

# PHASE RETRIEVAL BY BINARY QUESTIONS: WHICH COMPLEMENTARY SUBSPACE IS CLOSER?

DYLAN DOMEL-WHITE AND BERNHARD G. BODMANN

**ABSTRACT.** Phase retrieval in real or complex Hilbert spaces is the task of recovering a vector, up to an overall unimodular multiplicative constant, from magnitudes of linear measurements. In this paper, we assume that the vector is normalized, but retain only qualitative, binary information about the measured magnitudes by comparing them with a threshold. In more specific, geometric terms, we choose a sequence of subspaces in a real or complex Hilbert space and only record whether a given vector is closer to the subspace than to the complementary subspace. The subspaces have half the dimension of the Hilbert space and are independent, uniformly distributed with respect to the action of the orthogonal or unitary groups. The main goal of this paper is to find a feasible algorithm for approximate recovery based on the information gained about the vector from these binary questions and to establish error bounds for its approximate recovery. We provide a pointwise bound for fixed input vectors and a uniform bound that controls the worst-case scenario among all inputs. Both bounds hold with high probability with respect to the choice of the subspaces. For real or complex vectors of dimension  $n$ , the pointwise bound requires  $m \geq C\delta^{-2}n \log(n)$  and the uniform bound  $m \geq C\delta^{-2}n^2 \log(\delta^{-1}n)$  binary questions in order to achieve an accuracy of  $\delta$ . The accuracy  $\delta$  is measured by the operator norm of the difference between the rank-one orthogonal projections corresponding to the normalized input vector and its approximate recovery.

## 1. INTRODUCTION

This paper is concerned with approximate phase retrieval from measuring qualitative, binary measurements. Phase retrieval is the task of recovering a vector in a real or complex Hilbert space up to an overall multiplicative unimodular constant from magnitudes of linear quantities. Motivated by applications from diffraction imaging [19, 26, 35], or from studying properties of the Fourier transform [2, 3], results on phase retrieval first focused on the case where measurements consist of magnitudes of linear functionals [7, 8, 17]. Phase retrieval with quantized measurements was studied as well [27, 30], see also the preceding works [1, 11, 29]. In this context, quantization means the magnitudes are replaced by values from a finite alphabet. Coarse, one-bit quantization represents the extreme case, for example when only qualitative information is obtained such as how each measured magnitude compares to a single given threshold. Another example in which coarsely quantized measurements appear is quantum state tomography, where the outcomes of experiments are recorded in order to estimate the state of a quantum system [21, 33]. In this case, the probability of an outcome is given by the squared norm of the projection of a (normalized) state vector onto a subspace associated with the outcome. Estimating this quantum state is then, up to the known normalization, equivalent to phase retrieval for the state vector. Phase retrieval based on norms of

---

2010 *Mathematics Subject Classification.* 94A12 (primary); 90C22, 15A83 (secondary).

*Key words and phrases.* Phase retrieval, coarse quantization, semidefinite program, measure concentration.

The research in this paper was supported in part by NSF grant DMS-1715735. Both authors would like to thank ICERM and HKUST for supporting a workshop where this project was started.

projections onto subspaces has also been studied outside of the context of quantum theory [6, 13, 16]. It may be viewed as a fusion-frame version of phase retrieval, where higher rank maps replace linear functionals and the norm replaces the absolute value. The recovery of matrices rather than vectors is yet another higher rank generalization of phase retrieval [14, 30].

The main goal of the present paper is to combine coarse quantization with phase retrieval from norms of projections under the assumption that the input vector  $x$  is normalized. In our setup, each measured quantity is the answer to a binary question: Is the input vector  $x$  closer to a given subspace or to its orthogonal complement? Hence, a measurement results in a binary string that encodes the orientation of  $x$  in terms of the answers to the binary questions associated with a collection of subspaces. This reduction to binary quantities is a dramatic loss of information compared to phase-insensitive, real-valued measurements. Since the outcome of a measurement is unchanged by rescaling the input vector, we are only obtaining information about the one-dimensional subspace spanned by it. The restriction of  $x$  being a unit vector permits us to perform phase retrieval from its proximity to subspaces. In analogy with the unresolvable ambiguity in phase retrieval, we only seek to recover the one-dimensional subspace spanned by  $x$ , or equivalently, the orthogonal rank-one projection  $X$  onto the span of  $x$ .

The motivating practical setting for the type of signal acquisition and recovery considered here is performing phase retrieval with an optical device related to the so-called single-pixel camera [18]. This camera has a lens and a segmented mirror array that can selectively reflect part of a light beam. The reflected portion is then captured by a photomultiplier. Mathematically, the lens approximately implements a Fourier transform and the mirror array is modeled by a mask that is applied to the Fourier transform of the image. The photomultiplier integrates the intensities, the squared amplitudes of the masked light wave, to obtain each measured quantity. This integrated intensity can indeed be interpreted as the norm of a signal that has been projected by applying the mask. The binary measurements we consider could be generated from the intensities by comparing the photomultiplier output with a threshold value for each choice of the mask. The single-pixel camera was developed to produce images of light intensities reflected by an object [18]. To retrieve phase information, one could replace the lens with an optical implementation of fractional Fourier transforms [24, 25]. In our paper, we do not consider fractional Fourier transforms but apply random unitaries before measuring the norm of a vector projected onto a subspace. We leave it to future work to see if the techniques developed here can be adapted to a more realistic, optics-related setting.

To achieve phase retrieval, we use measure concentration arguments and show that measurements coming from randomly selected subspaces allow approximate recovery via a semidefinite program. The recovery strategy in this paper can be outlined as follows: We specialize to even-dimensional real or complex Hilbert spaces and to randomized one-bit measurements based on subspaces of half the dimension. For each random subspace  $V_j$  in a sequence  $\{V_1, V_2, \dots, V_m\}$ , we determine whether the given input vector  $x$  is closer to  $V_j$  or to its orthogonal complement  $V_j^\perp$ . The outcome of the binary measurement is thus encoded in a sequence of orthogonal projections  $\{\hat{P}_1, \hat{P}_2, \dots, \hat{P}_m\}$  such that the range of each  $\hat{P}_j$  is the subspace  $\hat{V}_j \in \{V_j, V_j^\perp\}$  that is closest to  $x$ . The answer to each binary question is equivalently determined by comparing the squared norm  $\|P_j x\|_2^2 = \text{tr}[P_j x x^*]$  to a threshold. For the approximate recovery of the subspace spanned by  $x$  we then simply average over these orthogonal projections  $\{\hat{P}_j\}_{j=1}^m$  and find the eigenspace corresponding to the largest eigenvalue of this average. We denote the orthogonal projection onto this eigenspace by  $\hat{X}$ . This operator is, in fact, the solution of a semidefinite program which maximizes  $\sum_{j=1}^m \text{tr}[\hat{P}_j Y]$  in the

convex set of all positive semidefinite  $Y$  with  $\text{tr}[Y] \leq 1$  [22, Section 4.2]. This strategy is motivated by earlier results of Plan and Vershynin in the more general setting of one-bit low rank matrix recovery [30].

Randomized constructions and associated algorithms for recovery based on measure concentration have been studied previously in the contexts of matrix recovery, compressed sensing, and other problems in phase retrieval [1, 6, 11, 15, 20, 23, 29, 30]. In contrast to the low-rank matrix recovery treated by Plan and Vershynin [29, 30], we use measure concentration in operator norm to achieve our error bounds. For a related result based on measure concentration in operator norm but without a low-rank prior, see the work by Guta and others [21] on approximate quantum state tomography from measurements associated with projections onto subspaces.

In this paper, we show results that control the accuracy of the approximate recovery, in particular the decay of the error as the number of random subspaces grows. There are two types of error estimates, pointwise and uniform in the input vector.

*Pointwise Bound.* For a rank-one orthogonal projection  $X$  on a real or complex  $2n$ -dimensional Hilbert space and a desired recovery accuracy  $\delta > 0$ , we show that using

$$m \geq C\delta^{-2}n \log(n)$$

random subspaces for a binary measurement and the algorithm we described yields  $\hat{X}$  such that the operator norm difference is bounded by  $\|\hat{X} - X\| < \delta$  with high probability. Here  $C$  is a constant independent of  $n$  and  $\delta$ . See Theorem 2.3.3 for the exact statement and proof of this result, along with an exact value for  $C$ . One may compare this to a similar result from one-bit compressed sensing which says that  $m = C\delta^{-4}n$  random one-bit measurements (of the form  $X \mapsto \text{sign}(\text{tr}[G_j X])$  for  $\{G_j\}_{j=1}^m$  independent matrices with independent standard normal entries) are sufficient to recover  $\hat{X}$  with nuclear norm  $\text{tr}[\|\hat{X}\|] = 1$  and  $\text{tr}[\|\hat{X}\hat{X}^*\|^2] \leq 1$  such that the

Hilbert-Schmidt norm  $\|\hat{X} - X\|_{HS} = \left( \text{tr}[\|\hat{X} - X\|^2] \right)^{\frac{1}{2}} < \delta$  [30, Section 3.3]. Another result on one-bit phase retrieval [27] also gives comparable asymptotics when using measurements based on rank-two Gaussian random matrices.

*Uniform Bound.* We also establish an error bound that holds uniformly for all rank-one projections as input with one fixed choice of subspaces for measurement. For a desired recovery accuracy  $\delta > 0$ , we show that using

$$m \geq C\delta^{-2}n^2 \log(\delta^{-1}n)$$

random subspaces for a binary measurement ensures with high probability that for each rank-one orthogonal projection  $X$  we obtain  $\hat{X}$  such that  $\|\hat{X} - X\| < \delta$ . See Theorem 3.3.1 for details.

We note that for fixed  $n$ , the asymptotic dependence of  $m$  on  $\delta$  improves on results derived by Plan and Vershynin in a more general setting. To compare the results, we recall that their bounds are formulated for the *squared* error norm. In our units, even their pointwise result for 1-bit phase retrieval requires  $m$  to be proportional to  $\delta^{-4}$ . The fact that Plan and Vershynin use the Hilbert-Schmidt norm and not the operator norm as in the present paper is a minor difference, because the error in the rank-one case is computed as the norm of a Hermitian of rank two, so changing between the two types of norms only introduces a factor of  $\sqrt{2}$ . Our improvement in the asymptotics can be attributed to our choice of measurements which are constructed with random orthogonal projections, not Gaussian matrices. One expects that the Lipschitz regularity of the function  $X \mapsto \text{tr}[PX]$  is better than that of  $X \mapsto \text{tr}[GX]$ , at least in a set of large measure among

all rank-one projections. This is advantageous, in particular in combination with perturbation arguments as in Section 3.2.

We also include pointwise and uniform error bounds for faulty measurements. In this case, up to a fixed fraction of the answers to the binary questions have been flipped, possibly in an adversarial manner. For fixed dimension  $n$ , the faulty measurements contribute with an additional term in the bound for the recovery error that is proportional to the fraction of bit flips. We hope to improve the  $n$ -dependence of our bound further, but at least with our current technique the goal of achieving overwhelming success probability for recovering all inputs with a uniform accuracy using a fixed measurement architecture whose measured quantities grow proportionally to  $n$  is not within reach.

The remainder of this paper is organized as follows: After fixing some notation, the remainder of Section 1 describes our one-bit phaseless measurement model in more detail; we explain how we generate random projections for each binary measurement, and how we approximately recover a signal based on such a binary measurement of it. In Section 2 we prove the error bound for our pointwise recovery, Theorem 2.3.3. Lastly, in Section 3 we establish the uniform accuracy for recovery, Theorem 3.3.1. Each of the main theorems in Sections 2 and 3 is followed by a corollary that provides error bounds in the presence of faulty measurements. Both error bounds are also illustrated with plots showing empirical data from reconstruction using (PEP) in  $\mathbb{R}^{16}$ .

**Notation:** Since we are interested in both real and complex signals, we let  $\mathbb{F}$  stand for either  $\mathbb{R}$  or  $\mathbb{C}$ , and define  $\beta = \frac{1}{2}$  when  $\mathbb{F} = \mathbb{R}$  and  $\beta = 1$  when  $\mathbb{F} = \mathbb{C}$  in order to simplify some expressions which depend on the underlying field. We consider only unit norm signals, and so denote the unit sphere in  $\mathbb{F}^d$  by  $\mathbb{S}_{\mathbb{F}}^{d-1}$ . As mentioned previously, both our input signals and binary measurement can be defined in terms of orthogonal projections, so we let  $\text{Proj}_{\mathbb{F}}(k, d)$  denote the space of rank- $k$  orthogonal projections on  $\mathbb{F}^d$ . For a vector  $x \in \mathbb{S}_{\mathbb{F}}^{d-1}$ ,  $xx^* \in \text{Proj}_{\mathbb{F}}(1, d)$  is the rank-one projection onto the span of  $x$ . We write  $\|x\|$  for the euclidean norm of a vector  $x \in \mathbb{F}^d$  and  $\|A\|$  for the operator norm of a matrix  $A \in \mathbb{F}^{d \times d}$ .

**1.1. One-Bit Phaseless Measurement Model.** Our measurements are constructed from qualitative information about the proximity of  $x \in \mathbb{S}_{\mathbb{F}}^{d-1}$  to subspaces in  $\mathbb{F}^d$ . We formulate the measurements in terms of the orthogonal projections onto these subspaces.

For a projection  $P \in \text{Proj}_{\mathbb{F}}(k, d)$ , we define its associated binary question as the map  $\varphi_P : \mathbb{S}_{\mathbb{F}}^{d-1} \rightarrow \{0, 1\}$  given by

$$\varphi_P(x) = \begin{cases} 1 & \text{if } \|Px\|_2^2 \geq \frac{k}{d} \\ 0 & \text{else.} \end{cases}$$

The choice of  $k/d$  as the cut-off value for quantization is natural since it is the average of  $x \mapsto \|Px\|_2^2$  over all unit vectors. Equivalently,  $k/d$  is the average of  $P \mapsto \|Px\|_2^2$  when  $x$  is a fixed unit vector and  $P$  is chosen uniformly at random in  $\text{Proj}_{\mathbb{F}}(k, d)$ , as discussed further below in Section 1.2.

These binary questions are in fact phaseless, since  $\varphi_P(x) = \varphi_P(\alpha x)$  for any  $\alpha \in \mathbb{F}$  with  $|\alpha| = 1$ . Additionally, for any such  $\alpha$  and any  $x \in \mathbb{S}_{\mathbb{F}}^{d-1}$  we have  $\alpha x(\alpha x)^* = xx^*$ , and  $\|Px\|_2^2 = \text{tr}[Pxx^*]$ , so these binary questions can be recast as maps on the set of rank-one orthogonal projections. In this framework — thinking of input signals as rank-one projections — the binary question associated to  $P$  is the map  $\phi_P : \text{Proj}_{\mathbb{F}}(1, d) \rightarrow \{0, 1\}$  defined by

$$\phi_P(X) = \begin{cases} 1 & \text{if } \text{tr}[PX] > \frac{k}{d} \\ 0 & \text{else.} \end{cases}$$



Reformulating  $\varphi_P$  as  $\phi_P$  encapsulates the fact that the map  $\varphi_P$  is constant on the set of unit vectors that differ from  $x$  by a unimodular multiplicative constant. Henceforth, we will use this latter framework and speak of measuring and reconstructing rank-one orthogonal projections rather than unit vectors.

The binary question  $\phi_P$  measures qualitative proximity information about the input signal. For projections  $P \in \text{Proj}_{\mathbb{F}}(k, d)$  and  $X \in \text{Proj}_{\mathbb{F}}(1, d)$ ,  $\text{tr}[PX] = \cos^2(\theta)$ , where  $\theta$  is the principal angle between the one-dimensional subspace  $\text{Ran}(X)$  and the  $k$ -dimensional subspace  $\text{Ran}(P)$ . Thus,  $\phi_P(X) = 1$  if and only if  $\text{Ran}(X)$  is closer to  $\text{Ran}(P)$  than the average for a random one-dimensional subspace, and if this occurs we say  $P$  is *proximal* to  $X$ .

Our goal is to achieve accurate phase retrieval with the qualitative proximity information gained from a sufficiently large set of these binary questions from projections  $\{P_j\}_{j=1}^m$ . For such a collection, we define a corresponding binary measurement map.

**Definition 1.1.1.** *Given a sequence of orthogonal projections  $\mathcal{P} = \{P_j\}_{j=1}^m$  on  $\mathbb{F}^d$ , the binary measurement map associated with  $\mathcal{P}$  is  $\Phi_{\mathcal{P}} : \text{Proj}(1, 2n) \rightarrow \{0, 1\}^m$  defined by*

$$\Phi_{\mathcal{P}}(X) := (\phi_{P_j}(X))_{j=1}^m.$$

*We also define the measurement Hamming distance (associated with  $\mathcal{P}$ ) between  $X$  and  $Y$  to be*

$$d_{\mathcal{P}}(X, Y) := d_H(\Phi_{\mathcal{P}}(X), \Phi_{\mathcal{P}}(Y))$$

*where  $d_H$  denotes the normalized Hamming distance on  $\{0, 1\}^m$ .*

In other words:  $\Phi_{\mathcal{P}}(X)$  is a binary vector where each one-bit entry encodes the proximity of  $X$  to a projection in  $\mathcal{P}$ . The value  $d_{\mathcal{P}}(X, Y)$  gives the relative frequency of measurement projections that separate  $X$  and  $Y$ , i.e. the number of binary questions in the measurement that yield different answers for  $X$  and  $Y$  as inputs.

**1.2. Measurement by Random Projections.** In the absence of an intuitive way to construct “optimal” collections of projections for our one-bit measurements, we instead consider projections chosen uniformly at random. The uniform probability measure on  $\text{Proj}_{\mathbb{F}}(k, d)$  is induced by the Haar measure of the unitary group  $\mathcal{U}_{\mathbb{F}}(d)$ , and is characterized by the property of being rotationally invariant, see [6]. In other words, if  $P$  is uniformly distributed in  $\text{Proj}_{\mathbb{F}}(k, d)$  then for any  $U \in \mathcal{U}_{\mathbb{F}}(d)$  we have  $UPU^* \stackrel{(d)}{=} P$  (where  $\stackrel{(d)}{=}$  denotes equality in distribution).

In practice, there are many equivalent ways to generate a uniformly distributed rank- $k$  projection. For example, one can take  $k$  Gaussian random vectors in  $\mathbb{F}^d$  and then form the projection onto their span. A second way is to take a fixed rank- $k$  projection and conjugate it by a Haar distributed random unitary  $U \in \mathcal{U}_{\mathbb{F}}(d)$ . It can be helpful to think of a “uniformly distributed rank- $k$  projection” as just a “projection onto a uniformly distributed  $k$ -dimensional subspace”.

For most of the paper, we work with the binary measurement map associated to a collection  $\mathcal{P} = \{P_j\}_{j=1}^m$  of independent uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . The reason for using half-dimensioned projections is because their associated one-bit measurements  $\phi_P$  have a geometrically intuitive meaning: for a fixed  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$ ,  $\phi_P(X) = 1$  if and only if  $\text{tr}[PX] > \frac{1}{2} \geq \text{tr}[(I - P)X]$ , i.e. the subspace  $\text{Ran}(X)$  is closer to  $\text{Ran}(P)$  than to its orthogonal complement  $\text{Ran}(I - P)$ .

**1.3. Approximate Phase Retrieval by Semidefinite Programming.** A main goal of this paper is to use the outcomes of a random binary measurement to estimate the input accurately. Suppose we have measured an unknown vector  $x \in \mathbb{S}_{\mathbb{F}}^{2n-1}$  with the binary measurement map

$\Phi_{\mathcal{P}}$  associated with a random collection of projections  $\mathcal{P} \subset \text{Proj}_{\mathbb{F}}(n, 2n)$  and obtained the binary vector  $\Phi_{\mathcal{P}}(xx^*)$ . The information we gain from these measurements will not in general completely determine the rank-1 projection  $X = xx^*$  corresponding to the input vector  $x$ , but with enough measured quantities we can deduce a projection  $\hat{X}$  which approximates  $X$  in some metric. A consistent reconstruction would seek an element  $\hat{X}$  in the feasible set, that is, the set of all  $Y$  consistent with the binary measurement in the sense that  $\Phi_{\mathcal{P}}(Y) = \Phi_{\mathcal{P}}(X)$  [12]. A natural error bound for such a reconstruction strategy would then result from the diameter of the feasible set, which intuitively will be small if  $\mathcal{P}$  is suitably large.

In this paper, we relax the perfect consistency condition, but still achieve approximate recovery with a computationally feasible, semidefinite programming algorithm investigated in other works [22, Section 4.2]. The approximate recovery of  $X$  is conveniently described in terms of projections obtained from the binary measurement  $\Phi_{\mathcal{P}}(X)$ .

**Definition 1.3.1.** *Given  $X \in \text{Proj}_{\mathbb{F}}(1, d)$  and  $P \in \text{Proj}_{\mathbb{F}}(k, d)$  we define the proximally flipped projection*

$$\hat{P}(X) := \begin{cases} P & \text{if } \text{tr}[PX] \geq \frac{k}{d} \\ I - P & \text{if } \text{tr}[PX] < \frac{k}{d}. \end{cases}$$

*Next, for a sequence of orthogonal projections  $\mathcal{P} = \{P_j\}_{j=1}^m$ , the empirical average of the proximally flipped projections is*

$$\hat{Q}_{\mathcal{P}}(X) := \frac{1}{m} \sum_{j=1}^m \hat{P}_j(X).$$

The recovery algorithm we study takes the binary measurement  $\Phi_{\mathcal{P}}(X)$  and produces  $\hat{X}$  that solves the semidefinite program

$$\begin{aligned} \text{(PEP)} \quad & \underset{Y}{\text{maximize}} && \text{tr}[\hat{Q}_{\mathcal{P}}(X)Y] \\ & \text{subject to} && Y \succeq 0, \text{tr}[Y] \leq 1. \end{aligned}$$

We call this the Principal Eigenspace Program (PEP) because it amounts to maximizing the Rayleigh quotient [22, Section 4.2] for  $\hat{Q}_{\mathcal{P}}(X)$ . This special class of semidefinite programs can be implemented efficiently [28, Chapter 4].

Since  $\hat{Q}_{\mathcal{P}}(X)$  is a positive self-adjoint operator, it may be decomposed according to the spectral theorem as a linear combination of mutually orthogonal rank-1 projections  $\hat{Q}_{\mathcal{P}}(X) = \sum_{i=1}^{2n} \lambda_i E_i$ , where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{2n} \geq 0$ . Thus, any positive self-adjoint trace normalized operator with range contained in the principal eigenspace of  $\hat{Q}_{\mathcal{P}}(X)$  is a solution to (PEP). If in addition  $\lambda_1$  is strictly larger than  $\lambda_2$  (which happens with probability 1 for our random measurement model), then its principal eigenspace is one-dimensional, and so  $\hat{X} = E_1$  is the unique solution to (PEP). Proposition 2.1.2 will show that  $\mathbb{E}[\hat{Q}_{\mathcal{P}}(X)] = \mu_1 X + \mu_2(I - X)$  with  $\mu_1 > \mu_2$ , and so for large  $m$  we might expect  $\hat{X} \approx X$  by a measure concentration argument.

Section 2 of this paper shows the following pointwise result: for any fixed  $X \in \text{Proj}(1, 2n)$  and any  $\delta > 0$ , we can choose  $m$  large enough so that a collection of independent uniformly distributed half-dimensioned projections  $\mathcal{P} = \{P_j\}_{j=1}^m$  will, with high probability, yield a measurement  $\Phi_{\mathcal{P}}(X)$  for which the solution  $\hat{X}$  to (PEP) satisfies  $\|\hat{X} - X\| < \delta$ . See Theorem 2.3.3 for details.

Much of the effort in Section 3 is directed toward getting *uniform* results from the above *pointwise* one. The uniform result we derive says: for any  $\delta > 0$ , we can choose  $m$  large enough so that a

collection of independent uniformly distributed half-dimensioned projections  $\mathcal{P} = \{P_j\}_{j=1}^m$  will, with high probability, yield measurements  $\Phi_{\mathcal{P}}(X)$  for every  $X \in \text{Proj}(1, 2n)$  for which the solution  $\hat{X}$  to (PEP) satisfies  $\|\hat{X} - X\| < \delta$ . See Theorem 3.3.1 for details.

According to the uniform result, we can generate a collection of projections for which *every* signal is approximately recoverable up to an error of  $\delta$  from the one-bit questions using those projections. The pointwise result can be thought of as an averaged performance guarantee, whereas the uniform bound controls even the worst case input.

## 2. ERROR BOUND FOR THE APPROXIMATE RECOVERY OF FIXED INPUT SIGNALS

We begin deriving results on the statistics of signal recovery using (PEP) and our one-bit phaseless measurement model by considering a fixed unit-norm input vector  $x \in \mathbb{F}^{2n}$  while the binary measurement map  $\Phi_{\mathcal{P}}$  is chosen randomly. As outlined before, we identify vectors that differ by a unimodular multiplicative constant, and when considering only unit-norm vectors as input signals we represent these equivalence classes by rank-one projection matrices. The random binary measurement map is determined by a sequence of random projections  $\mathcal{P} = \{P_j\}_{j=1}^m$  whose rank is half the dimension of the signal space, and provides information whether the input signal is closer to the range of each projection or to its orthogonal complement. The main goal of this section is to prove that (PEP) provides accurate recovery of an input signal  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  when sufficiently many random projections are used for the binary measurement, i.e. when  $m$  is large enough. The derivation of the results proceeds in three steps:

- (1) If the orthogonal projections for the measurement of  $X$  are chosen uniformly at random and proximally flipped, then their empirical average has the expectation  $Q(X) := \mathbb{E} [\hat{Q}_{\mathcal{P}}(X)] = \mu_1 X + \mu_2 (I - X)$  where  $0 < \mu_2 < \mu_1$  are constants. In particular,  $X$  is the projection onto the eigenspace corresponding to the largest eigenvalue of  $Q(X)$ .
- (2) The empirical average  $\hat{Q}_{\mathcal{P}}(X)$  concentrates near its expectation  $Q(X)$ .
- (3) The eigenspace of  $\hat{Q}_{\mathcal{P}}(X)$  corresponding to its largest eigenvalue concentrates near  $X$ .

**2.1. Expectation of  $\hat{Q}_{\mathcal{P}}(X)$ .** Before we can investigate the accuracy of (PEP), we need a simple fact about the distribution of the principal angle between a random  $n$ -dimensional subspace and a fixed one-dimensional subspace in  $\mathbb{F}^{2n}$ .

**Lemma 2.1.1.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  be fixed and  $P \in \text{Proj}_{\mathbb{F}}(n, 2n)$  be uniformly distributed. Then  $\text{tr}[PX] \sim \text{Beta}(\beta n, \beta n)$ , i.e.  $\text{tr}[PX]$  has probability density function*

$$p(t) = B(\beta n, \beta n)^{-1} [t(1-t)]^{\beta n-1},$$

where  $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$  is the Beta function. In particular,  $\mathbb{E}[\text{tr}[PX]] = \frac{1}{2}$  and the distribution of  $\text{tr}[PX]$  is symmetric about  $\frac{1}{2}$ .

*Proof.* Recall that if  $U \in \mathcal{U}_{\mathbb{F}}(2n)$  is uniformly distributed and  $E$  is the orthogonal projection onto the first  $n$  standard basis vectors, then  $UEU^* \stackrel{(d)}{=} P$ . Thus

$$\text{tr}[PX] \stackrel{(d)}{=} \text{tr}[UEU^*X] = \text{tr}[EU^*XU].$$

Observe that  $U^*XU$  is a uniformly distributed rank-1 projection, which has the same distribution as  $uu^*$  where  $u \in \mathbb{S}_{\mathbb{F}}^{2n-1}$  is a uniformly distributed unit vector. Furthermore,  $u \stackrel{(d)}{=} \frac{g}{\|g\|_2}$  where

$g \sim N(0, I_{2n})$  is the standard Gaussian random vector in  $\mathbb{F}^{2n}$ . So we have

$$(1) \quad \operatorname{tr}[EU^*XU] \stackrel{(d)}{=} \operatorname{tr}[Euu^*] = \|Eu\|_2^2 \stackrel{(d)}{=} \frac{\|Eg\|_2^2}{\|g\|_2^2} = \frac{\sum_{k=1}^n |g_k|^2}{\sum_{k=1}^{2n} |g_k|^2}.$$

If  $\mathbb{F} = \mathbb{R}$ , then the  $g_k$ 's are independent standard Gaussian random variables, so the right hand side of equation (1) has the form  $\frac{A}{A+B}$  where the random variables  $A, B \sim \chi^2(n)$  are independent. Thus, Equation (1) is a Beta( $\frac{n}{2}, \frac{n}{2}$ ) random variable.

If  $\mathbb{F} = \mathbb{C}$ , then each  $g_k = a_k + ib_k$  where all the  $a_k$  and  $b_k$ 's are independent standard random variables. In this case, since  $|g_k|^2 = |a_k|^2 + |b_k|^2$ , the right hand side of Equation 1 has the form  $\frac{A}{A+B}$  where  $A, B \sim \chi^2(2n)$  are independent, and thus is a Beta( $n, n$ ) random variable.  $\square$

Next we compute the expectation of the empirical average of the proximally flipped projections.

**Proposition 2.1.2.** *Let  $X \in \operatorname{Proj}_{\mathbb{F}}(1, 2n)$  and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\operatorname{Proj}_{\mathbb{F}}(n, 2n)$ , then*

$$Q(X) = \mu_1 X + \mu_2 (I - X),$$

where

$$\mu_1 = \frac{1}{2} + \frac{1}{\beta n 4^{\beta n} B(\beta n, \beta n)}, \quad \mu_2 = \frac{1}{2} - \frac{1}{\beta n (2n-1) 4^{\beta n} B(\beta n, \beta n)}.$$

*Proof.* We begin with some manipulation and reasoning that does not depend on whether  $\mathbb{F}$  is  $\mathbb{R}$  or  $\mathbb{C}$ , which only makes a difference when computing the values of  $\mu_1$  and  $\mu_2$ .

Since the  $P_j$ 's are identically distributed, we know that  $\mathbb{E}[\hat{P}_i(X)] = \mathbb{E}[\hat{P}_j(X)]$  for all  $i$  and  $j$ .

Thus, by linearity of expectation we have  $Q(X) = \mathbb{E}[\hat{P}_1(X)]$ .

Also, the distribution of  $\hat{P}_1(X)$  is invariant under conjugation with a unitary that fixes  $X$ . In other words, for a unitary  $U \in \mathcal{U}_{\mathbb{F}}(2n)$  such that  $UXU^* = X$ , then  $U\hat{P}_1(X)U^* \stackrel{(d)}{=} \hat{P}_1(X)$ . To verify this, we use the rotational invariance of  $P_1$  and the cyclic property of the trace to obtain

$$U\hat{P}_1(X)U^* \stackrel{(d)}{=} U(\widehat{U^*P_1U})(X)U^* = \hat{P}_1(UXU^*) = \hat{P}_1(X).$$

Using the linearity of expectation once more, it follows that  $Q(X)$  is also invariant under conjugation by unitaries that fix  $X$ . This implies that every eigenspace of  $Q(X)$  is preserved under rotations by all such unitaries, hence  $\operatorname{Ran}(X)$  and  $\operatorname{Ran}(X)^\perp$  are the eigenspaces of  $Q(X)$ . Letting  $\mu_1$  and  $\mu_2$  denote the respective eigenvalues, we write

$$Q(X) = \mu_1 X + \mu_2 (I - X).$$

In order to determine the value of  $\mu_1$ , we use linearity of expectation to see

$$\mu_1 = \operatorname{tr}[Q(X)X] = \mathbb{E}\left[\operatorname{tr}\left[\hat{P}_1(X)X\right]\right].$$

By the law of total probability we have,

$$\begin{aligned} \mathbb{E}\left[\operatorname{tr}\left[\hat{P}_1(X)X\right]\right] &= \mathbb{E}\left[\operatorname{tr}\left[\hat{P}_1(X)X\right] \mid \operatorname{tr}[P_1X] \geq \frac{1}{2}\right] \mathbb{P}\left\{\operatorname{tr}[P_1X] \geq \frac{1}{2}\right\} \\ &\quad + \mathbb{E}\left[\operatorname{tr}\left[\hat{P}_1(X)X\right] \mid \operatorname{tr}[P_1X] < \frac{1}{2}\right] \mathbb{P}\left\{\operatorname{tr}[P_1X] < \frac{1}{2}\right\}, \end{aligned}$$

so by the definition of  $\hat{P}_1(X)$  and the symmetry of the distribution of  $\text{tr}[P_1 X]$  — a consequence of Lemma 2.1.1 — it follows that

$$\mathbb{E} \left[ \text{tr} \left[ \hat{P}_1(X) X \right] \right] = \mathbb{E} \left[ \text{tr} [PX] \mid \text{tr} [PX] \geq \frac{1}{2} \right].$$

We can compute this conditional expectation using integration by parts with the probability density function given in Lemma 2.1.1, yielding

$$\mu_1 = 2B(\beta n, \beta n)^{-1} \int_{\frac{1}{2}}^1 [t(1-t)]^{\beta n-1} dt = \frac{1}{2} + \frac{1}{\beta n 4^{\beta n} B(\beta n, \beta n)}.$$

Since  $\text{tr}[Q(X)] = n$  by linearity of expectation, we know  $\mu_1 + (2n-1)\mu_2 = n$ , from which we get the desired expression for  $\mu_2$ .  $\square$

**2.2. Concentration of  $\hat{Q}_{\mathcal{P}}(X)$  near  $Q(X)$ .** Since the empirical average of the proximally flipped projections  $\hat{Q}_{\mathcal{P}}(X)$  is, after all, an empirical average, by the law of large numbers it should concentrate tightly around its expectation  $Q(X)$  as the number of measurements  $m$  goes to infinity. To make this precise, we use the Matrix Bernstein Inequality [34, Theorem 1.6.2].

**Lemma 2.2.1.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then*

$$\mathbb{E} \left[ \left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \right] \leq \sqrt{\frac{\log(4n)}{2m}} + \frac{\log(4n)}{3m},$$

and for any  $t > 0$ ,

$$\mathbb{P} \left\{ \left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \geq t \right\} \leq 4n \exp \left( -\frac{6t^2 m}{7} \right).$$

In particular, if  $m \geq \frac{7}{6}t^{-2}(\log(4n) + D)$  then

$$\mathbb{P} \left\{ \left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \geq t \right\} \leq \exp(-D).$$

*Proof.* Let  $S_j = \frac{1}{m}(\hat{P}_j(X) - Q(X))$ . Then  $\mathbb{E}[S_j] = 0$  and  $\|S_j\| \leq \frac{1}{m}$  for all  $j = 1, \dots, m$ . Note that  $Z := \sum_{j=1}^m S_j = \hat{Q}_{\mathcal{P}}(X) - Q(X)$ . Additionally, since  $\hat{P}_j(X)$  is a projection and  $\mathbb{E}[\hat{P}_j(X)] = Q(X)$  for all  $j$ , we may bound the matrix variance

$$v(Z) := \left\| \sum_{j=1}^m \mathbb{E}[S_j^2] \right\| = \frac{1}{m} \|Q(X) - Q(X)^2\| \leq \frac{1}{4m}.$$

The expectation bound and tail bound now follow from applying the Matrix Bernstein Inequality as in [34, Theorem 1.6.2]. Additionally, if  $m \geq \frac{7}{6}t^{-2}(\log(4n) + D)$  then  $\log(4n) - \frac{6t^2 m}{7} \leq -D$ , which yields (2.2.1).  $\square$

**2.3. Concentration of  $\hat{X}$  near  $X$  (Pointwise Result).** From Lemma 2.2.1 we know that, with enough measurement projections, with high probability  $\hat{Q}_{\mathcal{P}}(X)$  is close to  $Q(X)$  in operator norm. When it is sufficiently close, then the eigenspace of  $\hat{Q}_{\mathcal{P}}(X)$  corresponding to its maximum eigenvalue will also be close to  $X$ . To see this, we first need the following lemma. It is based on the fact that for two rank-one projections  $X$  and  $Y$ , the difference  $X - Y$  is a zero-trace self-adjoint operator of

rank two, and hence has a spectral representation of the form  $X - Y = \|X - Y\| (A - B)$  with two mutually orthogonal rank-one projections  $A$  and  $B$ .

**Lemma 2.3.1.** *Let  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$ , then*

$$\|X - Y\| = (\mu_1 - \mu_2)^{-1} \text{tr}[Q(X)(A - B)]$$

where  $A, B \in \text{Proj}_{\mathbb{F}}(1, 2n)$  are the mutually orthogonal projections in the spectral decomposition  $X - Y = \|X - Y\| (A - B)$ .

*Proof.* Let  $\theta$  be the principal angle between the subspaces associated to  $X$  and  $Y$ . Then we can pick  $x, y, z \in \mathbb{S}_{\mathbb{F}}^{2n-1}$  with  $x \perp z$  such that  $X = xx^*$ ,  $Y = yy^*$  and  $y = \cos(\theta)x + \sin(\theta)z$ . Then

$$Y = yy^* = \cos^2(\theta)xx^* + \sin^2(\theta)zz^* + \sin(\theta)\cos(\theta)(xz^* + zx^*).$$

Since  $Q(X) = \mu_1 X + \mu_2(I - X)$ ,

$$\text{tr}[Q(X)(X - Y)] = \mu_1 - \cos^2(\theta)\mu_1 - \sin^2(\theta)\mu_2 = (\mu_1 - \mu_2)\sin^2(\theta).$$

Lastly, since  $\sin(\theta) = \|X - Y\|$ , rewriting the left hand side using the spectral decomposition  $X - Y = \|X - Y\| (A - B)$  and cancelling the common factor of  $\|X - Y\|$  yields the desired equality.  $\square$

The spectral gap  $\mu_1 - \mu_2$  of  $Q(X)$  appears in the sufficient number of binary questions in both our pointwise and uniform result. The following lemma bounds this quantity in terms of the dimension  $n$ .

**Lemma 2.3.2.** *Let  $\mu_1$  and  $\mu_2$  be as in Proposition 2.1.2, then for  $\beta n \geq 2$*

$$\frac{(n-1)\sqrt{2\beta n-1}}{\sqrt{2\pi}\beta n(2n-1)} \leq \mu_1 - \mu_2 \leq \frac{4(n-1)\sqrt{2\beta n-1}}{e\sqrt{2\pi}\beta n(2n-1)}$$

In particular,  $(\mu_1 - \mu_2)^{-1} = O(\sqrt{n})$ .

*Proof.* From the expressions derived in Proposition 2.1.2 we have

$$(2) \quad \mu_1 - \mu_2 = \frac{2(n-1)}{\beta n(2n-1)4^{\beta n}B(\beta n, \beta n)}.$$

Since  $B(\alpha_1, \alpha_2) = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}$ , we may use Stirling's formula to approximate the Beta function. In particular, from [32] we have for all real numbers  $k \geq 2$

$$\exp\left(\frac{1}{12k-11}\right) \leq \frac{\Gamma(k)}{\sqrt{2\pi}(k-1)^{k-\frac{1}{2}}\exp(-(k-1))} \leq \exp\left(\frac{1}{12k-12}\right).$$

In particular, when  $\beta n \geq 2$  these inequalities for the Gamma function yield the bounds

$$(3) \quad \frac{e\sqrt{2\pi}}{2 \cdot 4^{\beta n}\sqrt{2\beta n-1}} \leq B(\beta n, \beta n) \leq \frac{2\sqrt{2\pi}}{4^{\beta n}\sqrt{2\beta n-1}}.$$

Using these bounds for the Beta function in (2) gives the desired inequalities for  $\mu_1 - \mu_2$ .  $\square$

Now we have the tools to prove the pointwise error bound for approximate recovery of a fixed input signal using (PEP).

**Theorem 2.3.3.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and  $\delta > 0$  be fixed. If*

$$m \geq \frac{14}{3}(\mu_1 - \mu_2)^{-2}\delta^{-2}(\log(4n) + D),$$

*and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then with probability at least  $1 - \exp(-D)$*

$$(4) \quad \|\hat{X} - X\| < \delta,$$

*where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$ .*

*Proof.* From Lemma 2.3.1, we know that

$$\|\hat{X} - X\| = (\mu_1 - \mu_2)^{-1} \text{tr}[Q(X)(A - B)],$$

where  $A, B \in \text{Proj}_{\mathbb{F}}(1, 2n)$  are the orthogonal projections from the spectral decomposition of the difference  $X - \hat{X} = \|X - \hat{X}\| (A - B)$ .

Since  $\hat{X}$  is the projection onto the principal eigenspace of  $\hat{Q}_{\mathcal{P}}(X)$ , we see

$$\text{tr}[\hat{Q}_{\mathcal{P}}(X)(\hat{X} - X)] \geq 0 \implies (\mu_1 - \mu_2)^{-1} \text{tr}[\hat{Q}_{\mathcal{P}}(X)(B - A)] \geq 0,$$

and so

$$(5) \quad \|\hat{X} - X\| \leq (\mu_1 - \mu_2)^{-1} \text{tr}[(Q(X) - \hat{Q}(X))(A - B)] \leq 2(\mu_1 - \mu_2)^{-1} \|Q(X) - \hat{Q}(X)\|.$$

We have chosen  $m$  such that  $m \geq \frac{7}{6}t^{-2}(\log(4n) + D)$  for  $t = \frac{1}{2}(\mu_1 - \mu_2)\delta$ , so the tail bound in Lemma 2.2.1 says with probability at least  $1 - \exp(-D)$  we have  $\|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| < t$ . If this occurs, then from (5) we see

$$\|\hat{X} - X\| \leq 2(\mu_1 - \mu_2)^{-1} \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| < \delta.$$

□

See Figure 1 for a plot showing how our bound on the sufficient number of measurements to achieve an accuracy of  $\delta$  relates to experimental results.

Our proof lets us fine tune the probability of successful recovery by adjusting the value of  $D$  in (4). By increasing  $D$  we increase the probability of success, but also increase the sufficient number of measurements. In particular, we can take  $D = \alpha \log(n)$  to ensure success with high probability, i.e. the failure rate decays on the order of  $n^{-\alpha}$ . To do so, we gain a constant factor that depends on  $\alpha$  in the number of sufficient measurement projections  $m$ .

**Corollary 2.3.4.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and  $\delta > 0$  be fixed. If  $\alpha > 0$ ,  $m \geq C_{\alpha}\delta^{-2}n \log(n)$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then with probability at least  $1 - n^{-\alpha}$*

$$\|\hat{X} - X\| < \delta,$$

*where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$  and  $C$  is a constant that only depends on  $\alpha$ .*

We can also take  $D = n$  in (4) to ensure success with overwhelming probability, i.e. the failure rate decays on the order of  $\exp(-n)$ . In this case, we gain an additional factor of  $n$  in the number of sufficient measurement projections  $m$ .

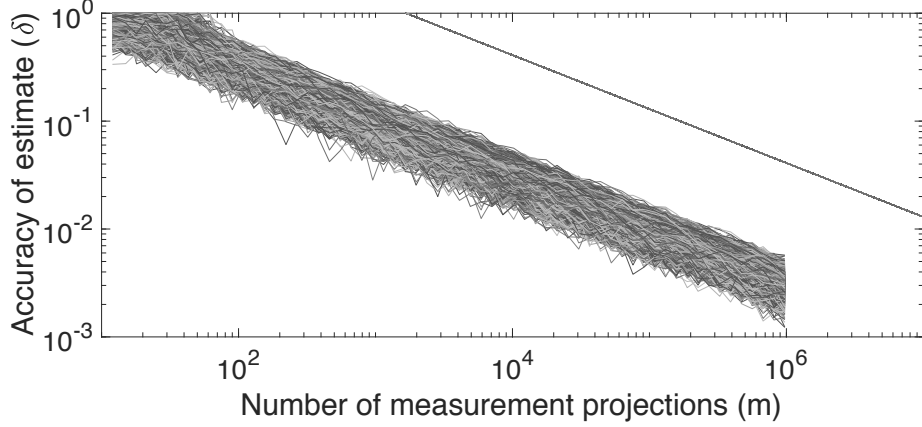


FIGURE 1. Plot showing the accuracy of recovery using (PEP) for a fixed input and 7200 independent collections of up to  $10^6$  measurement projections on  $\mathbb{R}^{16}$ . The single line separate from the cluster represents the upper bound on  $\delta$  given by Theorem 2.3.3.

**Corollary 2.3.5.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and  $\delta > 0$  be fixed. If  $m \geq C\delta^{-2}n^2 \log(n)$  and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then with probability at least  $1 - \exp(-n)$*

$$\|\hat{X} - X\| < \delta,$$

where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$  and  $C$  is a constant.

The pointwise accuracy guarantee of Theorem 2.3.3 can also be thought of as an “average case” error bound with respect to the random sequence of measurement projections  $\mathcal{P}$ . The following corollary makes this explicit.

**Corollary 2.3.6.** *Let  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and  $\delta > 0$  be fixed. If  $m \geq C\delta^{-2}n \log(n)$  and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then*

$$\mathbb{E} \left[ \|\hat{X} - X\| \right] < \delta,$$

where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$  and  $C$  is a constant.

*Proof.* Take the expectation on both sides of (5) with respect to the random sequence of projections  $\mathcal{P}$  and use the expectation bound from Lemma 2.2.1.  $\square$

We also remark that recovery using (PEP) is robust to bit-flip errors in the binary measurement, which can be seen via a small addition to the proof of Theorem 2.3.3. To this end, we consider a faulty measurement  $\tilde{\Phi}_{\mathcal{P}}$  with the property that the normalized Hamming distance between the faulty and correct measurements is bounded by a fixed fraction  $d_H(\tilde{\Phi}_{\mathcal{P}}(X), \Phi_{\mathcal{P}}(X)) \leq \tau$ .

**Corollary 2.3.7.** *Let  $X$ ,  $\delta$ ,  $m$ , and  $\{P_j\}_{j=1}^m$  be as in Theorem 2.3.3, and fix  $0 < \tau < 1$ . Then with probability at least  $1 - \exp(-D)$ , for all  $\tilde{\Phi}_{\mathcal{P}}(X) \in \{0, 1\}^m$  such that*

$$d_H(\Phi_{\mathcal{P}}(X), \tilde{\Phi}_{\mathcal{P}}(X)) \leq \tau,$$



we have

$$\|\tilde{X} - X\| \leq \delta + 2(\mu_1 - \mu_2)^{-1}\tau,$$

where  $\tilde{X}$  denote the solution to (PEP) with input  $\tilde{\Phi}_{\mathcal{P}}(X)$  and  $\mu_1 - \mu_2$  is controlled by Lemma 2.3.2.

*Proof.* Let  $\tilde{Q}_{\mathcal{P}}(X)$  denote the empirical average of the (faulty) flipped projections, i.e. flipped using  $\tilde{\Phi}_{\mathcal{P}}(X)$  rather than  $\Phi_{\mathcal{P}}(X)$ . Then as before, we have

$$\|\tilde{X} - X\| \leq 2(\mu_1 - \mu_2)^{-1} \|\tilde{Q}_{\mathcal{P}}(X) - Q(X)\|.$$

By the triangle inequality, it follows that

$$\|\tilde{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \|\tilde{Q}_{\mathcal{P}}(X) - \hat{Q}_{\mathcal{P}}(X)\| + \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\|.$$

Since the normalized Hamming distance between  $\Phi_{\mathcal{P}}(X)$  and  $\tilde{\Phi}_{\mathcal{P}}(X)$  is bounded by  $\tau$ , we see

$$\|\tilde{Q}_{\mathcal{P}}(X) - \hat{Q}_{\mathcal{P}}(X)\| \leq \tau.$$

Since  $\|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \delta$  with probability at least  $1 - \exp(-D)$  by the same proof as in Theorem 2.3.3, the result follows.  $\square$

We expect that a deeper analysis will reveal a better dependence on the error rate, or perhaps eliminate the dimension dependent factor  $(\mu_1 - \mu_2)^{-1}$ .

### 3. FROM POINTWISE TO UNIFORMLY ACCURATE RECOVERY

In this section we extend the result from Theorem 2.3.3 to show that the recovery error using (PEP) is small uniformly across all input vectors  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  for a single random binary measurement  $\Phi_{\mathcal{P}}$ . Our strategy consists of the following steps:

- (1) Using sufficiently many random projections,  $\hat{Q}_{\mathcal{P}}(X)$  concentrates near  $Q(X)$  for all  $X$  in an  $\epsilon$ -net of  $\text{Proj}_{\mathbb{F}}(1, 2n)$ .
- (2) With high probability the measurement Hamming distance between a pair  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  is not much larger than  $\|X - Y\|$ , uniformly for all such pairs.
- (3) The eigenspace of  $\hat{Q}_{\mathcal{P}}(X)$  corresponding to its largest eigenvalue concentrates near  $X$  uniformly for all  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$ .

**3.1. Concentration of  $\hat{Q}_{\mathcal{P}}(X)$  near  $Q(X)$  uniformly on a net.** First, we show an inequality relating the Euclidean distance between unit vectors to the operator norm distance between their associated rank-one projections.

**Lemma 3.1.1.** *Let  $d \in \mathbb{N}$ . Then for all  $x, y \in \mathbb{S}_{\mathbb{F}}^{d-1}$ ,*

$$\|xx^* - yy^*\| \leq \|x - y\|_2.$$

*Proof.* Let  $\theta$  be the principal angle between the subspaces associated to  $xx^*$  and  $yy^*$ , and recall  $\|xx^* - yy^*\| = \sin(\theta)$ . Thus

$$\|x - y\|_2^2 = \langle x - y, x - y \rangle = 2 - 2\Re\langle x, y \rangle \geq 2 - 2|\langle x, y \rangle| = 2 - 2\cos(\theta).$$

Since  $\theta \in [0, \frac{\pi}{2}]$  we know  $0 \leq \cos(\theta) \leq 1$  and so

$$2 - 2\cos(\theta) = 2(1 - \cos) \geq (1 + \cos(\theta))(1 - \cos(\theta)) = \sin^2(\theta) = \|xx^* - yy^*\|^2.$$

$\square$

Next, we use Lemma 3.1.1 to prove the existence of  $\epsilon$ -nets of  $\text{Proj}_{\mathbb{F}}(1, 2n)$  with explicit cardinality bounds. This follows from the analogous results for  $\epsilon$ -nets of  $\mathbb{S}_{\mathbb{F}}^{2n-1}$ .

**Lemma 3.1.2.** *For any  $\epsilon > 0$ , there exists an  $\epsilon$ -net  $\mathcal{N}_{\epsilon}$  for  $\text{Proj}_{\mathbb{F}}(1, 2n)$  with respect to the operator norm with cardinality satisfying*

$$\log |\mathcal{N}_{\epsilon}| \leq 4\beta n \log(1 + 2\epsilon^{-1}).$$

*Proof.* By the standard volume bound for the covering number of the sphere in real euclidean space [9], and the fact that  $\mathbb{S}_{\mathbb{C}}^{2n-1}$  is naturally isometric to  $\mathbb{S}_{\mathbb{R}}^{4n-1}$ , for every  $\epsilon > 0$  there exists an  $\epsilon$ -net  $\mathcal{N}'_{\epsilon}$  for  $\mathbb{S}_{\mathbb{F}}^{2n-1}$  (with respect to the Euclidean distance) with cardinality satisfying

$$|\mathcal{N}'_{\epsilon}| \leq \left(1 + \frac{2}{\epsilon}\right)^{4\beta n}.$$

By Lemma 3.1.1,  $\mathcal{N}_{\epsilon} := \{xx^* : x \in \mathcal{N}'_{\epsilon}\}$  is an  $\epsilon$ -net for  $\text{Proj}_{\mathbb{F}}(1, 2n)$  with the desired cardinality bound.  $\square$

Now that we have existence of epsilon-nets with control on their cardinality, we use a union bound and Lemma 2.2.1 to show that with sufficiently many measurements,  $\hat{Q}_{\mathcal{P}}(X)$  concentrates near  $Q(X)$  uniformly for all  $X$  in an epsilon-net of  $\text{Proj}_{\mathbb{F}}(1, 2n)$ .

**Lemma 3.1.3.** *Let  $\epsilon > 0$  and  $\mathcal{N}_{\epsilon}$  be an  $\epsilon$ -net of  $\text{Proj}_{\mathbb{F}}(1, 2n)$  such that  $\log |\mathcal{N}_{\epsilon}| \leq 4\beta n \log(1 + 2\epsilon^{-1})$ . Also, let  $\delta > 0$ ,  $m \geq \frac{7}{6}\delta^{-2} [\log(4n) + 4\beta n \log(1 + 2\epsilon^{-1}) + D]$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . Then with probability at least  $1 - \exp(-D)$  we have*

$$\left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \leq \delta$$

for all  $X \in \mathcal{N}_{\epsilon}$ .

*Proof.* By Lemma 2.2.1 and our assumption on  $m$ , for each  $X \in \mathcal{N}_{\epsilon}$  we know

$$\mathbb{P} \left\{ \left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \geq \delta \right\} \leq \exp(-4\beta n \log(1 + 2\epsilon^{-1}) - D).$$

By taking a union bound over all  $X \in \mathcal{N}_{\epsilon}$  it follows that

$$\mathbb{P} \left\{ \left\| \hat{Q}_{\mathcal{P}}(X) - Q(X) \right\| \leq \delta \text{ for all } X \in \mathcal{N}_{\epsilon} \right\} \geq 1 - |\mathcal{N}_{\epsilon}| \exp(-4\beta n \log(1 + 2\epsilon^{-1}) - D).$$

The claim follows from our upper bound on  $|\mathcal{N}_{\epsilon}|$ .  $\square$

**3.2. Relation between the measurement Hamming distance and operator norm distance.** The main goal of this section is to prove our guarantee for uniformly accurate recovery, Theorem 3.2.10: With sufficiently many measurements, with high probability the measurement Hamming distance between any pair  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  is not much larger than the operator norm of their difference. It is relatively simple to show that this happens for fixed  $X$  and  $Y$ , but showing that it holds uniformly for all such pairs requires more complicated techniques. To this end, we will define the *t-soft Hamming distance* similarly as in Plan and Vershynin's *Dimension reduction by random hyperplane tessellations* [31]. We establish a continuity property and concentration results for the *t-soft Hamming distance*, which allow us to show uniform concentration of the measurement Hamming distance near its expected value over all of  $\text{Proj}_{\mathbb{F}}(1, 2n)$ . We then show that  $\mathbb{E}[d_{\mathcal{P}}(X, Y)]$  can be bounded in terms of  $\|X - Y\|$ , after which Theorem 3.2.10 follows.

**3.2.1. The  $t$ -soft Hamming distance and its continuity properties.** For any  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  let  $\mathcal{S}_{X,Y} := \{P \in \text{Proj}_{\mathbb{F}}(n, 2n) : \phi_P(X) \neq \phi_P(Y)\}$ , i.e. the set of projections that yield different measurements of  $X$  and  $Y$ . If  $P \in \mathcal{S}_{X,Y}$ , then we say that  $P$  *separates*  $X$  and  $Y$ . For a sequence  $\mathcal{P} = \{P_j\}_{j=1}^m \subset \text{Proj}_{\mathbb{F}}(n, 2n)$ , notice that  $d_{\mathcal{P}}(X, Y) = \frac{1}{m} \sum_{j=1}^m \mathbb{1}_{\mathcal{S}_{X,Y}}(P_j)$ .

With this expression for the measurement Hamming distance in mind, we define

$$\begin{aligned} \mathcal{S}_{X,Y}^t &:= \{P \in \text{Proj}_{\mathbb{F}}(n, 2n) : \text{tr}[PX] + t < \frac{1}{2} \leq \text{tr}[PY] - t\} \\ &\cup \{P \in \text{Proj}_{\mathbb{F}}(n, 2n) : \text{tr}[PY] + t < \frac{1}{2} \leq \text{tr}[PX] - t\} \end{aligned}$$

for all  $t \in \mathbb{R}$ , and if  $P \in \mathcal{S}_{X,Y}^t$  then we say  $P$   $t$ -separates  $X$  and  $Y$ .

**Definition 3.2.1.** Given a sequence of orthogonal projections  $\mathcal{P} = \{P_j\}_{j=1}^m$  in  $\text{Proj}_{\mathbb{F}}(n, 2n)$  and  $t \in \mathbb{R}$ , we define the  $t$ -soft Hamming distance between input projections  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  to be

$$d_{\mathcal{P}}^t(X, Y) := \frac{1}{m} \sum_{j=1}^m \mathbb{1}_{\mathcal{S}_{X,Y}^t}(P_j).$$

Ultimately we want to prove uniform results for the measurement Hamming distance, but its discontinuity causes problems with standard  $\epsilon$ -net arguments. The  $t$ -soft Hamming distance helps us work around this discontinuity, where the parameter  $t$  determines how strict the criteria should be for determining if the measurements of two vectors are different. This is reflected in the fact that for  $t_1 \leq 0 \leq t_2$  we have  $\mathcal{S}_{X,Y}^{t_2} \subset \mathcal{S}_{X,Y} \subset \mathcal{S}_{X,Y}^{t_1}$ .

The addition of this extra parameter lets us formulate a type of continuity for  $d_{\mathcal{P}}^t(X, Y)$  where both  $t$  and the projections  $X$  and  $Y$  are allowed to vary. If we want to perturb the projections  $X, Y$  by a small amount in operator norm, then we can make up for it by slightly increasing/decreasing the parameter  $t$ .

**Proposition 3.2.2.** Let  $\mathcal{P} = \{P_j\}_{j=1}^m$  be a sequence of projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ ,  $t \in \mathbb{R}$ ,  $\epsilon > 0$ , and  $X_0, Y_0, X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  such that  $\|X - X_0\| < \epsilon$  and  $\|Y - Y_0\| < \epsilon$ , then

$$d_{\mathcal{P}}^{t+\epsilon}(X, Y) \leq d_{\mathcal{P}}^t(X_0, Y_0) \leq d_{\mathcal{P}}^{t-\epsilon}(X, Y)$$

*Proof.* Suppose  $P \in \mathcal{S}_{X,Y}^{t+\epsilon}$ . Then, without loss of generality, we may assume that

$$\text{tr}[PY] + t + \epsilon < \frac{1}{2} < \text{tr}[PX] - t - \epsilon.$$

Since  $P$  is a projection we have  $|\text{tr}[P(Y_0 - Y)]| \leq \|Y - Y_0\| < \epsilon$ , so

$$\text{tr}[PY_0] + t = \text{tr}[PY] - \text{tr}[P(Y - Y_0)] + t \leq \text{tr}[PY] + t + \epsilon < \frac{1}{2}$$

and also

$$\text{tr}[PX_0] - t = \text{tr}[PX] - \text{tr}[P(X - X_0)] - t \geq \text{tr}[PX] - t - \epsilon > \frac{1}{2}.$$

Thus  $\mathcal{S}_{X,Y}^{t+\epsilon} \subset \mathcal{S}_{X_0,Y_0}^t$ , and so  $d_{\mathcal{P}}^{t+\epsilon}(X, Y) \leq d_{\mathcal{P}}^t(X_0, Y_0)$ .

The second inequality follows from above by swapping the roles of  $X, Y$  with  $X_0, Y_0$  and replacing  $t$  with  $t - \epsilon$ .  $\square$

**3.2.2. Concentration of  $t$ -soft Hamming distance.** In this section, we state a basic concentration result for the  $t$ -soft Hamming distance between two fixed vectors, and then extend it to a uniform result over an  $\epsilon$ -net.

**Lemma 3.2.3.** *Let  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ ,  $t \in \mathbb{R}$ ,  $\delta > 0$ , and  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  be fixed, then*

$$\mathbb{P} \left\{ \left| d_{\mathcal{P}}^t(X, Y) - \mathbb{E} [d_{\mathcal{P}}^t(X, Y)] \right| > \delta \right\} \leq 2 \exp(-2\delta^2 m).$$

*Proof.* From the way that we defined the  $t$ -soft Hamming distance,  $m \cdot d_{\mathcal{P}}^t(X, Y) \sim \text{Bin}(m, p)$  where  $p = \mathbb{E} [d_{\mathcal{P}}^t(X, Y)]$ . The result then follows from a standard Chernoff bound for binomial random variables (see [4]).  $\square$

We can now use Proposition 3.2.3 and the bounds on the size of  $\epsilon$ -nets of  $\text{Proj}_{\mathbb{F}}(1, 2n)$  from Lemma 3.1.2 to take a union bound. The result is a bound for the probability that the  $t$ -soft Hamming distance is close to its expectation for all pairs of projections in an  $\epsilon$ -net simultaneously.

**Proposition 3.2.4.** *Let  $\epsilon > 0$  and  $\mathcal{N}_{\epsilon}$  be an  $\epsilon$ -net of  $\text{Proj}_{\mathbb{F}}(1, 2n)$  such that  $\log |\mathcal{N}_{\epsilon}| \leq 4\beta n \log(1 + 2\epsilon^{-1})$ . Also, let  $t \in \mathbb{R}$ ,  $\delta > 0$ ,  $m \geq \frac{1}{2}\delta^{-2} (8\beta n \log(1 + 2\epsilon^{-1}) + D)$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . Then with probability at least  $1 - \exp(-D)$  we have*

$$\left| d_{\mathcal{P}}^t(X, Y) - \mathbb{E} [d_{\mathcal{P}}^t(X, Y)] \right| \leq \delta$$

for all  $X, Y \in \mathcal{N}_{\epsilon}$ .

*Proof.* By Proposition 3.2.3 and taking a union bound over all  $\binom{|\mathcal{N}_{\epsilon}|}{2} \leq \frac{1}{2} |\mathcal{N}_{\epsilon}|^2$  pairs in  $\mathcal{N}_{\epsilon} \times \mathcal{N}_{\epsilon}$ , we have that

$$\mathbb{P} \left\{ \left| d_{\mathcal{P}}^t(X, Y) - \mathbb{E} [d_{\mathcal{P}}^t(X, Y)] \right| \leq \delta, \forall (X, Y) \in \mathcal{N}_{\epsilon} \times \mathcal{N}_{\epsilon} \right\} \geq 1 - |\mathcal{N}_{\epsilon}|^2 \exp(-2\delta^2 m).$$

Using our bound on the cardinality of  $|\mathcal{N}_{\epsilon}|$  and our assumption about  $m$  we have

$$|\mathcal{N}_{\epsilon}|^2 \exp(-2\delta^2 m) \leq \exp(8\beta n \log(1 + 2\epsilon^{-1}) - 2\delta^2 m) = \exp(-D).$$

$\square$

The following proposition addresses how varying  $t$  affects the expected difference of the  $t$ -soft Hamming distance from the measurement Hamming distance.

**Proposition 3.2.5.** *Let  $\mathcal{P} = \{P_j\}_{j=1}^m$  be an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ ,  $t \in \mathbb{R}$ , and  $X, Y \in \text{Proj}(1, 2n)$  be fixed, then*

$$\left| \mathbb{E} [d_{\mathcal{P}}^t(X, Y) - d_{\mathcal{P}}(X, Y)] \right| = \left| \mathbb{P} \{P_1 \in \mathcal{S}_{X,Y}^t\} - \mathbb{P} \{P_1 \in \mathcal{S}_{X,Y}\} \right| \leq \frac{32\sqrt{2\beta n - 1}}{e\sqrt{2\pi}} |t|.$$

*Proof.* Because the  $t$ -soft and regular Hamming distances are linear combinations of indicator functions, and the fact that the  $P_j$  are i.i.d., we have

$$\left| \mathbb{E} [d_{\mathcal{P}}^t(X, Y) - d_{\mathcal{P}}(X, Y)] \right| = \left| \mathbb{E} \left[ \mathbb{1}_{\mathcal{S}_{X,Y}^t}(P_1) - \mathbb{1}_{\mathcal{S}_{X,Y}}(P_1) \right] \right|,$$

and by Jensen's inequality it follows that

$$(6) \quad \left| \mathbb{E} \left[ \mathbb{1}_{\mathcal{S}_{X,Y}^t}(P_1) - \mathbb{1}_{\mathcal{S}_{X,Y}}(P_1) \right] \right| \leq \mathbb{E} \left[ \left| \mathbb{1}_{\mathcal{S}_{X,Y}^t}(P_1) - \mathbb{1}_{\mathcal{S}_{X,Y}}(P_1) \right| \right] = \mathbb{P} \{P_1 \in \mathcal{S}_{X,Y}^t \triangle \mathcal{S}_{X,Y}\}.$$

We break up this symmetric difference into two disjoint pieces

$$\mathbb{P} \{P_1 \in \mathcal{S}_{X,Y}^t \triangle \mathcal{S}_{X,Y}\} = \mathbb{P} \{P_1 \in \mathcal{S}_{X,Y}^t \setminus \mathcal{S}_{X,Y}\} + \mathbb{P} \{P_1 \in \mathcal{S}_{X,Y} \setminus \mathcal{S}_{X,Y}^t\}$$

and look at two cases. First, if  $t > 0$  then  $\mathcal{S}_{X,Y}^t \setminus \mathcal{S}_{X,Y}$  is empty, and

$$\mathcal{S}_{X,Y} \setminus \mathcal{S}_{X,Y}^t \subset \left\{ \left| \operatorname{tr}[P_1 X] - \frac{1}{2} \right| < t \right\} \cup \left\{ \left| \operatorname{tr}[P_1 Y] - \frac{1}{2} \right| < t \right\}.$$

Similarly, if  $t < 0$  then  $\mathcal{S}_{X,Y} \setminus \mathcal{S}_{X,Y}^t$  is empty and again

$$\mathcal{S}_{X,Y}^t \setminus \mathcal{S}_{X,Y} \subset \left\{ \left| \operatorname{tr}[P_1 X] - \frac{1}{2} \right| < -t \right\} \cup \left\{ \left| \operatorname{tr}[P_1 Y] - \frac{1}{2} \right| < -t \right\},$$

Since  $\operatorname{tr}[P_1 X] \stackrel{(d)}{=} \operatorname{tr}[P_1 Y]$ , in both cases we have

$$(7) \quad \mathbb{P} \{ P_1 \in \mathcal{S}_{X,Y}^t \triangle \mathcal{S}_{X,Y} \} \leq 2\mathbb{P} \left\{ \left| \operatorname{tr}[P_1 X] - \frac{1}{2} \right| < |t| \right\}.$$

By Lemma 2.1.1 we know  $\operatorname{tr}[P_1 X] \sim \text{Beta}(\beta n, \beta n)$ , and so we can bound this probability using the the probability density function of the beta distribution. To begin with, we see

$$(8) \quad \begin{aligned} \mathbb{P} \left\{ \left| \operatorname{tr}[P_1 X] - \frac{1}{2} \right| < |t| \right\} &= \frac{2}{B(\beta n, \beta n)} \int_{\frac{1}{2}}^{\frac{1}{2}+|t|} x^{\beta n-1} (1-x)^{\beta n-1} dx \\ &= \frac{2}{B(\beta n, \beta n)} \int_0^{|t|} \left( \frac{1}{4} - x^2 \right)^{\beta n-1} dx \\ &\leq \frac{2}{B(\beta n, \beta n)} \int_0^{|t|} 4^{1-\beta n} dx \\ &= \frac{8|t|}{4^{\beta n} B(\beta n, \beta n)}. \end{aligned}$$

Using the lower bound for the Beta function in (3) then yields

$$(9) \quad \mathbb{P} \left\{ \left| \operatorname{tr}[P_1 X] - \frac{1}{2} \right| < |t| \right\} \leq \frac{16\sqrt{2\beta n-1}}{e\sqrt{2\pi}} |t|.$$

The result follows from combining equation (6) with inequalities (7) and (9).  $\square$

**3.2.3. Uniform concentration of Hamming distance.** We now have all the tools we need to prove that with sufficiently many measurements the Hamming distance concentrates near its expected value for all pairs in  $\text{Proj}_{\mathbb{F}}(1, 2n)$ .

**Theorem 3.2.6.** *Let  $\delta > 0$ ,  $m \geq 2\delta^{-2} \left( 8\beta n \log \left( 1 + 128 \frac{\sqrt{2\beta n-1}}{2\sqrt{2\pi}} \delta^{-1} \right) + \log(2) + D \right)$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be a collection of independent uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . Then with probability at least  $1 - \exp(-D)$  we have*

$$|d_{\mathcal{P}}(X, Y) - \mathbb{E}[d_{\mathcal{P}}(X, Y)]| < \delta$$

for all  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$ .

*Proof.* Let  $\epsilon = \frac{e\sqrt{2\pi}}{64\sqrt{2\beta n-1}}\delta$  and let  $\mathcal{N}_{\epsilon}$  be an  $\epsilon$ -net of  $\text{Proj}_{\mathbb{F}}(1, 2n)$  with  $\log |\mathcal{N}_{\epsilon}| \leq 4\beta n \log(1 + 2\epsilon^{-1})$  as in Lemma 3.1.2. By our assumption on  $m$ , Proposition 3.2.4 says that

$$\mathbb{P} \left\{ |d_{\mathcal{P}}^{\epsilon}(X, Y) - \mathbb{E}[d_{\mathcal{P}}^{\epsilon}(X, Y)]| > \frac{\delta}{2} \text{ for some } X, Y \in \mathcal{N}_{\epsilon} \right\} \leq \exp(-\log(2) - D)$$

and also

$$\mathbb{P} \left\{ \left| d_{\mathcal{P}}^{-\epsilon}(X, Y) - \mathbb{E} [d_{\mathcal{P}}^{-\epsilon}(X, Y)] \right| > \frac{\delta}{2} \text{ for some } X, Y \in \mathcal{N}_{\epsilon} \right\} \leq \exp(-\log(2) - D),$$

and so with probability at least  $1 - \exp(-D)$  we have  $|d_{\mathcal{P}}^{\pm\epsilon}(X, Y) - \mathbb{E} [d_{\mathcal{P}}^{\pm\epsilon}(X, Y)]| \leq \delta$  for all  $X, Y \in \mathcal{N}_{\epsilon}$  (call this event  $\mathcal{A}$ ).

Suppose that  $\mathcal{A}$  occurs. Consider an arbitrary pair  $X, Y \in \text{Proj}(1, 2n)$  and let  $X_0, Y_0 \in \mathcal{N}_{\epsilon}$  such that  $\|X - X_0\| < \epsilon$  and  $\|Y - Y_0\| < \epsilon$ . By Proposition 3.2.2 we know that  $d_{\mathcal{P}}(X, Y) \leq d_{\mathcal{P}}^{\epsilon}(X_0, Y_0) \leq d_{\mathcal{P}}^{2\epsilon}(X, Y)$ . These inequalities together with  $\mathcal{A}$  holding imply

$$d_{\mathcal{P}}(X, Y) \leq d_{\mathcal{P}}^{\epsilon}(X_0, Y_0) \leq \mathbb{E} [d_{\mathcal{P}}^{\epsilon}(X_0, Y_0)] + \frac{\delta}{2} \leq \mathbb{E} [d_{\mathcal{P}}^{2\epsilon}(X, Y)] + \frac{\delta}{2}.$$

By Proposition 3.2.5 we have  $|\mathbb{E} [d_{\mathcal{P}}^{2\epsilon}(X, Y)] - \mathbb{E} [d_{\mathcal{P}}(X, Y)]| \leq \frac{32\sqrt{2\beta n-1}}{e\sqrt{2\pi}} |\epsilon| = \frac{\delta}{2}$ , hence

$$d_{\mathcal{P}}(X, Y) \leq \mathbb{E} [d_{\mathcal{P}}(X, Y)] + \delta.$$

Similarly, using Proposition 3.2.2 again shows that  $d_{\mathcal{P}}(X, Y) \geq d_{\mathcal{P}}^{-\epsilon}(X_0, Y_0) \geq d_{\mathcal{P}}^{-2\epsilon}(X, Y)$ , and since  $\mathcal{A}$  holds we have

$$d_{\mathcal{P}}(X, Y) \geq d_{\mathcal{P}}^{-\epsilon}(X_0, Y_0) \geq \mathbb{E} [d_{\mathcal{P}}^{-\epsilon}(X_0, Y_0)] - \frac{\delta}{2} \geq \mathbb{E} [d_{\mathcal{P}}^{-2\epsilon}(X, Y)] - \frac{\delta}{2}.$$

Using Proposition 3.2.5 as above but for  $t = -\epsilon$  yields

$$d_{\mathcal{P}}(X, Y) \geq \mathbb{E} [d_{\mathcal{P}}(X, Y)] - \delta.$$

□

We have just shown that when the measurement projections are chosen uniformly and independently, then  $d_{\mathcal{P}}(X, Y)$  concentrates near  $\mathbb{E} [d_{\mathcal{P}}(X, Y)] = \mathbb{P} \{P \in \mathcal{S}_{X,Y}\}$  for all  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$ , where  $P$  is a single uniformly distributed projection in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . When  $n = 1$ , then  $\mathbb{P} \{P \in \mathcal{S}_{X,Y}\} = \frac{2}{\pi} \theta \leq \sin(\theta) = \|X - Y\|$ , where  $\theta$  is the principal angle between  $\text{Ran}(X)$  and  $\text{Ran}(Y)$ . In the remainder of section, we show that this upper bound holds for arbitrary  $n$ , see Proposition 3.2.9. To achieve this, we need to investigate the joint distribution of  $(\text{tr}[PX], \text{tr}[PY])$ .

By rotational invariance of the distribution of  $P$  we may assume that  $\text{Ran}(X)$  and  $\text{Ran}(Y)$  are in the two-dimensional subspace spanned by  $e_1$  and  $e_2$ , the first two standard basis vectors. Viewed as matrices, this means that all entries of  $X$  and  $Y$  are zero outside of the top-left  $2 \times 2$  submatrix. Furthermore, if  $\tilde{P}, \tilde{X}$ , and  $\tilde{Y}$  are the top-left  $2 \times 2$  submatrices of their respective matrices then  $(\text{tr}[PX], \text{tr}[PY]) = (\text{tr}[\tilde{P}\tilde{X}], \text{tr}[\tilde{P}\tilde{Y}])$ . We study the joint distribution of  $(\text{tr}[PX], \text{tr}[PY])$  through the submatrix  $\tilde{P}$  acting on  $\mathbb{F}^2$ .

Since  $P$  is Hermitian, so is  $\tilde{P}$ . Thus we may write  $\tilde{P} = \lambda_1 E_1 + \lambda_2 E_2$  where  $\lambda_1 \geq \lambda_2$  are the eigenvalues of  $\tilde{P}$  and  $E_1 \perp E_2$  are the projections onto their corresponding eigenspaces. We write  $\lambda(\tilde{P}) := (\lambda_1, \lambda_2)$ , and  $E(\tilde{P}) := (E_1, E_2)$ . By the rotational invariance of  $P$ ,  $E_1$  is uniformly distributed in  $\text{Proj}_2(1, 2)$  and  $E_2 = I - E_1$  since Hermitian matrices have mutually orthogonal eigenspaces. Note also that  $\lambda(\tilde{P})$  and  $E(\tilde{P})$  are independent of each other. The distribution of  $\lambda(\tilde{P})$  is given in the following lemma.

**Lemma 3.2.7.** *Let  $n \geq 2$  and  $P \in \text{Proj}_{\mathbb{F}}(n, 2n)$  be uniformly distributed. Then  $\lambda(\tilde{P})$  has probability density function  $p_n$  on  $\mathcal{D} := \{(x, y) \in [0, 1]^2 : y \leq x\}$  defined by*

$$p_n(x, y) := M_n^{-1} (x - y)^{2\beta} [x(1 - x)y(1 - y)]^{\beta(n-1)-1},$$

with the normalization constant

$$M_n = \begin{cases} \frac{2}{n-1} B(n-1, n-1) & \text{if } \mathbb{F} = \mathbb{R} \\ \frac{1}{8n-4} B(n-1, n-1)^2 & \text{if } \mathbb{F} = \mathbb{C}. \end{cases}$$

*Proof.* The probability density functions are given by [5, Proposition 4.1.4] with  $p = 2$ ,  $q = 2n - 2$ ,  $r = n - 2$  and  $s = n - 2$ . It only remains to compute the normalization constants  $M_n$ .

Suppose  $\mathbb{F} = \mathbb{R}$ . Then  $p_n(x, y) = M_n^{-1} (x - y) [x(1 - x)y(1 - y)]^{\frac{n-3}{2}}$ . Define the functions

$$\begin{aligned} f_n(x, y) &= -\frac{1}{n-1} [x(1-x)]^{\frac{n-3}{2}} [y(1-y)]^{\frac{n-1}{2}} \\ g_n(x, y) &= -\frac{1}{n-1} [x(1-x)]^{\frac{n-1}{2}} [y(1-y)]^{\frac{n-3}{2}}. \end{aligned}$$

With these definitions, we have  $p_n = M_n^{-1} (\frac{\partial g_n}{\partial x} - \frac{\partial f_n}{\partial y})$  on  $\mathcal{D}$ . So by Green's theorem,

$$1 = \iint_{\mathcal{D}} p_n(x, y) dx dy = M_n^{-1} \oint_{\partial \mathcal{D}} f_n dx + g_n dy,$$

where  $\partial \mathcal{D}$  is the boundary of  $\mathcal{D}$ . Note that  $f_n$  and  $g_n$  both vanish on the boundary of  $\mathcal{D}$  except for the diagonal  $\Delta := \{(x, y) \in \mathcal{D} : x = y\}$ , so we only need to compute the line integral over  $\Delta$ . Parameterizing  $\Delta$  by  $x(t) = y(t) = 1 - t$  for  $t \in [0, 1]$ , we see

$$\begin{aligned} M_n &= \oint_{\partial \mathcal{D}} f_n dx + g_n dy = - \int_0^1 [f_n(x(t), y(t)) + g_n(x(t), y(t))] dt \\ &= \frac{2}{n-1} \int_0^1 t^{n-2} (1-t)^{n-2} dt \\ &= \frac{2}{n-1} B(n-1, n-1). \end{aligned}$$

Next, we consider the case when  $\mathbb{F} = \mathbb{C}$ . Then  $p_n(x, y) = M_n^{-1} (x - y)^2 [x(1 - x)y(1 - y)]^{n-2}$ . By symmetry,  $1 = \frac{1}{2} \iint_{[0,1]^2} p_n(x, y) dx dy$ , so by expanding this integral and facts about the Beta distribution, we see

$$1 = \frac{1}{2} \iint_{[0,1]^2} p_n(x, y) dx dy = M_n^{-1} \text{var}(b) \cdot B(n-1, n-1)^2,$$

where  $b \sim \text{Beta}(n-1, n-1)$ . This beta-distributed random variable has variance  $\text{var}(b) = \frac{1}{4(2n-1)}$ , which determines  $M_n$ .  $\square$

Let  $\mathcal{D}_{\text{Sep}} := \{(x, y) \in \mathcal{D} : y < \frac{1}{2} < x\}$ . Then  $\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}$  if and only if there exist projections  $A, B \in \text{Proj}_{\mathbb{F}}(1, 2)$  such that  $\tilde{P} \in \mathcal{S}_{A, B}$ . This is true because  $\lambda_1 = \max_{A' \in \text{Proj}(1, 2)} \text{tr}[PA']$  and  $\lambda_2 = \max_{B' \in \text{Proj}(1, 2n)} \text{tr}[PB']$ . In particular,  $P \in \mathcal{S}_{X, Y}$  requires  $\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}$ . For this reason, we compute the probability that  $\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}$ .

**Lemma 3.2.8.** *Let  $n \geq 2$ , and  $P \in \text{Proj}_{\mathbb{F}}(n, 2n)$  be uniformly distributed, then*

$$\mathbb{P} \left\{ \lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}} \right\} = \begin{cases} \frac{B(\frac{n-1}{2}, \frac{n-1}{2})}{2^n B(n-1, n-1)} \rightarrow \frac{1}{\sqrt{2}} & \text{if } \mathbb{F} = \mathbb{R} \\ \frac{1}{2} + \frac{8n-4}{(n-1)^2 2^{4n-3} B(n-1, n-1)^2} \rightarrow \frac{1}{2} + \frac{1}{\pi} & \text{if } \mathbb{F} = \mathbb{C}. \end{cases}$$

*Proof.* First, suppose  $\mathbb{F} = \mathbb{R}$ , so  $p_n(x, y) = M_n^{-1}(x - y)[x(1 - x)y(1 - y)]^{\frac{n-3}{2}}$ . Then,

$$\mathbb{P}\left\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}\right\} = M_n^{-1} \int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 (x - y)[x(1 - x)y(1 - y)]^{\frac{n-3}{2}} dx dy.$$

By linearity and Fubini's theorem, we get

$$\int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 (x - y)[x(1 - x)y(1 - y)]^{\frac{n-3}{2}} dx dy = \frac{1}{4} \left[ \mathbb{E}\left[b \mid b \geq \frac{1}{2}\right] - \mathbb{E}\left[b \mid b \leq \frac{1}{2}\right] \right] B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)^2,$$

where  $b \sim \text{Beta}\left(\frac{n-1}{2}, \frac{n-1}{2}\right)$ . Calculating these conditional expectations we get

$$\mathbb{E}\left[b \mid b \geq \frac{1}{2}\right] - \mathbb{E}\left[b \mid b \leq \frac{1}{2}\right] = \frac{1}{(n-1)2^{n-3}B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)},$$

and combining this with Lemma 3.2.7 yields

$$\mathbb{P}\left\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}\right\} = \frac{B\left(\frac{n-1}{2}, \frac{n-1}{2}\right)}{2^n B(n-1, n-1)}.$$

Next, suppose  $\mathbb{F} = \mathbb{C}$ , so  $p_n(x, y) = M_n^{-1}(x - y)^2[x(1 - x)y(1 - y)]^{n-2}$ . Then,

$$\mathbb{P}\left\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}\right\} = M_n^{-1} \int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 (x - y)^2[x(1 - x)y(1 - y)]^{n-2} dx dy.$$

Expanding  $(x - y)^2$  and rewriting integrals in terms of expectations of beta-distributed random variables, we see

$$\int_0^{\frac{1}{2}} \int_{\frac{1}{2}}^1 (x - y)^2[x(1 - x)y(1 - y)]^{n-2} dx dy = \frac{1}{2} \left[ \mathbb{E}[b^2] - \mathbb{E}\left[b \mid b \geq \frac{1}{2}\right] \cdot \mathbb{E}\left[b \mid b \leq \frac{1}{2}\right] \right] B(n-1, n-1)^2,$$

where  $b \sim \text{Beta}(n-1, n-1)$ . We know that  $\mathbb{E}[b^2] = \frac{n}{4n-2} = \frac{1}{4} + \frac{1}{8n-4}$ , and also

$$\begin{aligned} \mathbb{E}\left[b \mid b \geq \frac{1}{2}\right] \cdot \mathbb{E}\left[b \mid b \leq \frac{1}{2}\right] &= \left(\frac{1}{2} + \frac{1}{(n-1)2^{2n-2}B(n-1, n-1)}\right) \left(\frac{1}{2} - \frac{1}{(n-1)2^{2n-2}B(n-1, n-1)}\right) \\ &= \frac{1}{4} - \frac{1}{(n-1)^2 2^{4n-4} B(n-1, n-1)^2}. \end{aligned}$$

Putting this all together yields

$$\mathbb{P}\left\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}\right\} = \frac{1}{2} + \frac{8n-4}{(n-1)^2 2^{4n-3} B(n-1, n-1)^2}.$$

The asymptotic limit of  $\mathbb{P}\left\{\lambda(\tilde{P})\right\}$  as  $n \rightarrow \infty$  follows from Stirling's approximation as in (3), see [32].  $\square$

Now we are prepared to bound  $\mathbb{P}\{P \in \mathcal{S}_{X,Y}\}$  in terms of the operator norm distance  $\|X - Y\|$ .

**Proposition 3.2.9.** *Let  $P \in \text{Proj}_{\mathbb{F}}(n, 2n)$  be uniformly distributed, then*

$$\mathbb{P}\{P \in \mathcal{S}_{X,Y}\} \leq \|X - Y\|.$$



*Proof.* The case when  $n = 1$  is simple and was mentioned previously, so we consider here  $n \geq 2$ . Further, without loss of generality, assume  $\text{Ran}(X), \text{Ran}(Y) \subset \text{Ran}(E)$  where  $E$  is the orthogonal projection onto  $\text{span}\{e_1, e_2\}$ . By conditioning,  $\mathbb{P}\{P \in \mathcal{S}_{X,Y}\} = \mathbb{E}\left[\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\}\right]$ . By the definition of  $\mathcal{D}_{\text{Sep}}$  we see that  $\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} = 0$  if  $\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}^c$ . Hence

$$\mathbb{E}\left[\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\}\right] = \mathbb{E}\left[\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} \mathbb{1}_{\mathcal{D}_{\text{Sep}}}(\lambda(\tilde{P}))\right].$$

Suppose now that  $\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}$ , and first consider the case when  $\mathbb{F} = \mathbb{R}$ . Then  $\text{Proj}_{\mathbb{R}}(1, 2)$  can be viewed as  $\mathbb{S}_{\mathbb{R}}^1$  with its opposite points identified, and  $E(\tilde{P})$  is a (uniformly distributed) random pair of antipodal points in this quotient space. Letting  $E_1 = v_1 v_1^*$  and  $E_2 = v_2 v_2^*$  where  $v_1$  and  $v_2$  are normalized eigenvectors corresponding to eigenvalues  $\lambda_1$  and  $\lambda_2$  of  $\tilde{P}$ , we may parameterize  $\text{Proj}_{\mathbb{R}}(1, 2)$  by  $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$  via  $\phi \mapsto Z_\phi := (\cos(\phi)v_1 + \sin(\phi)v_2)(\cos(\phi)v_1 + \sin(\phi)v_2)^* = \cos^2(\phi)E_1 + \sin^2(\phi)E_2 + \sin(\phi)(\cos(\phi)(v_1 v_2^* + v_2 v_1^*))$ . We see that  $\text{tr}[\tilde{P}Z_\phi] = \lambda_1 \cos^2(\phi) + \lambda_2 \sin^2(\phi) = \lambda_1 - (\lambda_1 - \lambda_2)\sin^2(\phi)$ . Since  $\text{tr}[\tilde{P}Z_0] = \lambda_1 > \frac{1}{2}$  and  $\text{tr}[\tilde{P}Z_{\frac{\pi}{2}}] = \lambda_2 < \frac{1}{2}$ , there exists some  $\phi_h \in (0, \frac{\pi}{2})$  such that  $\text{tr}[\tilde{P}Z_{\phi_h}] = \text{tr}[\tilde{P}Z_{-\phi_h}] = \frac{1}{2}$ . In fact,  $\phi_h = \arcsin\left(\sqrt{\frac{\lambda_1 - \frac{1}{2}}{\lambda_1 - \lambda_2}}\right)$ . We see that  $\text{tr}[\tilde{P}Z_\phi] > \frac{1}{2}$  for  $\phi \in (-\phi_h, \phi_h)$ , and  $\text{tr}[\tilde{P}Z_\phi] < \frac{1}{2}$  for  $\phi \in [-\frac{\pi}{2}, -\phi_h) \cup (\phi_h, \frac{\pi}{2}]$ .

All of this goes to show that  $\lambda(\tilde{P})$  determines  $\phi_h$ , which along with the orientation of  $E_1$  determines which rank-1 projections in  $\text{Ran}(E)$  that  $P$  separates. In the quotient space picture, the open arc between  $E_{\phi_h}$  and  $E_{-\phi_h}$  containing  $E_1$  represents the rank-1 projections with measurements greater than  $\frac{1}{2}$ , and the complementary arc represents those with measurements less than  $\frac{1}{2}$ . Let  $w = \min\{2\phi_h, \pi - 2\phi_h\}$ , which is the length of the smallest of these two arcs. If  $w \leq \theta$ , then  $\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} = \frac{2w}{\pi} \leq \frac{2\theta}{\pi}$ . If  $w > \theta$ , then  $\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} = \frac{2\theta}{\pi}$ . So

$$\begin{aligned} \mathbb{E}\left[\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} \mathbb{1}_{\mathcal{D}_{\text{Sep}}}(\lambda(\tilde{P}))\right] &\leq \mathbb{E}\left[\frac{2\theta}{\pi} \mathbb{1}_{\mathcal{D}_{\text{Sep}}}(\lambda(\tilde{P}))\right] \\ &= \frac{2\theta}{\pi} \mathbb{P}\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{Sep}}\} \\ &\leq \|X - Y\|. \end{aligned}$$

Next, we consider the case when  $\mathbb{F} = \mathbb{C}$ , in which case  $\text{Proj}_{\mathbb{C}}(1, 2)$  can be identified with the Bloch sphere [10]. By rotational invariance,  $E(\tilde{P})$  is a pair of (uniformly distributed) antipodal points on the sphere, and  $\lambda(\tilde{P})$  determines which pairs of projections are separated by  $P$ . If  $v_1$  and  $v_2$  are eigenvectors of  $\tilde{P}$  as above,  $E_1 = v_1 v_1^*$  and  $E_2 = v_2 v_2^*$ , and  $v_{\phi,\psi} := \cos(\frac{\phi}{2})v_1 + e^{i\psi} \sin(\frac{\phi}{2})v_2$  for  $\phi \in [0, \pi], \psi \in [0, 2\pi]$ , then  $Z_{\phi,\psi} = v_{\phi,\psi} v_{\phi,\psi}^*$  lies on the circle of points in the Bloch sphere at an angle of  $\phi$  from  $E_1$ . Moreover, this representation shows that  $\text{tr}[\tilde{P}Z_{\phi,\psi_1}] = \text{tr}[\tilde{P}Z_{\phi,\psi_2}]$  for all  $\phi, \psi_1$ , and  $\psi_2$ . By continuity, there must exist some  $\phi_h \in [0, \pi]$  such that  $\text{tr}[\tilde{P}Z_{\phi_h,\psi}] = \frac{1}{2}$  for all  $\psi \in [0, 2\pi]$ . In fact, we can calculate  $\phi_h = 2\arcsin(\sqrt{\frac{\lambda_1 - \frac{1}{2}}{\lambda_1 - \lambda_2}})$ . The open spherical cap centered at  $E_1$  of angle  $\phi_h$  consists exactly of those projections  $Z \in \text{Proj}_{\mathbb{C}}(1, 2)$  such that  $\text{tr}[\tilde{P}Z] > \frac{1}{2}$ , and the complementary cap consists of those for which  $\text{tr}[\tilde{P}Z] < \frac{1}{2}$ .

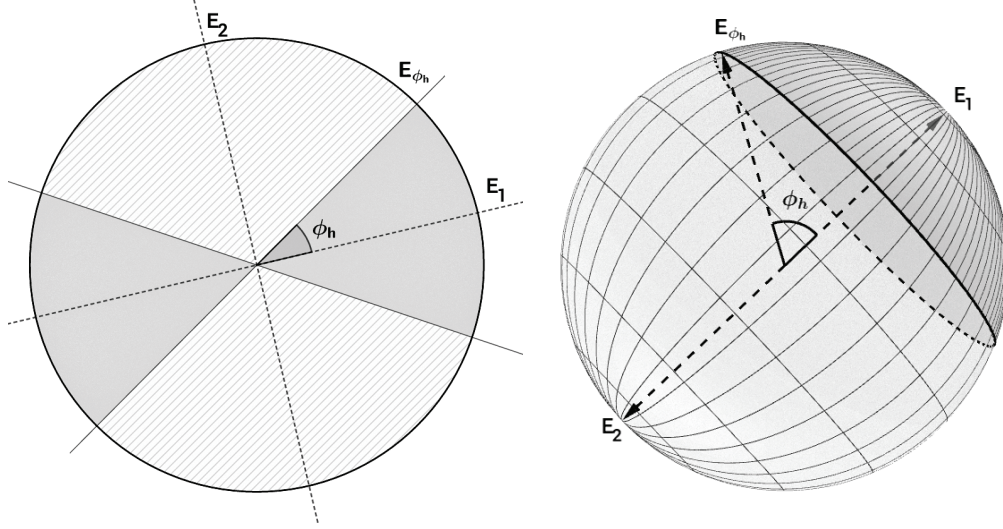


FIGURE 2. The  $2 \times 2$  principal submatrix  $\tilde{P}$  of  $P$  divides  $\text{Proj}_{\mathbb{F}}(1, 2)$  into two disjoint sets based on whether the Hilbert-Schmidt inner product of a rank-one orthogonal projection with  $\tilde{P}$  is greater or less than  $\frac{1}{2}$  (Left:  $\mathbb{F} = \mathbb{R}$ ; Right:  $\mathbb{F} = \mathbb{C}$ ). If  $P$  separates two points  $X$  and  $Y$ , then  $\tilde{P} = \lambda_1 E_1 + \lambda_2 E_2$  with eigenvalues  $\lambda_1 > 1/2 > \lambda_2$  and mutually orthogonal eigenprojectors  $E_1$  and  $E_2$ . The subset shaded in darker gray contains the points for which the Hilbert-Schmidt inner product with  $\tilde{P}$  is greater than  $1/2$ .

Conditioning on  $\lambda(\tilde{P})$  determines the opening angles of these two spherical caps, which are oriented along a random diameter determined by  $E(\tilde{P})$ . The projections  $X, Y$  are two fixed points on the Bloch sphere at an angle of  $2\theta$ , and are separated if and only if they are not in the same cap. Let  $w = \min\{\phi_h, \pi - \phi_h\}$ , which is the smallest opening angle of these two caps. If  $w \leq \theta$ , then any cap of angle  $w$  containing  $X$  cannot contain  $Y$  (and vice versa), so  $\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\}$  is just twice the normalized area of a cap of angle  $w$  (which is just its normalized height), i.e.

$$\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} = 1 - \cos(w) \leq 1 - \cos(\theta) \leq \sin(\theta) = \|X - Y\|.$$

If  $w > \theta$ , then it is possible for both  $X$  and  $Y$  to be in a cap of opening angle  $w$ . In this case,  $\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\}$  is just the normalized area of the symmetric difference of spherical caps of angle  $w$  centered at  $X$  and  $Y$ . The intersection of these two caps contains a spherical cap of angle  $w - \theta$  centered at the geodesic midpoint of  $X$  and  $Y$ , so for this case

$$\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} \leq \cos(w - \theta) - \cos(w) \leq \sin(\theta) = \|X - Y\|.$$

where the last inequality follows since  $w \leq \frac{\pi}{2}$ . Thus we have

$$\begin{aligned} \mathbb{E}\left[\mathbb{P}\{P \in \mathcal{S}_{X,Y} \mid \lambda(\tilde{P})\} \mathbb{1}_{\mathcal{D}_{\text{sep}}}(\lambda(\tilde{P}))\right] &\leq \|X - Y\| \mathbb{P}\{\lambda(\tilde{P}) \in \mathcal{D}_{\text{sep}}\} \\ &\leq \|X - Y\|. \end{aligned}$$

□

The uniform bound for the measurement Hamming distance in terms of the operator norm distance now follows directly by combining Theorem 3.2.6 with Proposition 3.2.9.

**Theorem 3.2.10.** *Let  $\delta > 0$ ,  $m \geq 2\delta^{-2} \left( 8\beta n \log \left( 1 + 128 \frac{\sqrt{2\beta n - 1}}{2\sqrt{2\pi}} \delta^{-1} \right) + \log(2) + D \right)$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  be a collection of independent uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ . Then with probability at least  $1 - \exp(-D)$*

$$d_{\mathcal{P}}(X, Y) \leq \|X - Y\| + \delta$$

for all  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$ .

**3.3. Uniform guarantees for accurate recovery.** With the results from Sections 3.1 and 3.2 we are ready to extend the pointwise result given in Theorem 2.3.3 to a uniform result that controls the behavior of our recovery procedure for all input vectors simultaneously.

**Theorem 3.3.1.** *Let  $\delta > 0$  and set  $\epsilon = \frac{(\mu_1 - \mu_2)}{8} \delta$ . If*

$$(10) \quad m \geq 2\epsilon^{-2} \left( 8\beta n \log \left( 1 + 128 \frac{\sqrt{2\beta n - 1}}{2\sqrt{2\pi}} \epsilon^{-1} \right) + 2\log(2) + D \right)$$

and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then with probability at least  $1 - \exp(-D)$

$$\|\hat{X} - X\| < \delta$$

for all  $X \in \text{Proj}(1, 2n)$ , where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$ .

*Proof.* Let  $\mathcal{N}_{\epsilon}$  be an  $\epsilon$ -net for  $\text{Proj}_{\mathbb{F}}(1, 2n)$  such that  $\log |\mathcal{N}_{\epsilon}| \leq 4\beta n \log(1 + 2\epsilon^{-1})$  as in Lemma 3.1.2. By our choice of  $m$ , Lemma 3.1.3 says that with probability greater than  $1 - \exp(-\log(2) - D)$  we have  $\|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \epsilon$  for all  $X \in \mathcal{N}_{\epsilon}$  (call this event  $\mathcal{A}$ ). Also by our choice of  $m$ , Theorem 3.2.10 says that with probability at least  $1 - \exp(-\log(2) - D)$  we have  $d_{\mathcal{P}}(X, Y) \leq \|X - Y\| + \epsilon$  for all  $X, Y \in \text{Proj}_{\mathbb{F}}(1, 2n)$  (call this event  $\mathcal{B}$ ).

Suppose that  $\mathcal{A}$  and  $\mathcal{B}$  both occur, which happens with probability at least  $1 - \exp(-D)$ , and consider an arbitrary  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$ . We know from (5) that

$$(11) \quad \|\hat{X} - X\| \leq 2(\mu_1 - \mu_2)^{-1} \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\|.$$

To bound the right-hand side of this last inequality we pass to the  $\epsilon$ -net  $\mathcal{N}_{\epsilon}$  by picking  $X_0 \in \mathcal{N}_{\epsilon}$  with  $\|X - X_0\| < \epsilon$ . Then

$$(12) \quad \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \|\hat{Q}_{\mathcal{P}}(X) - \hat{Q}_{\mathcal{P}}(X_0)\| + \|\hat{Q}_{\mathcal{P}}(X_0) - Q(X_0)\| + \|Q(X_0) - Q(X)\|.$$

Next, we examine each of the three terms on the right side of (112). To bound the first term, note that  $\left| \left\{ j : \hat{P}_j(X) \neq \hat{P}_j(X_0) \right\} \right| = m \cdot d_{\mathcal{P}}(X, X_0)$ . Using this and the assumption that  $\mathcal{A}$  holds yields

$$\|\hat{Q}_{\mathcal{P}}(X) - \hat{Q}_{\mathcal{P}}(X_0)\| = \left\| \frac{1}{m} \sum_{j: \hat{P}_j(X) \neq \hat{P}_j(X_0)} \hat{P}_j(X) - \hat{P}_j(X_0) \right\| \leq d_{\mathcal{P}}(X, X_0) \leq 2\epsilon.$$

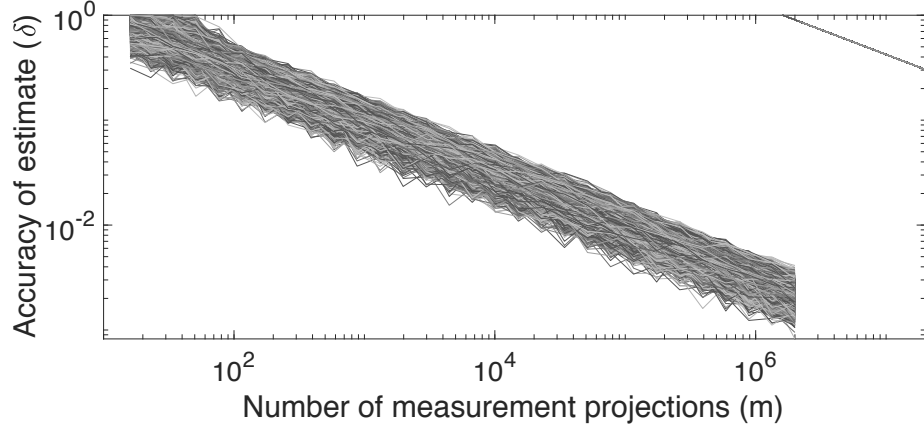


FIGURE 3. Plot showing the accuracy for the recovery of 15000 random inputs using (PEP) with a fixed collection of measurement projections on  $\mathbb{R}^{16}$ . The single line separate from the cluster represents the upper bound on  $\delta$  given by Theorem 3.3.1.

Since  $\mathcal{B}$  holds, we can bound the second term by  $\|\hat{Q}_{\mathcal{P}}(X_0) - Q(X_0)\| \leq \epsilon$ . Lastly, using Proposition 2.1.2 gives  $Q(X) - Q(X_0) = (\mu_1 - \mu_2)(X - X_0)$ , and so we can bound the third term by

$$\|Q(X_0) - Q(X)\| = (\mu_1 - \mu_2) \|X - X_0\| \leq (\mu_1 - \mu_2)\epsilon.$$

Using these three bounds together in (12) gives

$$\|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| \leq 3\epsilon + (\mu_1 - \mu_2)\epsilon \leq \frac{1}{2}(\mu_1 - \mu_2)\delta,$$

which combined with (11) yields  $\|\hat{X} - X\| < \delta$ .  $\square$

See Figure 3 for a plot showing how our bound on the sufficient number of measurements to achieve a uniform accuracy of  $\delta$  relates to experimental results.

As in the pointwise case, our proof allows us to fine tune the probability that a sequence of measurement projections provides uniformly accurate recovery by adjusting the value of  $D$  in (10). In particular, we can take  $D = n$  to ensure success with overwhelming probability, i.e. the failure rate decays exponentially in  $n$ . In the pointwise case, this resulted in gaining an additional factor of  $n$  in the number of measurement projections, see Corollary 2.3.5. In the uniform case, however, the asymptotics remain the same.

**Corollary 3.3.2.** *Let  $\delta > 0$ . If  $\alpha > 0$ ,  $m \geq C\delta^{-2}n^2 \log(\delta^{-1}n)$ , and  $\mathcal{P} = \{P_j\}_{j=1}^m$  is an independent sequence of uniformly distributed projections in  $\text{Proj}_{\mathbb{F}}(n, 2n)$ , then with probability at least  $1 - \exp(-n)$*

$$\|\hat{X} - X\| < \delta$$

for all  $X \in \text{Proj}(1, 2n)$ , where  $\hat{X}$  is the solution to (PEP) with input  $\Phi_{\mathcal{P}}(X)$  and  $C$  is a constant.

As in the pointwise case, we can modify the proof of Theorem 3.3.1 to show that uniform recovery using (PEP) is robust to bit-flip errors occurring in a faulty measurement  $\tilde{\Phi}_{\mathcal{P}}$ .

**Corollary 3.3.3.** *Let  $\delta$ ,  $m$ , and  $\{P_j\}_{j=1}^m$  be as in Theorem 3.3.1, and additionally let  $0 < \tau < 1$ . Then with probability at least  $1 - \exp(-D)$ , for all  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  and all  $\tilde{\Phi}_{\mathcal{P}}(X) \in \{0, 1\}^m$  with*

$$d_H(\Phi_{\mathcal{P}}(X), \tilde{\Phi}_{\mathcal{P}}(X)) \leq \tau$$

*we have*

$$\|\tilde{X} - X\| \leq \delta + 2(\mu_1 - \mu_2)^{-1}\tau,$$

*where  $\tilde{X}$  is the solution to (PEP) with input  $\tilde{\Phi}_{\mathcal{P}}(X)$  and  $\mu_1 - \mu_2$  is bounded by Lemma 2.3.2.*

*Proof.* Let  $\tilde{Q}_{\mathcal{P}}(X)$  denote the empirical average of the (faulty) flipped projections, i.e. flipped using  $\tilde{\Phi}_{\mathcal{P}}(X)$  rather than  $\Phi_{\mathcal{P}}(X)$ . As before, for all  $X \in \text{Proj}_{\mathbb{F}}(1, 2n)$  we have

$$\|\tilde{X} - X\| \leq 2(\mu_1 - \mu_2)^{-1} \|\tilde{Q}_{\mathcal{P}}(X) - Q(X)\|.$$

Using the triangle inequality, we expand

$$\|\tilde{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \|\tilde{Q}_{\mathcal{P}}(X) - \hat{Q}_{\mathcal{P}}(X)\| + \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\| \leq \tau + \|\hat{Q}_{\mathcal{P}}(X) - Q(X)\|.$$

Bounding  $\|\hat{Q}_{\mathcal{P}}(X) - Q(X)\|$  with high probability proceeds exactly as in Theorem 3.3.1.  $\square$

## REFERENCES

- [1] Albert Ai, Alex Lapanowski, Yaniv Plan, and Roman Vershynin. One-bit compressed sensing with non-Gaussian measurements. *Linear Algebra Appl.*, 441:222–239, 2014.
- [2] Edwin J. Akutowicz. On the determination of the phase of a Fourier integral. I. *Trans. Amer. Math. Soc.*, 83:179–192, 1956.
- [3] Edwin J. Akutowicz. On the determination of the phase of a Fourier integral. II. *Proc. Amer. Math. Soc.*, 8:234–238, 1957.
- [4] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley Publishing, 4th edition, 2016.
- [5] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [6] Christine Bachoc and Martin Ehler. Signal reconstruction from the magnitude of subspace components. *IEEE Trans. Inform. Theory*, 61(7):4015–4027, 2015.
- [7] Radu Balan, Pete Casazza, and Dan Edidin. On signal reconstruction without phase. *Appl. Comput. Harmon. Anal.*, 20(3):345–356, 2006.
- [8] Afonso S. Bandeira, Jameson Cahill, Dustin G. Mixon, and Aaron A. Nelson. Saving phase: injectivity and stability for phase retrieval. *Appl. Comput. Harmon. Anal.*, 37(1):106–125, 2014.
- [9] Richard Baraniuk, Mark Davenport, Ronald DeVore, and Michael Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, 28(3):253–263, Dec 2008.
- [10] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, Oct 1946.
- [11] Petros T. Boufounos and Richard G. Baraniuk. 1-bit compressive sensing. In *2008 42nd Annual Conference on Information Sciences and Systems*, pages 16–21, March 2008.
- [12] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.
- [13] Jameson Cahill, Peter G. Casazza, Jesse Peterson, and Lindsey Woodland. Phase retrieval by projections. *Houston J. Math.*, 42(2):537–558, 2016.
- [14] Emmanuel J. Candès, Yonina C. Eldar, Thomas Strohmer, and Vladislav Voroninski. Phase retrieval via matrix completion. *SIAM J. Imaging Sci.*, 6(1):199–225, 2013.
- [15] Emmanuel J. Candès, Thomas Strohmer, and Vladislav Voroninski. PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming. *Comm. Pure Appl. Math.*, 66(8):1241–1274, 2013.
- [16] Peter G. Casazza and Lindsey M. Woodland. Phase retrieval by vectors and projections. In *Operator methods in wavelets, tilings, and frames*, volume 626 of *Contemp. Math.*, pages 1–17. Amer. Math. Soc., Providence, RI, 2014.
- [17] Aldo Conca, Dan Edidin, Milena Hering, and Cynthia Vinzant. An algebraic characterization of injectivity in phase retrieval. *Appl. Comput. Harmon. Anal.*, 38(2):346–356, 2015.

- [18] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):83–91, 2008.
- [19] James R. Fienup. Reconstruction of an object from the modulus of its fourier transform. *Opt. Lett.*, 3(1):27–29, Jul 1978.
- [20] David Gross, Yi-Kai Liu, Steven T Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Physical review letters*, 105(15):150401, 2010.
- [21] Madalin Guta, Jonas Kahn, Richard Kueng, and Joel A Tropp. Fast state tomography with optimal error bounds. *arXiv preprint arXiv:1809.11162*, 2018.
- [22] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 2nd edition, 2013.
- [23] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Transactions on Information Theory*, 59(4):2082–2102, April 2013.
- [24] Philippe Jaming. Uniqueness results in an extension of pauli’s phase retrieval problem. *Applied and Computational Harmonic Analysis*, 37(3):413 – 441, 2014.
- [25] David Mendlovic and Haldun M. Ozaktas. Fractional fourier transforms and their optical implementation: I. *J. Opt. Soc. Am. A*, 10(9):1875–1881, Sep 1993.
- [26] Rick P. Millane. Phase retrieval in crystallography and optics. *J. Opt. Soc. Am. A*, 7(3):394–411, Mar 1990.
- [27] Youssef Mroueh and Lorenzo Rosasco. On efficiency and low sample complexity in phase retrieval. In *2014 IEEE International Symposium on Information Theory*, pages 931–935, June 2014.
- [28] Beresford N. Parlett. *The symmetric eigenvalue problem*, volume 20 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Corrected reprint of the 1980 original.
- [29] Yaniv Plan and Roman Vershynin. One-bit compressed sensing by linear programming. *Comm. Pure Appl. Math.*, 66(8):1275–1297, 2013.
- [30] Yaniv Plan and Roman Vershynin. Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach. *IEEE Trans. Inform. Theory*, 59(1):482–494, 2013.
- [31] Yaniv Plan and Roman Vershynin. Dimension reduction by random hyperplane tessellations. *Discrete Comput. Geom.*, 51(2):438–461, 2014.
- [32] Herbert Robbins. A remark on Stirling’s formula. *Amer. Math. Monthly*, 62:26–29, 1955.
- [33] A. J. Scott. Tight informationally complete quantum measurements. *Journal of Physics A-Mathematical and General*, 39(43):13507–13530, October 2006.
- [34] Joel A. Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1-2):1–230, 2015.
- [35] Adriaan Walther. The question of phase retrieval in optics. *Optica Acta*, 10:41–49, 1963.