Polarimetric phase retrieval : uniqueness and algorithms

Julien Flamant, Konstantin Usevich, Marianne Clausel, David Brie

Polarimetric phase retrieval: uniqueness and algorithms*

Julien Flamant[a,*], Konstantin Usevich[a], Marianne Clausel[b], David Brie[a]

[a] *CNRS, Université de Lorraine, CRAN, F-54000 Nancy France*
[b] *CNRS, Université de Lorraine, Institut Elie Cartan de Lorraine, F-54000 Nancy France*

Abstract

This work introduces a novel Fourier phase retrieval model, called *polarimetric phase retrieval* that enables a systematic use of polarization information in Fourier phase retrieval problems. We provide a complete characterization of uniqueness properties of this new model by unraveling equivalencies with a peculiar polynomial factorization problem. We introduce two different but complementary categories of reconstruction methods. The first one is algebraic and relies on the use of approximate greatest common divisor computations using Sylvester matrices. The second one carefully adapts existing algorithms for Fourier phase retrieval, namely semidefinite positive relaxation and Wirtinger-Flow, to solve the polarimetric phase retrieval problem. Finally, a set of numerical experiments permits a detailed assessment of the numerical behavior and relative performances of each proposed reconstruction strategy. We further highlight a reconstruction strategy that combines both approaches for scalable, computationally efficient and asymptotically MSE optimal performance.

Keywords: Fourier phase retrieval, polarization, approximate greatest common divisor, semidefinite positive relaxation, Wirtinger Flow

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 [1]. The Fourier phase retrieval problem has been – and continues to be – of tremendous importance for many applications areas involving optics, such as crystallography [2, 3, 4], astronomy [5, 6], coherent diffraction imaging (also known as lensless imaging) [7, 8, 9], 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). This means that in such imaging applications, only intensity measurements can be performed, and that the phase should be recovered numerically afterwards. Moreover, during the last decade, the phase retrieval problem has gained a lot of interest in the signal processing and applied mathematics community [10, 11, 12, 13]. However, it is important to note that most of works in this community focus on generalized phase retrieval problems, where Fourier measurements are replaced or combined with random projections. While this allows the derivation of several important results using probabilistic considerations, e.g. uniqueness or stability guarantees, these results are not directly applicable to the original (deterministic) Fourier phase retrieval problem. Indeed, it is well known that one-dimensional univariate Fourier phase retrieval does not admit a unique solution in

*This work was funded by CNRS and GdR ISIS under the 2019-2021 OPENING exploratory research project grant.
*Corresponding author
Email addresses: julien.flamant@cnrs.fr (Julien Flamant), konstantin.usevich@univ-lorraine.fr (Konstantin Usevich), marianne.clausel@univ-lorraine.fr (Marianne Clausel), david.brie@univ-lorraine.fr (David Brie)

general [14]. We refer the reader to [15] for a recent review of proposed (deterministic) strategies to recover uniqueness of Fourier phase retrieval as well as associated algorithms.

Just like color (wavelength), *polarization* is a fundamental property of light. It encodes 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 [16]. As polarized light propagates in media, its polarization can change, thus revealing key properties, such as medium anisotropy or architectural order that are inaccessible to conventional, non-polarized light [17]. As a result, polarized light imaging has found many applications such as in material characterization [18], remote sensing [19] or bio-imaging [20]. 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 [21, 22] pioneered the use of polarization in Fourier phase retrieval in the context of ultrashort (e.g. attosecond, $\sim 10^{-18}s$) laser pulse characterization. Figure 1 depicts a simplified experimental setup of such an experiment. The goal here is to recover the polarized pulse after the medium, given *polarimetric* projections of its Fourier transform – recorded by the spectrometer. Comparing the reconstructed output pulse with the input laser pulse, one is able to recover key anisotropic properties of the studied medium. More recently, authors have developed *vectorial ptychography* [23, 24], a promising lensless imaging technique that simultaneously uses polarization and tilted measurements. This allows quantitative imaging of complex anisotropic media, such as biominerals [25, 26].

This work introduces a novel Fourier phase retrieval model, called *polarimetric phase retrieval* that enables a systematic use of polarization information in Fourier phase retrieval problems. The rationale is the following: we consider the *polarimetric phase retrieval* problem as the problem of recovering the 1D bivariate signal representing polarized light from scalar quadratic Fourier magnitude measurements. Importantly, this new model leverages physical acquisition schemes relevant to polarization measurement. Notably, it encompasses as special cases previous models proposed in the literature [27, 28] originally developed to account for vectorial-like diversity in phase retrieval. Our contributions can be stated as follows. We first provide a complete characterization of uniqueness properties of the polarimetric phase retrieval model, by unraveling equivalencies with a peculiar polynomial factorization problem. Notably, we show that unlike standard 1D Fourier phase retrieval, almost all bivariate signals can be uniquely recovered from polarimetric Fourier measurements. We also introduce two different but complementary categories of reconstruction methods. The first one is algebraic and relies on the use of approximate greatest common divisor computations using Sylvester matrices. The second one carefully adapts existing algorithms for Fourier phase retrieval, namely semidefinite positive relaxation and Wirtinger-Flow, to solve the polarimetric phase retrieval problem. Finally, a set of numerical experiments permits a thorough assessment of the numerical behavior and relative performances of each proposed reconstruction strategy. We further highlight a reconstruction strategy that combines both approaches for scalable, computationally efficient and asymptotically MSE optimal performance.

This paper is organized as follows. Section 2 introduces the polarimetric phase retrieval model and discusses its equivalent formulations as well as trivial ambiguities. Section 3 provides a complete study of the uniqueness properties of the polarimetric phase retrieval model, by leveraging a polynomial factorization representation of the problem. Section 4 exploits uniqueness results to propose two algebraic reconstruction methods based on approximate greatest common divisor computations. Section 5 takes a complementary path, by developing two iterative algorithms to solve the polarimetric phase retrieval problem. Section 6 details several numerical experiments to illustrate and assess the practical performances of the proposed reconstruction algorithms. Section 7 collects concluding remarks and Appendices gather technical details and supplementary results.

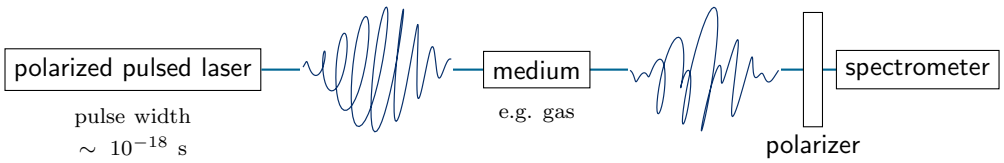

Figure 1: Simplified experimental setup for measuring ultrashort (e.g. a few attoseconds 10^{-18} s) electromagnetic polarized pulses. Inspired by [21, 22].

2. Polarimetric phase retrieval model

2.1. General formulation

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

$$\mathbf{X} \triangleq \begin{bmatrix} x_1[0] & x_2[0] \\ x_1[1] & x_2[1] \\ \vdots & \vdots \\ x_1[N-1] & x_2[N-1] \end{bmatrix}. \tag{1}$$

We define the *polarimetric phase retrieval* (PPR) problem as the problem of recovering the bivariate signal $\{\mathbf{x}[n]\}_{n=0,1,\dots,N-1}$ given scalar quadratic Fourier measurements. Formally,

$$\begin{aligned} &\text{find } \mathbf{X} \in \mathbb{C}^{N\times 2} \text{ given measurements } y_{m,p} = \left|\mathbf{a}_m^{\mathsf{H}} \mathbf{X} \mathbf{b}_p\right|^2 \\ &\qquad\qquad m = 0, 1, \dots M-1, \quad p = 0, 1, \dots P-1 \end{aligned} \tag{PPR}$$

where $\mathbf{a}_m \in \mathbb{C}^N$ is the discrete Fourier vector corresponding to frequency $\frac{2\pi m}{M}$, *i.e.* $a_m[n] = \exp(\jmath \frac{2\pi m}{M} n)$ for $n = 0, 1, \dots N-1$, and the $\mathbf{b}_p \in \mathbb{C}^2$ denote P arbitrary projection vectors which are supposed to be normalized such that $\|\mathbf{b}_p\|_2^2 = 1$. PPR measurements encode the physics of the acquisition in coherent diffraction imaging, where only intensity measurements can be performed: Fourier vectors $\mathbf{a}_m$ model Fraunhofer diffraction, whereas the $\mathbf{b}_p$'s represent the different polarizers (or polarization analysers) required to measure polarization. This agrees completely with the experimental setup described in Figure 1 in the context of ultra-short polarized electromagnetic pulse characterization. Note also that while we focus here on *physically realizable* measurement schemes, there is no obstacle from a mathematical viewpoint to extend the measurement scheme of PPR can be extended to arbitrary $\mathbf{a}_m$ sensing vectors. This includes, for instance, random gaussian vectors as in *generalized phase retrieval* problems, see *e.g.* [11, 10, 29] to cite only a few.

2.2. Relation with Fourier matrix measurements

A closely related problem to PPR is the *bivariate phase retrieval* (BPR) problem. Let us introduce the discrete Fourier transform of the bivariate signal $\{\mathbf{x}[n]\}_{n=0,1,\dots,N-1}$ as

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

for $m = 0, 1, \dots M-1$. Then let $\mathfrak{F}[m]$ denote the rank-one 2-by-2 complex spectral matrix at frequency indexed by m,

$$\mathfrak{F}[m] \triangleq \mathfrak{X}[m]\mathfrak{X}[m]^{\mathsf{H}} = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m]\overline{\mathfrak{X}_2[m]} \\ \mathfrak{X}_2[m]\overline{\mathfrak{X}_1[m]} & |\mathfrak{X}_2[m]|^2 \end{bmatrix} \in \mathbb{C}^{2\times 2}. \tag{3}$$

For each m, the spectral matrix $\mathfrak{F}[m]$ collects the squared amplitude of Fourier transforms of the two components $x_1[n]$ and $x_2[n]$ of the bivariate signal $\mathbf{x}[n]$ as well as their relative Fourier phase. BPR is then formulated as the problem of recovering the original bivariate signal from its spectral matrices, that is:

$$\text{find } \mathbf{X} \in \mathbb{C}^{N\times 2} \text{ given measurements } \mathfrak{F}[m], \quad m = 0, 1, \dots M-1\,. \tag{BPR}$$

Proposition 1 below shows that BPR and PPR are equivalent in the noiseless setting under very general assumptions on the projection vectors $\mathbf{b}_p$.

Proposition 1. *Suppose that the collection of projection vectors $\mathbf{b}_0, \mathbf{b}_1, \dots \mathbf{b}_{P-1} \in \mathbb{C}^2$ satisfies the condition*

$$\operatorname{span}_{\mathbb{R}} \left\{\mathbf{b}_p \mathbf{b}_p^{\mathsf{H}}\right\}_{p=0,1,\dots P-1} = \left\{\mathbf{M} \in \mathbb{C}^{2\times 2} \mid \mathbf{M}^{\mathsf{H}} = \mathbf{M}\right\} \tag{$\mathcal{H}$}$$

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

Proof. We first show that that under ($\mathcal{H}$), measurements $y_{m,p}$ in PPR can be directly expressed in terms of spectral matrices $\mathfrak{F}[m]$, which shall imply that any solution of the problem PPR is a solution of the problem BPR. Let us fix $0 \le m \le M-1$. Then one has for every $0 \le p \le P-1$

$$y_{m,p} = |\mathbf{a}_m^{\mathsf{H}} \mathbf{X} \mathbf{b}_p|^2 = \mathfrak{X}[m]^\top \mathbf{b}_p \mathbf{b}_p^{\mathsf{H}} \overline{\mathfrak{X}[m]} = \operatorname{Tr} \overline{\mathbf{b}_p}\mathbf{b}_p^\top \mathfrak{F}[m].$$

Conversely, let us assume that $\{\overline{\mathbf{b}_p}\mathbf{b}_p^\top\}_{p=0,1,\dots,P-1}$ (or equivalently, the set $\left\{\mathbf{b}_p\mathbf{b}_p^{\mathsf{H}}\right\}_{p=0,1,\dots P-1}$) is a generating family of the space of 2-by-2 Hermitian matrices. Then there exist $\lambda_{m,p} \in \mathbb{R}$, $= 0, 1, \dots, M-1$, $p = 0, 1, \dots P-1$ such that

$$\mathfrak{F}[m] = \sum_{p=0}^{P-1} \lambda_{m,p} \overline{\mathbf{b}_p}\mathbf{b}_p^\top,$$

meaning that the spectral matrices $\mathfrak{F}[m]$ can be readily obtained from scalar polarimetric projections $y_{m,p}$. It implies that any solution of the problem BPR is a solution of the problem PPR. Gathering the two parts of the proof yields Proposition 1. □

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

$$\mathbf{b}_0 = \begin{bmatrix}1\\0\end{bmatrix}, \ \mathbf{b}_1 = \begin{bmatrix}0\\1\end{bmatrix}, \ \mathbf{b}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \ \mathbf{b}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\\jmath\end{bmatrix}. \tag{4}$$

A direct check shows that rank-one matrices $\mathbf{b}_0\mathbf{b}_0^{\mathsf{H}}$, $\mathbf{b}_1\mathbf{b}_1^{\mathsf{H}}$, $\mathbf{b}_2\mathbf{b}_2^{\mathsf{H}}$, $\mathbf{b}_3\mathbf{b}_3^{\mathsf{H}}$ form a basis over the real vector space of 2-by-2 Hermitian matrices, and as a result, they are a generating family of such matrices. Polarimetric measurements in PPR read explicitly

$$\begin{aligned} y_{m,0} &= |\mathfrak{X}_1[m]|^2\,, \quad y_{m,1} = |\mathfrak{X}_2[m]|^2 \\ y_{m,2} &= \frac{1}{2}|\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2\,, \quad y_{m,3} = \frac{1}{2}|\mathfrak{X}_1[m] + \jmath\mathfrak{X}_2[m]|^2 \end{aligned} \tag{5}$$

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

$$\operatorname{real}\left(\mathfrak{X}_1[m]\overline{\mathfrak{X}_2[m]}\right) = \frac{1}{2}\left(|\mathfrak{X}_1[m] + \mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2\right) = y_{m,2} - \frac{1}{2}\left(y_{m,0} + y_{m,1}\right)\,, \tag{6}$$

$$\operatorname{imag}\left(\mathfrak{X}_1[m]\overline{\mathfrak{X}_2[m]}\right) = \frac{1}{2}\left(|\mathfrak{X}_1[m] + \jmath\mathfrak{X}_2[m]|^2 - |\mathfrak{X}_1[m]|^2 - |\mathfrak{X}_2[m]|^2\right) = y_{m,3} - \frac{1}{2}\left(y_{m,0} + y_{m,1}\right)\,. \tag{7}$$

Remark that the measurement scheme (4) yield the same quadratic measurements (5) as proposed by the authors in [28, 27]. This shows that PPR encompasses existing measurements strategies as a special case, while bringing extra flexibility in the experimental design of measurements.

2.3. Trivial ambiguities

Thanks to Proposition 1, we can now give a characterization of trivial ambiguities of PPR model by leveraging the equivalent BPR problem. Indeed, one can investigate in a rather simple way the trivial ambiguities that characterize BPR. For ease of presentation, let us extend to all $n \in \mathbb{Z}$ the bivariate signal $\{\mathbf{x}[n] = (x_1[n], x_2[n])^\top\}_{n=0,\dots,N-1}$ by zero padding for $n < 0$ and $n \geq N$. Consider the BPR measurement matrix $\mathfrak{F}[m]$ defined in (3) for an arbitrary frequency indexed by $m \in \mathbb{Z}$. Our goal in this section consists in identifying trivial operations on $\{\mathbf{x}[n]\}_{n\in\mathbb{Z}}$ that leave the measurements $\{\mathfrak{F}[m]\}_{m\in\mathbb{Z}}$ unchanged.

Global phase ambiguity. Let $\alpha \in \mathbb{R}$ and consider the bivariate signal $\mathbf{x}'$ such that $\mathbf{x}'[n] = \exp(\jmath\alpha)\mathbf{x}[n]$ for every n. Then, one has, for any $m \in \mathbb{Z}$, $\mathfrak{F}'[m] = \mathfrak{F}[m]$ since $\mathfrak{X}'_i[m] = \exp(\jmath\alpha)\mathfrak{X}_i[m]$ for $i = 1, 2$.

Time shift. Consider the time shifted signal $\mathbf{x}'$ such that $x'_1[n] = x_1[n - n_1]$ and $x'_2[n] = x'_2[n - n_2]$. Then $\mathfrak{F}'[m] = \mathfrak{F}[m]$ iff time-shifts are equal $n_1 = n_2$.

Conjugate reflection. Consider now $\mathbf{x}'$ such that $x'_1[n] = \overline{x_1[N - n]}$ and $x'_2[n] = \overline{x_2[N - n]}$. Then one has for every m

$$\mathfrak{F}'[m] = \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_2[m]\overline{\mathfrak{X}_1}[m] \\ \mathfrak{X}_1[m]\overline{\mathfrak{X}_2}[m] & |\mathfrak{X}_2[m]|^2 \end{bmatrix} = \mathfrak{F}[m]^\top. \tag{8}$$

This shows that conjugate reflection is not, in general, a trivial ambiguity for complex bivariate phase retrieval. This contrasts with standard univariate phase retrieval, see [15, 14].

Conjugate reflection can still be a trivial ambiguity provided that the measurement matrix is symmetric for every m, that is $\mathfrak{F}[m] = \mathfrak{F}[m]^\top$. Equivalently, $\mathfrak{F}[m]$ is symmetric iff $\mathfrak{X}_1[m]\overline{\mathfrak{X}_2}[m] = \mathfrak{X}_2[m]\overline{\mathfrak{X}_1}[m]$. This means that $\operatorname{imag}\left(\mathfrak{X}_1[m]\overline{\mathfrak{X}_2}[m]\right) = 0$, *i.e.* components $\mathfrak{X}_1[m]$, $\mathfrak{X}_2[m]$ are in phase at every frequency. Interestingly, this condition is interpreted in physical terms as: conjugate reflection is a trivial ambiguity for bivariate phase retrieval iff $\mathbf{x}$ is linearly polarized at all frequencies.

2.4. 1D equivalent model for PPR

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

Let us denote by $\boldsymbol{\xi} \triangleq \operatorname{vec}\mathbf{X} \in \mathbb{C}^{2N}$ the long vector obtained by stacking the two columns of $\mathbf{X}$. Using standard properties of matrix products vectorization, one can rewrite PPR measurements as

$$y_{m,p} = |\mathbf{a}_m^{\mathsf{H}}\mathbf{X}\mathbf{b}_p|^2 = |(\mathbf{b}_p^\top \otimes \mathbf{a}_m^{\mathsf{H}})\boldsymbol{\xi}|^2 = |(\overline{\mathbf{b}_p} \otimes \mathbf{a}_m)^{\mathsf{H}}\boldsymbol{\xi}|^2 \tag{9}$$

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 of vectors $\mathbf{a}$ and $\mathbf{b}$. Letting $\mathbf{c}_{m,p} \triangleq \overline{\mathbf{b}_p} \otimes \mathbf{a}_m \in \mathbb{C}^{2N}$, the PPR problem is equivalent to

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

This shows that PPR can be rewritten as a specific instance of 1D phase retrieval with structured measurements vectors $\mathbf{c}_{m,p} \in \mathbb{C}^{2N}$ called PPR-1D. While being mathematically sound, this equivalent 1D problem brings almost no insights about the bivariate nature of the signal to be recovered. Moreover, PPR-1D cannot be interpreted as a Fourier phase retrieval problem with masks [13, 30], since measurements vectors $\mathbf{c}_{m,p}$ intertwine Fourier measurements $\mathbf{a}_m$ and polarimetric projections $\mathbf{b}_p$ using Kronecker product. Thus, the study of the theoretical properties of PPR (and BPR) cannot be inferred from standard phase retrieval properties applied to PPR-1D. This requires a dedicated study, which is described in detail in Section 3. Nonetheless, as we shall see in Section 5, the equivalent formulation PPR-1D can still be particularly useful for designing algorithms to solve the original PPR problem.

3. Uniqueness and polynomial formulation

This section studies the uniqueness properties of noiseless PPR under the set of assumptions ($\mathcal{H}$) defined in Section 2.2. Thanks to Proposition 1, we see that any solution of the problem PPR is a solution of the problem BPR, and vice-versa. This formal equivalence permits to study uniqueness properties of the original PPR problem by leveraging its bivariate equivalent BPR. Pushing this idea further, we reformulate BPR using a polynomial formalism. Thus, the problem BPR appears as a crucial bridge that enables us to study the uniqueness conditions as uniqueness of a certain factorization of polynomials. This idea is classic [15, 14], but unlike the previous works, we allow for having roots at infinity, which enables us to establish the one-to-one correspondence between the two formulations, as well as a complete characterization of the uniqueness properties of BPR (and by equivalence, to those of PPR) for any signals of bounded support. In particular, this gives another view on the ambiguities in the phase retrieval problem.

3.1. Polynomials with the roots at infinity, and operations with them

In this paper, we work with polynomials which may have possibly roots at ∞. We briefly review properties of such polynomials, leaving more details in Appendix A (we also refer an interested reader to [31, §I.0], where such polynomials are used to build the algebraic theory of Hankel matrices). Formally, we define $\mathbb{C}_{\leq D}[z]$ to be the space of polynomials of degree at most D

$$A(z) = \sum_{n=0}^{D} a[n] z^n.$$

defined by a vector of coefficients $\mathbf{a} = \begin{bmatrix} a[0] & a[1] & \cdots & a[D] \end{bmatrix}^\top \in \mathbb{C}^{D+1}$. We will say that the polynomial has root at ∞ (with multiplicity μ_k) if its leading coefficient vanishes (*i.e.* if $a[D] = \cdots = a[D-\mu_k+1] = 0$). With such a convention, the following extended version of the fundamental theorem of algebra holds true: any nonzero polynomial $A \in \mathbb{C}_{\leq D}[z]$ can be uniquely (up to permutation of roots) factorized as

$$A(z) = \lambda \prod_{i=1}^{m} (z - \alpha_i)^{\mu_i} \tag{10}$$

where $\lambda \in \mathbb{C}$, $\alpha_i \in \mathbb{C} \cup \{\infty\}$ are distinct roots and μ_i are the multiplicities of α_i, so that their sum is

$$\mu_1 + \cdots + \mu_m = D;$$

in addition, the multiplication by $(z-\infty)^d$ formally means that d leading zero coefficients are appended (see also Appendix A for a formal definition).

Example 2. *Consider the following polynomial from* $\mathbb{C}_{\leq 5}[z]$*:*

$$A(z) = 0 \cdot z^5 + 0 \cdot z^4 + \frac{1}{2} z^3 + \frac{1}{2} z^2 - z \in \mathbb{C}_{\leq 5}[z]. \tag{11}$$

This polynomial has roots $\{\infty, -2, 1, 0\}$*, where the root* ∞ *has multiplicity* 2*. Hence it has the following factorization*

$$A(z) = \frac{1}{2}(z-\infty)^2 (z-1)(z+2)z.$$

We will also use the operation of conjugate reflection of $A(z) \in \mathbb{C}_{\leq D}[z]$, defined as

$$\widetilde{A}(z) = z^D \overline{A(\overline{z}^{-1})} = \sum_{n=0}^{D} \overline{a[D-n]} z^n. \tag{12}$$

Then it is easy to see that the conjugate reflection of the polynomial (10) admits a factorization

$$\widetilde{A}(z) = \widetilde{\lambda} \prod_{i=1}^{m} \left(z - \overline{\alpha_i^{-1}}\right)^{\mu_i}, \quad \text{where } \widetilde{\lambda} \triangleq \overline{\lambda} \prod_{\substack{i=1 \\ \alpha_i \neq \infty}}^{m} (-\overline{\alpha_i})^{\mu_i},$$

i.e. the roots α_i are mapped to $\overline{\alpha_i^{-1}}$, where 0 is formally assumed to be the inverse of ∞ and vice versa.

Example 3. *For Example 2, the conjugate reflection $\widetilde{A}(z) \in \mathbb{C}_{\leq 5}[z]$, as well as its factorization becomes:*

$$\widetilde{A}(z) = 0 \cdot z^5 - z^4 + \frac{1}{2}z^3 + \frac{1}{2}z^2 = (-1)(z-\infty)(z+\frac{1}{2})(z-1)z^2.$$

which has roots $\{\infty, 1, -\frac{1}{2}, 0\}$, where the root 0 has multiplicity 2.

Graphically, the conjugate reflection of the roots has a nice interpretation in terms of the Riemann sphere: the mapping of the root under conjugate reflection becomes simply a reflection with respect to the plane passing through the equator, see Fig. 2.

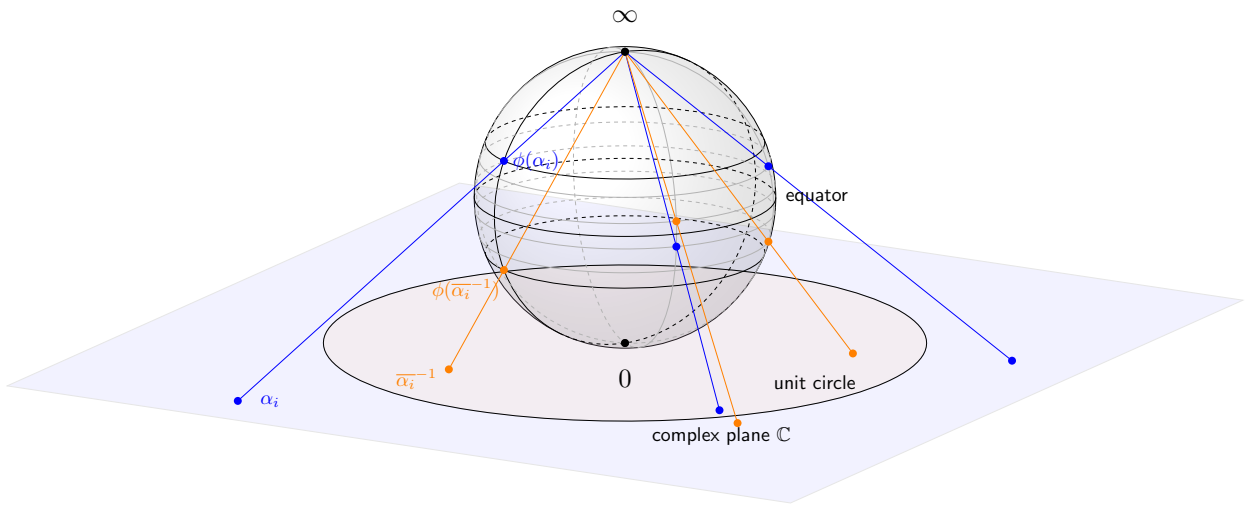

Figure 2: Complex plane and the Riemann sphere (the preimage under the stereographic projection). The conjugate inversion corresponds to reflection with respect to the equator on the Riemann sphere. Here $\phi : \mathbb{C} \to \mathcal{S}^2$ denote the inverse stereographic mapping onto the sphere $\mathcal{S}^2$.

Finally, we need to be careful when speaking about multiplication and divisors in such polynomial spaces, see also Appendix A for more details. The multiplication of polynomials $A \in \mathbb{C}_{\leq D}[z]$ and $B \in \mathbb{C}_{\leq D'}[z]$ is the polynomial $C(z) = A(z)B(z) \in \mathbb{C}_{\leq (D+D')}[z]$. Conversely, the polynomial $C(z)$ has a divisor $A(z)$ if it can be represented as $C(z) = A(z)B(z)$ in this sense. Similarly to the standard case, the greatest common divisor exists and is unique (up to a multiplicative constant) if at least one polynomial is non-zero. Note that divisibility takes into account the roots at ∞ and their multiplicities.

Example 4. *The polynomial $0 \cdot z^3 + 0 \cdot z^2 + z + 2 = (z-\infty)^2(z+2) \in \mathbb{C}_{\leq 3}[z]$ is a divisor of the polynomial $A(z)$ in from Example 2, but the polynomial $0 \cdot z^4 + 0 \cdot z^3 + 0 \cdot z^2 + z + 2 = (z-\infty)^3(z+2) \in \mathbb{C}_{\leq 4}[z]$ is not, because there are not enough infinite roots in the expansion of $A(z)$.*

3.2. Phase retrieval as a polynomial factorization problem

In this subsection, we are going to provide a polynomial reformulation of BPR. First, we define the following four polynomials as generating polynomials of the components of the bivariate signal $\mathbf{x}[n] =$

$(x_1[n], x_2[n])^\top \in \mathbb{C}^2$, $n = 0, 1, \dots N-1$ and their conjugate reflections, all belonging to $\mathbb{C}_{\leq N-1}[z]$

$$X_1(z) = \sum_{n=0}^{N-1} x_1[n] z^n, \quad \widetilde{X}_2(z) = \sum_{n=0}^{N-1} \overline{x_1}[N-n+1] z^n,$$
$$X_2(z) = \sum_{n=0}^{N-1} x_2[n] z^n, \quad \widetilde{X}_2(z) = \sum_{n=0}^{N-1} \overline{x_2}[N-n+1] z^n.$$

Then we define the following matrix polynomial

$$\mathbf{\Gamma}(z) = \begin{bmatrix} \Gamma_{11}(z) & \Gamma_{12}(z) \\ \Gamma_{21}(z) & \Gamma_{22}(z) \end{bmatrix} = \begin{bmatrix} X_1(z)\widetilde{X}_1(z) & X_1(z)\widetilde{X}_2(z) \\ X_2(z)\widetilde{X}_1(z) & X_2(z)\widetilde{X}_2(z) \end{bmatrix} = \begin{bmatrix} X_1(z) \\ X_2(z) \end{bmatrix} \begin{bmatrix} \widetilde{X}_1(z) & \widetilde{X}_2(z) \end{bmatrix}, \tag{13}$$

where each element is a polynomial $\Gamma_{ij} \in \mathbb{C}_{\leq 2N-2}[z]$, and we will use the notation $\mathbf{\Gamma} \in \mathbb{C}^{2\times 2}_{\leq 2N-2}[z]$.

The coefficients of these polynomials are nothing but the covariance functions (auto-correlation and cross-correlation) of the signals X_1 and X_2; in addition, the spectral matrices $\mathfrak{F}[m]$ appearing in BPR are linked to the evaluations of the polynomial $\mathbf{\Gamma}(z)$.

Lemma 1. *The coefficients Γ_{ij} of the matrix polynomial can be expressed as*

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

where $x_i[n] = 0$ for $n < 0$ and $n \geq N$ by convention, and $\gamma_{ij}[n]$ are defined for $n \in \{-N+1, \dots, N-1\}$. Moreover, the spectral matrices $\mathfrak{F}[m]$ appearing in BPR can be expressed as

$$\mathfrak{F}[m] = e^{\jmath 2\pi \frac{m(N-1)}{M}} \mathbf{\Gamma}(e^{-\jmath 2\pi \frac{m}{M}}). \tag{14}$$

This result is well-known and follows from the spectral (Fourier) representation of the signals, but we give a formal proof in Appendix B, because we consider finite support signals and extended polynomials.

Therefore, we will refer to $\Gamma_{ij}(z)$ as *measurement polynomials*. Note that the coefficients of the measurement polynomials $\Gamma_{ij}(z)$ can be uniquely identified from the measurement polynomials from the Fourier measurements $\{\mathfrak{F}[m]\}_{m=0,\dots,M-1}$, if the number of measurements M exceeds the degree of these polynomials, *i.e.* $2N-2$, by at least one:

$$M \geq 2N-1 \tag{15}$$

which is the well-known oversampling condition in standard univariate Fourier phase retrieval, see e.g. [15].

Therefore, the problem BPR is equivalent to the following recovery problem, which we refer to as Polynomial Autocorrelation Factorization (PAF) as to emphasize that we factorize the autocovariance measurements in the polynomial form. The polynomial reformulation of the problem is very helpful for establishing the uniqueness conditions for BPR. Notably it enables a complete characterization of its uniqueness properties in terms of algebraic properties of complex polynomials.

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

find the polynomials $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$ given measurement polynomials $\Gamma_{ij}(z)$ *defined as* (13) , (PAF)

i.e. *there is a one-to-one correspondence between the data ($\mathbf{\Gamma}(z)$ and $\mathfrak{F}[m]$) as well as the sets of solutions of the problems (polynomials X_1, X_2 and bivariate signal components $x_1[n], x_2[n]$).*

The proof can be summarized in Figure 3, and the formal proof can be found in Appendix B.

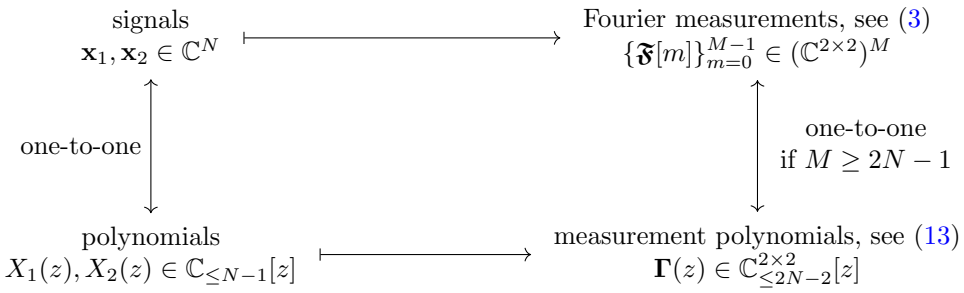

Figure 3: Equivalences of data and solutions in problems BPR and PAF.

3.3. General uniqueness result

We now derive a full characterization of the uniqueness properties of the polynomial factorization problem PAF. It is important to keep in mind that by the set of equivalences given in Figure 3, these results also provide a complete characterization of uniqueness properties of PPR and BPR. The following fundamental lemma establishes that the uniqueness properties of the PAF problem essentialy boil down to a specific spectral factorization problem.

Lemma 2. *Let* $Q(z) \triangleq \gcd(X_1, X_2)$ *where* $Q \in \mathbb{C}_{\leq D}[z]$ *and* $R_1, R_2 \in \mathbb{C}_{\leq N-D-1}[z]$ *be the corresponding quotients,* i.e. $X_1(z) = Q(z)R_1(z)$ *and* $X_2(z) = Q(z)R_2(z)$ *with* $\gcd(R_1, R_2) = 1$. *Then we have that*

1) the GCD of the autocovariance polynomials $H(z) \triangleq \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22})$ *must have the form*

$$H(z) = cQ(z)\tilde{Q}(z), \quad c \neq 0; \tag{16}$$

2) given the quotients $R_{ij}(z)$ *(*i.e. $\Gamma_{ij}(z) = H(z)R_{ij}(z)$, $\gcd(R_{11}, R_{12}, R_{21}, R_{22}) = 1$*), the quotients of* X_1, X_2 *are determined up to a multiplicative constant as*

$$R_1(z) = \gcd\{R_{11}(z), R_{12}(z)\}, \qquad R_2(z) = \gcd\{R_{21}(z), R_{22}(z)\}. \tag{17}$$

Proof. Start by 1). Direct calculations show that, for $i, j = 1, 2$, $\Gamma_{ij}(z) = X_i(z)\tilde{X}_j(z) = R_i(z)\tilde{R}_j(z)Q(z)\tilde{Q}(z)$. Then the GCD of autocovariance polynomials can be explicitly computed as

$$\begin{aligned} H(z) &= \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}) \\ &= \gcd\left(\gcd\left(\Gamma_{11}, \Gamma_{12}\right), \gcd\left(\Gamma_{21}, \Gamma_{22}\right)\right) \\ &= \gcd\Big(R_1(z)Q(z)\tilde{Q}(z), R_2(z)Q(z)\tilde{Q}(z)\Big) && \text{since } \gcd(\tilde{R}_1, \tilde{R}_2) = 1 \\ &= cQ(z)\tilde{Q}(z) && \text{since } \gcd(R_1, R_2) = 1 \end{aligned}$$

Proof of 2). From the previous point, we have that $R_{ij}(z) = c^{-1}R_i(z)\tilde{R}_j(z)$. The determination (17) of R_1 and R_2 is then straightforward using that $\gcd(R_1, R_2) = \gcd(\tilde{R}_1, \tilde{R}_2) = 1$ by assumption. □

Lemma 2 shows that the study of the uniqueness properties of PAF is directly related to uniqueness of the spectral factorization (16), *i.e.* the recovery of $Q(z)$ given $H(z)$. Indeed, if $Q(z)$ can be uniquely recovered from $H(z)$, then the polynomials $X_1(z), X_2(z)$ can be found by mutiplying $Q(z)$ with the quotients $R_1(z)$ and $R_2(z)$ obtained in (17). Before giving the sufficient and necessary uniqueness condition, we make a remark about the roots of the product $Q(z)\widetilde{Q}(z)$ which are key to understanding uniqueness.

Remark 1. *Let* $Q(z) = \lambda \prod_{i=1}^{D}(z-\alpha_i)$ *(with possibly repeating* α_i*). Then* $H(z) = cQ(z)\tilde{Q}(z)$ *has the following factorization*

$$H(z) = c\lambda\widetilde{\lambda}\prod_{i=1}^{D}(z-\alpha_i)(z-\overline{\alpha_i^{-1}}). \tag{18}$$

Furthermore, if $|\alpha| = 1$*, then* $\alpha = \overline{\alpha^{-1}}$*. Therefore, a unit-modulus* α *is a root of* $Q(z)$ *of multiplicity* μ *if and only if it is a root of* $H(z)$ *of multiplicity* 2μ*.*

From Remark 1, we see that the unit-modulus roots of $H(z)$ do not contribute to ambiguity of PAF. Indeed, all unit-modulus roots of $Q(z)$ can be uniquely retrieved from $H(z)$. This helps us to establish a necessary and sufficient condition for uniqueness.

Theorem 2 (Uniqueness of PAF)**.** *The following equivalences are true:*

$$\text{PAF } \textit{admits a unique solution} \Leftrightarrow X_1(z) \textit{ and } X_2(z) \textit{ have no common roots outside the unit circle.} \tag{19}$$

$$\Leftrightarrow H(z) = \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}) \textit{ has no roots outside the unit circle.} \tag{20}$$

Proof. The last equivalence being trivial by Lemma 2, we prove only the first equivalence.

- $\boxed{\Rightarrow}$ Suppose that the solution of PAF is essentially unique, but the polynomial $Q(z)$ has a root α outside the unit circle. Then easy calculations show that polynomial $Q'(z) = \frac{Q(z)(z-\overline{\alpha^{-1}})}{(z-\alpha)}$ satisfies

  $$Q'(z)\widetilde{Q'}(z) = Q(z)\widetilde{Q}(z).$$

  On the other hand $Q'(z)$ is not proportional to $Q(z)$, and therefore polynomials $X_1'(z) \triangleq Q'(z)R_1(z)$, $X_2'(z) \triangleq Q'(z)R_2(z)$ are not proportional to $X_1(z)$ and $X_2(z)$, but give an alternative pair of polynomials such that $\Gamma_{ij}'(z) = \Gamma_{ij}(z)$ (a contradiction).

- $\boxed{\Leftarrow}$ Assume that $H(z)$ has only unit-modulus roots. Then there is a unique monic polynomial $Q(z)$ such that $H(z) = cQ(z)\widetilde{Q}(z) = c(Q(z))^2$. Therefore, by Lemma 2, we can find polynomials $R_1'(z)$ and $R_2'(z)$ such that

  $$X_1(z) = c_1R_1'(z)Q(z), \quad X_2 = c_2R_2'(z)Q(z), \quad c_1, c_2 \in \mathbb{C}.$$

  The pair (c_1, c_2) can be determined from the relations (13) up to a common unit-modulus factor accounting for the global rotation ambiguity, see Section 2.3. □

Remark 2. *Note that the uniqueness condition given in Theorem 2 clarifies previous statements made in the literature [28, 27]. In particular, in [28, Theorem 1] it was claimed that a necessary and sufficient for uniqueness of the solution of* BPR *is the coprimeness of the polynomials* $X_1(z)$ *and* $X_2(z)$*. Our Theorem 2 shows that it was just a sufficient condition, because unimodular roots do not affect uniqueness. This agrees with a similar behavior observed for standard univariate 1D phase retrieval, see [14].*

Remark 3 (Almost everywhere uniqueness of PAF)**.** *Whereas the analysis of non-uniqueness properties of* PAF *follows closely that of the standard phase retrieval problem, a distinctive feature of the bivariate/polarimetric phase retrieval problem is that it is almost everywhere unique. This can be seen by observing that the set of polynomials* $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$ *with at least one common root is an algebraic variety of dimension smaller than* $2N-1$*; , hence it is of measure zero. Put it differently, this shows that* PAF *has the appealing property that almost all polynomials* $X_1, X_2 \in \mathbb{C}_{\leq N-1}[z]$ *can be uniquely recovered by the set of measurement polynomials* $\Gamma_{ij}(z)$*.*

3.4. Ambiguities and counting the number of solutions

In this section, we refine Theorem 2 by providing the number of solutions and describing the set of solution of PAF. Note that, that Lemma 2 implies that the uniqueness properties of PAF resumes in essence to that of a standard univariate phase retrieval problem (taking into account specificites related to the bivariate/polarimetric phase retrieval setting). Indeed the uniqueness of the univariate phase retrieval is determined by the uniqueness [14, 15] of the factorization $H(z) = Q(z)\tilde{Q}(z)$. In principle, we could invoke the existing results from [14], however, we prefer to give a complete characterization that relies upon the formalism with 0 and ∞ roots used in this paper. In particular, this formalism allows us to treat the time shift ambiguities in a unified manner.

Theorem 3 (Number of solutions of PAF)**.** *Let $H(z) = Q(z)\tilde{Q}(z)$ and $\mu_1, \ldots, \mu_{N_D}$ be the respective multiplicities of the N_D non-unit-modulus roots pairs $(\delta_1, \overline{\delta}_1^{-1}), \ldots, (\delta_{N_D}, \overline{\delta}_{N_D}^{-1})$ of $H(z)$. Then the problem PAF admits exactly*

$$\prod_{i=1,|\delta_i|\neq 1}^{N_D} (\mu_i + 1) \tag{21}$$

different solutions, where only non-unimodular common roots of X_1 and X_2 contribute to the total number of solutions. In particular, when common roots are all simple and outside the unit circle, there is exactly 2^{N_D} different solutions.

Proof. Lemma 2 shows that the number of solutions of PAF is exactly the number of different (up to multiplication by a scalar) polynomials $Q(z)$ such that $H(z) = Q(z)\tilde{Q}(z)$. This *spectral factorization* problem is equivalent to selecting the roots of $Q(z)$ amongst the root pairs $(\delta_i, \overline{\delta}_i^{-1})$ of $H(z)$. Since $\mu_1, \ldots, \mu_{N_D}$ are the multiplicities of the root pairs of $H(z)$, then for each root pair $(\delta_i, \overline{\delta}_i^{-1})$ one has to select exactly μ_i roots among those pairs, leading to a polynomial $Q(z)$ of degree $D = \mu_1 + \ldots + \mu_{N_D}$.

Consider a non-unit root pair $(\delta_i, \overline{\delta}_i^{-1})$ with multiplicity μ_i; then the number of different combinations of μ_i roots is that of a random draw of $k = \mu_i$ items with replacement in a set of $n = 2$ elements, *i.e.* $(n+k-1)!/(k!(n-1)!) = \mu_i + 1$. Repeating the same process for each root pair gives the total number (21) of solutions to the PAF problem. □

Remark 4 (Counting multiplicities)**.** *The number of solutions (21) depends in fact on the multiplicities of pairs of roots of $Q(z)$. This means in particular that if δ and $\overline{\delta}^{-1}$ are roots of $Q(z)$ with multiplicities μ_1 and μ_2, then the multiplicity of the pair $(\delta, \overline{\delta}^{-1})$ of $H(z)$ is equal to $\mu_1 + \mu_2$. The same applies to 0 and ∞ roots.*

Example 5. *Consider $Q(z) = A(z)$ that is the polynomial from Example 2 having double ∞ root and simple roots $\{-2, 1, 0\}$. Then the polynomial $H(z) = A(z)\tilde{A}(z)$, where is*

$$H(z) = Q(z)\tilde{Q}(z) = -\frac{1}{2}(z-\infty)^3\left(z+\frac{1}{2}\right)(z+2)(z-1)^2z^3.$$

The multiplicity of the root pair $(0, \infty)$ is 3, while the root pair $(-2, -\frac{1}{2})$ have multiplicity 1. This yields a total of $4 \cdot 2 = 8$ solutions, where the other factorizations are given by permuting 0 and ∞ roots or/and replacing root -2 with $-\frac{1}{2}$, For example, some of possible alternative factorisations are given by $Q(z) = \frac{1}{2}(z+2)(z-1)z^3$ or $Q(z) = (z+\frac{1}{2})(z-1)z^3$.

Remark 5 (time shifts and common roots at 0 or ∞)**.** *One striking benefit of the use of polynomials with roots at the infinity is that it permits to recover the trivial shift ambiguity as part of the set of solutions to PAF. To illustrate this, let us consider the particular example of a bivariate signal $\{\mathbf{x}[n]\}_{n=0,\ldots N-1}$ where $\mathbf{x}[0] = \mathbf{x}[N-1] = \mathbf{0}$. We further assume that the polynomials associated to the sub-signal $\{\mathbf{x}[n]\}_{n=1,\ldots N-2}$ have no roots in common. Thus, the polynomials $X_1(z)$ and $X_2(z)$ share only two roots: 0 and ∞, which turn out to be "conjugate reflected" from one another. The polynomial $H(z) = Q(z)\tilde{Q}(z)$ has then a double*

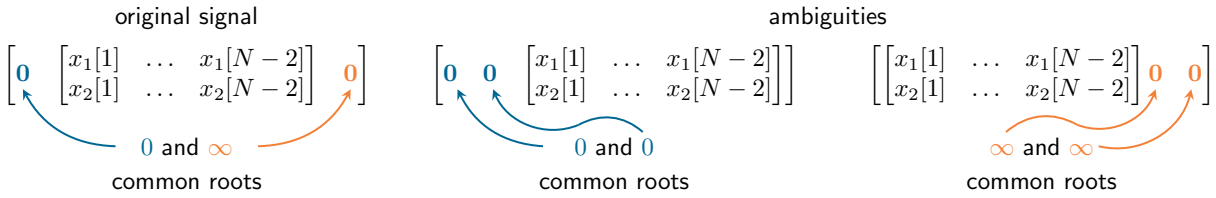

Figure 4: The three solutions of PAF obtained by selecting either the root 0 or ∞ for each single root pair $(0, \infty)$ of $H(z)$. These three solutions are simple time-shifted versions of one-another, showing that Theorem 3 encompasses the time-shift trivial ambiguity as a special case.

root pair $(0, \infty)$ *meaning that by* (21) *the* PAF *problem has exactly three different solutions – which are simply trivially shifted versions from one another. Figure 4 depicts these three solutions obtained by selecting either the root* 0 *or* ∞ *for each single root pair* $(0, \infty)$ *of* $H(z)$.

Remark 6 (On conjugate reflection ambiguity)**.** *Assume that* $\mathbf{x}[0], \mathbf{x}[N-1] \neq \mathbf{0}$*, so that time-shifts are not part of the total number of solutions* (21)*. Then, in the non-unique case, the number of non-trivially different solutions* (21) *for* PAF *(and by equivalence, for* PPR *and* BPR*) is twice that of the standard univariate 1D phase retrieval problem [15, 14]. This can be explained by the fact that conjugate reflection is not, in general, a trivial ambiguity for the bivariate case (see Section 2.3). More precisely, this means that unlike the univariate case, exchanging common roots* $(\delta_1, \ldots, \delta_{N_D})$ *with their conjugated reflected versions* $(\overline{\delta_1}^{-1}, \ldots, \overline{\delta_{N_D}}^{-1})$ *do not yield to a trivial ambiguity.*

We conclude the study of uniqueness properties of PAF in Theorem 4 below. It provides an explicit expression of PAF solutions in the simplified case where there are no 0 or ∞ roots in common, meaning that $\mathbf{x}[0] \neq \mathbf{0}$ and $\mathbf{x}[N-1] \neq \mathbf{0}$.

Theorem 4 (Expression of PAF solutions)**.** *Suppose that* $H(z) = \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_2) = Q(z)\tilde{Q}(z)$ *with* D *pair roots* $(\delta_i, \overline{\delta}_i^{-1})$*. Let* $X_1(z) = Q(z)R_1(z)$ *and* $X_2(z) = Q(z)R_2(z)$*. Denote by* α_{1i} *and* α_{2i} *the roots of* $R_1(z)$ *and* $R_2(z)$*, respectively, Then all solutions* $X_1'(z)$ *and* $X_2'(z)$ *to the* PAF *problem can be expressed as*

$$X_1'(z) = e^{\jmath\theta}\lambda_1 \prod_{i=1}^{D}(z-\beta_i) \prod_{i=1}^{N-D-1}(z-\alpha_{1i}) \tag{22}$$

$$X_2'(z) = e^{\jmath\theta}\lambda_2 \prod_{i=1}^{D}(z-\beta_i) \prod_{i=1}^{N-D-1}(z-\alpha_{2i}) \tag{23}$$

where each β_i *is chosen amongst zeros pairs* $(\delta_i, \overline{\delta}_i^{-1})$ *of* $H(z)$*; the angle* $\theta \in (-\pi, \pi)$ *accounts for the global phase trivial ambiguity. The constants* $\lambda_1, \lambda_2 \in \mathbb{C}$ *are given by*

$$\lambda_1 = \sqrt{|\gamma_{11}[N-1]| \prod_{i=1}^{D}|\beta_i|^{-1} \prod_{i=1}^{N-D-1}|\alpha_{1i}|^{-1}} \tag{24}$$

$$\lambda_2 = e^{\jmath\Delta}\sqrt{|\gamma_{22}[N-1]| \prod_{i=1}^{D}|\beta_i|^{-1} \prod_{i=1}^{N-D-1}|\alpha_{2i}|^{-1}}\,, \tag{25}$$

where Δ *reads*

$$\Delta = \pi(N-1) + \arg\gamma_{12}[N-1] + \sum_{i=1}^{D}\arg\beta_i + \sum_{i=1}^{N-D-1}\arg\alpha_{2i}\,. \tag{26}$$

Proof. See Appendix C. □

3.5. Uniqueness in practice

We provide in this section a small numerical study which illustrates how uniqueness of PAF can quickly occur in practice. Let us consider the case of a bivariate signal with constant polarization state, which is one of the simplest models for bivariate signals. Formally, such signal $\{\mathbf{x}[n]\}_{n=0,1,\dots,N-1}$ can be written as

$$\mathbf{x}[n] = \begin{bmatrix} a \\ b \end{bmatrix} s[n], \quad \mathbf{s} \in \mathbb{C}^N, a, b \in \mathbb{C}\,. \tag{27}$$

In (27), the complex-valued signal $\mathbf{s}$ defines the time or 1D-spatial dynamics of the bivariate signal $\mathbf{x}[n]$, whereas complex constants a and b define its polarized state. For instance, when $a \neq 0, b = 0$, the signal $\mathbf{x}[n]$ is said to be linearly horizontally polarized; similarly for $a = 0, b \neq 0$, it is said to be linearly horizontally polarized. Finally, when e.g. $a = 1$ and $b = \pm\jmath$, (27) is that of a circularly polarized signal, since the two entries of $\mathbf{x}[n]$ are in quadrature with one another.

Regarding the uniqueness of the bivariate signal defined by (27), we observe that the polynomials $X_1(z)$ and $X_2(z)$ associated with entries of $\mathbf{x}[n]$ share the same $N-1$ roots[1], those of the complex polynomial $S(z) \triangleq \sum_{n=0}^{N-1} s[n]z^n$. Thus, according to Theorem 3 there are up to 2^{N-1} different solutions for the bivariate phase retrieval problem. Notably, if $S(z)$ has no roots on the unit circle, and if they are all distinct from one another, then there are exactly 2^{N-1} different solutions for the bivariate phase retrieval problem. We now consider two perturbations models of the constant polarized signal (27):

- *single entry perturbation*: we consider a perturbated signal $\{\mathbf{x}_{n_0,\boldsymbol{\epsilon}}[n]\}_{n=0,1,\dots,N-1}$, where n_0 is a randomly selected index from the uniform discrete distribution on $[\![0, N-1]\!]$ and

$$\mathbf{x}_{n_0,\boldsymbol{\epsilon}}[n_0] = \mathbf{x}[n] + \boldsymbol{\epsilon}, \boldsymbol{\epsilon} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2\mathbf{I}_2), \quad \mathbf{x}_{n_0,\boldsymbol{\epsilon}}[n] = \mathbf{x}[n] \quad \text{for } n \neq n_0 \tag{28}$$

  with $\mathcal{N}_{\mathbb{C}}(0, \sigma^2\mathbf{I}_2)$ the complex normal distribution of variance σ^2.

- *full signal perturbation*: this time the perturbation is applied to all entries of $\{\mathbf{x}[n]\}_{n=0,1,\dots,N-1}$ such that

$$\mathbf{x}_{\boldsymbol{\epsilon}}[n] = \mathbf{x}[n] + \boldsymbol{\epsilon}[n], \qquad \boldsymbol{\epsilon}[n] \sim \mathcal{N}_{\mathbb{C}}\left(0, \frac{\sigma^2}{N}\mathbf{I}_2\right) \tag{29}$$

  for every $n \in [\![0, N-1]\!]$. Note that we normalized by N the variance of the perturbation to ensure proper comparison between the two models.

To assess numerically whether these perturbated signals remains non-uniquely recoverable by PAF, we use two different but complementary figures of merits. The first one is the rank deficiency of the standard Sylvester matrix of polynomials $X_1(z)$ and $X_2(z)$: it is upper bounded by $N/2$ when $\exists c \in \mathbb{C}$ such that $X_2(z) = cX_1(z)$, and lower bounded by 0 when $\gcd(X_1(z), X_2(z)) = 1$, *i.e.* the Sylvester matrix is full-rank. See also details in Section 4.2 further below. The second figure of merit is *root separation*, *i.e.* the minimal Euclidean distance between roots of $X_1(z)$ and $X_2(z)$.

Figure 5 displays the evolution of two non-uniqueness metrics with respect to the perturbation standard deviation σ arising in models (28) and (29). Results have been obtained by averaging 1000 independent realizations for each value of σ. We used an arbitrary randomly generated bivariate signal with constant polarization (27) for $N = 32$ and unit norm. Regarding the rank deficiency criteria, both perturbation models lead to maximum rank-deficiency (non-unique case where roots of $X_1(z)$ and $X_2(z)$ all coincide) for very small values of $\sigma < 10^{-13}$. For $\sigma > 10^{-13}$ we observe a rapid phase transition to zero-rank deficiency (uniqueness) which appears from $\sigma > 10^{-10}$ for the full signal perturbation model (29). A similar but slower phenomenon arises for the single-entry perturbation model (28), where we observe uniqueness for $\sigma \geq 10^{-5}$.

[1] we assume here for simplicity that nor $a = 0$ nor $b = 0$, so that polynomials are properly defined.

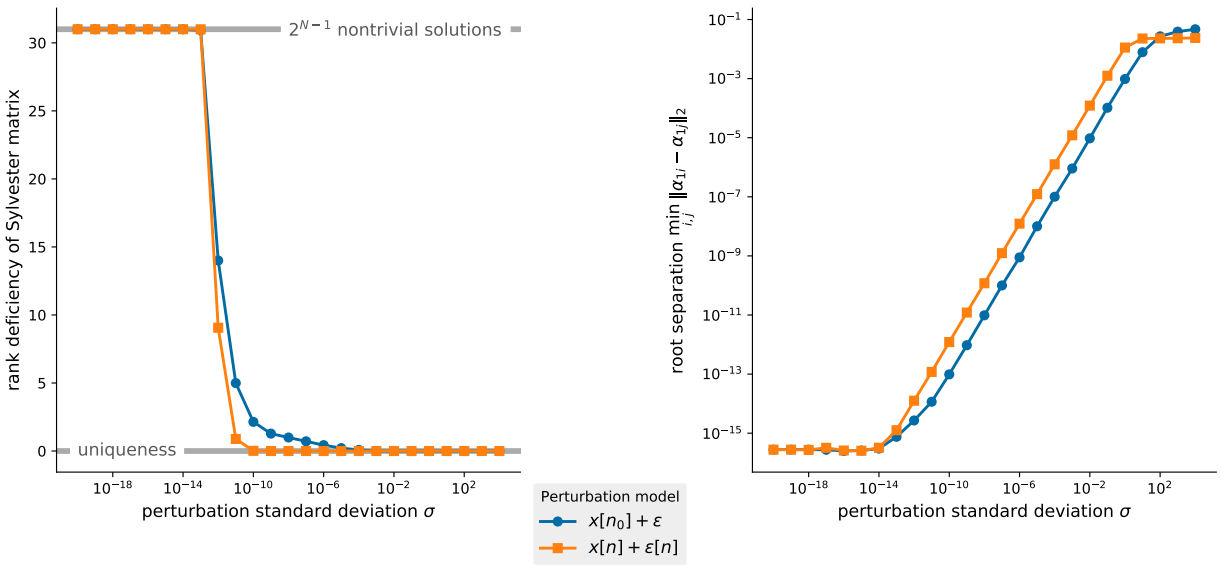

Figure 5: Illustration of the uniqueness properties of PAF by studying the evolution of two figures of merit as the standard deviation σ of the perturbation increases for models (28) and (29).

It is interesting to note that the start of the transition coincides with that of root separation criteria, which starts to increase for $\sigma > 10^{-14}$. Moreover, it appears that an average root separation of the order of 10^{-11} is sufficient to ensure uniqueness with model (29), whereas it should be of the order of 10^{-8} with model (28).

In summary, we see that very small perturbations (say, roughly $\sigma \approx 10^{-6}$) can already transform a non-uniquely recoverable signal from PAF into a uniquely recoverable one. This provides a numerical illustration of the fact that PAF is almost everywhere unique, since the set of complex polynomials with at least one common roots is of measure zero. Compared to standard phase retrieval, where additional measurements are needed to ensure uniqueness (see [15] and references therein), BPR and its equivalent PPR inherently ship with nice uniqueness properties.

4. Solving PPR: algebraic methods

We propose two different but complementary strategies for solving PPR in practice. This first section describes algebraic approaches which exploits the peculiarities of PPR through its equivalent BPR formulation. In particular, it leverages the polynomial representation PAF which underpins uniqueness results of Section 3. Iterative algorithms for solving PPR are described later on in Section 5.

In the sequel, we assume that measurements are corrupted by additive i.i.d. Gaussian noise such that for $m = 0, 1, \dots M-1$ and $p = 0, 1, \dots P-1$

$$y_{m,p} = |\mathbf{a}_m^{\mathsf{H}} \mathbf{X} \mathbf{b}_p|^2 + n_{m,p}, \quad n_{m,p} \sim \mathcal{N}(0, \sigma^2) \tag{30}$$

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

$$\mathrm{SNR} = \frac{\sum_{m=0}^{M-1} \sum_{p=0}^{P-1} |\mathbf{a}_m^{\mathsf{H}} \mathbf{X} \mathbf{b}_p|^4}{MP\sigma^2} \tag{31}$$

Algebraic approaches exploits two key properties of PPR in the noiseless case: *(i)* it is equivalent as BPR under the nonrestrictive hypothesis set ($\mathcal{H}$) and *(ii)* BPR itself can be equivalently formulated as a

polynomial factorization problem PAF, see Section 3.2 for details. Notably, a key result from Section 3 is that polynomials $X_1(z)$ and $X_2(z)$ can be uniquely recovered (up to trivial ambiguities) as greatest common divisors when BPR is unique.

In practice, such an approach can be performed in two stages. One first needs to reconstruct the measurement polynomials $\Gamma_{ij}(z)$, $i = 1, 2$ given scalar, possibly noisy PPR measurements $y_{m,p}$, $m = 0, 1, \ldots, M-1$, $p = 0, 1, \ldots, P-1$. Second, using techniques from approximate GCD computations [32], we recover $X_1(z)$ and $X_2(z)$ from the measurements polynomials $\Gamma_{ij}(z)$.

4.1. Reconstruction of measurement polynomials

Recall that by Lemma 1 measurement polynomials $\Gamma_{ij}(z)$ can be readily expressed in terms of correlation functions $\gamma_{11}[n], \gamma_{12}[m], \gamma_{21}[m], \gamma_{22}[m]$. Thus, recovery of polynomials $\Gamma_{ij}(z)$'s is identical to the recovery of $\{\gamma_{ij}[n]\}_{n\in\mathbb{Z}}$ for $i, j = 1, 2$. Equivalently, by discrete Fourier transformation, one must retrieve the spectral matrix $\mathfrak{F}[m]$ for $m = 0, 1, \ldots, M-1$ from PPR measurements.

Consider noisy measurements given by (30). Since $|\mathbf{a}_m^{\mathsf{H}}\mathbf{X}\mathbf{b}_p|^2 = \operatorname{Tr}\overline{\mathbf{b}_p}\mathbf{b}_p^{\top}\mathfrak{F}[m]$, an estimate $\hat{\mathbf{\Gamma}}[m]$ of $\mathfrak{F}[m]$ is found for every m by minimizing the following quadratic-loss

$$\hat{\mathbf{\Gamma}}[m] = \underset{\substack{\tilde{\mathbf{\Gamma}}[m]=\tilde{\mathbf{\Gamma}}[m]^{\mathsf{H}} \\ \operatorname{rank}\tilde{\mathbf{\Gamma}}[m]=1}}{\arg\min} \sum_{p=0}^{P-1} \left(y_{m,p} - \operatorname{Tr}\overline{\mathbf{b}_p}\mathbf{b}_p^{\top}\tilde{\mathbf{\Gamma}}[m] \right)^2 , \tag{32}$$

where the Hermitian and rank-one constraint ensures the estimated spectral matrix $\hat{\mathbf{\Gamma}}[m]$ to have the right structure for future polynomial gcd computations.

To solve (32), we adopt a heuristic but simple strategy similar to practical polarimetric reconstruction techniques used in optics [33, 34]. First, we exploit the *Stokes parameters* representation of 2-by-2 Hermitian matrices, which read for an arbitrary Hermitian matrix $\mathbf{M} \in \mathbb{C}^{2\times 2}$

$$\mathbf{M} = \frac{1}{2}\begin{bmatrix} S_0 + S_1 & S_2 + \jmath S_3 \\ S_2 - \jmath S_3 & S_0 - S_1 \end{bmatrix} \qquad S_0, S_1, S_2, S_3 \in \mathbb{R}\,. \tag{33}$$

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

$$\operatorname{Tr}\overline{\mathbf{b}_p}\mathbf{b}_p^{\top}\tilde{\mathbf{\Gamma}}[m] = \left[\mathcal{S}\left(\overline{\mathbf{b}_p}\mathbf{b}_p^{\top}\right)\right]^{\top}\mathcal{S}\left(\tilde{\mathbf{\Gamma}}[m]\right)\,. \tag{34}$$

Thus, for m fixed, we can set $\mathbf{y}_{m,:} \triangleq (y_{m,0}, y_{m,1}, \ldots, y_{m,P-1})^{\top} \in \mathbb{R}_+^P$ and the polarization measurement matrix $\mathbf{D} \in \mathbb{R}^{P\times 4}$ such that its p-th row reads $\mathbf{D}_p = \left[\mathcal{S}\left(\overline{\mathbf{b}_p}\mathbf{b}_p^{\top}\right)\right]^{\top}$, leading us to rewrite problem (32) as

$$\hat{\mathbf{\Gamma}}[m] = \underset{\substack{\tilde{\mathbf{\Gamma}}[m]=\tilde{\mathbf{\Gamma}}[m]^{\mathsf{H}} \\ \operatorname{rank}\tilde{\mathbf{\Gamma}}[m]=1}}{\arg\min} \left\| \mathbf{y}_{m,:} - \mathbf{D}\mathcal{S}\left(\tilde{\mathbf{\Gamma}}[m]\right) \right\|_2^2 . \tag{35}$$

A possibly sub-optimal yet very simple solution to (35) consists in finding the best rank-one approximation of the classical least square estimator of Stokes parameters, *i.e.*

$$\hat{\mathbf{\Gamma}}[m] = \operatorname{rank1}\left\{\mathcal{S}^{-1}\left(\mathbf{D}^{\dagger}\mathbf{y}_{m,:}\right)\right\} \tag{36}$$

where $\mathbf{D}^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of $\mathbf{D}$, $\mathcal{S}^{-1}$ is the inverse Stokes mapping defined by (33). The operator rank1 finds the best rank-one approximation of a given matrix with respect to the Frobenius norm. More precisely, consider the SVD of a 2-by-2 Hermitian matrix $\mathbf{M} = \mathbf{U}\operatorname{diag}(\sigma_0, \sigma_1)\mathbf{U}^{\mathsf{H}}$, one has $\operatorname{rank1}(\mathbf{M}) = \sigma_0\mathbf{u}_0\mathbf{u}_0^{\mathsf{H}}$, where σ_0 and $\mathbf{u}_0$ are respectively the largest singular value and its corresponding singular vector.

4.2. Sylvester matrices and GCD

After estimating the matrices $\hat{\mathbf{\Gamma}}[m]$, we will build the (estimated) matrix polynomial $\hat{\mathbf{\Gamma}}(z)$, and then our goal is to solve (approximately) the problem PAF. Thanks to Lemma 2 and Theorem 4, the solution can be found through the GCD (or an approximate common divisors) of the polynomials. The GCD (or approximate GCD) can be found thanks to the correspondence between (greatest) common divisors and low-rank Sylvester matrices, reviewed below (see also [32] for more details).

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

$$\mathcal{S}_D(A,B) \triangleq \left[\begin{array}{ccc|ccc} a_0 & & & b_0 & & \\ \vdots & \ddots & & \vdots & \ddots & \\ a_L & & a_0 & b_L & & b_0 \\ & \ddots & \vdots & & \ddots & \vdots \\ & & a_L & & & b_L \end{array}\right] \in \mathbb{C}^{(2L-D+1)\times 2(L-D+1)}. \tag{37}$$

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

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

For an overview of such matrices and the corresponding literature, we refer to [32] (note that unlike [32] we use the same notation for subresultant and extended Sylvester matrices). The following theorem is classic.

Theorem 5 (Sylvester)**.** *Two polynomials $A, B \in \mathbb{C}_{\leq L}[z]$ have a non-trivial common divisor if and only if $\mathcal{S}_1(A,B)$ is rank deficient. Moreover the (extended) degree K of $\gcd(A(z), B(z))$ is equal to the rank defect of $\mathcal{S}_1(A,B)$,* i.e.

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

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

Remark 7. *Note that we use the term "extended degree" of $\gcd(A(z), B(z))$ to highlight the fact that the polynomials may have common ∞ roots. (And therefore the degree in the usual sense may be lower, see also remarks at the end of Appendix A.)*

The GCD itself can be retrieved from the left or right kernel of the Sylvester matrix $\mathcal{S}_D(A,B)$, as summarized in the following propositions (which can be viewed as extensions of Theorem 5). In what follows, we assume that the GCD has (extended) degree K and note $H(z) = \gcd(A(z), B(z)) \in \mathbb{C}_{\leq K}[z]$. Moreover, we define

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

the corresponding quotient polynomials. We begin with the result on the right kernel of Sylverster subresultant matrices.

Proposition 2 (Right kernel, see e.g. [32, Lemma 4.6])**.** *The rank of the Sylvester subresultant matrix $\mathcal{S}_K(A,B)$ is equal to $2(L-K+1)-1$ (*i.e. *it has rank defect equal to 1). Moreover, for the (unique up to scalar factor) nonzero vector in the right kernel*

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

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

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

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

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

Proposition 3 (Left kernel)**.** *Let* $D \leq 1$ *(*i.e. *$\mathcal{S}_D(A, B)$ is fat with $2L - D + 1$ rows). Then the rank of $\mathcal{S}_D(A, B)$ is equal to*

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

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

$$\mathbf{h}^\top \underbrace{\begin{bmatrix} u[0] & u[1] & \cdots & u[2L-D-K] \\ u[1] & u[2] & \cdots & u[2L-D-K+1] \\ \vdots & \vdots & & \vdots \\ u[K] & u[K+1] & \cdots & u[2L-D] \end{bmatrix}}_{\mathcal{H}_{K+1}(\mathbf{u})} = 0,$$

i.e. $\mathbf{h}$ *is in the (left) kernel of the Hankel matrix built from* $\mathbf{u}$*.*

The next section exploits these properties of the kernel of Sylvester matrices to obtain algebraic reconstruction techniques for the PAF problem.

4.3. Algebraic algorithms for PAF

In this section, we propose two algorithms for solving PAF using the results of Section 4.2. The intuition behind the algorithms is that generic polynomials (informally speaking, with probability 1 if drawn from an absolutely continuous probability distribution) $X_1(z)$ and $X_2(z)$ are coprime, and therefore $\gcd(\Gamma_{11}(z), \Gamma_{12}(z)) = X_1(z)$ and $\gcd(\Gamma_{21}(z), \Gamma_{22}(z)) = X_2(z)$.

As a consequence, we now assume without loss of generality that the problem PAF (or equivalently, PPR or BPR) admits a unique solution up to trivial ambiguities. In all the algorithms, we use the singular value decomposition (SVD) as to find the approximate kernels of the matrix. Thus the proposed reconstuction methods may appear as suboptimal since the Sylvester structure is not taken account when computing the (low-rank) kernels. This limitation could be overcome with structured low-rank approximations [35], to be specifically tailored for the PAF problem. Such a study would fall outside the scope of the present work. Still, as demonstrated by the numerical experiments presented in Section 6, the SVD already provides excellent reconstruction performance in many scenarios, while maintaining a reasonable computational burden.

4.3.1. Right kernel Sylvester

The first algorithm is based on the properties of the right kernel of Sylvester matrices described in Proposition 2. It uses the fact that $X_1(z)$ and $X_2(z)$ are (without noise) quotient polynomials of

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

Note that $X_1(z)$ and $X_2(z)$ are also quotient polynomials of $\Gamma_{12}(z) = X_1(z)\widetilde{X}_2(z)$ and $\Gamma_{22}(z) = X_1(z)\widetilde{X}_2(z)$, which adds some freedom in the choice of measurement polynomials. For the sake of simplicity, we will work with estimated polynomials $\Gamma_{11}(z)$ and $\Gamma_{21}(z)$ in the following.

The complete right kernel Sylvester approach for solving PAF is summarized in Algorithm 1. It estimates the (one-dimensional) right kernel by computing the last nontrivial singular value of the Sylvester matrix

Algorithm 1: Sylvester right kernel for PAF

Input: Matrix polynomial $\mathbf{\Gamma}(z) \in \mathbb{C}^{2\times 2}_{\leq 2(N-1)}$.
Build the matrix $\mathbf{S} = \mathcal{S}_{N-1}(\Gamma_{11}, \Gamma_{21}) \in \mathbb{C}^{(3N-2)\times 2N}$;
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 $\mathbf{v}_1 = c\widehat{\mathbf{x}}_1$ and $\mathbf{v}_2 = c\widehat{\mathbf{x}}_2$ with $c \in \mathbb{C}$;
Determine $|c|$ by proper norm scaling as

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

Set $\widehat{\mathbf{x}}_1 = \mathbf{v}_1/|c|$ and $\widehat{\mathbf{x}}_2 = \mathbf{v}_2/|c|$;
Result: Estimates $\widehat{\mathbf{x}}_1$ and $\widehat{\mathbf{x}}_2$

$\mathcal{S}_{N-1}(\Gamma_{11}, \Gamma_{21})$. According to Proposition 2, this directly gives the vectors of coefficients $\mathbf{x}_1, \mathbf{x}_2$ of the polynomials $X_1(z)$ and $X_2(z)$ (or simply, the columns of the matrix $\mathbf{X}$ in PPR and BPR) up to one complex multiplicative constant. This constant is then computed (up to one unit-modulus factor due to the trivial rotation ambiguity) by scaling the 2-norm of $\mathbf{x}_1$ and $\mathbf{x}_2$ thanks to the value at the origin of estimated autocovariance functions $\gamma_{11}[0]$ and $\gamma_{22}[0]$.

One of the key advantages of this algorithm lies in its simplicity. Indeed, it only requires a single SVD of a $(3N-2) \times 2N$ matrix and thus has computational complexity $\mathcal{O}(N^3)$ (with a naive implementation by computing the full SVD). Potentially, iterative algorithms for truncated SVD can bring this complexity down to $\mathcal{O}(N^2)$, or even lower (by using techniques based of the FFT). For the sake of simplicity, we only considered the naive SVD implementation in our experiments.

4.3.2. Left kernel Sylvester

The second algorithm exploits the properties of the left kernel of extended (fat) Sylvester matrices (*i.e.* $\mathcal{S}_D$ for $D \leq 1$) detailed in Proposition 3. For simplicity and to reduce the size of the involved matrices we set $D = 1$ in what follows. Nonetheless, the proposed approach can be adapted to any value of $D \leq 1$ if needed.

Algorithm 2 summarizes the complete procedure. Compared to the right kernel Sylvester approach, the polynomials $X_1(z)$ and $X_2(z)$ are obtained by two separate GCD computations, that is

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

where c_1, c_2 are complex constants. The computation of each GCD requires three steps: a first SVD to determine the $N-1$ last left singular vectors of $\mathcal{S}_1$; the construction of a fat, horizontally stacked Hankel matrix $\mathbf{H}$ with N rows from these $N-1$ singular vectors; a second SVD to obtain the N coefficients of the GCD as the last left singular vector of $\mathbf{H}$. We refer the reader to [32] for further details on this procedure. Once GCDs have been obtained, determination of constants c_1 and c_2 (up to a common global phase factor) is carried out by properly scaling the norms of estimated coefficients $\mathbf{x}_1$ and $\mathbf{x}_2$ (using $\gamma_{11}[0]$ and $\gamma_{22}[0]$) and adjusting the phase factor $\arg c_1\overline{c_2}$ thanks to the value at origin of the estimated cross-covariance function $\gamma_{12}[0]$.

Importantly, the complexity of the left kernel Sylvester method described in Algorithm 2 is higher for two main reasons. First, as explained above, it requires the computations of 2 SVD for each of the two GCD determinations. Moreover, while the first SVD has a cost of $\mathcal{O}(N^3)$, the second SVD is performed on a large fat stacked Hankel matrix $\mathbf{H}$, with complexity $\mathcal{O}(N^4)$ for a naive implementation. This can potentially reduced by computing the SVD of $\mathbf{H}\mathbf{H}^{\mathsf{H}}$ instead, eventhough we do not consider such refinement here.

Algorithm 2: Sylvester left kernel for PAF

Input: Matrix polynomial $\mathbf{\Gamma}(z) \in \mathbb{C}^{2\times 2}_{\leq 2(N-1)}$

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

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

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

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

Stack the Hankel matrices with N rows in the following matrix

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_N(\mathbf{u}_{3N-2}) & \cdots & \mathbf{H}_N(\mathbf{u}_{4N-4}) \end{bmatrix} \in \mathbb{C}^{N\times(N-1)(3N-3)}$$

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

end

Determine constants c_1, c_2 as

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

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

Result: Estimates $\widehat{\mathbf{x}}_1$ and $\widehat{\mathbf{x}}_2$

5. Solving PPR: iterative algorithms

We now address the design of iterative algorithms to solve the noisy PPR problem. Section 5.1 and Section 5.2 exploit the PPR-1D representation of the original problem to provide a semidefinite programming (SDP) relaxation and Wirtinger flow algorithm, respectively.

5.1. SDP relaxation

Semidefinite programming (SDP) approaches for phase retrieval have been increasingly popular for over a decade [12, 11]. In the classical 1D phase retrieval case, SDP approaches exploit that eventhough measurements are quadratic in the unknown signal $\mathbf{x} \in \mathbb{C}^N$, they are linear in the rank-one matrix $\mathbf{x}\mathbf{x}^{\mathsf{H}}$. For PPR, the 1D equivalent representation PPR-1D enables to formulate a SDP relaxation of the original problem, by observing that

$$|\mathbf{c}_{m,p}^{\mathsf{H}}\boldsymbol{\xi}|^2 = \operatorname{Tr} \mathbf{c}_{m,p}\mathbf{c}_{m,p}^{\mathsf{H}}\boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{H}} \triangleq \operatorname{Tr} \mathbf{C}_{m,p}\mathbf{\Xi} \tag{38}$$

i.e. , noiseless measurements can be rewritten as a linear functions of the lifted positive semidefinite rank-one matrix $\mathbf{\Xi} \triangleq \boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{H}} \in \mathbb{C}^{2N\times 2N}$. Following the classical PhaseLift methodology [11, 12], the original nonconvex PPR problem can be relaxed into a SDP convex program as

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\sum_{m=0}^{M-1}\sum_{p=0}^{P-1}(y_{m,p} - \operatorname{Tr}\mathbf{C}_{m,p}\mathbf{\Xi})^2 + \lambda\|\mathbf{\Xi}\|_\star \\ \text{subject to} \quad & \mathbf{\Xi} \succeq 0 \end{aligned} \tag{39}$$

where $\lambda \geq 0$ is an hyperparameter that allows to control the trade-off between the likelihood of observations and the nuclear norm regularization $\|\cdot\|_\star$. Note that since $\mathbf{\Xi}$ is constrained to be positive semidefinite, the nuclear norm regularization is equivalent to the trace-norm regularization used in [12] since $\|\mathbf{\Xi}\|_\star = \operatorname{Tr}\mathbf{\Xi}$ in this case.

The SDP program (39) takes a standard form: thus it can be solved in many ways, including interior point methods [36], first-order methods [37] or using disciplined convex programming solvers such as `CVXPY`[2].

[2] https://www.cvxpy.org/

For completeness, we provide below an explicit algorithm to solve (39) using a proximal gradient approach [38, Chapter 10]. It closely follows the approach described in [12, 39].

The objective function in (39) can be rewritten as the sum $f(\boldsymbol{\Xi}) + g(\boldsymbol{\Xi})$, where

$$f(\boldsymbol{\Xi}) = \frac{1}{2}\sum_{m=0}^{M-1}\sum_{p=0}^{P-1}\left(y_{m,p} - \operatorname{Tr}\mathbf{C}_{m,p}\boldsymbol{\Xi}\right)^2, \quad g(\boldsymbol{\Xi}) = \lambda\|\boldsymbol{\Xi}\|_\star + \iota_{\succeq 0}(\boldsymbol{\Xi}) \tag{40}$$

where $\iota_{\succeq 0}(\cdot)$ denote the indicator function on the positive semidefinite cone. This ensures the formal equivalence between (39) and the unconstrained minimization problem

$$\min_{\boldsymbol{\Xi}\in\mathbb{C}^{2N\times 2N}} f(\boldsymbol{\Xi}) + g(\boldsymbol{\Xi})\,. \tag{41}$$

The convex optimization problem (41) can be efficiently solved by proximal gradient methods, which take advantage of the splitting between f and g of the objective function. More precisely, we use the fast proximal gradient method which consist, at iteration k:

$$\boldsymbol{\Xi}^{(k+1)} = \underset{t_k g}{\mathbf{prox}}\left(\boldsymbol{\Psi}^{(k)} - t_k\nabla f(\boldsymbol{\Psi}^{(k)})\right) \tag{42}$$

$$\eta_{k+1} = \frac{1+\sqrt{1+4\eta_k^2}}{2} \tag{43}$$

$$\boldsymbol{\Psi}^{(k+1)} = \boldsymbol{\Xi}^{(k+1)} + \left(\frac{\eta_k - 1}{\eta_{k+1}}\right)\left(\boldsymbol{\Xi}^{(k+1)} - \boldsymbol{\Xi}^{(k)}\right) \tag{44}$$

where t_k is a step-size which is chosen such that the proximal gradient step (42) obey some sufficient decrease condition; see e.g. [38, p. 271] for details. Our choice for the function g in (41) enables a simple expression for the associated proximal operator (see [39]):

$$\begin{aligned}\underset{\tau g}{\mathbf{prox}}(\mathbf{X}) &\triangleq \min_{\mathbf{Z}\succeq 0} \tau\lambda\|\mathbf{Z}\|_\star + \|\mathbf{Z}-\mathbf{X}\|_2^2 \\ &= \mathbf{Ushrink}(\Sigma, \tau\lambda)\mathbf{U}^{\mathsf{H}}\end{aligned} \tag{45}$$

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

Choice of regularization parameter λ. In this work, we fix the value of the regularization parameter to $\lambda = 1/\mathrm{SNR}$: we found empirically that this choice provides good results in most scenarios, as it provides a reasonable tradeoff between likelihood of observations and the nuclear norm regularization in the objective function of (39).

Convergence. Obviously, as our algorithm is a convex SDP program, the precision towards the optimal cost value can become arbitrarily good as one increases the number of iterations. In practice, one needs to stop the algorithm when a prescribed tolerance ε is reached. To this aim we implemented stopping criteria that carefully monitor a normalized residual, see [39] for details. Moreover, it may happen that the estimated lifted matrix $\hat{\boldsymbol{\Xi}}$ generated by the sequence of $\boldsymbol{\Xi}^{(k)}$ is not rank one: in this case, one first computes the rank-one approximation of $\hat{\boldsymbol{\Xi}}$ (e.g. using SVD) to obtain the estimated signal $\hat{\boldsymbol{\xi}}$.

Complexity. The computational cost of the proposed algorithm concentrates on the proximal gradient step (42), where the evaluation of the proximal operator and the computation ∇f share the computational burden. More precisely, the eigenvalue decomposition of a $2N \times 2N$ matrix together with the shrink operator leads to $\mathcal{O}(N^3)$ calculations. The computation of the gradient leads to MP trace evaluations of order $\mathcal{O}(N^2)$ flops, meaning that the number of flops per iteration is of order $\mathcal{O}(MPN^2 + N^3)$.

The full procedure is summarized in Algorithm 3.

Algorithm 3: 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 $\boldsymbol{\Xi}^{(0)}$;

$\boldsymbol{\Psi}^{(0)} \leftarrow \boldsymbol{\Xi}^{(0)}$;

$k \leftarrow 0$;

while *stopping criterion is not satisfied* **do**

$\boldsymbol{\Xi}^{(k+1)} = \mathbf{prox}_{t_k g}\left(\boldsymbol{\Psi}^{(k)} - t_k \nabla f(\boldsymbol{\Psi}^{(k)})\right)$ where the proximal operator is given by (45);

$\eta_{k+1} = \frac{1+\sqrt{1+4\eta_k^2}}{2}$;

$\boldsymbol{\Psi}^{(k+1)} = \boldsymbol{\Xi}^{(k+1)} + \left(\frac{\eta_k - 1}{\eta_{k+1}}\right)\left(\boldsymbol{\Xi}^{(k+1)} - \boldsymbol{\Xi}^{(k)}\right)$;

$k \leftarrow k+1$;

end

$\hat{\boldsymbol{\xi}} \leftarrow \mathrm{rank1}\left(\boldsymbol{\Xi}^{(k)}\right)$;

Result: estimate $\hat{\boldsymbol{\xi}}$

5.2. Wirtinger flow for PPR

Exploiting further the 1D equivalent representation PPR-1D of the PPR problem, another approach consists in minimizing directly the following nonconvex quadratic objective

$$\min_{\boldsymbol{\xi}\in\mathbb{C}^{2N}} F(\boldsymbol{\xi}) \triangleq \frac{1}{2}\|\mathbf{y} - |\mathbf{C}\boldsymbol{\xi}|^2\|_2^2 \tag{46}$$

where $\mathbf{y} \in \mathbb{R}^{MP}$ gather PPR measurements and where the rows of $\mathbf{C} \in \mathbb{C}^{MP\times 2N}$ are given by $\mathbf{c}_{m,p}^{\mathsf{H}}$, see Section 2.4. Provided that one can find a initial point $\boldsymbol{\xi}^{(0)}$ close enough from the global minimizer of (46), a simple strategy based on gradient descent can be used to solve PPR. However, such an approach requires special care since the optimization variable $\boldsymbol{\xi}$ is complex-valued. In fact, the objective function in (46) is real-valued, and thus it is not differentiable with respect to complex analysis. Instead, one needs to resort to the so-called $\mathbb{CR}$ or *Wirtinger*-calculus [40] to provide a meaningful extension of gradient-descent-type algorithms to the complex case. This is precisely the approach proposed in [41] to solve standard phase retrieval, where the complex gradient descent is called *Wirtinger flow*.

Leveraging the original Wirtinger flow approach, we propose below a complex-gradient descent algorithm which solves the nonconvex problem (46). Compared to the original paper [41], we incorporate optimal step size selection [42] together with a proposed acceleration scheme [43]. We further propose an efficient strategy for initialization based on the algebraic methods for PPR described in Section 4. The superiority of these initializations over standard ones (e.g. spectral initialization as in [41]) will be demonstrated in Section 6.2.

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

$$\beta_k = \frac{k+1}{k+3} \tag{47}$$

$$\boldsymbol{\psi}^{(k)} = \boldsymbol{\xi}^{(k)} + \beta_k\left(\boldsymbol{\xi}^{(k)} - \boldsymbol{\xi}^{(k-1)}\right) \tag{48}$$

$$\boldsymbol{\xi}^{(k+1)} = \boldsymbol{\psi}^{(k)} - \mu_k \nabla F\left(\boldsymbol{\psi}^{(k)}\right) \tag{49}$$

where β_k is a sequence of accelerated parameters and μ_k is a carefully chosen stepsize, see further below. Compared to the standard WF algorithm, PPR-WF takes advantage of the acceleration procedure first proposed in [43] in the context of ptychographic phase retrieval (but using a magnitude loss function instead

of a square magnitude loss function as used here). Note that the complex gradient of F can be computed explicitly as

$$\nabla F(\boldsymbol{\psi}) = \mathbf{C}^{\mathsf{H}} \left(|\mathbf{C}\boldsymbol{\psi}|^2 - \mathbf{y} \right). \tag{50}$$

Optimal step-size selection. We combine acceleration for WF with the optimal step-size selection proposed in [42] for the standard WF algorithm. For completeness, we reproduce here the main ingredients underpinning optimal step size selection in (49) and refer the reader to [42] for further details. At iteration k, the optimal stepsize μ_k is defined by line search, *i.e.*

$$\mu_k \triangleq \arg\min_{\mu} F\left(\boldsymbol{\xi}^{(k+1)}\right) \triangleq F\left(\boldsymbol{\psi}^{(k)} - \mu \nabla F\left(\boldsymbol{\psi}^{(k)}\right)\right) \tag{51}$$

The authors in [42] showed that the 1D optimization problem (51) boils down to finding the roots of a univariate cubic polynomial with real coefficients, the latter being completely determined by the knowledge of $\boldsymbol{\psi}^{(k)}$, $\nabla F\left(\boldsymbol{\psi}^{(k)}\right)$ and $\mathbf{y}$, see [42, Eq. (17)]. Roots can be determined in closed-form, and two cases can occur: *(a)* there is only one real root, and thus it gives the optimal step-size μ_k; *(b)* there are three real roots, and in this case μ_k is set to the real root associated to the minimum objective value. Note that optimal selection for WF is somewhat inexpensive, with computational cost dominated by the calculation of the cubic polynomial coefficients scaling as $\mathcal{O}(MP)$.

Initialization. Since PPR-WF attempts a minimizing a nonconvex quadratic objective (46), the choice of initial points $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\xi}^{(1)}$ is crucial to hope that PPR-WF will be able to recover a global minimizer of the objective function. For simplicity, in this work we set $\boldsymbol{\xi}^{(1)} = \boldsymbol{\xi}^{(0)}$, so that we only discuss the selection of $\boldsymbol{\xi}^{(0)}$. In this work we consider four different initialization strategies for PPR-WF:

- *spectral initialization* [41]: this standard approach consists in computing the eigenvector $\mathbf{v}$ corresponding to the largest eigenvalue of the matrix

$$\mathbf{Y} \triangleq \frac{1}{MP} \sum_{r=0}^{MP-1} y_r \mathbf{c}_r \mathbf{c}_r^{\mathsf{H}} \tag{52}$$

  and to rescale it properly to set

$$\boldsymbol{\xi}^{(0)} = \frac{\mathbf{v}}{\lambda}, \quad \lambda = \left(N \frac{\sum_{r=0}^{MP-1} y_r}{\sum_{r=0}^{MP-1} \|\mathbf{c}_r\|^2} \right)^{1/2} \tag{53}$$

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

$$\boldsymbol{\xi}^{(0)} = \mathbf{C}^{\dagger} \tilde{\mathbf{y}}, \quad \tilde{\mathbf{y}} \triangleq \mathbf{y} \odot \exp(\jmath \boldsymbol{\phi}) \tag{54}$$

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

- *left kernel Sylvester initialization*: we simply set $\boldsymbol{\xi}^{(0)}$ as the result of the left-kernel Sylvester method.

- *right kernel Sylvester initialization* we simply set $\boldsymbol{\xi}^{(0)}$ as the result of the right-kernel Sylvester method.

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

Complexity. The computational cost per iteration of PPR-WF is dominated by the evaluation of the complex gradient (50), which scales as $\mathcal{O}(MPN)$. Note that the optimal step-size selection procedure scales as $\mathcal{O}(MP)$, so meaning that the whole cost of PPR-WF remains $\mathcal{O}(MPN)$ per iteration. Algorithm 4 summarizes the proposed PPR-WF algorithm.

Algorithm 4: Wirtinger Flow for PPR: PPR-WF

Input: measurements $\mathbf{y} \in \mathbb{R}^{MP}$, measurment matrix $\mathbf{C} \in \mathbb{C}^{MP \times 2N}$, tolerance ε
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**
  $\beta_k \leftarrow \frac{k+1}{k+3}$;
  $\boldsymbol{\psi}^{(k)} \leftarrow \boldsymbol{\xi}^{(k)} + \beta_k \left(\boldsymbol{\xi}^{(k)} - \boldsymbol{\xi}^{(k-1)}\right)$;
  compute optimal step-size μ_k (51);
  $\boldsymbol{\xi}^{(k+1)} \leftarrow \boldsymbol{\psi}^{(k)} - \mu_k \nabla F\left(\boldsymbol{\psi}^{(k)}\right)$;
  $k \leftarrow k + 1$;
end
$\hat{\boldsymbol{\xi}} \leftarrow \boldsymbol{\xi}^{(k)}$;
Result: estimate $\hat{\boldsymbol{\xi}}$

6. Numerical experiments

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

Just like in standard phase retrieval, the global phase ambiguity in PPR requires to properly realign any estimated signal $\hat{\mathbf{X}}$ with the ground truth $\mathbf{X}$ in order to provide a meaningful MSE value. Thus, the MSE is defined as

$$\text{MSE} \triangleq \mathbf{E}\|\tilde{\hat{\mathbf{X}}} - \mathbf{X}\|_F^2 \quad \text{where } \tilde{\hat{\mathbf{X}}} \triangleq e^{\jmath\Phi_0}\hat{\mathbf{X}} \text{ with } \Phi_0 \triangleq \underset{\phi\in[0,2\pi)}{\arg\min} \|e^{\jmath\phi}\hat{\mathbf{X}} - \mathbf{X}\|_F^2 \,. \tag{55}$$

Note that in practice, the minimization involved in the realignment procedure can simply be performed by evaluating the complex phase of the standard inner product between the vectors $\hat{\boldsymbol{\xi}}$ and $\boldsymbol{\xi}$ obtained from matrices $\hat{\mathbf{X}}$ and $\mathbf{X}$, respectively.

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

6.1. Reconstruction of bivariate pulse

As a first experiment, we consider the reconstruction of a bivariate pulse from noiseless PPR measurements. The signal to be recovered defines a typical complex-valued bivariate analytic signal associated to the bivariate electromagnetic field to be estimated in ultra-short electromagnetic pulses experiments, see e.g. [21, 44]. It is defined for $N = 64$ points and we consider the simple noise-free measurement scheme (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 several methods introduced in Section 4 and Section 5 to properly recover the bivariate signal of interest. Note that for Wirtinger Flow, we consider two initialization strategies, one using spectral initialization and the other one based on the solution given by the right kernel Sylvester approach.

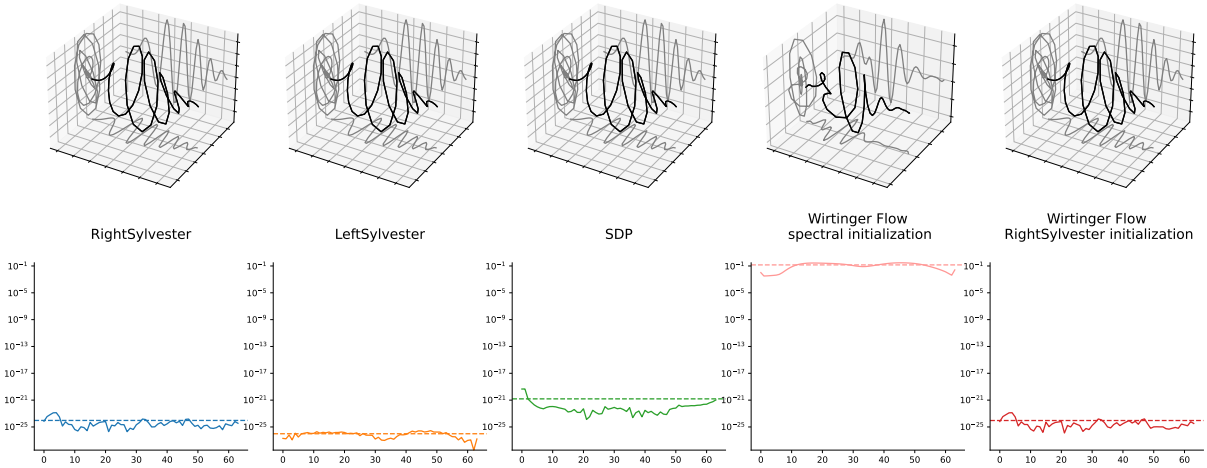

Figure 6: Reconstruction of a bivariate pulse ($N = 64$) from noiseless PPR measurements ($M = 2N - 1, P = 4$) using the different methods described in this paper. The reconstructed signal trace and squared error per time index n are shown for each approach.

Figure 6 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 $\mathbf{x}$ beforehand. Excepted Wirtinger Flow with spectral initialization, all methods successfully recover the original bivariate signal, where successful recovery in the noiseless context is decided whenever $\|\hat{\mathbf{X}} - \mathbf{X}\|_2^2 < 10^{-20}$. Left and right-kernel Sylvester and Wirtinger Flow with right-Sylvester initialization 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 MSE 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 6. To assess the role played by initial points in PPR-WF, we carefully benchmark the four initialization methods described in Section 5.2, that is spectral initialization, random phase initialization, 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 (30) were considered for the simple measurement scheme (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 four aforementioned initialization procedures.

Figure 7 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

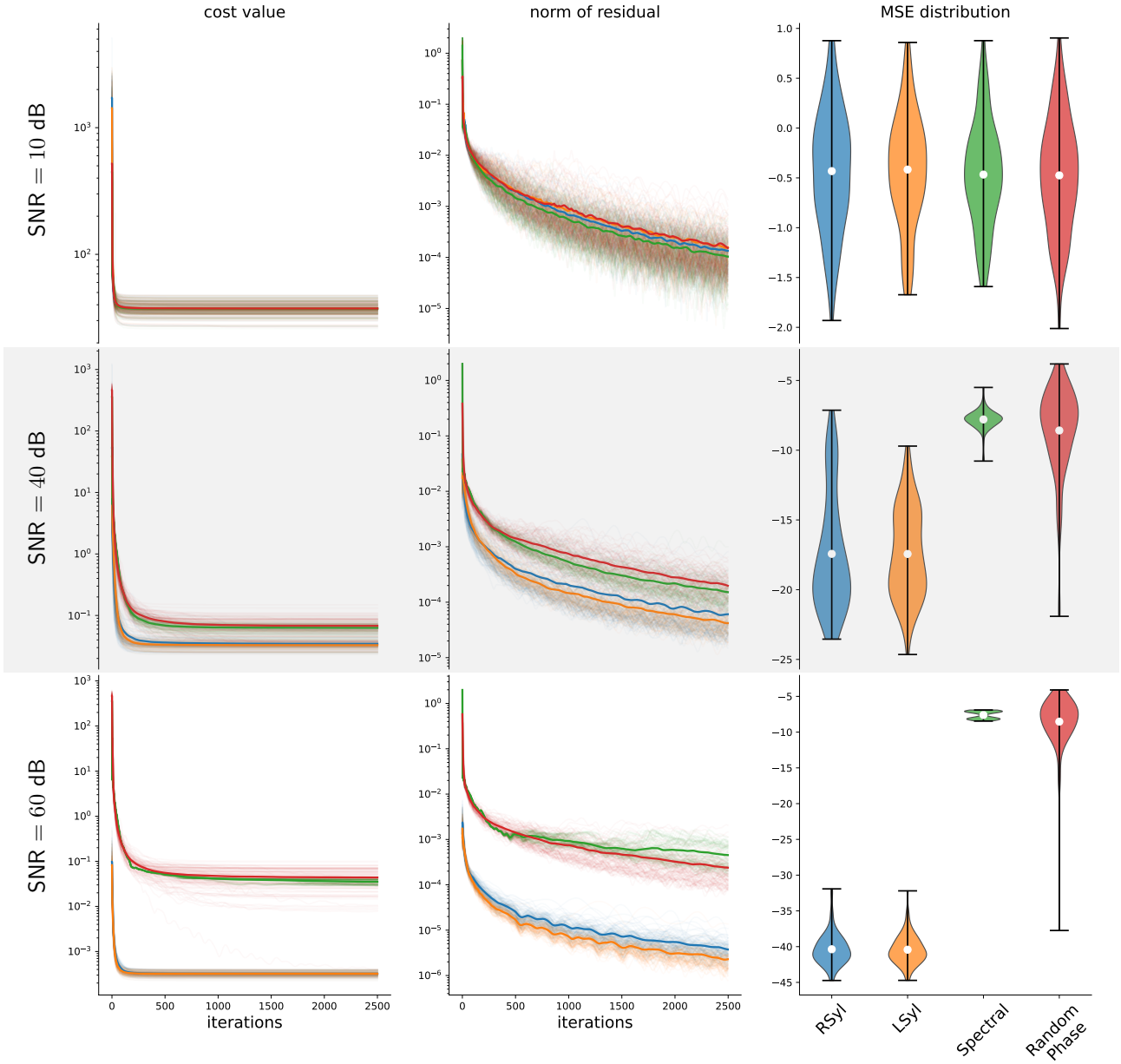

Figure 7: Comparison of initialization strategies for PPR-WF for the recovery of an arbitrary random bivariate signal of length $N = 32$ with $M = 2N - 1$ and $P = 4$ noisy measurements. We benchmark spectral initialization, random phase initialization, left and right-kernel Sylvester initialization strategies in terms of cost function evolution, normed residual decrease and MSE 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 MSE distribution associated to each initialization strategy. White dots indicate mean-values and horizontal bars extreme values for each MSE distribution.

ensure fair comparisons. We also plot the empirical distribution of MSE values for each initialization for further comparison of the quality of the reconstructed signal (recall that MSE 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 provide similar results in terms of cost value decrease, residual evolution and MSE distribution. For SNR = 40 dB, one starts to observe significant differences between Sylvester-based approaches and spectral/random phase initializations. On average, Sylvester-based initial points provides smaller optimal values, faster decrease of the residual and better reconstruction results in terms of MSE. This behavior is accentuated for SNR = 60 dB, where spectral and random phase initialization are unable to ensure convergence of PPR-WF to the global optimum. This agrees with the observations made in Figure 6 in the noiseless case for spectral initialization.

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

6.3. Recovery performance with noisy measurements

We now investigate the recovery performances of the different proposed algorithms for PPR when dealing with noisy measurements. We consider an additive white Gaussian noise model (30) for which the SNR is defined in (31). We generated a ground truth signal $\mathbf{X} \in \mathbb{C}^{N\times 2}$ with i.i.d. Gaussian entries of length $N = 32$ such that $\|\mathbf{X}\|_F = 1$ and selected the simple, $M = 2N-1, P = 4$ measurement scheme (5). For a given SNR value, the MSE associated with each one of the proposed methods to solve PPR was obtained by averaging of 100 independent reconstructions. Note that we used our recommended right-kernel Sylvester initialization with PPR-WF, as explained in Section 6.2.

Figure 8 displays the evolution of MSE for values of SNR ranging from 0 dB to 80 dB. As expected, the MSE decreases as the SNR increases, independently from the considered method. Overall, algorithmic methods (PPR-WF and SDP) outperform algebraic (left and right-kernel Sylvester) ones in terms of MSE values. More precisely, algebraic methods are not informative in the "low-SNR" regime (SNR $\leq$ 30 dB) as they provide (relative) MSE values above 0 dB, meaning that they do not provide a better reconstruction than a simple i.i.d. random guess scaled to the ground truth norm. Furthermore we observe that SDP is more robust to noise than PPR-WF. This agrees with the fact that SDP methods are known to be robust to noise in general. Remarkably, the high-SNR regime ($\geq$ 60 dB) highlights several distinctive behaviors. First, we observe that beyond SNR = 40 dB, PPR-WF outperforms all other methods, including SDP, by a few dB up to about 10 dB of relative MSE in the asymptotic regime. Second, SDP do not longer outperforms left-kernel Sylvester, and only improves from the right-kernel Sylvester approach by a small margin. This shows that, in this high-SNR regime, the computational burden associated to the SDP approach becomes prohibitive as 1) it provides no clear advantage over computationally cheaper algebraic methods and 2) it clearly underperforms PPR-WF.

For completeness, we also provide the Cramèr-Rao lower bound (CRLB) for the PRR measurement model (30) to characterize a lower bound on the MSE of any unbiased estimator of the ground truth signal. An analytical derivation of the resulting CRLB is given in Appendix E. Figure 8 displays the CRLB on top of MSE values obtained for each reconstruction method. We observe that the CRLB is not informative below SNR $\leq$ 20 dB as all methods provide smaller MSE values – it simply means that the CRLB is particularly pessimistic in this regime. On the contrary, the CRLB provides a meaningful lower bound in the high-SNR regime. Importantly, it demonstrates that PPR-WF is an asymptotically optimal reconstruction method for PPR since it attains the CRLB for SNR $\geq$ 60 dB.

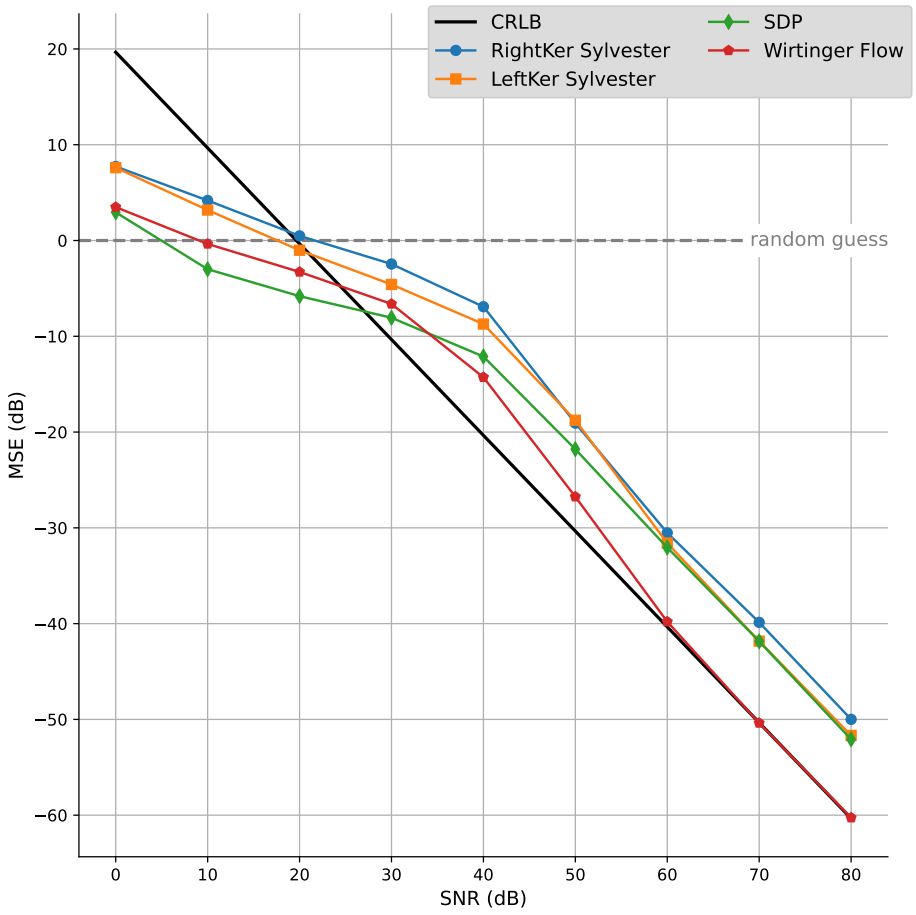

Figure 8: Evolution of the MSE with the SNR for the four PPR reconstruction methods proposed in this paper. Ground truth is randomly generated bivariate signal with $N = 32$. Simple measurement scheme for $M = 2N - 1$ and $P = 4$ was used.Thick black line indicate the corresponding Cramèr-Rao lower bound analytically derived in Appendix E.

6.4. Influence of number of measurements

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

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

- $M = 2N-1, P = 12$ case: we use the correspondence between the 2-sphere and $\mathbb{C}^2$ to take advantage of optimal spherical tesselations such as HEALPix [45]. In physical terms, it can interpreted as finding one of the many possible Jones vector $\mathbf{b}_p$ corresponding to the Stokes parameters defining

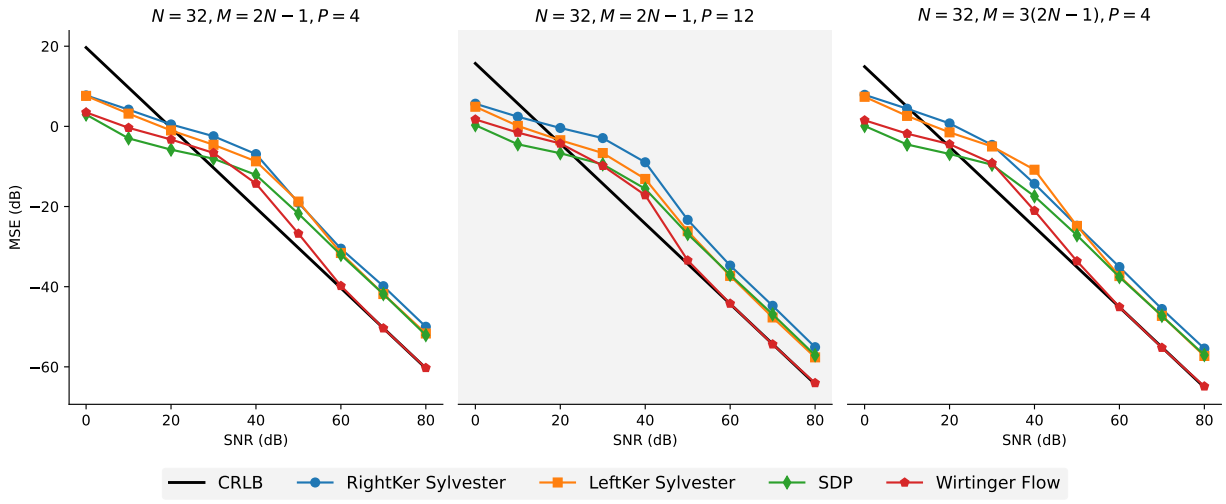

Figure 9: Comparison of the evolution of the MSE with respect to SNR for three measurements scheme $M = 2N - 1, P = 4$ (left), $M = 2N - 1, P = 12$ (center) and $M = 3(2N - 1), P = 4$ (right). Experiments follow the same protocol as described in Section 6.3.

the rank-one matrix $\mathbf{b}_p\mathbf{b}_p^{\mathsf{H}}$. 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, we define the projection vector $\mathbf{b}_p$ as:

$$\mathbf{b}_p \triangleq \frac{1}{\sqrt{2}\sqrt{1+s_p^z}} \begin{bmatrix} \jmath s_p^x \\ s_p^y + (1+s_p^z)\jmath \end{bmatrix} \quad \text{if } s_p^z \neq -1, \quad \mathbf{b}_p \triangleq \begin{bmatrix} \jmath \\ 0 \end{bmatrix} \quad \text{if } s_p^z = -1\,. \tag{56}$$

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

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

Figure 9 depicts MSE as a function of SNR for the two measurement setups described above, where results from the experiment in Section 6.3 have been reproduced for better comparison. As expected, increasing the total number of measurements MP improves overall performance: this can be directly checked by remarking that the CRLB corresponding to $M = 2N - 1, P = 12$ and $M = 3(2N - 1), P = 4$ cases is lower that of the $M = 2N - 1, P = 4$ setup presented in Figure 8. Moreover, the different proposed reconstructions method for PPR behave similarly with one another as in our description made in Section 6.3. In particular, we note that PPR-WF also attains the CRLB in these two new setups, proving again that it establishes a versatile approach to solve PPR.

Figure 10 provides a side-by-side comparison of these three measurement schemes for each reconstruction method. First, remark that $M = 2N-1, P = 12$ and $M = 3(2N-1), P = 4$ scheme have similar CRLB MSE bounds, with a slight advantage to the $M = 3(2N - 1), P = 4$ case which can be observed on the PPR-WF panel. Second, we note that for algorithmic approaches (SDP and PPR-WF), the difference concentrates in the mid-SNR regime, *i.e.* between 30 dB and 50 dB, where oversampling in the Fourier domain offers slightly MSE improvement over increasing the number of polarimetric projections. On the other hand, for algebraic approaches we observe that performing more polarimetric measurements usually improves the performance in the the low-SNR regime (SNR $\leq$ 30 dB), eventhough algebraic approaches do not perform well in this scenario. This performance improvement can be explained by the two-step nature of algebraic methods, which first need to reconstruct autocorrelation polynomials from polarimetric projections: in this case more polarimetric projections enable to reduce the reconstruction error in this first step.

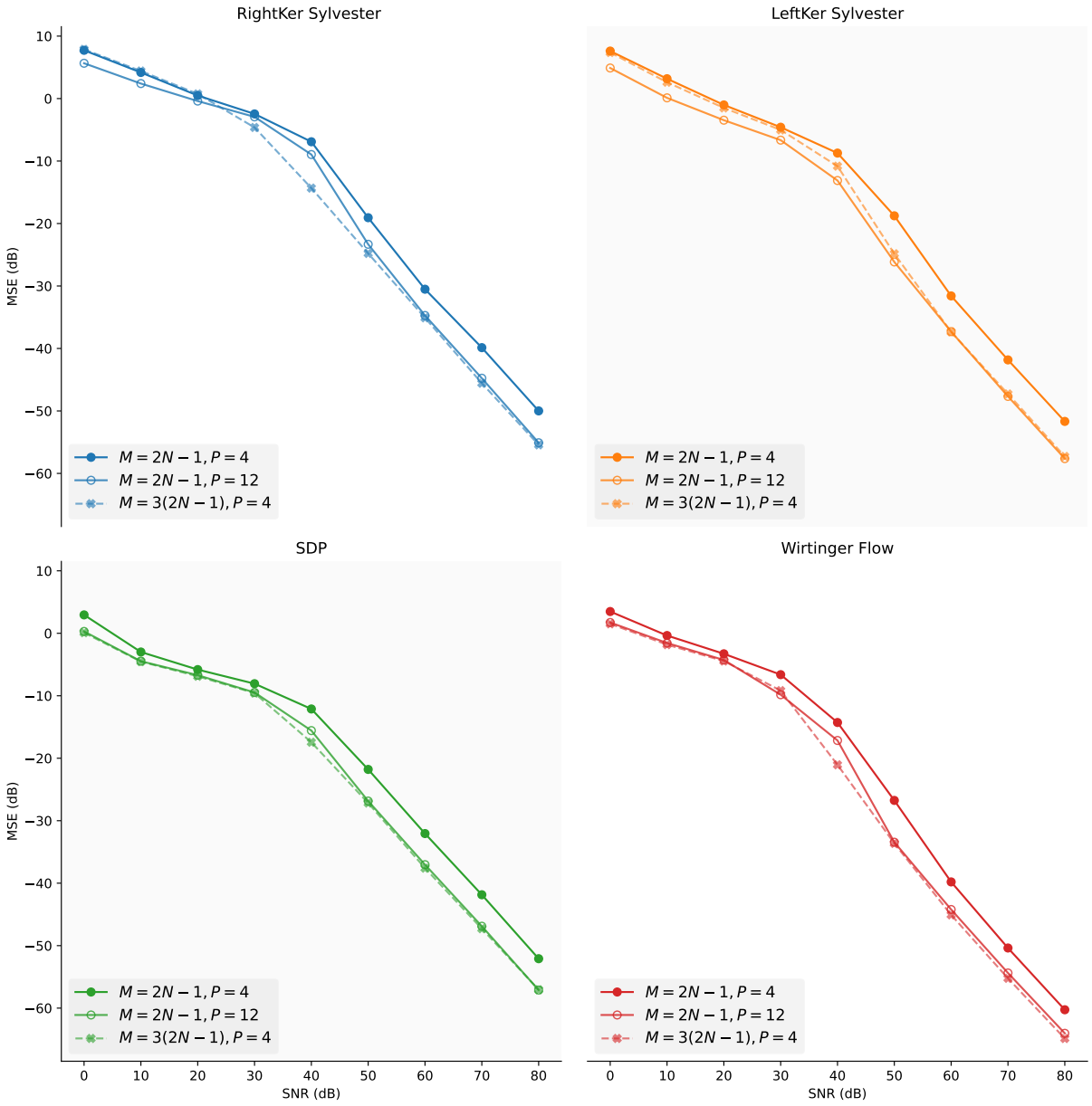

Figure 10: Side-by-side comparison of the behavior of each proposed reconstruction method for the three measurements scheme $M = 2N - 1, P = 4$, $M = 2N - 1, P = 12$ and $M = 3(2N - 1), P = 4$.

7. Conclusion

In this paper, we have introduced a new model for Fourier phase retrieval called PPR that takes advantage of polarization measurements in applications involving polarized light. Remarkably, we have shown that the problem of reconstructing a 1D bivariate signal from Fourier magnitude enjoys nice uniqueness properties which contrast with the non-uniqueness of standard 1D Fourier phase retrieval. The theoretical study of PPR has been made possible by carefully drawing equivalences with two related problems, namely BPR and PAF. Moreover, the derivation of uniqueness thanks to the polynomial factorization representation PAF motivated the development of two algebraic reconstructions methods for PPR using approximate greatest common divisor computations and Sylvester-like matrices. We also carefully adapted SDP and Wirtinger-Flow methods to solve the PPR problem. The extensive numerical experiments demonstrated the pros and cons of each approach. It also allowed us to establish a scalable, computationaly efficient and robust to noise reconstruction strategy that combines both algebraic (right-kernel Sylvester initialization) and algorithmic (Wirtinger-Flow iterates) approaches.

Our conviction is that the new model PPR opens promising new avenues in for exploiting polarization in Fourier phase retrieval problems. For instance, an important challenge to be addressed lies in improving the performance of algebraic methods at low SNR, e.g. with more robust estimation of the measurement polynomials or adding some prior information about the signal to be recovered (e.g. smoothness). These questions will be addressed in future work.

Appendix A. Polynomials and roots at infinity

Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ (or, in general, a field). The vector space $\mathbb{C}_{\leq D}[z]$ of polynomials of degree less than D in subsection 3.1 is, in fact, isomorphic to the vector space $\mathbb{C}^{D+1}$ via the following one-to-one map:

$$\mathbf{a} = \begin{bmatrix} a[0] & a[1] & \cdots & a[D] \end{bmatrix}^\top \mapsto A(z) = a[0] + za[1] + \cdots + z^D a[D].$$

In the literature (for example in algebraic geometry) a common way is to deal with bivariate homogeneous polynomials, that is polynomials of the form

$$y^D a[0] + y^{D-1} z a[1] + \cdots + z^D a[D],$$

which belong to the space $\mathbb{C}_{=D}[z, y]$. The roots of such polynomials belong to the projective space $\mathbb{P}^1$, which corresponds to the extended complex plane $\mathbb{C} \cup \{\infty\}$. However, for simplicity, we prefer to work with the polynomials in $\mathbb{F}_{\leq D}[z]$ instead, and we refer the reader to [46, Ch. 8] for details on homogeneous polynomials.

When working with the space $\mathbb{F}_{\leq D}[z]$, the multiplication operation becomes the map from $\mathbb{F}_{\leq D_1}[z] \times \mathbb{F}_{\leq D_2}[z]$ to $\mathbb{F}_{\leq (D_1+D_2)}[z]$:

$$(A(z), B(z)) \mapsto C(z) = A(z)B(z),$$

which in coordinates (i.e., for the vectors of coefficients $\mathbf{a} \in \mathbb{F}^{D_1+1}$ and $\mathbf{b} \in \mathbb{F}^{D_2+1}$) can be expressed as

$$(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{c} = \mathbf{M}_{D_1}(\mathbf{b})\mathbf{a} = \mathbf{M}_{D_2}(\mathbf{a})\mathbf{b}, \tag{A.1}$$

where $\mathbf{M}_L(\mathbf{a})$ is the multiplication matrix defined for any vector of coefficients

$$\mathbf{a} = \begin{bmatrix} a[0] & a[1] & \cdots & a[D] \end{bmatrix},$$

as the following matrix $(D + L + 1) \times (L + 1)$

$$\mathbf{M}_L(\mathbf{a}) \triangleq \underbrace{\begin{bmatrix} a_0 & & \\ \vdots & \ddots & \\ a_D & & a_0 \\ & \ddots & \vdots \\ & & a_D \end{bmatrix}}_{L+1 \text{ columns}}. \tag{A.2}$$

Armed with the definition of the multiplication, we can now give a formal justification to (10). Let us formally define

$$(z-\infty) \triangleq 0 \cdot z + 1 \in \mathbb{C}_{\leq 1}[z].$$

Then, thanks to the fundamental theorem of algebra, any nonzero polynomial $A \in \mathbb{C}_{\leq D}[z]$ can be uniquely (up to permutation of roots) factorized as

$$A(z) = \lambda(z-\alpha_1)\cdots(z-\alpha_D),$$

where for any k, $\alpha_k \in \mathbb{C} \cup \{\infty\}$.

Finally, we remark on the notion of the greatest common divisor, which, for two nonzero polynomials $A_1, A_2 \in \mathbb{C}_{\leq D}[z]$ is a polynomial $H \in \mathbb{C}_{\leq D'}[z]$ with highest possible D', which is a divisor of both $A_1(z)$ and $A_2(z)$. The GCD is defined uniquely up to a multiplication by a scalar in $\mathbb{C} \setminus \{0\}$. The same notion can be defined for several polynomials, see [32, Section 2] for more details.

Appendix B. Relation between Fourier measurements and correlation polynomials

Proof of Lemma 1. The first part of lemma follows from the correspondence between multiplication of polynomials and discrete convolution of two vectors, see (A.1). Next, recall that the discrete Fourier transform of $\{\mathbf{x}[n]\}_{n=0,1,\ldots,N-1}$ is denoted by $\mathfrak{X}[m] = [\mathfrak{X}_1[m], \mathfrak{X}_2[m]]^\top$ for $m = 0, 1, \ldots, M-1$, see (2). Then the Fourier entries can be related to polynomials $X_1(z)$ and $X_2(z)$ as follows:

$$\mathfrak{X}_1[m] = X_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right), \quad \mathfrak{X}_2[m] = X_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right),$$

for any $m = 0, 1, \ldots, M-1$. Similarly, thanks to (12), their conjugates can be expressed through the conjugate reflection polynomials $\widetilde{X}_1(z)$ and $\widetilde{X}_2(z)$

$$\overline{\mathfrak{X}_1[m]} = \overline{X_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right)} = \sum_{n=0}^{N-1} \overline{x_1[n]} e^{2\pi\jmath \frac{nm}{M}} = e^{\jmath 2\pi \frac{m(N-1)}{M}} \widetilde{X}_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right),$$
$$\overline{\mathfrak{X}_2[m]} = \overline{X_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right)} = \sum_{n=0}^{N-1} \overline{x_2[n]} e^{2\pi\jmath \frac{nm}{M}} = e^{\jmath 2\pi \frac{m(N-1)}{M}} \widetilde{X}_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right).$$

As a result, thanks to (3), BPR measurements can be expressed in terms of measurement polynomials $\Gamma_{ij}(z)$ as follows:

$$\begin{aligned}\mathfrak{F}[m] &= \begin{bmatrix} |\mathfrak{X}_1[m]|^2 & \mathfrak{X}_1[m]\overline{\mathfrak{X}_2[m]} \\ \mathfrak{X}_2[m]\overline{\mathfrak{X}_1[m]} & |\mathfrak{X}_2[m]|^2 \end{bmatrix} \\ &= e^{\jmath 2\pi \frac{m(N-1)}{M}} \begin{bmatrix} X_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \widetilde{X}_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right) & X_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \widetilde{X}_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \\ X_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \widetilde{X}_1\left(e^{-\jmath 2\pi \frac{m}{M}}\right) & X_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \widetilde{X}_2\left(e^{-\jmath 2\pi \frac{m}{M}}\right) \end{bmatrix} = e^{\jmath 2\pi \frac{m(N-1)}{M}} \mathbf{\Gamma}(e^{-\jmath 2\pi \frac{m}{M}}),\end{aligned}$$

which completes the proof. □

Proof of Theorem 1. Here, we make use of the two one-to-one correspondences. Note that the mapping between $\mathbb{C}^{N+1}$ and $\mathbb{C}_{\leq N}$ is a linear one-to-one map (and is an isomorphism), see Appendix A. Hence, the signals X_1, X_2 can be uniquely recovered from the polynomials and vice versa.

Similarly, thanks to (14), the Fourier covariance measurements $\mathfrak{F}[m]$ are a linear transformation of the sequence

$$\{\mathbf{\Gamma}(e^{-\jmath 2\pi \frac{m}{M}})\}_{m=0,\ldots,M-1}$$

of evaluations of the matrix polynomial $\mathbf{\Gamma}(z)$ at a set of M distinct points $\{e^{-\jmath 2\pi \frac{m}{M}}\}_{m=0,\ldots,M-1}$ on the complex plane. If $M \geq 2N-1$ (the degree of the polynomials + 1), then it is known that the coefficients

of the polynomials can be uniquely recovered from the evaluations at M distinct points, and therefore the following map is an injection

$$\begin{aligned}\mathbb{C}^{2\times 2}_{\leq 2N-2} &\to (\mathbb{C}^{2\times 2})^M\\ \mathbf{\Gamma}(z) &\mapsto \{\mathfrak{F}[m]\}_{m=0,\ldots,M-1},\end{aligned}$$

which completes the proof.

□

Appendix C. Proof of Theorem 4

Suppose that $H(z) = \gcd(\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_2) = Q(z)\tilde{Q}(z)$ with D pair roots $(\delta_i, \overline{\delta_i}^{-1})$. We further assume that $(0, \infty)$ is not a pair root of $H(z)$. Let $X_1(z) = Q(z)R_1(z)$ and $X_2(z) = Q(z)R_2(z)$. By Lemma 2, the polynomials R_1 and R_2 can be determined up to one multiplicative constant by (17). Let us denote by α_{1i} (resp. α_{2i}) the $L - D - 1$ roots of $R_1(z)$ (resp. $R_2(z)$), such that

$$R_1(z) = \lambda_1 \prod_{i=1}^{N-D-1}(z - \alpha_{1i}), \qquad R_2(z) = \lambda_2 \prod_{i=1}^{N-D-1}(z - \alpha_{2i})$$

where $\lambda_1, \lambda_2 \in \mathbb{C}$ are constants, to be derived hereafter. The recovery of $Q(z)$ from $Q(z)\tilde{Q}(z)$ is identical to the univariate case, see e.g. [14, 15]. Denoting by $(\delta_i, \overline{\delta_i}^{-1})$ the pair roots of $Q(z)\tilde{Q}(z)$, $Q(z)$ can be written as

$$Q(z) = \prod_{i=1}^{D}(z - \beta_i), \quad \beta_i \in (\delta_i, \overline{\delta_i}^{-1}) \tag{C.1}$$

As explained in Theorem 3, the number of different solutions for $Q(z)$ dictates the number of solutions for the PAF problem. Thus, if polynomials $(X_1'(z), X_2'(z))$ are solutions to PAF then they can be expressed as

$$X_1'(z) = \lambda_1 \prod_{i=1}^{D}(z - \beta_i) \prod_{i=1}^{N-D-1}(z - \alpha_{1i}) \tag{C.2}$$

$$X_2'(z) = \lambda_2 \prod_{i=1}^{D}(z - \beta_i) \prod_{i=1}^{N-D-1}(z - \alpha_{2i}) \tag{C.3}$$

where λ_1, λ_2 remain to be determined.

To this aim, one writes the expression of the measurements polynomials in terms of $X_1'(z)$ and $X_2'(z)$ above. For instance:

$$\begin{aligned}\Gamma_{11}(z) &= X_1'(z) z^{N-1} \overline{X_1'(\overline{z}^{-1})}\\ &= |\lambda_1|^2 \prod_{i=1}^{D}(z - \beta_i) \prod_{i=1}^{N-D-1}(z - \alpha_{1i}) \prod_{i=1}^{D}(1 - \overline{\beta_i} z) \prod_{i=1}^{N-D-1}(1 - \overline{\alpha_{1i}} z)\end{aligned} \tag{C.4}$$

Using that $\Gamma_{11}(z) := \sum_{n=0}^{2N-2} \gamma_{11}[n - N + 1] z^n$, identifying leading order coefficients yields

$$\gamma_{11}[N-1] = |\lambda_1|^2 (-1)^{N-1} \prod_{i=1}^{D} \overline{\beta_i} \prod_{i=1}^{N-D-1} \overline{\alpha_{1i}} \tag{C.5}$$

Similarly, one gets

$$\gamma_{22}[N-1] = |\lambda_2|^2 (-1)^{N-1} \prod_{i=1}^{D} \overline{\beta_i} \prod_{i=1}^{N-D-1} \overline{\alpha_{2i}} \tag{C.6}$$

$$\gamma_{12}[N-1] = \lambda_1 \overline{\lambda}_2 (-1)^{N-1} \prod_{i=1}^{D} \overline{\beta_i} \prod_{i=1}^{N-D-1} \overline{\alpha_{2i}} \tag{C.7}$$

These relations determine uniquely the amplitudes of λ_1, λ_2 as well as the phase difference between λ_1 and λ_2. Thus λ_1, λ_2 are unique up to a global phase factor $\exp(\jmath\theta)$, $\theta \in [-\pi, \pi)$. One obtains eventually the following expressions

$$\lambda_1 = e^{\jmath\theta} \left(|\gamma_{11}[N-1]| \prod_{i=1}^{D} |\beta_i|^{-1} \prod_{i=1}^{N-D-1} |\alpha_{1i}|^{-1} \right)^{1/2} \tag{C.8}$$

$$\lambda_2 = e^{\jmath(\theta-\Delta)} \left(|\gamma_{22}[N-1]| \prod_{i=1}^{D} |\beta_i|^{-1} \prod_{i=1}^{N-D-1} |\alpha_{2i}|^{-1} \right)^{1/2} \tag{C.9}$$

with

$$\Delta = \arg(\lambda_1 \overline{\lambda_2}) \tag{C.10}$$

$$= \pi(N-1) + \arg \gamma_{12}[N-1] + \sum_{i=1}^{D} \arg \beta_i + \sum_{i=1}^{N-D-1} \arg \alpha_{2i} \, . \tag{C.11}$$

Appendix D. Sylvester matrices and greatest common divisors

Proof of Proposition 3. We first note that the result on the rank of $\mathcal{S}_D(A, B)$ is known (see, for example, [32, Theorem 4.7]). Thus, we are left prove the second part, which is somewhat related to [32, Remark 4.8].

We write $A(z) = F(z)H(z)$, $B(z) = G(z)H(z)$, so that $\gcd(A, B) = 1$ and $F, G \in \mathbb{C}_{\leq L-K}[z]$. Consider the following multiplication matrix

$$\mathbf{M}_{2L-D-K}(\mathbf{h}) = \underbrace{\begin{bmatrix} h_0 & & \\ \vdots & \ddots & \\ h_K & & h_0 \\ & \ddots & \vdots \\ & & h_K \end{bmatrix}}_{2L-D-K+1 \text{ columns}},$$

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

$$R(z) = U(z)A(z) + V(z)B(z) = H(z)(U(z)F(z) + V(z)G(z)), \tag{D.1}$$

and therefore any element in the range of $\mathcal{S}_D(A, B)$ belongs to the range of $\mathbf{M}_{2L-D-K}(\mathbf{h})$ (since the range of $\mathbf{M}_{2L-D-K}(\mathbf{h})$ corresponds to all polynomials of the form $H(z)Q(z)$ with $Q \in \mathbb{C}_{\leq 2L-D-K}[z]$).

Next we note that $\mathbf{M}_{2L-D-K}(\mathbf{h})$ is full column rank and therefore the ranks of $\operatorname{colspan}(\mathcal{S}_D(A, B))$ and $\operatorname{colspan}(\mathbf{M}_{2L-D-K}(\mathbf{h}))$ are equal. Hence the ranges of the two matrices coincide, as well as the left kernels; in particular the following equivalence holds true

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

Finally, easy algebraic calculations (see also, for instance, [32, eqn. (33)]) show that

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

which completes the proof. □

Appendix E. Cramèr-Rao bound for PPR

Several authors have considered Cramèr-Rao bounds for the classical phase retrieval problem with additive white gaussian noise [47, 29, 48]. These results directly apply to the additive Gaussian noise PPR model (30) since it can be equivalently rewritten as a particular one-dimensional noise model thanks to PPR-1D model introduced in Section 2.4. For completeness, we provide below an alternative derivation of the Cramèr-rao bound described in [48], where we use a full complex-domain approach instead of considering separate Cramèr-Rao bounds on amplitude and phase. Since measurement noise $n_{m,p}$ is i.i.d. Gaussian distributed with variance σ^2, the pdf of the vector of observations $\mathbf{y}$ is given by

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

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

where we recall that $\mathbf{C}_{m,p} \triangleq \mathbf{c}_{m,p}\mathbf{c}_{m,p}^{\mathsf{H}}$ with $\mathbf{c}_{m,p} = \overline{\mathbf{b}}_p \otimes \mathbf{a}_m$ by definition. One obtains the log-likelihood of observations as

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

Since one wants to estimate the complex parameter vector $\boldsymbol{\xi}$, it is necessary to use the complex Fisher Information Matrix (FIM) [49, 50, 51], which reads

$$\mathcal{J}_{\boldsymbol{\xi}} = \begin{bmatrix} \mathcal{I}_{\boldsymbol{\xi}} & \mathcal{P}_{\boldsymbol{\xi}} \\ \overline{\mathcal{P}_{\boldsymbol{\xi}}} & \overline{\mathcal{I}_{\boldsymbol{\xi}}} \end{bmatrix} \in \mathbb{C}^{4N\times 4N} \tag{E.4}$$

where entries are defined using Wirtinger derivatives [40] since $\boldsymbol{\xi}$ is a complex vector:

$$\mathcal{I}_{\boldsymbol{\xi}} = \mathbf{E}\left[\left(\nabla_{\overline{\boldsymbol{\xi}}}\log p(\mathbf{y}|\boldsymbol{\xi})\right)\left(\nabla_{\overline{\boldsymbol{\xi}}}\log p(\mathbf{y}|\boldsymbol{\xi})\right)^{\mathsf{H}}\right] \tag{E.5}$$

$$\mathcal{P}_{\boldsymbol{\xi}} = \mathbf{E}\left[\left(\nabla_{\overline{\boldsymbol{\xi}}}\log p(\mathbf{y}|\boldsymbol{\xi})\right)\left(\nabla_{\overline{\boldsymbol{\xi}}}\log p(\mathbf{y}|\boldsymbol{\xi})\right)^{\top}\right] \tag{E.6}$$

Note that the FIM $\mathcal{J}_{\boldsymbol{\xi}}$ defined in (E.4) is isomorphic to the real FIM which would have been obtained by stacking the real and imaginary parts of $\boldsymbol{\xi}$ in a single long vector [50]. This explains why $\mathcal{J}_{\boldsymbol{\xi}}$ has dimension $4N \times 4N$. Using properties of Wirtinger derivatives, we obtain

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

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 independence, one gets

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

$$= \frac{1}{\sigma^4}\sum_{m,p,m',p'}\mathbf{E}\left[n_{m,p}n_{m',p'}\right]\mathbf{C}_{m,p}\boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{H}}\mathbf{C}_{m',p'} \tag{E.9}$$

$$= \frac{1}{\sigma^2}\sum_{m,p}\mathbf{C}_{m,p}\boldsymbol{\xi}\boldsymbol{\xi}^{\mathsf{H}}\mathbf{C}_{m,p} \tag{E.10}$$

$$= \frac{1}{\sigma^2}\sum_{m,p}|\mathbf{c}_{m,p}^{\mathsf{H}}\boldsymbol{\xi}|^2\mathbf{c}_{m,p}\mathbf{c}_{m,p}^{\mathsf{H}} \tag{E.11}$$

Similar calculations leads to:

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

A key result [51] is that the inverse of the complex FIM (E.4) provides a lower bound on the covariance and pseudo-covariance of any unbiased estimator $\hat{\boldsymbol{\xi}}$ of the complex parameter $\boldsymbol{\xi}$:

$$\begin{bmatrix}\operatorname{cov}\hat{\boldsymbol{\xi}} & \operatorname{pcov}\hat{\boldsymbol{\xi}} \\ \overline{\operatorname{pcov}\hat{\boldsymbol{\xi}}} & \overline{\operatorname{cov}\hat{\boldsymbol{\xi}}}\end{bmatrix} \succeq \mathcal{J}_{\boldsymbol{\xi}}^{-1} \tag{E.13}$$

When the complex FIM is singular – as in phase retrieval [47, 29] –, one can show its pseudo-inverse remains a valid lower bound for the MSE; following the discussion in [48], we still refer to the resultant bound as the CRB with little abuse. In particular, we obtain the following bound on the MSE on any unbiased PPR estimator $\hat{\mathbf{X}}$ for the model (30):

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

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

References

[1] D. Sayre, Some implications of a theorem due to Shannon, Acta Crystallographica 5 (6) (1952) 843–843. doi:10.1107/S0365110X52002276.
URL http://scripts.iucr.org/cgi-bin/paper?S0365110X52002276

[2] V. Elser, Phase retrieval by iterated projections, JOSA A 20 (1) (2003) 40–55.

[3] V. Elser, T.-Y. Lan, T. Bendory, Benchmark problems for phase retrieval, SIAM Journal on Imaging Sciences 11 (4) (2018) 2429–2455.

[4] R. P. Millane, Phase retrieval in crystallography and optics, JOSA A 7 (3) (1990) 394–411.

[5] J. R. Fienup, Reconstruction of an object from the modulus of its Fourier transform, Optics Letters 3 (1) (1978) 27. doi:10.1364/OL.3.000027.
URL https://www.osapublishing.org/abstract.cfm?URI=ol-3-1-27

[6] J. R. Fienup, J. C. Marron, T. J. Schulz, J. H. Seldin, Hubble space telescope characterized by using phase-retrieval algorithms, Appl. Opt. 32 (10) (1993) 1747–1767. doi:10.1364/AO.32.001747.
URL http://opg.optica.org/ao/abstract.cfm?URI=ao-32-10-1747

[7] J. Miao, P. Charalambous, J. Kirz, D. Sayre, Extending the methodology of x-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens, Nature 400 (6742) (1999) 342–344.

[8] A. M. Maiden, J. M. Rodenburg, An improved ptychographical phase retrieval algorithm for diffractive imaging, Ultramicroscopy 109 (10) (2009) 1256–1262.

[9] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, M. Segev, Phase Retrieval with Application to Optical Imaging: A contemporary overview, IEEE Signal Processing Magazine 32 (3) (2015) 87–109. doi:10.1109/MSP.2014.2352673.
URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7078985

[10] R. Balan, P. Casazza, D. Edidin, On signal reconstruction without phase, Applied and Computational Harmonic Analysis 20 (3) (2006) 345–356. doi:10.1016/j.acha.2005.07.001.
URL https://linkinghub.elsevier.com/retrieve/pii/S1063520305000667

[11] E. J. Candès, Y. C. Eldar, T. Strohmer, V. Voroninski, Phase retrieval via matrix completion, SIAM Journal on Imaging Sciences 6 (1) (2013) 199–225. arXiv:https://doi.org/10.1137/110848074, doi:10.1137/110848074.
URL https://doi.org/10.1137/110848074

[12] E. Candes, X. Li, M. Soltanolkotabi, Phase Retrieval from Coded Diffraction Patterns, arXiv:1310.3240 [cs, math, stat]ArXiv: 1310.3240.
URL http://arxiv.org/abs/1310.3240

[13] A. S. Bandeira, Y. Chen, D. G. Mixon, Phase retrieval from power spectra of masked signals, Information and Inference: a Journal of the IMA 3 (2) (2014) 83–102.

[14] R. Beinert, G. Plonka, Ambiguities in one-dimensional discrete phase retrieval from fourier magnitudes, Journal of Fourier Analysis and Applications 21 (6) (2015) 1169–1198.

[15] T. Bendory, R. Beinert, Y. C. Eldar, Fourier Phase Retrieval: Uniqueness and Algorithms, in: H. Boche, G. Caire, R. Calderbank, M. März, G. Kutyniok, R. Mathar (Eds.), Compressed Sensing and its Applications, Springer International Publishing, Cham, 2017, pp. 55–91, series Title: Applied and Numerical Harmonic Analysis. doi:10.1007/978-3-319-69802-1_2.
URL http://link.springer.com/10.1007/978-3-319-69802-1_2

[16] R. A. Chipman, G. Young, W. S. T. Lam, Polarized light and optical systems, Optical sciences and applications of light, Taylor & Francis, CRC Press, Boca Raton, 2018.

[17] J. J. G. Perez, R. Ossikovski, Polarized Light and the Mueller Matrix Approach, Polarized Light (2018) 398.

[18] J. P. Gordon, H. Kogelnik, PMD fundamentals: Polarization mode dispersion in optical fibers, Proceedings of the National Academy of Sciences 97 (9) (2000) 4541–4550. doi:10.1073/pnas.97.9.4541.
URL http://www.pnas.org/content/97/9/4541.abstract

[19] J. S. Tyo, D. L. Goldstein, D. B. Chenault, J. A. Shaw, Review of passive imaging polarimetry for remote sensing applications, Applied optics 45 (22) (2006) 5453–5469.

[20] S.-M. Guo, L.-H. Yeh, J. Folkesson, I. E. Ivanov, A. P. Krishnan, M. G. Keefe, E. Hashemi, D. Shin, B. B. Chhun, N. H. Cho, M. D. Leonetti, M. H. Han, T. J. Nowakowski, S. B. Mehta, Revealing architectural order with quantitative label-free imaging and deep learning, eLife 9 (2020) e55502. doi:10.7554/eLife.55502.
URL https://elifesciences.org/articles/55502

[21] O. Smirnova, S. Patchkovskii, Y. Mairesse, N. Dudovich, D. Villeneuve, P. Corkum, M. Y. Ivanov, Attosecond Circular Dichroism Spectroscopy of Polyatomic Molecules, Physical Review Letters 102 (6) (2009) 063601. doi:10.1103/PhysRevLett.102.063601.
URL https://link.aps.org/doi/10.1103/PhysRevLett.102.063601

[22] O. Raz, O. Schwartz, D. Austin, A. S. Wyatt, A. Schiavi, O. Smirnova, B. Nadler, I. A. Walmsley, D. Oron, N. Dudovich, Vectorial Phase Retrieval for Linear Characterization of Attosecond Pulses, Physical Review Letters 107 (13) (2011) 133902, arXiv: 1104.5086. doi:10.1103/PhysRevLett.107.133902.
URL http://arxiv.org/abs/1104.5086

[23] P. Ferrand, M. Allain, V. Chamard, Ptychography in anisotropic media, Optics Letters 40 (22) (2015) 5144. doi:10.1364/OL.40.005144.
URL https://www.osapublishing.org/abstract.cfm?URI=ol-40-22-5144

[24] P. Ferrand, A. Baroni, M. Allain, V. Chamard, Quantitative imaging of anisotropic material properties with vectorial ptychography, Optics Letters 43 (4) (2018) 763, arXiv: 1712.00260. doi:10.1364/OL.43.000763.
URL http://arxiv.org/abs/1712.00260

[25] A. Baroni, V. Chamard, P. Ferrand, Extending Quantitative Phase Imaging to Polarization-Sensitive Materials, Physical Review Applied 13 (5) (2020) 054028. doi:10.1103/PhysRevApplied.13.054028.
URL https://link.aps.org/doi/10.1103/PhysRevApplied.13.054028

[26] A. Baroni, P. Ferrand, Reference-free quantitative microscopic imaging of coherent arbitrary vectorial light beams, Optics Express 28 (23) (2020) 35339. doi:10.1364/OE.408665.
URL https://www.osapublishing.org/abstract.cfm?URI=oe-28-23-35339

[27] K. Jaganathan, B. Hassibi, Reconstruction of Signals From Their Autocorrelation and Cross-Correlation Vectors, With Applications to Phase Retrieval and Blind Channel Estimation, IEEE Transactions on Signal Processing 67 (11) (2019) 2937–2946. doi:10.1109/TSP.2019.2911254.
URL https://ieeexplore.ieee.org/document/8691612/

[28] O. Raz, N. Dudovich, B. Nadler, Vectorial Phase Retrieval of 1-D Signals, IEEE Transactions on Signal Processing 61 (7) (2013) 1632–1643. doi:10.1109/TSP.2013.2239994.
URL http://ieeexplore.ieee.org/document/6410442/

[29] A. S. Bandeira, J. Cahill, D. G. Mixon, A. A. Nelson, Saving phase: Injectivity and stability for phase retrieval, Applied and Computational Harmonic Analysis 37 (1) (2014) 106–125. doi:10.1016/j.acha.2013.10.002.
URL https://linkinghub.elsevier.com/retrieve/pii/S1063520313000936

[30] K. Jaganathan, Y. Eldar, B. Hassibi, Phase retrieval with masks using convex optimization, in: 2015 IEEE International

Symposium on Information Theory (ISIT), IEEE, 2015, pp. 1655–1659.
[31] G. Heinig, K. Rost, Algebraic methods for Toeplitz-like matrices and operators, Birkhäuser, Basel, 1984.
[32] K. Usevich, I. Markovsky, Variable projection methods for approximate (greatest) common divisor computations, Theoretical Computer Science 681 (2017) 176–198. doi:10.1016/j.tcs.2017.03.028.
URL https://linkinghub.elsevier.com/retrieve/pii/S0304397517302505
[33] B. Schaefer, E. Collett, R. Smyth, D. Barrett, B. Fraher, Measuring the Stokes polarization parameters, Am. J. Phys. 75 (2) (2015) 6.
[34] D. H. Goldstein, Polarized light, CRC press, 2017.
[35] I. Markovsky, Structured low-rank approximation and its applications, Automatica 44 (4) (2008) 891–909.
[36] L. Vandenberghe, S. Boyd, Semidefinite programming, SIAM review 38 (1) (1996) 49–95.
[37] R. Monteiro, et al., First-and second-order methods for semidefinite programming, Mathematical Programming 97 (1) (2003) 209–244.
[38] A. Beck, First-order methods in optimization, SIAM, 2017.
[39] T. Goldstein, C. Studer, R. Baraniuk, A field guide to forward-backward splitting with a fasta implementation, arXiv preprint arXiv:1411.3406.
[40] K. Kreutz-Delgado, The Complex Gradient Operator and the CR-Calculus, arXiv:0906.4835 [math]ArXiv: 0906.4835.
URL http://arxiv.org/abs/0906.4835
[41] E. Candes, X. Li, M. Soltanolkotabi, Phase Retrieval via Wirtinger Flow: Theory and Algorithms, IEEE Transactions on Information Theory 61 (4) (2015) 1985–2007, arXiv: 1407.1065. doi:10.1109/TIT.2015.2399924.
URL http://arxiv.org/abs/1407.1065
[42] X. Jiang, S. Rajan, X. Liu, Wirtinger Flow Method With Optimal Stepsize for Phase Retrieval, IEEE Signal Processing Letters 23 (11) (2016) 1627–1631. doi:10.1109/LSP.2016.2611940.
URL http://ieeexplore.ieee.org/document/7572075/
[43] R. Xu, M. Soltanolkotabi, J. P. Haldar, W. Unglaub, J. Zusman, A. F. J. Levi, R. M. Leahy, Accelerated Wirtinger Flow: A fast algorithm for ptychography, arXiv:1806.05546 [eess, math]ArXiv: 1806.05546.
URL http://arxiv.org/abs/1806.05546
[44] I. A. Walmsley, C. Dorrer, Characterization of ultrashort electromagnetic pulses, Advances in Optics and Photonics 1 (2) (2009) 308. doi:10.1364/AOP.1.000308.
URL https://www.osapublishing.org/aop/abstract.cfm?uri=aop-1-2-308
[45] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, M. Bartelmann, Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere, The Astrophysical Journal 622 (2) (2005) 759.
[46] D. Cox, J. Little, D. O'Shea, Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 2nd Edition, Springer, 1997.
[47] R. Balan, Reconstruction of signals from magnitudes of redundant representations: The complex case, Foundations of Computational Mathematics 16 (3) (2016) 677–721.
[48] C. Qian, N. D. Sidiropoulos, K. Huang, L. Huang, H. C. So, Phase retrieval using feasible point pursuit: Algorithms and cramér–rao bound, IEEE Transactions on Signal Processing 64 (20) (2016) 5282–5296.
[49] A. Van den Bos, A cramér-rao lower bound for complex parameters, IEEE Transactions on Signal Processing [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on], 42 (10).
[50] B. Loesch, B. Yang, Cramér-rao bound for circular and noncircular complex independent component analysis, IEEE transactions on signal processing 61 (2) (2012) 365–379.
[51] E. Ollila, V. Koivunen, J. Eriksson, On the cramér-rao bound for the constrained and unconstrained complex parameters, in: 2008 5th IEEE Sensor Array and Multichannel Signal Processing Workshop, IEEE, 2008, pp. 414–418.