Signal reconstruction from the magnitude of subspace components

Christine Bachoc and Martin Ehler

***Abstract*—We consider signal reconstruction from the norms of subspace components generalizing standard phase retrieval problems. In the deterministic setting, a closed reconstruction formula is derived when the subspaces satisfy certain cubature conditions, that require at least a quadratic number of subspaces. Moreover, we address reconstruction under the erasure of a subset of the norms; using the concepts of p-fusion frames and list decoding, we propose an algorithm that outputs a finite list of candidate signals, one of which is the correct one. In the random setting, we show that a set of subspaces chosen at random and of cardinality scaling linearly in the ambient dimension allows for exact reconstruction with high probability by solving the feasibility problem of a semidefinite program.**

***Index Terms*—phase retrieval, Grassmannian cubature, fusion frame.**

I. Introduction

The phase retrieval problem, which refers to the task of recovering a signal from the absolute values of linear measurements, has received much attention recently: see [7], [8], [16], [18], [21], [50] to mention only few. We are dealing with a generalization, in which the measurements consist of norms of projections of the signal onto k-dimensional subspaces. For $k = 1$, our setting reduces to the classical phase retrieval problem.

Here, we pose the following questions: Under which properties of the subspaces can we reconstruct the original signal from the norms of its k-dimensional subspace components by means of a closed formula? Also, given that requiring a closed formula for reconstruction is too costly, can we develop strategies to reduce the number of required subspace components, under a numerical reconstruction? We shall provide positive xanswers for a deterministic choice and a random choice of subspaces.

Deterministic setting: Given k-dimensional linear subspaces $\{V_j\}_{j=1}^n$ in $\mathbb{R}^d$, we aim to reconstruct the signal $x \in \mathbb{R}^d$ from $\{\|P_{V_j}(x)\|\}_{j=1}^n$, where P_{V_j} denotes the orthogonal projector onto V_j. Clearly, x can only be recovered up to its sign. In [15], several characterizations of subspaces such that the mapping $\{\pm x\} \mapsto \{\|P_{V_j}(x)\|\}_{j=1}^n$ is injective are given, see also [6]. Our aim is to showcase properties of the subspaces that moreover allow for an explicit reconstruction formula.

If there are positive weights $\{\omega_j\}_{j=1}^n$ such that $\{(V_j, \omega_j)\}_{j=1}^n$ yields a so-called cubature of strength 4 as defined in Section II-C of the present paper, then we shall obtain a closed reconstruction formula for xx^* enabling us to extract $\pm x$. Thus, we extend the 1-dimensional results in [7] to k-dimensional projections. Note that the authors in [7] require cubatures for the projective space whose weights are $\omega_j = 1/n$, i.e., so-called projective designs. In practice, however, the choice of subspaces may underlie restrictions that prevent them from being a design. Therefore, our results are a significant improvement for 1-dimensional projections already.

To address subspace erasures, we suppose that we are only given the values of $n - p$ norms and we need to reconstruct the missing p norms. Notice that our input are not the subspace components but their norms, as opposed to signal reconstruction under the erasures discussed in [13], [42], [43]. If there are positive weights $\{\omega_j\}_{j=1}^n$ such that $\{(V_j, \omega_j)\}_{j=1}^n$ forms a tight p-fusion frame as recently introduced in [5], then the computation of the erased norms up to permutations amounts to solving a system of algebraic equations. We can then reconstruct $\pm x$ from the entire set of n magnitude subspace components. In other words, we found conditions on subspaces, so that we can compute a finite list of candidate signals, one of which is the correct one. The latter is a form of list decoding as introduced in [34].

The limit of this deterministic approach stands in the required number of subspaces. Indeed, it is known that the cardinality of a cubature formula of strength 4 scales at least quadratically with the ambient dimension d. In the random setting, it will be possible to reduce the number of subspaces to linear size:

Random setting: We shall extend to k-dimensional subspaces the results obtained for $k = 1$ in a recent series of papers [18], [29], [17]. In [18] it was shown that semidefinite programming yields signal recovery with high probability when the 1-dimensional subspaces are chosen at random and that the cardinality of the subspaces can scale linearly in the ambient dimension up to a logarithmic factor. Numerical stability in the presence of noise was also verified. The underlying semidefinite program was shown in [29] to afford (with high probability) a unique feasible solution, and the logarithmic factor was removed in [17].

Our proof for k-dimensional subspaces (see Theorem V.1) follows the approach in [17], [18]. We verify that randomly selected subspaces satisfy a near isometry property and ensure the existence of a so-called dual certificate, which implies that the solution of the semidefinite program indeed recovers the signal with high probability. However, the generalization to k-dimensional projections raised additional difficulties. Indeed, the case $k = 1$ relies on random vectors whose entries are

C. Bachoc is with the Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France, e-mail: christine.bachoc@math.u-bordeaux1.fr.

M. Ehler is with the University of Vienna, Department of Mathematics, Oskar-Morgenstern-Platz 1 A-1090 Vienna, e-mail: martin.ehler@univie.ac.at.

i.i.d. Gaussian modeling the measurements. For $k > 1$, we must deal with measurement matrices having orthogonal rows, so that entries from one row are stochastically dependent on those in any other row. Hence, the extension from $k = 1$ to $k > 1$ is not obvious and requires special care. In particular, we apply weak convergence results from random matrix theory to derive the lower estimate for the near isometry property, see our Proposition V.7. Moreover, we present numerical experiments to illustrate the practical feasibility of the method in small dimensions.

Complex case: Although we present our results for real signals and subspaces exclusively, the agenda can also be followed in the complex setting. We shall discuss the required modifications at the end of the present paper.

We would like to mention that signal reconstruction from phaseless measurements is a common problem in optical physics such as X-ray crystallography and diffraction imaging, where coherent light sources correspond to magnitude measurements of Fourier frame coefficients. Crystal twinning [30], on the other hand, involves signal reconstruction from averaged diffraction patterns by means of incoherent addition of k wavefields, where usually $k = 1, 2, 3$ [35]. Indeed, incoherent light sources involve (weighted) sums of k squared moduli of Fourier coefficients. The latter is essentially the squared norm of the orthogonal projection onto the associated k-dimensional subspace. Hence, our mathematical setting of rank-k orthogonal projectors, where k is independent of the ambient dimension d, relates to measurements in optical physics. Nonetheless, we should mention that we do not focus on Fourier type measurements and do not incorporate any add itional information and side constraints that are commonly available in optical physics measurements and that are used in standard reconstruction algorithms, see [36], [38] and also [10].

Outline: In Section II, we recall fusion frames, state the phase retrieval problem, and introduce tight p-fusion frames and cubature formulas. We present the closed reconstruction formula in Section III and our reconstruction algorithm in presence of erasures in Section IV. The random subspace selection is addressed in Section V. Numerical experiments are presented in Section VI, and we discuss the complex setting in Section VII.

II. Fusion frames, phase retrieval, and cubature formulas

A. Fusion frames and the problem of reconstruction without phase

Let $\mathcal{G}_{k,d} = \mathcal{G}_{k,d}(\mathbb{R})$ denote the *real Grassmann space*, i.e., the k-dimensional subspaces of $\mathbb{R}^d$. Each $V \in \mathcal{G}_{k,d}$ can be identified with the orthogonal projector onto V, denoted by P_V. Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ and let $\{\omega_j\}_{j=1}^n$ be a collection of positive weights. Then $\{(V_j, \omega_j)\}_{j=1}^n$ is called a *fusion frame* if there are positive constants A and B such that

$$A\|x\|^2 \leq \sum_{j=1}^n \omega_j \|P_{V_j}(x)\|^2 \leq B\|x\|^2, \text{ for all } x \in \mathbb{R}^d, \quad (1)$$

cf. [20]. The condition (1) is equivalent to

$$A \leq \sum_{j=1}^n \omega_j \langle P_x, P_{V_j}\rangle \leq B, \text{ for all } x \in S^{d-1}, \quad (2)$$

where P_x is short for $P_{x\mathbb{R}}$ and $\langle P_x, P_{V_j}\rangle := \operatorname{trace}(P_x P_{V_j})$ is the standard inner product between self-adjoint operators. If $A = B$, then $\{(V_j, \omega_j)\}_{j=1}^n$ is called a *tight fusion frame*, and any signal $x \in S^{d-1}$ can be reconstructed from its subspace components by the simple formula

$$x = \frac{1}{A}\sum_{j=1}^n \omega_j P_{V_j}(x). \quad (3)$$

If, however, instead of $\{P_{V_j}(x)\}_{j=1}^n$ we only observe the norms $\{\|P_{V_j}(x)\|\}_{j=1}^n$ and, worse, we even lose some of these norms, can we still reconstruct x? Clearly, x can be determined up to its sign at best. In the present paper, we find conditions on $\{(V_j, \omega_j)\}_{j=1}^n$ together with a computationally feasible algorithm that enable us to determine $\pm x$.

Remark II.1. We want to point out that 1-bit compressed sensing, cf. [14], [45], deals with a problem that is complementary to phase retrieval. There, the magnitudes are unknown and signals are reconstructed from the signs of the frame coefficients.

B. Tight p-fusion frames

Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ and let $\{\omega_j\}_{j=1}^n$ be a collection of positive weights and p a positive integer. Then $\{(V_j, \omega_j)\}_{j=1}^n$ is called a *p-fusion frame* in [5] if there exist positive constants A_p and B_p such that

$$A_p\|x\|^{2p} \leq \sum_{j=1}^n \omega_j \|P_{V_j}(x)\|^{2p} \leq B_p\|x\|^{2p}, \text{ for all } x \in \mathbb{R}^d, \quad (4)$$

see also [33] for related concepts. If $A_p = B_p$, then $\{(V_j, \omega_j)\}_{j=1}^n$ is called a *tight p-fusion frame*. As with (1) and (2), the condition (4) is equivalent to

$$A_p \leq \sum_{j=1}^n \omega_j \langle P_x, P_{V_j}\rangle^p \leq B_p, \text{ for all } x \in S^{d-1}. \quad (5)$$

If $\{(V_j, \omega_j)\}_{j=1}^n$ is a tight p-fusion frame, then it is also a tight ℓ-fusion frame for all integers $1 \leq \ell \leq p$, and the tight ℓ-fusion frame bounds are

$$A_\ell = \frac{(k/2)_\ell}{(d/2)_\ell}\sum_{j=1}^n \omega_j, \quad (6)$$

where we used $(a)_\ell = a(a+1)\cdots(a+\ell-1)$, cf. [5]. We also refer to [5] for constructions and general existence results.

C. Cubature formulas

The real orthogonal group $O(\mathbb{R}^d)$ acts transitively on $\mathcal{G}_{k,d}$, and the Haar measure on $O(\mathbb{R}^d)$ induces a probability measure σ_k on $\mathcal{G}_{k,d}$. Let $L^2(\mathcal{G}_{k,d})$ denote the complex valued functions on $\mathcal{G}_{k,d}$, whose squared module is integrable with respect to σ_k. The complex irreducible representations of $O(\mathbb{R}^d)$ are associated to partitions $\mu = (\mu_1, \ldots, \mu_d)$, $\mu_1 \geq \ldots \geq \mu_d \geq 0$,

denoted by V_d^μ, cf. [39]. Let $l(\mu)$ be the number of nonzero entries in μ so that

$$L^2(\mathcal{G}_{k,d}) = \bigoplus_{l(\mu)\le k} H_{k,d}^{2\mu}, \quad \text{where } H_{k,d}^{2\mu} \simeq V_d^{2\mu}, \tag{7}$$

see [39]. The space of polynomial functions on $\mathcal{G}_{k,d}$ of degree bounded by $2p$ is

$$\mathrm{Pol}_{\le 2p}(\mathcal{G}_{k,d}) := \bigoplus_{l(\mu)\le k,\ |\mu|\le p} H_{k,d}^{2\mu}, \tag{8}$$

and we additionally define the subspace

$$\mathrm{Pol}^1_{\le 2p}(\mathcal{G}_{k,d}) := \bigoplus_{l(\mu)\le 1,\ |\mu|\le p} H_{k,d}^{2\mu}. \tag{9}$$

These spaces are explicitly given by

$$\mathrm{Pol}_{\le 2p}(\mathcal{G}_{k,d}) = \mathrm{span}\{V\mapsto\langle P_{x_1},P_V\rangle\cdots\langle P_{x_p},P_V\rangle : \{x_i\}_{i=1}^p \subset S^{d-1}\},$$
$$\mathrm{Pol}^1_{\le 2p}(\mathcal{G}_{k,d}) = \mathrm{span}\{V\mapsto\langle P_x,P_V\rangle^p : x\in S^{d-1}\},$$

cf. [5, Remark 5.4, proof of Theorem 5.3]. Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ and $\{\omega_j\}_{j=1}^n$ be a collection of positive weights normalized such that $\sum_{j=1}^n \omega_j = 1$. Then $\{(V_j,\omega_j)\}_{j=1}^n$ is called a *cubature of strength* $2p$ *for* $\mathcal{G}_{k,d}$ if

$$\int_{\mathcal{G}_{k,d}} f(V)d\sigma_k(V) = \sum_{j=1}^n \omega_j f(V_j) \quad \text{for all } f\in \mathrm{Pol}_{\le 2p}(\mathcal{G}_{k,d}). \tag{10}$$

Grassmannian designs, i.e., cubatures with constant weights, have been studied in [1], [2], [3], [4]. For existence results on cubatures and the relations between p and n, we refer to [27]. It was verified in [5] that $\{(V_j,\omega_j)\}_{j=1}^n$ is a tight p-fusion frame if and only if

$$\int_{\mathcal{G}_{k,d}} f(V)d\sigma_k(V) = \sum_{j=1}^n \omega_j f(V_j) \quad \text{for all } f\in \mathrm{Pol}^1_{\le 2p}(\mathcal{G}_{k,d}).$$

Thus, any cubature of strength $2p$ is a tight p-fusion frame. The converse implication does not hold in general except for p or k equals 1.

Remark II.2. Note that the case $k=1$ with constant weights corresponds to projective designs. Spherical designs have been widely studied in the literature [9], [28], [47] and any antipodal spherical $2p$-design induces a projective $2p$-design by choosing the lines along the antipodal points.

III. Signal reconstruction in the case of a cubature of strength 4

Let $\mathscr{H}$ denote the collection of symmetric matrices in $\mathbb{R}^{d\times d}$. If $\{P_{V_j}\}_{j=1}^n$ spans $\mathscr{H}$, then standard results in frame theory imply that $S:\mathscr{H}\to\mathscr{H}$ given by $X\mapsto \sum_{j=1}^n\langle X,P_{V_j}\rangle P_{V_j}$ is invertible and

$$xx^* = \sum_{j=1}^n \|P_{V_j}(x)\|^2 S^{-1}(P_{V_j}), \quad \text{for all } x\in\mathbb{R}^d.$$

By imposing stronger conditions on $\{P_{V_j}\}_{j=1}^n$, the operator S can be inverted explicitly. To that end, we establish the following result that generalizes the case $k=1$ treated in [7]. We point out that we allow for cubatures as opposed to projective designs in [7] that require the cubature weights to be constant:

Proposition III.1. *Let $\{(V_j,\omega_j)\}_{j=1}^n$ be a cubature of strength 4 for $\mathcal{G}_{k,d}$. If $x\in S^{d-1}$, then*

$$P_x = a_1\sum_{j=1}^n \omega_j\|P_{V_j}(x)\|^2 P_{V_j} - a_2 I, \tag{11}$$

where $a_1 = \frac{d(d+2)(d-1)}{2k(d-k)}$ and $a_2 = \frac{kd+k-2}{2(d-k)}$.

Proof. For any $x,y\in S^{d-1}$, the function $V\mapsto\langle P_x,P_V\rangle\langle P_y,P_V\rangle$ belongs to $\mathrm{Pol}_{\le 4}(\mathcal{G}_{k,d})$. Applying the cubature formula yields

$$\sum_{j=1}^n \omega_j\langle P_x,P_{V_j}\rangle\langle P_y,P_{V_j}\rangle = \int_{\mathcal{G}_{k,d}}\langle P_x,P_V\rangle\langle P_y,P_V\rangle d\sigma_k(V). \tag{12}$$

The function

$$G:(\mathbb{R}x,\mathbb{R}y)\mapsto \int_{\mathcal{G}_{k,d}}\langle P_x,P_V\rangle\langle P_y,P_V\rangle d\sigma_k(V) \tag{13}$$

belongs to $L^2(\mathcal{G}_{1,d}\times\mathcal{G}_{1,d})$ and is zonal. For each variable, it has the form $\mathbb{R}x\mapsto\langle P_x,A(y)\rangle$, where $A(y) = \int_{\mathcal{G}_{k,d}}\langle P_y,P_V\rangle P_V d\sigma_k(V)$, and $\mathbb{R}y\mapsto\langle P_y,A(x)\rangle$, respectively. Since $A(y)$ is self-adjoint and hence a linear combination of projections, $G(\cdot,\mathbb{R}y)$ and $G(\mathbb{R}x,\cdot)$ belong to $\mathrm{Pol}_{\le 2}(\mathcal{G}_{1,d})$. The zonal functions on the projective space are polynomials in the variable $\langle P_x,P_y\rangle = (x,y)^2$, so that G must be of the form $\alpha_1(x,y)^2+\alpha_2$. Thus, (12) yields

$$\sum_{j=1}^n \omega_j\langle P_x,P_{V_j}\rangle\langle P_y,P_{V_j}\rangle = \alpha_1\langle P_x,P_y\rangle + \alpha_2\langle I,P_y\rangle. \tag{14}$$

Since (14) holds for every y, we derive

$$\sum_{j=1}^n \omega_j\langle P_x,P_{V_j}\rangle P_{V_j} = \alpha_1 P_x + \alpha_2 I. \tag{15}$$

Taking traces in (15) leads to $k\sum_{j=1}^n\omega_j\langle P_x,P_{V_j}\rangle = \alpha_1 + d\alpha_2$, and the property of tight 1-fusion frames gives $\sum_{j=1}^n\omega_j\langle P_x,P_{V_j}\rangle = A_1 = k/d$, so we obtain

$$\alpha_1 + d\alpha_2 = k^2/d. \tag{16}$$

Taking $x=y$ in (14) implies $\sum_{j=1}^n\omega_j\langle P_x,P_{V_j}\rangle^2 = \alpha_1+\alpha_2$, and the tight 2-fusion frame property leads to $\sum_{j=1}^n\omega_j\langle P_x,P_{V_j}\rangle^2 = A_2 = k(k+2)/(d(d+2))$, so that we obtain

$$\alpha_1+\alpha_2 = k(k+2)/(d(d+2)). \tag{17}$$

Solving for α_1 and α_2 in (16) and (17) yields the required identity with $a_1 = 1/\alpha_1$ and $a_2 = \alpha_2/\alpha_1$. □

Remark III.2. Since any $X\in\mathscr{H}$ can be written as a sum of weighted orthogonal projectors, (11) can be extended to

$$X = a_1\sum_{j=1}^n\omega_j\langle X,P_{V_j}\rangle P_{V_j} - a_2\,\mathrm{trace}(X)I. \tag{18}$$

For $x\in\mathbb{R}^d$ and $X=xx^*$, the tight-1 fusion frame property yields $\mathrm{trace}(X) = \|x\|^2 = \frac{d}{k}\sum_{j=1}^n\omega_j\|P_{V_j}(x)\|^2$, so that

the entire right-hand side of (18) can be computed from $\{\|P_{V_j}(x)\|^2\}_{j=1}^n$ and hence $\pm x$ can be recovered.

We can conclude from (18) that $\{\omega_j P_{V_j}\}_{j=1}^n$ and $\{Q_j\}_{j=1}^n$, where $Q_j = a_1 P_{V_j} - a_2 \frac{d}{k} I$, are pairs of dual frames for $\mathscr{H}$, i.e.,

$$X = \sum_{j=1}^n \langle X, \omega_j P_{V_j}\rangle Q_j, \quad \text{for all } X \in \mathscr{H}.$$

Moreover, if V is a random subspace, uniformly distributed in $\mathcal{G}_{k,d}$, i.e., distributed according to σ_k, then the proof of Proposition III.1 yields that

$$a_1 \mathbb{E}(\langle X, P_V\rangle P_V) - a_2 \operatorname{trace}(X) I = X, \tag{19}$$

for all $X \in \mathscr{H}$. Thus, if $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ are independent copies of V, then the law of large numbers implies

$$\frac{a_1}{n} \sum_{j=1}^n \langle X, P_{V_j}\rangle P_{V_j} - a_2 \operatorname{trace}(X) I \to X \quad \text{almost surely.} \tag{20}$$

However, n must be chosen large to obtain an accurate representation of X. In Sections V and VI, we shall see that the random choice of subspaces can be efficient when the algebraic reconstruction formula is replaced with a semidefinite program.

IV. Algorithm for signal reconstruction from magnitudes of incomplete subspace components

In this section, we consider the situation where $x \in S^{d-1}$, and we aim to reconstruct $\pm x$ from any $n-p$ elements of the set $\{\|P_{V_j}(x)\|^2\}_{j=1}^n$. Indeed, for fixed p, we are aiming at a reconstruction scheme valid for any subset of p missing norms. Without loss of generality, we can assume that the first p norms have been erased, so we want to recover $\pm x$ from the knowledge of $\{\|P_{V_j}(x)\|\}_{j=p+1}^n$.

In a first step, we attempt to compute the missing values

$$t_j := \|P_{V_j}(x)\|^2, \quad 1 \le j \le p.$$

This will be made possible by the property that $\{(V_j, \omega_j)\}_{j=1}^n$ is a tight p-fusion frame with $\sum_{j=1}^n \omega_j = 1$. The second step is dedicated to reconstructing $\pm x$ from $\{\|P_{V_j}(x)\|^2\}_{j=1}^n$.

A. Step 1: reconstruction of the erased norms

The tight p-fusion frame $\{(V_j, \omega_j)\}_{j=1}^n$ is also a tight ℓ-fusion frame for $1 \le \ell \le p$, cf. [5, Proposition 5.1], so that (6) yields

$$\sum_{j=1}^n \omega_j \|P_{V_j}(x)\|^{2\ell} = A_\ell = \frac{(k/2)_\ell}{(d/2)_\ell}, \quad 1 \le \ell \le p.$$

Therefore, $(t_1, \dots, t_p)$ is a solution of the algebraic system of equations

$$\sum_{j=1}^p \omega_j T_j^\ell = \frac{(k/2)_\ell}{(d/2)_\ell} - \sum_{j=p+1}^n \omega_j \|P_{V_j}(x)\|^{2\ell}, \quad 1 \le \ell \le p, \tag{AE}$$

in the unknowns $(T_1, \dots, T_p)$. To start with, let us consider the special case of equal weights; then, (AE) gives the values of the symmetric powers $\sum_{j=1}^p t_j^\ell$, for $\ell = 0, \dots, p$, which, as polynomial expressions, generate the ring of symmetric polynomials up to degree p. Vieta's formula yields

$$\prod_{i=1}^p (T - t_i) = \sum_{j=0}^p (-1)^j e_j T^{p-j},$$

where $e_0 = 1$ and $e_j = \sum_{1 \le i_1 < \dots < i_j \le p} t_{i_1} \cdots t_{i_p}$, and Newton's identity leads to

$$e_j = \frac{1}{j} \sum_{\ell=1}^j (-1)^{\ell-1} e_{j-\ell} \sum_{j=1}^p t_j^\ell, \quad \text{for } j = 1, \dots, p.$$

Therefore, we can compute the coefficients of $\prod_{i=1}^p (T - t_i)$ as a polynomial in T and solve for its roots; we see that $(t_1, \dots, t_p)$ is determined up to a permutation so that we obtain at most $p!$ distinct solutions to (AE).

If the weights are not equal, one can still show that (AE) has at most $p!$ solutions. The issue is to verify that the associated affine variety is zero-dimensional. Results from intersection theory and the refined Bézout theorem, cf. [37], [48] and [12], then imply that the variety's cardinality is at most the product of the degrees of the p polynomials, i.e., there are at most $p!$ solutions:

Proposition IV.1. *Let $\{b_\ell\}_{\ell=1}^p$ be complex numbers and define*

$$f_\ell(T) = \sum_{j=1}^p \omega_j T_j^\ell - b_\ell, \quad \ell = 1, \dots, p.$$

If $\{\omega_j\}_{j=1}^p$ are positive numbers, then the affine variety $\mathcal{V} := \{T \in \mathbb{C}^p : f_1(T) = 0, \dots, f_p(T) = 0\}$ is zero-dimensional.

Proof. We proceed by induction on p. The assertion is certainly true for $p = 1$. Next, we observe that the Jacobian determinant satisfies

$$\det\Big(\frac{\partial(f_1, \dots, f_p)}{\partial(T_1, \dots, T_p)}\Big) = \det \begin{pmatrix} \omega_1 & \dots & \omega_p \\ \vdots & & \vdots \\ \omega_1 p T_1^{p-1} & \dots & \omega_1 p T_p^{p-1} \end{pmatrix}.$$

By applying the multilinearity of the determinant once to the columns and another time to the rows, we obtain

$$\det\Big(\frac{\partial(f_1, \dots, f_p)}{\partial(T_1, \dots, T_p)}\Big) = \omega_1 \cdots \omega_p \cdot p! \cdot \det \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ T_1^{p-1} & \dots & T_p^{p-1} \end{pmatrix}.$$

The well-known formula for the Vandermonde determinant yields

$$\det\Big(\frac{\partial(f_1, \dots, f_p)}{\partial(T_1, \dots, T_p)}\Big) = \omega_1 \cdots \omega_p \cdot p! \cdot \prod_{1 \le i < j \le p} (T_j - T_i).$$

For $i < j$, let $\Delta_{i,j} := \{T \in \mathbb{C}^d : T_i = T_j\}$ denote the diagonals. The Jacobian determinant is apparently nonzero for $T \notin \bigcup_{i<j} \Delta_{i,j}$. Therefore, every $T \in \mathcal{V} \setminus \bigcup_{i<j} \Delta_{i,j}$ is a nonsingular point of $\mathcal{V}$, and the dimension of $\mathcal{V}$ at T is $p - p = 0$, cf. [25, Theorem 9.9] and [12, Lemma 11.5.1]. It remains to consider the intersection of $\mathcal{V}$ with $\Delta_{i,j}$. To fix ideas, let

us consider the case $i = 1$, $j = 2$. The intersection $\mathcal{V} \cap \Delta_{1,2}$ is given by the system of equations

$$(\omega_1 + \omega_2)T_2^\ell + \sum_{j=3}^{p} \omega_j T_j^\ell = b_\ell, \quad \ell = 1, \dots, p.$$

Because $\omega_1 + \omega_2 > 0$, by induction the first $p-1$ of these equations have only finitely many solutions. Thus, $\mathcal{V} \cap \Delta_{1,2}$ is finite, too. □

Remark IV.2. The above proof shows that the positivity assumption on the weights in Proposition IV.1 can be replaced with $\omega_{j_1} + \ldots + \omega_{j_i} \neq 0$, for all $1 \leq j_1 < \ldots < j_i \leq p$ and $i = 1 \ldots, p$.

We have proved that the system of algebraic equations (AE) has at most $p!$ complex solutions. In order to compute these solutions, standard algorithmic methods can be applied [24], [25]. The construction of a Gröbner basis of the ideal $\mathcal{I}$ generated by the p equations allows to compute the algebraic operations in the quotient ring $\mathbb{R}[T_1, \dots, T_p]/\mathcal{I}$, which is finite dimensional and of dimension at most $p!$. The computation of the solutions then boils down to linear algebra in this space.

B. Step 2: reconstruction from the magnitude of subspace components

In this second step, we try to compute P_x from each of the possible candidates for $\{\|P_{V_j}(x)\|^2\}_{j=1}^n$ derived from a solution $(t_1, \dots, t_p)$ of (AE). For this, we assume that $\{(V_j, \omega_j)\}_{j=1}^n$ is also a cubature of strength 4, and we apply formula (11) where we replace $\|P_{V_j}(x)\|^2$ by t_j for $1 \leq i \leq p$.

To summarize, we have proved:

Theorem IV.3. *Let $\{(V_j, \omega_j)\}_{j=1}^n$ be a tight p-fusion frame that is also a cubature of strength 4 for $\mathcal{G}_{k,d}$. If $x \in S^{d-1}$, then Algorithm 1 outputs a list L of at most $2p!$ elements of S^{d-1} containing x.*

Algorithm 1 List reconstruction

Input: $\{t_j := \|P_{V_j}(x)\|^2\}_{j=p+1}^n$.
Output: L, $x \in L$.

1: Initialize $L = \emptyset$.
2: Compute the set $\mathcal{S}$ of solutions of the algebraic system of equations in the unknowns $T_1, \dots, T_p$:

$$\sum_{j=1}^{p} \omega_j T_j^\ell = \frac{(k/2)_\ell}{(d/2)_\ell} - \sum_{j=p+1}^{n} \omega_j t_j^\ell, \quad 1 \leq \ell \leq p. \quad \text{(AE)}$$

3: For every $(t_1, \dots, t_p) \in \mathcal{S}$, and α, β defined in Proposition III.1, compute

$$P = a_1 \sum_{j=1}^{n} \omega_j t_j P_{V_j} - a_2 I.$$

4: If P is a projection of rank 1, compute a unit vector ξ spanning its image and add $\pm\xi$ to L.
5: **return** L

Note that the cardinality of the list L is triggered by the number of erasures p. The number of measurements depends on p and on the ambient signal dimension d. The weighted subspaces are supposed to form a cubature of strength 4 for $\mathcal{G}_{k,d}$, therefore, we must have at least $n \geq \frac{1}{2}d(d+1)$ many subspaces, see [27]. Hence, the cardinality scales at least quadratically in the ambient dimension d already for $p = 2$. In general, the minimal number of measurements is a growing function of p and d. We refer to [5] for some explicit constructions of tight p-fusion frames.

We also note that the actual output list L can be much shorter than $2p!$ because many solutions of the algebraic system of equations will not lead to a candidate for the signal x. In the first place, we can exclude those solutions of (AE) that are not real or have negative entries. Moreover, one can expect that, for most solutions of (AE), the symmetric operator P in step 3 is not a rank-one projector. Also, the solutions $(t_1, \dots, t_p)$ of (AE) that do not satisfy $|t_i^{1/2} - t_j^{1/2}|^2 \leq \|P_{V_i} - P_{V_j}\|^2$, for every $1 \leq i < j \leq n$, can be removed because they violate the consistency conditions

$$|\|P_{V_i}(x)\| - \|P_{V_j}(x)\||^2 \leq \|P_{V_i}(x) - P_{V_j}(x)\|^2 \leq \|P_{V_i} - P_{V_j}\|_\infty^2,$$

where $\|P_{V_i} - P_{V_j}\|_\infty$ denotes the operator norm of $P_{V_i} - P_{V_j}$.

Remark IV.4. For $p = 2$, the assumptions in Theorem IV.3 reduce to $\{(V_j, \omega_j)\}_{j=1}^n$ being a cubature of strength 4 for $\mathcal{G}_{k,d}$. Even for $k = 1$, our result extends [7] since we only need $n-2$ elements of the collection $\{\|P_{V_j}(x)\|^2\}_{j=1}^n$ as opposed to all n elements in [7]. This additional flexibility is not for free: We must assume that $x \in S^{d-1}$, and, instead of the two possibilities $\pm x$ in [7], we obtain a list L of 4 elements, one of which is x.

V. Replacing the algebraic reconstruction formula with semidefinite programming

We assume in Proposition III.1 that the weighted subspaces form a cubature of strength 4 for $\mathcal{G}_{k,d}$. However, any real cubature of strength 4 requires at least $\frac{1}{2}d(d+1)$ subspaces, see [27]. Hence, the cardinality scales at least quadratically in the ambient dimension d. In this section, we replace the algebraic reconstruction formula with a feasibility problem of a semidefinite program similar to the approach in [18], [29], where the case $k = 1$ was discussed.

Recall that $\mathscr{H}$ denotes the collection of symmetric matrices in $\mathbb{R}^{d \times d}$. For $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$, we define the operator

$$\mathcal{F}_n : \mathscr{H} \to \mathbb{R}^n, \quad X \mapsto \frac{d}{k}(\langle X, P_{V_j} \rangle)_{j=1}^n. \tag{21}$$

For $x \in \mathbb{R}^d$, let $f := \frac{d}{k}(\|P_{V_j}(x)\|^2)_{j=1}^n = \mathcal{F}_n(xx^*) \in \mathbb{R}^n$, and we now aim to reconstruct $\pm x$ from f. By assuming that the union of the subspaces $\{V_j\}_{j=1}^n$ spans $\mathbb{R}^d$, clearly, xx^* is a solution of

$$\min_{X \in \mathscr{H}} (\mathrm{rank}(X)), \quad \text{subject to} \quad \mathcal{F}_n(X) = f, \ X \succeq 0. \tag{22}$$

The notation $X \succeq 0$ stands for X being positive semidefinite. Rank minimization is in general NP-hard, and in convex optimization it is standard to replace (22) with

$$\min_{X \in \mathscr{H}} (\mathrm{trace}(X)), \quad \text{subject to} \quad \mathcal{F}_n(X) = f, \ X \succeq 0, \tag{23}$$

a semidefinite program, for which efficient algorithms based on interior point methods are available. The NEOS Server [26] provides online solvers for semidefinite programs. We know that the solution of (22) has rank 1, so there is more structure to it and, as in [29], we can consider the underlying feasibility problem, i.e.,

$$\text{find } X \in \mathscr{H}, \quad \text{subject to} \quad \mathcal{F}_n(X) = f,\ X \succeq 0. \tag{24}$$

For $k = 1$, there is a constant $c > 0$, such that the random choice of at least cd subspaces yields that, with high probability, xx^* is the only solution to (24), i.e., the only feasible point of (22) and (23), cf. [17], [18], [29]. Here, we extend the result to $k > 1$:

Theorem V.1. *There are constants $c_1, c_2 > 0$ such that, if $n \geq c_1 d$ and $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ are chosen independently identically distributed according to σ_k, then, for all $x \in \mathbb{R}^d$, the matrix xx^* is the unique solution to* (24) *with probability at least $1 - e^{-c_2 n}$.*

Note that the probability of exact recovery in Theorem V.1 holds simultaneously over all input signals $x \in \mathbb{R}^d$, and the constants are independent of the ambient dimension d but may depend on the subspace dimension k.

To verify Theorem V.1, we shall first derive deterministic conditions serving uniqueness in (24). Later, we shall verify that these conditions are satisfied with high probability when the subspaces are chosen in the appropriate random fashion. After having assembled all ingredients, the proof of Theorem V.1 is presented in Appendix E.

A simple rescaling allows us to restrict the considerations to $x \in S^{d-1}$. Let $T := T_x := \{xy^* + yx^* : y \in \mathbb{R}^d\} \subset \mathscr{H}$, and, for $Z \in \mathbb{R}^{d\times d}$, denote Z_T its orthogonal projection onto T and $Z_{T^\perp}$ its orthogonal projection onto the orthogonal complement of T. The term $\|\cdot\|_1$ denotes the nuclear norm and $\|\cdot\|_\infty$ the operator norm:

Theorem V.2. *Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ and $f = (\|P_{V_j}(x)\|^2)_{j=1}^n$. Assume that $0 < A, B$ and $\gamma < A/B$ are fixed numbers, such that the following three points are satisfied:*

(a) *For all positive semidefinite matrices $X \in \mathscr{H}$,*

$$\frac{1}{n}\|\mathcal{F}_n(X)\|_{\ell_1} \leq B\|X\|_1. \tag{25}$$

(b) *For all $X \in T$,*

$$A\|X\|_\infty \leq \frac{1}{n}\|\mathcal{F}_n(X)\|_{\ell_1}. \tag{26}$$

(c) *There exists Y in the range of $\mathcal{F}_n^*$ such that*

$$\|Y_T\|_1 \leq \gamma, \qquad Y_{T^\perp} \succeq I_{T^\perp}. \tag{27}$$

Then xx^* is the unique solution to (24).

The matrix Y in (27) was called a dual certificate in [18]. To verify Theorem V.2, we can straightforwardly follow the lines of the proof in [17], [29] while keeping track of the constants, see also [18], [22]. The complete proof is in Appendix A.

Remark V.3. If $\{V_j\}_{j=1}^n$ is a design of strength 4, then the conditions in Theorem V.2 can be satisfied. Indeed, we can choose $B = 1$ and there is $A_k > 0$ satisfying (26) that is even allowed to depend on d in this case. Since $\mathcal{F}_n^*$ is onto, the certificate $Y = 2I - 2P_x$ is admissible and γ can be zero.

In the subsequent sections, we shall verify that the conditions of Theorem V.2 are satisfied with high probability when the subspaces $\{V_j\}_{j=1}^n$ are selected at random.

A. Nuclear norm estimates on $\|\mathcal{F}_n(X)\|_{\ell_1}$ for $X \succeq 0$

We shall verify that $\mathcal{F}_n$ is close to an isometry with high probability:

Theorem V.4. *Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ be independently chosen random subspaces with identical distribution σ_k. For $0 < r < 1$ fixed, there are constants $c(r), C(r) > 0$, such that, for all positive semidefinite matrices X and $n \geq c(r)d$,*

$$(1-r)\|X\|_1 \leq \frac{1}{n}\|\mathcal{F}_n(X)\|_{\ell_1} \leq (1+r)\|X\|_1 \tag{28}$$

holds with probability at least $1 - e^{-C(r)n}$.

By using the spectral decomposition of X, we see that condition (28) is equivalent to

$$(1-r)\frac{nk}{d}\|x\|^2 \leq \sum_{j=1}^n \|P_{V_j}(x)\|^2 \leq (1+r)\frac{nk}{d}\|x\|^2, \text{ for all } x \in \mathbb{R}^d.$$

In other words, $\{V_j\}_{j=1}^n$ is a fusion frame that is not too far from being tight. It turns out that we can follow the lines in [18] to prove Theorem V.4 after having established some analogy between $k = 1$ and $k > 1$. If $k = 1$, the random variable $d\|P_V(x)\|^2$, for $x \in S^{d-1}$, is sub-exponential. We can verify the analogue result for $k > 1$:

Lemma V.5. *If V is a random subspace distributed according to σ_k on $\mathcal{G}_{k,d}$, then, for any $x \in S^{d-1}$,*

$$\sup_{p\geq 1} p^{-1}\big(\mathbb{E}(\frac{d}{k}\|P_V(x)\|^2)^p\big)^{1/p} \leq 1. \tag{29}$$

Proof of Lemma V.5. The distribution of $\|P_V(x)\|^2$ does not depend on the particular choice of $x \in S^{d-1}$ and is beta distributed with parameters $(\frac{k}{2}, \frac{d-k}{2})$. Thus, its moments are given by

$$\mathbb{E}\|P_V(x)\|^{2p} = \frac{(k/2)_p}{(d/2)_p} = \frac{k(k+2)\cdots(k+2p-2)}{d(d+2)\cdots(d+2p-2)}, \tag{30}$$

which coincide with the tight p-fusion frame bounds (6) when the weights are constant. An induction over p yields (29). □

Note that Lemma V.5 says that $\frac{d}{k}\|P_V(x)\|^2$ is a sub-exponential random variable with a bound in (29) that does not depend on d. The latter is one of the main ingredients to verify Theorem V.4 along the lines in [18], see Appendix B for the details.

B. Operator norm estimates on $\|\mathcal{F}_n(X)\|_{\ell_1}$ for symmetric rank-2 matrices

We shall verify the condition (26):

Theorem V.6. *Let k be fixed. There is a constant $u > 0$ such that, for $0 < r < 1$ fixed, there exist constants $c, C > 0$, such that, for all $n \geq cd$ and $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ independently*

chosen random subspaces with identical distribution σ_k, the inequality

$$\frac{1}{n}\|\mathcal{F}_n(X)\|_{\ell_1} \geq u(1-r)\|X\|_\infty,$$

for all symmetric rank-2 matrices X, holds with probability at least $1-e^{-Cn}$.

Note that the probability in the estimate in Theorem V.6 is uniform in X. The proof of Theorem V.6 is based on the following Proposition that was derived for $k=1$ in [18]. Our proof for $k>1$ is original:

Proposition V.7. *Let k be fixed and $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ be independently chosen random subspaces with identical distribution σ_k. There is a constant $u>0$ such that, for all $-1 \leq t \leq 1$ and $z_1, z_2 \in S^{d-1}$ with $z_1 \perp z_2$,*

$$\frac{d}{k}\mathbb{E}\big|\|P_V(z_1)\|^2 - t\|P_V(z_2)\|^2\big| \geq u.$$

Proof. The sphere is two-point homogeneous and σ_k is invariant under orthogonal transformation so that we can restrict the analysis to the first two canonical basis vectors e_1 and e_2. Since the integral is always nonzero, we only need to take care of the limit $d\to\infty$. We first see that

$$\begin{aligned}\frac{d}{k}&\mathbb{E}\big|\|P_V(e_1)\|^2 - t\|P_V(e_2)\|^2\big| \\ &= \frac{d}{k}\int_{\mathcal{G}_{k,d}} \big|\|P_V(e_1)\|^2 - t\|P_V(e_2)\|^2\big| d\sigma_k(V) \\ &= \frac{d}{k}\int_{\mathcal{V}_{2,d}} \big|\sum_{i=1}^k m_{i,1}^2 - t\sum_{i=1}^k m_{i,2}^2\big| d\nu_2(M),\end{aligned}$$

where $\mathcal{V}_{2,d} = \{M=(m_{i,j}) \in \mathbb{R}^{d\times 2} : M^*M = I\}$ denotes the Stiefel-manifold endowed with the standard probability measure ν_2. If M is a random matrix, distributed according to ν_2, then, according to [31, Proposition 7.5], the upper $k\times 2$ block of M multiplied by d converges in distribution (for $d\to\infty$) towards a random $k\times 2$ matrix whose entries are standard normal i.i.d.. Let us denote the underlying probability measure on $\mathbb{R}^{k\times 2}$ by $\mathcal{N}(0, I_k\otimes I_2)$. The convergence in distribution implies that, for $d\to\infty$,

$$\begin{aligned}d&\mathbb{E}\big|\|P_V(e_1)\|^2 - t\|P_V(e_2)\|^2\big| \\ &\longrightarrow \int_{\mathbb{R}^{k\times 2}} \big|\|N(e_1)\|^2 - t\|N(e_2)\|^2\big| d\mathcal{N}(0, I_k\otimes I_2)(N).\end{aligned}$$

Since the right-hand side is bigger than 0, for all $-1\leq t\leq 1$, compactness and continuity arguments suffice to conclude the proof. □

For the complete proof of Theorem V.6 that is based on Proposition V.7, we refer to Appendix C.

C. The dual certificate Y

To derive the dual certificate Y, we can follow the ideas in [17], [18], [29] adapted to $k>1$. We will use Proposition III.1 from the deterministic setting and the Remark III.2. Let $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ be independently chosen random subspaces with identical distribution σ_k. The choice

$$Y_1 := 2I - 2P_x$$

would satisfy both conditions in (27) but may not lie in the range of

$$\mathcal{F}_n^* : \mathbb{R}^n \to \mathscr{H}, \qquad (\lambda_j)_{j=1}^n \mapsto \frac{d}{k}\sum_{j=1}^n \lambda_j P_{V_j}.$$

Thus, we aim to determine an appropriate sequence $(\lambda_j)_{j=1}^n$ such that $\frac{d}{k}\sum_{j=1}^n \lambda_j P_{V_j}$ "approximates" Y_1. First, let us rewrite

$$Y_1 = (k+2)I - (2P_x + kI).$$

For $a := \frac{2d(d-k)}{(d+2)(d-1)}$ and $b := \frac{d(kd+k-2)}{(d+2)(d-1)}$, we observe that $a\to 2$ and $b\to k$ when d tends to infinity, so that we can approximate Y_1 by

$$Y_2 := (k+2)I - (aP_x + bI).$$

Since Proposition III.1 implies

$$aP_x + bI = \frac{d^2}{k}\mathbb{E}\|P_V(x)\|^2 P_V \tag{31}$$

and $\frac{d}{k}\mathbb{E}P_V = I$ holds, we obtain

$$Y_2 = \frac{d}{k}\mathbb{E}((k+2-d\|P_V(x)\|^2)P_V).$$

The sample mean converges towards the population mean, so

$$Y_3 := \frac{d}{nk}\sum_{j=1}^n ((k+2-d\|P_{V_j}(x)\|^2)P_{V_j})$$

approximates Y_2, and we observe that Y_3 lies in the range of $\mathcal{F}_n^*$. In view of tail bound estimates, it will be advantageous to use an additional cut-off similar to the one in [18]: keeping in mind that (30) yields $\frac{d^2}{k}\mathbb{E}\|P_{V_j}(x)\|^4 \to k+2$, when d tends to infinity, we define the dual certificate by

$$Y := \frac{d}{nk}\sum_{j=1}^n \lambda_j P_{V_j}, \quad \text{where} \quad \lambda_j = \alpha - d\|P_{V_j}(x)\|^2 1_{E_j}, \tag{32}$$

$\alpha = \frac{d^2}{k}\mathbb{E}(\|P_{V_j}(x)\|^4 1_{E_j})$, and $E_j = \{\sqrt{\frac{d}{k}}\|P_{V_j}(x)\| \leq 2\beta_\gamma\}$ for some constant $\beta_\gamma > 0$. Obviously, Y is in the range of $\mathcal{F}_n^*$ and, as outlined above, can be considered as an approximation to $Y_1 = 2I - 2P_x$.

The above definitions will be used throughout the remaining part of this paper.

1) Dual certificate: Y_T: We shall verify that the dual certificate defined by (32) satisfies the first condition in (27). The following theorem is the analogy to [29, Lemma 1] and [17, Lemma 2.3]:

Theorem V.8. *Let $x\in S^{d-1}$ be fixed. There are constants $c, C>0$ such that, for $n\geq cd$,*

$$\|Y_T\|_1 \leq \gamma \tag{33}$$

with probability at least $1-e^{-Cn}$.

Proof. First, we suppose that $x=e_1$ and take care of the general case later. We observe that $\|Y_T\|_1 \leq \sqrt{2}\|Y_T\|_{HS} \leq 2\sqrt{2}\|y\|^2)$, where $y\in\mathbb{R}^d$ is the first column of Y and $\|\cdot\|_{HS}$

denotes the Frobenius norm. We split $P_{V_j} = Q_j Q_j^*$, such that $Q_j \in \mathbb{R}^{d\times k}$ with orthonormal columns. By using

$$Z = \sqrt{\frac{d}{k}}(Q_1, \dots, Q_n) \in \mathbb{R}^{d\times kn},\ h = \sqrt{\frac{d}{k}} \begin{pmatrix} \lambda_1 Q_1^* e_1 \\ \vdots \\ \lambda_n Q_n^* e_1 \end{pmatrix} \in \mathbb{R}^{kn},$$

and $h_j = \sqrt{\frac{d}{k}}\lambda_j Q_j^* e_1 \in \mathbb{R}^k$, for $j = 1, \dots, n$, we see that $\|y\|^2 = \frac{1}{n^2}\|Zh\|^2$. According to Lemma V.5, $\|h_j\|^2 = \lambda_j^2 \frac{d}{k}\|P_{V_j}(e_1)\|^2$ is sub-exponential, and [49, Corollary 5.17] implies

$$\mathbb{P}\big(\|h\|^2 - \mathbb{E}\|h\|^2 \geq n\big) \leq 2e^{-C_1 n}, \tag{34}$$

for some constant $C_1 > 0$. Since $\alpha \leq k+2$, we observe that there is a constant $C_2 > 0$ such that $\mathbb{E}\|h\|^2 \leq C_2 n$. Thus, the above estimate (34) implies that there is a constant C_3 such that

$$\mathbb{P}\big(\|h\|^2 \geq n\big) \leq 2e^{-C_3 n}. \tag{35}$$

For $n > \log(2)/C_3$, the factor 2 can be put into a constant in the exponential, say $C > 0$.

For $q \in \mathbb{R}^{kn}$ with $\|q\| = 1$ and $q = (q_j)_{j=1}^n$, where $q_j \in \mathbb{R}^k$, we obtain

$$\|Zq\|^2 \leq \frac{d}{k}\sum_{j=1}^n \|Q_j q_j\|^2 = \frac{d}{k}\sum_{j=1}^n \|q_j\|^2 = d/k, \tag{36}$$

where we have used that the columns of Q_j are orthonormal. By combining (35) with (36), we obtain

$$\|y\|^2 = \frac{1}{n^2}\|Zh\|^2 = \frac{1}{n^2}\|h\|^2 \|Z\frac{h}{\|h\|}\|^2 \leq \frac{d}{kn}$$

with probability at least $1 - e^{-Cn}$. Thus, for sufficiently large $c > 0$, the condition $n \geq cd$ implies (33).

To conclude the proof, we need to allow general $x \in S^{d-1}$. Note that there exists an orthogonal matrix U such that $x = Ue_1$. We observe that $T_x = UT_{e_1}U^*$ and $P_{UV_j} = U^* P_{V_j} U$. Therefore, the definition $Y_x := UY_{e_1}U^*$, where Y_{e_1} is the dual certificate w.r.t. e_1, is in the range of the map $\tilde{\mathcal{F}}_n$ that corresponds to $\{UV_j\}_{j=1}^n$. The latter subspaces are also i.i.d. according to σ_k. Since $(Y_x)_{T_x} = UY_{T_{e_1}}U^*$, we also derive $\|(Y_x)_{T_x}\|_1 = \|Y_{T_{e_1}}\|_1$.

□

2) Dual certificate: $Y_{T^\top}$: Let us verify that the dual certificate Y in (32) satisfies the second condition in (27). Indeed, we prove a slightly stronger result:

Theorem V.9. *Let $x \in S^{d-1}$ be fixed. For all $0 < \varepsilon < 1/2$, there is $\delta \geq 3/2$ and constants $c, C > 0$ such that, for $n \geq cd$,*

$$\|Y_{T^\perp} - \delta I_{T^\perp}\|_\infty \leq \varepsilon \tag{37}$$

with probability at least $1 - e^{-Cn}$.

Note that (37) implies $Y_{T^\perp} \succeq I_{T^\perp}$. The proof follows the analogous results in [29, Lemma 2] and [17, Lemma 2.3], where $k = 1$ is addressed, see our Appendix D for the details.

D. Proof of Theorem V.1

After having generalized the intermediate results from $k = 1$ to the general case $k \geq 1$, we can assemble these findings as in [17], [18], [29] to prove Theorem V.1. The details are presented in Appendix E.

Remark V.10. Note that our proof for involving semidefinite programming in the phase retrieval problem is guided by the ideas in [18], [17], [29]. Meanwhile the golfing scheme as originally proposed in [40] has been used for constructing dual certificates in rank-1 phase retrieval with semidefinite programming enabling a partial derandomization [41], so that $\sigma_{1,d}$ can be replaced with a probability distribution of smaller support. However, signal recovery probability decreases as well. Analogous results also hold for rank-k phase retrieval [32]. It is worth mentioning that the golfing scheme was also used to treat phase retrieval with sparse signals [44] and with coded diffraction patterns [19].

E. Stability

In many applications of interest, we may have access to the exact subspaces $\{V_j\}_{j=1}^n$ but the actual measurements are noisy, so that we need to reconstruct the signal from observations of the form

$$f_j = \|P_{V_j}(x)\|^2 + \omega_j, \quad j = 1, \dots, n, \tag{38}$$

where ω_j is some distortion term. If we replace the feasibility problem of the semi-definite program with the constrained ℓ_1-minimization

$$\arg\min_{X \in \mathscr{H}} \|\mathcal{F}_n(X) - f\|_{\ell_1}, \quad \text{subject to} \quad X \succeq 0, \tag{39}$$

then we obtain the same stability properties as in [17]. Indeed, we can straightforwardly follow the lines of the proof in [17, Theorem 1.3] for $k = 1$ to derive our next statement, which covers $k \geq 1$:

Theorem V.11. *There are constants $c_0, c_1, c_2 > 0$ such that, if $n \geq c_1 d$ and $\{V_j\}_{j=1}^n \subset \mathcal{G}_{k,d}$ are chosen independently identically distributed according to σ_k, then, for all $x \in \mathbb{R}^d$ and f given by* (38)*, the solution $\hat{X}$ to* (39) *obeys*

$$\|\hat{X} - xx^*\|_{HS} \leq c_0 \frac{\|\omega\|_{\ell_1}}{n} \tag{40}$$

with probability at least $1 - e^{-c_2 n}$.

It was also pointed out in [17] that (40) implies

$$\min\big(\|\hat{x} - x\|, \|\hat{x} + x\|\big) \leq c_0 \min\big(\|x\|, \frac{\|\omega\|_{\ell_1}}{n\|x\|}\big),$$

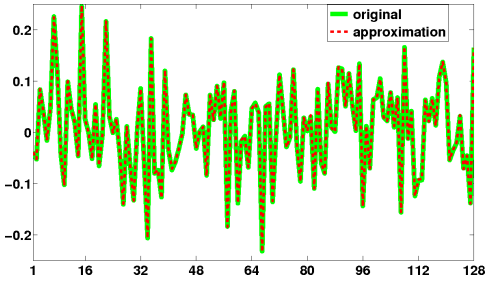
where $\hat{x} = \sqrt{\alpha}x_0$ and α is the largest eigenvalue of $\hat{X}$ with normalized eigenvector x_0. Hence, we also have a bound on the deviation to the exact signal when the measurements are noisy and $k > 1$.

VI. Numerical experiments

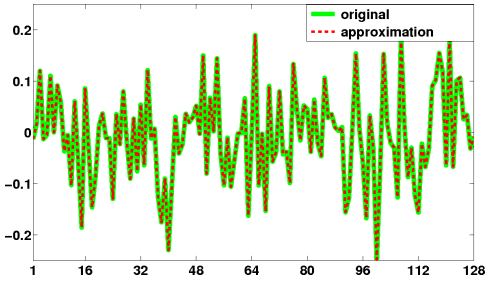
We shall present some numerical experiments illustrating Theorem V.1 and the choice of k. Let $x \in S^{d-1}$ and observe that $V \in \mathcal{G}_{k,d}$ is uniformly distributed if and only if $P_V = Z(Z^*Z)^{-1}Z^*$ for some $Z \in \mathbb{R}^{d\times k}$ with independent standard normal entries, cf. [23, Theorem 2.2.2]. Thus, we can easily generate pseudo-random orthogonal projectors $\{P_{V_j}\}_{j=1}^n$. Since $\|P_{V_j^\perp}(x)\|^2 + \|P_{V_j}(x)\|^2 = 1$, we shall restrict us to $k \le d/2$. We follow the numerical experiments in [18], where the measurement vector is $f = (\|P_{V_j}(x)\|^2)_{j=1}^n$. As in [18], we use the software package Templates for First-Order Conic Solvers (TFOCS) [11]. If $\hat{X}$ is the solution, then we define $\pm\hat{x} \in S^{d-1}$ as the normalized eigenvector corresponding to the largest eigenvalue of $\hat{X}$. If x is not supposed to lie on the sphere, then we can use the largest eigenvalue to rescale the normalized eigenvector.

A. Examples of signal reconstruction

We illustrate Theorem V.1 by following a numerical test from [16]. As in [16] for $k = 1$, the computed approximation is visually indistinguishable from the test signal when $k = 10$ and $k = 20$, where $d = 128$ and $n = 6d$, cf. Fig. 1.



(a) $k = 10$

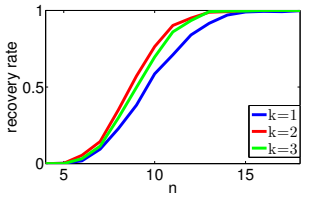


(b) $k = 20$

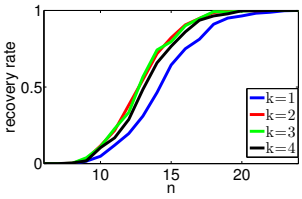
Fig. 1. We choose the original signal x uniformly distributed on the sphere S^{d-1}. As in [16], where $k = 1$ was used, the approximation is computed for $d = 128$ and $n = 6d$. Here, also for $k > 1$, we see that original and computed signal are visually indistinguishable.

B. Optimal choice of k

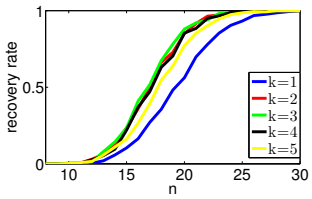
We investigate on the optimal choice of k. Indeed, for $d = 6, 8, 10, 12$, we check on the reconstruction rate in dependence of the number of subspaces n when k varies between 1 and $d/2$. We see in Figure 2 that, for small n, the proposed algorithm yields higher recovery rates when k is selected bigger than 1, and the choice $k = \lceil d/4 \rceil$ appears to be optimal. Here, the recovery rate is computed as the number of reconstructions deviating less than 10^{-2} from the original signal divided by the number of repeats (1000).



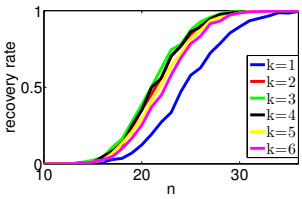
(a) $d = 6$



(b) $d = 8$



(c) $d = 10$



(d) $d = 12$

Fig. 2. When the subspace number n is small but the subspace dimension k can be selected freely, then $k = 1$ is clearly not the optimal choice. It appears that $k = \lceil d/4 \rceil$ yields the best results.

VII. Brief outline of the complex case

If we deal with complex signals $x \in \mathbb{C}^d$ and complex k-dimensional subspaces $\{V_j\}_{j=1}^n$, then there is again a canonical notion of cubature, cf. [46], and the complex analogue of Proposition III.1 holds with adjusted constants $a_1 = \frac{(d-1)d(d+1)}{k(d-k)}$ and $a_2 = \frac{kd-1}{d-k}$.

For random subspaces, Theorem V.1 can also be derived in the complex setting. The underlying Theorem V.2 holds the same way for complex signals and subspaces, so that we need to verify the respective conditions as in the real case. If the subspaces are chosen i.i.d. from the Haar measure on the complex Grassmann space, then $\frac{d}{k}\|P_{V_j}(x)\|^2$ is unitarily invariant in x and sub-exponential since $\mathbb{E}\|P_{V_j}(x)\|^{2p} = \frac{(k)_p}{(d)_p}$, cf. [5]. Thus, the analogue of Lemma V.5 holds. Proposition V.7 can be extended to the complex case, because the underlying result from [31, Proposition 7.5] has a complex version too. The formula (31) still holds, only the constants a and b need adjustments, so that the dual certificate Y can be defined the same way as in (32). Thus, we can follow the same proof strategy to cover the complex phase retrieval problem.

Acknowledgements

The authors would like to thank Thomas Bauer for valuable advice on the proof of Proposition IV.1 and Pierre Thibault for discussions on diffraction imaging. The authors are also thankful to Frank Vallentin and Christian Reiher for discussions on the proof of Theorem V.1. M. E. has been funded by the Vienna Science and Technology Fund (WWTF) through project VRG12-009.

References

[1] C. Bachoc, *Designs, groups and lattices*, J. Theor. Nombres Bordeaux (2005), 25–44.

[2] ———, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52** (2006), no. 5, 2111–2125.
[3] C. Bachoc, E. Bannai, and R. Coulangeon, *Codes and designs in Grassmannian spaces*, Discrete Mathematics **277** (2004), 15–28.
[4] C. Bachoc, R. Coulangeon, and G. Nebe, *Designs in Grassmannian spaces and lattices*, J. Algebraic Combinatorics **16** (2002), 5–19.
[5] C. Bachoc and M. Ehler, *Tight p-fusion frames*, Appl. Comput. Harmon. Anal. **35** (2013), no. 1, 1–15.
[6] S. Bahmanpour, J. Cahill, P. G. Casazza, J. Jasper, and L. M. Woodland, *Phase retrieval and norm retrieval*, arXiv:1409.8266.
[7] R. Balan, B. G. Bodmann, P. G. Casazza, and D. Edidin, *Painless reconstruction from magnitudes of frame coefficients*, J. Fourier Anal. Appl. **15** (2009), no. 4, 488–501.
[8] R. Balan, P. Casazza, and D. Edidin, *On signal reconstruction without phase*, Appl. Comput. Harmon. Anal **20** (2006), 345–356.
[9] E. Bannai and R. Damerell, *Tight spherical designs I*, J. Math. Soc. Japan **31** (1979), 199–207.
[10] H. H. Bauschke, P. L. Combettes, and D. R. Luke, *Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization*, J. Opt. Soc. Amer. A **19** (2002), 1334–1345.
[11] S. Becker, E. J. Candès, and M. Grant, *Templates for convex cone problems with applications to sparse signal recovery*, Mathematical Programming Computation **3** (2011), no. 3, 165–218.
[12] J. Bochnak, M. Coste, and M. Roy, *Real algebraic geometry*, Springer-Verlag, 1998.
[13] B. G. Bodmann and V. I. Paulsen, *Frames, graphs and erasures*, Lin. Alg. Appl. **404** (2005), 118–146.
[14] P. T. Boufounos and R. G. Baraniuk, *1-bit compressive sensing*, Proceedings of Conference on Information Science and Systems (CISS), Princeton, NJ, 2008.
[15] J. Cahill, P. G. Casazza, J. Peterson, and L. Woodland, *Phase retrieval by projections*, arXiv:1305.6226v3 (2013).
[16] E. J. Candès, Y. Eldar, T. Strohmer, and V. Voroninski, *Phase retrieval via matrix completion*, arXiv:1109.0573v2 (2011).
[17] E. J. Candès and X. Li., *Solving quadratic equations via PhaseLift when there are about as many equations as unknowns*, Foundations of Computational Mathematics **14** (2014), 1017–1026.
[18] E. J. Candès, T. Strohmer, and V. Voroninski, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, DOI:10.1002/cpa.21432 **66** (2013), no. 8, 1241–1274.
[19] E. J. Candѐs, X. Li, and M. Soltanolkotabi, *Phase retrieval from coded diffraction patterns*, to appear in Appl. Comput. Harmon. Anal. (2014).
[20] P. G. Casazza, G. Kutyniok, and S. Li, *Fusion frames and distributed processing*, Appl. Comput. Harmon. Anal. **25** (2008), no. 1, 114–132.
[21] A. Chai, M. Moscoso, and G. Papanicolaou, *Array imaging using intensity-only measurements*, Inverse Problems **27** (2011), no. 015005.
[22] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, *The convex geometry of linear inverse problems*, arXiv:1012.0621v3 (2012).
[23] Y. Chikuse, *Statistics on special manifolds*, Lecture Notes in Statistics, Springer, New York, 2003.
[24] A. M. Cohen, H. Cuypers, and H. Sterk (eds.), *Some tapas of computer algebra*, Algorithms and Computation in Mathematics, vol. 4, Springer, 1999.
[25] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer, 1996.
[26] J. Czyzyk, M. Mesnier, and J. Moré, *The NEOS server*, IEEE Computational Science & Engineering **5** (1998), no. 3, 68–75.
[27] P. de la Harpe and C. Pache, *Cubature formulas, geometrical designs, reproducing kernels, and Markov operators*, Infinite groups: geometric, combinatorial and dynamical aspects (Basel), vol. 248, Birkhäuser, 2005, pp. 219–267.
[28] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.
[29] L. Demanet and P. Hand, *Stable optimizationless recovery from phaseless linear measurements*, J. Fourier Anal. Appl. **20** (2014), 199–221.
[30] J. Drenth, *Principles of protein x-ray crystallography*, Springer, 2010.
[31] M. L. Eaton, *Group invariance applications in statistics*, Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, 1989.
[32] M. Ehler, M. Gräf, and F. Kiraly, *Phase retrieval using cubatures of positive semidefinite matrices*, arXiv (2015).
[33] M. Ehler and K. Okoudjou, *Minimization of the probabilistic p-frame potential*, J. Stat. Plann. Inference **142** (2012), no. 3, 645–659.
[34] P. Elias, *Error-correcting codes for list decoding*, IEEE Trans. Inform. Theory **37** (1991), no. 1, 5–12.
[35] V. Elser and R. P. Millane, *Reconstruction of an object from its symmetry-averaged diffraction pattern*, Acta Crystallographica Section A **64** (2008), no. 2, 273–279.
[36] J. R. Fienup, *Phase retrieval algorithms: a comparison*, Applied Optics **21** (1982), no. 15, 2758–2769.
[37] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, Springer-Verlag, 1984.
[38] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik **35** (1972), no. 2, 237–246.
[39] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*, Encyclopedia of Mathematics and its Applications, vol. 68, Cambridge University Press, 1998.
[40] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Trans. Inform. Theory **57** (2011), 1548–1566.
[41] D. Gross, F. Krahmer, and R. Kueng, *A partial derandomization of PhaseLift using spherical designs*, J. Fourier Anal. Appl. **21** (2015), no. 2, 229–266.
[42] R. B. Holmes and V. Paulsen, *Optimal frames for erasures*, Lin. Alg. Appl. **377** (2004), 31–51.
[43] G. Kutyniok, A. Pezeshki, R. Calderbank, and T. Liu, *Robust dimension reduction, fusion frames, and Grassmannian packings*, Appl. Comput. Harmon. Anal. **26** (2009), no. 1, 64–76.
[44] X. Li and V. Voroninski, *Sparse signal recovery from quadratic measurements via convex programming*, SIAM J. Math. Anal **45** (2013), no. 5, 3019–3033.
[45] Y. Plan and R. Vershynin, *One-bit compressed sensing by linear programming*, (2011).
[46] A. Roy, *Bounds for codes and designs in complex subspaces*, Journal of Algebraic Combinatorics **31** (2010), no. 1, 1–32.
[47] J. J. Seidel, *Definitions for spherical designs*, J. Statist. Plann. Inference **95** (2001), no. 1-2, 307–313.
[48] F. Sottile, *From enumerative geometry to solving systems of polynomial equations*, vol. 2, Springer-Verlag, 2001.
[49] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, Compressed sensing, Theory and Applications (Y. Eldar and G. Kutyniok, eds.), Cambridge University Press, 2012, pp. 210–268.
[50] I. Waldspurger, A. d'Aspremont, and S. Mallat, *Phase recovery, maxcut and complex semidefinite programming*, arXiv:1206.0102v2 (2012).

APPENDIX A
PROOF OF THEOREM V.2

Proof. For $Z \in \mathscr{H}$ being positive semi-definite and satisfying $\mathcal{F}_n(Z) = f$, we choose $H := Z - xx^*$ and aim to verify that $H = 0$. Since

$$0 = \mathcal{F}_n(H) = \mathcal{F}_n(H_T) + \mathcal{F}_n(H_{T^\perp}),$$

and $H_{T^\perp} = Z_{T^\perp}$ is positive semi-definite, the Conditions (26) and (25) yield

$$A\|H_T\|_\infty \leq \frac{1}{n}\|\mathcal{F}_n(H_T)\|_{\ell_1} = \frac{1}{n}\|\mathcal{F}_n(H_{T^\perp})\|_{\ell_1} \leq B\|H_{T^\perp}\|_1. \tag{41}$$

The range of $\mathcal{F}_n^*$ is orthogonal to the nullspace of $\mathcal{F}_n$, so that we derive

$$0 = \langle H, Y\rangle = \langle H_{T^\perp}, Y_{T^\perp}\rangle + \langle H_T, Y_T\rangle.$$

The left-hand inequality of (27) yields

$$0 \geq \langle H_{T^\perp}, Y_{T^\perp}\rangle - \gamma\|H_T\|_\infty,$$

and the right-hand inequality of (27) leads to $\|H_{T^\perp}\|_1 = \langle H_{T^\perp}, I_{T^\perp}\rangle \leq \langle H_{T^\perp}, Y_{T^\perp}\rangle$, so that we obtain

$$0 \geq \|H_{T^\perp}\|_1 - \gamma\|H_T\|_\infty \geq (\frac{A}{B} - \gamma)\|H_T\|_\infty,$$

where we have used (41). Thus, $H_T = 0$ must hold and hence also $H_{T^\perp} = 0$, so that we have $Z = xx^*$. □

APPENDIX B
PROOF OF THEOREM V.4

The following result extends findings on the smallest and largest singular values $s_{\min}(P)$ and $s_{\max}(P)$ of a random matrix P with independent sub-exponential rows in [49, Theorem 5.39]. Here, we consider independent blocks but there are dependent rows within each block:

Proposition B.1. *Let $P := \sqrt{\frac{d}{k}}\big(P_{V_1}, \dots, P_{V_n}\big)^* \in \mathbb{R}^{nd\times d}$, in which $\{V_j\}_{j=1}^n$ are identically and independently distributed according to σ_k on $\mathcal{G}_{k,d}$. Then, for every $t \geq 0$, we have with probability at least $1 - 2\exp(-ct^2)$*

$$\sqrt{n} - C\sqrt{d} - t \leq s_{\min}(P) \leq s_{\max}(P) \leq \sqrt{n} + C\sqrt{d} + t,$$

where $c, C > 0$ are absolute constant.

The proof of Proposition B.1 requires two lemmas for preparation:

Lemma B.2 ([49, Lemma 5.36])**.** *If $B \in \mathbb{R}^{n\times d}$ satisfies $\|B^*B - I\|_\infty \leq \max(\delta, \delta^2)$, for some $\delta > 0$, then*

$$1 - \delta \leq s_{\min}(B) \leq s_{\max}(B) \leq 1 + \delta. \tag{42}$$

Conversely, if B satisfies (42)*, then $\|B^*B - I\|_\infty \leq 3\max(\delta, \delta^2)$.*

An ε-net $\mathcal{N}_\varepsilon$ is a finite subset of S^{d-1} such that to any element $x \in S^{d-1}$, there is an element in $\mathcal{N}_\varepsilon$ at distance less than or equals ε.

Lemma B.3 ([49, Lemma 5.4])**.** *Let $A \in \mathbb{R}^{d\times d}$ be symmetric, and let $\mathcal{N}_\varepsilon$ be an ε-net of S^{d-1} for some $\varepsilon \in [0, \frac{1}{2})$. Then*

$$\|A\|_\infty = \sup_{x\in S^{d-1}} |\langle Ax, x\rangle| \leq (1-2\varepsilon)^{-1} \sup_{x\in\mathcal{N}_\varepsilon} |\langle Ax, x\rangle|.$$

Proof of Proposition B.1. To verify Proposition B.1, we want to apply Lemma B.2 with $B = \frac{1}{\sqrt{n}}P$. We shall explicitly derive the upper estimate on $s_{\max}(P)$. The lower estimate on $s_{\min}(P)$ follows from similar arguments. We must check that

$$\|\frac{1}{n}\frac{d}{k}\sum_{j=1}^n P_{V_j} - I\|_\infty \leq \max(\delta, \delta^2) =: \varepsilon, \text{ where } \delta = C\sqrt{\frac{d}{n}} + \frac{t}{\sqrt{n}}. \tag{43}$$

Let $\mathcal{N}$ be a $\frac{1}{4}$-net, so that an application of Lemma B.3 yields

$$\begin{aligned}\|\frac{1}{n}P^*P - I\|_\infty &= \|\frac{1}{n}\frac{d}{k}\sum_{j=1}^n P_{V_j} - I\|_\infty \\ &\leq 2\max_{x\in\mathcal{N}} |\langle(\frac{1}{n}\frac{d}{k}\sum_{j=1}^n P_{V_j} - I)x, x\rangle| \\ &= 2\max_{x\in\mathcal{N}} |\frac{1}{n}\|Px\|^2 - 1|.\end{aligned}$$

Thus, we must verify with the required probability that

$$\max_{x\in\mathcal{N}} |\frac{1}{n}\|Px\|^2 - 1| \leq \frac{\varepsilon}{2}.$$

To derive this estimate, we define random variables $Z_j = \sqrt{\frac{d}{k}}\|P_{V_j}(x)\|$ so that $\sum_{j=1}^n Z_j^2 = \|P(x)\|^2$. Since $\mathbb{E}(Z_j^2) = 1$, we can estimate

$$\begin{aligned}\|Z_j^2\|_{\psi_1} &:= \sup_{p\geq 1} p^{-1}(\mathbb{E}Z_j^{2p})^{1/p} \\ &= \frac{d}{k}\sup_{p\geq 1} p^{-1}(\mathbb{E}\|P_{V_j}(x)\|^{2p})^{1/p} \\ &= \frac{d}{k}\sup_{p\geq 1} p^{-1}(\frac{(k/2)_p}{(d/2)_p})^{1/p} \leq 1,\end{aligned}$$

where the last inequality is due to Lemma V.5. According to [49, Remark 5.18], $\|Z_j^2 - 1\|_{\psi_1} \leq 2$, and we obtain from the Bernstein type inequality [49, Corollary 5.17]

$$\begin{aligned}\mathbb{P}(|\frac{1}{n}\sum_{j=1}^n \frac{d}{k}\|P_{V_j}(x)\|^2 - 1| \geq \varepsilon/2) &= \mathbb{P}(|\frac{1}{n}\sum_{j=1}^n Z_j^2 - 1| \geq \varepsilon/2) \\ &\leq 2\exp(-cn\min(\frac{\varepsilon^2}{16}, \frac{\varepsilon}{4})) \\ &= 2\exp(-n\frac{c}{16}\delta^2) \\ &\leq 2\exp(-\frac{c}{16}(C^2d + t^2)),\end{aligned}$$

where the last line follows from (43). Since the net can be chosen such that $|\mathcal{N}| \leq 9^d$, cf. [49], we obtain

$$\begin{aligned}\mathbb{P}(\max_{x\in\mathcal{N}} |\frac{1}{n}\sum_{j=1}^n \frac{d}{k}\|P_{V_j}(x)\|^2 - 1| \geq \varepsilon/2) &\leq 9^d 2\exp(-\frac{c}{16}(C^2d + t^2)) \\ &\leq 2\exp(-\frac{c}{16}t^2),\end{aligned}$$

where we assume $C \geq 4\sqrt{\ln(9)/c}$. The latter does not cause any trouble because c is a constant independent of ε. This finally yields

$$\mathbb{P}(s_{\max}(P) \geq \sqrt{n} + C\sqrt{d} + t) \leq 2\exp(-\frac{c}{16}t^2).$$

The estimates on $s_{\min}(P)$ are derived analogously. □

We can now prove Theorem V.4:

Proof of Theorem V.4. Since any positive semidefinite matrix X can be written by means of its projectors on eigenspaces, it is sufficient to verify

$$1 - r \leq \frac{1}{n}\|\mathcal{F}_n(xx^*)\|_{\ell_1} \leq 1 + r, \quad \forall x \in S^{d-1},$$

in place of (28). We observe that $\|Px\|^2 = \|\mathcal{F}_n(xx^*)\|_{\ell_1}$, so that

$$s_{\min}^2(P) \leq \|\mathcal{F}_n(xx^*)\|_{\ell_1} \leq s_{\max}^2(P)$$

holds. First, we take care of the upper bound. According to Proposition B.1, we have

$$\frac{1}{n}\|\mathcal{F}_n(P_x)\|_{\ell_1} \leq \frac{1}{n}s_{\max}^2(P) \leq (1 + \frac{1}{\sqrt{n}}(C\sqrt{d} + t))^2,$$

with probability at least $1 - 2e^{-ct^2}$. Choose $\varepsilon > 0$ such that $r/4 = \varepsilon^2 + \varepsilon$ and observe that $\varepsilon \geq \frac{r}{5}$, so that $n \geq c_1 r^{-2}d$ with $c_1 = 25C^2$ implies $n \geq \varepsilon^{-2}C^2d$. For $t = \sqrt{n}\varepsilon$, we obtain that

$$\frac{1}{\sqrt{n}}s_{\max} \leq (1 + 2\varepsilon)$$

holds with probability at least $1-2e^{-cn\varepsilon^2}$. Hence, we have $\frac{1}{n}s_{\max}^2 \le (1+r)$ with the same probability. Since $\varepsilon^2 \ge r^2/25$, we can adjust c_1 such that $n \ge \frac{25}{cr^2}\ln(2)$ so that $c_2 > 0$ exists and the required upper estimate holds with probability $1-e^{-c_2 r^2 n}$, The lower estimate can be derived in an analogous way. □

APPENDIX C
PROOF OF THEOREM V.6

Proof of Theorem V.6. It is sufficient to consider $\|X\|_\infty = 1$, so that $X = P_{z_1} - tP_{z_2}$, where $z_1, z_2 \in S^{d-1}$ and $z_1 \perp z_2$ and $t \in [-1,1]$. We observe

$$\frac{1}{n}\|\mathcal{F}_n(X)\|_1 = \frac{1}{n}\sum_{j=1}^n \frac{d}{k}|\|P_{V_j}(z_1)\|^2 - t\|P_{V_j}(z_2)\|^2| = \frac{1}{n}\sum_{j=1}^n \xi_j,$$

where $\xi_j = \frac{d}{k}|\|P_{V_j}(z_1)\|^2 - t\|P_{V_j}(z_2)\|^2|$. Since $|t|$ is bounded, Lemma V.5 implies that ξ_j is sub-exponential. Therefore, the Bernstein inequality as stated in [49] yields

$$\mathbb{P}(|\frac{1}{n}\|\mathcal{F}_n(X)\|_1 - \mathbb{E}\xi| \ge \varepsilon) \le 2\exp(-cn\min(\frac{\varepsilon^2}{4}, \frac{\varepsilon}{2})),$$

where $c > 0$ is an absolute constant. Proposition V.7 yields $\mathbb{E}\xi_j \ge u$, and, for $\varepsilon < 2$, we derive

$$\frac{1}{n}\|\mathcal{F}_n(X)\|_1 \ge u - \varepsilon,$$

with probability at least $1-2\exp(-C_1 n\varepsilon^2)$, where $C_1 = c/4$. The choice $\varepsilon = ur$ establishes the required estimate at least for fixed $X \in T$ with probability at least $1-2\exp(-C_2 n r^2)$, where $C_2 = C_1 u^2$. The remaining part of the proof is the same covering argument as in [18], so we omit this. □

APPENDIX D
PROOF OF THEOREM V.9

Proof of Theorem V.9. As in the proof of Theorem V.8, we first consider $x = e_1$. Let us split $Y = Y^{(0)} - Y^{(1)}$ into

$$Y^{(0)} = \frac{1}{n}\sum_{j=1}^n \alpha\frac{d}{k}P_{V_j}, \qquad Y^{(1)} = \frac{1}{n}\sum_{j=1}^n d\|P_{V_j}(e)\|^2 1_{E_j}\frac{d}{k}P_{V_j}.$$

First, we shall estimate $\|Y^{(0)}_{T^\perp} - \alpha I_{T^\perp}\|_\infty$, later also $\|Y^{(1)}_{T^\perp} - b_0 I_{T^\perp}\|_\infty$ for some special number b_0. We observe that $\mathbb{E}Y^{(0)} = \alpha I$. By using $P := \sqrt{\frac{d}{k}}\big(P_{V_1}, \dots, P_{V_n}\big)^*$ as in Proposition B.1 and squaring the estimates there, we see that

$$(\sqrt{n} - C_1\sqrt{d} - t)^2 \le s_{\min}^2(P) \le s_{\max}^2(P) \le (\sqrt{n} + C_1\sqrt{d} + t)^2$$

with probability at least $1-2e^{-c_1 t^2}$. Since $\frac{\alpha}{n}P^*P = Y^{(0)}$, the latter implies at least for sufficiently small $t/\sqrt{n}$:

$$\|Y^{(0)} - \alpha I\|_\infty \le \alpha(C_1^2 d + t^2 + 2\sqrt{nd} + 2\sqrt{n}t + 2C_1 t\sqrt{d})$$

with the same probability. For all $\varepsilon_1 > 0$, there is c_2 sufficiently large and $\varepsilon_2 > 0$ sufficiently small such that $t = \varepsilon_2\sqrt{n}$ yields

$$\|Y^{(0)} - \alpha I\|_\infty \le \alpha\varepsilon_1,$$

for all $n \ge c_2 d$ with probability $1-e^{-c_3 n}$. In particular, we have

$$\|Y^{(0)}_{T^\perp} - \alpha I_{T^\perp}\|_\infty \le \alpha\varepsilon_1 \tag{44}$$

with the same probability.

Let us now take care of $Y^{(1)}_{T^\perp}$. Due to the unitary invariance of σ_k, (31) for $X = P_{e_1}$ yields

$$\mathbb{E}(d\|P_{V_j}(e_1)\|^2 1_{E_j}\frac{d}{k}P_{V_j}) = a_0 P_{e_1} + b_0 I,$$

for some constants $a_0, b_0 > 0$ that depend on β_γ. Therefore, we have $\mathbb{E}Y^{(1)}_{T^\perp} = b_0 I$. The random matrix

$$X_j = \frac{d^2}{k}\|P_{V_j}(e_1)\|^2 1_{E_j}(P_{V_j})_{T^\perp} - b_0 I_{T^\perp}$$

is bounded, say by K. We find a constant $C_2 > 0$ such that $\|\mathbb{E}X_j^* X_j\|_\infty \le C_2$ implying $\|\sum_{j=1}^n \mathbb{E}X_j^* X_j\|_\infty \le nC_1$. According to [49, Theorem 5.29], we have, for all $t > 0$,

$$\mathbb{P}\big(\|\frac{1}{n}\sum_{j=1}^n X_j\|_\infty \ge \frac{t}{n}\big) \le 2de^{\frac{-t^2/2}{nC_2 + Kt/3}}.$$

By choosing $\varepsilon_2 > 0$ and $t = \varepsilon_3 n$, we derive

$$\mathbb{P}\big(\|\frac{1}{n}\sum_{j=1}^n X_j\|_\infty \ge \varepsilon_2\big) \le 2de^{-c_4 n} \le e^{-c_5 n},$$

for all $n \ge c_6\ln(d)$. Thus, we obtain

$$\|Y^{(1)}_{T^\perp} - b_0 I_{T^\perp}\|_\infty \le \varepsilon_2, \tag{45}$$

with probability $1-e^{-c_5 n}$, for all $n \ge c_6\ln(d)$.

Combining (44) and (45) implies

$$\|Y_{T^\perp} - (\alpha - b_0)I_{T^\perp}\| \le \alpha\varepsilon_1 + \varepsilon_2$$

with probability at least $1-e^{-Cn}$, for all $n \ge cd$. We can now choose $\varepsilon_1, \varepsilon_2$ sufficiently small, such that $\alpha\varepsilon_1 + \varepsilon_2 \le \varepsilon$. The term α is bounded by $k+2$. According to Vershynin's lecture note on nonasymptotic random matrix theory (Lemma 9 in Lecture 4 on dimension reduction), we have, for all $\beta_\gamma \ge 1/2$ that $\mathbb{P}\big(E_j^c\big) \le 2e^{k/2}e^{-k\beta_\gamma}$. Since $\mathbb{E}(\frac{d^4}{k^2}\|P_{V_j}(e_1)\|^8)$ is bounded independently of d, see (30), the term $k+2-\alpha = \mathbb{E}(\frac{d^2}{k}\|P_{V_j}(e_1)\|^4 1_{E_j^c})$ can be made arbitrarily small by choosing β_γ sufficiently large. Thus, we can derive $\alpha \ge k + 5/3$. Similar arguments yield that b_0 gets closer to b when we increase β_γ. With $b \le k$ we can assume that $b_0 \le k + 1/6$, so that $\delta = \alpha - b_0 \ge 3/2$.

We still need to address general vectors $x \in S^{d-1}$. With the notation and arguments at the end of the proof of Theorem V.8, we observe that $\|(Y_x)_{T_x^\perp} - \delta I_{T_x^\perp}\|_\infty = \|(Y_{e_1})_{T_{e_1}^\perp} - \delta I_{T_{e_1}^\perp}\|_\infty$, which concludes the proof. □

APPENDIX E
PROOF OF THEOREM V.1

We can now assemble all of our findings to verify that the conditions in Theorem V.2 hold with the required probability:

Proof of Theorem V.1. We first fix $x \in S^{d-1}$. Then we choose $r \in (0,1)$ and $\gamma < u\frac{1-r}{1+r}$, where $u \in (0,1)$ as in Proposition V.7. Let c_i and C_i, $i = 1, \dots, 4$, be suitable positive constants. Theorem V.4 yields that Condition (25) holds with probability of failure at most $e^{-C_1 n}$, for all $n \ge c_1 d$. Theorem V.6 implies that Condition (26) holds with probability of failure at most $e^{-C_2 n}$, for all $n \ge c_2 d$. According to Theorem V.8, the first

condition in (27) holds with probability of failure at most $e^{-C_3 n}$, for all $n \geq c_3 d$. Theorem V.9 yields that the second condition in (27) is satisfied with probability of failure at most $e^{-C_4 n}$, for all $n \geq c_4 d$.

Finally, there are constants $c, C > 0$ such that, for all $n \geq cd$, we can estimate $\sum_{i=1}^{4} e^{-C_i n} \leq e^{-Cn}$, so that all conditions in Theorem V.2 are satisfied with probability at least $1 - e^{-Cn}$. In order to turn the latter into a uniform estimate in x, we take an ϵ-net $\mathcal{N}_\epsilon$ on the sphere of cardinality less or equals $(1 + \frac{2}{\epsilon})^d$, cf. [49, Lemma 5.2]. Since $(1 + \frac{2}{\epsilon})^d e^{-Cn} \leq e^{-\tilde{C}n}$, for all $n \geq \tilde{c}d$ when $\tilde{C}$ is sufficiently small and $\tilde{c}$ sufficiently large, we have a uniform estimate for the net $\mathcal{N}_\epsilon$. Now, to any arbitrary $x \in S^{d-1}$, we find $x_0 \in \mathcal{N}_\epsilon$ with $\|x - x_0\| \leq \epsilon$. By following the lines in [17, Proof of Theorem 1.2], one can derive that the certificate for x_0 also works for x, so that we can conclude the proof of Theorem V.1. □

REFERENCES

[1] C. Bachoc, *Designs, groups and lattices*, J. Theor. Nombres Bordeaux (2005), 25–44.

[2] ———, *Linear programming bounds for codes in Grassmannian spaces*, IEEE Trans. Inf. Th. **52** (2006), no. 5, 2111–2125.

[3] C. Bachoc, E. Bannai, and R. Coulangeon, *Codes and designs in Grassmannian spaces*, Discrete Mathematics **277** (2004), 15–28.

[4] C. Bachoc, R. Coulangeon, and G. Nebe, *Designs in Grassmannian spaces and lattices*, J. Algebraic Combinatorics **16** (2002), 5–19.

[5] C. Bachoc and M. Ehler, *Tight p-fusion frames*, Appl. Comput. Harmon. Anal. **35** (2013), no. 1, 1–15.

[6] S. Bahmanpour, J. Cahill, P. G. Casazza, J. Jasper, and L. M. Woodland, *Phase retrieval and norm retrieval*, arXiv:1409.8266.

[7] R. Balan, B. G. Bodmann, P. G. Casazza, and D. Edidin, *Painless reconstruction from magnitudes of frame coefficients*, J. Fourier Anal. Appl. **15** (2009), no. 4, 488–501.

[8] R. Balan, P. Casazza, and D. Edidin, *On signal reconstruction without phase*, Appl. Comput. Harmon. Anal **20** (2006), 345–356.

[9] E. Bannai and R. Damerell, *Tight spherical designs I*, J. Math. Soc. Japan **31** (1979), 199–207.

[10] H. H. Bauschke, P. L. Combettes, and D. R. Luke, *Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization*, J. Opt. Soc. Amer. A **19** (2002), 1334–1345.

[11] S. Becker, E. J. Candès, and M. Grant, *Templates for convex cone problems with applications to sparse signal recovery*, Mathematical Programming Computation **3** (2011), no. 3, 165–218.

[12] J. Bochnak, M. Coste, and M. Roy, *Real algebraic geometry*, Springer-Verlag, 1998.

[13] B. G. Bodmann and V. I. Paulsen, *Frames, graphs and erasures*, Lin. Alg. Appl. **404** (2005), 118–146.

[14] P. T. Boufounos and R. G. Baraniuk, *1-bit compressive sensing*, Proceedings of Conference on Information Science and Systems (CISS), Princeton, NJ, 2008.

[15] J. Cahill, P. G. Casazza, J. Peterson, and L. Woodland, *Phase retrieval by projections*, arXiv:1305.6226v3 (2013).

[16] E. J. Candès, Y. Eldar, T. Strohmer, and V. Voroninski, *Phase retrieval via matrix completion*, arXiv:1109.0573v2 (2011).

[17] E. J. Candès and X. Li., *Solving quadratic equations via PhaseLift when there are about as many equations as unknowns*, Foundations of Computational Mathematics **14** (2014), 1017–1026.

[18] E. J. Candès, T. Strohmer, and V. Voroninski, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Communications on Pure and Applied Mathematics, DOI:10.1002/cpa.21432 **66** (2013), no. 8, 1241–1274.

[19] E. J. Candès, X. Li, and M. Soltanolkotabi, *Phase retrieval from coded diffraction patterns*, to appear in Appl. Comput. Harmon. Anal. (2014).

[20] P. G. Casazza, G. Kutyniok, and S. Li, *Fusion frames and distributed processing*, Appl. Comput. Harmon. Anal. **25** (2008), no. 1, 114–132.

[21] A. Chai, M. Moscoso, and G. Papanicolaou, *Array imaging using intensity-only measurements*, Inverse Problems **27** (2011), no. 015005.

[22] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, *The convex geometry of linear inverse problems*, arXiv:1012.0621v3 (2012).

[23] Y. Chikuse, *Statistics on special manifolds*, Lecture Notes in Statistics, Springer, New York, 2003.

[24] A. M. Cohen, H. Cuypers, and H. Sterk (eds.), *Some tapas of computer algebra*, Algorithms and Computation in Mathematics, vol. 4, Springer, 1999.

[25] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, Springer, 1996.

[26] J. Czyzyk, M. Mesnier, and J. Moré, *The NEOS server*, IEEE Computational Science & Engineering **5** (1998), no. 3, 68–75.

[27] P. de la Harpe and C. Pache, *Cubature formulas, geometrical designs, reproducing kernels, and Markov operators*, Infinite groups: geometric, combinatorial and dynamical aspects (Basel), vol. 248, Birkhäuser, 2005, pp. 219–267.

[28] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geom. Dedicata **6** (1977), 363–388.

[29] L. Demanet and P. Hand, *Stable optimizationless recovery from phaseless linear measurements*, J. Fourier Anal. Appl. **20** (2014), 199–221.

[30] J. Drenth, *Principles of protein x-ray crystallography*, Springer, 2010.

[31] M. L. Eaton, *Group invariance applications in statistics*, Regional Conference Series in Probability and Statistics, Institute of Mathematical Statistics, 1989.

[32] M. Ehler, M. Gräf, and F. Kiraly, *Phase retrieval using cubatures of positive semidefinite matrices*, arXiv (2015).

[33] M. Ehler and K. Okoudjou, *Minimization of the probabilistic p-frame potential*, J. Stat. Plann. Inference **142** (2012), no. 3, 645–659.

[34] P. Elias, *Error-correcting codes for list decoding*, IEEE Trans. Inform. Theory **37** (1991), no. 1, 5–12.

[35] V. Elser and R. P. Millane, *Reconstruction of an object from its symmetry-averaged diffraction pattern*, Acta Crystallographica Section A **64** (2008), no. 2, 273–279.

[36] J. R. Fienup, *Phase retrieval algorithms: a comparison*, Applied Optics **21** (1982), no. 15, 2758–2769.

[37] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 2, Springer-Verlag, 1984.

[38] R. W. Gerchberg and W. O. Saxton, *A practical algorithm for the determination of the phase from image and diffraction plane pictures*, Optik **35** (1972), no. 2, 237–246.

[39] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*, Encyclopedia of Mathematics and its Applications, vol. 68, Cambridge University Press, 1998.

[40] D. Gross, *Recovering low-rank matrices from few coefficients in any basis*, IEEE Trans. Inform. Theory **57** (2011), 1548–1566.

[41] D. Gross, F. Krahmer, and R. Kueng, *A partial derandomization of PhaseLift using spherical designs*, J. Fourier Anal. Appl. **21** (2015), no. 2, 229–266.

[42] R. B. Holmes and V. Paulsen, *Optimal frames for erasures*, Lin. Alg. Appl. **377** (2004), 31–51.

[43] G. Kutyniok, A. Pezeshki, R. Calderbank, and T. Liu, *Robust dimension reduction, fusion frames, and Grassmannian packings*, Appl. Comput. Harmon. Anal. **26** (2009), no. 1, 64–76.

[44] X. Li and V. Voroninski, *Sparse signal recovery from quadratic measurements via convex programming*, SIAM J. Math. Anal **45** (2013), no. 5, 3019–3033.

[45] Y. Plan and R. Vershynin, *One-bit compressed sensing by linear programming*, (2011).

[46] A. Roy, *Bounds for codes and designs in complex subspaces*, Journal of Algebraic Combinatorics **31** (2010), no. 1, 1–32.

[47] J. J. Seidel, *Definitions for spherical designs*, J. Statist. Plann. Inference **95** (2001), no. 1-2, 307–313.

[48] F. Sottile, *From enumerative geometry to solving systems of polynomial equations*, vol. 2, Springer-Verlag, 2001.

[49] R. Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, Compressed sensing, Theory and Applications (Y. Eldar and G. Kutyniok, eds.), Cambridge University Press, 2012, pp. 210–268.

[50] I. Waldspurger, A. d'Aspremont, and S. Mallat, *Phase recovery, maxcut and complex semidefinite programming*, arXiv:1206.0102v2 (2012).

Christine Bachoc received the Ph.D. degree in 1989 and the Habilitationà Diriger des Recherches in 1994 from the Université de Bordeaux I, France. She has held a permanent research position of Centre National de la Recherche Scientifique from 1990 until 2002, then a professor position since 2002, from the Université de Bordeaux I. She is a member of the Institut de

Mathématiques de Bordeaux. Her research interests include number theory, coding theory, combinatorics, and optimization.

Martin Ehler received his doctoral degree in 2007 from the Philipps-Universität Marburg, Germany. He has held postdoctoral research positions at the National Institutes of Health, Bethesda MD, USA, and at the Helmholtz Zentrum München, Germany. Since 2013 he is the head of a Vienna Research Group for Young Investigators at the University of Vienna. His research interests are guided by high-dimensional signal analysis and include the fields of applied harmonic analysis and numerical analysis.