \times 2} \text{ given spectral matrix measurements } \{\mathfrak{F}[m]\}_{m=0}^{M-1}.$$

171 The following proposition shows that BPR and PPR are equivalent in the noiseless setting  
 172 under very general assumptions on the projection vectors  $\{\mathbf{b}_p\}_{p=0}^{P-1}$ .

173 **Proposition 2.1 (Equivalence between BPR and PPR).** *Suppose that the collection of pro-*  
 174 *jection vectors  $\mathbf{b}_0, \mathbf{b}_1, \dots, \mathbf{b}_{P-1} \in \mathbb{C}^2$  satisfies the condition*

$$175 \quad (\mathcal{H}) \quad \text{span}_{\mathbb{R}} \left\{ \mathbf{b}_p \mathbf{b}_p^H \right\}_{p=0}^{P-1} = \left\{ \mathbf{M} \in \mathbb{C}^{2 \times 2} \mid \mathbf{M}^H = \mathbf{M} \right\},$$

176 *i.e., the set of  $P$  rank-1 matrices  $\mathbf{b}_p \mathbf{b}_p^H$  is a generating family (over  $\mathbb{R}$ ) of the space of 2-by-2*  
 177 *Hermitian matrices. Then, under assumption  $(\mathcal{H})$ , the problem PPR is equivalent to BPR in*  
 178 *the sense that  $\mathbf{X}$  is a solution of the problem PPR if and only if  $\mathbf{X}$  is solution of BPR.*

179 *Proof.* It is sufficient to show that, under assumption  $(\mathcal{H})$ , there is a one-to-one corre-  
 180 spondence between the data of BPR (spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$ ) and that of PPR (Fourier  
 181 polarimetric measurements  $\{y_{m,p}\}_{m,p=0}^{M-1, P-1}$ ). In particular, we prove that for  $m$  fixed, the  
 182 spectral matrix  $\mathfrak{F}[m]$  can be obtained from  $\{y_{m,p}\}_{p=0}^{P-1}$  and vice-versa. First, remark that

$$183 \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 = \mathfrak{X}[m]^\top \mathbf{b}_p \mathbf{b}_p^H \mathfrak{X}^*[m] = \text{Tr } \mathbf{b}_p^* \mathbf{b}_p^\top \mathfrak{F}[m],$$

184 *i.e., measurements  $y_{m,p}$  are linear measurements of  $\mathfrak{F}[m]$  through sensing matrices  $\{\mathbf{b}_p^* \mathbf{b}_p^\top\}_{p=0}^{P-1}$ .*  
 185 *Conversely, since  $\{\mathbf{b}_p \mathbf{b}_p^H\}_{p=0}^{P-1}$  (and equivalently,  $\{\mathbf{b}_p^* \mathbf{b}_p^\top\}_{p=0}^{P-1}$ ) is a generating family of the*  
 186 *space of 2-by-2 Hermitian by matrices by assumption  $(\mathcal{H})$ , the spectral matrix  $\mathfrak{F}[m]$  can be*  
 187 *uniquely determined from  $\{y_{m,p}\}_{p=0}^{P-1}$  by linear combinations. This concludes the proof. ■*

188 It is worth noting that the assumption  $(\mathcal{H})$  is not restrictive at all. In fact, for  $P \geq 4$ , the set  
 189  $\{\mathbf{b}_p\}_{p=0}^{P-1}$  where vectors are i.i.d. Gaussian distributed on  $\mathbb{C}^2$  almost surely satisfies  $(\mathcal{H})$ . The  
 190 following example gives an explicit choice of projection vectors  $\mathbf{b}_p$  for  $P = 4$ , which has a nice  
 191 physical interpretation in terms of polarization optics.

192 **Example 1.** Let  $P = 4$  and consider the following projection vectors

$$193 \quad (2.4) \quad \mathbf{b}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ j \end{bmatrix}.$$

194 The projection vectors  $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$  and  $\mathbf{b}_3$  correspond to Jones vectors of standard polarizers used  
 195 in optics [19], which are, respectively: horizontal linear polarizer, vertical linear polarizer, 45°  
 196 linear polarizer and left circular polarizer. See Figure 2 for an illustration. A direct check  
 197 shows that rank-one matrices  $\mathbf{b}_0\mathbf{b}_0^H, \mathbf{b}_1\mathbf{b}_1^H, \mathbf{b}_2\mathbf{b}_2^H, \mathbf{b}_3\mathbf{b}_3^H$  form a basis over the real vector space  
 198 of 2-by-2 Hermitian matrices, and as a result, they are a generating family of such matrices.  
 199 **PPR** measurements read explicitly

$$200 \quad (2.5) \quad \begin{aligned} y_{m,0} &= |\mathfrak{X}_1[m]|^2, & y_{m,1} &= |\mathfrak{X}_2[m]|^2, \\ y_{m,2} &= \frac{1}{2} |\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2, & y_{m,3} &= \frac{1}{2} |\mathfrak{X}_1[m] + j\mathfrak{X}_2[m]|^2. \end{aligned}$$

201 These expressions directly give the diagonal terms of  $\mathfrak{F}[m]$  as  $y_{m,0}$  and  $y_{m,1}$ . The off-diagonals  
 202 terms can be recovered easily using polarization identities in the complex case, such that

$$\begin{aligned} 203 \quad \text{real}(\mathfrak{X}_1[m]\mathfrak{X}_2[m]^*) &= \frac{1}{2} \left( |\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2 \right) \\ 204 &= y_{m,2} - \frac{1}{2} (y_{m,0} + y_{m,1}), \\ 205 \quad \text{imag}(\mathfrak{X}_1[m]\mathfrak{X}_2[m]^*) &= \frac{1}{2} \left( |\mathfrak{X}_1[m] + j\mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2 \right) \\ 206 &= y_{m,3} - \frac{1}{2} (y_{m,0} + y_{m,1}). \end{aligned}$$

208 Remark that the measurement scheme (2.4) yields the same quadratic measurements (2.5)  
 209 as proposed by several authors [53, 38, 42, 54, 50]. Because of that, **BPR** is equivalent to  
 210 the vectorial phase retrieval problem originally introduced in [53]. This shows that **PPR**  
 211 encompasses existing measurements strategies as a special case, while bringing extra flexibility  
 212 in the experimental design of measurements. One of the key benefits of the **PPR** model is  
 213 that additional polarimetric measurements can be generated at will using simple off-the-shelf  
 214 optical components such as linear polarizers or waveplates.

215 **2.3. Trivial ambiguities.** Thanks to Proposition 2.1, we can now give a characterization  
 216 of trivial ambiguities of **PPR** model by leveraging the equivalent **BPR** problem. Indeed, one  
 217 can investigate in a rather simple way the trivial ambiguities that characterize **BPR**. Formally,  
 218 these trivial ambiguities correspond to elementary transformations  $\{\mathbf{x}[n]\}_{n=0}^{N-1} \rightarrow \{\mathbf{x}'[n]\}_{n=0}^{N-1}$   
 219 that leave **BPR** measurements (spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  defined in (2.3)) unchanged.

220 **Global phase ambiguity.** Let  $\alpha \in \mathbb{R}$  and consider the bivariate signal  $\{\mathbf{x}'[n]\}_{n=0}^{N-1}$  such that  
 221  $\mathbf{x}'[n] = \exp(j\alpha)\mathbf{x}[n]$  for every  $n$ . Then for any  $m$ ,  $\mathfrak{F}'[m] = \mathfrak{X}'[m]\mathfrak{X}'[m]^H = \mathfrak{X}[m]\mathfrak{X}[m]^H = \mathfrak{F}[m]$   
 222 since  $\mathfrak{X}'[m] = \exp(j\alpha)\mathfrak{X}[m]$  by linearity properties of the Fourier transform.

223 **Shifts.** This trivial ambiguity only appears when the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  has not full  
 224 support, i.e., when there exist  $n_a, n_b$  with  $0 \leq n_a \leq n_b \leq N-1$  such that  $\mathbf{x}[n] = \mathbf{0}$  for  $n \leq n_a$   
 225 and  $n \geq n_b$ . Assuming this is the case, define the shifted signal  $\{\mathbf{x}'[n]\}_{n=0}^{N-1}$  as  $\mathbf{x}'[n] = \mathbf{x}[n+n_0]$



226 where  $n_0$  is a relative integer between  $(n_b - N)$  and  $(n_a + 1)$  as to ensure proper support. Then,  
 227 using standard Fourier transform properties one gets that  $\mathfrak{X}'[m] = \exp(j2\pi n_0 m/M)\mathfrak{X}[m]$ , so  
 228 that in turn  $\mathfrak{F}'[m] = \mathfrak{F}[m]$  for every  $m$ .

229 **Conjugate reflection.** Consider now  $\{\mathbf{x}'[n]\}_{n=0}^{N-1}$  such that  $\mathbf{x}'[n] = \mathbf{x}^*[N-1-n]$ . Then for  
 230 every  $m$ ,  $\mathfrak{X}'[m] = \exp[-j2\pi(N-1)m/M]\mathfrak{X}^*[m]$ . As a result

$$231 \quad (2.6) \quad \mathfrak{F}'[m] = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_2[m]\mathfrak{X}_1^*[m] \\ \mathfrak{X}_1[m]\mathfrak{X}_2^*[m] & |\mathfrak{X}_2[m]|^2 \end{bmatrix} = \mathfrak{F}[m]^\top.$$

232 This shows that conjugate reflection is not, in general, a trivial ambiguity for **BPR**. This  
 233 contrasts with standard univariate Fourier phase retrieval, see [10, 12].

234 Conjugate reflection can still be a trivial ambiguity provided that the spectral matrix is  
 235 symmetric for every  $m$ , that is  $\mathfrak{F}[m] = \mathfrak{F}[m]^\top$ . Equivalently,  $\mathfrak{F}[m]$  is symmetric if and only if  
 236  $\mathfrak{X}_1[m]\mathfrak{X}_2^*[m] = \mathfrak{X}_2[m]\mathfrak{X}_1^*[m]$ . This means that  $\text{imag}(\mathfrak{X}_1[m]\mathfrak{X}_2^*[m]) = 0$ , i.e., components  $\mathfrak{X}_1[m]$ ,  
 237  $\mathfrak{X}_2[m]$  are in phase at every frequency (they have the same complex argument). Interestingly,  
 238 this condition is interpreted in physical terms as: conjugate reflection is a trivial ambiguity  
 239 for bivariate phase retrieval if and only if the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  is linearly polarized  
 240 at all frequencies.

241 **2.4. 1D equivalent model for PPR.** Back to the original **PPR** problem, we see that it  
 242 defines a new measurement model that performs quadratic scalar projections of the matrix  
 243 representation  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  of the bivariate signal of interest. This *matrix representation* of  
 244 the underlying signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  can be confusing at first: indeed, the bivariate signal is  
 245 intrinsically one-dimensional, in the sense that it is a function of a single index  $n$  – which  
 246 can represent time or 1D spatial coordinates, for instance. Thus, a natural question is the  
 247 following: can **PPR** be equivalently rewritten as a one-dimensional phase retrieval problem?  
 248 If so, what is the physical interpretation of such problem?

249 Let us denote by  $\boldsymbol{\xi} = \text{vec } \mathbf{X} \in \mathbb{C}^{2N}$  the long vector obtained by stacking the two col-  
 250 umns of  $\mathbf{X}$ . Using standard vectorization properties of matrix products, one can rewrite **PPR**  
 251 measurements as

$$252 \quad (2.7) \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 = |(\mathbf{b}_p^\top \otimes \mathbf{a}_m^H) \boldsymbol{\xi}|^2 = |(\mathbf{b}_p^* \otimes \mathbf{a}_m)^H \boldsymbol{\xi}|^2$$

253 for  $m = 0, 1, \dots, M-1$ ,  $p = 0, 1, \dots, P-1$  and where  $\mathbf{a} \otimes \mathbf{b}$  stands for the Kronecker product  
 254 of vectors  $\mathbf{a}$  and  $\mathbf{b}$ . Letting  $\mathbf{c}_{m,p} = \mathbf{b}_p^* \otimes \mathbf{a}_m \in \mathbb{C}^{2N}$ , the **PPR** problem is equivalent to

$$255 \quad (\text{PPR-1D}) \quad \begin{aligned} &\text{find } \boldsymbol{\xi} \in \mathbb{C}^{2N} \text{ given measurements } y_{m,p} = \left| \mathbf{c}_{m,p}^H \boldsymbol{\xi} \right|^2 \\ & \quad m = 0, 1, \dots, M-1, \quad p = 0, 1, \dots, P-1 \end{aligned}$$

256 This shows that **PPR** can be rewritten as a specific instance of 1D phase retrieval with struc-  
 257 tured measurements vectors  $\mathbf{c}_{m,p} \in \mathbb{C}^{2N}$ . While being mathematically sound, the equivalent  
 258 **PPR-1D** problem brings almost no insights about the bivariate nature of the signal to be  
 259 recovered. Moreover, **PPR-1D** cannot be interpreted as a Fourier phase retrieval problem  
 260 with masks [3, 36], since measurements vectors  $\mathbf{c}_{m,p}$  intertwine Fourier measurements  $\mathbf{a}_m$   
 261 and polarimetric projections  $\mathbf{b}_p$  using Kronecker products. Thus, the study of the theoretical

properties of PPR cannot be inferred from standard phase retrieval properties applied to PPR-1D. This requires a dedicated study, which is described in detail in Section 3 and exploited in Section 4 to formulate algebraic solutions to the PPR problem. Nonetheless, as we shall see in Section 5, the equivalent formulation PPR-1D remains particularly useful for designing (iterative) algorithms to solve the original PPR problem.

**3. Uniqueness and polynomial formulation.** This section studies the uniqueness properties of noiseless PPR under the set of assumptions ( $\mathcal{H}$ ) defined in Section 2.2. Thanks to Proposition 2.1, we see that any solution of the problem PPR is a solution of the problem BPR, and vice-versa. This formal equivalence permits to study uniqueness properties of the original PPR through BPR. Following standard practice in Fourier phase retrieval problems, Section 3.1 reformulates BPR using a polynomial formalism. Theorem 3.2 shows that under the usual oversampling condition  $M \geq 2N - 1$ , BPR is equivalent to a polynomial autocorrelation factorization (PAF) problem. Section 3.2 then provides general uniqueness results for PAF and demonstrates that it can be solved using simple greatest common divisor computations.

**3.1. Bivariate phase retrieval as a polynomial factorization problem.** This section follows standard practice in Fourier phase retrieval problems [10, 12, 8, 11, 9] and adopts the polynomial representation of Fourier transforms to study the uniqueness properties of the BPR problem. Formally, let  $\mathbb{C}_{\leq N-1}[z]$  be the space of polynomials of degree at most  $N - 1$ . First, let us define the polynomials  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  as generating polynomials of the components of the bivariate signal  $\mathbf{x}[n] = (x_1[n], x_2[n])^\top \in \mathbb{C}^2$ ,  $n = 0, 1, \dots, N - 1$

$$(3.1) \quad X_1(z) = \sum_{n=0}^{N-1} x_1[n]z^n, \quad X_2(z) = \sum_{n=0}^{N-1} x_2[n]z^n.$$

Similarly, define their conjugate reflections  $\tilde{X}_1, \tilde{X}_2 \in \mathbb{C}_{\leq N-1}[z]$ , obtained by reversing the order and conjugating the coefficients of  $X_1(z)$  and  $X_2(z)$ :

$$(3.2) \quad \tilde{X}_1(z) = \sum_{n=0}^{N-1} x_1^*[N - n - 1]z^n, \quad \tilde{X}_2(z) = \sum_{n=0}^{N-1} x_2^*[N - n - 1]z^n.$$

Then we define the following matrix polynomial  $\mathbf{\Gamma} \in \mathbb{C}_{\leq 2N-2}^{2 \times 2}[z]$

$$(3.3) \quad \mathbf{\Gamma}(z) = \begin{bmatrix} \Gamma_{11}(z) & \Gamma_{12}(z) \\ \Gamma_{21}(z) & \Gamma_{22}(z) \end{bmatrix} = \begin{bmatrix} X_1(z)\tilde{X}_1(z) & X_1(z)\tilde{X}_2(z) \\ X_2(z)\tilde{X}_1(z) & X_2(z)\tilde{X}_2(z) \end{bmatrix} = \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \begin{bmatrix} \tilde{X}_1(z) & \tilde{X}_2(z) \end{bmatrix},$$

where each element of the matrix is a polynomial  $\Gamma_{ij} \in \mathbb{C}_{\leq 2N-2}[z]$ . The coefficients of these polynomials are simply the covariance functions (auto-covariances and cross-covariances) of the vector components  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^N$  that define the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$ . Moreover, the spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  of BPR are linked to the evaluations of the polynomial  $\mathbf{\Gamma}(z)$ .

**Lemma 3.1.** *The coefficients  $\Gamma_{ij}$  of the matrix polynomial  $\mathbf{\Gamma} \in \mathbb{C}_{\leq 2N-2}^{2 \times 2}[z]$  are given by*

$$(3.4) \quad \Gamma_{ij}(z) = \sum_{n=0}^{2N-2} \gamma_{ij}[n - N + 1]z^n \text{ with } \gamma_{ij}[n] = \sum_{k \in \mathbb{Z}} x_i[k + n]x_j^*[k],$$

294 where  $x_i[n] = 0$  for  $n < 0$  and  $n \geq N$  by convention, and the covariance functions  $\gamma_{ij}[n]$  are  
 295 defined for  $n = -N + 1, \dots, N - 1$ . Moreover, the spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  of BPR can  
 296 be expressed for  $m = 0, 1, \dots, M - 1$  as

$$297 \quad (3.5) \quad \mathfrak{F}[m] = e^{j2\pi \frac{m(N-1)}{M}} \mathbf{\Gamma}(e^{-j2\pi \frac{m}{M}}).$$

298 Lemma 3.1 extends to the bivariate case the well-known correspondence between autocovari-  
 299 ance polynomials and Fourier amplitude in univariate Fourier phase retrieval (see for instance  
 300 [10, 12]). For completeness, we give a formal proof in Appendix A.

301 We will refer to  $\mathbf{\Gamma}(z)$  and its entries  $\Gamma_{ij}(z)$  as *measurement polynomials*. Eq. (3.5) shows  
 302 that the coefficients of  $\Gamma_{ij} \in \mathbb{C}_{\leq 2N-2}[z]$  can be uniquely identified from the spectral matrix  
 303 measurements  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  of BPR provided that the number of Fourier measurements  $M$   
 304 exceeds the degree of these polynomials by at least one, i.e.,

$$305 \quad (3.6) \quad M \geq 2N - 1.$$

307 This is the well-known oversampling condition in standard univariate Fourier phase retrieval,  
 308 see e.g. [12]. As a result, one can establish the equivalence between BPR and a polynomial  
 309 recovery problem called Polynomial Autocorrelation Factorization (PAF).

310 **Theorem 3.2.** For  $M \geq 2N - 1$ , BPR is equivalent to the following problem

311 (PAF) find  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  given measurement polynomial  $\mathbf{\Gamma}(z)$  defined as (3.3) .

312 In other terms, there is a one-to-one correspondence between the data ( $\mathbf{\Gamma}(z)$  and  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$ )  
 313 as well as the sets of solutions of the problems (polynomials  $X_1(z), X_2(z)$  and bivariate signal  
 314 components  $\mathbf{x}_1, \mathbf{x}_2$ ).

315 Appendix A provides a proof of this result for completeness. Figure 1 summarizes this equiva-  
 316 lence between BPR and PAF problems, and recall how data and solutions of respective problems  
 317 connect to the initial PPR problem.

318 **3.2. General uniqueness result.** The PAF formulation is very helpful for establishing the  
 319 uniqueness conditions of BPR and, in turn, that of PPR under the nonrestrictive assumption  
 320 ( $\mathcal{H}$ ). Notably, PAF enables a complete characterization of uniqueness properties in terms of  
 321 algebraic properties of complex polynomials. To simplify the presentation in the following,  
 322 uniqueness properties refer jointly to PPR, BPR and PAF problems.

323 In this section, we reproduce several important results from [63] regarding the uniqueness  
 324 of polynomial autocorrelation factorizations problems. The notion of greatest common divisor  
 325 (GCD) of complex polynomials plays a pivotal role in establishing and interpreting these  
 326 statements. Consider two non-zero polynomials  $A_1, A_2 \in \mathbb{C}_{\leq D}[z]$ . The GCD of  $A_1(z)$  and  
 327  $A_2(z)$  is denoted  $\gcd(A_1, A_2)$ . It is a polynomial in  $\mathbb{C}_{\leq K}[z]$ , with highest possible  $K$ , which is  
 328 a divisor of both  $A_1(z)$  and  $A_2(z)$ . Moreover, it is defined up to a multiplication by a scalar in  
 329  $\mathbb{C} \setminus \{0\}$ . We denote  $Q(z) = \gcd(A_1, A_2)$  then there exists two polynomials  $R_1, R_2 \in \mathbb{C}_{\leq D-K}[z]$   
 330 such that  $A_1(z) = Q(z)R_1(z)$  and  $A_2(z) = Q(z)R_2(z)$ . The polynomials  $R_1(z)$  and  $R_2(z)$  are  
 331 called *quotient polynomials*. They are *co-prime* since  $\gcd(R_1, R_2) = 1$ .

332 **Theorem 3.3** ([63], Theorem 2). *The following equivalences are true:*

- 333 1. **PAF** admits a unique solution (up to trivial ambiguities);
- 334 2.  $X_1(z)$  and  $X_2(z)$  have no common roots outside the unit circle;
- 335 3.  $Q(z) = \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22})$  has no roots outside the unit circle.

336 The proof of this result can be found after [63] for the generalization of **PAF** to the case of  
 337  $R$  polynomials. Note that the uniqueness condition given in **Theorem 3.3** clarifies previous  
 338 statements made in the literature [53, 38]. In particular, in [53, Theorem 1] it was claimed that  
 339 coprimeness of the polynomials  $X_1(z)$  and  $X_2(z)$  was a necessary and sufficient for uniqueness  
 340 of the solution. **Theorem 3.3** shows that it was just a sufficient condition, because unimodular  
 341 roots do not affect uniqueness. This agrees with a similar behavior observed for univariate  
 342 one-dimensional Fourier phase retrieval [10], where unimodular roots do not contribute to the  
 343 number of non-trivial solutions. However, unlike univariate one-dimensional Fourier phase  
 344 retrieval, the bivariate case is almost everywhere unique, as shown in the following corollary.

345 **Corollary 3.4** ([63], Corollary 2). *The **PAF** problem admits a unique solution for almost  
 346 every polynomials  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$ .*

347 The proof essentially comes down to observing that the set of polynomials  $X_1, X_2 \in$   
 348  $\mathbb{C}_{\leq N-1}[z]$  with at least one common root is an algebraic variety of dimension smaller than  
 349  $2N - 1$ ; hence it is of measure zero. Put it differently, this shows that **PAF** has the appealing  
 350 property that almost all polynomials  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  can be uniquely recovered from  
 351 measurement polynomials  $\Gamma_{11}(z), \Gamma_{12}(z), \Gamma_{21}(z)$  and  $\Gamma_{22}(z)$ .

352 In practice, if one picks polynomials  $X_1(z)$  and  $X_2(z)$  at random from some continuous  
 353 probability distribution, then they can be almost surely uniquely recovered through **PAF**.  
 354 Moreover, they are almost surely co-prime, i.e.,  $\gcd(X_1, X_2) = 1$ . In this very general case,  
 355 the following proposition shows that recovery is possible through simple GCD computations.

356 **Proposition 3.5** (GCD-based recovery). *Let  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  such that  $\gcd(X_1, X_2) = 1$ .  
 357 Then  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered as*

$$358 \quad (3.7) \quad X_1(z) = \gcd(\Gamma_{11}, \Gamma_{12}) \text{ and } X_2(z) = \gcd(\Gamma_{21}, \Gamma_{22}).$$

359 *Proof.* Suppose that  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  such that  $\gcd(X_1, X_2) = 1$ . This implies that  
 360  $\gcd(\tilde{X}_1, \tilde{X}_2) = 1$ . Therefore,  $\gcd(\Gamma_{11}, \Gamma_{12}) = \gcd(X_1 \tilde{X}_1, X_1 \tilde{X}_2) = X_1(z)$  since  $\tilde{X}_1(z)$  and  
 361  $\tilde{X}_2(z)$  are co-prime. The same argument yields  $\gcd(\Gamma_{21}, \Gamma_{22}) = X_2(z)$ . ■

362 **Proposition 3.5** is a central result. It indicates that the **PAF** problem, and by extension,  
 363 **BPR** and **PPR** can be solved using polynomial algebraic techniques. This distinctive fea-  
 364 ture arises as a direct consequence of accounting for polarization in Fourier phase retrieval  
 365 problems. This original direction is further explored in **Section 4**, where we devise algebraic  
 366 approaches to solve the noisy **PPR** problem using approximate GCD computations.

367 **4. Solving PPR with algebraic methods.** A central result of the previous section is  
 368 **Proposition 3.5**, which states that polynomials  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered (up  
 369 to trivial ambiguities) as GCDs of measurements polynomials  $\Gamma_{11}(z), \Gamma_{12}(z), \Gamma_{21}(z)$  and  $\Gamma_{22}(z)$ .  
 370 The set of equivalencies summarized in **Figure 1** further demonstrates that, in absence of noise,  
 371 such *algebraic approaches* can be readily used to solve the initial **PPR** problem. In the context

**Algorithm 1:** Algebraic approaches for noisy PPR

---

**Input:** polarimetric measurements  $y_{m,p}$ ,  $m = 0, 1, \dots, M-1$ ,  $p = 0, 1, \dots, P-1$   
**Step 1:** reconstruction of measurements polynomials (Section 4.1);  
**for**  $m = 0, \dots, M-1$  **do**  
    | use  $P$  polarimetric measurements to obtain an estimate  $\hat{\mathfrak{F}}[m]$  as (4.7);  
**end**  
Obtain estimates  $\{\hat{\gamma}_{ij}[n]\}_{n=1-N}^{N-1}$  of covariance functions for  $i, j = 1, 2$  by inverse FFT  
of entries of  $\{\hat{\mathfrak{F}}[m]\}_{m=0}^{M-1}$  (possibly resampled to  $2N-1$  points if  $M > 2N-1$ );  
Define measurement polynomials  $\hat{\Gamma}_{ij}(z)$  with coefficients  $\{\hat{\gamma}_{ij}[n-N+1]\}_{n=0}^{2N-2}$ , see  
(3.4);  
**Step 2:** approximate GCD computations (Section 4.2 and Section 4.3);  
Construct the estimated matrix polynomial  $\hat{\mathbf{\Gamma}}(z)$  using step 1;  
Obtain  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  as outputs of one the following methods: right-kernel Sylvester  
(Algorithm 2) or left-kernel Sylvester (Algorithm 3);  
**Result:** estimates  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$

---

372 of noisy PPR measurements, this section shows how to leverage the notion of *approximate*  
373 GCD [64] for solving the polarimetric phase retrieval problem thanks to computational linear  
374 algebra methods. In the sequel, we assume that PPR measurements are corrupted by additive  
375 i.i.d. Gaussian noise such that for  $m = 0, 1, \dots, M-1$  and  $p = 0, 1, \dots, P-1$ ,

$$376 \quad (4.1) \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 + n_{m,p}, \quad n_{m,p} \sim \mathcal{N}(0, \sigma^2),$$

377 where  $\sigma^2$  is the Gaussian noise variance. The signal-to-noise ratio (SNR) is then defined as

$$378 \quad (4.2) \quad \text{SNR} = \frac{\sum_{m=0}^{M-1} \sum_{p=0}^{P-1} |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^4}{MP\sigma^2}.$$

379 Algorithm 1 summarizes the use of algebraic approaches to solve noisy PPR. They operate in  
380 two steps. First, one first needs to obtain an estimate  $\hat{\mathbf{\Gamma}}(z)$  of the measurement polynomial  
381 matrix  $\mathbf{\Gamma}(z)$  given noisy scalar PPR measurements  $y_{m,p}$ ,  $m = 0, 1, \dots, M-1$ ,  $p = 0, 1, \dots, P-1$ .  
382 Section 4.1 addresses this question. The second step exploits approximate GCDs computations  
383 of measurement polynomials to recover estimates  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  of the coefficients of polynomials  
384  $X_1(z)$  and  $X_2(z)$  (or equivalently, the two components of the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$ ).  
385 Section 4.2 introduces the main theoretical tools for this task, namely the notion of Sylvester  
386 matrices and their (left or right) kernel properties, in a general context. Section 4.3 then  
387 introduces two practical algebraic algorithms to recover estimates of the bivariate signal of  
388 interest.

389 **4.1. Reconstruction of measurement polynomials.** Recall that by Lemma 3.1 mea-  
390 surement polynomials  $\Gamma_{ij}(z)$  can be readily expressed in terms of auto-covariance functions  
391  $\{\gamma_{11}[n]\}$ ,  $\{\gamma_{22}[n]\}$  and cross-covariance functions  $\{\gamma_{12}[n]\}$ ,  $\{\gamma_{21}[n]\}$ . Thus, recovery of polyno-  
392 mials  $\Gamma_{ij}(z)$  is identical to the recovery of  $\{\gamma_{ij}[n]\}_{n \in \mathbb{Z}}$  for  $i, j = 1, 2$ . Equivalently, by discrete

393 Fourier transformation, one must retrieve the spectral matrix  $\mathfrak{F}[m]$  for  $m = 0, 1, \dots, M - 1$   
 394 from PPR measurements.

395 Consider noisy measurements given by (4.1). Since  $|\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 = \text{Tr} \mathbf{b}_p^* \mathbf{b}_p^\top \mathfrak{F}[m]$ , an estimate  
 396  $\hat{\mathfrak{F}}[m]$  of  $\mathfrak{F}[m]$  is found for every  $m$  by minimizing the following quadratic-loss

$$397 \quad (4.3) \quad \hat{\mathfrak{F}}[m] = \underset{\substack{\tilde{\mathfrak{F}}[m] = \tilde{\mathfrak{F}}[m]^H \\ \text{rank } \tilde{\mathfrak{F}}[m] = 1}}{\arg \min} \sum_{p=0}^{P-1} \left( y_{m,p} - \text{Tr} \mathbf{b}_p^* \mathbf{b}_p^\top \tilde{\mathfrak{F}}[m] \right)^2,$$

398 where the Hermitian and rank-one constraint ensures the estimated spectral matrix  $\hat{\mathfrak{F}}[m]$  has  
 399 the right structure for future polynomial GCD computations.

400 To solve (4.3), we adopt a heuristic but simple strategy similar to practical polarimetric  
 401 reconstruction techniques used in optics [57, 28]. First, we exploit the *Stokes parameters*  
 402 representation of 2-by-2 Hermitian matrices, which read for an arbitrary Hermitian matrix  
 403  $\mathbf{M} \in \mathbb{C}^{2 \times 2}$

$$404 \quad (4.4) \quad \mathbf{M} = \frac{1}{2} \begin{bmatrix} S_0 + S_1 & S_2 + jS_3 \\ S_2 - jS_3 & S_0 - S_1 \end{bmatrix} \quad S_0, S_1, S_2, S_3 \in \mathbb{R}.$$

405 This set of four real-valued parameters are widely used in optics to describe the different  
 406 polarization states of light. Formally, Stokes parameters define a bijective map  $\mathcal{S} : \{\mathbf{M} \in$   
 407  $\mathbb{C}^{2 \times 2} | \mathbf{M} = \mathbf{M}^H\} \rightarrow \mathbb{R}^4$  such that  $\mathcal{S}(\mathbf{M}) = (S_0, S_1, S_2, S_3)^\top$ . This allows to express the  
 408 noiseless measurements as a simple scalar product between Stokes vectors, i.e.,

$$409 \quad (4.5) \quad \text{Tr} \mathbf{b}_p^* \mathbf{b}_p^\top \tilde{\mathfrak{F}}[m] = \left[ \mathcal{S} \left( \mathbf{b}_p^* \mathbf{b}_p^\top \right) \right]^\top \mathcal{S} \left( \tilde{\mathfrak{F}}[m] \right).$$

410 Therefore, for  $m$  fixed, we can set  $\mathbf{y}_{m,:} = (y_{m,0}, y_{m,1}, \dots, y_{m,P-1})^\top \in \mathbb{R}_+^P$  as the vector collect-  
 411 ing the  $P$  polarimetric measurements. Then one defines the polarization measurement matrix  
 412  $\mathbf{D} \in \mathbb{R}^{P \times 4}$  such that its  $p$ -th row reads  $\mathbf{D}_p = \left[ \mathcal{S} \left( \mathbf{b}_p^* \mathbf{b}_p^\top \right) \right]^\top$ . Note that the matrix  $\mathbf{D}$  does not  
 413 depend on Fourier frequency index  $m$ . This leads to rewriting problem (4.3) as

$$414 \quad (4.6) \quad \hat{\mathfrak{F}}[m] = \underset{\substack{\tilde{\mathfrak{F}}[m] = \tilde{\mathfrak{F}}[m]^H \\ \text{rank } \tilde{\mathfrak{F}}[m] = 1}}{\arg \min} \left\| \mathbf{y}_{m,:} - \mathbf{D} \mathcal{S} \left( \tilde{\mathfrak{F}}[m] \right) \right\|_2^2.$$

415 A possibly sub-optimal yet very simple solution to (4.6) consists in finding the best rank-one  
 416 approximation of the classical least square estimator of Stokes parameters, i.e.,

$$417 \quad (4.7) \quad \hat{\mathfrak{F}}[m] = \text{rank1} \left\{ \mathcal{S}^{-1} \left( \mathbf{D}^\dagger \mathbf{y}_{m,:} \right) \right\},$$

418 where  $\mathbf{D}^\dagger$  denotes the Moore-Penrose pseudo-inverse of  $\mathbf{D}$  and  $\mathcal{S}^{-1}$  is the inverse Stokes  
 419 mapping defined by (4.4). The operator  $\text{rank1}\{\mathbf{M}\}$  finds the best rank-one approximation  
 420 of a given matrix  $\mathbf{M}$  with respect to the Frobenius norm. For the present 2-by-2 Hermitian  
 421 matrix case, the solution is given by keeping the first singular vector of  $\mathbf{M}$ , that is  $\text{rank1}(\mathbf{M}) =$   
 422  $\sigma_0 \mathbf{u}_0 \mathbf{u}_0^H$ , where  $\sigma_0$  and  $\mathbf{u}_0$  are respectively the largest singular value and its corresponding

423 singular vector. Then, estimates  $\{\hat{\gamma}_{ij}[n]\}_{n=1-N}^{N-1}$  of covariance functions for  $i, j = 1, 2$  are  
 424 directly obtained by inverse discrete Fourier transformation of entries of the spectral matrices  
 425  $\{\hat{\mathfrak{F}}[m]\}_{m=0}^{M-1}$  (possibly resampled to  $2N - 1$  points if  $M > 2N - 1$ ). Finally, Eq. (3.4) permits  
 426 to define estimated polynomials  $\hat{\Gamma}_{ij}(z)$  as polynomials in  $\mathbb{C}_{\leq 2N-2}[z]$  with vector of coefficients  
 427  $[\hat{\gamma}_{ij}[1 - N] \ \hat{\gamma}_{ij}[2 - N] \ \dots \ \hat{\gamma}_{ij}[N - 1]]$ .

428 **4.2. Sylvester matrices and GCD.** Proposition 3.5 shows that, in the noiseless case, poly-  
 429 nomials  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered as GCDs of the measurement polynomial  
 430 matrix  $\Gamma(z)$ . In the noisy PPR measurement case, it further suggests that polynomials  $X_1(z)$   
 431 and  $X_2(z)$  can be estimated, or *approximately recovered* from the estimated matrix polyno-  
 432 mial  $\hat{\Gamma}(z)$  computed in Section 4.1. Due to noise, exact GCDs computations are replaced with  
 433 *approximate* GCD computations, which are carried using kernel (or null-space) properties of  
 434 Sylvester matrices. The following section reviews the relevant theory. Practical use of these  
 435 results in the context of PPR is given in Section 4.3.

436 For simplicity, we assume polynomials  $A, B \in \mathbb{C}_{\leq L}[z]$  of same degree  $L$ . Then we define  
 437 the Sylvester-like matrices, parameterized by an integer  $D \leq L$  (possibly negative) as

$$438 \quad (4.8) \quad \mathcal{S}_D(A, B) = \left[ \begin{array}{ccc|ccc} a_0 & & & b_0 & & \\ \vdots & \ddots & & \vdots & \ddots & \\ a_L & & a_0 & b_L & & b_0 \\ & \ddots & \vdots & & \ddots & \vdots \\ & & a_L & & & b_L \end{array} \right] \in \mathbb{C}^{(2L-D+1) \times 2(L-D+1)}.$$

439 When  $D = 1$  (i.e., the matrix is square  $2L \times 2L$ ), the matrix is the well-known Sylvester  
 440 matrix. There are, however, two important extensions of the classic case:

- 441 • When  $1 \leq D \leq L$ , the matrix is tall (the number of columns does not exceed the  
 442 number of rows), and it is called the *Sylvester subresultant* matrix.
- 443 • If  $D \leq 1$  (in general, chosen to be negative), the matrix is fat (the number of rows  
 444 does not exceed the number of columns), and such a matrix is called *extended Sylvester*  
 445 *matrix*.

446 For an overview of such matrices and the corresponding literature, we refer to [64] (note that  
 447 unlike [64] we use the same notation for subresultant and extended Sylvester matrices). The  
 448 following theorem is classic.

449 **Theorem 4.1 (Sylvester).** *Two polynomials  $A, B \in \mathbb{C}_{\leq L}[z]$  have a non-trivial common*  
 450 *divisor if and only if  $\mathcal{S}_1(A, B)$  is rank deficient. Moreover the degree  $K$  of  $\gcd(A, B)$  is equal*  
 451 *to the rank defect of  $\mathcal{S}_1(A, B)$ , i.e.,*

$$452 \quad K = 2L - \text{rank } \mathcal{S}_1(A, B)$$

453 *and  $\gcd(A, B) \in \mathbb{C}_{\leq K}[z]$ .*

454 Unfortunately, Theorem 4.1 does not give an explicit way to compute  $\gcd(A, B)$ . In fact,  
 455 explicit determination of the GCD requires the use of Sylvester matrices  $\mathcal{S}_D(A, B)$  in the  
 456 general case  $D \neq 1$ . More precisely, Proposition 4.2 and Proposition 4.3 below show that the  
 457 GCD can be retrieved from the left or right kernel of carefully constructed Sylvester matrices.

458 In what follows, we assume that the GCD has degree  $K$  and note  $Q(z) = \gcd(A, B) \in \mathbb{C}_{\leq K}[z]$ .  
 459 Moreover, we define

$$460 \quad F(z) = \frac{A(z)}{Q(z)}, \quad G(z) = \frac{B(z)}{Q(z)}$$

461 the corresponding quotient polynomials. We begin with the result on the right kernel of  
 462 Sylvester subresultant matrices.

463 **Proposition 4.2 (Right kernel, see e.g. [64, Lemma 4.6]).** *The rank of the Sylvester sub-*  
 464 *resultant matrix  $\mathcal{S}_K(A, B)$  is equal to  $2(L - K + 1) - 1$  (i.e., it has rank defect equal to 1).*  
 465 *Moreover, for the (unique up to scalar factor) nonzero vector in the right kernel*

$$466 \quad \mathcal{S}_K(A, B) \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = 0;$$

467 with  $\mathbf{u}, \mathbf{v} \in \mathbb{C}^{L-K+1}$ , the corresponding polynomials are multiples of the quotient polynomials:

$$468 \quad U(z) = -cG(z), \quad V(z) = cF(z),$$

469 where  $c \in \mathbb{C}$  is some constant.

470 For the case of extended Sylvester matrices ( $D \leq 1$ ), the result on the left kernel matrices  
 471 is less known in the form that we are using here. This is the reason why we also provide a  
 472 short proof in [Appendix B](#).

473 **Proposition 4.3 (Left kernel).** *Let  $D \leq 1$  (i.e.,  $\mathcal{S}_D(A, B)$  is fat with  $2L - D + 1$  rows).*  
 474 *Then the rank of  $\mathcal{S}_D(A, B)$  is equal to*

$$475 \quad \text{rank } \mathcal{S}_D(A, B) = 2L - D + 1 - K;$$

476 therefore the dimension of the left kernel (i.e., the rank defect) is equal to  $K$  (the degree of the  
 477 GCD). Moreover, a vector  $\mathbf{u} \in \mathbb{C}^{2L-D+1}$  is in the left kernel ( $\mathbf{u}^\top \mathcal{S}_D(A, B) = \mathbf{0}$ ) if and only if  
 478 the vector of coefficients  $\mathbf{q} \in \mathbb{C}^{K+1}$  of the GCD  $Q(z)$  satisfies

$$479 \quad \mathbf{q}^\top \begin{bmatrix} u[0] & u[1] & \cdots & u[2L - D - K] \\ u[1] & u[2] & \cdots & u[2L - D - K + 1] \\ \vdots & \vdots & & \vdots \\ u[K] & u[K + 1] & \cdots & u[2L - D] \end{bmatrix} = 0,$$

480 i.e.,  $\mathbf{q}$  is in the (left) kernel of the Hankel matrix with  $K + 1$  rows built from  $\mathbf{u}$ .

481 The next section exploits these properties of the kernel of Sylvester matrices to formulate  
 482 algebraic algorithms for the [PPR](#) problem.

483 **4.3. Algebraic algorithms.** In this section, we propose two algorithms for estimating coef-  
 484 ficients of polynomials  $X_1(z)$  and  $X_2(z)$  from the estimated matrix polynomial  $\hat{\mathbf{\Gamma}}(z)$  computed  
 485 in [Section 4.1](#). Both algorithms rely on the use of the singular value decomposition (SVD)  
 486 to find the left or right kernels of Sylvester matrices constructed from  $\hat{\mathbf{\Gamma}}(z)$ . Thus the pro-  
 487 posed reconstruction methods may appear as suboptimal since the Sylvester structure is not



**Algorithm 2:** Right kernel Sylvester

**Input:** estimated matrix polynomial  $\hat{\Gamma}(z) \in \mathbb{C}_{<2N-2}^{2 \times 2}$ .

Build the matrix  $\mathbf{S} = \mathcal{S}_{N-1}(\hat{\Gamma}_{11}, \hat{\Gamma}_{21}) \in \mathbb{C}^{(3N-2) \times 2N}$ ;

Take  $\mathbf{v} = \mathbf{v}_{2N} \in \mathbb{C}^{2N}$  to be the  $2N$ -th right singular vector of  $\mathbf{S}$  (corresponding to the last nontrivial singular value);

Partition  $\mathbf{v}$  as  $\mathbf{v} = (-\mathbf{v}_2, \mathbf{v}_1)$ , where  $\mathbf{v}_1 = c\hat{\mathbf{x}}_1$  and  $\mathbf{v}_2 = c\hat{\mathbf{x}}_2$  with  $c \in \mathbb{C}$ ;

Determine  $|c|$  by proper norm scaling as

$$|c| = \left( \frac{\|\mathbf{v}_1\|_2^2 + \|\mathbf{v}_2\|_2^2}{\hat{\gamma}_{11}[0] + \hat{\gamma}_{22}[0]} \right)^{\frac{1}{2}}$$

Set  $\hat{\mathbf{x}}_1 = \mathbf{v}_1/|c|$  and  $\hat{\mathbf{x}}_2 = \mathbf{v}_2/|c|$ ;

**Result:** estimates  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$

488 taken account when computing the (low-rank) kernels. This limitation could be overcome  
 489 with structured low-rank approximations [45], to be specifically tailored for the PPR prob-  
 490 lem. Such a study would fall outside the scope of the present work. Still, as demonstrated  
 491 by the numerical experiments presented in Section 6, the SVD already provides excellent re-  
 492 construction performance in many scenarios, while maintaining a reasonable computational  
 493 burden.

494 **4.3.1. Right kernel Sylvester.** The first algorithm is based on the properties of the right  
 495 kernel of Sylvester matrices described in Proposition 4.2. It uses the fact that  $X_1(z)$  and  
 496  $X_2(z)$  are (without noise) quotient polynomials of

$$497 \quad \Gamma_{11}(z) = X_1(z)\tilde{X}_1(z) \text{ and } \Gamma_{21}(z) = X_2(z)\tilde{X}_1(z).$$

498 One can remark that  $X_1(z)$  and  $X_2(z)$  are also quotient polynomials of  $\Gamma_{12}(z) = X_1(z)\tilde{X}_2(z)$   
 499 and  $\Gamma_{22}(z) = X_2(z)\tilde{X}_2(z)$ , which adds some freedom in the choice of measurement polynomi-  
 500 als. For the sake of simplicity, we will work with estimated polynomials  $\hat{\Gamma}_{11}(z)$  and  $\hat{\Gamma}_{21}(z)$  in  
 501 the following.

502 The complete right kernel Sylvester approach is summarized in Algorithm 2. It estimates  
 503 the (one-dimensional) right kernel by computing the last nontrivial singular value of the  
 504 Sylvester matrix  $\mathcal{S}_{N-1}(\hat{\Gamma}_{11}, \hat{\Gamma}_{21})$ . According to Proposition 4.2, this directly gives, up to one  
 505 complex multiplicative constant, an estimation  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  of the vectors of coefficients defining  
 506 polynomials  $X_1(z)$  and  $X_2(z)$ . This constant is then computed (up to one unit-modulus factor  
 507 due to the trivial rotation ambiguity) by scaling the 2-norm of  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  thanks to the value  
 508 at the origin ( $n = 0$ ) of estimated auto-covariance functions  $\hat{\gamma}_{11}[n]$  and  $\hat{\gamma}_{22}[n]$ .

509 One of the key advantages of this algorithm lies in its simplicity. Indeed, it only requires a  
 510 single SVD of a  $(3N - 2) \times 2N$  matrix and thus has overall computational complexity  $\mathcal{O}(N^3)$ .

511 **4.3.2. Left kernel Sylvester.** The second algorithm exploits the properties of the left  
 512 kernel of extended (fat) Sylvester matrices (i.e.,  $\mathcal{S}_D$  for  $D \leq 1$ ) detailed in Proposition 4.3.  
 513 For simplicity and to reduce the size of the involved matrices we set  $D = 1$  in what follows.  
 514 Nonetheless, the proposed approach can be adapted to any value of  $D \leq 1$  if needed.

**Algorithm 3:** Left kernel Sylvester

**Input:** estimated matrix polynomial  $\hat{\Gamma}(z) \in \mathbb{C}_{<2N-2}^{2 \times 2}$ .

**for**  $j = 1, 2$  **do**

Build the matrix  $\mathbf{S} = \mathcal{S}_1(\hat{\Gamma}_{j1}, \hat{\Gamma}_{j2}) \in \mathbb{C}^{(4N-4) \times (4N-4)}$ ;

Take the last  $N - 1$  left singular vectors of  $\mathbf{S}$ , i.e.,

$$\mathbf{u}_{3N-2}, \dots, \mathbf{u}_{4N-4}.$$

Stack the Hankel matrices with  $N$  rows in the following matrix

$$\mathbf{H} = [\mathbf{H}_N(\mathbf{u}_{3N-2}) \quad \dots \quad \mathbf{H}_N(\mathbf{u}_{4N-4})] \in \mathbb{C}^{N \times (N-1)(3N-3)}$$

Retrieve  $\mathbf{w}_j = c_j \hat{\mathbf{x}}_j$ ,  $c_j \in \mathbb{C}$  as the last left singular vector of  $\mathbf{H}$ .

**end**

Determine constants  $c_1, c_2$  as

$$c_1 = \frac{\|\mathbf{w}_1\|_2}{\sqrt{\hat{\gamma}_{11}[0]}} \quad \text{and} \quad c_2 = \frac{\|\mathbf{w}_2\|_2}{\sqrt{\hat{\gamma}_{22}[0]}} \exp \left[ j(\arg \hat{\gamma}_{12}[0] - \arg \mathbf{w}_2^H \mathbf{w}_1) \right]$$

Set  $\hat{\mathbf{x}}_1 = \mathbf{w}_1/c_1$  and  $\hat{\mathbf{x}}_2 = \mathbf{w}_2/c_2$ ;

**Result:** estimates  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$

515 Algorithm 3 summarizes the complete procedure. In essence, it follows the theoretical  
 516 result of Proposition 3.5. In particular, compared to the right kernel Sylvester approach,  
 517 estimated coefficients  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  are obtained by two separate GCD computations: the vector  
 518 of coefficients  $\hat{\mathbf{x}}_1$  is obtained by computing the GCD of estimated measurement polynomials  
 519  $\hat{\Gamma}_{11}(z)$  and  $\hat{\Gamma}_{12}(z)$ , whereas  $\hat{\mathbf{x}}_2$  is obtained by computing the GCD of  $\hat{\Gamma}_{21}(z)$  and  $\hat{\Gamma}_{22}(z)$ .  
 520 Importantly, the two GCDs are determined up to a multiplicative complex constant, say  $c_1$   
 521 and  $c_2$ , which can be determined jointly using PPR measurements.

522 The computation of each GCD requires three steps [64]: a first SVD to determine the  $N - 1$   
 523 last left singular vectors of the Sylvester matrix  $\mathcal{S}_1$ ; the construction of a fat, horizontally  
 524 stacked Hankel matrix  $\mathbf{H}$  with  $N$  rows from these  $N - 1$  singular vectors; a second SVD to  
 525 obtain the  $N$  coefficients of the GCD as the last left singular vector of  $\mathbf{H}$ . Once GCDs have  
 526 been obtained, determination of constants  $c_1$  and  $c_2$  (up to a common global phase factor) is  
 527 carried out by properly scaling the norms of estimated coefficients  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  (using  $\hat{\gamma}_{11}[0]$  and  
 528  $\hat{\gamma}_{22}[0]$ ) and adjusting the phase factor  $\arg c_1 c_2^*$  thanks to the value at  $n = 0$  of the estimated  
 529 cross-covariance function  $\hat{\gamma}_{12}[n]$ .

530 The complexity of the left kernel Sylvester method described in Algorithm 3 is higher for  
 531 two main reasons. First, as explained above, it requires the computations of two SVDs for  
 532 each one of the two GCDs determinations. Moreover, while the first SVD has a cost of  $\mathcal{O}(N^3)$ ,  
 533 the second SVD is performed on a large fat stacked Hankel matrix  $\mathbf{H}$ , with complexity  $\mathcal{O}(N^4)$ ,  
 534 which dominates the overall computational burden of Algorithm 3.

**Algorithm 4:** SDP relaxation for PPR

---

**Input:** measurements  $\mathbf{y} \in \mathbb{R}^{MP}$ , lifted measurement matrices  $\mathbf{C}_{m,p} \in \mathbb{C}^{2N \times 2N}$ , regularization parameter  $\lambda \geq 0$ .  
 set arbitrary  $\Xi^{(0)}$ ;  
 $\Psi^{(0)} \leftarrow \Xi^{(0)}$ ;  
 $k \leftarrow 0$ ;  
**while** *stopping criterion is not satisfied* **do**  
      $\Xi^{(k+1)} = \text{prox}_{t_k g} \left( \Psi^{(k)} - t_k \nabla f(\Psi^{(k)}) \right)$  where the proximal operator is given by  
     (5.8);  
      $\eta_{k+1} = \frac{1 + \sqrt{1 + 4\eta_k^2}}{2}$ ;  
      $\Psi^{(k+1)} = \Xi^{(k+1)} + \left( \frac{\eta_k - 1}{\eta_{k+1}} \right) \left( \Xi^{(k+1)} - \Xi^{(k)} \right)$ ;  
      $k \leftarrow k + 1$ ;  
**end**  
 $\hat{\xi} \leftarrow \text{rank1} \left( \Xi^{(k)} \right)$ ;  
**Result:** estimate  $\hat{\xi}$

---

535 **5. Solving PPR with iterative algorithms.** We now address the design of iterative al-  
 536 gorithms to solve the noisy PPR problem. Section 5.1 and Section 5.2 exploit the PPR-1D  
 537 representation of the original problem to provide a semidefinite programming (SDP) relaxation  
 538 and Wirtinger flow algorithm, respectively.

539 **5.1. SDP relaxation.** Semidefinite programming (SDP) approaches for phase retrieval  
 540 have been increasingly popular for over a decade [15, 16, 67]. In the classical 1D phase retrieval  
 541 case, SDP approaches exploit that even though measurements are quadratic in the unknown  
 542 signal  $\mathbf{x} \in \mathbb{C}^N$ , they are linear in the rank-one matrix  $\mathbf{x}\mathbf{x}^H$ . For PPR, the 1D equivalent  
 543 representation PPR-1D enables to formulate a SDP relaxation of the original problem, by  
 544 observing that

$$545 \quad (5.1) \quad |\mathbf{c}_{m,p}^H \xi|^2 = \text{Tr} \mathbf{c}_{m,p} \mathbf{c}_{m,p}^H \xi \xi^H = \text{Tr} \mathbf{C}_{m,p} \Xi,$$

546 i.e., noiseless measurements can be rewritten as a linear function of the lifted positive semi-  
 547 definite rank-one matrix  $\Xi = \xi \xi^H \in \mathbb{C}^{2N \times 2N}$ . Following the classical PhaseLift methodology  
 548 [16, 15], the original nonconvex PPR problem can be relaxed into a SDP convex program as

$$549 \quad (5.2) \quad \begin{aligned} & \text{minimize} && \frac{1}{2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \text{Tr} \mathbf{C}_{m,p} \Xi)^2 + \lambda \|\Xi\|_* , \\ & \text{subject to} && \Xi \succeq 0 \end{aligned}$$

550 where  $\lambda \geq 0$  is an hyperparameter that allows to control the trade-off between the likelihood  
 551 of observations and the nuclear norm regularization  $\|\cdot\|_*$ . Note that since  $\Xi$  is constrained  
 552 to be positive semidefinite, the nuclear norm regularization is equivalent to the trace-norm

553 regularization used in [15] since  $\|\Xi\|_* = \text{Tr } \Xi$  in this case. The SDP program (5.2) takes a  
 554 standard form: therefore it can be solved in many ways, including interior point methods [66],  
 555 first-order methods [48] or using disciplined convex programming solvers such as CVXPY<sup>1</sup>. For  
 556 completeness, we provide below an explicit algorithm to solve (5.2) using a proximal gradient  
 557 approach [7, Chapter 10]. It closely follows the approach described in [15, 29].

558 The objective function in (5.2) can be rewritten as the sum  $f(\Xi) + g(\Xi)$  with

$$559 \quad (5.3) \quad f(\Xi) = \frac{1}{2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \text{Tr } \mathbf{C}_{m,p} \Xi)^2, \quad g(\Xi) = \lambda \|\Xi\|_* + \iota_{\succeq 0}(\Xi),$$

560 where  $\iota_{\succeq 0}(\cdot)$  denotes the indicator function on the positive semidefinite cone. This ensures  
 561 the formal equivalence between (5.2) and the unconstrained minimization problem

$$562 \quad (5.4) \quad \min_{\Xi \in \mathbb{C}^{2N \times 2N}} f(\Xi) + g(\Xi).$$

563 The convex optimization problem (5.4) can be efficiently solved by proximal gradient methods,  
 564 which take advantage of the splitting between  $f$  and  $g$  of the objective function. More precisely,  
 565 we use the fast proximal gradient method which consist, at iteration  $k$ :

$$566 \quad (5.5) \quad \Xi^{(k+1)} = \underset{t_k g}{\text{prox}} \left( \Psi^{(k)} - t_k \nabla f(\Psi^{(k)}) \right),$$

$$567 \quad (5.6) \quad \eta_{k+1} = \frac{1 + \sqrt{1 + 4\eta_k^2}}{2},$$

$$568 \quad (5.7) \quad \Psi^{(k+1)} = \Xi^{(k+1)} + \left( \frac{\eta_k - 1}{\eta_{k+1}} \right) \left( \Xi^{(k+1)} - \Xi^{(k)} \right),$$

570 where  $t_k$  is a step-size which is chosen such that the proximal gradient step (5.5) obey some  
 571 sufficient decrease condition; see e.g. [7, p. 271] for details. Our choice for the function  $g$  in  
 572 (5.4) enables a simple expression for the associated proximal operator (see [29]):

$$573 \quad (5.8) \quad \underset{\tau g}{\text{prox}}(\mathbf{X}) = \min_{\mathbf{Z} \succeq 0} \tau \lambda \|\mathbf{Z}\|_* + \|\mathbf{Z} - \mathbf{X}\|_2^2 \\ = \mathbf{U} \text{shrink}(\Sigma, \tau \lambda) \mathbf{U}^H,$$

574 where in the last equation,  $\mathbf{U} \Sigma \mathbf{U}^H$  is the eigenvalue decomposition of  $\mathbf{X}$  and the shrink  
 575 operator is defined entry-wise by  $\text{shrink}(\sigma_i, \tau \lambda) = \text{sign}(\sigma_i) \max\{|\sigma_i| - \tau \lambda, 0\}$ .

576 *Choice of regularization parameter  $\lambda$ .* In this work, we fix the value of the regularization  
 577 parameter to  $\lambda = 1/\text{SNR}$ : we found empirically that this choice provides good results in  
 578 most scenarios, as it provides a reasonable tradeoff between likelihood of observations and the  
 579 nuclear norm regularization in the objective function of (5.2).

---

<sup>1</sup><https://www.cvxpy.org/>

**Algorithm 5:** Wirtinger Flow for PPR: PPR-WF

---

**Input:** measurements  $\mathbf{y} \in \mathbb{R}^{MP}$ , measurement matrix  $\mathbf{C} \in \mathbb{C}^{MP \times 2N}$ , tolerance  $\varepsilon$   
 set  $\boldsymbol{\xi}^{(0)}$  using the desired initialization method;  
 $\boldsymbol{\xi}^{(1)} \leftarrow \boldsymbol{\xi}^{(0)}$ ;  
 $k \leftarrow 1$ ;  
**while**  $\|\boldsymbol{\xi}^{(i+1)} - \boldsymbol{\xi}^{(i)}\|_2 > \varepsilon \|\boldsymbol{\xi}^{(i)}\|_2$  **do**  
      $\beta_k \leftarrow \frac{k+1}{k+3}$ ;  
      $\boldsymbol{\psi}^{(k)} \leftarrow \boldsymbol{\xi}^{(k)} + \beta_k (\boldsymbol{\xi}^{(k)} - \boldsymbol{\xi}^{(k-1)})$ ;  
     compute optimal step-size  $\mu_k$  (5.14);  
      $\boldsymbol{\xi}^{(k+1)} \leftarrow \boldsymbol{\psi}^{(k)} - \mu_k \nabla F(\boldsymbol{\psi}^{(k)})$ ;  
      $k \leftarrow k + 1$ ;  
**end**  
 $\hat{\boldsymbol{\xi}} \leftarrow \boldsymbol{\xi}^{(k)}$ ;  
**Result:** estimate  $\hat{\boldsymbol{\xi}}$

---

580 **Convergence.** Obviously, as (5.2) is a convex program, the precision towards the optimal  
 581 cost value can become arbitrarily good as one increases the number of iterations. In practice,  
 582 one needs to stop the algorithm when a prescribed tolerance  $\varepsilon$  is reached. To this aim we  
 583 implemented stopping criteria that carefully monitor a normalized residual, see [29] for details.  
 584 Moreover, it may happen that the estimated lifted matrix  $\hat{\boldsymbol{\Xi}}$  generated by the sequence of  $\boldsymbol{\Xi}^{(k)}$   
 585 is not rank one: in this case, one first computes the rank-one approximation of  $\hat{\boldsymbol{\Xi}}$  (e.g. using  
 586 SVD) to obtain the estimated signal  $\hat{\boldsymbol{\xi}}$ .

587 **Complexity.** The computational cost of the proposed algorithm concentrates on the proxi-  
 588 mal gradient step (5.5), where the evaluation of the proximal operator and the computation of  
 589 the gradient  $\nabla f$  share the computational burden. More precisely, the eigenvalue decomposi-  
 590 tion of a  $2N \times 2N$  matrix together with the shrink operator leads to  $\mathcal{O}(N^3)$  calculations. The  
 591 computation of the gradient leads to  $MP$  trace evaluations of order  $\mathcal{O}(N^2)$  flops, meaning  
 592 that the number of flops per iteration is of order  $\mathcal{O}(MPN^2 + N^3)$ .

593 The full procedure is summarized in Algorithm 4.

594 **5.2. Wirtinger flow for PPR.** Exploiting further the 1D equivalent representation PPR-  
 595 1D of the PPR problem, another approach consists in minimizing directly the following non-  
 596 convex quadratic objective

$$597 \quad (5.9) \quad \min_{\boldsymbol{\xi} \in \mathbb{C}^{2N}} F(\boldsymbol{\xi}) = \frac{1}{2} \|\mathbf{y} - |\mathbf{C}\boldsymbol{\xi}|^2\|_2^2,$$

598 where  $\mathbf{y} \in \mathbb{R}^{MP}$  gathers PPR measurements and where the rows of  $\mathbf{C} \in \mathbb{C}^{MP \times 2N}$  are given  
 599 by  $\mathbf{c}_{m,p}^H$ , see Section 2.4. Provided that one can find a initial point  $\boldsymbol{\xi}^{(0)}$  close enough from  
 600 the global minimizer of (5.9), a simple strategy based on gradient descent can be used to  
 601 solve PPR. However, such an approach requires special care since the optimization variable

602  $\boldsymbol{\xi}$  is complex-valued. In fact, the objective function in (5.9) is real-valued, and thus it is not  
 603 differentiable with respect to complex analysis. Instead, one needs to resort to the so-called  
 604  $\mathbb{C}\mathbb{R}$  or *Wirtinger*-calculus [41] to provide a meaningful extension of gradient-descent-type  
 605 algorithms to the complex case. This is precisely the approach proposed in [17] to solve  
 606 standard phase retrieval, where the complex gradient descent is called *Wirtinger flow* (WF).

607 Leveraging the original WF approach, we propose below a complex-gradient descent al-  
 608 gorithm which solves the nonconvex problem (5.9). Compared to the original paper [17], we  
 609 incorporate optimal step size selection [40] together with a proposed acceleration scheme [69].  
 610 We further propose an efficient strategy for initialization based on the algebraic methods for  
 611 PPR described in Section 4. The superiority of these initializations over standard ones (e.g.  
 612 spectral initialization as proposed in [17]) will be demonstrated in Section 6.2.

613 The proposed PPR-WF algorithm is as follows. Starting from two initial points  $\boldsymbol{\xi}^{(0)}, \boldsymbol{\xi}^{(1)}$ ,  
 614 the  $k$ -th iteration reads

$$615 \quad (5.10) \quad \beta_k = \frac{k+1}{k+3},$$

$$616 \quad (5.11) \quad \boldsymbol{\psi}^{(k)} = \boldsymbol{\xi}^{(k)} + \beta_k (\boldsymbol{\xi}^{(k)} - \boldsymbol{\xi}^{(k-1)}),$$

$$617 \quad (5.12) \quad \boldsymbol{\xi}^{(k+1)} = \boldsymbol{\psi}^{(k)} - \mu_k \nabla F(\boldsymbol{\psi}^{(k)}),$$

619 where  $\beta_k$  is a sequence of accelerated parameters and  $\mu_k$  is a carefully chosen stepsize, see  
 620 further below. Compared to the standard WF algorithm, PPR-WF takes advantage of the  
 621 acceleration procedure first proposed in [69] in the context of ptychographic phase retrieval  
 622 (but using a magnitude loss function instead of a square magnitude loss function as used here).  
 623 Note that the complex gradient of  $F$  can be computed explicitly as

$$624 \quad (5.13) \quad \nabla F(\boldsymbol{\psi}) = \mathbf{C}^H (|\mathbf{C}\boldsymbol{\psi}|^2 - \mathbf{y}).$$

625 *Optimal step-size selection.* We combine acceleration for WF with the optimal step-size  
 626 selection proposed in [40] for the standard WF algorithm. For completeness, we reproduce  
 627 here the main ingredients underpinning optimal step size selection in (5.12) and refer the  
 628 reader to [40] for further details. At iteration  $k$ , the optimal stepsize  $\mu_k$  is defined by line  
 629 search, i.e.,

$$630 \quad (5.14) \quad \mu_k = \arg \min_{\mu} F(\boldsymbol{\xi}^{(k+1)}) = F(\boldsymbol{\psi}^{(k)} - \mu \nabla F(\boldsymbol{\psi}^{(k)})).$$

631 The authors in [40] showed that the 1D optimization problem (5.14) boils down to finding  
 632 the roots of a univariate cubic polynomial with real coefficients, the latter being completely  
 633 determined by the knowledge of  $\boldsymbol{\psi}^{(k)}, \nabla F(\boldsymbol{\psi}^{(k)})$  and  $\mathbf{y}$ , see [40, Eq. (17)]. Roots can be  
 634 determined in closed-form, and two cases can occur: (a) there is only one real root, and thus  
 635 it gives the optimal step-size  $\mu_k$ ; (b) there are three real roots, and in this case  $\mu_k$  is set to  
 636 the real root associated to the minimum objective value. Note that optimal selection for WF  
 637 is somewhat inexpensive, with computational cost dominated by the calculation of the cubic  
 638 polynomial coefficients scaling as  $\mathcal{O}(MP)$ .

639 *Initialization.* Since PPR-WF attempts at minimizing a nonconvex quadratic objective  
 640 (5.9), the choice of initial points  $\boldsymbol{\xi}^{(0)}$ ,  $\boldsymbol{\xi}^{(1)}$  is crucial to hope that PPR-WF will be able to  
 641 recover a global minimizer of the objective function. For simplicity, we set  $\boldsymbol{\xi}^{(1)} = \boldsymbol{\xi}^{(0)}$ , so that  
 642 we only discuss the selection of  $\boldsymbol{\xi}^{(0)}$ . Four different initialization strategies for PPR-WF are  
 643 considered:

- 644 • *spectral initialization* [17]: this standard approach consists in computing the eigenvector  
 645  $\mathbf{v}$  corresponding to the largest eigenvalue of the matrix

$$646 \quad (5.15) \quad \mathbf{Y} = \frac{1}{MP} \sum_{r=0}^{MP-1} y_r \mathbf{c}_r \mathbf{c}_r^H$$

647 and to rescale it properly to set

$$648 \quad (5.16) \quad \boldsymbol{\xi}^{(0)} = \frac{\mathbf{v}}{\lambda}, \quad \lambda = \left( N \frac{\sum_{r=0}^{MP-1} y_r}{\sum_{r=0}^{MP-1} \|\mathbf{c}_r\|^2} \right)^{1/2}.$$

- 649 • *random phase initialization*: we first generate a random measurement phase vector  
 650  $\boldsymbol{\phi} \in \mathbb{R}^{MP}$  with i.i.d. entries  $\phi_r \sim \mathcal{U}([0, 2\pi])$ . Then, we set

$$651 \quad (5.17) \quad \boldsymbol{\xi}^{(0)} = \mathbf{C}^\dagger \tilde{\mathbf{y}}, \quad \tilde{\mathbf{y}} = \mathbf{y} \odot \exp(j\boldsymbol{\phi}),$$

652 where  $\mathbf{C}^\dagger$  is the pseudo-inverse of  $\mathbf{C}$  and  $\odot$  denotes entry-wise product between vectors.

- 653 • *right kernel Sylvester initialization*: we set  $\boldsymbol{\xi}^{(0)}$  as the result of Algorithm 1 where ap-  
 654 proximate GCDs computations are performed using the right kernel Sylvester method  
 655 (Algorithm 2).
- 656 • *left kernel Sylvester initialization*: we set  $\boldsymbol{\xi}^{(0)}$  as the result of Algorithm 1 where  
 657 approximate GCDs computations are performed using the left kernel Sylvester method  
 658 (Algorithm 3).

659 *Convergence monitoring.* We monitor convergence of PPR-WF by computing at each it-  
 660 eration  $k$ , the normed residual  $\|\boldsymbol{\xi}^{(k+1)} - \boldsymbol{\xi}^{(k)}\|_2 / \|\boldsymbol{\xi}^{(k)}\|_2$  and stop the algorithm when it goes  
 661 below a prescribed tolerance  $\varepsilon \ll 1$ .

662 *Complexity.* The computational cost per iteration of PPR-WF is dominated by the eval-  
 663 uation of the complex gradient (5.13), which scales as  $\mathcal{O}(MPN)$ . Note that the optimal  
 664 step-size selection procedure scales as  $\mathcal{O}(MP)$ , meaning that the whole cost of PPR-WF  
 665 remains  $\mathcal{O}(MPN)$  per iteration. Algorithm 5 summarizes the proposed PPR-WF algorithm.

666 **6. Numerical experiments.** We provide in this section several numerical experiments that  
 667 address how PPR can be solved in practice using both algebraic and algorithmic approaches  
 668 described in Section 4 and Section 5, respectively. Importantly, we demonstrate that the use  
 669 of Wirtinger Flow together with a right-Sylvester initial point achieves the best performance  
 670 in terms of mean squared error (MSE) with limited computational burden. This combination  
 671 of algorithmic and algebraic reconstruction methods provides a scalable, asymptotically MSE  
 672 optimal, and parameter free inversion procedure for PPR.

673 Just like in standard phase retrieval, the global phase ambiguity in PPR requires to prop-  
 674 erly realign any estimated signal  $\hat{\mathbf{X}}'$  with the ground truth  $\mathbf{X}$  in order to provide a meaningful

675 squared reconstruction error value. We define the realigned estimated signal  $\hat{\mathbf{X}}$  as

$$676 \quad (6.1) \quad \hat{\mathbf{X}} = e^{j\Phi_0} \hat{\mathbf{X}}' \text{ with } \Phi_0 = \arg \min_{\phi \in [0, 2\pi)} \|e^{j\phi} \hat{\mathbf{X}}' - \mathbf{X}\|_F^2.$$

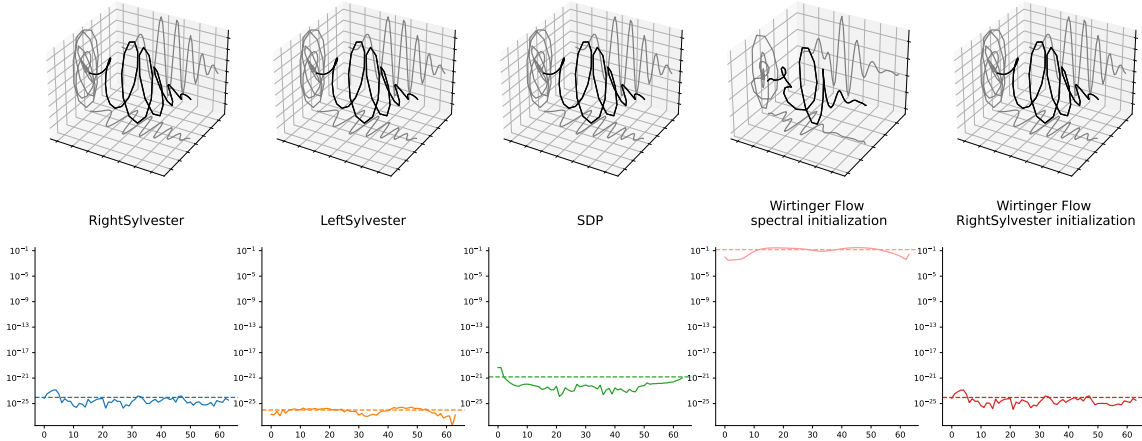
677 The squared reconstruction error is then defined in terms of the Frobenius norm as  $\|\hat{\mathbf{X}} - \mathbf{X}\|_F^2$ .  
 678 Note that in practice, the minimization involved in the realignment procedure can simply be  
 679 performed by evaluating the complex phase of the standard inner product between the vectors  
 680  $\hat{\boldsymbol{\xi}}'$  and  $\boldsymbol{\xi}$  obtained from matrices  $\hat{\mathbf{X}}'$  and  $\mathbf{X}$ , respectively.

681 This section is organized as follows. Section 6.1 presents the reconstruction of a realistic  
 682 bivariate pulse from noiseless PPR measurements using the different approaches presented in  
 683 the paper. Section 6.2 then discusses the choice of initialization in PPR-WF. Section 6.3  
 684 benchmarks the robustness to noise of proposed reconstructions methods. Finally, Section 6.4  
 685 provides a first study of the impact of the number of PPR measurements on reconstruction  
 686 performances.

687 **6.1. Reconstruction of bivariate pulse.** As a first experiment, we consider the reconstruc-  
 688 tion of a bivariate pulse from noiseless PPR measurements. The signal to be recovered defines  
 689 a typical complex-valued bivariate analytic signal associated to the bivariate electromagnetic  
 690 field to be estimated in ultra-short electromagnetic pulses experiments, see e.g. [59, 68]. It  
 691 is defined for  $N = 64$  points and we consider the simple noise-free measurement scheme (2.5)  
 692 with  $M = 2N - 1$  and  $K = 4$ . The bivariate pulse exhibits slow variations of the instantaneous  
 693 polarization state, ensuring uniqueness of the PPR solution. We investigate the capacity of  
 694 the methods introduced in Section 4 and Section 5 to properly recover the bivariate signal  
 695 of interest. Note that for Wirtinger Flow, we consider two initialization strategies, one us-  
 696 ing spectral initialization and the other one based on the solution given by the right kernel  
 697 Sylvester approach.

698 Figure 3 depicts the different reconstructed bivariate signals obtained by each method  
 699 along with the associated squared error  $(\hat{\mathbf{x}}[n] - \mathbf{x}[n])^2$  for every time index  $n$ , where the  
 700 estimated signal  $\hat{\mathbf{x}}$  is realigned with the ground truth  $\mathbf{x}$  using (6.1). Except Wirtinger Flow  
 701 with spectral initialization, all methods successfully recover the original bivariate signal, where  
 702 successful recovery in the noiseless context is decided whenever  $\|\hat{\mathbf{X}} - \mathbf{X}\|_F^2 < 10^{-20}$ . Left and  
 703 right kernel Sylvester and Wirtinger Flow with right Sylvester initialization provide similar  
 704 reconstruction quality, with a slight advantage to left kernel Sylvester. The SDP approach  
 705 performs also well, yet three or four order of magnitude of squared error above the previous  
 706 approaches. Due to the very low error levels involved here, this has little consequence; however,  
 707 compared to the aforementioned methods SDP exhibits both larger memory usage and overall  
 708 computational cost, which makes it a less attractive option to solve this PPR problem in the  
 709 noiseless scenario. Strikingly, one can observe that the Wirtinger Flow approach relying on  
 710 spectral initialization is not able to recover the ground truth signal. Intuitively, it may be  
 711 explained by the fact that spectral initialization provides an initial point too far from the  
 712 global optimum, resulting in Wirtinger Flow to get stuck in a local minima instead. This first  
 713 experiment suggests that the performance of WF-based methods for PPR is tightly related to  
 714 the quality of initial points, which we will investigate in detail in the next section.

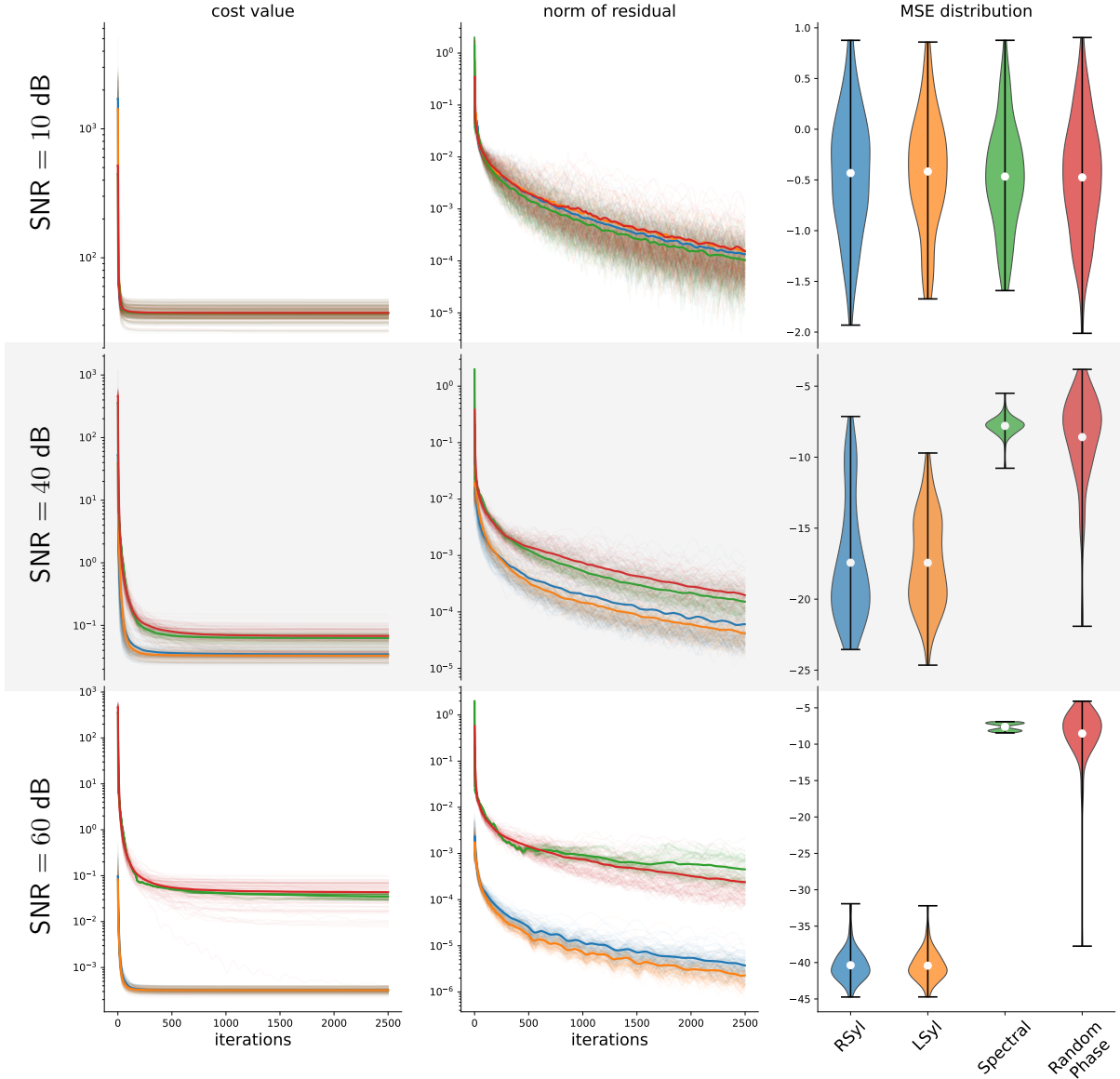




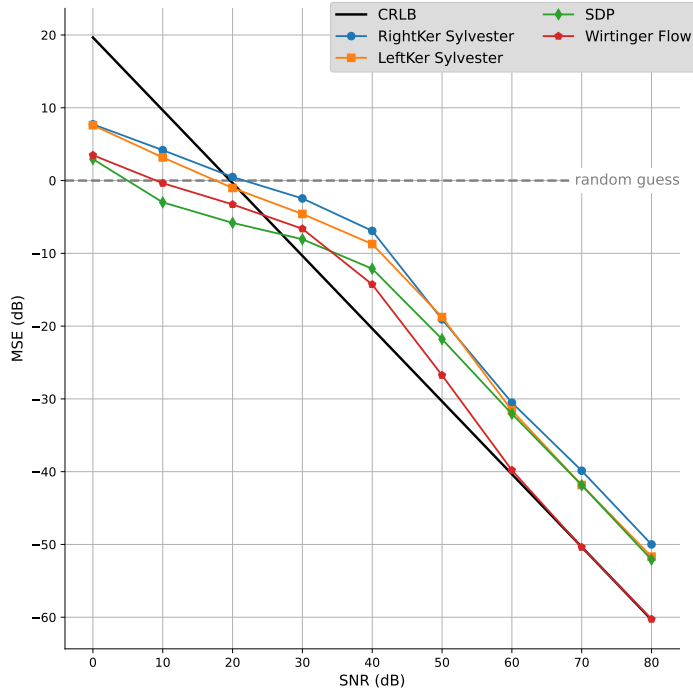
**Figure 3.** Reconstruction of a bivariate pulse ( $N = 64$ ) from noiseless PPR measurements ( $M = 2N - 1, P = 4$ ) using the different methods described in this paper. The reconstructed signal trace and squared error per time index  $n$  are shown for each approach.

715 **6.2. Comparison of initialization strategies for PPR-WF.** Choice of initial points in  
 716 nonconvex problems is usually a difficult but crucial task, as it directly impacts whether or  
 717 not the considered algorithm will be able to recover the global optimum of the problem. The  
 718 proposed PPR-WF algorithm does not avoid this key bottleneck, as already illustrated by  
 719 the bivariate pulse recovery experiment depicted in Figure 3. To assess the role played by  
 720 initial points in PPR-WF, we carefully benchmark the four initialization methods described  
 721 in Section 5.2, that is spectral initialization, random phase initialization, left and right kernel  
 722 Sylvester. We generated a random Gaussian complex-valued signal  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  with i.i.d.  
 723 entries of length  $N = 32$  such that  $\|\mathbf{X}\|_F = 1$  which was fixed for all experiments. PPR  
 724 noisy measurements (4.1) were considered for the simple measurement scheme (2.5) with  
 725  $M = 2N - 1, P = 4$ . We investigated three values of SNR, of 10, 40 and 60 dB respectively.  
 726 For each SNR value, we generated 100 independent noisy measurements and run the proposed  
 727 PPR-WF algorithm using the four aforementioned initialization procedures.

728 Figure 4 depicts obtained reconstruction results for the three SNR scenarios, where we  
 729 compare initialization methods in terms of cost function evolution  $F(\boldsymbol{\xi}^{(k)})$  and normed residual  
 730  $\|\boldsymbol{\xi}^{(k+1)} - \boldsymbol{\xi}^{(k)}\|_2 / \|\boldsymbol{\xi}^{(k)}\|_2$  decrease. Note that we imposed a identical number of 2500 iterations  
 731 of PPR-WF for each approach to ensure fair comparisons. We also plot the empirical distribu-  
 732 tion of squared error values for each initialization for further comparison of the quality of the  
 733 reconstructed signal (recall that squared error values are calculated after proper realignment  
 734 of the estimated signal with the ground truth). For SNR = 10 dB (which is a very challenging  
 735 scenario for PPR), there are no noticeable difference between initialization strategies: they  
 736 provide similar results in terms of cost value decrease, residual evolution and error distribu-  
 737 tion. For SNR = 40 dB, one starts to observe significant differences between Sylvester-based  
 738 approaches and spectral or random phase initializations. On average, Sylvester-based initial  
 739 points provides smaller optimal values, faster decrease of the residual and better reconstruc-



**Figure 4.** Comparison of initialization strategies for PPR-WF for the recovery of an arbitrary random bivariate signal of length  $N = 32$  with  $M = 2N - 1$  and  $P = 4$  noisy measurements. We benchmark spectral initialization, random phase initialization, left and right-kernel Sylvester initialization strategies in terms of cost function evolution, normed residual decrease and squared error distribution. Rows corresponds to values of SNR of 10, 40 and 60 dB, respectively. For each SNR value, left and middle panels present the evolution of the cost function and residual value with iterations, respectively. For each initialization method, thin colored lines indicate trajectories for each one of the 100 independent trials, and thick colored lines display their average respective average. The right panel provides violin plots representing a kernel density estimate of squared error distribution associated to each initialization strategy. White dots indicate MSE values and horizontal bars give extreme values for each squared error distribution.



**Figure 5.** Evolution of the MSE with the SNR for the four PPR reconstruction methods proposed in this paper. Ground truth is randomly generated bivariate signal with  $N = 32$ . Simple measurement scheme for  $M = 2N - 1$  and  $P = 4$  was used. Thick black line indicate the corresponding Cramér-Rao lower bound analytically derived in [Appendix C](#).

740 tion results in terms of squared error. This behavior is accentuated for  $\text{SNR} = 60$  dB, where  
 741 spectral and random phase initialization are unable to ensure convergence of PPR-WF to the  
 742 global optimum. This agrees with the observations made in [Figure 3](#) in the noiseless case for  
 743 spectral initialization.

744 These results demonstrate the importance of the choice of the initial point in PPR-WF  
 745 towards good convergence properties and recovery performance. Overall, left and right ker-  
 746 nel Sylvester initializations systematically outperform spectral and random phase strategies.  
 747 While the left kernel approach displays a slight advantage over the right kernel approach in  
 748 terms of residual decrease, it involves a much more important computational cost than its  
 749 right kernel counterpart. This explains why we recommend to use right kernel Sylvester ini-  
 750 tialization with PPR-WF for the best trade-off between algorithmic recovery performance and  
 751 computational time.

752 **6.3. Recovery performance with noisy measurements.** We now investigate the recovery  
 753 performances of the different proposed algorithms for PPR when dealing with noisy measure-  
 754 ments. We consider an additive white Gaussian noise model [\(4.1\)](#) for which the SNR is defined  
 755 in [\(4.2\)](#). We generated a ground truth signal  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  with i.i.d. Gaussian entries of length  
 756  $N = 32$  such that  $\|\mathbf{X}\|_F = 1$  and selected the simple,  $M = 2N - 1$ ,  $P = 4$  measurement scheme  
 757 [\(2.5\)](#). For a given SNR value, the MSE associated with each one of the proposed methods to

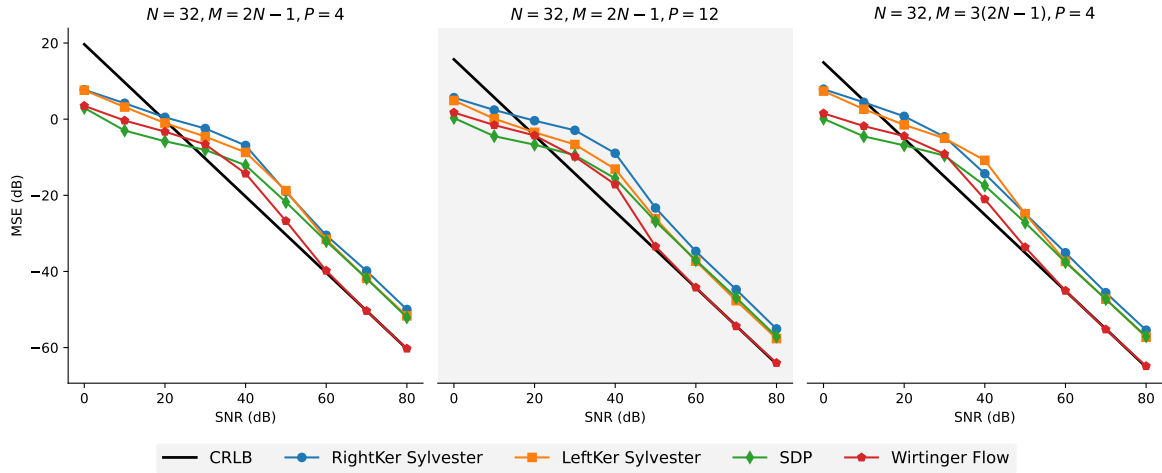
758 solve PPR was obtained by averaging of 100 independent reconstructions. Note that PPR-WF  
 759 uses the right kernel Sylvester initialization, following our analysis of initialization strategies  
 760 in Section 6.2.

761 Figure 5 displays the evolution of MSE for values of SNR ranging from 0 dB to 80 dB.  
 762 As expected, the MSE decreases as the SNR increases, independently from the considered  
 763 method. Overall, algorithmic methods (PPR-WF and SDP) outperform algebraic ones (left  
 764 and right kernel Sylvester) in terms of MSE values. More precisely, algebraic methods are not  
 765 informative in the “low-SNR” regime ( $\text{SNR} \leq 30$  dB) as they provide (relative) MSE values  
 766 above 0 dB, meaning that they do not provide a better reconstruction than a simple i.i.d.  
 767 random guess scaled to the ground truth norm. Furthermore we observe that SDP is more  
 768 robust to noise than PPR-WF. This agrees with the fact that SDP methods are known to be  
 769 robust to noise in general. Remarkably, the high-SNR regime ( $\geq 60$  dB) highlights several  
 770 distinctive behaviors. First, we observe that beyond  $\text{SNR} = 40$  dB, PPR-WF outperforms  
 771 all other methods, including SDP, by a few dB up to about 10 dB of relative MSE in the  
 772 asymptotic regime. Second, SDP do not longer outperforms left-kernel Sylvester, and only  
 773 improves from the right-kernel Sylvester approach by a small margin. This shows that, in this  
 774 high-SNR regime, the computational burden associated to the SDP approach becomes pro-  
 775 hibitive as 1) it provides no clear advantage over computationally cheaper algebraic methods  
 776 and 2) it clearly underperforms PPR-WF.

777 For completeness, we also provide the Cramèr-Rao lower bound (CRLB) for the noisy  
 778 PPR measurement model (4.1) to characterize a lower bound on the MSE of any unbiased  
 779 estimator of the ground truth signal. An analytical derivation of the resulting CRLB is  
 780 given in Appendix C. Figure 5 displays the CRLB on top of MSE values obtained for each  
 781 reconstruction method. We observe that the CRLB is not informative below  $\text{SNR} \leq 20$  dB  
 782 as all methods provide smaller MSE values – it simply means that the CRLB is particularly  
 783 pessimistic in this regime. On the contrary, the CRLB provides a meaningful lower bound  
 784 in the high-SNR regime. Importantly, it demonstrates that PPR-WF is an asymptotically  
 785 optimal reconstruction method for PPR since it attains the CRLB for  $\text{SNR} \geq 60$  dB.

786 **6.4. Influence of number of measurements.** One of the key advantages of the polarimet-  
 787 ric measurement model in PPR is that one can easily increase the number of measurements  
 788  $MP$  by performing more polarimetric projections, i.e., by increasing  $P$ . In fact, in practical  
 789 experiments it may be oftentimes easier to set up a new polarizer state  $\mathbf{b}_p$  than changing  
 790 the actual detector, which would be required if one desires to increase the number of Fourier  
 791 measurements  $M$ . Therefore, a natural question is the following: if one desires to increase the  
 792 total number of measurements  $MP$ , is it better – in terms of MSE – to increase the number of  
 793 Fourier measurements  $M$  or to increase the number of polarimetric projections  $P$ ? This is a  
 794 vast topic related to the question of experimental design, which requires a specific treatment  
 795 which is outside the scope of the present paper. Nonetheless, we provide in the sequel a first  
 796 study of the influence of the number of measurements in PPR for completeness.

797 Following the MSE performance analysis in Section 6.3, we use the same randomly gen-  
 798 erated ground truth signal  $N = 32$  and investigate the performances for two cases, i.e.,  $M =$   
 799  $2N - 1, P = 12$  and  $M = 3(2N - 1), P = 4$ , which lead to the same total number of measure-  
 800 ments  $MP$ . More precisely, the measurement scheme corresponding to each case is:



**Figure 6.** Comparison of the evolution of the MSE with respect to SNR for three measurements scheme  $M = 2N - 1, P = 4$  (left),  $M = 2N - 1, P = 12$  (center) and  $M = 3(2N - 1), P = 4$  (right). Experiments follow the same protocol as described in Section 6.3.

- 801 •  $M = 2N - 1, P = 12$  case: we use the correspondence between the 2-sphere and unit  
 802 vectors of  $\mathbb{C}^2$  to take advantage of optimal spherical tessellations such as HEALPix  
 803 [32]. In physical terms, it can be interpreted as finding one of the many possible Jones  
 804 vector  $\mathbf{b}_p$  corresponding to the Stokes parameters defining the rank-one matrix  $\mathbf{b}_p \mathbf{b}_p^H$ .  
 805 Formally, given Cartesian coordinates  $(s_p^x, s_p^y, s_p^z) \in \mathbb{R}^3$  of a point on the unit 2-sphere,  
 806 we define the projection vector  $\mathbf{b}_p$  as:

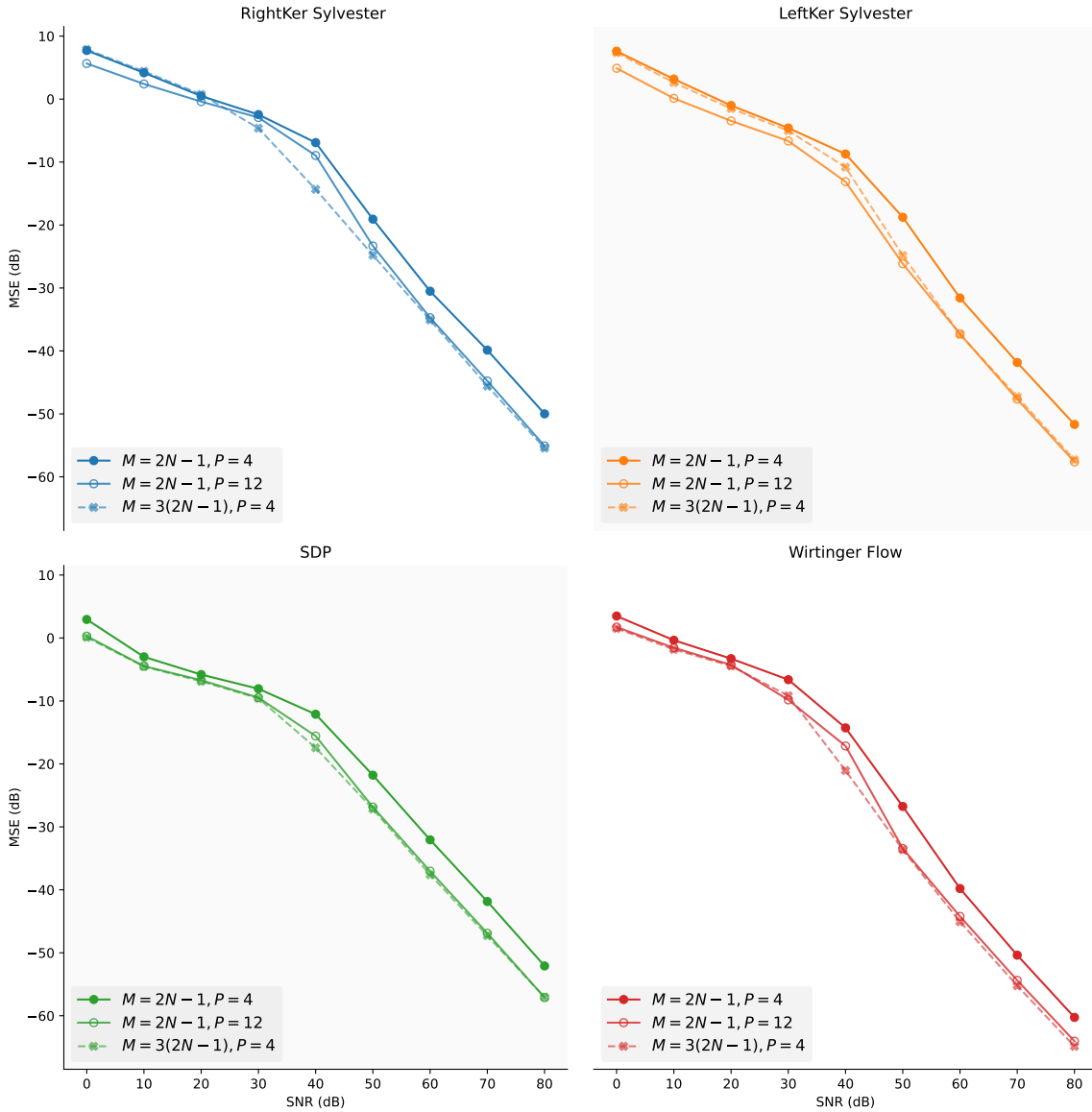
807 (6.2) 
$$\mathbf{b}_p = \frac{1}{\sqrt{2}\sqrt{1+s_p^z}} \begin{bmatrix} js_p^x \\ s_p^y + (1+s_p^z)j \end{bmatrix} \quad \text{if } s_p^z \neq -1, \quad \mathbf{b}_p = \begin{bmatrix} j \\ 0 \end{bmatrix} \quad \text{if } s_p^z = -1.$$

808 Note that our choice of  $P = 12$  corresponds to the first level of HEALPix sphere  
 809 discretization.

- 810 •  $M = 3(2N - 1), P = 4$  case: we keep the simple polarimetric measurement scheme  
 811 (2.4) and increase the number  $M$  of Fourier domain measurements.

812 Figure 6 depicts the MSE as a function of SNR for the two measurement setups described  
 813 above, where results from the experiment in Section 6.3 have been reproduced for better  
 814 comparison. As expected, increasing the total number of measurements  $MP$  improves overall  
 815 performance: this can be directly checked by remarking that the CRLB corresponding to  
 816  $M = 2N - 1, P = 12$  and  $M = 3(2N - 1), P = 4$  cases is lower than that of the  $M = 2N - 1, P = 4$   
 817 setup presented in Figure 5. Moreover, the different proposed reconstruction methods for PPR  
 818 behave similarly with one another as in our description made in Section 6.3. In particular,  
 819 we note that PPR-WF also attains the CRLB in these two new setups, proving again that it  
 820 establishes a versatile approach to solve PPR.

821 Figure 7 provides a side-by-side comparison of these three measurement schemes for each  
 822 reconstruction method. First, remark that  $M = 2N - 1, P = 12$  and  $M = 3(2N - 1), P =$



**Figure 7.** Side-by-side comparison of the behavior of each proposed reconstruction method for the three measurements scheme  $M = 2N - 1, P = 4$ ,  $M = 2N - 1, P = 12$  and  $M = 3(2N - 1), P = 4$ .

823 4 scheme have similar CRLB MSE bounds, with a slight advantage to the  $M = 3(2N -$   
 824  $1), P = 4$  case which can be observed on the PPR-WF panel. Second, we note that for  
 825 algorithmic approaches (SDP and PPR-WF), the difference concentrates in the mid-SNR  
 826 regime, i.e., between 30 dB and 50 dB, where oversampling in the Fourier domain offers slightly  
 827 MSE improvement over increasing the number of polarimetric projections. On the other  
 828 hand, for algebraic approaches we observe that performing more polarimetric measurements

usually improves the performance in the the low-SNR regime ( $\text{SNR} \leq 30$  dB), even though algebraic approaches do not perform well in this scenario. This performance improvement can be explained by the two-step nature of algebraic methods, which first need to reconstruct autocorrelation polynomials from polarimetric projections: in this case more polarimetric projections enable to reduce the reconstruction error in this first step.

**7. Conclusion.** This paper introduces a new model for Fourier phase retrieval called polarimetric phase retrieval (PPR), which takes advantage of polarization measurements in applications involving polarized light. The theoretical study of PPR relies on drawing careful equivalences with two other problems, namely bivariate phase retrieval (BPR) and polynomial autocorrelation factorization (PAF). In the noiseless case, these problems are found to be equivalent under very general conditions, which are summarized in Figure 1. A crucial result is Theorem 3.3: it shows that PAF admits a unique solution under very general conditions. Therefore, the original PPR problem admits a unique solution for almost every signals. Moreover, the PAF representation enables the use of algebraic reconstruction strategies for PPR based on GCD computations (Proposition 3.5). This original research direction is explored in detail in Section 4, where we propose two fully algebraic (i.e., non-iterative) algorithms for PPR relying on SVDs of Sylvester-like matrices. For completeness, Section 5 carefully adapts classical phase retrieval algorithms (SDP relaxation and Wirtinger-Flow) to solve the PPR problem. Section 6 provides extensive numerical experiments to benchmark the performances of each approach. These results demonstrate that, if one is interested in a scalable, computationally efficient and robust to noise reconstruction strategy, then both algebraic and iterative approaches should be combined. In practice, the best method for PPR appears to combine Wirtinger Flow (PPR-WF, Algorithm 5) with a carefully designed initialization based on right kernel Sylvester (Algorithm 1 with GCDs computations performed using Algorithm 2).

We believe that PPR opens promising new avenues for the exploitation of light polarization in Fourier phase retrieval problems. It enables the use of algebraic methods based on GCDs computations to solve the Fourier phase retrieval problem. While this research direction is particularly exciting, it also raises important challenges. For instance, an important issue to be addressed lies in improving the performance of algebraic methods at low SNR, e.g. with more robust estimation of the measurement polynomials or adding some prior information about the signal to be recovered (e.g. smoothness). A second challenge lies in extending the presented approaches to the case of polarized images, which is not straightforward at all since properties of polynomials with multiple variables (and their GCDs) differ considerably from their single variable counterpart. These questions will be addressed in future work.

## Appendix A. Relation between Fourier measurements and measurements polynomials.

*Proof of Lemma 3.1.* Recall that the discrete Fourier transform of  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  is denoted by  $\mathfrak{X}[m] = [\mathfrak{X}_1[m], \mathfrak{X}_2[m]]^\top \in \mathbb{C}^2$  for  $m = 0, 1, \dots, M-1$ , see (2.2). Then the Fourier entries can be related to polynomials  $X_1(z)$  and  $X_2(z)$  by comparing (2.2) with (3.1):

$$\mathfrak{X}_1[m] = X_1\left(e^{-j2\pi\frac{m}{M}}\right), \quad \mathfrak{X}_2[m] = X_2\left(e^{-j2\pi\frac{m}{M}}\right),$$

for any  $m = 0, 1, \dots, M-1$ . Similarly, comparing (2.2) with (3.2), their conjugates can be

870 expressed through the conjugate reflection polynomials  $\tilde{X}_1(z)$  and  $\tilde{X}_2(z)$

$$871 \quad \mathfrak{X}_1^*[m] = X_1^* \left( e^{-j2\pi \frac{m}{M}} \right) = \sum_{n=0}^{N-1} x_1[n]^* e^{2\pi j \frac{nm}{M}} = e^{j2\pi \frac{m(N-1)}{M}} \tilde{X}_1 \left( e^{-j2\pi \frac{m}{M}} \right),$$

$$872 \quad \mathfrak{X}_2^*[m] = X_2^* \left( e^{-j2\pi \frac{m}{M}} \right) = \sum_{n=0}^{N-1} x_2[n]^* e^{2\pi j \frac{nm}{M}} = e^{j2\pi \frac{m(N-1)}{M}} \tilde{X}_2 \left( e^{-j2\pi \frac{m}{M}} \right).$$

873

874 As a result, thanks to (2.3), BPR measurements can be expressed in terms of measurement  
875 polynomials  $\Gamma_{ij}(z)$  as follows:

$$876 \quad \mathfrak{F}[m] = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m]\mathfrak{X}_2[m]^* \\ \mathfrak{X}_2[m]\mathfrak{X}_1[m]^* & |\mathfrak{X}_2[m]|^2 \end{bmatrix}$$

$$877 \quad = e^{j2\pi \frac{m(N-1)}{M}} \begin{bmatrix} X_1 \left( e^{-j2\pi \frac{m}{M}} \right) \tilde{X}_1 \left( e^{-j2\pi \frac{m}{M}} \right) & X_1 \left( e^{-j2\pi \frac{m}{M}} \right) \tilde{X}_2 \left( e^{-j2\pi \frac{m}{M}} \right) \\ X_2 \left( e^{-j2\pi \frac{m}{M}} \right) \tilde{X}_1 \left( e^{-j2\pi \frac{m}{M}} \right) & X_2 \left( e^{-j2\pi \frac{m}{M}} \right) \tilde{X}_2 \left( e^{-j2\pi \frac{m}{M}} \right) \end{bmatrix}$$

$$878 \quad = e^{j2\pi \frac{m(N-1)}{M}} \mathbf{\Gamma} \left( e^{-j2\pi \frac{m}{M}} \right),$$

879

880 which completes the proof. ■

881 *Proof of Theorem 3.2.* The proof essentially comes down to showing the one-to-one corre-  
882 spondences summarized in Figure 1. More precisely, we show the one-to-one correspondence  
883 between the data (measurement matrix polynomial  $\mathbf{\Gamma}(z)$  in PAF, spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$   
884 in BPR) as well as the one-to-one correspondence between sets of solutions (polynomials  $X_1(z)$   
885 and  $X_2(z)$  in PAF, vectors components  $\mathbf{x}_1$  and  $\mathbf{x}_2$  in BPR). First note that the mapping be-  
886 tween  $\mathbb{C}^N$  and  $\mathbb{C}_{\leq N-1}[z]$  is a linear one-to-one map (and is an isomorphism):

$$887 \quad \mathbf{a} = [a[0] \ a[1] \ \cdots \ a[N-1]]^\top \mapsto A(z) = a[0] + za[1] + \cdots + z^{N-1}a[N-1].$$

888 Hence, the signals  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{C}^N$  can be uniquely recovered from the polynomials  $X_1, X_2 \in$   
889  $\mathbb{C}_{\leq N-1}[z]$  and vice versa. Similarly, thanks to (3.5), the Fourier covariance measurements  
890  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  are a linear transformation of the sequence

$$891 \quad \left\{ \mathbf{\Gamma} \left( e^{-j2\pi \frac{m}{M}} \right) \right\}_{m=0}^{M-1}$$

892 of evaluations of the matrix polynomial  $\mathbf{\Gamma}(z)$  at a set of  $M$  distinct points  $\{e^{-j2\pi \frac{m}{M}}\}_{m=0}^{M-1}$  on  
893 the complex plane. If  $M \geq 2N - 1$  (the degree of the polynomials plus one), then it is known  
894 that the coefficients of the polynomials can be uniquely recovered from the evaluations at  $M$   
895 distinct points, and therefore the following map is an injection

$$896 \quad \mathbb{C}_{\leq 2N-2}^{2 \times 2} \rightarrow (\mathbb{C}^{2 \times 2})^M$$

$$897 \quad \mathbf{\Gamma}(z) \mapsto \{\mathfrak{F}[m]\}_{m=0}^{M-1}, \quad \blacksquare$$

898

899 which completes the proof.



## Appendix B. Sylvester matrices and greatest common divisors.

*Proof of Proposition 4.3.* We first note that the result on the rank of  $\mathcal{S}_D(A, B)$  is known (see, for example, [64, Theorem 4.7]). Thus, we are left prove the second part, which is somewhat related to [64, Remark 4.8]. We write  $A(z) = F(z)Q(z)$ ,  $B(z) = G(z)Q(z)$ , so that  $\gcd(A, B) = Q(z)$  and  $F, G \in \mathbb{C}_{\leq L-K}[z]$ . Consider the following multiplication matrix

$$\mathbf{M}_{2L-D-K}(\mathbf{q}) = \underbrace{\begin{bmatrix} q_0 & & & & \\ \vdots & \ddots & & & \\ q_K & & q_0 & & \\ & & & \ddots & \\ & & & & q_K \end{bmatrix}}_{2L-D-K+1 \text{ columns}},$$

and our first goal is to show that the range of  $\mathcal{S}_D(A, B)$  is a subset of the range of  $\mathbf{M}_{2L-D-K}(\mathbf{q})$ . Indeed, the range of  $\mathcal{S}_D(A, B)$  corresponds to all polynomials  $R(z) \in \mathbb{C}_{\leq 2L-D}[z]$  that can be represented as

$$(B.1) \quad R(z) = U(z)A(z) + V(z)B(z) = Q(z)(U(z)F(z) + V(z)G(z)),$$

and therefore any element in the range of  $\mathcal{S}_D(A, B)$  belongs to the range of  $\mathbf{M}_{2L-D-K}(\mathbf{q})$  (since the range of  $\mathbf{M}_{2L-D-K}(\mathbf{q})$  corresponds to all polynomials of the form  $Q(z)H(z)$  with  $H \in \mathbb{C}_{\leq 2L-D-K}[z]$ ). Next we note that  $\mathbf{M}_{2L-D-K}(\mathbf{q})$  is full column rank and therefore the ranks of  $\text{colspan}(\mathcal{S}_D(A, B))$  and  $\text{colspan}(\mathbf{M}_{2L-D-K}(\mathbf{q}))$  are equal. Hence the ranges of the two matrices coincide, as well as the left kernels; in particular the following equivalence holds true

$$\mathbf{u}^\top \mathcal{S}_D(A, B) = 0 \iff \mathbf{u}^\top \mathbf{M}_{2L-D-K}(\mathbf{q}) = 0.$$

Finally, easy algebraic calculations (see also, for instance, [64, Eq. (33)]) show that

$$\mathbf{u}^\top \mathbf{M}_{2L-D-K}(\mathbf{q}) = \mathbf{q}^\top \mathcal{H}_{K+1}(\mathbf{u}),$$

where  $\mathcal{H}_{K+1}(\mathbf{u})$  is the Hankel matrix built from  $\mathbf{u}$  with  $K+1$  rows. This completes the proof.  $\blacksquare$

**Appendix C. Cramèr-Rao bound for PPR.** Several authors have considered Cramèr-Rao bounds for the classical phase retrieval problem with additive white gaussian noise [1, 2, 51]. These results directly apply to the additive Gaussian noise PPR model (4.1) since it can be equivalently rewritten as a particular one-dimensional noise model (the PPR-1D model introduced in Section 2.4). For completeness, we provide below an alternative derivation of the Cramèr-rao bound described in [51], where we use a full complex-domain approach instead of considering separate Cramèr-Rao bounds on amplitude and phase. Since measurement noise  $n_{m,p}$  in (4.1) is i.i.d. Gaussian distributed with variance  $\sigma^2$ , the probability density function

929 of the vector of observations  $\mathbf{y}$  is given by

$$930 \quad (\text{C.1}) \quad p(\mathbf{y}|\boldsymbol{\xi}) = \prod_{m=0}^{M-1} \prod_{p=0}^{P-1} p(y_{m,p}|\boldsymbol{\xi})$$

$$931 \quad (\text{C.2}) \quad = \prod_{m=0}^{M-1} \prod_{p=0}^{P-1} \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi})^2}{2\sigma^2} \right],$$

932  
933 where we recall that  $\mathbf{C}_{m,p} = \mathbf{c}_{m,p} \mathbf{c}_{m,p}^H$  with  $\mathbf{c}_{m,p} = \mathbf{b}_p^* \otimes \mathbf{a}_m$  by definition. The log-likelihood  
934 of observations reads

$$935 \quad (\text{C.3}) \quad \log p(\mathbf{y}|\mathbf{x}_{\text{vec}}) = -\frac{MP}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi})^2.$$

936 Since one wants to estimate the complex parameter vector  $\boldsymbol{\xi}$ , it is necessary to use the complex  
937 Fisher Information Matrix (FIM) [65, 43, 49], which reads

$$938 \quad (\text{C.4}) \quad \mathcal{J}_{\boldsymbol{\xi}} = \begin{bmatrix} \mathcal{I}_{\boldsymbol{\xi}} & \mathcal{P}_{\boldsymbol{\xi}} \\ \mathcal{P}_{\boldsymbol{\xi}}^* & \mathcal{I}_{\boldsymbol{\xi}}^* \end{bmatrix} \in \mathbb{C}^{4N \times 4N},$$

939 where entries are defined using Wirtinger derivatives [41] since  $\boldsymbol{\xi}$  is a complex vector:

$$940 \quad (\text{C.5}) \quad \mathcal{I}_{\boldsymbol{\xi}} = \mathbf{E} \left[ (\nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi})) (\nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}))^H \right],$$

$$941 \quad (\text{C.6}) \quad \mathcal{P}_{\boldsymbol{\xi}} = \mathbf{E} \left[ (\nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi})) (\nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}))^T \right].$$

942  
943 Note that the FIM  $\mathcal{J}_{\boldsymbol{\xi}}$  defined in (C.4) is isomorphic to the real FIM which would have been  
944 obtained by stacking the real and imaginary parts of  $\boldsymbol{\xi}$  in a single long vector [43]. This  
945 explains why  $\mathcal{J}_{\boldsymbol{\xi}}$  has dimensions  $4N \times 4N$ . Using properties of Wirtinger derivatives, we get

$$946 \quad (\text{C.7}) \quad \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) = -\frac{1}{\sigma^2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi}) \mathbf{C}_{m,p} \boldsymbol{\xi}.$$

947 This allows to compute explicitly the block terms  $\mathcal{I}_{\boldsymbol{\xi}}$  and  $\mathcal{P}_{\boldsymbol{\xi}}$  that define  $\mathcal{J}_{\boldsymbol{\xi}}$ . Using noise  
948 independence, one gets

$$949 \quad (\text{C.8}) \quad \mathcal{I}_{\boldsymbol{\xi}} = \frac{1}{\sigma^4} \mathbf{E} \left[ \left( \sum_{m,p} (y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi}) \mathbf{C}_{m,p} \boldsymbol{\xi} \right) \left( \sum_{m',p'} (y_{m',p'} - \boldsymbol{\xi}^H \mathbf{C}_{m',p'} \boldsymbol{\xi}) \boldsymbol{\xi}^H \mathbf{C}_{m',p'} \right) \right]$$

$$950 \quad (\text{C.9}) \quad = \frac{1}{\sigma^4} \sum_{m,p,m',p'} \mathbf{E} [n_{m,p} n_{m',p'}^*] \mathbf{C}_{m,p} \boldsymbol{\xi} \boldsymbol{\xi}^H \mathbf{C}_{m',p'}$$

$$951 \quad (\text{C.10}) \quad = \frac{1}{\sigma^2} \sum_{m,p} \mathbf{C}_{m,p} \boldsymbol{\xi} \boldsymbol{\xi}^H \mathbf{C}_{m,p}$$

$$952 \quad (\text{C.11}) \quad = \frac{1}{\sigma^2} \sum_{m,p} |\mathbf{c}_{m,p}^H \boldsymbol{\xi}|^2 \mathbf{c}_{m,p} \mathbf{c}_{m,p}^H.$$

953

954 Similar calculations leads to:

$$955 \quad (C.12) \quad \mathcal{P}_\xi = \frac{1}{\sigma^2} \sum_{ij} \mathbf{C}_{m,p} \xi(\xi)^\top \mathbf{C}_{m,p}^\top = \frac{1}{\sigma^2} \sum_{m,p} \left( \mathbf{c}_{m,p}^H \xi \right)^2 \mathbf{c}_{m,p} \mathbf{c}_{m,p}^\top.$$

956 A key result [49] is that the inverse of the complex FIM (C.4) provides a lower bound on  
 957 the covariance and pseudo-covariance of any unbiased estimator  $\hat{\xi}$  of the complex parameter  
 958  $\xi$ :

$$959 \quad (C.13) \quad \begin{bmatrix} \text{cov } \hat{\xi} & \text{pcov } \hat{\xi} \\ \text{pcov } \hat{\xi}^* & \text{cov } \hat{\xi}^* \end{bmatrix} \succeq \mathcal{J}_\xi^{-1}.$$

960 When the complex FIM is singular – as in phase retrieval [1, 2] –, one can show its pseudo-  
 961 inverse remains a valid lower bound for the MSE; following the discussion in [51], we still refer  
 962 to the resultant bound as the CRB with little abuse. In particular, we obtain the following  
 963 bound on the MSE on any unbiased PPR estimator  $\hat{\mathbf{X}}$  for the model (4.1):

$$964 \quad (C.14) \quad \text{MSE}(\hat{\mathbf{X}}) = \mathbf{E} \|\hat{\mathbf{X}} - \mathbf{X}\|_F^2 = \mathbf{E} \|\hat{\xi} - \xi\|_2^2 = \text{Tr} \text{cov } \hat{\xi} \geq \text{Tr} \left( \left[ \mathcal{J}_\xi^\dagger \right]_{[:2N, :2N]} \right)$$

965 where the subscript  $[:2N, :2N]$  denotes the restriction to the upper-left block of  $\mathcal{J}_\xi^\dagger$ .

966

## REFERENCES

- 967 [1] R. BALAN, *Reconstruction of signals from magnitudes of redundant representations: The complex*  
 968 *case*, Foundations of Computational Mathematics, 16 (2016), pp. 677–721, [https://doi.org/10.1007/](https://doi.org/10.1007/s10208-015-9261-0)  
 969 [s10208-015-9261-0](https://doi.org/10.1007/s10208-015-9261-0).
- 970 [2] A. S. BANDEIRA, J. CAHILL, D. G. MIXON, AND A. A. NELSON, *Saving phase: Injectivity and stability*  
 971 *for phase retrieval*, Applied and Computational Harmonic Analysis, 37 (2014), pp. 106–125, <https://doi.org/10.1016/j.acha.2013.10.002>.
- 972 [3] A. S. BANDEIRA, Y. CHEN, AND D. G. MIXON, *Phase retrieval from power spectra of masked signals*,  
 973 *Information and Inference: a Journal of the IMA*, 3 (2014), pp. 83–102, [https://doi.org/10.1093/](https://doi.org/10.1093/imaiai/iau002)  
 974 [imaiai/iau002](https://doi.org/10.1093/imaiai/iau002).
- 975 [4] A. H. BARNETT, C. L. EPSTEIN, L. GREENGARD, AND J. MAGLAND, *Geometry of the Phase Retrieval*  
 976 *Problem: Graveyard of Algorithms*, Cambridge University Press, 1 ed., Apr. 2022, [https://doi.org/](https://doi.org/10.1017/9781009003919)  
 977 [10.1017/9781009003919](https://doi.org/10.1017/9781009003919).
- 978 [5] A. BARONI, V. CHAMARD, AND P. FERRAND, *Extending Quantitative Phase Imaging to Polarization-*  
 979 *Sensitive Materials*, Physical Review Applied, 13 (2020), p. 054028, [https://doi.org/10.1103/](https://doi.org/10.1103/PhysRevApplied.13.054028)  
 980 [PhysRevApplied.13.054028](https://doi.org/10.1103/PhysRevApplied.13.054028).
- 981 [6] A. BARONI AND P. FERRAND, *Reference-free quantitative microscopic imaging of coherent arbitrary vec-*  
 982 *torial light beams*, Optics Express, 28 (2020), p. 35339, <https://doi.org/10.1364/OE.408665>.
- 983 [7] A. BECK, *First-order methods in optimization*, SIAM, 2017, <https://doi.org/10.1137/1.9781611974997>.
- 984 [8] R. BEINERT, *Non-negativity constraints in the one-dimensional discrete-time phase retrieval problem*,  
 985 *Information and Inference: A Journal of the IMA*, 6 (2017), pp. 213–224, <https://doi.org/https://doi.org/10.1093/imaiai/iaw018>.
- 986 [9] R. BEINERT, *One-dimensional phase retrieval with additional interference intensity measurements*, Results  
 987 *in Mathematics*, 72 (2017), pp. 1–24, <https://doi.org/10.1007/s00025-016-0633-9>.
- 988 [10] R. BEINERT AND G. PLONKA, *Ambiguities in one-dimensional discrete phase retrieval from fourier mag-*  
 989 *nitudes*, Journal of Fourier Analysis and Applications, 21 (2015), pp. 1169–1198, [https://doi.org/10.](https://doi.org/10.1007/s00041-015-9405-2)  
 990 [1007/s00041-015-9405-2](https://doi.org/10.1007/s00041-015-9405-2).

- 993 [11] R. BEINERT AND G. PLONKA, *Enforcing uniqueness in one-dimensional phase retrieval by additional sig-*  
994 *nal information in time domain*, Applied and Computational Harmonic Analysis, 45 (2018), pp. 505–  
995 525, <https://doi.org/10.1016/j.acha.2016.12.002>.
- 996 [12] T. BENDORY, R. BEINERT, AND Y. C. ELДАР, *Fourier Phase Retrieval: Uniqueness and Algorithms*, in  
997 Compressed Sensing and its Applications, H. Boche, G. Caire, R. Calderbank, M. März, G. Kutyniok,  
998 and R. Mathar, eds., Springer International Publishing, Cham, 2017, pp. 55–91, [https://doi.org/10.](https://doi.org/10.1007/978-3-319-69802-1_2)  
999 [1007/978-3-319-69802-1\\_2](https://doi.org/10.1007/978-3-319-69802-1_2). Series Title: Applied and Numerical Harmonic Analysis.
- 1000 [13] T. BENDORY AND D. EDIDIN, *Algebraic Theory of Phase Retrieval*, Notices of the American Mathematical  
1001 Society, 69 (2022), p. 1, <https://doi.org/10.1090/noti2540>.
- 1002 [14] T. BENDORY, Y. C. ELДАР, AND N. BOUMAL, *Non-convex phase retrieval from STFT measurements*,  
1003 IEEE Transactions on Information Theory, 64 (2017), pp. 467–484, [https://doi.org/10.1109/TIT.](https://doi.org/10.1109/TIT.2017.2745623)  
1004 [2017.2745623](https://doi.org/10.1109/TIT.2017.2745623).
- 1005 [15] E. J. CANDÈS, Y. C. ELДАР, T. STROHMER, AND V. VORONINSKI, *Phase retrieval via matrix completion*,  
1006 SIAM Journal on Imaging Sciences, 6 (2013), pp. 199–225, <https://doi.org/10.1137/110848074>.
- 1007 [16] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase retrieval from coded diffraction patterns*, Applied  
1008 and Computational Harmonic Analysis, 39 (2015), pp. 277–299, [https://doi.org/10.1016/j.acha.2014.](https://doi.org/10.1016/j.acha.2014.09.004)  
1009 [09.004](https://doi.org/10.1016/j.acha.2014.09.004).
- 1010 [17] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase Retrieval via Wirtinger Flow: Theory and*  
1011 *Algorithms*, IEEE Transactions on Information Theory, 61 (2015), pp. 1985–2007, [https://doi.org/](https://doi.org/10.1109/TIT.2015.2399924)  
1012 [10.1109/TIT.2015.2399924](https://doi.org/10.1109/TIT.2015.2399924).
- 1013 [18] H. N. CHAPMAN AND K. A. NUGENT, *Coherent lensless X-ray imaging*, Nature photonics, 4 (2010),  
1014 pp. 833–839, <https://doi.org/10.1038/nphoton.2010.240>.
- 1015 [19] R. A. CHIPMAN, G. YOUNG, AND W. S. T. LAM, *Polarized light and optical systems*, Optical sciences  
1016 and applications of light, Taylor & Francis, CRC Press, Boca Raton, 2018, [https://doi.org/10.1201/](https://doi.org/10.1201/9781351129121)  
1017 [9781351129121](https://doi.org/10.1201/9781351129121).
- 1018 [20] V. ELSER, *Phase retrieval by iterated projections*, Journal of the Optical Society of America A, 20 (2003),  
1019 pp. 40–55, <https://doi.org/10.1364/JOSAA.20.000040>.
- 1020 [21] V. ELSER, T.-Y. LAN, AND T. BENDORY, *Benchmark problems for phase retrieval*, SIAM Journal on  
1021 Imaging Sciences, 11 (2018), pp. 2429–2455, <https://doi.org/10.1137/18M1170364>.
- 1022 [22] A. FANNJIANG AND T. STROHMER, *The numerics of phase retrieval*, Acta Numerica, 29 (2020), pp. 125–  
1023 228, <https://doi.org/10.1017/S0962492920000069>.
- 1024 [23] P. FERRAND, M. ALLAIN, AND V. CHAMARD, *Ptychography in anisotropic media*, Optics Letters, 40  
1025 (2015), p. 5144, <https://doi.org/10.1364/OL.40.005144>.
- 1026 [24] P. FERRAND, A. BARONI, M. ALLAIN, AND V. CHAMARD, *Quantitative imaging of anisotropic material*  
1027 *properties with vectorial ptychography*, Optics Letters, 43 (2018), p. 763, [https://doi.org/10.1364/OL.](https://doi.org/10.1364/OL.43.000763)  
1028 [43.000763](https://doi.org/10.1364/OL.43.000763).
- 1029 [25] J. R. FIENUP, *Reconstruction of an object from the modulus of its Fourier transform*, Optics Letters, 3  
1030 (1978), pp. 27–29, <https://doi.org/10.1364/OL.3.000027>.
- 1031 [26] J. R. FIENUP, J. C. MARRON, T. J. SCHULZ, AND J. H. SELDIN, *Hubble space telescope characterized by*  
1032 *using phase-retrieval algorithms*, Applied Optics, 32 (1993), pp. 1747–1767, [https://doi.org/10.1364/](https://doi.org/10.1364/AO.32.001747)  
1033 [AO.32.001747](https://doi.org/10.1364/AO.32.001747).
- 1034 [27] J. J. GIL AND R. OSSIKOVSKI, *Polarized light and the Mueller matrix approach*, CRC press, Boca Raton,  
1035 2 ed., 2022, [https://doi.org/doi.org/10.1201/9780367815578](https://doi.org/10.1201/9780367815578).
- 1036 [28] D. H. GOLDSTEIN, *Polarized light*, CRC press, Boca Raton, 3 ed., 2017, <https://doi.org/10.1201/b10436>.
- 1037 [29] T. GOLDSTEIN, C. STUDER, AND R. BARANIUK, *A field guide to forward-backward splitting with a FASTA*  
1038 *implementation*, 2014, <https://doi.org/10.48550/arXiv.1411.3406>.
- 1039 [30] J. W. GOODMAN, *Introduction to Fourier optics*, Roberts and Company publishers, Englewood, CO,  
1040 3 ed., 2005.
- 1041 [31] J. P. GORDON AND H. KOGELNIK, *PMD fundamentals: Polarization mode dispersion in optical fibers*,  
1042 Proceedings of the National Academy of Sciences, 97 (2000), pp. 4541–4550, [https://doi.org/10.1073/](https://doi.org/10.1073/pnas.97.9.4541)  
1043 [pnas.97.9.4541](https://doi.org/10.1073/pnas.97.9.4541).
- 1044 [32] K. M. GORSKI, E. HIVON, A. J. BANDAY, B. D. WANDELT, F. K. HANSEN, M. REINECKE, AND  
1045 M. BARTELMANN, *HEALPix: A framework for high-resolution discretization and fast analysis of data*  
1046 *distributed on the sphere*, The Astrophysical Journal, 622 (2005), p. 759, <https://doi.org/10.1086/>

- 1047 427976.
- 1048 [33] P. GROHS, S. KOPPENSTEINER, AND M. RATHMAIR, *Phase retrieval: uniqueness and stability*, SIAM
- 1049 Review, 62 (2020), pp. 301–350, <https://doi.org/10.1137/19M1256865>.
- 1050 [34] S.-M. GUO, L.-H. YEH, J. FOLKESSON, I. E. IVANOV, A. P. KRISHNAN, M. G. KEEFE, E. HASHEMI,
- 1051 D. SHIN, B. B. CHHUN, N. H. CHO, M. D. LEONETTI, M. H. HAN, T. J. NOWAKOWSKI, AND S. B.
- 1052 MEHTA, *Revealing architectural order with quantitative label-free imaging and deep learning*, eLife, 9
- 1053 (2020), p. e55502, <https://doi.org/10.7554/eLife.55502>.
- 1054 [35] K. HUANG, Y. C. ELДАР, AND N. D. SIDIROPOULOS, *Phase Retrieval from 1D Fourier Mea-*
- 1055 *surements: Convexity, Uniqueness, and Algorithms*, IEEE Transactions on Signal Processing,
- 1056 64 (2016), pp. 6105–6117, <https://doi.org/10.1109/TSP.2016.2601291>, [http://ieeexplore.ieee.org/](http://ieeexplore.ieee.org/document/7547374/)
- 1057 [document/7547374/](http://ieeexplore.ieee.org/document/7547374/) (accessed 2020-04-03).
- 1058 [36] K. JAGANATHAN, Y. ELДАР, AND B. HASSIBI, *Phase retrieval with masks using convex optimization*,
- 1059 in 2015 IEEE International Symposium on Information Theory (ISIT), IEEE, 2015, pp. 1655–1659,
- 1060 <https://doi.org/10.1109/ISIT.2015.7282737>.
- 1061 [37] K. JAGANATHAN, Y. C. ELДАР, AND B. HASSIBI, *STFT phase retrieval: Uniqueness guarantees and*
- 1062 *recovery algorithms*, IEEE Journal of selected topics in signal processing, 10 (2016), pp. 770–781,
- 1063 <https://doi.org/10.1109/JSTSP.2016.2549507>.
- 1064 [38] K. JAGANATHAN AND B. HASSIBI, *Reconstruction of Signals From Their Autocorrelation and Cross-*
- 1065 *Correlation Vectors, With Applications to Phase Retrieval and Blind Channel Estimation*, IEEE
- 1066 Transactions on Signal Processing, 67 (2019), pp. 2937–2946, [https://doi.org/10.1109/TSP.2019.](https://doi.org/10.1109/TSP.2019.2911254)
- 1067 [2911254](https://doi.org/10.1109/TSP.2019.2911254).
- 1068 [39] K. JAGANATHAN, S. OYMAK, AND B. HASSIBI, *Sparse phase retrieval: Uniqueness guarantees and recovery*
- 1069 *algorithms*, IEEE Transactions on Signal Processing, 65 (2017), pp. 2402–2410, [https://doi.org/10.](https://doi.org/10.1109/TSP.2017.2656844)
- 1070 [1109/TSP.2017.2656844](https://doi.org/10.1109/TSP.2017.2656844).
- 1071 [40] X. JIANG, S. RAJAN, AND X. LIU, *Wirtinger Flow Method With Optimal Stepsize for Phase Re-*
- 1072 *trieval*, IEEE Signal Processing Letters, 23 (2016), pp. 1627–1631, [https://doi.org/10.1109/LSP.2016.](https://doi.org/10.1109/LSP.2016.2611940)
- 1073 [2611940](https://doi.org/10.1109/LSP.2016.2611940).
- 1074 [41] K. KREUTZ-DELGADO, *The Complex Gradient Operator and the CR-Calculus*, June 2009, [http://arxiv.](http://arxiv.org/abs/0906.4835)
- 1075 [org/abs/0906.4835](http://arxiv.org/abs/0906.4835).
- 1076 [42] B. LESHEM, R. XU, Y. DALLAL, J. MIAO, B. NADLER, D. ORON, N. DUDOVICH, AND O. RAZ, *Direct*
- 1077 *single-shot phase retrieval from the diffraction pattern of separated objects*, Nature Communications,
- 1078 7 (2016), p. 10820, <https://doi.org/10.1038/ncomms10820>.
- 1079 [43] B. LOESCH AND B. YANG, *Cramér-rao bound for circular and noncircular complex independent component*
- 1080 *analysis*, IEEE transactions on signal processing, 61 (2012), pp. 365–379, [https://doi.org/10.1109/](https://doi.org/10.1109/TSP.2012.2226166)
- 1081 [TSP.2012.2226166](https://doi.org/10.1109/TSP.2012.2226166).
- 1082 [44] A. M. MAIDEN AND J. M. RODENBURG, *An improved ptychographical phase retrieval algorithm for diffrac-*
- 1083 *tive imaging*, Ultramicroscopy, 109 (2009), pp. 1256–1262, [https://doi.org/10.1016/j.ultramic.2009.](https://doi.org/10.1016/j.ultramic.2009.05.012)
- 1084 [05.012](https://doi.org/10.1016/j.ultramic.2009.05.012).
- 1085 [45] I. MARKOVSKY, *Structured low-rank approximation and its applications*, Automatica, 44 (2008), pp. 891–
- 1086 909, <https://doi.org/10.1016/j.automatica.2007.09.011>.
- 1087 [46] J. MIAO, P. CHARALAMBOUS, J. KIRZ, AND D. SAYRE, *Extending the methodology of X-raycrystallography*
- 1088 *to allow imaging of micrometre-sized non-crystalline specimens*, Nature, 400 (1999), pp. 342–344,
- 1089 <https://doi.org/10.1038/22498>.
- 1090 [47] R. P. MILLANE, *Phase retrieval in crystallography and optics*, Journal of the Optical Society of America
- 1091 A, 7 (1990), pp. 394–411, <https://doi.org/10.1364/JOSAA.7.000394>.
- 1092 [48] R. MONTEIRO ET AL., *First-and second-order methods for semidefinite programming*, Mathematical Pro-
- 1093 *gramming*, 97 (2003), pp. 209–244, <https://doi.org/10.1007/s10107-003-0451-1>.
- 1094 [49] E. OLLILA, V. KOIVUNEN, AND J. ERIKSSON, *On the Cramér-Rao bound for the constrained and un-*
- 1095 *constrained complex parameters*, in 2008 5th IEEE Sensor Array and Multichannel Signal Processing
- 1096 *Workshop*, IEEE, 2008, pp. 414–418, <https://doi.org/10.1109/SAM.2008.4606902>.
- 1097 [50] O. PEDATZUR, A. TRABATTONI, B. LESHEM, H. SHALMONI, M. CASTROVILLI, M. GALLI, M. LUCCHINI,
- 1098 E. MÅNSSON, F. FRASSETTO, L. POLETTA, ET AL., *Double-blind holography of attosecond pulses*,
- 1099 Nature Photonics, 13 (2019), pp. 91–95, <https://doi.org/10.1038/s41566-018-0308-z>.
- 1100 [51] C. QIAN, N. D. SIDIROPOULOS, K. HUANG, L. HUANG, AND H. C. SO, *Phase retrieval using feasible*

- 1101 *point pursuit: Algorithms and Cramér–Rao bound*, IEEE Transactions on Signal Processing, 64 (2016),  
1102 pp. 5282–5296, <https://doi.org/10.1109/TSP.2016.2593688>.
- 1103 [52] J. RANIERI, A. CHEBIRA, Y. M. LU, AND M. VETTERLI, *Phase retrieval for sparse signals: Uniqueness*  
1104 *conditions*, 2013, <https://arxiv.org/abs/1308.3058>.
- 1105 [53] O. RAZ, N. DUDOVICH, AND B. NADLER, *Vectorial Phase Retrieval of 1-D Signals*, IEEE Transactions  
1106 on Signal Processing, 61 (2013), pp. 1632–1643, <https://doi.org/10.1109/TSP.2013.2239994>.
- 1107 [54] O. RAZ, B. LESHEM, J. MIAO, B. NADLER, D. ORON, AND N. DUDOVICH, *Direct phase retrieval in double*  
1108 *blind Fourier holography*, Optics Express, 22 (2014), p. 24935, <https://doi.org/10.1364/OE.22.024935>.
- 1109 [55] O. RAZ, O. SCHWARTZ, D. AUSTIN, A. S. WYATT, A. SCHIAVI, O. SMIRNOVA, B. NADLER, I. A. WALM-  
1110 SLEY, D. ORON, AND N. DUDOVICH, *Vectorial Phase Retrieval for Linear Characterization of Attosec-*  
1111 *ond Pulses*, Physical Review Letters, 107 (2011), p. 133902, [https://doi.org/10.1103/PhysRevLett.](https://doi.org/10.1103/PhysRevLett.107.133902)  
1112 [107.133902](https://doi.org/10.1103/PhysRevLett.107.133902).
- 1113 [56] D. SAYRE, *Some implications of a theorem due to Shannon*, Acta Crystallographica, 5 (1952), pp. 843–843,  
1114 <https://doi.org/10.1107/S0365110X52002276>.
- 1115 [57] B. SCHAEFER, E. COLLETT, R. SMYTH, D. BARRETT, AND B. FRAHER, *Measuring the Stokes polarization*  
1116 *parameters*, American Journal of Physics, 75 (2015), p. 6, <https://doi.org/10.1119/1.2386162>.
- 1117 [58] Y. SHECHTMAN, Y. C. ELДАР, O. COHEN, H. N. CHAPMAN, J. MIAO, AND M. SEGEV, *Phase Retrieval*  
1118 *with Application to Optical Imaging: A contemporary overview*, IEEE Signal Processing Magazine,  
1119 32 (2015), pp. 87–109, <https://doi.org/10.1109/MSP.2014.2352673>.
- 1120 [59] O. SMIRNOVA, S. PATCHKOVSKII, Y. MAIRESSE, N. DUDOVICH, D. VILLENEUVE, P. CORKUM, AND  
1121 M. Y. IVANOV, *Attosecond Circular Dichroism Spectroscopy of Polyatomic Molecules*, Physical Review  
1122 Letters, 102 (2009), p. 063601, <https://doi.org/10.1103/PhysRevLett.102.063601>.
- 1123 [60] Q. SONG, A. BARONI, R. SAWANT, P. NI, V. BRANDLI, S. CHENOT, S. VÉZIAN, B. DAMILANO,  
1124 P. DE MIERRY, S. KHADIR, ET AL., *Ptychography retrieval of fully polarized holograms from*  
1125 *geometric-phase metasurfaces*, Nature communications, 11 (2020), p. 2651, [https://doi.org/10.1038/](https://doi.org/10.1038/s41467-020-16437-9)  
1126 [s41467-020-16437-9](https://doi.org/10.1038/s41467-020-16437-9).
- 1127 [61] R. TREBINO, *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*, Springer  
1128 Science & Business Media, New York, 2000, <https://doi.org/10.1007/978-1-4615-1181-6>.
- 1129 [62] J. S. TYO, D. L. GOLDSTEIN, D. B. CHENAULT, AND J. A. SHAW, *Review of passive imaging polarimetry*  
1130 *for remote sensing applications*, Applied optics, 45 (2006), pp. 5453–5469, [https://doi.org/10.1364/](https://doi.org/10.1364/AO.45.005453)  
1131 [AO.45.005453](https://doi.org/10.1364/AO.45.005453).
- 1132 [63] K. USEVICH, J. FLAMANT, M. CLAUSEL, AND D. BRIE, *Uniqueness of rank-one auto-correlation matrix*  
1133 *polynomials factorization*. submitted, 2023, <https://hal.science/hal-04062934>.
- 1134 [64] K. USEVICH AND I. MARKOVSKY, *Variable projection methods for approximate (greatest) common divisor*  
1135 *computations*, Theoretical Computer Science, 681 (2017), pp. 176–198, [https://doi.org/10.1016/j.tcs.](https://doi.org/10.1016/j.tcs.2017.03.028)  
1136 [2017.03.028](https://doi.org/10.1016/j.tcs.2017.03.028).
- 1137 [65] A. VAN DEN BOS, *A Cramér-Rao lower bound for complex parameters*, IEEE Transactions on Signal  
1138 Processing, 42 (1994), <https://doi.org/10.1109/78.324755>.
- 1139 [66] L. VANDENBERGHE AND S. BOYD, *Semidefinite programming*, SIAM review, 38 (1996), pp. 49–95, <https://doi.org/10.1137/1038003>.
- 1140 [67] I. WALDSPURGER, A. D’ASPREMONT, AND S. MALLAT, *Phase recovery, maxcut and complex semidef-*  
1141 *inite programming*, Mathematical Programming, 149 (2015), pp. 47–81, [https://doi.org/10.1007/](https://doi.org/10.1007/s10107-013-0738-9)  
1142 [s10107-013-0738-9](https://doi.org/10.1007/s10107-013-0738-9).
- 1143 [68] I. A. WALMSLEY AND C. DORRER, *Characterization of ultrashort electromagnetic pulses*, Advances in  
1144 Optics and Photonics, 1 (2009), pp. 308–437, <https://doi.org/10.1364/AOP.1.000308>.
- 1145 [69] R. XU, M. SOLTANOLKOTABI, J. P. HALDAR, W. UNGLAUB, J. ZUSMAN, A. F. J. LEVI, AND R. M.  
1146 LEAHY, *Accelerated Wirtinger Flow: A fast algorithm for ptychography*, June 2018, [http://arxiv.org/](http://arxiv.org/abs/1806.05546)  
1147 [abs/1806.05546](http://arxiv.org/abs/1806.05546).
- 1148