



**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. SIAM Journal on Imaging Sciences, 2024, pp.632-671. 10.1137/23M1570971. hal-03613352v4

**HAL Id: hal-03613352**

**<https://hal.science/hal-03613352v4>**

Submitted on 11 Mar 2024

**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.** 49N30, 94A12, 12D05

**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 [61]. It has been – and continues to be – of tremendous importance for many applications areas involving optics, such as crystallography [23, 24, 50], astronomy [28, 29], coherent diffraction imaging (also known as lensless imaging) [49, 47], 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 [33]. Examples include one-dimensional (1D) temporal Fourier transforms performed by spectrometers in ultra-short laser pulse characterization [76] 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 [63] 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 [20]. 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 [30]. As a result, polarized light imaging has found many applications such as material characterization [34], remote sensing [70] or bio-imaging [37]. 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 [64, 59] 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) [69] 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* [26, 27] combines spatially redundant measurements with polarimetric measurements. This allows quantitative imaging of complex anisotropic media, such as biominerals [5, 6]. More generally, recent years have seen a growing interest in the experimental development of computational imaging systems exploiting vectorial and polarization properties of light, such as polarization-sensitive Fourier ptychography [22, 67], polarization-sensitive diffraction tomography [68, 60] or vectorial holography [66], among others.

*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 [36]; see also [13] for a discussion of its algebraic properties. A comprehensive tour of existing algorithms is given in [25]; 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 [42, 56] or minimum phase [38]. Another approach consists in generating additional measurements, e.g., using deterministic masks [39, 15], (randomly) coded diffraction patterns [16] or using redundant, overlapping measurements inspired by ptychography [14, 40].

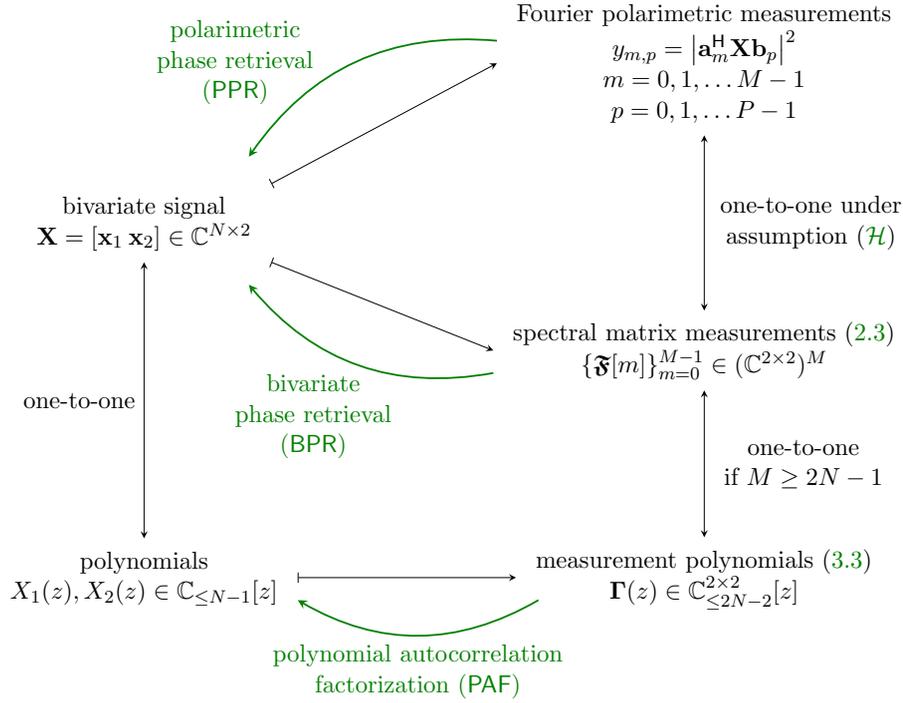
More closely related to the present work is the use of additional, interference-like measurements in Fourier phase retrieval. The main idea roots in a imaging technique known as *holography*, which involves the coherent interference of the object of interest  $\mathbf{x}$  with some reference signal  $\mathbf{y}$ . Pushing this idea further, authors have developed a strategy ensuring uniqueness in Fourier phase retrieval, called *vectorial phase retrieval* [57] or double-blind holography [45, 58, 54]. 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 [41], 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,

82 which limits their use in contexts where one is interested in reconstructing the polarized (or  
83 bivariate) electromagnetic field (such as in polarized coherent diffraction imaging techniques  
84 [65]). In particular, we will show that the proposed polarimetric Fourier phase retrieval  
85 model encompasses vectorial phase retrieval as a special case, for a specific choice of *four*  
86 *polarimetric projections*. In addition, while the connection between vectorial phase retrieval  
87 and greatest common divisor of polynomials was observed in [41], it was not investigated in  
88 detail as the authors focused on a semidefinite programming relaxation. In contrast, algebraic  
89 approaches based on greatest common divisor computations are a cornerstone of the proposed  
90 methodology for the polarimetric Fourier phase retrieval model.

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

107 *Organization of the paper.* A crucial feature of the present paper is the extensive use of  
108 equivalences between the polarimetric Fourier phase retrieval (*PPR*) problem and two other  
109 problems, namely bivariate Fourier phase retrieval (*BPR*) and polynomial autocorrelation fac-  
110 torization (*PAF*). For reference, these equivalences are stated in [Figure 1](#), with pointers to  
111 relevant definitions and equations. [Section 2](#) introduces the *PPR* model and discusses its physi-  
112 cal interpretations in terms of polarimetric measurement. Under some very general conditions,  
113 the equivalence with *BPR* is then established, which permits the study of trivial ambiguities.  
114 The relation of *PPR* with a standard 1D Fourier phase retrieval problem is also discussed.  
115 [Section 3](#) starts by reformulating the *BPR* problem using a polynomial representation, leading  
116 to *PAF*. Then, we leverage uniqueness results on multivariate spectral representations [71]  
117 to establish a necessary and sufficient characterization of uniqueness in *PAF* ([Theorem 3.5](#)).  
118 [Corollary 3.6](#) states that *PAF* is almost everywhere unique, and as a result, an algebraic  
119 solution can be found using greatest common divisors of measurement polynomials ([Propo-  
120 sition 3.7](#)). [Section 4](#) goes back to *PPR* and exploits uniqueness results to propose a fully  
121 algebraic reconstruction method for *PPR* ([Algorithm 1](#)) based on two variations of approxi-  
122 mate greatest common divisor computations. [Section 5](#) focus instead on iterative algorithms  
123 for *PPR*, by tailoring semidefinite relaxation ([Algorithm 4](#)) and Wirtinger Flow ([Algorithm  
124 5](#)). [Section 6](#) presents several numerical experiments to illustrate and assess the practical  
125 performances of the proposed reconstruction strategies. [Section 7](#) collects concluding remarks



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

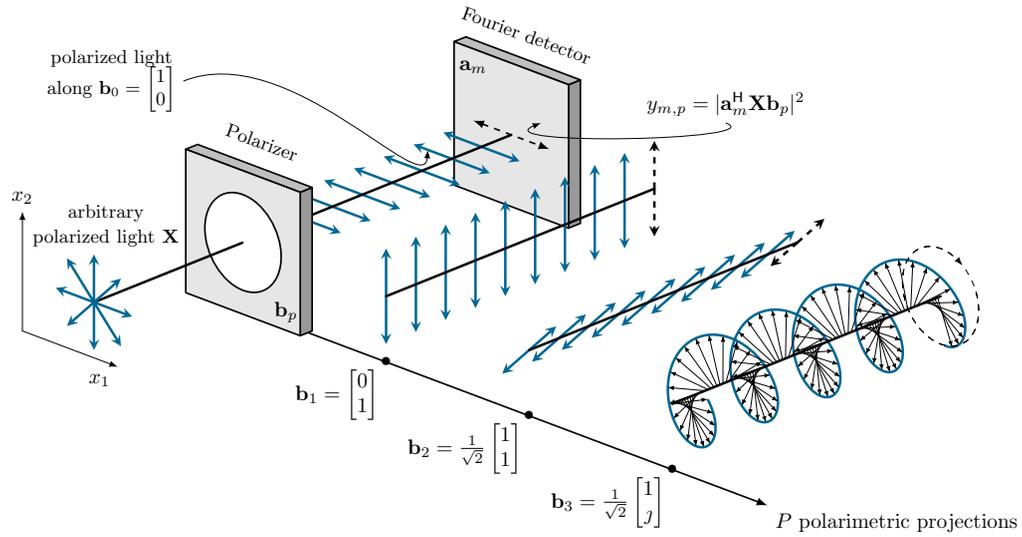
126 and Appendices gather technical details and proofs.

127 **Notations.** In this paper, we denote by  $\mathbb{R}$  the set of real numbers and by  $\mathbb{C}$  the set of  
 128 complex numbers with imaginary unit  $j$  such that  $j^2 = -1$ . Vectors and matrices are denoted  
 129 in bold lowercase letters and bold capital letters, respectively. Dependence of quantities in  
 130 terms of a discrete index are indicated by brackets, i.e.,  $\mathbf{x}[n]$  denotes the  $n$ -th entry of the set  
 131 of vectors  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$ . Notation  $\mathbf{a}^*$ ,  $\mathbf{A}^*$  indicate the complex conjugate of vector  $\mathbf{a}$  and matrix  
 132  $\mathbf{A}$ , respectively. The transpose of a matrix  $\mathbf{A}$  is  $\mathbf{A}^T$  and its conjugate transpose is given  
 133 by  $\mathbf{A}^H$ . Fourier domain quantities are denoted using capital gothic letters, i.e., the vector  
 134  $\mathfrak{X}[m] \in \mathbb{C}^2$  denotes the  $m$ -th entry of the (one-dimensional) discrete Fourier transform of the  
 135 vector signal  $\{\mathbf{x}[n] \in \mathbb{C}^2\}_{n=0}^{N-1}$ , evaluated at a frequency indexed by integer  $m$ .

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

138 **2.1. General formulation.** Consider a discrete bivariate signal  $\mathbf{x}[n] = (x_1[n], x_2[n])^T \in \mathbb{C}^2$   
 139 defined for  $n = 0, 1, \dots, N-1$ . Let  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  be the matrix representation of  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$   
 140 obtained by stacking samples row-wise such that

$$141 \quad (2.1) \quad \mathbf{X} = \begin{bmatrix} x_1[0] & x_2[0] \\ x_1[1] & x_2[1] \\ \vdots & \vdots \\ x_1[N-1] & x_2[N-1] \end{bmatrix} = [\mathbf{x}_1 \ \mathbf{x}_2],$$



**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.

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

$$145 \quad (\text{PPR}) \quad \text{find } \mathbf{X} \in \mathbb{C}^{N \times 2} \text{ given measurements } y_{m,p} = \left| \mathbf{a}_m^H \mathbf{X} \mathbf{b}_p \right|^2, \\ m = 0, 1, \dots, M-1, \quad p = 0, 1, \dots, P-1$$

146 where  $\mathbf{a}_m \in \mathbb{C}^N$  is the discrete Fourier vector corresponding to frequency  $f_m = (2\pi m)/M$ ,  
 147 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  
 148  $\|\mathbf{b}_p\|_2^2 = 1$ , denotes an arbitrary projection acting on the two vector components of  $\mathbf{X}$ .

149 Figure 2 permits to attach precise physical interpretations of PPR measurements in terms  
 150 of polarization optics. The matrix  $\mathbf{X}$  represents the one-dimensional bivariate electromagnetic  
 151 field, where each row is a vector of  $\mathbb{C}^2$  describing an arbitrary polarization state (the so-called  
 152 Jones vector [30]). This states passes through a polarizer defined by  $\mathbf{b}_p \in \mathbb{C}^2$ , evaluating the  
 153 projection of polarization states of  $\mathbf{X}$  onto  $\mathbf{b}_p$ . Finally, light impinges on a Fourier detector  
 154 described by  $\mathbf{a}_m \in \mathbb{C}^N$ , leading to squared magnitude PPR measurements  $y_{m,p}$ .

155 The measurement model PPR can be easily implemented experimentally. Indeed, Fourier  
 156 vectors  $\{\mathbf{a}_m\}_{m=0}^{M-1}$  correspond to far-field measurements in optics, as encountered in coherent  
 157 diffraction imaging techniques (for the case of 2D/3D images) or in spectrometry (for the  
 158 1D case of ultra-short pulses). On the other hand, the set  $\{\mathbf{b}_p\}_{p=0}^{P-1}$  describes the different  
 159 polarizers (or polarization analysers) required to measure polarization of light. Any arbitrary  
 160 polarizer (in mathematical terms, any unit-norm vector  $\mathbf{b}_p \in \mathbb{C}^2$ ) can be constructed as as  
 161 combination of standard optical components, such as linear polarizers or waveplates [30].

162 Therefore, polarimetric measurements are very flexible: their number, as well as the reference  
163 polarization states  $\{\mathbf{b}_p\}_{p=0}^{P-1}$  can be tailored at will depending on the context.

164 **2.2. Relation with Fourier matrix measurements.** A closely related problem to PPR is  
165 the *bivariate phase retrieval* (BPR) problem. Let us introduce the discrete Fourier transform  
166 of the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  as

$$167 \quad (2.2) \quad \mathfrak{X}[m] = \sum_{n=0}^{N-1} \mathbf{x}[n] \exp\left(-2\pi j \frac{mn}{M}\right) = \begin{bmatrix} \mathfrak{X}_1[m] \\ \mathfrak{X}_2[m] \end{bmatrix} = (\mathbf{a}_m^H \mathbf{X})^\top \in \mathbb{C}^2$$

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

$$169 \quad (2.3) \quad \mathfrak{F}[m] = \mathfrak{X}[m] \mathfrak{X}[m]^H = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m] \mathfrak{X}_2[m]^* \\ \mathfrak{X}_2[m] \mathfrak{X}_1[m]^* & |\mathfrak{X}_2[m]|^2 \end{bmatrix} \in \mathbb{C}^{2 \times 2}.$$

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

$$174 \quad (\text{BPR}) \quad \text{find } \mathbf{X} \in \mathbb{C}^{N \times 2} \text{ given spectral matrix measurements } \{\mathfrak{F}[m]\}_{m=0}^{M-1}.$$

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

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

$$179 \quad (\mathcal{H}) \quad \text{span}_{\mathbb{R}} \left\{ \mathbf{b}_p \mathbf{b}_p^H \right\}_{p=0}^{P-1} = \left\{ \mathbf{M} \in \mathbb{C}^{2 \times 2} \mid \mathbf{M}^H = \mathbf{M} \right\},$$

180 *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*  
181 *Hermitian matrices. Then, under assumption  $(\mathcal{H})$ , the problem PPR is equivalent to BPR in*  
182 *the sense that  $\mathbf{X}$  is a solution of the problem PPR if and only if  $\mathbf{X}$  is solution of BPR.*

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

$$187 \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 = \mathfrak{X}[m]^\top \mathbf{b}_p \mathbf{b}_p^H \mathfrak{X}^*[m] = \text{Tr } \mathbf{b}_p^* \mathbf{b}_p^\top \mathfrak{F}[m],$$

188 *i.e., measurements  $y_{m,p}$  are linear measurements of  $\mathfrak{F}[m]$  through sensing matrices  $\{\mathbf{b}_p^* \mathbf{b}_p^\top\}_{p=0}^{P-1}$ .*  
189 *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*  
190 *space of 2-by-2 Hermitian by matrices by assumption  $(\mathcal{H})$ , the spectral matrix  $\mathfrak{F}[m]$  can be*  
191 *uniquely determined from  $\{y_{m,p}\}_{p=0}^{P-1}$  by linear combinations. This concludes the proof. ■*

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

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

$$197 \quad (2.4) \quad \mathbf{b}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{b}_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{b}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{b}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ j \end{bmatrix}.$$

198 The projection vectors  $\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2$  and  $\mathbf{b}_3$  correspond to Jones vectors of standard polarizers used  
 199 in optics [20], which are, respectively: horizontal linear polarizer, vertical linear polarizer,  $45^\circ$   
 200 linear polarizer and left circular polarizer. See Figure 2 for an illustration. A direct check  
 201 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  
 202 of 2-by-2 Hermitian matrices, and as a result, they are a generating family of such matrices.  
 203 **PPR** measurements read explicitly

$$204 \quad (2.5) \quad \begin{aligned} y_{m,0} &= |\mathfrak{X}_1[m]|^2, & y_{m,1} &= |\mathfrak{X}_2[m]|^2, \\ y_{m,2} &= \frac{1}{2} |\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2, & y_{m,3} &= \frac{1}{2} |\mathfrak{X}_1[m] + j\mathfrak{X}_2[m]|^2. \end{aligned}$$

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

$$\begin{aligned} 207 \quad \text{real}(\mathfrak{X}_1[m]\mathfrak{X}_2[m]^*) &= \frac{1}{2} \left( |\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2 \right) \\ 208 &= y_{m,2} - \frac{1}{2} (y_{m,0} + y_{m,1}), \\ 209 \quad \text{imag}(\mathfrak{X}_1[m]\mathfrak{X}_2[m]^*) &= \frac{1}{2} \left( |\mathfrak{X}_1[m] + j\mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2 \right) \\ 210 &= y_{m,3} - \frac{1}{2} (y_{m,0} + y_{m,1}). \end{aligned}$$

212 Remark that the measurement scheme (2.4) yields the same quadratic measurements (2.5)  
 213 as proposed by several authors [57, 41, 45, 58, 54]. Because of that, **BPR** is equivalent to  
 214 the vectorial phase retrieval problem originally introduced in [57]. This shows that **PPR**  
 215 encompasses existing measurements strategies as a special case, while bringing extra flexibility  
 216 in the experimental design of measurements. One of the key benefits of the **PPR** model is  
 217 that additional polarimetric measurements can be generated at will using simple off-the-shelf  
 218 optical components such as linear polarizers or waveplates.

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

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

227 **Shifts.** This trivial ambiguity only appears when the bivariate signal  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  has not full  
 228 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  $n \leq n_a$   
 229 and  $n \geq n_b$ . Assuming this is the case, define the shifted signal  $\{\mathbf{x}'[n]\}_{n=0}^{N-1}$  as  $\mathbf{x}'[n] = \mathbf{x}[n+n_0]$   
 230 where  $n_0$  is a relative integer between  $(n_b - N)$  and  $(n_a + 1)$  as to ensure proper support. Then,  
 231 using standard Fourier transform properties one gets that  $\mathfrak{X}'[m] = \exp(j2\pi n_0 m/M) \mathfrak{X}[m]$ , so  
 232 that in turn  $\mathfrak{F}'[m] = \mathfrak{F}[m]$  for every  $m$ .

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

$$235 \quad (2.6) \quad \mathfrak{F}'[m] = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_2[m] \mathfrak{X}_1^*[m] \\ \mathfrak{X}_1[m] \mathfrak{X}_2^*[m] & |\mathfrak{X}_2[m]|^2 \end{bmatrix} = \mathfrak{F}[m]^\top.$$

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

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

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

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

$$256 \quad (2.7) \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 = |(\mathbf{b}_p^\top \otimes \mathbf{a}_m^H) \boldsymbol{\xi}|^2 = |(\mathbf{b}_p^* \otimes \mathbf{a}_m)^H \boldsymbol{\xi}|^2$$

257 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  
 258 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

$$259 \quad (\text{PPR-1D}) \quad \text{find } \boldsymbol{\xi} \in \mathbb{C}^{2N} \text{ given measurements } y_{m,p} = \left| \mathbf{c}_{m,p}^H \boldsymbol{\xi} \right|^2 \\ m = 0, 1, \dots, M-1, \quad p = 0, 1, \dots, P-1$$

260 This shows that **PPR** can be rewritten as a specific instance of 1D phase retrieval with struc-  
 261 tured measurements vectors  $\mathbf{c}_{m,p} \in \mathbb{C}^{2N}$ . While being mathematically sound, the equivalent  
 262 **PPR-1D** problem brings almost no insights about the bivariate nature of the signal to be

263 recovered. Moreover, **PPR-1D** cannot be interpreted as a Fourier phase retrieval problem  
 264 with masks [3, 39], since measurements vectors  $\mathbf{c}_{m,p}$  intertwine Fourier measurements  $\mathbf{a}_m$   
 265 and polarimetric projections  $\mathbf{b}_p$  using Kronecker products. Thus, the study of the theoretical  
 266 properties of **PPR** cannot be inferred from standard phase retrieval properties applied to **PPR-**  
 267 **1D**. This requires a dedicated study, which is described in detail in **Section 3** and exploited  
 268 in **Section 4** to formulate algebraic solutions to the **PPR** problem. Nonetheless, as we shall  
 269 see in **Section 5**, the equivalent formulation **PPR-1D** remains particularly useful for designing  
 270 (iterative) algorithms to solve the original **PPR** problem.

271 **3. Uniqueness and polynomial formulation.** This section studies the uniqueness prop-  
 272 erties of noiseless **PPR** under the set of assumptions ( $\mathcal{H}$ ) defined in **Section 2.2**. Thanks to  
 273 **Proposition 2.1**, we see that any solution of the problem **PPR** is a solution of the problem **BPR**,  
 274 and vice-versa. This formal equivalence permits to study uniqueness properties of the original  
 275 **PPR** through **BPR**. Following standard practice in Fourier phase retrieval problems, **Section**  
 276 **3.1** reformulates **BPR** using a polynomial formalism. **Theorem 3.2** shows that under the usual  
 277 oversampling condition  $M \geq 2N - 1$ , **BPR** is equivalent to a polynomial autocorrelation fac-  
 278 torization (**PAF**) problem. **Section 3.2** then provides general uniqueness results for **PAF** and  
 279 demonstrates that it can be solved using simple greatest common divisor computations.

280 **3.1. Bivariate phase retrieval as a polynomial factorization problem.** This section fol-  
 281 lows standard practice in Fourier phase retrieval problems [10, 12, 8, 11, 9] and adopts the  
 282 polynomial representation of Fourier transforms to study the uniqueness properties of the **BPR**  
 283 problem. Formally, let  $\mathbb{C}_{\leq N-1}[z]$  be the space of polynomials of degree at most  $N - 1$ . First,  
 284 let us define the polynomials  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  as generating polynomials of the components  
 285 of the bivariate signal  $\mathbf{x}[n] = (x_1[n], x_2[n])^\top \in \mathbb{C}^2$ ,  $n = 0, 1, \dots, N - 1$

$$286 \quad (3.1) \quad X_1(z) = \sum_{n=0}^{N-1} x_1[n]z^n, \quad X_2(z) = \sum_{n=0}^{N-1} x_2[n]z^n.$$

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

$$289 \quad (3.2) \quad \tilde{X}_1(z) = \sum_{n=0}^{N-1} x_1^*[N - n - 1]z^n, \quad \tilde{X}_2(z) = \sum_{n=0}^{N-1} x_2^*[N - n - 1]z^n.$$

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

$$291 \quad (3.3) \quad \mathbf{\Gamma}(z) = \begin{bmatrix} \Gamma_{11}(z) & \Gamma_{12}(z) \\ \Gamma_{21}(z) & \Gamma_{22}(z) \end{bmatrix} = \begin{bmatrix} X_1(z)\tilde{X}_1(z) & X_1(z)\tilde{X}_2(z) \\ X_2(z)\tilde{X}_1(z) & X_2(z)\tilde{X}_2(z) \end{bmatrix} = \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \begin{bmatrix} \tilde{X}_1(z) & \tilde{X}_2(z) \end{bmatrix},$$

292 where each element of the matrix is a polynomial  $\Gamma_{ij} \in \mathbb{C}_{\leq 2N-2}[z]$ . The coefficients of these  
 293 polynomials are simply the covariance functions (auto-covariances and cross-covariances) of  
 294 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  
 295 spectral matrices  $\{\mathfrak{F}[m]\}_{m=0}^{M-1}$  of **BPR** are linked to the evaluations of the polynomial  $\mathbf{\Gamma}(z)$ .

296 **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*

$$297 \quad (3.4) \quad \Gamma_{ij}(z) = \sum_{n=0}^{2N-2} \gamma_{ij}[n-N+1]z^n \text{ with } \gamma_{ij}[n] = \sum_{k \in \mathbb{Z}} x_i[k+n]x_j^*[k],$$

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

$$301 \quad (3.5) \quad \mathfrak{F}[m] = e^{j2\pi \frac{m(N-1)}{M}} \mathbf{\Gamma}(e^{-j2\pi \frac{m}{M}}).$$

302 Lemma 3.1 extends to the bivariate case the well-known correspondence between autocovari-  
 303 ance polynomials and Fourier amplitude in univariate Fourier phase retrieval (see for instance  
 304 [10, 12]). For completeness, we give a formal proof in Appendix A.

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

$$309 \quad (3.6) \quad M \geq 2N - 1.$$

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

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

$$315 \quad (\text{PAF}) \quad \text{find } X_1, X_2 \in \mathbb{C}_{\leq N-1}[z] \text{ given measurement polynomial } \mathbf{\Gamma}(z) \text{ defined as (3.3)} .$$

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

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

322 **3.2. General uniqueness result.** The PAF formulation is very helpful for establishing the  
 323 uniqueness conditions of BPR and, in turn, that of PPR under the nonrestrictive assumption  
 324 ( $\mathcal{H}$ ). Notably, PAF enables a complete characterization of uniqueness properties in terms  
 325 of algebraic properties of complex polynomials. To simplify the presentation in the follow-  
 326 ing, uniqueness properties refer jointly to PPR, BPR and PAF problems. In this section, we  
 327 reproduce several important results from [71] regarding the uniqueness of polynomial autocor-  
 328 relation factorizations problems. The notion of greatest common divisor (GCD) of complex  
 329 polynomials plays a pivotal role in establishing and interpreting these statements. In what  
 330 follows, we use the following definition of the GCD of polynomials taken from [71, 72].

331 **Definition 3.3.** Consider two polynomials  $A_1, A_2 \in \mathbb{C}_{\leq D}[z]$ , where at least one of them is  
 332 nonzero. The GCD of  $A_1(z)$  and  $A_2(z)$  is a polynomial  $Q(z) \in \mathbb{C}_{\leq K}[z]$ , with highest possible  
 333  $K$ , such that there exists two polynomials  $R_1, R_2 \in \mathbb{C}_{\leq D-K}[z]$  satisfying

$$334 \quad (3.7) \quad A_1(z) = Q(z)R_1(z) \text{ and } A_2(z) = Q(z)R_2(z),$$

335 and  $Q(z)$  has exactly  $\min(L_1, L_2)$  zero leading coefficients (where  $L_1$  and  $L_2$  is the number  
 336 of zero leading coefficients in  $A_1$  and  $A_2$  respectively).

337 **Remark 3.4.** The GCD exists and is unique up to a multiplication by a scalar in  $\mathbb{C} \setminus \{0\}$ .  
 338 Therefore by writing  $Q(z) = \gcd(A_1, A_2)$  we mean that  $Q(z)$  is a GCD up to this ambiguity.  
 339 All the usual properties of the GCD apply, despite the special treatment of zero leading coef-  
 340 ficients, see [71]. In particular, the polynomials  $R_1(z)$  and  $R_2(z)$ , called quotient polynomials,  
 341 satisfy  $\gcd(R_1, R_2) = 1$  and are said to be *co-prime*. Extension of the notion of GCD to  
 342 multiple polynomials is straightforward.

343 **Theorem 3.5 ([71]).** The following equivalences are true:

- 344 1. **PAF** admits a unique solution (up to trival ambiguities);
- 345 2.  $H(z) = \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22})$  has no roots outside the unit circle;
- 346 3.  $X_1(z)$  and  $X_2(z)$  have no common roots outside the unit circle and the leading coeffi-  
 347 cient of  $\gcd(X_1, X_2)$  is nonzero.

348 The proof of this result can be found in [71], where the generalization of **PAF** to the case of  $R$   
 349 polynomials is considered. Note that the uniqueness condition given in **Theorem 3.5** clarifies  
 350 previous statements made in the literature [57, 41]. In particular, in [57, Theorem 1] it was  
 351 claimed that coprimeness of the polynomials  $X_1(z)$  and  $X_2(z)$  was a necessary and sufficient  
 352 for uniqueness of the solution. **Theorem 3.5** shows that it was just a sufficient condition,  
 353 because unimodular roots do not affect uniqueness. This agrees with a similar behavior  
 354 observed for univariate one-dimensional Fourier phase retrieval [10], where unimodular roots  
 355 do not contribute to the number of non-trivial solutions. However, unlike univariate one-  
 356 dimensional Fourier phase retrieval, the bivariate case is almost everywhere unique, as shown  
 357 in the following corollary.

358 **Corollary 3.6 ([71]).** The **PAF** problem admits a unique solution for almost every poly-  
 359 nomials  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$ .

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

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

369 **Proposition 3.7 (GCD-based recovery).** Let  $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$  such that  $\gcd(X_1, X_2) = 1$ .  
 370 Then  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered as

$$371 \quad (3.8) \quad X_1(z) = c_1 \gcd(\Gamma_{11}, \Gamma_{12}) \text{ and } X_2(z) = c_2 \gcd(\Gamma_{21}, \Gamma_{22}).$$

372 where  $c_1, c_2 \in \mathbb{C}$  can be determined explicitly (up to one global phase) from measurement  
373 polynomials.

374 *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  
375  $\gcd(\tilde{X}_1, \tilde{X}_2) = 1$ . Therefore,  $\gcd(\Gamma_{11}, \Gamma_{12}) = \gcd(X_1 \tilde{X}_1, X_1 \tilde{X}_2) = c_1 X_1(z)$  since  $\tilde{X}_1(z)$  and  
376  $\tilde{X}_2(z)$  are co-prime. The same argument yields  $\gcd(\Gamma_{21}, \Gamma_{22}) = c_2 X_2(z)$ . Constants  $c_1$  and  
377  $c_2$  can be determined from the knowledge of correlation functions  $\gamma_{ij}[n]$ , or equivalently,  
378 measurement polynomials  $\Gamma_{ij}(z)$ ; see Section 4.3 for explicit expressions. ■

379 **Proposition 3.7** is a central result. It indicates that the **PAF** problem, and by extension, **BPR**  
380 and **PPR** can be solved using polynomial algebraic techniques. This distinctive feature arises  
381 as a direct consequence of accounting for polarization in Fourier phase retrieval problems.  
382 This original direction is further explored in Section 4, where we devise algebraic approaches  
383 to solve the noisy **PPR** problem using approximate GCD computations.

384 **4. Solving PPR with algebraic methods.** A central result of the previous section is  
385 **Proposition 3.7**, which states that polynomials  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered (up  
386 to trivial ambiguities) as GCDs of measurements polynomials  $\Gamma_{11}(z), \Gamma_{12}(z), \Gamma_{21}(z)$  and  $\Gamma_{22}(z)$ .  
387 The set of equivalencies summarized in Figure 1 further demonstrates that, in absence of noise,  
388 such *algebraic approaches* can be readily used to solve the initial **PPR** problem. In the context  
389 of noisy **PPR** measurements, this section shows how to leverage the notion of *approximate*  
390 GCD [72] for solving the polarimetric phase retrieval problem thanks to computational linear  
391 algebra methods. In the sequel, we assume that **PPR** measurements are corrupted by additive  
392 i.i.d. Gaussian noise such that for  $m = 0, 1, \dots, M-1$  and  $p = 0, 1, \dots, P-1$ ,

$$393 \quad (4.1) \quad y_{m,p} = |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^2 + n_{m,p}, \quad n_{m,p} \sim \mathcal{N}(0, \sigma^2),$$

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

$$395 \quad (4.2) \quad \text{SNR} = \frac{\sum_{m=0}^{M-1} \sum_{p=0}^{P-1} |\mathbf{a}_m^H \mathbf{X} \mathbf{b}_p|^4}{MP\sigma^2}.$$

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

405 **4.1. Reconstruction of measurement polynomials.** Recall that by Lemma 3.1 mea-  
406 surement polynomials  $\Gamma_{ij}(z)$  can be readily expressed in terms of auto-covariance functions  
407  $\{\gamma_{11}[n]\}, \{\gamma_{22}[n]\}$  and cross-covariance functions  $\{\gamma_{12}[n]\}, \{\gamma_{21}[n]\}$ . Thus, recovery of polyno-  
408 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  
409 Fourier transformation, one must retrieve the spectral matrix  $\mathfrak{F}[m]$  for  $m = 0, 1, \dots, M-1$   
410 from **PPR** measurements.

**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  $\hat{\Gamma}(z)$  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  $\hat{\mathbf{x}}_2$

---

411 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  
412  $\hat{\mathfrak{F}}[m]$  of  $\mathfrak{F}[m]$  is found for every  $m$  by minimizing the following quadratic-loss

$$413 \quad (4.3) \quad \hat{\mathfrak{F}}[m] = \arg \min_{\substack{\tilde{\mathfrak{F}}[m] = \tilde{\mathfrak{F}}[m]^H \\ \text{rank } \tilde{\mathfrak{F}}[m] = 1}} \sum_{p=0}^{P-1} \left( y_{m,p} - \text{Tr} \mathbf{b}_p^* \mathbf{b}_p^\top \tilde{\mathfrak{F}}[m] \right)^2,$$

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

416 To solve (4.3), we adopt a heuristic but simple strategy similar to practical polarimetric  
417 reconstruction techniques used in optics [62, 31]. First, we exploit the *Stokes parameters*  
418 representation of 2-by-2 Hermitian matrices, which read for an arbitrary Hermitian matrix  
419  $\mathbf{M} \in \mathbb{C}^{2 \times 2}$

$$420 \quad (4.4) \quad \mathbf{M} = \frac{1}{2} \begin{bmatrix} S_0 + S_1 & S_2 + jS_3 \\ S_2 - jS_3 & S_0 - S_1 \end{bmatrix} \quad S_0, S_1, S_2, S_3 \in \mathbb{R}.$$

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

$$425 \quad (4.5) \quad \text{Tr} \mathbf{b}_p^* \mathbf{b}_p^\top \tilde{\mathfrak{F}}[m] = \frac{1}{2} \left[ \mathcal{S} \left( \mathbf{b}_p^* \mathbf{b}_p^\top \right) \right]^\top \mathcal{S} \left( \tilde{\mathfrak{F}}[m] \right).$$

426 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-  
427 ing the  $P$  polarimetric measurements. Then one defines the polarization measurement matrix

428  $\mathbf{D} \in \mathbb{R}^{P \times 4}$  such that its  $p$ -th row reads  $\mathbf{D}_p = \frac{1}{2} [\mathcal{S}(\mathbf{b}_p^* \mathbf{b}_p^\top)]^\top$ . Note that the matrix  $\mathbf{D}$  does  
 429 not depend on Fourier frequency index  $m$ . This leads to rewriting problem (4.3) as

$$430 \quad (4.6) \quad \hat{\mathfrak{F}}[m] = \arg \min_{\substack{\tilde{\mathfrak{F}}[m] = \tilde{\mathfrak{F}}[m]^H \\ \text{rank } \tilde{\mathfrak{F}}[m] = 1}} \left\| \mathbf{y}_{m,:} - \mathbf{D} \mathcal{S}(\tilde{\mathfrak{F}}[m]) \right\|_2^2.$$

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

$$433 \quad (4.7) \quad \hat{\mathfrak{F}}[m] = \text{rank1} \left\{ \mathcal{S}^{-1} \left( \mathbf{D}^\dagger \mathbf{y}_{m,:} \right) \right\},$$

434 where  $\mathbf{D}^\dagger$  denotes the Moore-Penrose pseudo-inverse of  $\mathbf{D}$  and  $\mathcal{S}^{-1}$  is the inverse Stokes  
 435 mapping defined by (4.4). The operator  $\text{rank1}\{\mathbf{M}\}$  finds the best rank-one approximation  
 436 of a given matrix  $\mathbf{M}$  with respect to the Frobenius norm. For the present 2-by-2 Hermitian  
 437 matrix case, the solution is given by keeping the first singular vector of  $\mathbf{M}$ , that is  $\text{rank1}(\mathbf{M}) =$   
 438  $\sigma_0 \mathbf{u}_0 \mathbf{u}_0^H$ , where  $\sigma_0$  and  $\mathbf{u}_0$  are respectively the largest singular value and its corresponding  
 439 singular vector. Then, estimates  $\{\hat{\gamma}_{ij}[n]\}_{n=1-N}^{N-1}$  of covariance functions for  $i, j = 1, 2$  are  
 440 directly obtained by inverse discrete Fourier transformation of entries of the spectral matrices  
 441  $\{\tilde{\mathfrak{F}}[m]\}_{m=0}^{M-1}$  (possibly resampled to  $2N - 1$  points if  $M > 2N - 1$ ). Finally, Eq. (3.4) permits  
 442 to define estimated polynomials  $\hat{\Gamma}_{ij}(z)$  as polynomials in  $\mathbb{C}_{\leq 2N-2}[z]$  with vector of coefficients  
 443  $[\hat{\gamma}_{ij}[1 - N] \quad \hat{\gamma}_{ij}[2 - N] \quad \dots \quad \hat{\gamma}_{ij}[N - 1]]$ .

444 **4.2. Sylvester matrices and GCD.** Proposition 3.7 shows that, in the noiseless case, poly-  
 445 nomials  $X_1(z)$  and  $X_2(z)$  can be uniquely recovered as GCDs of the measurement polynomial  
 446 matrix  $\mathbf{\Gamma}(z)$ . In the noisy PPR measurement case, it further suggests that polynomials  $X_1(z)$   
 447 and  $X_2(z)$  can be estimated, or *approximately recovered* from the estimated matrix polynomial  
 448  $\hat{\mathbf{\Gamma}}(z)$  computed in Section 4.1. Importantly, the term *approximate* refers here to the practi-  
 449 cal infeasibility of computing exact GCDs due to numerical instabilities related to machine  
 450 precision, so that approximate GCD computations must be used instead. There are many  
 451 possible approaches for approximate GCD, see e.g., [52] for a recent review. In this work, we  
 452 follow [72] and carry approximate GCD computations by taking advantage of the kernel (or  
 453 null-space) properties of Sylvester matrices.

454 The following section reviews the relevant theory. Practical use of these results in the  
 455 context of PPR is given in Section 4.3.

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

$$458 \quad (4.8) \quad \mathcal{S}_D(A, B) = \left[ \begin{array}{cc|cc} a_0 & & b_0 & \\ \vdots & \ddots & \vdots & \ddots \\ a_L & & b_L & b_0 \\ & \ddots & & \vdots \\ & & a_L & b_L \end{array} \right] \in \mathbb{C}^{(2L-D+1) \times 2(L-D+1)}.$$

459 The Sylvester-like matrices are tightly linked with multiplication matrices of polynomials, see  
 460 Appendix B for more details. When  $D = 1$  (i.e., the matrix is square  $2L \times 2L$ ), the matrix is

461 the well-known Sylvester matrix. There are, however, two important extensions of the classic  
462 case:

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

468 For an overview of such matrices and the corresponding literature, we refer to [72] (note that  
469 unlike [72] we use the same notation for subresultant and extended Sylvester matrices). The  
470 following theorem is classic.

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

$$474 \quad K = 2L - \text{rank } \mathcal{S}_1(A, B)$$

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

476 Unfortunately, **Theorem 4.1** does not give an explicit way to compute  $\gcd(A, B)$ . In fact,  
477 explicit determination of the GCD requires the use of Sylvester matrices  $\mathcal{S}_D(A, B)$  in the  
478 general case  $D \neq 1$ . More precisely, **Proposition 4.2** and **Proposition 4.4** below show that the  
479 GCD can be retrieved from the left or right kernel of carefully constructed Sylvester matrices.  
480 In what follows, we assume that the GCD has degree  $K$  and note  $Q(z) = \gcd(A, B) \in \mathbb{C}_{\leq K}[z]$ .  
481 Moreover, we define

$$482 \quad F(z) = \frac{A(z)}{Q(z)}, \quad G(z) = \frac{B(z)}{Q(z)}$$

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

485 **Proposition 4.2 (Right kernel, see e.g. [72, Lemma 4.6]).** *The rank of the Sylvester sub-  
486 resultant matrix  $\mathcal{S}_K(A, B)$  is equal to  $2(L - K + 1) - 1$  (i.e., it has rank defect equal to 1).  
487 Moreover, for the (unique up to scalar factor) nonzero vector in the right kernel*

$$488 \quad (4.9) \quad \mathcal{S}_K(A, B) \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = 0;$$

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

$$490 \quad (4.10) \quad U(z) = -cG(z), \quad V(z) = cF(z),$$

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

492 **Remark 4.3.** In view of the connection between Sylvester and multiplication matrices (see  
493 **Appendix B**), **Proposition 4.2** implies that the polynomials  $U, V \in \mathbb{C}_{\leq L-K}[z]$  defined in (4.10)  
494 are the only ones (up to multiplication by a nonzero scalar) that satisfy  $A(z)U(z) + B(z)V(z) =$   
495  $0$  (see (4.9)).

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

499 **Proposition 4.4 (Left kernel).** *Let  $D \leq 1$  (i.e.,  $\mathcal{S}_D(A, B)$  is fat with  $2L - D + 1$  rows).  
 500 Then the rank of  $\mathcal{S}_D(A, B)$  is equal to*

$$501 \quad (4.11) \quad \text{rank } \mathcal{S}_D(A, B) = 2L - D + 1 - K;$$

502 *therefore the dimension of the left kernel (i.e., the rank defect) is equal to  $K$  (the degree of  
 503 the GCD). Moreover, let  $\mathbf{u}_{-1}, \dots, \mathbf{u}_{-K} \in \mathbb{C}^{2L-D+1}$  be a basis of the left kernel of  $\mathcal{S}_D(A, B)$   
 504 (for example, its last  $K$  left singular vectors) and define the following matrix*

$$505 \quad (4.12) \quad \mathbf{H} = [\mathcal{H}_{K+1}(\mathbf{u}_{-1}) \quad \cdots \quad \mathcal{H}_{K+1}(\mathbf{u}_{-K})] \in \mathbb{C}^{(K+1) \times K(2L-D+1-K)},$$

506 *where each block is a Hankel matrix, i.e.,  $\mathcal{H}_{K+1}(\mathbf{u})$  denotes the Hankel matrix with  $K+1$  rows  
 507 built from a vector  $\mathbf{u} \in \mathbb{C}^{2L-D+1}$  such that*

$$508 \quad \mathcal{H}_{K+1}(\mathbf{u}) = \begin{bmatrix} u[0] & u[1] & \cdots & u[2L-D-K] \\ u[1] & u[2] & \cdots & u[2L-D-K+1] \\ \vdots & \vdots & & \vdots \\ u[K] & u[K+1] & \cdots & u[2L-D] \end{bmatrix} \in \mathbb{C}^{(K+1) \times (2L-D+1-K)}.$$

509 *Then we have  $\text{rank } \mathbf{H} = K$  and the left kernel of  $\mathbf{H}$  is spanned by the the vector of coefficients  
 510  $\mathbf{q} \in \mathbb{C}^{K+1}$  of the GCD  $Q(z)$ .*

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

513 **4.3. Algebraic algorithms.** In this section, we propose two algorithms for estimating coef-  
 514 ficients of polynomials  $X_1(z)$  and  $X_2(z)$  from the estimated matrix polynomial  $\hat{\mathbf{\Gamma}}(z)$  computed  
 515 in [Section 4.1](#). Both algorithms rely on the use of the singular value decomposition (SVD) to  
 516 find the left or right kernels of Sylvester matrices constructed from  $\hat{\mathbf{\Gamma}}(z)$ . Thus the proposed  
 517 reconstruction methods may appear as suboptimal since the Sylvester structure is not taken  
 518 into account when computing the (low-rank) kernels. This limitation could be overcome with  
 519 structured low-rank approximations [\[48\]](#), to be specifically tailored for the [PPR](#) problem. Such  
 520 a study would fall outside the scope of the present work. Still, as demonstrated by the numer-  
 521 ical experiments presented in [Section 6](#), the SVD already provides excellent reconstruction  
 522 performance in many scenarios, while maintaining a reasonable computational burden.

523 **4.3.1. Right kernel Sylvester.** The first algorithm is based on the properties of the right  
 524 kernel of Sylvester matrices described in [Proposition 4.2](#). It uses the fact that  $X_1(z)$  and  
 525  $X_2(z)$  are (without noise) quotient polynomials of

$$526 \quad \Gamma_{11}(z) = X_1(z)\tilde{X}_1(z) \text{ and } \Gamma_{21}(z) = X_2(z)\tilde{X}_1(z).$$

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

**Algorithm 2:** Right kernel Sylvester

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

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  $\hat{\mathbf{x}}_2$

**Algorithm 3:** Left kernel Sylvester

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

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

    Construct the Hankel matrix  $\mathbf{H}$  as (4.12);

    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

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

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  $\hat{\mathbf{x}}_2$

531 The complete right kernel Sylvester approach is summarized in Algorithm 2. It exploits the  
 532 fact that  $\gcd(\Gamma_{11}, \Gamma_{21}) = X_1(z)$ , i.e., it is a polynomial of degree  $N - 1$ . Therefore according  
 533 to Proposition 4.2, the last nontrivial singular vector of the Sylvester matrix  $\mathcal{S}_{N-1}(\hat{\Gamma}_{11}, \hat{\Gamma}_{21})$   
 534 provides an estimate of the (one-dimensional) right kernel, which in turn gives, up to one  
 535 complex multiplicative constant, an estimation  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  of the vectors of coefficients defining  
 536 polynomials  $X_1(z)$  and  $X_2(z)$ . This constant is then computed (up to one unit-modulus factor  
 537 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  
 538 at the origin ( $n = 0$ ) of estimated auto-covariance functions  $\hat{\gamma}_{11}[n]$  and  $\hat{\gamma}_{22}[n]$ .

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

541 **4.3.2. Left kernel Sylvester.** The second algorithm exploits the properties of the left  
 542 kernel of extended (fat) Sylvester matrices (i.e.,  $\mathcal{S}_D$  for  $D \leq 1$ ) detailed in [Proposition 4.4](#).  
 543 For simplicity and to reduce the size of the involved matrices we set  $D = 1$  in what follows.  
 544 Nonetheless, the proposed approach can be adapted to any value of  $D \leq 1$  if needed.

545 [Algorithm 3](#) summarizes the complete procedure. In essence, it follows the theoretical  
 546 result of [Proposition 3.7](#). In particular, compared to the right kernel Sylvester approach,  
 547 estimated coefficients  $\hat{\mathbf{x}}_1$  and  $\hat{\mathbf{x}}_2$  are obtained by two separate GCD computations: the vector  
 548 of coefficients  $\hat{\mathbf{x}}_1$  is obtained by computing the GCD of estimated measurement polynomials  
 549  $\hat{\Gamma}_{11}(z)$  and  $\hat{\Gamma}_{12}(z)$ , whereas  $\hat{\mathbf{x}}_2$  is obtained by computing the GCD of  $\hat{\Gamma}_{21}(z)$  and  $\hat{\Gamma}_{22}(z)$ .  
 550 Importantly, the two GCDs are determined up to a multiplicative complex constant, say  $c_1$   
 551 and  $c_2$ , which can be determined jointly using [PPR](#) measurements.

552 The computation of each GCD requires three steps [[72](#)]: a first SVD to determine the  $N-1$   
 553 last left singular vectors of the Sylvester matrix  $\mathcal{S}_1$ ; the construction of a fat, horizontally  
 554 stacked Hankel matrix  $\mathbf{H}$  with  $N$  rows from these  $N-1$  singular vectors; a second SVD to  
 555 obtain the  $N$  coefficients of the GCD as the last left singular vector of  $\mathbf{H}$ . Once GCDs have  
 556 been obtained, determination of constants  $c_1$  and  $c_2$  (up to a common global phase factor) is  
 557 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  
 558  $\hat{\gamma}_{22}[0]$ ) and adjusting the phase factor  $\arg c_1 c_2^*$  thanks to the value at  $n = 0$  of the estimated  
 559 cross-covariance function  $\hat{\gamma}_{12}[n]$ .

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

565 **5. Solving PPR with iterative algorithms.** We now address the design of iterative al-  
 566 gorithms to solve the noisy [PPR](#) problem. [Section 5.1](#) and [Section 5.2](#) exploit the [PPR-1D](#)  
 567 representation of the original problem to provide a semidefinite programming (SDP) relaxation  
 568 and Wirtinger flow algorithm, respectively.

569 **5.1. SDP relaxation.** Semidefinite programming (SDP) approaches for phase retrieval  
 570 have been increasingly popular for over a decade [[15](#), [16](#), [75](#)]. In the classical 1D phase retrieval  
 571 case, SDP approaches exploit that even though measurements are quadratic in the unknown  
 572 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  
 573 representation [PPR-1D](#) enables to formulate a SDP relaxation of the original problem, by  
 574 observing that

$$575 \quad (5.1) \quad |\mathbf{c}_{m,p}^H \boldsymbol{\xi}|^2 = \text{Tr } \mathbf{c}_{m,p} \mathbf{c}_{m,p}^H \boldsymbol{\xi} \boldsymbol{\xi}^H = \text{Tr } \mathbf{C}_{m,p} \boldsymbol{\Xi},$$

576 i.e., noiseless measurements can be rewritten as a linear function of the lifted positive semi-  
 577 definite rank-one matrix  $\boldsymbol{\Xi} = \boldsymbol{\xi} \boldsymbol{\xi}^H \in \mathbb{C}^{2N \times 2N}$ . Following the classical PhaseLift methodology  
 578 [[16](#), [15](#)], the original nonconvex [PPR](#) problem can be relaxed into a SDP convex program as

$$579 \quad (5.2) \quad \begin{aligned} & \text{minimize} && \frac{1}{2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \text{Tr } \mathbf{C}_{m,p} \boldsymbol{\Xi})^2 + \lambda \|\boldsymbol{\Xi}\|_* \\ & \text{subject to} && \boldsymbol{\Xi} \succeq 0 \end{aligned}$$

**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  $\hat{\xi}$

580 where  $\lambda \geq 0$  is an hyperparameter that allows to control the trade-off between the likelihood  
 581 of observations and the nuclear norm regularization  $\|\cdot\|_*$ . Note that since  $\Xi$  is constrained  
 582 to be positive semidefinite, the nuclear norm regularization is equivalent to the trace-norm  
 583 regularization used in [15] since  $\|\Xi\|_* = \text{Tr} \Xi$  in this case. The SDP program (5.2) takes a  
 584 standard form: therefore it can be solved in many ways, including interior point methods [74],  
 585 first-order methods [51] or using disciplined convex programming solvers such as CVXPY<sup>1</sup>. For  
 586 completeness, we provide below an explicit algorithm to solve (5.2) using a proximal gradient  
 587 approach [7, Chapter 10]. It closely follows the approach described in [15, 32].

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

$$589 \quad (5.3) \quad f(\Xi) = \frac{1}{2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \text{Tr} \mathbf{C}_{m,p} \Xi)^2, \quad g(\Xi) = \lambda \|\Xi\|_* + \iota_{\succeq 0}(\Xi),$$

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

$$592 \quad (5.4) \quad \min_{\Xi \in \mathbb{C}^{2N \times 2N}} f(\Xi) + g(\Xi).$$

593 The convex optimization problem (5.4) can be efficiently solved by proximal gradient methods,  
 594 which take advantage of the splitting between  $f$  and  $g$  of the objective function. More precisely,

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

595 we use the fast proximal gradient method which consist, at iteration  $k$ :

$$596 \quad (5.5) \quad \Xi^{(k+1)} = \underset{t_k g}{\text{prox}} \left( \Psi^{(k)} - t_k \nabla f(\Psi^{(k)}) \right),$$

$$597 \quad (5.6) \quad \eta_{k+1} = \frac{1 + \sqrt{1 + 4\eta_k^2}}{2},$$

$$598 \quad (5.7) \quad \Psi^{(k+1)} = \Xi^{(k+1)} + \left( \frac{\eta_k - 1}{\eta_{k+1}} \right) \left( \Xi^{(k+1)} - \Xi^{(k)} \right),$$

599

600 where  $t_k$  is a step-size which is chosen such that the proximal gradient step (5.5) obey some  
 601 sufficient decrease condition; see e.g. [7, p. 271] for details. Our choice for the function  $g$  in  
 602 (5.4) enables a simple expression for the associated proximal operator (see [32]):

$$603 \quad (5.8) \quad \underset{\tau g}{\text{prox}}(\mathbf{X}) = \min_{\mathbf{Z} \succeq 0} \tau \lambda \|\mathbf{Z}\|_* + \|\mathbf{Z} - \mathbf{X}\|_2^2$$

$$= \mathbf{U} \text{shrink}(\Sigma, \tau \lambda) \mathbf{U}^H,$$

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

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

610 *Convergence.* Obviously, as (5.2) is a convex program, the precision towards the optimal  
 611 cost value can become arbitrarily good as one increases the number of iterations. In practice,  
 612 one needs to stop the algorithm when a prescribed tolerance  $\varepsilon$  is reached. To this aim we  
 613 implemented stopping criteria that carefully monitor a normalized residual, see [32] for details.  
 614 Moreover, it may happen that the estimated lifted matrix  $\hat{\Xi}$  generated by the sequence of  $\Xi^{(k)}$   
 615 is not rank one: in this case, one first computes the rank-one approximation of  $\hat{\Xi}$  (e.g. using  
 616 SVD) to obtain the estimated signal  $\hat{\xi}$ .

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

623 The full procedure is summarized in Algorithm 4.

624 **5.2. Wirtinger flow for PPR.** Exploiting further the 1D equivalent representation PPR-  
 625 1D of the PPR problem, another approach consists in minimizing directly the following non-  
 626 convex quadratic objective

$$627 \quad (5.9) \quad \min_{\xi \in \mathbb{C}^{2N}} F(\xi) = \frac{1}{2} \|\mathbf{y} - |\mathbf{C}\xi|^2\|_2^2,$$

**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  $\hat{\boldsymbol{\xi}}$

---

628 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  
 629 by  $\mathbf{c}_{m,p}^H$ , see Section 2.4. Provided that one can find a initial point  $\boldsymbol{\xi}^{(0)}$  close enough from  
 630 the global minimizer of (5.9), a simple strategy based on gradient descent can be used to  
 631 solve PPR. However, such an approach requires special care since the optimization variable  
 632  $\boldsymbol{\xi}$  is complex-valued. In fact, the objective function in (5.9) is real-valued, and thus it is not  
 633 differentiable with respect to complex analysis. Instead, one needs to resort to the so-called  
 634  $\mathbb{C}\mathbb{R}$  or *Wirtinger*-calculus [44] to provide a meaningful extension of gradient-descent-type  
 635 algorithms to the complex case. This is precisely the approach proposed in [17] to solve  
 636 standard phase retrieval, where the complex gradient descent is called *Wirtinger flow* (WF).

637 Leveraging the original WF approach, we propose below a complex-gradient descent al-  
 638 gorithm which solves the nonconvex problem (5.9). Compared to the original paper [17], we  
 639 incorporate optimal step size selection [43] together with a proposed acceleration scheme [77].  
 640 We further propose an efficient strategy for initialization based on the algebraic methods for  
 641 PPR described in Section 4. The superiority of these initializations over standard ones (e.g.  
 642 spectral initialization as proposed in [17]) will be demonstrated in Section 6.2.

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

$$645 \quad (5.10) \quad \beta_k = \frac{k+1}{k+3},$$

$$646 \quad (5.11) \quad \boldsymbol{\psi}^{(k)} = \boldsymbol{\xi}^{(k)} + \beta_k (\boldsymbol{\xi}^{(k)} - \boldsymbol{\xi}^{(k-1)}),$$

$$647 \quad (5.12) \quad \boldsymbol{\xi}^{(k+1)} = \boldsymbol{\psi}^{(k)} - \mu_k \nabla F(\boldsymbol{\psi}^{(k)}),$$

649 where  $\beta_k$  is a sequence of accelerated parameters and  $\mu_k$  is a carefully chosen stepsize, see  
 650 further below. Compared to the standard WF algorithm, PPR-WF takes advantage of the

651 acceleration procedure first proposed in [77] in the context of ptychographic phase retrieval  
 652 (but using a magnitude loss function instead of a square magnitude loss function as used here).  
 653 Note that the complex gradient of  $F$  can be computed explicitly as

$$654 \quad (5.13) \quad \nabla F(\boldsymbol{\psi}) = \mathbf{C}^H [(|\mathbf{C}\boldsymbol{\psi}|^2 - \mathbf{y}) \odot \mathbf{C}\boldsymbol{\psi}],$$

655 where the symbol  $\odot$  denotes entry-wise product between vectors.

656 *Optimal step-size selection.* We combine acceleration for WF with the optimal step-size  
 657 selection proposed in [43] for the standard WF algorithm. For completeness, we reproduce  
 658 here the main ingredients underpinning optimal step size selection in (5.12) and refer the  
 659 reader to [43] for further details. At iteration  $k$ , the optimal stepsize  $\mu_k$  is defined by line  
 660 search, i.e.,

$$661 \quad (5.14) \quad \mu_k = \arg \min_{\mu} F(\boldsymbol{\xi}^{(k+1)}) = F(\boldsymbol{\psi}^{(k)} - \mu \nabla F(\boldsymbol{\psi}^{(k)})).$$

662 The authors in [43] showed that the 1D optimization problem (5.14) boils down to finding  
 663 the roots of a univariate cubic polynomial with real coefficients, the latter being completely  
 664 determined by the knowledge of  $\boldsymbol{\psi}^{(k)}$ ,  $\nabla F(\boldsymbol{\psi}^{(k)})$  and  $\mathbf{y}$ , see [43, Eq. (17)]. Roots can be  
 665 determined in closed-form, and two cases can occur: (a) there is only one real root, and thus  
 666 it gives the optimal step-size  $\mu_k$ ; (b) there are three real roots, and in this case  $\mu_k$  is set to  
 667 the real root associated to the minimum objective value. Note that optimal selection for WF  
 668 is somewhat inexpensive, with computational cost dominated by the calculation of the cubic  
 669 polynomial coefficients scaling as  $\mathcal{O}(MP)$ .

670 *Initialization.* Since PPR-WF attempts at minimizing a nonconvex quadratic objective  
 671 (5.9), the choice of initial points  $\boldsymbol{\xi}^{(0)}$ ,  $\boldsymbol{\xi}^{(1)}$  is crucial to hope that PPR-WF will be able to  
 672 recover a global minimizer of the objective function. For simplicity, we set  $\boldsymbol{\xi}^{(1)} = \boldsymbol{\xi}^{(0)}$ , so that  
 673 we only discuss the selection of  $\boldsymbol{\xi}^{(0)}$ . Five different initialization strategies for PPR-WF are  
 674 considered:

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

$$677 \quad (5.15) \quad \mathbf{Y} = \frac{1}{MP} \sum_{r=0}^{MP-1} y_r \mathbf{c}_r \mathbf{c}_r^H$$

678 and to rescale it properly to set

$$679 \quad (5.16) \quad \boldsymbol{\xi}^{(0)} = \frac{\mathbf{v}}{\lambda}, \quad \lambda = \left( N \frac{\sum_{r=0}^{MP-1} y_r}{\sum_{r=0}^{MP-1} \|\mathbf{c}_r\|^2} \right)^{1/2}.$$

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

$$682 \quad (5.17) \quad \boldsymbol{\xi}^{(0)} = \mathbf{C}^\dagger \tilde{\mathbf{y}}, \quad \tilde{\mathbf{y}} = \mathbf{y} \odot \exp(j\boldsymbol{\phi}),$$

683 where  $\mathbf{C}^\dagger$  is the pseudo-inverse of  $\mathbf{C}$ .

- 684 • *right kernel Sylvester initialization*: we set  $\xi^{(0)}$  as the result of Algorithm 1 where ap-  
685 proximate GCDs computations are performed using the right kernel Sylvester method  
686 (Algorithm 2).
- 687 • *left kernel Sylvester initialization*: we set  $\xi^{(0)}$  as the result of Algorithm 1 where  
688 approximate GCDs computations are performed using the left kernel Sylvester method  
689 (Algorithm 3).
- 690 • *SDP initialization*: we set  $\xi^{(0)}$  as the output the SDP approach described in Section  
691 5.1 and summarized in Algorithm 4.

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

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

699 **6. Numerical experiments.** We provide in this section several numerical experiments that  
700 address how PPR can be solved in practice using both algebraic and algorithmic approaches  
701 described in Section 4 and Section 5, respectively. Importantly, we demonstrate that the  
702 use of the Wirtinger Flow algorithm together with a right-Sylvester initialization provides  
703 an excellent trade-off between mean squared error (MSE) and computational burden. This  
704 combination of algorithmic and algebraic reconstruction methods provides a scalable, asymp-  
705 totically MSE optimal, and parameter free inversion procedure for PPR.

706 Just like in standard phase retrieval, the global phase ambiguity in PPR requires to prop-  
707 erly realign any estimated signal  $\hat{\mathbf{X}}'$  with the ground truth  $\mathbf{X}$  in order to provide a meaningful  
708 squared reconstruction error value. We define the realigned estimated signal  $\hat{\mathbf{X}}$  as

$$709 \quad (6.1) \quad \hat{\mathbf{X}} = e^{j\Phi_0} \hat{\mathbf{X}}' \text{ with } \Phi_0 = \arg \min_{\phi \in [0, 2\pi)} \|e^{j\phi} \hat{\mathbf{X}}' - \mathbf{X}\|_F^2.$$

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

714 This section is organized as follows. Section 6.1 presents the reconstruction of a realistic  
715 bivariate pulse from noiseless PPR measurements using the different approaches presented in  
716 the paper. Section 6.2 then discusses the choice of initialization in PPR-WF. Section 6.3  
717 benchmarks the robustness to noise of the proposed reconstructions methods. Finally, Section  
718 6.4 provides a first study of the impact of the number of PPR measurements on reconstruction  
719 performances.

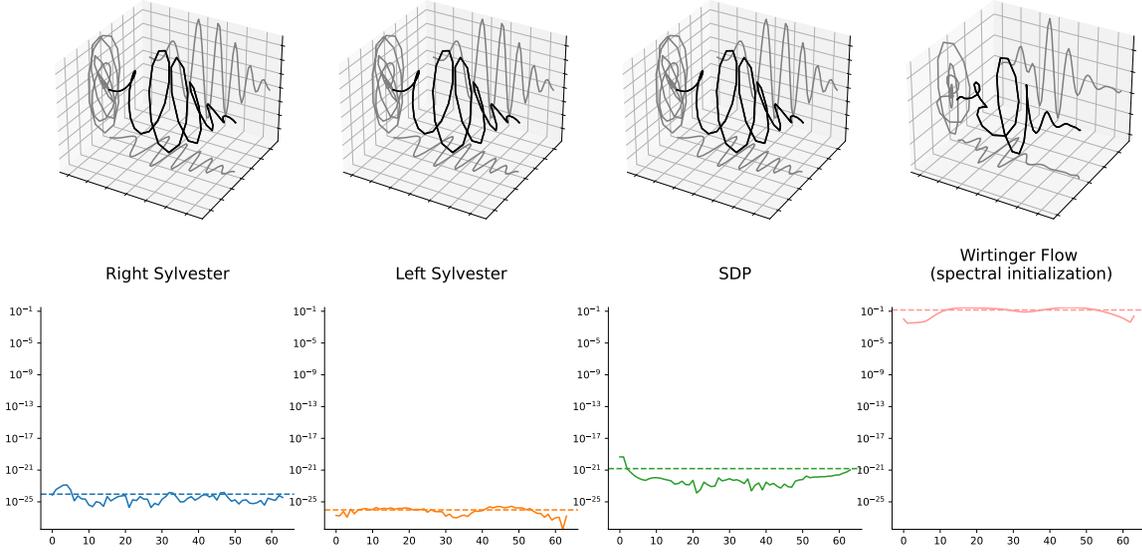
720 **6.1. Reconstruction of bivariate pulse.** As a first experiment, we consider the reconstruc-  
721 tion of a bivariate pulse from noiseless PPR measurements. The signal to be recovered defines  
722 a typical complex-valued bivariate analytic signal associated to the bivariate electromagnetic  
723 field to be estimated in ultra-short electromagnetic pulses experiments, see e.g. [64, 76]. It  
724 is defined for  $N = 64$  points and we consider the simple noise-free measurement scheme (2.5)

with  $M = 2N - 1$  and  $K = 4$ . The bivariate pulse exhibits slow variations of the instantaneous polarization state, ensuring uniqueness of the PPR solution. We investigate the capacity of the methods introduced in Section 4 and Section 5 to properly recover the bivariate signal of interest. Note that for Wirtinger Flow, we only consider spectral initialization—remaining strategies will be extensively benchmarked in Section 6.2 below.

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

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

Figure 4 depicts obtained reconstruction results for the three SNR scenarios, where we compare initialization methods in terms of cost function evolution  $F(\boldsymbol{\xi}^{(k)})$  and normed residual  $\|\boldsymbol{\xi}^{(k+1)} - \boldsymbol{\xi}^{(k)}\|_2 / \|\boldsymbol{\xi}^{(k)}\|_2$  decrease. Note that we imposed a identical number of 2500 iterations of PPR-WF for each approach to ensure fair comparisons. We also plot the empirical distribution of squared error values for each initialization for further comparison of the quality of the reconstructed signal (recall that squared error values are calculated after proper realignment of the estimated signal with the ground truth). For SNR = 10 dB (which is a very challenging scenario for PPR), there are no noticeable difference between initialization strategies: they

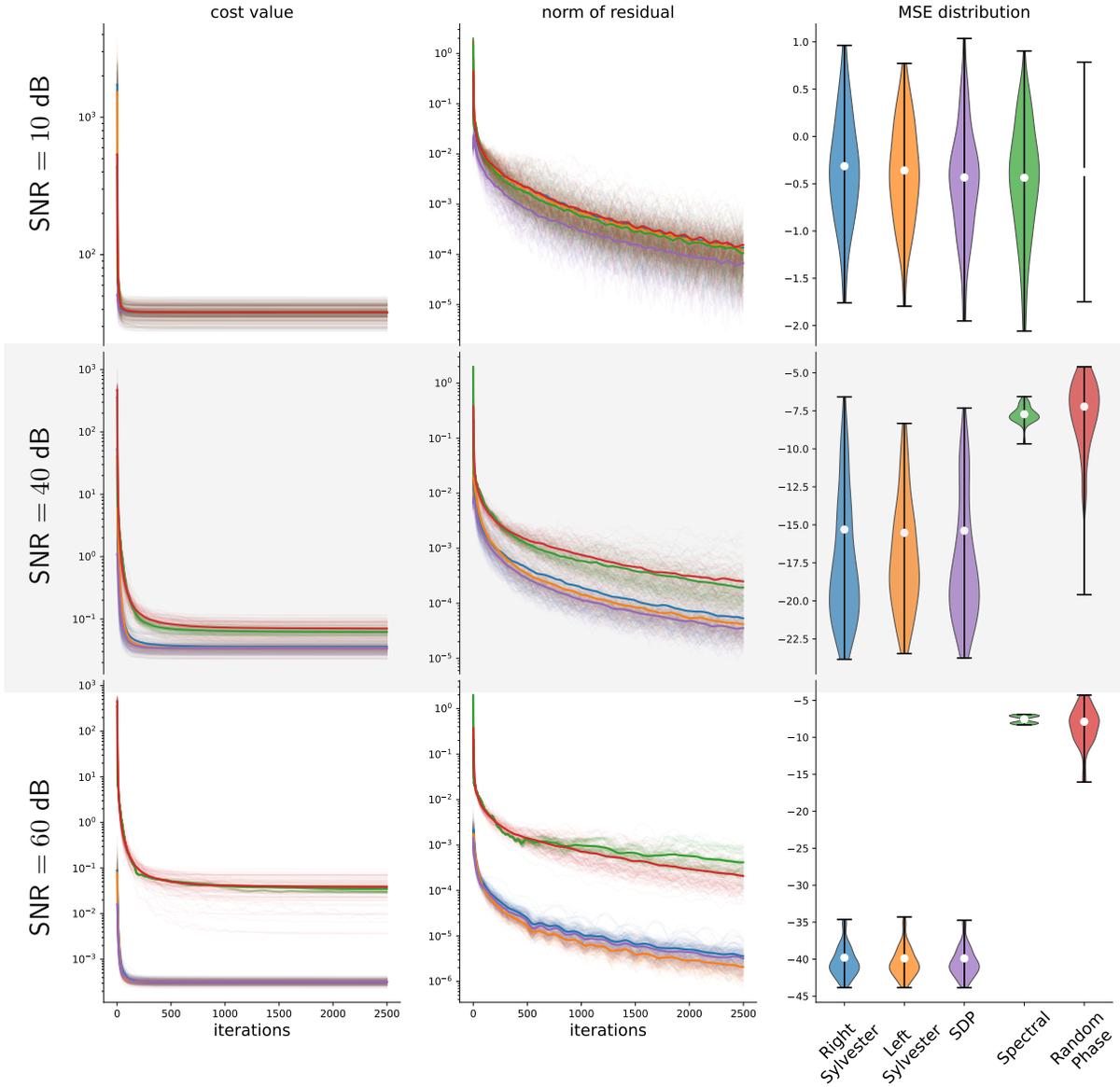


**Figure 3.** Reconstruction of a bivariate pulse ( $N = 64$ ) from noiseless PPR measurements ( $M = 2N - 1$ ,  $P = 4$ ) using the different methods described in this paper. The reconstructed signal trace and squared error per index  $n$  are shown for each approach. Dashed lines show the corresponding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

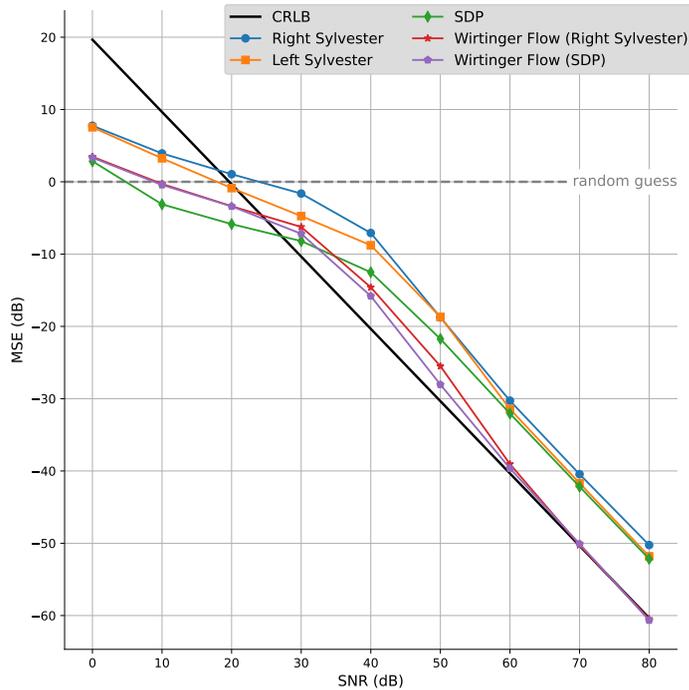
768 provide similar results in terms of cost value decrease, residual evolution and error distribu-  
 769 tion. For SNR = 40 dB, two categories of initializations with distinctive behaviors start to  
 770 appear: SDP and Sylvester-based approaches on one side, and on the other side, spectral and  
 771 random phase initializations. On average, SDP and Sylvester-based initial points provides  
 772 smaller optimal values, faster decrease of the residual and better reconstruction results in  
 773 terms of squared error. This behavior is accentuated for SNR = 60 dB, where spectral and  
 774 random phase initialization are unable to ensure convergence of PPR-WF to the global opti-  
 775 mum. This agrees with the observations made in Figure 3 in the noiseless case for spectral  
 776 initialization. Note that the poor performance of spectral initialization is not that surprising  
 777 since it was originally designed for random phase retrieval measurements [17], and not for the  
 778 case of deterministic Fourier measurements as in the present PPR setting.

779 These results demonstrate the importance of the choice of the initial point in PPR-WF  
 780 towards good convergence properties and recovery performance. Overall, spectral and ran-  
 781 dom phase strategies are systematically outperformed by SDP and left/right kernel Sylvester  
 782 initializations. The latter initialization strategies provide similar performances in terms of  
 783 the three figures-of-merit used here; however they exhibit very different computational costs  
 784 (see Figure 6 and next section). We shall see that the very limited cost of the right kernel  
 785 Sylvester approach (only requiring  $\mathcal{O}(N^3)$  operations, i.e., that of a single iteration of the  
 786 SDP Algorithm 4) makes it an excellent initialization for PPR-WF in many scenarios.

787 **6.3. Recovery performance with noisy measurements.** We now investigate the recovery  
 788 performances (in terms of MSE and computation time) of the different proposed algorithms



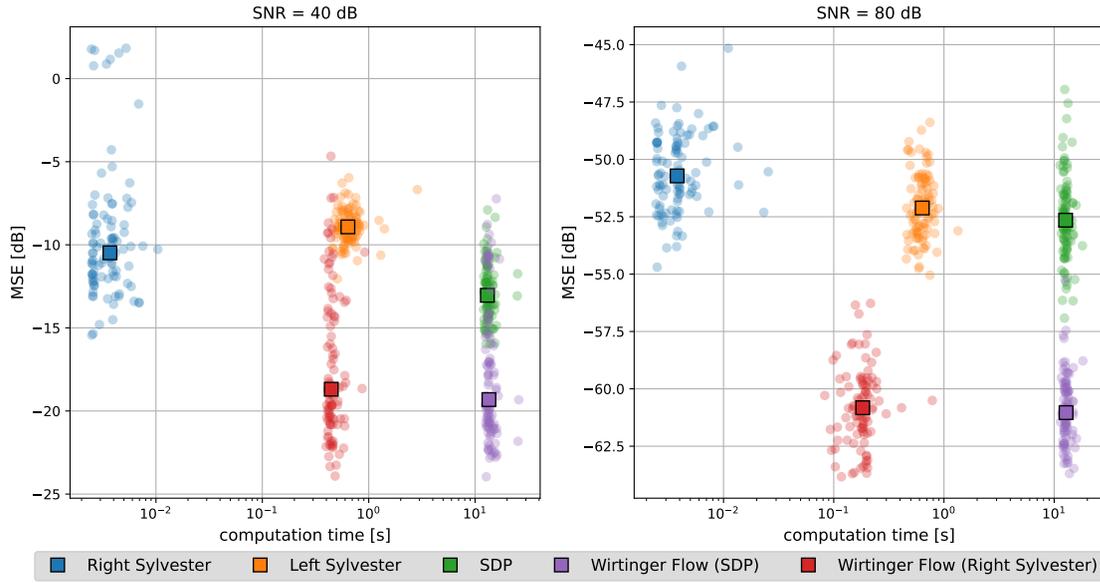
**Figure 4.** Comparison of initialization strategies for PPR-WF for the recovery of an arbitrary random bivariate signal of length  $N = 32$  with  $M = 2N - 1$  and  $P = 4$  noisy measurements. We benchmark spectral initialization, random phase initialization, left and right-kernel Sylvester initialization, and SDP 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 different PPR reconstruction methods proposed in this paper. For Wirtinger Flow (PPR-WF), two initialization strategies (SDP and right kernel Sylvester) are considered. The ground truth signal is a randomly generated bivariate signal with  $N = 32$ . The simple measurement scheme for  $M = 2N - 1$  and  $P = 4$  was used. The thick black line indicates the corresponding Cramér-Rao lower bound analytically derived in [Appendix C](#).

789 for PPR when dealing with noisy measurements. We consider an additive white Gaussian  
 790 noise model (4.1) for which the SNR is defined in (4.2). We generated a ground truth signal  
 791  $\mathbf{X} \in \mathbb{C}^{N \times 2}$  with i.i.d. Gaussian entries of length  $N = 32$  such that  $\|\mathbf{X}\|_F = 1$  and selected  
 792 the simple,  $M = 2N - 1, P = 4$  measurement scheme (2.5). For a given SNR value, the MSE  
 793 associated with each one of the proposed methods to solve PPR was obtained by averaging of  
 794 100 independent reconstructions. Following our analysis of initialization strategies in Section  
 795 6.2, we considered two initializations for PPR-WF (Algorithm 5): an algebraic one using right  
 796 kernel Sylvester, and a second one exploiting the output of the SDP approach (Algorithm 4).

797 Figure 5 displays the evolution of MSE for values of SNR ranging from 0 dB to 80 dB.  
 798 As expected, the MSE decreases as the SNR increases, independently from the considered  
 799 method. Overall, algorithmic methods (SDP and PPR-WF, independently from the initial-  
 800 ization method) outperform algebraic ones (left and right kernel Sylvester) in terms of MSE  
 801 values. More precisely, algebraic methods are not informative in the “low-SNR” regime (SNR  
 802  $\leq 30$  dB) as they provide (relative) MSE values above 0 dB, meaning that they do not pro-  
 803 vide a better reconstruction than a simple i.i.d. random guess scaled to the ground truth  
 804 norm. Furthermore we observe that SDP is more robust to noise than PPR-WF, even when  
 805 PPR-WF is initialized with the SDP output. This agrees with the fact that SDP methods  
 806 are known to be robust to noise in general. In terms of initialization strategies of PPR-WF,



**Figure 6.** Comparison of performances of the different reconstruction strategies in the MSE - computation time plane. Results corresponds to the same data plotted in Figure 5, for SNR = 40 dB (left panel) and SNR = 80 dB (right panel). For each reconstruction strategy, points represent one of the 100 independent trials; the boxed markers depicts the median error - median computation time of each method.

807 SDP provides a small advantage over the right kernel Sylvester initialization for SNR values  
 808 ranging from 30 to 60dB; in the low-SNR or high-SNR regime, both initialization strategies  
 809 are equivalent in terms of MSE performance. The high-SNR regime ( $\geq 60$  dB) highlights  
 810 several other distinctive behaviors. First, we observe that beyond SNR = 40 dB, PPR-WF  
 811 outperforms all other methods, including SDP, by a few dB up to about 10 dB of relative MSE  
 812 in the asymptotic regime. Second, left-kernel Sylvester and SDP provide similar performance,  
 813 only improving the right-kernel Sylvester method by a small margin.

814 For completeness, we also provide the Cramèr-Rao lower bound (CRLB) for the noisy  
 815 PPR measurement model (4.1) to characterize a lower bound on the MSE of any unbiased  
 816 estimator of the ground truth signal. An analytical derivation of the resulting CRLB is  
 817 given in Appendix C. Figure 5 displays the CRLB on top of MSE values obtained for each  
 818 reconstruction method. We observe that the CRLB is not informative below SNR  $\leq 20$  dB  
 819 as all methods provide smaller MSE values—it simply means that the CRLB is particularly  
 820 pessimistic in this regime. On the contrary, the CRLB provides a meaningful lower bound  
 821 in the high-SNR regime. Importantly, it demonstrates that PPR-WF is an asymptotically  
 822 optimal reconstruction method for PPR—independently from the initialization strategy—since  
 823 it attains the CRLB for SNR  $\geq 60$  dB.

824 Figure 6 depicts the performances of the different reconstruction strategies in the MSE -  
 825 computation time plane. Two SNR values of SNR = 40 dB (“mid” SNR regime) and SNR = 80  
 826 dB (high-SNR regime) are considered. For each reconstruction method, a single point corre-  
 827 spond to one of the 100 independent trials. Execution runtimes have been obtained on a 2021

828 Apple M1 Max MacBookPro with 32Gb RAM, using Python 3.11.1 and NumPy 1.24.1.  
 829 PPR-WF runtimes include the computation of the initialization (right kernel Sylvester or  
 830 SDP). In terms of computation burden, the right kernel Sylvester approach outperforms all  
 831 other methods by at least 2 orders of magnitude. SDP is the slowest method tested, with a  
 832 computational cost greater than 4 orders of magnitude for right kernel Sylvester and 2 orders  
 833 of magnitude for left kernel Sylvester and PPR-WF with the right kernel Sylvester initializa-  
 834 tion. In addition, when using the SDP initialization, the cost of PPR-WF is almost entirely  
 835 dominated by the cost of SDP, which prevents one from benefiting of the low-complexity of  
 836 the PPR-WF algorithm. Finally, Figure 6 shows that the best trade-off between MSE perfor-  
 837 mance and computation time is attained by combining PPR-WF with the inexpensive right  
 838 kernel Sylvester initialization.

839 **6.4. Influence of number of measurements.** One of the key advantages of the polarimet-  
 840 ric measurement model in PPR is that one can easily increase the number of measurements  
 841  $MP$  by performing more polarimetric projections, i.e., by increasing  $P$ . In fact, in practical  
 842 experiments it may be oftentimes easier to set up a new polarizer state  $\mathbf{b}_p$  than changing  
 843 the actual detector, which would be required if one desires to increase the number of Fourier  
 844 measurements  $M$ . Therefore, a natural question is the following: if one seeks to increase  
 845 the total number of measurements  $MP$ , is it better—in terms of MSE—to increase the num-  
 846 ber of Fourier measurements  $M$  or to increase the number of polarimetric projections  $P$ ?  
 847 Alternatively, can performance be simply improved by averaging over repeated independent  
 848 measurements? This is a vast topic related to the question of experimental design, which  
 849 requires a specific treatment which is outside the scope of the present paper. Nonetheless, we  
 850 provide in the sequel a first study of the influence of the number of measurements in PPR for  
 851 completeness.

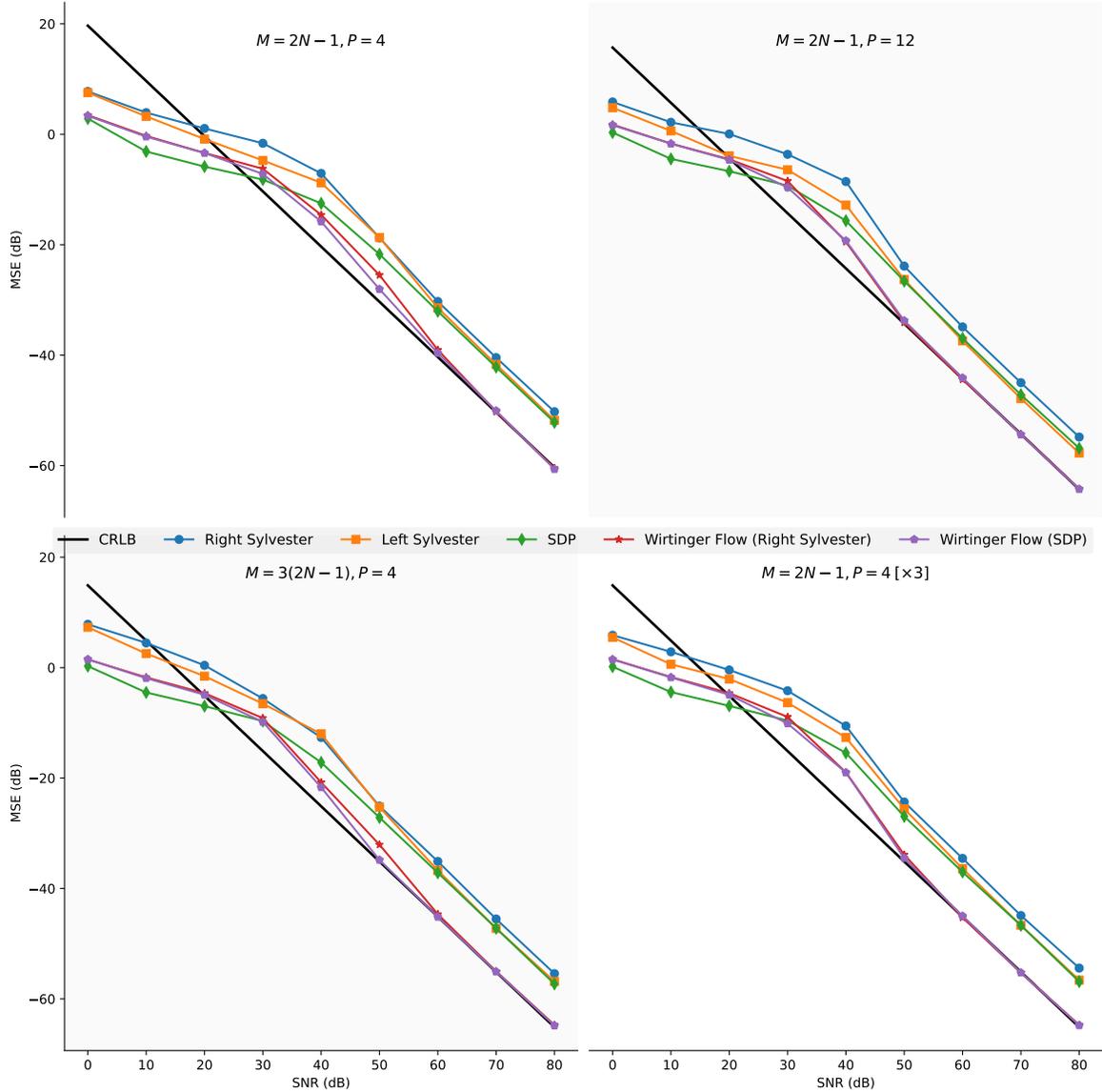
852 Following the MSE performance analysis in Section 6.3, we use the same randomly gener-  
 853 ated ground truth signal  $N = 32$  and investigate the performances for three cases leading to  
 854 the same total number of measurements  $MP$ :

- 855 •  $M = 2N - 1, P = 12$  case: we use the correspondence between the 2-sphere and unit  
 856 vectors of  $\mathbb{C}^2$  to take advantage of optimal spherical tessellations such as HEALPix  
 857 [35]. In physical terms, it can be interpreted as finding one of the many possible Jones  
 858 vector  $\mathbf{b}_p$  corresponding to the Stokes parameters defining the rank-one matrix  $\mathbf{b}_p \mathbf{b}_p^H$ .  
 859 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,  
 860 we define the projection vector  $\mathbf{b}_p$  as:

$$861 \quad (6.2) \quad \mathbf{b}_p = \frac{1}{\sqrt{2}\sqrt{1+s_p^z}} \begin{bmatrix} J s_p^x \\ s_p^y + (1+s_p^z)J \end{bmatrix} \quad \text{if } s_p^z \neq -1, \quad \mathbf{b}_p = \begin{bmatrix} J \\ 0 \end{bmatrix} \quad \text{if } s_p^z = -1.$$

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

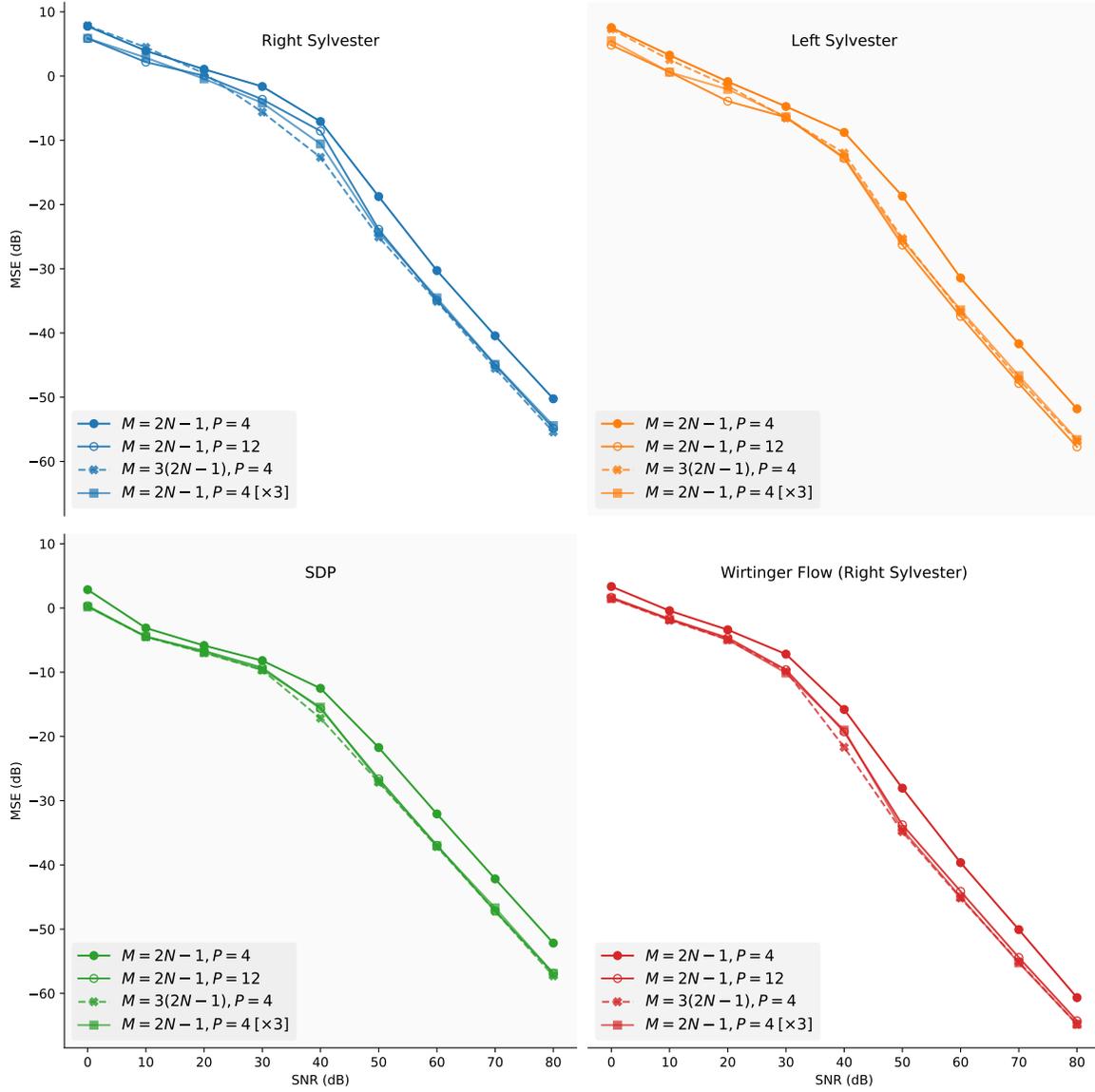
- 864 •  $M = 3(2N - 1), P = 4$  case: we keep the simple polarimetric measurement scheme  
 865 (2.4) and increase the number  $M$  of Fourier domain measurements.
- 866 •  $M = 2N - 1, P = 4 [\times 3]$  case: we keep the simple minimal measurement scheme for  
 867 PPR—which was used in Section 6.3—and repeat independently each measurement  
 868 sample three times, therefore leading to the same total number of measurements as



**Figure 7.** Comparison of the evolution of the MSE with respect to SNR for the four measurements schemes  $M = 2N - 1, P = 4$  (top left),  $M = 2N - 1, P = 12$  (top right),  $M = 3(2N - 1), P = 4$  (bottom left) and  $M = 2N - 1, P = 4 [\times 3]$  (i.e., the first scheme where each measurement sample is repeated independently three times). Experiments all use  $N = 32$  and follow the same protocol as described in Section 6.3.

869 the two strategies above.

870 Figure 7 depicts the MSE as a function of SNR for the three measurement setups described  
 871 above, where results from the experiment in Section 6.3 have been reproduced for better  
 872 comparison. As expected, increasing the total number of measurements  $MP$  improves overall



**Figure 8.** Side-by-side comparison of each proposed reconstruction method for the four measurements schemes  $M = 2N - 1, P = 4$ ,  $M = 2N - 1, P = 12$ ,  $M = 3(2N - 1), P = 4$  and  $M = 2N - 1, P = 4[\times 3]$ .

873 performance: this can be directly checked by remarking that the CRLB corresponding to  
 874  $M = 2N - 1, P = 12$ ,  $M = 3(2N - 1), P = 4$  and  $M = 2N - 1, P = 4 [\times 3]$  cases is lower  
 875 that of the  $M = 2N - 1, P = 4$  setup presented in Figure 5. Moreover, the different proposed  
 876 reconstructions methods for PPR behave similarly with one another as in our description made  
 877 in Section 6.3. In particular, we note that PPR-WF—for both initialization strategies—also  
 878 attains the CRLB in these three new setups, proving again that it establishes a convenient

879 approach to solve PPR.

880 Figure 8 provides a side-by-side comparison of these three measurement schemes for each  
 881 reconstruction method along with the minimal  $M = 2N - 1, P = 4$  scheme for comparisons.  
 882 For PPR-WF, only the right kernel Sylvester initialization is depicted to simplify the pre-  
 883 sentation, without affecting conclusions. First, on the PPR-WF panel, we can remark that  
 884 all three measurement schemes lead to similar asymptotic performance (i.e., similar CRLB),  
 885 with a slight disadvantage to the  $M = 2N - 1, P = 12$  case. Second, we note that for algo-  
 886 rithmic approaches (SDP and PPR-WF), the difference concentrates in the mid-SNR regime,  
 887 i.e., between 30 dB and 50 dB, where oversampling in the Fourier domain offers slight MSE  
 888 improvement over increasing the number of polarimetric projections or repeating measure-  
 889 ments. On the other hand, for algebraic approaches we observe two slight different behaviors  
 890 for  $\text{SNR} \geq 20$  dB: while oversampling in the Fourier domain gives the best MSE results for  
 891 the left kernel Sylvester approach, the strategy with increased polarimetric measurements pro-  
 892 vides the best asymptotic MSE results for the right kernel Sylvester approach. Nonetheless,  
 893 differences between measurement strategies remain tenuous, so that—in first approximation—  
 894 they can be considered all equivalent for this signal example. The choice of one strategy over  
 895 the other ones may be decided from experimental constraints or by exploiting some known  
 896 properties of the solution, e.g., narrow-band or average polarization properties.

897 **7. Conclusion.** This paper introduces a new model for Fourier phase retrieval called po-  
 898 larimetric phase retrieval (PPR), which takes advantage of polarization measurements in ap-  
 899 plications involving polarized light. The theoretical study of PPR relies on drawing careful  
 900 equivalences with two other problems, namely bivariate phase retrieval (BPR) and polynomial  
 901 autocorrelation factorization (PAF). In the noiseless case, these problems are found to be  
 902 equivalent under very general conditions, which are summarized in Figure 1. A crucial result  
 903 is Theorem 3.5: it shows that PAF admits a unique solution under very general conditions.  
 904 Therefore, the original PPR problem admits a unique solution for almost every signals. More-  
 905 over, the PAF representation enables the use of algebraic reconstruction strategies for PPR  
 906 based on GCD computations (Proposition 3.7). This original research direction is explored  
 907 in detail in Section 4, where we propose two fully algebraic (i.e., non-iterative) algorithms for  
 908 PPR relying on SVDs of Sylvester-like matrices. For completeness, Section 5 carefully adapts  
 909 classical phase retrieval algorithms (SDP relaxation and Wirtinger-Flow) to solve the PPR  
 910 problem. Section 6 provides extensive numerical experiments to benchmark the performances  
 911 of each approach. These results demonstrate that, if one is interested in a scalable, computa-  
 912 tionally efficient and robust to noise reconstruction strategy, then both algebraic and iterative  
 913 approaches should be combined. In practice, the best trade-off between reconstruction per-  
 914 formance and computational burden for PPR combines Wirtinger Flow (PPR-WF, Algorithm  
 915 5) with a carefully designed initialization based on right kernel Sylvester (Algorithm 1 with  
 916 GCDs computations performed using Algorithm 2).

917 We believe that PPR opens promising new avenues for the exploitation of light polarization  
 918 in Fourier phase retrieval problems. It enables the use of algebraic methods based on GCDs  
 919 computations to solve the Fourier phase retrieval problem. While this research direction is  
 920 particularly exciting, it also raises important challenges. For instance, an important issue to  
 921 be addressed lies in improving the performance of algebraic methods at low SNR, e.g. with

922 more robust estimation of the measurement polynomials or adding some prior information  
 923 about the signal to be recovered (e.g., smoothness). A second challenge lies in extending  
 924 the presented approaches to the case of polarized images, which is not straightforward at all  
 925 since properties of polynomials with multiple variables (and their GCDs) differ considerably  
 926 from their single variable counterpart. Combined with recent advances in non-convex gra-  
 927 dient descent optimization (e.g., using partial normalizations of the gradient [19]), algebraic  
 928 approaches pave the way for computationally efficient reconstruction strategies in polarimetric  
 929 Fourier phase retrieval problems.

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

931

932 *Proof of Lemma 3.1.* Recall that the discrete Fourier transform of  $\{\mathbf{x}[n]\}_{n=0}^{N-1}$  is denoted  
 933 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  
 934 can be related to polynomials  $X_1(z)$  and  $X_2(z)$  by comparing (2.2) with (3.1):

$$935 \quad \mathfrak{X}_1[m] = X_1\left(e^{-j2\pi\frac{m}{M}}\right), \quad \mathfrak{X}_2[m] = X_2\left(e^{-j2\pi\frac{m}{M}}\right),$$

936 for any  $m = 0, 1, \dots, M-1$ . Similarly, comparing (2.2) with (3.2), their conjugates can be  
 937 expressed through the conjugate reflection polynomials  $\tilde{X}_1(z)$  and  $\tilde{X}_2(z)$

$$938 \quad \mathfrak{X}_1^*[m] = X_1^*\left(e^{-j2\pi\frac{m}{M}}\right) = \sum_{n=0}^{N-1} x_1[n]^* e^{2\pi j\frac{nm}{M}} = e^{j2\pi\frac{m(N-1)}{M}} \tilde{X}_1\left(e^{-j2\pi\frac{m}{M}}\right),$$

$$939 \quad \mathfrak{X}_2^*[m] = X_2^*\left(e^{-j2\pi\frac{m}{M}}\right) = \sum_{n=0}^{N-1} x_2[n]^* e^{2\pi j\frac{nm}{M}} = e^{j2\pi\frac{m(N-1)}{M}} \tilde{X}_2\left(e^{-j2\pi\frac{m}{M}}\right).$$

940

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

$$943 \quad \mathfrak{F}[m] = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m]\mathfrak{X}_2[m]^* \\ \mathfrak{X}_2[m]\mathfrak{X}_1[m]^* & |\mathfrak{X}_2[m]|^2 \end{bmatrix}$$

$$944 \quad = e^{j2\pi\frac{m(N-1)}{M}} \begin{bmatrix} X_1\left(e^{-j2\pi\frac{m}{M}}\right) \tilde{X}_1\left(e^{-j2\pi\frac{m}{M}}\right) & X_1\left(e^{-j2\pi\frac{m}{M}}\right) \tilde{X}_2\left(e^{-j2\pi\frac{m}{M}}\right) \\ X_2\left(e^{-j2\pi\frac{m}{M}}\right) \tilde{X}_1\left(e^{-j2\pi\frac{m}{M}}\right) & X_2\left(e^{-j2\pi\frac{m}{M}}\right) \tilde{X}_2\left(e^{-j2\pi\frac{m}{M}}\right) \end{bmatrix}$$

$$945 \quad = e^{j2\pi\frac{m(N-1)}{M}} \mathbf{\Gamma}\left(e^{-j2\pi\frac{m}{M}}\right),$$

946

947 which completes the proof. ■

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

$$954 \quad \mathbf{a} = [a[0] \quad a[1] \quad \dots \quad a[N-1]]^\top \mapsto A(z) = a[0] + za[1] + \dots + z^{N-1}a[N-1].$$

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

$$958 \quad \left\{ \mathbf{\Gamma} \left( e^{-j2\pi \frac{m}{M}} \right) \right\}_{m=0}^{M-1}$$

959 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  
 960 the complex plane. If  $M \geq 2N - 1$  (the degree of the polynomials plus one), then it is known  
 961 that the coefficients of the polynomials can be uniquely recovered from the evaluations at  $M$   
 962 distinct points, and therefore the following map is an injection

$$963 \quad \mathbb{C}_{\leq 2N-2}^{2 \times 2} \rightarrow (\mathbb{C}^{2 \times 2})^M$$

$$964 \quad \mathbf{\Gamma}(z) \mapsto \{\mathfrak{F}[m]\}_{m=0}^{M-1}, \quad \blacksquare$$

966 which completes the proof.

### 967 **Appendix B. Multiplication matrices, Sylvester matrices and greatest common divisors.**

968 For  $\mathbf{q} \in \mathbb{C}^{D+1}$  and an integer  $T$ , we define the following  $(D + T + 1) \times (T + 1)$  matrix

$$969 \quad \mathbf{M}_T(\mathbf{q}) = \underbrace{\begin{bmatrix} q_0 & & & & \\ \vdots & \ddots & & & \\ q_K & & q_0 & & \\ & \ddots & & \vdots & \\ & & & & q_K \end{bmatrix}}_{T+1 \text{ columns}}.$$

970 This matrix is called the *multiplication matrix*, as it represents multiplication by the  
 971 polynomial  $Q \in \mathbb{C}_{\leq D}[z]$ , whose coefficients are in  $\mathbf{q}$ . Indeed, for any  $F \in \mathbb{C}_{\leq T}[z]$  and  
 972  $A(z) = F(z)Q(z)$ , the coefficient vectors  $\mathbf{f} \in \mathbb{C}^{T+1}$  and  $\mathbf{a} \in \mathbb{C}^{T+D+1}$  are linked with

$$973 \quad \mathbf{a} = \mathbf{M}_T(\mathbf{q})\mathbf{f},$$

974 see also [21] for more details on multiplication matrices.

975 *Remark B.1.* Note that for  $A, B \in \mathbb{C}_{\leq L}[z]$  the Sylvester matrix (4.8) is nothing but a  
 976 horizontal stack of multiplication matrices

$$977 \quad \mathcal{S}_D(A, B) = [\mathbf{M}_{L-D}(\mathbf{a}) \quad \mathbf{M}_{L-D}(\mathbf{b})].$$

978 Hence for the vectors  $\mathbf{u} \in \mathbb{C}^{L-D+1}$  and  $\mathbf{v} \in \mathbb{C}^{L-D+1}$ , the product

$$979 \quad \mathbf{w} = \mathcal{S}_D(A, B) \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix},$$

980 corresponds to the coefficients of the polynomial

$$981 \quad W(z) = A(z)U(z) + B(z)V(z).$$

982 *Proof of Proposition 4.4.* We first note that (4.11) is known (see, for example, [72, The-  
 983 orem 4.7]). Thus, we are left prove the second part, which is somewhat related to [72, Re-  
 984 mark 4.8]. We write  $A(z) = F(z)Q(z)$ ,  $B(z) = G(z)Q(z)$ , so that  $\gcd(A, B) = Q(z)$  and  
 985  $F, G \in \mathbb{C}_{\leq L-K}[z]$ . Consider the multiplication matrix

$$986 \quad \mathbf{M}_{2L-D-K}(\mathbf{q}) \in \mathbb{C}^{(2L-K+1) \times (2L-D-K+1)}.$$

987 Note that  $\mathbf{M}_{2L-D-K}(\mathbf{q})$  is full column rank (equal to  $2L-D-K+1$ ) and therefore, by (4.11)  
 988 (see also [72, Theorem 4.7]), we have

$$989 \quad (\text{B.1}) \quad \text{rank } \mathcal{S}_D(A, B) = \text{rank } \mathbf{M}_{2L-D-K}(\mathbf{q})$$

990 We will show that not only the ranks, but also the ranges of  $\mathcal{S}_D(A, B)$  and  $\mathbf{M}_{2L-D-K}(\mathbf{q})$   
 991 coincide. Note that, by Remark B.1, the range of  $\mathcal{S}_D(A, B)$  corresponds to all polynomials  
 992  $R \in \mathbb{C}_{\leq 2L-D}[z]$  that can be represented as

$$993 \quad (\text{B.2}) \quad R(z) = U(z)A(z) + V(z)B(z) = Q(z)(U(z)F(z) + V(z)G(z)),$$

994 and therefore any element in the range of  $\mathcal{S}_D(A, B)$  belongs to the range of  $\mathbf{M}_{2L-D-K}(\mathbf{q})$ .  
 995 Hence, by (B.1), the ranges of the two matrices coincide, which implies that their left kernels  
 996 coincide as well. In particular, the following equivalence holds true

$$997 \quad \mathbf{u}^\top \mathcal{S}_D(A, B) = 0 \iff \mathbf{u}^\top \mathbf{M}_{2L-D-K}(\mathbf{q}) = 0.$$

998 Finally, easy algebraic calculations (see also, for instance, [72, Eq. (33)]) show that

$$999 \quad \mathbf{u}^\top \mathbf{M}_{2L-D-K}(\mathbf{q}) = \mathbf{q}^\top \mathcal{H}_{K+1}(\mathbf{u}).$$

1000 Therefore,  $\mathbf{q}$  is in the left kernel of  $\mathbf{H}$  (4.12). Conversely, assume that there is a polynomial  
 1001  $\mathbf{q}'$  such that  $\mathbf{q}'^\top \mathbf{H} = 0$ . Then the image of  $\mathbf{M}_{2L-D-K}(\mathbf{q}')$  must be a subspace of the image of  
 1002  $\mathbf{M}_{2L-D-K}(\mathbf{q})$ , which is only possible if  $\mathbf{q}$  is a divisor of  $\mathbf{q}'$  (which implies that  $\mathbf{q}'$  is proportional  
 1003 to  $\mathbf{q}$ ). This shows that the left kernel of  $\mathbf{H}$  has dimension 1, hence  $\text{rank } \mathbf{H} = K+1-1 = K$ . ■

1004 **Appendix C. Cramèr-Rao bound for PPR.** Several authors have considered Cramèr-Rao  
 1005 bounds for the classical phase retrieval problem with additive white gaussian noise [1, 2, 55].  
 1006 These results directly apply to the additive Gaussian noise PPR model (4.1) since it can  
 1007 be equivalently rewritten as a particular one-dimensional noise model (the PPR-1D model  
 1008 introduced in Section 2.4). For completeness, we provide below an alternative derivation of  
 1009 the Cramèr-rao bound described in [55], where we use a full complex-domain approach instead  
 1010 of considering separate Cramèr-Rao bounds on amplitude and phase. Since measurement noise  
 1011  $n_{m,p}$  in (4.1) is i.i.d. Gaussian distributed with variance  $\sigma^2$ , the probability density function  
 1012 of the vector of observations  $\mathbf{y}$  is given by

$$1013 \quad (\text{C.1}) \quad p(\mathbf{y}|\boldsymbol{\xi}) = \prod_{m=0}^{M-1} \prod_{p=0}^{P-1} p(y_{m,p}|\boldsymbol{\xi})$$

$$1014 \quad (\text{C.2}) \quad = \prod_{m=0}^{M-1} \prod_{p=0}^{P-1} \frac{1}{\sqrt{2\pi\sigma}} \exp \left[ -\frac{(y_{m,p} - \boldsymbol{\xi}^\text{H} \mathbf{C}_{m,p} \boldsymbol{\xi})^2}{2\sigma^2} \right],$$

1015

1016 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  
1017 of observations reads

$$1018 \quad (\text{C.3}) \quad \log p(\mathbf{y}|\mathbf{x}_{\text{vec}}) = -\frac{MP}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} \left( y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi} \right)^2.$$

1019 Since one wants to estimate the complex parameter vector  $\boldsymbol{\xi}$ , it is necessary to use the complex  
1020 Fisher Information Matrix (FIM) [73, 46, 53], which reads

$$1021 \quad (\text{C.4}) \quad \mathcal{J}_{\boldsymbol{\xi}} = \begin{bmatrix} \mathcal{I}_{\boldsymbol{\xi}} & \mathcal{P}_{\boldsymbol{\xi}} \\ \mathcal{P}_{\boldsymbol{\xi}}^* & \mathcal{I}_{\boldsymbol{\xi}}^* \end{bmatrix} \in \mathbb{C}^{4N \times 4N},$$

1022 where entries are defined using Wirtinger derivatives [44] since  $\boldsymbol{\xi}$  is a complex vector:

$$1023 \quad (\text{C.5}) \quad \mathcal{I}_{\boldsymbol{\xi}} = \mathbf{E} \left[ \left( \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) \right) \left( \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) \right)^H \right],$$

$$1024 \quad (\text{C.6}) \quad \mathcal{P}_{\boldsymbol{\xi}} = \mathbf{E} \left[ \left( \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) \right) \left( \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) \right)^{\top} \right].$$

1026 Note that the FIM  $\mathcal{J}_{\boldsymbol{\xi}}$  defined in (C.4) is isomorphic to the real FIM which would have been  
1027 obtained by stacking the real and imaginary parts of  $\boldsymbol{\xi}$  in a single long vector [46]. This  
1028 explains why  $\mathcal{J}_{\boldsymbol{\xi}}$  has dimensions  $4N \times 4N$ . Using properties of Wirtinger derivatives, we get

$$1029 \quad (\text{C.7}) \quad \nabla_{\boldsymbol{\xi}^*} \log p(\mathbf{y}|\boldsymbol{\xi}) = -\frac{1}{\sigma^2} \sum_{m=0}^{M-1} \sum_{p=0}^{P-1} (y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi}) \mathbf{C}_{m,p} \boldsymbol{\xi}.$$

1030 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  
1031 independence, one gets

$$1032 \quad (\text{C.8}) \quad \mathcal{I}_{\boldsymbol{\xi}} = \frac{1}{\sigma^4} \mathbf{E} \left[ \left( \sum_{m,p} (y_{m,p} - \boldsymbol{\xi}^H \mathbf{C}_{m,p} \boldsymbol{\xi}) \mathbf{C}_{m,p} \boldsymbol{\xi} \right) \left( \sum_{m',p'} (y_{m',p'} - \boldsymbol{\xi}^H \mathbf{C}_{m',p'} \boldsymbol{\xi}) \boldsymbol{\xi}^H \mathbf{C}_{m',p'} \right) \right]$$

$$1033 \quad (\text{C.9}) \quad = \frac{1}{\sigma^4} \sum_{m,p,m',p'} \mathbf{E} [n_{m,p} n_{m',p'}] \mathbf{C}_{m,p} \boldsymbol{\xi} \boldsymbol{\xi}^H \mathbf{C}_{m',p'}$$

$$1034 \quad (\text{C.10}) \quad = \frac{1}{\sigma^2} \sum_{m,p} \mathbf{C}_{m,p} \boldsymbol{\xi} \boldsymbol{\xi}^H \mathbf{C}_{m,p}$$

$$1035 \quad (\text{C.11}) \quad = \frac{1}{\sigma^2} \sum_{m,p} |\mathbf{c}_{m,p}^H \boldsymbol{\xi}|^2 \mathbf{c}_{m,p} \mathbf{c}_{m,p}^H.$$

1036  
1037 Similar calculations leads to:

$$1038 \quad (\text{C.12}) \quad \mathcal{P}_{\boldsymbol{\xi}} = \frac{1}{\sigma^2} \sum_{ij} \mathbf{C}_{m,p} \boldsymbol{\xi}(\boldsymbol{\xi})^{\top} \mathbf{C}_{m,p}^{\top} = \frac{1}{\sigma^2} \sum_{m,p} \left( \mathbf{c}_{m,p}^H \boldsymbol{\xi} \right)^2 \mathbf{c}_{m,p} \mathbf{c}_{m,p}^{\top}.$$

1039 A key result [53] is that the inverse of the complex FIM (C.4) provides a lower bound on the  
1040 covariance and pseudo-covariance of any unbiased estimator  $\hat{\boldsymbol{\xi}}$  of the complex parameter  $\boldsymbol{\xi}$ :

$$1041 \quad (\text{C.13}) \quad \begin{bmatrix} \text{cov } \hat{\boldsymbol{\xi}} & \text{pcov } \hat{\boldsymbol{\xi}} \\ \text{pcov } \hat{\boldsymbol{\xi}}^* & \text{cov } \hat{\boldsymbol{\xi}}^* \end{bmatrix} \succeq \mathcal{J}_{\boldsymbol{\xi}}^{-1}.$$

1042 When the complex FIM is singular—as in phase retrieval [1, 2]—one can show its pseudo-  
 1043 inverse remains a valid lower bound for the MSE; following the discussion in [55], we still refer  
 1044 to the resultant bound as the CRB with little abuse. In particular, we obtain the following  
 1045 bound on the MSE on any unbiased PPR estimator  $\hat{\mathbf{X}}$  for the model (4.1):

$$1046 \quad (\text{C.14}) \quad \text{MSE}(\hat{\mathbf{X}}) = \mathbf{E}\|\hat{\mathbf{X}} - \mathbf{X}\|_F^2 = \mathbf{E}\|\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_2^2 = \text{Tr} \text{cov} \hat{\boldsymbol{\xi}} \geq \text{Tr} \left( \left[ \mathcal{J}_{\boldsymbol{\xi}}^\dagger \right]_{[:2N;:2N]} \right)$$

1047 where the subscript  $[:2N;:2N]$  denotes the restriction to the upper-left block of  $\mathcal{J}_{\boldsymbol{\xi}}^\dagger$ .

1048

## REFERENCES

- 1049 [1] R. BALAN, *Reconstruction of signals from magnitudes of redundant representations: The complex*  
 1050 *case*, Foundations of Computational Mathematics, 16 (2016), pp. 677–721, [https://doi.org/10.1007/](https://doi.org/10.1007/s10208-015-9261-0)  
 1051 [s10208-015-9261-0](https://doi.org/10.1007/s10208-015-9261-0).
- 1052 [2] A. S. BANDEIRA, J. CAHILL, D. G. MIXON, AND A. A. NELSON, *Saving phase: Injectivity and stability*  
 1053 *for phase retrieval*, Applied and Computational Harmonic Analysis, 37 (2014), pp. 106–125, <https://doi.org/10.1016/j.acha.2013.10.002>.
- 1054 [3] A. S. BANDEIRA, Y. CHEN, AND D. G. MIXON, *Phase retrieval from power spectra of masked signals*,  
 1055 *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)  
 1056 [imaiai/iau002](https://doi.org/10.1093/imaiai/iau002).
- 1057 [4] A. H. BARNETT, C. L. EPSTEIN, L. GREENGARD, AND J. MAGLAND, *Geometry of the Phase Retrieval*  
 1058 *Problem: Graveyard of Algorithms*, Cambridge University Press, 1 ed., Apr. 2022, [https://doi.org/](https://doi.org/10.1017/9781009003919)  
 1059 [10.1017/9781009003919](https://doi.org/10.1017/9781009003919).
- 1060 [5] A. BARONI, V. CHAMARD, AND P. FERRAND, *Extending Quantitative Phase Imaging to Polarization-*  
 1061 *Sensitive Materials*, Physical Review Applied, 13 (2020), p. 054028, [https://doi.org/10.1103/](https://doi.org/10.1103/PhysRevApplied.13.054028)  
 1062 [PhysRevApplied.13.054028](https://doi.org/10.1103/PhysRevApplied.13.054028).
- 1063 [6] A. BARONI AND P. FERRAND, *Reference-free quantitative microscopic imaging of coherent arbitrary vec-*  
 1064 *torial light beams*, Optics Express, 28 (2020), p. 35339, <https://doi.org/10.1364/OE.408665>.
- 1065 [7] A. BECK, *First-order methods in optimization*, SIAM, 2017, <https://doi.org/10.1137/1.9781611974997>.
- 1066 [8] R. BEINERT, *Non-negativity constraints in the one-dimensional discrete-time phase retrieval problem*,  
 1067 *Information and Inference: A Journal of the IMA*, 6 (2017), pp. 213–224, <https://doi.org/https://doi.org/10.1093/imaiai/iaw018>.
- 1068 [9] R. BEINERT, *One-dimensional phase retrieval with additional interference intensity measurements*, Results  
 1069 *in Mathematics*, 72 (2017), pp. 1–24, <https://doi.org/10.1007/s00025-016-0633-9>.
- 1070 [10] R. BEINERT AND G. PLONKA, *Ambiguities in one-dimensional discrete phase retrieval from fourier mag-*  
 1071 *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)  
 1072 [1007/s00041-015-9405-2](https://doi.org/10.1007/s00041-015-9405-2).
- 1073 [11] R. BEINERT AND G. PLONKA, *Enforcing uniqueness in one-dimensional phase retrieval by additional sig-*  
 1074 *nal information in time domain*, Applied and Computational Harmonic Analysis, 45 (2018), pp. 505–  
 1075 [525](https://doi.org/10.1016/j.acha.2016.12.002), <https://doi.org/10.1016/j.acha.2016.12.002>.
- 1076 [12] T. BENDORY, R. BEINERT, AND Y. C. ELДАР, *Fourier Phase Retrieval: Uniqueness and Algorithms*, in  
 1077 *Compressed Sensing and its Applications*, H. Boche, G. Caire, R. Calderbank, M. März, G. Kutyniok,  
 1078 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)  
 1079 [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.
- 1080 [13] T. BENDORY AND D. EDIDIN, *Algebraic Theory of Phase Retrieval*, Notices of the American Mathematical  
 1081 Society, 69 (2022), p. 1, <https://doi.org/10.1090/noti2540>.
- 1082 [14] T. BENDORY, Y. C. ELДАР, AND N. BOUMAL, *Non-convex phase retrieval from STFT measurements*,  
 1083 *IEEE Transactions on Information Theory*, 64 (2017), pp. 467–484, [https://doi.org/10.1109/TIT.](https://doi.org/10.1109/TIT.2017.2745623)  
 1084 [2017.2745623](https://doi.org/10.1109/TIT.2017.2745623).
- 1085 [15] E. J. CANDÈS, Y. C. ELДАР, T. STROHMER, AND V. VORONINSKI, *Phase retrieval via matrix completion*,  
 1086 *SIAM Journal on Imaging Sciences*, 6 (2013), pp. 199–225, <https://doi.org/10.1137/110848074>.

- 1089 [16] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase retrieval from coded diffraction patterns*, Applied  
1090 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)  
1091 [09.004](https://doi.org/10.1016/j.acha.2014.09.004).
- 1092 [17] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase Retrieval via Wirtinger Flow: Theory and*  
1093 *Algorithms*, IEEE Transactions on Information Theory, 61 (2015), pp. 1985–2007, [https://doi.org/](https://doi.org/10.1109/TIT.2015.2399924)  
1094 [10.1109/TIT.2015.2399924](https://doi.org/10.1109/TIT.2015.2399924).
- 1095 [18] H. N. CHAPMAN AND K. A. NUGENT, *Coherent lensless X-ray imaging*, Nature photonics, 4 (2010),  
1096 pp. 833–839, <https://doi.org/10.1038/nphoton.2010.240>.
- 1097 [19] Z. CHEN, Y. ZHOU, Y. LIANG, AND Z. LU, *Generalized-smooth nonconvex optimization is as efficient as*  
1098 *smooth nonconvex optimization*, 2023, <https://arxiv.org/pdf/2303.02854.pdf>.
- 1099 [20] R. A. CHIPMAN, G. YOUNG, AND W. S. T. LAM, *Polarized light and optical systems*, Optical sciences  
1100 and applications of light, Taylor & Francis, CRC Press, Boca Raton, 2018, [https://doi.org/10.1201/](https://doi.org/10.1201/9781351129121)  
1101 [9781351129121](https://doi.org/10.1201/9781351129121).
- 1102 [21] R. M. CORLESS, P. M. GIANNI, B. M. TRAGER, AND S. M. WATT, *The singular value decomposition for*  
1103 *polynomial systems*, in Proceedings of the 1995 international symposium on Symbolic and algebraic  
1104 computation - ISSAC '95, Montreal, Quebec, Canada, 1995, ACM Press, pp. 195–207, [https://doi.](https://doi.org/10.1145/220346.220371)  
1105 [org/10.1145/220346.220371](https://doi.org/10.1145/220346.220371).
- 1106 [22] X. DAI, S. XU, X. YANG, K. C. ZHOU, C. GLASS, P. C. KONDA, AND R. HORSTMAYER, *Quantitative*  
1107 *Jones matrix imaging using vectorial Fourier ptychography*, Biomedical Optics Express, 13 (2022),  
1108 p. 1457, <https://doi.org/10.1364/BOE.448804>.
- 1109 [23] V. ELSER, *Phase retrieval by iterated projections*, Journal of the Optical Society of America A, 20 (2003),  
1110 pp. 40–55, <https://doi.org/10.1364/JOSAA.20.000040>.
- 1111 [24] V. ELSER, T.-Y. LAN, AND T. BENDORY, *Benchmark problems for phase retrieval*, SIAM Journal on  
1112 Imaging Sciences, 11 (2018), pp. 2429–2455, <https://doi.org/10.1137/18M1170364>.
- 1113 [25] A. FANNJIANG AND T. STROHMER, *The numerics of phase retrieval*, Acta Numerica, 29 (2020), pp. 125–  
1114 228, <https://doi.org/10.1017/S0962492920000069>.
- 1115 [26] P. FERRAND, M. ALLAIN, AND V. CHAMARD, *Ptychography in anisotropic media*, Optics Letters, 40  
1116 (2015), p. 5144, <https://doi.org/10.1364/OL.40.005144>.
- 1117 [27] P. FERRAND, A. BARONI, M. ALLAIN, AND V. CHAMARD, *Quantitative imaging of anisotropic material*  
1118 *properties with vectorial ptychography*, Optics Letters, 43 (2018), p. 763, [https://doi.org/10.1364/OL.](https://doi.org/10.1364/OL.43.000763)  
1119 [43.000763](https://doi.org/10.1364/OL.43.000763).
- 1120 [28] J. R. FIENUP, *Reconstruction of an object from the modulus of its Fourier transform*, Optics Letters, 3  
1121 (1978), pp. 27–29, <https://doi.org/10.1364/OL.3.000027>.
- 1122 [29] J. R. FIENUP, J. C. MARRON, T. J. SCHULZ, AND J. H. SELDIN, *Hubble space telescope characterized by*  
1123 *using phase-retrieval algorithms*, Applied Optics, 32 (1993), pp. 1747–1767, [https://doi.org/10.1364/](https://doi.org/10.1364/AO.32.001747)  
1124 [AO.32.001747](https://doi.org/10.1364/AO.32.001747).
- 1125 [30] J. J. GIL AND R. OSSIKOVSKI, *Polarized light and the Mueller matrix approach*, CRC press, Boca Raton,  
1126 2 ed., 2022, <https://doi.org/doi.org/10.1201/9780367815578>.
- 1127 [31] D. H. GOLDSTEIN, *Polarized light*, CRC press, Boca Raton, 3 ed., 2017, <https://doi.org/10.1201/b10436>.
- 1128 [32] T. GOLDSTEIN, C. STUDER, AND R. BARANIUK, *A field guide to forward-backward splitting with a FASTA*  
1129 *implementation*, 2014, <https://doi.org/10.48550/arXiv.1411.3406>.
- 1130 [33] J. W. GOODMAN, *Introduction to Fourier optics*, Roberts and Company publishers, Englewood, CO,  
1131 3 ed., 2005.
- 1132 [34] J. P. GORDON AND H. KOGELNIK, *PMD fundamentals: Polarization mode dispersion in optical fibers*,  
1133 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)  
1134 [pnas.97.9.4541](https://doi.org/10.1073/pnas.97.9.4541).
- 1135 [35] K. M. GORSKI, E. HIVON, A. J. BANDAY, B. D. WANDELT, F. K. HANSEN, M. REINECKE, AND  
1136 M. BARTELMANN, *HEALPix: A framework for high-resolution discretization and fast analysis of data*  
1137 *distributed on the sphere*, The Astrophysical Journal, 622 (2005), p. 759, [https://doi.org/10.1086/](https://doi.org/10.1086/427976)  
1138 [427976](https://doi.org/10.1086/427976).
- 1139 [36] P. GROHS, S. KOPPENSTEINER, AND M. RATHMAIR, *Phase retrieval: uniqueness and stability*, SIAM  
1140 Review, 62 (2020), pp. 301–350, <https://doi.org/10.1137/19M1256865>.
- 1141 [37] S.-M. GUO, L.-H. YEH, J. FOLKESSON, I. E. IVANOV, A. P. KRISHNAN, M. G. KEEFE, E. HASHEMI,  
1142 D. SHIN, B. B. CHHUN, N. H. CHO, M. D. LEONETTI, M. H. HAN, T. J. NOWAKOWSKI, AND S. B.

- 1143 MEHTA, *Revealing architectural order with quantitative label-free imaging and deep learning*, eLife, 9  
1144 (2020), p. e55502, <https://doi.org/10.7554/eLife.55502>.
- 1145 [38] K. HUANG, Y. C. ELДАР, AND N. D. SIDIROPOULOS, *Phase Retrieval from 1D Fourier Mea-*  
1146 *surements: Convexity, Uniqueness, and Algorithms*, IEEE Transactions on Signal Processing,  
1147 64 (2016), pp. 6105–6117, <https://doi.org/10.1109/TSP.2016.2601291>, [http://ieeexplore.ieee.org/](http://ieeexplore.ieee.org/document/7547374/)  
1148 [document/7547374/](http://ieeexplore.ieee.org/document/7547374/) (accessed 2020-04-03).
- 1149 [39] K. JAGANATHAN, Y. ELДАР, AND B. HASSIBI, *Phase retrieval with masks using convex optimization*,  
1150 in 2015 IEEE International Symposium on Information Theory (ISIT), IEEE, 2015, pp. 1655–1659,  
1151 <https://doi.org/10.1109/ISIT.2015.7282737>.
- 1152 [40] K. JAGANATHAN, Y. C. ELДАР, AND B. HASSIBI, *STFT phase retrieval: Uniqueness guarantees and*  
1153 *recovery algorithms*, IEEE Journal of selected topics in signal processing, 10 (2016), pp. 770–781,  
1154 <https://doi.org/10.1109/JSTSP.2016.2549507>.
- 1155 [41] K. JAGANATHAN AND B. HASSIBI, *Reconstruction of Signals From Their Autocorrelation and Cross-*  
1156 *Correlation Vectors, With Applications to Phase Retrieval and Blind Channel Estimation*, IEEE  
1157 Transactions on Signal Processing, 67 (2019), pp. 2937–2946, [https://doi.org/10.1109/TSP.2019.](https://doi.org/10.1109/TSP.2019.2911254)  
1158 [2911254](https://doi.org/10.1109/TSP.2019.2911254).
- 1159 [42] K. JAGANATHAN, S. OYMAK, AND B. HASSIBI, *Sparse phase retrieval: Uniqueness guarantees and recovery*  
1160 *algorithms*, IEEE Transactions on Signal Processing, 65 (2017), pp. 2402–2410, [https://doi.org/10.](https://doi.org/10.1109/TSP.2017.2656844)  
1161 [1109/TSP.2017.2656844](https://doi.org/10.1109/TSP.2017.2656844).
- 1162 [43] X. JIANG, S. RAJAN, AND X. LIU, *Wirtinger Flow Method With Optimal Stepsize for Phase Re-*  
1163 *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)  
1164 [2611940](https://doi.org/10.1109/LSP.2016.2611940).
- 1165 [44] K. KREUTZ-DELGADO, *The Complex Gradient Operator and the CR-Calculus*, June 2009, [http://arxiv.](http://arxiv.org/abs/0906.4835)  
1166 [org/abs/0906.4835](http://arxiv.org/abs/0906.4835).
- 1167 [45] B. LESHEM, R. XU, Y. DALLAL, J. MIAO, B. NADLER, D. ORON, N. DUDOVICH, AND O. RAZ, *Direct*  
1168 *single-shot phase retrieval from the diffraction pattern of separated objects*, Nature Communications,  
1169 7 (2016), p. 10820, <https://doi.org/10.1038/ncomms10820>.
- 1170 [46] B. LOESCH AND B. YANG, *Cramér-rao bound for circular and noncircular complex independent component*  
1171 *analysis*, IEEE transactions on signal processing, 61 (2012), pp. 365–379, [https://doi.org/10.1109/](https://doi.org/10.1109/TSP.2012.2226166)  
1172 [TSP.2012.2226166](https://doi.org/10.1109/TSP.2012.2226166).
- 1173 [47] A. M. MAIDEN AND J. M. RODENBURG, *An improved ptychographical phase retrieval algorithm for diffrac-*  
1174 *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)  
1175 [05.012](https://doi.org/10.1016/j.ultramic.2009.05.012).
- 1176 [48] I. MARKOVSKY, *Structured low-rank approximation and its applications*, Automatica, 44 (2008), pp. 891–  
1177 909, <https://doi.org/10.1016/j.automatica.2007.09.011>.
- 1178 [49] J. MIAO, P. CHARALAMBOUS, J. KIRZ, AND D. SAYRE, *Extending the methodology of X-ray crystallography*  
1179 *to allow imaging of micrometre-sized non-crystalline specimens*, Nature, 400 (1999), pp. 342–344,  
1180 <https://doi.org/10.1038/22498>.
- 1181 [50] R. P. MILLANE, *Phase retrieval in crystallography and optics*, Journal of the Optical Society of America  
1182 A, 7 (1990), pp. 394–411, <https://doi.org/10.1364/JOSAA.7.000394>.
- 1183 [51] R. MONTEIRO ET AL., *First- and second-order methods for semidefinite programming*, Mathematical Pro-  
1184 gramming, 97 (2003), pp. 209–244, <https://doi.org/10.1007/s10107-003-0451-1>.
- 1185 [52] K. NAGASAKA, *Toward the best algorithm for approximate gcd of univariate polynomials*, Journal of  
1186 Symbolic Computation, 105 (2021), pp. 4–27.
- 1187 [53] E. OLLILA, V. KOIVUNEN, AND J. ERIKSSON, *On the Cramér-Rao bound for the constrained and un-*  
1188 *constrained complex parameters*, in 2008 5th IEEE Sensor Array and Multichannel Signal Processing  
1189 Workshop, IEEE, 2008, pp. 414–418, <https://doi.org/10.1109/SAM.2008.4606902>.
- 1190 [54] O. PEDATZUR, A. TRABATTONI, B. LESHEM, H. SHALMONI, M. CASTROVILLI, M. GALLI, M. LUCCHINI,  
1191 E. MÅNSSON, F. FRASSETTO, L. POLETTI, ET AL., *Double-blind holography of attosecond pulses*,  
1192 Nature Photonics, 13 (2019), pp. 91–95, <https://doi.org/10.1038/s41566-018-0308-z>.
- 1193 [55] C. QIAN, N. D. SIDIROPOULOS, K. HUANG, L. HUANG, AND H. C. SO, *Phase retrieval using feasible*  
1194 *point pursuit: Algorithms and Cramér–Rao bound*, IEEE Transactions on Signal Processing, 64 (2016),  
1195 pp. 5282–5296, <https://doi.org/10.1109/TSP.2016.2593688>.
- 1196 [56] J. RANIERI, A. CHEBIRA, Y. M. LU, AND M. VETTERLI, *Phase retrieval for sparse signals: Uniqueness*

- 1197 conditions, 2013, <https://arxiv.org/abs/1308.3058>.
- 1198 [57] O. RAZ, N. DUDOVICH, AND B. NADLER, *Vectorial Phase Retrieval of 1-D Signals*, IEEE Transactions  
1199 on Signal Processing, 61 (2013), pp. 1632–1643, <https://doi.org/10.1109/TSP.2013.2239994>.
- 1200 [58] O. RAZ, B. LESHEM, J. MIAO, B. NADLER, D. ORON, AND N. DUDOVICH, *Direct phase retrieval in double*  
1201 *blind Fourier holography*, Optics Express, 22 (2014), p. 24935, <https://doi.org/10.1364/OE.22.024935>.
- 1202 [59] O. RAZ, O. SCHWARTZ, D. AUSTIN, A. S. WYATT, A. SCHIAVI, O. SMIRNOVA, B. NADLER, I. A. WALM-  
1203 SLEY, D. ORON, AND N. DUDOVICH, *Vectorial Phase Retrieval for Linear Characterization of Attosec-*  
1204 *ond Pulses*, Physical Review Letters, 107 (2011), p. 133902, [https://doi.org/10.1103/PhysRevLett.](https://doi.org/10.1103/PhysRevLett.107.133902)  
1205 [107.133902](https://doi.org/10.1103/PhysRevLett.107.133902).
- 1206 [60] A. SABA, J. LIM, A. B. AYOUB, E. E. ANTOINE, AND D. PSALTIS, *Polarization-sensitive optical diffrac-*  
1207 *tion tomography*, Optica, 8 (2021), p. 402, <https://doi.org/10.1364/OPTICA.415343>.
- 1208 [61] D. SAYRE, *Some implications of a theorem due to Shannon*, Acta Crystallographica, 5 (1952), pp. 843–843,  
1209 <https://doi.org/10.1107/S0365110X52002276>.
- 1210 [62] B. SCHAEFER, E. COLLETT, R. SMYTH, D. BARRETT, AND B. FRAHER, *Measuring the Stokes polarization*  
1211 *parameters*, American Journal of Physics, 75 (2015), p. 6, <https://doi.org/10.1119/1.2386162>.
- 1212 [63] Y. SHECHTMAN, Y. C. ELДАР, O. COHEN, H. N. CHAPMAN, J. MIAO, AND M. SEGEV, *Phase Retrieval*  
1213 *with Application to Optical Imaging: A contemporary overview*, IEEE Signal Processing Magazine,  
1214 32 (2015), pp. 87–109, <https://doi.org/10.1109/MSP.2014.2352673>.
- 1215 [64] O. SMIRNOVA, S. PATCHKOVSKII, Y. MAIRESSE, N. DUDOVICH, D. VILLENEUVE, P. CORKUM, AND  
1216 M. Y. IVANOV, *Attosecond Circular Dichroism Spectroscopy of Polyatomic Molecules*, Physical Review  
1217 Letters, 102 (2009), p. 063601, <https://doi.org/10.1103/PhysRevLett.102.063601>.
- 1218 [65] Q. SONG, A. BARONI, R. SAWANT, P. NI, V. BRANDLI, S. CHENOT, S. VÉZIAN, B. DAMILANO,  
1219 P. DE MIERRY, S. KHADIR, ET AL., *Ptychography retrieval of fully polarized holograms from*  
1220 *geometric-phase metasurfaces*, Nature communications, 11 (2020), p. 2651, [https://doi.org/10.1038/](https://doi.org/10.1038/s41467-020-16437-9)  
1221 [s41467-020-16437-9](https://doi.org/10.1038/s41467-020-16437-9).
- 1222 [66] Q. SONG, X. LIU, C.-W. QIU, AND P. GENEVET, *Vectorial metasurface holography*, Applied Physics  
1223 Reviews, 9 (2022), p. 011311, <https://doi.org/10.1063/5.0078610>.
- 1224 [67] S. SONG, J. KIM, S. HUR, J. SONG, AND C. JOO, *Large-area, high-resolution birefringence imaging*  
1225 *with polarization-sensitive fourier ptychographic microscopy*, ACS Photonics, 8 (2021), pp. 158–165,  
1226 <https://doi.org/10.1021/acsp Photonics.0c01695>.
- 1227 [68] S. SONG, J. KIM, T. MOON, B. SEONG, W. KIM, C.-H. YOO, J.-K. CHOI, AND C. JOO, *Polarization-*  
1228 *sensitive intensity diffraction tomography*, Light: Science & Applications, 12 (2023), p. 124, <https://doi.org/10.1038/s41377-023-01151-0>.
- 1229 [69] R. TREBINO, *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*, Springer  
1230 Science & Business Media, New York, 2000, <https://doi.org/10.1007/978-1-4615-1181-6>.
- 1231 [70] J. S. TYO, D. L. GOLDSTEIN, D. B. CHENAULT, AND J. A. SHAW, *Review of passive imaging polarimetry*  
1232 *for remote sensing applications*, Applied optics, 45 (2006), pp. 5453–5469, [https://doi.org/10.1364/](https://doi.org/10.1364/AO.45.005453)  
1233 [AO.45.005453](https://doi.org/10.1364/AO.45.005453).
- 1234 [71] K. USEVICH, J. FLAMANT, M. CLAUSEL, AND D. BRIE, *On factorization of rank-one auto-correlation*  
1235 *matrix polynomials*. preprint, 2023, <https://hal.science/hal-04062934>.
- 1236 [72] K. USEVICH AND I. MARKOVSKY, *Variable projection methods for approximate (greatest) common divisor*  
1237 *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)  
1238 [2017.03.028](https://doi.org/10.1016/j.tcs.2017.03.028).
- 1239 [73] A. VAN DEN BOS, *A Cramér-Rao lower bound for complex parameters*, IEEE Transactions on Signal  
1240 Processing, 42 (1994), <https://doi.org/10.1109/78.324755>.
- 1241 [74] L. VANDENBERGHE AND S. BOYD, *Semidefinite programming*, SIAM review, 38 (1996), pp. 49–95, <https://doi.org/10.1137/1038003>.
- 1242 [75] I. WALDSPURGER, A. D’ASPREMONT, AND S. MALLAT, *Phase recovery, maxcut and complex semidef-*  
1243 *inite programming*, Mathematical Programming, 149 (2015), pp. 47–81, [https://doi.org/10.1007/](https://doi.org/10.1007/s10107-013-0738-9)  
1244 [s10107-013-0738-9](https://doi.org/10.1007/s10107-013-0738-9).
- 1245 [76] I. A. WALMSLEY AND C. DORRER, *Characterization of ultrashort electromagnetic pulses*, Advances in  
1246 Optics and Photonics, 1 (2009), pp. 308–437, <https://doi.org/10.1364/AOP.1.000308>.
- 1247 [77] R. XU, M. SOLTANOLKOTABI, J. P. HALDAR, W. UNGLAUB, J. ZUSMAN, A. F. J. LEVI, AND R. M.  
1248 LEAHY, *Accelerated Wirtinger Flow: A fast algorithm for ptychography*, June 2018, <http://arxiv.org/>  
1249 <http://arxiv.org/>  
1250

1251

[abs/1806.05546](#).