

# Anisotropic twicing for single particle reconstruction using autocorrelation analysis

## Abstract

The missing phase problem in X-ray crystallography is commonly solved using the technique of molecular replacement (Rossmann & Blow, 1962; Rossmann, 2001; Scapin, 2013), which borrows phases from a previously solved homologous structure, and appends them to the measured Fourier magnitudes of the diffraction patterns of the unknown structure. More recently, molecular replacement has been proposed for solving the missing orthogonal matrices problem arising in Kam’s autocorrelation analysis (Kam, 1977; Kam, 1980) for single particle reconstruction using X-ray free electron lasers (Saldin *et al.*, 2009; Hosseinizadeh *et al.*, 2015; Starodub *et al.*, 2012) and cryo-EM (Bhamre *et al.*, 2015). In classical molecular replacement, it is common to estimate the magnitudes of the unknown structure as twice the measured magnitudes minus the magnitudes of the homologous structure, a procedure known as ‘twicing’ (Tukey, 1977). Mathematically, this is equivalent to finding an unbiased estimator for a complex-valued scalar (Main, 1979). We generalize this scheme for the case of estimating real or complex valued matrices arising in single particle autocorrelation analysis. We name this approach “Anisotropic Twicing” because unlike the scalar case,

the unbiased estimator is not obtained by a simple magnitude isotropic correction. We compare the performance of the least squares, twicing and anisotropic twicing estimators on synthetic and experimental datasets. We demonstrate 3D homology modeling in cryo-EM directly from experimental data without iterative refinement or class averaging, for the first time.

## 1. Introduction

The missing phase problem in crystallography entails recovering information about a crystal structure that is lost during the process of imaging. In X-ray crystallography, the measured diffraction patterns provide information about the modulus of the 3D Fourier transform of the crystal. The phases of the Fourier coefficients need to be recovered by other means, in order to reconstruct the 3D electron density map of the crystal. A popular method to solve the missing phase problem is Molecular Replacement (MR) (Rossmann & Blow, 1962; Rossmann, 2001; Scapin, 2013), which relies on a previously solved homologous structure which is similar to the unknown structure. The unknown structure is then estimated using the Fourier magnitudes of its diffraction data, along with phases from the homologous structure.

The missing phase problem can be formulated mathematically using matrix notation that enables generalization as follows. Each Fourier coefficient  $\mathbf{A}$  is a complex-valued scalar, i.e.,  $\mathbf{A} \in \mathbb{C}^{1 \times 1}$  that we wish to estimate, given measurements of  $\mathbf{C} = \mathbf{A}\mathbf{A}^*$  ( $\mathbf{A}^*$  denotes the complex conjugate transpose of  $\mathbf{A}$ , i.e.,  $\mathbf{A}_{ij}^* = \overline{A_{ji}}$ ), corresponding to the Fourier squared magnitudes, and  $\mathbf{B}$  corresponds to a previously solved homologous structure such that  $\mathbf{A} = \mathbf{B} + \mathbf{E}$ , where  $\mathbf{E}$  is a small perturbation. We denote an estimator of  $\mathbf{A}$  as  $\hat{\mathbf{A}}$ . There are many possible choices for such an estimator. One such choice is the solution to the least squares problem

$$\hat{\mathbf{A}}_{\text{LS}} = \arg \min_{\mathbf{A}} \|\mathbf{A} - \mathbf{B}\|_F, \text{ subject to } \mathbf{A}\mathbf{A}^* = \mathbf{C} \quad (1)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. However, it has been noticed that  $\hat{\mathbf{A}}_{\text{LS}}$  does not reveal the correct relative magnitude of the unknown part of the crystal structure, and the recovered magnitude is about half of the actual value. As a magnitude correction scheme, it was empirically found that setting the magnitude to be twice the experimentally measured magnitude minus the magnitude of the homologous structure has the desired effect of approximately resolving the issue. That is, the estimator  $2\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}$  is used instead. The theoretical advantage of this unbiased estimator for the case when  $\mathbf{A} \in \mathbb{C}^{1 \times 1}$  has been justified in (Main, 1979). Following (Tukey, 1977), we refer to this procedure as twicing.

The advantage of using twicing is demonstrated in the following illustrative toy experiment (Cowtan, 2014) for the 2D case. We start with an image of a cat with a tail, which is the unknown image that we want to recover. We are given the Fourier magnitudes of the unknown image, measured in an experiment. In analogy with a known homologous structure used in MR, we have access to a similar image, that of a cat, but with its tail missing. We show the results of retrieving the original image using least squares, with and without employing twicing for magnitude correction, and note that twicing restores the tail better than least squares (see Fig 1).

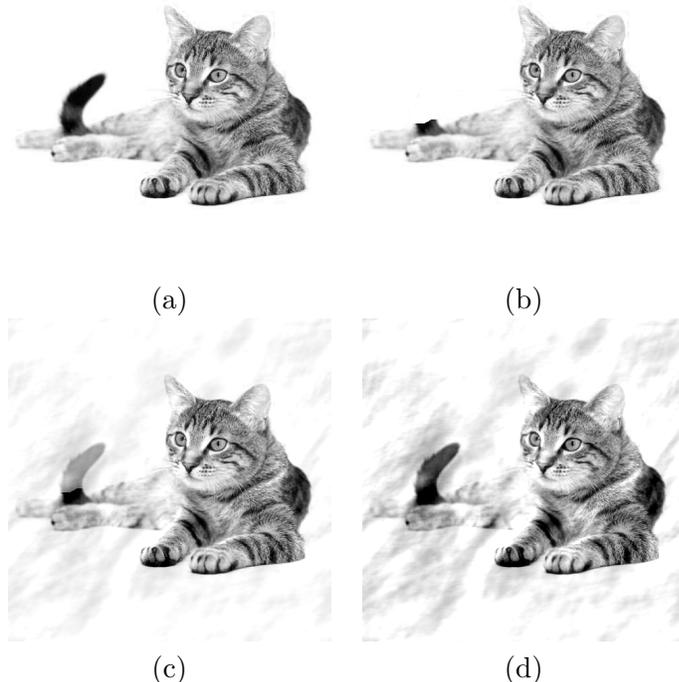


Fig. 1. Demonstrating twicing in MR through a toy example (Cowtan, 2014): given an unknown image whose Fourier magnitudes are known through measurements, but phases are missing, and a known similar image for which both the Fourier magnitudes and phases are completely known. (a) Original image: unknown phases, known magnitudes (b) Similar image: known phases and magnitudes (note that the tail is missing) (c) Least squares estimator of original image, no magnitude correction (d) Twicing for magnitude correction. Note that the tail is better restored when twicing is used.

As a natural generalization, one might wonder whether the estimator  $2\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}$  performs well for the non-scalar case  $\mathbf{A} \in \mathbb{R}^{N \times D}$  (or  $\mathbb{C}^{N \times D}$ ), where  $(N, D) \neq (1, 1)$ . In this paper, we consider the following problem: How to estimate  $\mathbf{A} \in \mathbb{R}^{N \times D}$  (or  $\mathbb{C}^{N \times D}$ ) from  $\mathbf{C}$  and  $\mathbf{B}$ , where  $\mathbf{C} = \mathbf{A}\mathbf{A}^*$  and  $\mathbf{A} = \mathbf{B} + \mathbf{E}$  for a matrix  $\mathbf{E}$  of small magnitude? When  $N = D$ , the result derived in this paper (see Sec. 4) for an “asymptotically consistent” estimator of  $\mathbf{A}$  is given by  $\hat{\mathbf{A}}_{\text{AT}} = \mathbf{B} + \mathbf{U}\mathbf{W}\mathbf{U}^*(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B})$  where  $\mathbf{U}$  and  $\mathbf{W}$  are defined in Sec. 4. In particular, when  $\mathbf{A} \in \mathbb{C}^{1 \times 1}$ , this result coincides with the result in (Main, 1979) and justifies the approach of twicing. A formal proof of the

result derived in this paper is provided later in Sec. 9.

The motivation to study this problem is 3D structure determination in single particle reconstruction (SPR) without estimating the viewing angle associated with each image. Although we focus on cryo-electron microscopy (cryo-EM) here, the methods in this paper can also be applied to SPR using X-ray free electron lasers (XFEL). In SPR using XFEL, short but intense pulses of X-rays are scattered from the molecule. The measured 2D diffraction patterns in random orientations are used to reconstruct the 3D diffraction volume by an iterative refinement procedure, akin to the approach in cryo-EM. Recently, there have been attempts to use Kam's theory for SPR using XFEL, to determine the 3D diffraction volume without any iterative refinement (Saldin *et al.*, 2009; Hosseinizadeh *et al.*, 2015; Starodub *et al.*, 2012).

In this paper, we revisit Orthogonal Extension (OE) in cryo-EM (Bhamre *et al.*, 2015) that combines ideas from MR and Kam's autocorrelation analysis (Kam, 1980; Kam & Gafni, 1985) for the purpose of 3D homology modeling, that is, for reconstruction of an unknown complex directly from its raw, noisy images when a previously solved similar complex exists. In SPR using cryo-EM (Kühlbrandt, 2014; Bai *et al.*, 2015; Henderson, 2004), the 3D structure of a macromolecule is reconstructed from its noisy, contrast transfer function (CTF) affected 2D projection images. Individual particle images are picked from micrographs, preprocessed, and used in further parts of the cryo-EM pipeline to obtain the 3D density map of the macromolecule. There exist many algorithms in popular cryo-EM software such as RELION, XMIPP, SPIDER, EMAN2, FREALIGN (Scheres, 2012; Marabini *et al.*, 1996; Shaikh *et al.*, 2008; Tang *et al.*, 2007; Grigorieff, 2007) that, given a starting 3D structure, refine it using the noisy 2D projection images. The result of the refinement procedure is often dependent on the choice of the initial model. It is therefore important to have a procedure to provide a good starting model for refinement. Also, a high quality

starting model may significantly reduce the computational time associated with the refinement procedure (although we note recent advances in fast refinement (Barnett *et al.*, 2016; Punjani *et al.*, 2017)). Such a high quality starting model can be obtained using OE. The main computational component of autocorrelation analysis is estimation of the covariance matrix of the 2D images. This computation requires only a single pass over the experimental images (Zhao *et al.*, 2016; Bhamre *et al.*, 2016). Autocorrelation analysis is therefore much faster than iterative refinement, which typically takes many iterations to converge. In fact, the computational cost of autocorrelation analysis is even lower than that of a single refinement iteration, as the latter involves comparison of image pairs (noisy raw images with volume projections). OE can also be used for the purpose of model validation, being a complementary method for structure prediction.

There are a few existing methods for ab-initio modeling. The random conical tilt method (Radermacher *et al.*, 1987) can be used when two electron micrographs, one tilted and one untilted, are acquired with the same field of view. There are two main approaches for ab-initio estimation that do not involve tilting. One approach is to use the method of moments, that leverages the second order moments of the unknown 3D volume to estimate the particle orientations, but it suffers from being very sensitive to errors in the data (Salzman, 1990; Goncharov, 1988). The other approach is based on using common-lines between images (van Heel, 1987; Vainshtein & Goncharov, 1986; Singer *et al.*, 2009; Singer & Shkolnisky, 2011). However, common-lines based approaches have not been successful in obtaining 3D ab-initio models directly from raw, noisy images without performing any class averaging to suppress the noise.

OE predicts the structure directly from the raw, noisy images without any averaging. The method is analogous to MR in X-ray crystallography for solving the missing phase problem. In OE, the homologous structure is used for estimating the missing

orthogonal matrices associated with the spherical harmonics expansion of the 3D structure in reciprocal space. It is important to note that the missing orthogonal matrices in OE are not associated with the unknown pose of the particles, but with the spherical harmonics expansion coefficients. The missing coefficient matrices are, in general, rectangular of size  $N \times D$ , which serves as the motivation to extend twicing to the general case of finding an estimator when  $(N, D) \neq (1, 1)$ .

The paper is organized as follows: First, we briefly review Kam’s theory for autocorrelation analysis and describe the problem of OE in cryo-EM in Sec. 2. Next, in Sec. 3, we describe the least squares solution to find an estimator to an unknown structure when we have noisy projection images of the unknown structure, and additional information about a homologous structure. In Sec. 4, we introduce Anisotropic Twicing as well as a family of estimators that interpolate between the least squares estimator and Anisotropic Twicing. We detail the procedure to estimate autocorrelation matrices and the algorithm of Orthogonal Extension with the Anisotropic Twicing estimator in Sec. 5. We benchmark the performance of these estimators through numerical experiments with synthetic and experimental datasets in Sec. 6. We provide a formal proof for asymptotic consistency of our Anisotropic Twicing correction scheme in the general case of  $(N, D) \neq (1, 1)$  in the appendix (see Sec. 9). The code for all the algorithms in this paper is available in the open source software toolbox, ASPIRE, available for download at `spr.math.princeton.edu`.

We apply anisotropic twicing to both synthetic and experimental cryo-EM datasets, and find that it recovers the unknown structure better than the least squares and twicing estimators on synthetic data. This is the first demonstration of reconstructing a starting 3D model in the presence of experimental conditions of CTF and noise without any class averaging, directly from raw images using the ‘Orthogonal Extension’ procedure (Bhamre *et al.*, 2015). While the anisotropic twicing estimator outperforms

other estimators on synthetic datasets, in the case of the experimental dataset the reconstructions from all estimators are similar in quality, and any of these reconstructions can be used as a good starting point for refinement.

## 2. Orthogonal Extension (OE) in Cryo-EM

In (Bhamre *et al.*, 2015), the authors presented two new approaches, collectively termed ‘Orthogonal Retrieval’ methods, for 3D homology modeling based on Kam’s theory (Kam, 1980). Orthogonal Retrieval can be regarded as a generalization of the MR method from X-ray crystallography to cryo-EM.

Let  $\Phi_A : \mathbb{R}^3 \rightarrow \mathbb{R}$  be the electron scattering density of the unknown structure, and let  $\mathcal{F}(\Phi_A) : \mathbb{R}^3 \rightarrow \mathbb{C}$  be its 3D Fourier transform. Consider the spherical harmonics expansion of  $\mathcal{F}(\Phi_A)$

$$\mathcal{F}(\Phi_A)(k, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l A_{lm}(k) Y_l^m(\theta, \varphi) \quad (2)$$

where  $k$  is the radial frequency and  $Y_l^m$  are the real spherical harmonics. Kam showed that the autocorrelation matrices

$$C_l(k_1, k_2) = \sum_{m=-l}^l A_{lm}(k_1) \overline{A_{lm}(k_2)}, \quad l = 0, 1, \dots \quad (3)$$

can be estimated from the covariance matrix of the 2D projection images whose viewing angles are uniformly distributed over the sphere. This can be achieved with both clean as well as noisy images, as long as the number of noisy images is large enough to allow estimation of the underlying population covariance matrix of the clean images to the desired level of accuracy, using (Bhamre *et al.*, 2016). The decomposition (3) suggests that the  $l$ ’th order autocorrelation matrix  $\mathbf{C}_l$  has a maximum rank of  $2l + 1$ , and the maximum rank is even smaller in the presence of symmetry.

While (2) is true if we want to represent the molecule to infinitely high resolution, in practice the images are sampled on a finite pixel grid and we cannot recover informa-

tion beyond the Nyquist frequency. In addition, the molecule is compactly supported in  $\mathbb{R}^3$ , and the support size can also be estimated from the images. It is therefore natural to expand the volume in a truncated basis of spherical Bessel functions or 3D prolates. This leads to

$$\mathcal{F}(\Phi_A)(k, \theta, \varphi) = \sum_{l=0}^L \sum_{m=-l}^l A_{lm}(k) Y_l^m(\theta, \varphi), \quad l = 0, 1, \dots, L \quad (4)$$

where the truncation  $L$  is based on the resolution limit that can be achieved by the reconstruction. Our specific choice of  $L$  is described after (8). We can expand  $A_{lm}(k)$  in a truncated basis of radial functions, chosen here as the spherical Bessel functions, as follows:

$$A_{lm}(k) = \sum_{s=1}^{S_l} a_{lms} j_{ls}(k). \quad (5)$$

Here the normalized spherical Bessel functions are

$$j_{ls}(k) = \frac{1}{c\sqrt{\pi}|j_{l+1}(R_{l,s})|} j_l(R_{l,s} \frac{k}{c}), \quad 0 < k < c, \quad s = 1, 2, \dots, S_l, \quad (6)$$

where  $c$  is the bandlimit of the images, and  $R_{l,s}$  is the  $s$ 'th positive root of the equation  $j_l(x) = 0$ . The functions  $j_{ls}$  are normalized such that

$$\int_0^c j_{ls}(k) j_{ls}^*(k) k^2 dk = 1 \quad (7)$$

The number of radial basis functions  $S_l$  in (5) is determined using the Nyquist criterion, similar to (Klug & Crowther, 1972; Zhao *et al.*, 2016), where it has been described for 2D images expanded in a Fourier-Bessel basis (rather than 3D volumes as done here). We assume that the 2D images, and hence the 3D volume, are compactly supported on a disk of radius  $R$  and have a bandlimit  $0 < c \leq 0.5$ . We require that the maximum of the inverse Fourier transform of the spherical Bessel function and its first zero after this maximum are both inside the sphere of compact support radius  $R$ . The truncation limit  $S_l$  in (5) is then defined by the sampling criterion as the largest

integer  $s$  that satisfies (Cheng, 2013)

$$R_{l,(s+1)} \leq 2\pi cR. \quad (8)$$

$L$  in (4) is the largest integer  $l$  for which (8) has only one solution, that is,  $S_l$  in (5) is at least 1. Each  $\mathbf{C}_l$  is a matrix of size  $S_l \times S_l$  when using the representation (5) in (3).  $S_l$  is a monotonically decreasing function of  $l$  with approximately linear decay that we compute numerically. In matrix notation, (3) can be written as

$$\mathbf{C}_l = \mathbf{A}_l \mathbf{A}_l^*, \quad (9)$$

where  $\mathbf{A}_l$  is a matrix of size  $S_l \times (2l + 1)$ , with  $A_l(s, m) = a_{lms}$  in (5). From (9), we note that  $\mathbf{A}_l$  can be obtained from the Cholesky decomposition of  $\mathbf{C}_l$  up to a unitary matrix  $\mathbf{U}_l \in \text{U}(2l + 1)$  (the group of unitary matrices of size  $(2l + 1) \times (2l + 1)$ ). Since  $\Phi_A$  is real-valued, one can show using properties of its Fourier transform together with properties of the real spherical harmonics, that  $A_{lm}(k)$  (and hence  $\mathbf{A}_l$ ) is real for even  $l$  and purely imaginary for odd  $l$ . So  $\mathbf{A}_l$  is unique up to an orthogonal matrix  $\mathbf{O}_l \in \text{O}(2l + 1)$  (the group of orthogonal matrices of size  $(2l + 1) \times (2l + 1)$ ). Determining  $\mathbf{O}_l$  is the orthogonal retrieval problem in (Bhamre *et al.*, 2015).

If  $S_l > 2l + 1$ , estimating the missing orthogonal matrix  $\mathbf{O}_l$  is equivalent to estimating  $\mathbf{A}_l$ . Since  $S_l$  is a decreasing function of  $l$ , for some large enough  $l$  we would have  $S_l < 2l + 1$ . For example, for the largest  $l = L$  where  $S_L = 1$ ,  $\mathbf{A}_L$  is of size  $1 \times (2L + 1)$ , that is, it has  $O(L)$  degrees of freedom. In such cases it does not make sense to estimate  $\mathbf{O}_L$  which has  $O(L^2)$  degrees of freedom. But we can still estimate  $\mathbf{A}_L$  closest to  $\mathbf{B}_L$  using (1).

### 3. The Least Squares Estimator

In this section we review the least squares estimator that was proposed in (Bhamre *et al.*, 2015). In order to determine the 3D Fourier transform  $\mathcal{F}(\Phi_A)$  and thereby

the 3D density  $\Phi_A$ , we need to determine the coefficient matrices  $\mathbf{A}_l$  of the spherical harmonic expansion. In OE, the coefficient matrices  $\mathbf{A}_l$  are estimated with the aid of a homologous structure  $\Phi_B$ . Suppose  $\Phi_B$  is a known homologous structure, whose 3D Fourier transform  $\mathcal{F}(\Phi_B)$  has the following spherical harmonic expansion:

$$\mathcal{F}(\Phi_B)(k, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l B_{lm}(k) Y_l^m(\theta, \varphi) \quad (10)$$

In practice, the homologous structure  $\Phi_B$  is available at some finite resolution, therefore only a finite number of coefficient matrices  $\mathbf{B}_l$  ( $l = 0, 1, \dots, L_B$ ) are given. We show how to estimate the unknown structure  $\Phi_A$  up to the resolution dictated by the input images and the resolution of the homologous structure through estimating the coefficient matrices  $\mathbf{A}_l$  for  $l = 0, 1, \dots, L_A$  where  $L_A = \min(L, L_B)$ .

Let  $\mathbf{F}_l$  be any matrix of size  $S_l \times 2l + 1$  satisfying  $\mathbf{C}_l = \mathbf{F}_l \mathbf{F}_l^*$ , determined from the Cholesky decomposition of  $\mathbf{C}_l$ . Then, using (9)

$$\mathbf{A}_l = \mathbf{F}_l \mathbf{O}_l \quad (11)$$

where  $\mathbf{O}_l \in O(2l + 1)$  (for  $S_l > 2l + 1$ ). Using the assumption that the structures are homologous,  $\mathbf{A}_l \approx \mathbf{B}_l$ , one can determine  $\mathbf{O}_l$  as the solution to the least squares problem

$$\mathbf{O}_l = \arg \min_{\mathbf{O} \in O(2l+1)} \|\mathbf{F}_l \mathbf{O} - \mathbf{B}_l\|_F^2, \quad (12)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm. Although the orthogonal group is non-convex, there is a closed form solution to (12) (see, e.g., (Keller, 1975)) given by

$$\mathbf{O}_l = \mathbf{V}_l \mathbf{U}_l^T, \quad (13)$$

where

$$\mathbf{B}_l^* \mathbf{F}_l = \mathbf{U}_l \mathbf{\Sigma}_l \mathbf{V}_l^T \quad (14)$$

is the singular value decomposition (SVD) of  $\mathbf{B}_l^* \mathbf{F}_l$ . Thus,  $\mathbf{A}_l$  can be estimated by the

following least squares estimator:

$$\hat{\mathbf{A}}_{l,\text{LS}} = \mathbf{F}_l \mathbf{V}_l \mathbf{U}_l^T. \quad (15)$$

Hereafter, we drop the subscript  $l$  for convenience, since the procedure can be applied to each  $l$  separately.

### 3.1. Algorithm 1: Orthogonal Extension by Least Squares

- 
- 1: **procedure** ORTHOGONAL EXTENSION BY LEAST SQUARES (OE-LS): ESTIMATE  $\mathbf{A}$  GIVEN  $\mathbf{B} \approx \mathbf{A}$ , SUBJECT TO  $\mathbf{C} = \mathbf{A}\mathbf{A}^*$
  - 2:
  - Input:**  $\mathbf{B} \in \mathbb{C}^{N \times D}$ ,  $\mathbf{C} \in \mathbb{C}^{N \times N}$
  - 3: Cholesky decomposition of  $\mathbf{C}$  to find an  $\mathbf{F} \in \mathbb{C}^{N \times D}$  such that  $\mathbf{C} = \mathbf{F}\mathbf{F}^*$
  - 4: Calculate  $\mathbf{B}^*\mathbf{F}$  and its singular value decomposition  $\mathbf{B}^*\mathbf{F} = \mathbf{U}_0 \mathbf{\Sigma}_0 \mathbf{V}_0^*$ .
  - 5: The estimator is  $\hat{\mathbf{A}}_{\text{LS}} = \mathbf{F}\mathbf{V}_0\mathbf{U}_0^*$ .
- 

## 4. Unbiased Estimator: Anisotropic Twicing

The case that  $\mathbf{A}$  is a complex-valued scalar, i.e.,  $\mathbf{A} \in \mathbb{C}^{1 \times 1}$  has been studied in X-ray crystallography. The theoretical advantage of the unbiased estimator  $2\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}$  for this case was elucidated in (Main, 1979). As a natural generalization, one may wonder whether the estimator  $2\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}$  is also unbiased for  $(N, D) \neq (1, 1)$ . We assume that  $\mathbf{A}$  is sampled from the model  $\mathbf{A} = \mathbf{F}\mathbf{V}$ , where  $\mathbf{F}\mathbf{F}^* = \mathbf{C}$  and  $\mathbf{V}$  is a random orthogonal matrix (or a random unitary matrix) sampled from the uniform distribution with Haar measure over the orthogonal group when  $\mathbf{A}$  is a real-valued matrix or the unitary group (when  $\mathbf{A}$  is a complex-valued matrix). This probabilistic model is reasonable for (1), because when  $\mathbf{A}\mathbf{A}^*$  is given,  $\mathbf{F}$  is known and  $\mathbf{V}$  is an unknown orthogonal or unitary matrix, that is, we have no prior information about  $\mathbf{V}$ . In addition, we assume that  $\mathbf{B}$  is a matrix close to  $\mathbf{A}$  such that  $\mathbf{A} - \mathbf{B}$  is fixed. Our goal is to find an unbiased estimator of  $\mathbf{A}$  which is an affine transformation of

$\hat{\mathbf{A}}_{\text{LS}}$ . The main result is as follows:

**Theorem 4.1.** *When  $N = D$ , assuming that the spectral decomposition of  $\mathbf{C}$  is given by  $\mathbf{C} = \mathbf{U} \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_D) \mathbf{U}^*$ , then using our probabilistic model we have*

$$\mathbb{E}[\mathbf{A} - \hat{\mathbf{A}}_{\text{LS}}] = \mathbf{UTU}^*(\mathbf{A} - \mathbf{B}) + o(\|\mathbf{A} - \mathbf{B}\|_F), \quad (16)$$

where  $\mathbf{T}$  is a diagonal matrix with  $i$ -th diagonal entry given by

$$\mathbf{T}_{ii} = \begin{cases} \frac{1}{D} \left[ -\frac{1}{2} + \sum_{1 \leq j \leq D} \frac{\lambda_i^2}{\lambda_i^2 + \lambda_j^2} \right] & \text{when } \mathbf{A}, \mathbf{C} \in \mathbb{R}^{D \times D}, \\ \frac{1}{D} \sum_{1 \leq j \leq D} \frac{\lambda_i^2}{\lambda_i^2 + \lambda_j^2} & \text{when } \mathbf{A}, \mathbf{C} \in \mathbb{C}^{D \times D}, \end{cases}$$

and  $f(\mathbf{X}) = o(\|\mathbf{X}\|_F)$  means that  $\limsup_{\|\mathbf{X}\|_F \rightarrow 0} f(\mathbf{X}) / \|\mathbf{X}\|_F \rightarrow 0$ .

From (16), we have

$$(-\mathbf{I} + \mathbf{UTU}^*)(\mathbf{A} - \mathbf{B}) = \mathbf{B} - \mathbb{E}[\hat{\mathbf{A}}_{\text{LS}}] + o(\|\mathbf{A} - \mathbf{B}\|_F)$$

and an ‘‘asymptotically consistent’’ estimator of  $\mathbf{A}$  is given by

$$\hat{\mathbf{A}}_{\text{AT}} = \mathbf{B} - (\mathbf{I} - \mathbf{UTU}^*)^{-1}(\mathbf{B} - \hat{\mathbf{A}}_{\text{LS}}) = \mathbf{B} + \mathbf{UWU}^*(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}), \quad (17)$$

where  $\mathbf{W} = (\mathbf{I} - \mathbf{T})^{-1}$ .

A formal proof of Theorem 4.1 is provided in the appendix (Sec. 9). In particular, when  $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{1 \times 1}$ , the matrices reduce to scalars:  $\mathbf{U} = 1$ ,  $\mathbf{T} = \frac{1}{2}$ ,  $\mathbf{W} = 2$  and  $\hat{\mathbf{A}}_{\text{AT}} = \mathbf{B} + 2(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B})$ . This result coincides with the result in (Main, 1979) and justifies the approach of ‘‘twicing’’.

#### 4.1. A family of estimators

From (16) it follows that

$$\mathbb{E}[\mathbf{A}] = \mathbb{E}[\hat{\mathbf{A}}_{\text{LS}}] + \mathbf{UTU}^*(\mathbf{A} - \mathbf{B}) + o(\|\mathbf{A} - \mathbf{B}\|_F). \quad (18)$$

Following the spirit of Tukey’s twicing, we could approximate  $\mathbf{A}$  in the RHS of (18)

by  $\hat{\mathbf{A}}_{\text{LS}}$ , which leads to a new estimator

$$\hat{\mathbf{A}}_{\text{T}}^{(1)} = \hat{\mathbf{A}}_{\text{LS}} + \mathbf{UTU}^*(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}).$$

In fact, there exists a family of estimators by approximating  $\mathbf{A}$  recursively in the RHS of (18) by  $\hat{\mathbf{A}}_{\mathbf{T}}^{(t-1)}$  (with  $\hat{\mathbf{A}}_{\mathbf{T}}^{(0)} = \hat{\mathbf{A}}_{\text{LS}}$ ):

$$\hat{\mathbf{A}}_{\mathbf{T}}^{(t)} = \hat{\mathbf{A}}_{\text{LS}} + \mathbf{U}\mathbf{T}\mathbf{U}^*(\hat{\mathbf{A}}_{\mathbf{T}}^{(t-1)} - \mathbf{B}). \quad (19)$$

This family of estimators can be explicitly written as

$$\hat{\mathbf{A}}_{\mathbf{T}}^{(t)} = \mathbf{B} + \mathbf{U}(\mathbf{I} + \mathbf{T} + \mathbf{T}^2 + \cdots + \mathbf{T}^t)\mathbf{U}^*(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B}). \quad (20)$$

Using  $\mathbf{W} = (\mathbf{I} - \mathbf{T})^{-1} = \sum_{i=0}^{\infty} \mathbf{T}^i$ , we have that  $\hat{\mathbf{A}}_{\mathbf{T}}^{(t)} \rightarrow \hat{\mathbf{A}}_{\text{AT}}$  as  $t \rightarrow \infty$ .

In general, this family of estimators has smaller variance than  $\hat{\mathbf{A}}_{\text{AT}}$ , but larger bias since they are not unbiased (see Fig. 2).

#### 4.2. Generalization to the setting $N \neq D$

If  $N > D$ , then the column space of  $\mathbf{A}$  is the same as the column space of  $\mathbf{C}$ . Let  $\mathbf{P}$  be the projector of size  $N \times D$  to this column space, then we have  $\mathbf{A} = \mathbf{P}\mathbf{P}^*\mathbf{A}$ . As a result, to find an unbiased estimator of  $\mathbf{A}$ , it is sufficient to find an unbiased estimator of  $\mathbf{P}^*\mathbf{A}$ , which is a square matrix. Since  $\mathbf{P}^*\mathbf{A}$  is close to  $\mathbf{P}^*\mathbf{B}$  and  $(\mathbf{P}^*\mathbf{A})(\mathbf{P}^*\mathbf{A})^* = \mathbf{P}^*\mathbf{C}\mathbf{P}^*$  is known, Theorem 4.1 is applicable, and an unbiased estimator of  $\mathbf{P}^*\mathbf{A}$  can be obtained through (17), with  $\mathbf{B}$  replaced by  $\mathbf{P}^*\mathbf{B}$  and  $\mathbf{C}$  replaced by  $\mathbf{P}^*\mathbf{C}\mathbf{P}^*$ . In summary, an unbiased estimator of  $\mathbf{A}$  can be obtained in two steps:

1. Find  $\hat{\mathbf{A}}_{\text{AT}}^{(0)}$ , an unbiased estimator of  $\mathbf{P}^*\mathbf{A}$ , by applying (17), with  $\mathbf{B}$  replaced by  $\mathbf{P}^*\mathbf{B}$  and  $\mathbf{C}$  replaced by  $\mathbf{P}^*\mathbf{C}\mathbf{P}^*$ .
2. An unbiased estimator of  $\mathbf{A}$  is obtained by  $\hat{\mathbf{A}}_{\text{AT}} = \mathbf{P}\hat{\mathbf{A}}_{\text{AT}}^{(0)}$ .

If  $N < D$ , we use the following heuristic estimator. Let  $\mathbf{P}$  be a matrix of size  $D \times N$  that is the projector to the row space of  $\mathbf{B}$ , and assuming that  $\hat{\mathbf{A}}$ , the estimator of  $\mathbf{A}$ , has the same row space as  $\mathbf{B}$ , then  $\hat{\mathbf{A}} = \hat{\mathbf{A}}\mathbf{P}\mathbf{P}^*$ , and it is sufficient to find  $\hat{\mathbf{A}}\mathbf{P}$ , an estimator of  $\mathbf{A}\mathbf{P}$ . With  $(\mathbf{A}\mathbf{P})(\mathbf{A}\mathbf{P})^* = \mathbf{A}\mathbf{A}^* = \mathbf{C}$  known and the fact that  $\mathbf{A}\mathbf{P}$

is close to  $\mathbf{BP}$ , we may use the estimator (17). In summary, we use the following procedure:

1. Find  $\hat{\mathbf{A}}_{\text{AT}}^{(0)}$ , an estimator of  $\mathbf{AP}$ , by applying the estimator (17), with  $\mathbf{B}$  replaced by  $\mathbf{BP}$ .
2. An estimator of  $\mathbf{A}$  is obtained by  $\hat{\mathbf{A}}_{\text{AT}} = \hat{\mathbf{A}}_{\text{AT}}^{(0)}\mathbf{P}^*$ .

We remark that for  $N < D$  there is no theoretical guarantee to show that it is an unbiased estimator, unlike the setting  $N \geq D$ . However, the assumption that  $\hat{\mathbf{A}}$  has the same column space as  $\mathbf{B}$  is reasonable, and the proposed estimator performs well in practice.

## 5. Estimation of the Covariance and Autocorrelation Matrices

The autocorrelation matrices  $\mathbf{C}_l$  in Kam's theory are derived from the covariance matrix  $\mathbf{\Sigma}$  of the 2D Fourier transformed projection images through (Kam, 1980)

$$\mathbf{C}_l(|k_1|, |k_2|) = 2\pi(2l + 1) \int_0^\pi \Sigma(|k_1|, |k_2|, \psi) P_l(\cos \psi) \sin \psi d\psi \quad (21)$$

where  $\psi$  is the angle between the vectors  $k_1$  and  $k_2$  in the x-y plane. We estimate the covariance matrix  $\mathbf{\Sigma}$  of the underlying 2D Fourier transformed clean projection images using the method described in (Bhamre *et al.*, 2016). This estimation method provides a more accurate covariance compared to the classical sample covariance matrix (van Heel & Frank, 1981; van Heel, 1984). First, it corrects for the CTF. Second, it performs eigenvalue shrinkage, which is critical for high dimensional statistical estimation problems. Third, it exploits the block diagonal structure of the covariance matrix in a steerable basis, a property that follows from the fact that any experimental image is just as likely to appear in different in-plane rotations. A steerable basis consists of outer products of radial functions (such as Bessel functions) and Fourier angular

modes. Each block along the diagonal corresponds to a different angular frequency (Zhao & Singer, 2013). Moreover, the special block diagonal structure facilitates fast computation of the covariance matrix (Zhao *et al.*, 2016).

Since the autocorrelation matrix  $\mathbf{C}_l$  estimated from projection images can have a rank exceeding  $2l+1$ , we first find its best rank  $2l+1$  approximation via singular value decomposition, before computing its Cholesky decomposition. In the case of symmetric molecules, we use the appropriate rank as dictated by classical representation theory of  $SO(3)$  (Klein, 1914; Cheng, 2013) (less than  $2l+1$ ).

### 5.1. Algorithm 2: Orthogonal Extension by Anisotropic Twicing

- 
- 1: **procedure** ORTHOGONAL EXTENSION BY ANISOTROPIC TWICING (OE-AT):  
 ESTIMATE  $\mathbf{A}$  GIVEN  $\mathbf{B} \approx \mathbf{A}$ , SUBJECT TO  $\mathbf{C} = \mathbf{A}\mathbf{A}^*$
- 2:
- Input:**  $\mathbf{B} \in \mathbb{C}^{N \times D}$ ,  $\mathbf{C} \in \mathbb{C}^{N \times N}$
- 3: Find any  $\mathbf{F} \in \mathbb{C}^{N \times D}$  such that  $\mathbf{C} = \mathbf{F}\mathbf{F}^*$
- 4: Calculate  $\mathbf{B}^*\mathbf{F}$  and calculate its singular value decomposition  $\mathbf{B}^*\mathbf{F} = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{V}_0^*$ .
- 5: Calculate the OE-LS estimator is  $\hat{\mathbf{A}}_{\text{LS}} = \mathbf{F}\mathbf{V}_0\mathbf{U}_0^*$ , (see Algorithm 1).
- 6: For  $N = D$ , the OE-AT estimator is given by  $\hat{\mathbf{A}}_{\text{AT}} = \mathbf{B} + \mathbf{U}\mathbf{W}\mathbf{U}^*(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{B})$ .
- 7: For  $N > D$ , assuming that  $\mathbf{P}$  is the projector of size  $N \times D$  to the  $D$ -dimensional subspace spanned by the columns of  $\mathbf{C}$ ,  $\hat{\mathbf{A}}_{\text{AT}} = \mathbf{P}\hat{\mathbf{A}}_{\text{AT}}^{(0)}$ .
- 8: For  $N < D$ , assuming that  $\mathbf{P}$  is the projector of size  $D \times N$  to the  $N$ -dimensional subspace in  $\mathbb{R}^D$  spanned by the rows of  $\mathbf{B}$ ,  $\hat{\mathbf{A}}_{\text{AT}} = \hat{\mathbf{A}}_{\text{AT}}^{(0)}\mathbf{P}^*$ .
- 

## 6. Numerical Experiments

### 6.1. Bias Variance Trade-off

For any parameter  $\theta$ , the performance of its estimator  $\hat{\theta}$  can be measured in terms of its mean squared error (MSE),  $\mathbb{E}||\theta - \hat{\theta}||^2$ . The MSE of any estimator can be decomposed into its bias and variance:

$$\text{MSE} = \mathbb{E}||\theta - \hat{\theta}||^2 = ||\text{Bias}||^2 + \text{Var} \quad (22)$$

where

$$\text{Bias} = \mathbb{E}[\hat{\theta}] - \theta \quad (23)$$

and

$$\text{Var} = \mathbb{E}[\|\hat{\theta} - \mathbb{E}[\hat{\theta}]\|^2] \quad (24)$$

Unbiased estimators are often not optimal in terms of MSE, but they can be valuable for being unbiased. We performed a numerical experiment starting with a fixed  $\mathbf{F} \in \mathbb{R}^{10 \times 10}$  and an unknown matrix  $\mathbf{A} = \mathbf{F}\mathbf{O}$  where  $\mathbf{O}$  is a random orthogonal matrix. We are given a known similar matrix  $\mathbf{B}$  such that  $\mathbf{A} = \mathbf{B} + \mathbf{E}$ . The goal is estimate  $\mathbf{A}$  given  $\mathbf{B}$  and  $\mathbf{F}$ . Figure 2 shows a comparison of the bias and root mean squared error (RMSE) of different estimators averaged over 10000 runs of the numerical experiment: the anisotropic twicing estimator, the twicing estimator, the least squares estimator, and estimators from the family of estimators for some values of  $t$  in Sec. 4.1. The figure demonstrates that the AT estimator is asymptotically unbiased at the cost of higher MSE.

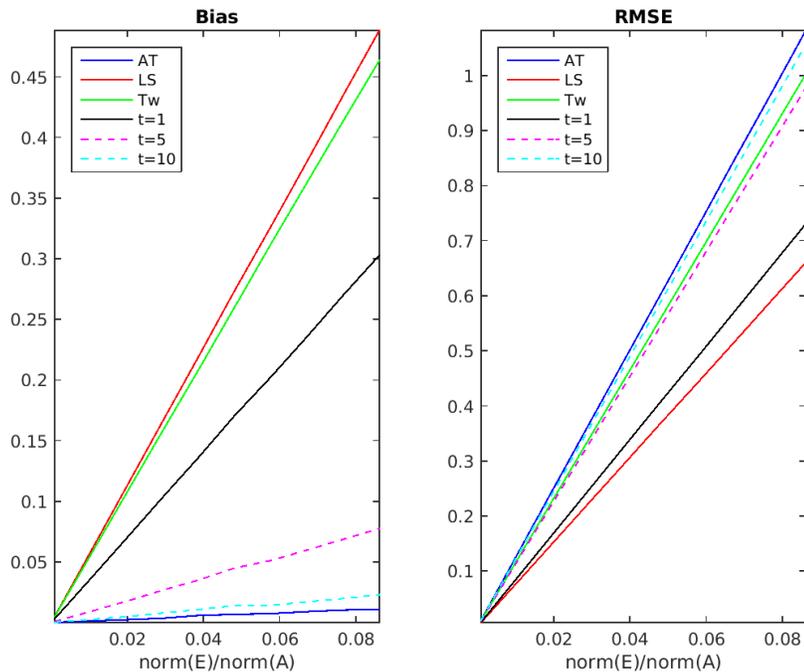


Fig. 2. Bias and RMSE of the Anisotropic Twicing (AT), Least Squares (LS), Twicing (Tw) estimators and also the family of estimators with  $t = 1, 5, 10$  averaged over 10000 experiments, as described in Sec. 6.1. The x-axis shows the relative perturbation  $\|\mathbf{E}\|/\|\mathbf{A}\|$ .

### 6.2. Synthetic Dataset: Toy Molecule

We perform numerical experiments with a synthetic dataset generated from an artificial ‘Mickey Mouse’ molecule. The molecule  $\mathbf{B}$  is made up of ellipsoids, and the density is set to 1 inside the ellipsoids and to 0 outside. Fig. 3 shows the artificial new volume  $\mathbf{A} = \mathbf{B} + \mathbf{E}$  created by adding a small ellipsoid  $\mathbf{E}$ , which we will refer to here as the “nose”, to the original mickey mouse volume  $\mathbf{B}$ . This represents the small perturbation  $\mathbf{E}$ . When the Fourier volume  $\mathbf{B} + \mathbf{E}$  is expanded in the truncated spherical Bessel basis described in Sec. 5, the average relative perturbation  $\|\mathbf{E}_l\|/\|\mathbf{A}_l\|$  for the first few coefficients in the truncated spherical harmonic expansion for  $l = 1, \dots, 10$  is 8%.

Next, we generate 10000 projection images from the volume  $\mathbf{A}$ . We then employ OE to reconstruct the volume  $\mathbf{A}$  from  $\mathbf{B}$  and the clean projection images of  $\mathbf{A}$ . Fig. 3 shows the reconstructions obtained using each of the three estimators in the OE framework, visualized in Chimera (Pettersen *et al.*, 2004). We note that while all three estimators are able to recover the additional subunit  $\mathbf{E}$ , the AT estimator best recovers the unknown subunit to its correct relative magnitude. The relative error in the region of the unknown subunit  $\mathbf{E}$  is 59% with least squares, 31% with twicing and 19% with anisotropic twicing.

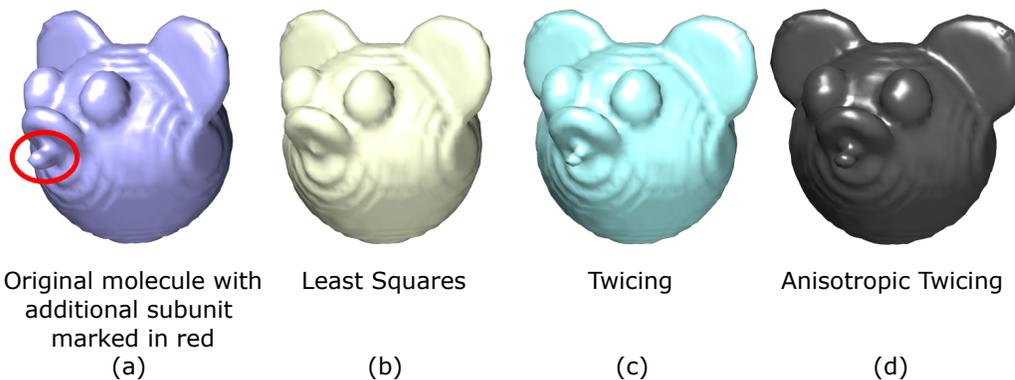


Fig. 3. A synthetic toy mickey mouse molecule with a small additional subunit, marked ‘E’ in (a). We reconstruct the molecule  $\mathbf{A}$  from its clean projection images, given  $\mathbf{B}$ . We show reconstructions obtained with the least squares estimator in (b), twicing estimator in (c), and AT estimator in (d).

### 6.3. Synthetic Dataset: TRPV1

We perform numerical experiments with a synthetic dataset generated from the TRPV1 molecule (with imposed  $C_4$  rotational symmetry) in complex with DkTx and RTX ( $\mathbf{A}$ ). This volume is available on EMDB as EMDB-8117. The small additional subunit is visible as an extension over the top of the molecule, shown in Fig. 4(i). This represents the small perturbation  $\mathbf{E}$ .

Next, we generate 26000 projection images from  $\mathbf{A}$ , add the effect of both the CTF (the images are divided into 10 defocus groups) and additive white Gaussian

noise (SNR=1/40) and use OE to reconstruct the volume  $\mathbf{A}$ . Fig. 4(ii) shows the reconstructions obtained using each of the three estimators in the OE framework, visualized in Chimera. The  $C_4$  symmetry was taken into account in the autocorrelation analysis by including in (2) only symmetry-invariant spherical harmonics  $Y_l^m$  for which  $m = 0 \pmod 4$ . As seen earlier with the synthetic case, all three estimators are able to recover the additional subunit  $\mathbf{E}$ , while the AT estimator best recovers the unknown subunit to its correct relative magnitude. The relative error in the unknown subunit is 43% with least squares, 56% with twicing and 30% with anisotropic twicing. We note that this is the first successful attempt, even with synthetic data, at using OE for 3D homology modeling directly from CTF-affected and noisy images (at experimentally relevant conditions). The numerical experiments using the Kv1.2 potassium channel in (Bhamre *et al.*, 2015) were at an unrealistically high SNR and did not include the effect of the CTF. The reason for this improvement is the improved covariance estimation (Bhamre *et al.*, 2016).

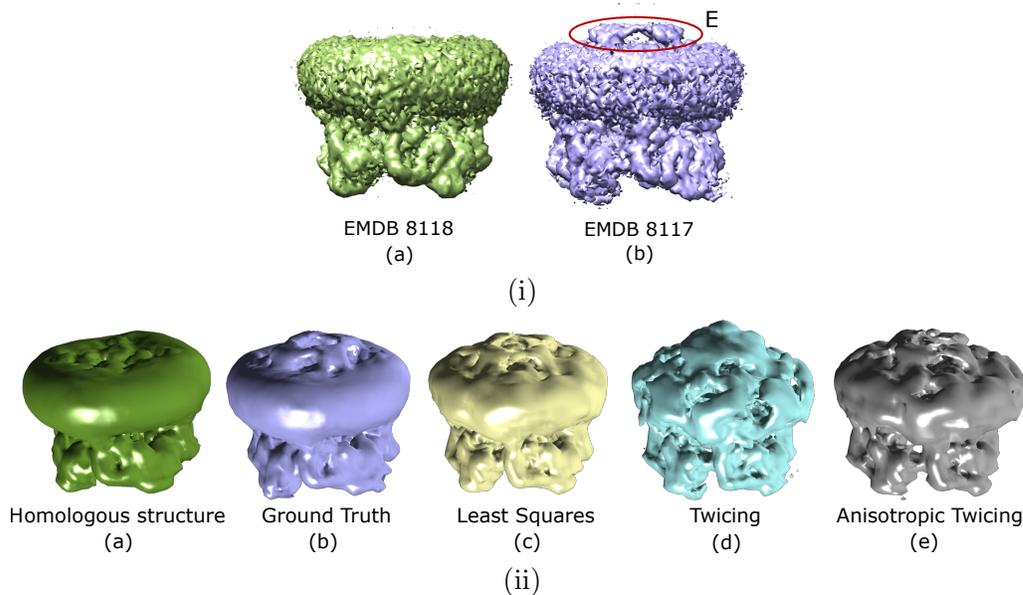


Fig. 4. A synthetic TRPV1 molecule (EMDB 8118), with a small additional subunit DxTx and RTX (EMDB 8117), marked ‘E’ in (i-b). We reconstruct the molecule from its noisy, CTF-affected images, and the homologous structure. In (ii), we show reconstructions obtained with the least squares, twicing and AT estimators using OE, along with the homologous structure and the ground truth projected on to the basis in (ii-a) and (ii-b).

#### 6.4. Experimental Dataset: TRPV1

We apply OE to an experimental data of the TRPV1 molecule in complex with DkTx and RTX, determined in lipid nanodisc, available on the public database Electron Microscopy Pilot Image Archive (EMPIAR) as EMPIAR-10059, and the 3D reconstruction is available on the electron microscopy data bank (EMDB) as EMDB-8117, courtesy of Y. Gao et al (Gao *et al.*, 2016). The dataset provided consists of 73000 motion corrected, picked particle images (which were used for the reconstruction in EMDB-8117) of size  $192 \times 192$  with a pixel size  $1.3\text{\AA}$ . We use the 3D structure of TRPV1 alone as the similar molecule. This is available on the EMDB as EMDB-8118. The two structures differ only by the small DkTx and RTX subunit at the top, which can be seen in Fig. 4(i).

Since the noise in experimental images is colored while our covariance estimation procedure requires white noise, we first preprocess the raw images in order to “whiten” the noise. We estimate the power spectrum of noise using the corner pixels of all images. The images are then whitened using the estimated noise power spectrum.

In the context of our mathematical model, the volume EMDB-8117 of TRPV1 with DkTx and RTX is the unknown volume  $\mathbf{A}$ , and the volume EMDB-8118 of TRPV1 alone is the known, similar volume  $\mathbf{B}$ . We use OE to estimate  $\mathbf{A}$  given  $\mathbf{B}$  and the raw, noisy projection images of  $\mathbf{A}$  from an experimental dataset.

The basis assumption in Kam’s theory is that the distribution of viewing angles is uniform. This assumption is difficult to satisfy in practice, since molecules in the sample can often have preference for certain orientations due to their shape and mass distribution. The viewing angle distribution in EMPIAR-10059 is non-uniform (see Fig. 6). As a robustness test of our methods, we attempt 3D reconstruction with (i) all images, such that the viewing angle distribution is non-uniform, as well as (ii) by sampling images such that the viewing angle distribution of the images is approximately uniform (as shown in Fig. 6). We obtained the final viewing angles estimated after refinement from the Cheng lab at UCSF (Gao *et al.*, 2016). Our sampling procedure is as follows: we choose 10000 points at random from the uniform distribution on the sphere and classify each image into these 10000 bins based on the point closest to it. We discard bins that have no images, and for the remaining bins we pick a maximum of 3 points per bin. We use the selected images (slightly less than 30000) for reconstruction with roughly uniform distributed viewing angles.

The reconstructed 3D volumes are shown in Fig. 5. We note that the additional subunit is recovered at the right location, and roughly to the expected size, using all three estimators. This is the first instance of reconstructing a 3D model directly from raw experimental images, without any class averaging or iterative refinement,

by employing OE. The Fourier cross resolution (FCR) of the reconstruction with the ‘ground truth’ EMDB-8117 is shown in Fig. 7.

The algorithm is implemented in the UNIX environment, on a machine with 60 cores, running at 2.3 GHz, with total RAM of 1.5TB. Using 20 cores, the total time taken here for preprocessing (whitening, background normalization. etc.) the 2D images and computing the covariance matrix was 1400 seconds. Calculating the autocorrelation matrices using (21) involves some numerical integration (eq. 7.15 in (Cheng, 2013)) which took 790 seconds, but for a fixed  $c$  and  $R$  (satisfied for datasets of roughly similar size and quality) these can be precomputed. Computing the basis functions and calculating the coefficient matrices  $\mathbf{A}_l$  of the homologous structure took 30 seconds and recovering the 3D structure by applying the appropriate estimator (AT, twicing, or LS) and computing the volume from the estimated coefficients took 10 seconds.

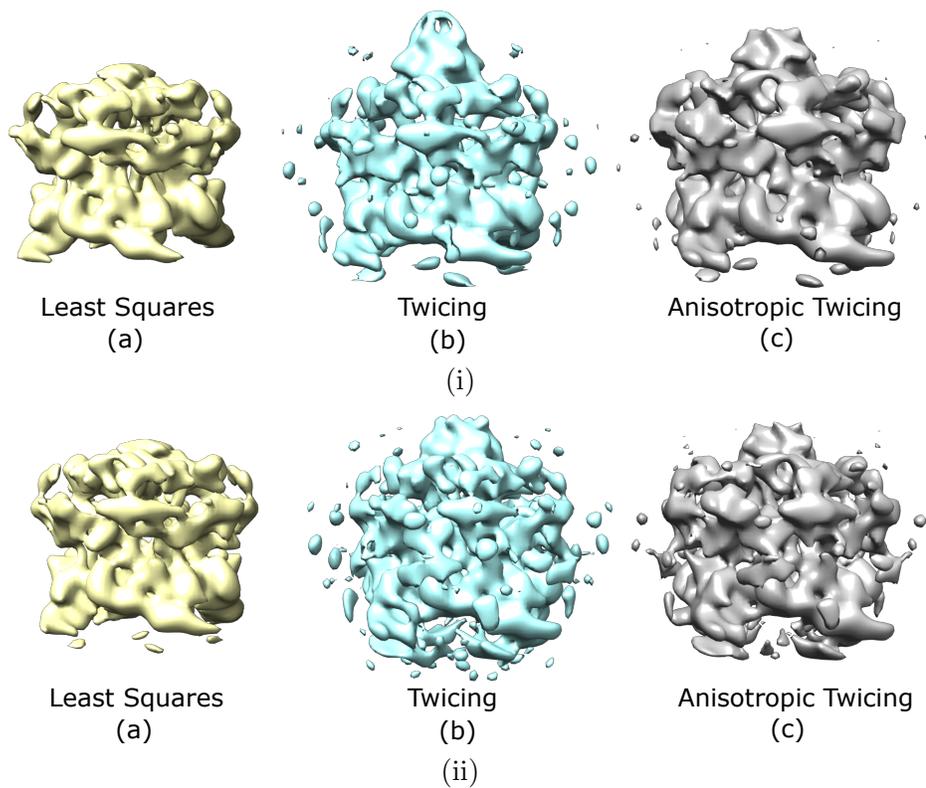


Fig. 5. OE with an experimental data of the TRPV1 in complex with DkTx and RTX (EMPIAR-10059) whose 3D reconstruction is available as EMD-8117. 3D reconstructions with OE using the least squares, anisotropic twicing, and twicing estimators: (i) With (slightly less than 30000) images selected by sampling to impose approximately uniform viewing angle distribution (ii) With all 73000 images such that the viewing angle distribution is non-uniform (see Fig. 6).

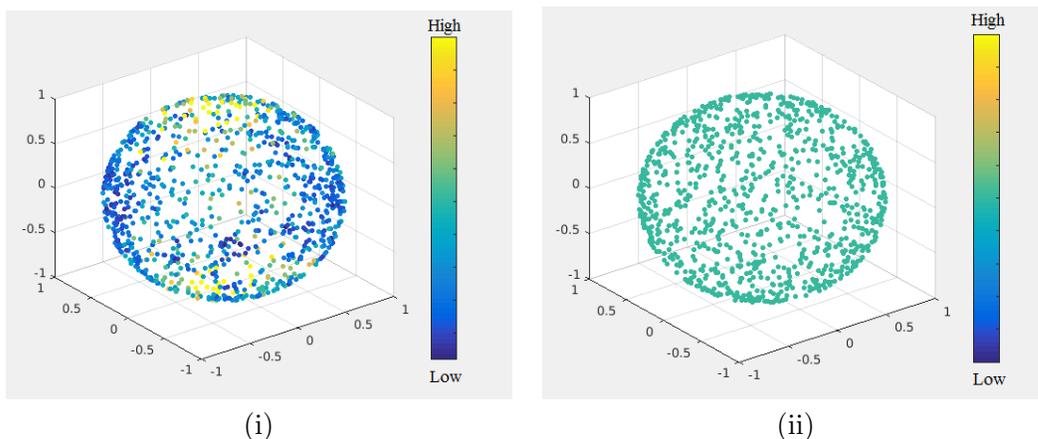


Fig. 6. Viewing angle distribution of images in the dataset EMPIAR-10059: (i) Non-uniform distribution in the raw dataset. The visualization here shows centroids of the bins that the sphere is divided into. The color of each point is assigned based on the number of points in the bin, yellow being the largest, representing the most dense bin, and blue being the smallest. (ii) Approximately uniform distribution after sampling.

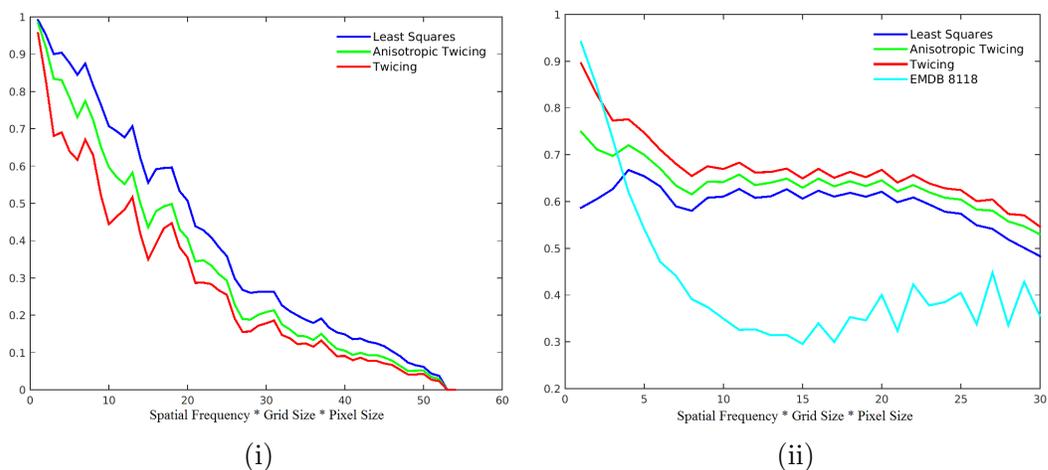


Fig. 7. (i) FCR curve for the reconstruction of the entire molecule obtained by OE using the least squares, twicing, anisotropic twicing estimators corresponding to Fig. 5(i). (ii) FCR curve for the reconstruction of the unknown subunit obtained by OE using the least squares, twicing, anisotropic twicing estimators corresponding to Fig. 5(i). We also show the FCR of the masked homologous volume (EMDB-8118) to show the improvement in FCR obtained using OE.

## 7. Conclusion

The orthogonal retrieval problem in SPR is akin to the phase retrieval problem (Burvall *et al.*, 2011; Liu *et al.*, 2008; Pfeiffer *et al.*, 2006) in X-ray crystallography. In crystallography, the measured diffraction patterns contain information about the modulus of the 3D Fourier transform of the structure but the phase information is missing and needs to be obtained by other means. In crystallography, the particle's orientations are known but the phase of the Fourier coefficients is missing, while in cryo-EM, the projection images contain phase information but the orientations of the particles are missing. Kam's autocorrelation analysis for SPR leads to an orthogonal retrieval problem which is analogous to the phase retrieval problem in crystallography. The phase retrieval problem is perhaps more challenging than the orthogonal matrix retrieval problem in cryo-EM. In crystallography each Fourier coefficient is missing its phase, while in cryo-EM only a single orthogonal matrix is missing per several radial components. For each  $l$ , the unknown coefficient matrix  $\mathbf{A}_l$  is of size  $S_l \times (2l + 1)$ , corresponding to  $(2l + 1)$  radial functions. Each  $\mathbf{A}_l$  is to be obtained from  $\mathbf{C}_l$ , which is a positive semidefinite matrix of size  $S_l \times S_l$  and rank at most  $2l + 1$ . For  $S_l > 2l + 1$ , instead of estimating  $S_l(2l + 1)$  coefficients, we only need to estimate an orthogonal matrix in  $O(2l + 1)$  which allows  $l(2l + 1)$  degrees of freedom. Therefore there are  $(S_l - l)(2l + 1)$  fewer parameters to be estimated.

It is important to note that the main requirement for OE to succeed is that there are sufficiently many images to estimate the covariance matrix to the desired level of accuracy, so it has a much greater chance of success for homology modeling from very noisy images than other ab-initio methods such as those based on common lines, which fail at very high noise levels.

In this paper, we find a general magnitude correction scheme for the class of 'phase-retrieval' problems, in particular, for Orthogonal Extension in cryo-EM. The mag-

nitude correction scheme is a generalization of ‘twicing’ that is commonly used in molecular replacement. We derive an asymptotically unbiased estimator and demonstrate 3D homology modeling using OE with synthetic and experimental datasets. We foresee this method as a good way to provide models to initialize refinement, directly from experimental images without performing class averaging and orientation estimation in cryo-EM and XFEL.

While Anisotropic Twicing outperforms least squares and twicing for synthetic data, the three estimation methods have similar performance for experimental data. One possible explanation is that the underlying assumption made by all estimation methods that  $\mathbf{C}_l$  are noiseless as implied by imposing the constraint  $\mathbf{C}_l = \mathbf{A}_l \mathbf{A}_l^*$ , is violated more severely for experimental data. Specifically, the  $\mathbf{C}_l$  matrices are derived from the 2D covariance matrix of the images, and estimation errors are the result of noise in the images, finite number of images available, non-uniformity of viewing directions, and imperfect estimation of individual image noise power spectrum, contrast transfer function, and centering. These effects are likely to be more pronounced in experimental data compared to synthetic data. As a result, the error in estimating the  $\mathbf{C}_l$  matrices from experimental data is larger. The error in the estimated  $\mathbf{C}_l$  can be taken into consideration by replacing the constrained least squares problem (1) with the regularized least squares problem

$$\min_{\mathbf{A}} \|\mathbf{A} - \mathbf{B}\|_F^2 + \lambda \|\mathbf{C} - \mathbf{A} \mathbf{A}^*\|_F^2 \quad (25)$$

where  $\lambda > 0$  is a regularization parameter that would depend on the spherical harmonic order  $l$ . A comprehensive analysis of (25) and its application to experimental datasets will be the subject of future work.

## 8. Acknowledgment

We are indebted to Garib Murshudov for motivating this work with his suggestion to explore unbiased estimators. We would like to thank Xiuyuan Cheng and Zhizhen Zhao for many helpful discussions about this work. We are grateful for discussions about experimental datasets with Daniel Asarnow, Xiaochen Bai, Adam Frost, Yuan Gao, David Julius, Eugene Palovcak, Yoel Shkolnisky, and Fred Sigworth. The authors were partially supported by Award Number R01GM090200 from the NIGMS, FA9550-12-1-0317 from AFOSR, the Simons Investigator Award and the Simons Collaboration on Algorithms and Geometry, and the Moore Foundation Data-Driven Discovery Investigator Award.

## References

- Bai, X., McMullan, G. & Scheres, S. H. (2015). *Trends in Biochemical Sciences*, **40**(1), 49 – 57.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S096800041400187X>
- Barnett, A., Greengard, L., Pataki, A. & Spivak, M. (2016). *ArXiv e-prints*.
- Bhamre, T., Zhang, T. & Singer, A. (2015). In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pp. 1048–1052. IEEE.
- Bhamre, T., Zhang, T. & Singer, A. (2016). *Journal of Structural Biology*, **195**(1), 72 – 81.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S104784771630082X>
- Burvall, A., Lundström, U., Takman, P. A., Larsson, D. H. & Hertz, H. M. (2011). *Optics express*, **19**(11), 10359–10376.
- Cheng, X. (2013).  
**URL:** <http://arks.princeton.edu/ark:/88435/dsp01wh246s26t>
- Cowtan, K., (2014). Kevin cowtan’s picture book of fourier transforms. Last updated: 2014-10-23.  
**URL:** <http://www.ysbl.york.ac.uk/cowtan/fourier/coeff.html>
- Gao, Y., Cao, E., Julius, D. & Cheng, Y. (2016). *Nature*, **534**(5), 347–51.
- Goncharov, A. B. (1988). *Acta Applicandae Mathematica*, **11**(3), 199–211.  
**URL:** <http://dx.doi.org/10.1007/BF00140118>
- Grigorieff, N. (2007). *Journal of Structural Biology*, **157**(1), 117 – 125. Software tools for macromolecular microscopy.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S1047847706001699>
- Henderson, R. (2004). *Quarterly Reviews of Biophysics*, **37**, 3–13.  
**URL:** [http://journals.cambridge.org/article\\_S0033583504003920](http://journals.cambridge.org/article_S0033583504003920)
- Hosseinizadeh, A., Dashti, A., Schwander, P., Fung, R. & Ourmazd, A. (2015). In *Structural dynamics*.
- Kam, Z. (1977). *Macromolecules*, **10**(5), 927–934.  
**URL:** <http://dx.doi.org/10.1021/ma60059a009>
- Kam, Z. (1980). *Journal of Theoretical Biology*, **82**(1), 15 – 39.  
**URL:** <http://idc311-www.sciencedirect.com/science/article/pii/0022519380900880>

- Kam, Z. & Gafni, I. (1985). *Ultramicroscopy*, **17**(3), 251–262.  
**URL:** <http://europepmc.org/abstract/MED/3003997>
- Keller, J. B. (1975). *Mathematics Magazine*, **48**(4), pp. 192–197.  
**URL:** <http://www.jstor.org/stable/2690338>
- Klein, F. (1914). *Lectures on the icosahedron and the solution of equations of the fifth degree*. London. 16, 289 p.  
**URL:** [//catalog.hathitrust.org/Record/012347870](http://catalog.hathitrust.org/Record/012347870)
- Klug, A. & Crowther, R. A. (1972). *Nature*, **238**(5365), 435–440.  
**URL:** <http://dx.doi.org/10.1038/238435a0>
- Kühlbrandt, W. (2014). *Science*, **343**, 1443–1444.
- Liu, Y. J., Chen, B., Li, E. R., Wang, J. Y., Marcelli, A., Wilkins, S. W., Ming, H., Tian, Y. C., Nugent, K. A., Zhu, P. P. & Wu, Z. Y. (2008). *Phys. Rev. A*, **78**, 023817.  
**URL:** <http://link.aps.org/doi/10.1103/PhysRevA.78.023817>
- Main, P. (1979). *Acta Crystallographica Section A*, **35**(5), 779–785.  
**URL:** <http://dx.doi.org/10.1107/S0567739479001789>
- Marabini, R., Masegosa, I., San Martín, M., Marco, S., Fernandez, J., De la Fraga, L., Vaque-rizo, C. & Carazo, J. (1996). *Journal of structural biology*, **116**(1), 237–240.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *Journal of Computational Chemistry*, **25**(13), 1605–1612.
- Pfeiffer, F., Weitkamp, T., Bunk, O. & David, C. (2006). *Nat Phys*, **2**(4), 258–261.  
**URL:** <http://dx.doi.org/10.1038/nphys265>
- Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. (2017). *Nature Methods*, **advanceonlinepublication**.  
**URL:** <http://dx.doi.org/10.1038/nmeth.4169>
- Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. (1987). *EMBO J*, **6**(4), 1107–14.
- Rossmann, M. G. (2001). *Acta Crystallographica Section D*, **57**(10), 1360–1366.  
**URL:** <https://doi.org/10.1107/S0907444901009386>
- Rossmann, M. G. & Blow, D. M. (1962). *Acta Crystallographica*, **15**(1), 24–31.  
**URL:** <http://dx.doi.org/10.1107/s0365110x62000067>
- Saldin, D. K., Shneerson, V. L., Fung, R. & Ourmazd, A. (2009). *Journal of Physics: Con-densed Matter*, **21**(13), 134014.  
**URL:** <http://stacks.iop.org/0953-8984/21/i=13/a=134014>
- Salzman, D. B. (1990). *Computer Vision, Graphics, and Image Processing*, **50**, 129–156.
- Scapin, G. (2013). *Acta Crystallographica Section D*, **69**(11), 2266–2275.  
**URL:** <https://doi.org/10.1107/S0907444913011426>
- Scheres, S. H. (2012). *Journal of Structural Biology*, **180**(3), 519 – 530.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S1047847712002481>
- Shaikh, T. R., Gao, H., Baxter, W. T., A., F. J., Boisset, N., Leith, A. & Frank, J. (2008). *Nature Protocols*, **3**(12), 1941–1974.  
**URL:** <http://www.nature.com/doifinder/10.1038/nprot.2008.156>
- Singer, A., Coifman, R. R., Sigworth, F. J., Chester, D. W. & Shkolnisky, Y. (2009). *Journal of Structural Biology*.
- Singer, A. & Shkolnisky, Y. (2011). *SIAM Journal on Imaging Sciences*, **4**(2), 543–572.  
**URL:** <http://dx.doi.org/10.1137/090767777>
- Starodub, D., Aquila, A., Bajt, S., Barthelmeß, M., Barty, A., Bostedt, C., Bozek, J. D., Coppola, N., Doak, R. B., Epp, S. W., Erk, B., Foucar, L., Gumprecht, L., Hampton, C. Y., Hartmann, A., Hartmann, R., Holl, P., Kassemeyer, S., Kimmel, N., Laksmono, H., Liang, M., Loh, N. D., Lomb, L., Martin, A. V., Nass, K., Reich, C., Rolles, D., Rudek, B., Rudenko, A., Schulz, J., Shoeman, R. L., Sierra, R. G., Soltau, H., Steinbrener, J., Stellato, F., Stern, S., Weidenspointner, G., Frank, M., Ullrich, J., Strüder, L., Schlichting, I., Chapman, H. N., Spence, J. C. H. & Bogan, M. J. (2012). *Nature Comm.* **3**, 1276+.  
**URL:** <http://dx.doi.org/10.1038/ncomms2288>

- Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I. & Ludtke, S. J. (2007). *Journal of Structural Biology*, **157**, 38 – 46. Software tools for macromolecular microscopy.  
**URL:** <http://www.sciencedirect.com/science/article/pii/S1047847706001894>
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.
- Vainshtein, B. & Goncharov, A. (1986). *Soviet Physics Doklady*, **31**, 278.
- van Heel, M. (1984). *Ultramicroscopy*, **13**(1), 165 – 183.  
**URL:** <http://www.sciencedirect.com/science/article/pii/0304399184900664>
- van Heel, M. (1987). *Ultramicroscopy*, **21**(2), 111 – 123.  
**URL:** <http://www.sciencedirect.com/science/article/pii/0304399187900787>
- van Heel, M. & Frank, J. (1981). *Ultramicroscopy*, **6**(2), 187 – 194.  
**URL:** <http://www.sciencedirect.com/science/article/pii/0304399181900590>
- Zhao, Z., Shkolnisky, Y. & Singer, A. (2016). *IEEE Transactions on Computational Imaging*, **2**(1), 1–12.
- Zhao, Z. & Singer, A. (2013). *J. Opt. Soc. Am. A*, **30**(5), 871–877.  
**URL:** <http://josaa.osa.org/abstract.cfm?URI=josaa-30-5-871>

## 9. Appendix: Proof of Theorem 4.1

### 9.1. Explicit expression of $\hat{\mathbf{A}}_{LS}$

Since  $\hat{\mathbf{A}}_{LS}$  is independent of the choice of  $\mathbf{F}$  in the algorithm, we may assume that  $\mathbf{F} = \mathbf{A}$  without loss of generality. Let  $\mathbf{E} = \mathbf{A} - \mathbf{B}$ , then by assumption,  $\mathbf{E}$  is fixed, and

$$\begin{aligned} \mathbf{V}_0 \mathbf{U}_0^* &= (\mathbf{V}_0 \boldsymbol{\Sigma}_0 \mathbf{U}_0^*) (\mathbf{U}_0 \boldsymbol{\Sigma}_0^{-1} \mathbf{U}_0^*) = (\mathbf{V}_0 \boldsymbol{\Sigma}_0 \mathbf{U}_0^*) [\mathbf{U}_0 \boldsymbol{\Sigma}_0^2 \mathbf{U}_0^*]^{-0.5} \\ &= \mathbf{A}^* (\mathbf{A} - \mathbf{E}) [(\mathbf{A} - \mathbf{E})^* \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E})]^{-0.5}. \end{aligned}$$

Therefore,

$$\hat{\mathbf{A}}_{LS} = \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E}) [(\mathbf{A} - \mathbf{E})^* \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E})]^{-0.5}. \quad (26)$$

Applying (26), we may simplify  $\hat{\mathbf{A}}_{LS}$  further as follows:

$$\begin{aligned} \hat{\mathbf{A}}_{LS} &= (\mathbf{A} - \mathbf{E})^{*-1} (\mathbf{A} - \mathbf{E})^* \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E}) [(\mathbf{A} - \mathbf{E})^* \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E})]^{-0.5} \\ &= (\mathbf{A} - \mathbf{E})^{*-1} [(\mathbf{A} - \mathbf{E})^* \mathbf{A} \mathbf{A}^* (\mathbf{A} - \mathbf{E})]^{0.5}. \end{aligned} \quad (27)$$

Since  $\mathbf{A} \mathbf{A}^* = \mathbf{C}$ , we may assume that the SVD decomposition of  $\mathbf{A}$  be given by  $\mathbf{A} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^*$ , where  $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_D)$ . Let  $\mathbf{E}_0 = \mathbf{U}^* \mathbf{E} \mathbf{V}$ , then applying the derivative of matrix inversion we have

$$\begin{aligned} (\mathbf{A} - \mathbf{E})^{*-1} &= \mathbf{A}^{*-1} + \mathbf{A}^{*-1} \mathbf{E}^* \mathbf{A}^{*-1} + O(\|\mathbf{E}\|_F^2) \\ &= \mathbf{U} \boldsymbol{\Sigma}^{-1} \mathbf{V}^* + (\mathbf{U} \boldsymbol{\Sigma}^{-1} \mathbf{V}^*) \mathbf{E}^* (\mathbf{U} \boldsymbol{\Sigma}^{-1} \mathbf{V}^*) + O(\|\mathbf{E}\|_F^2) \\ &= \mathbf{U} \boldsymbol{\Sigma}^{-1} \mathbf{V}^* + \mathbf{U} \boldsymbol{\Sigma}^{-1} \mathbf{E}_0^* \boldsymbol{\Sigma}^{-1} \mathbf{V}^* + O(\|\mathbf{E}\|_F^2). \end{aligned} \quad (28)$$

We also have

$$(\mathbf{A} - \mathbf{E})^* \mathbf{A} = \mathbf{A}^* \mathbf{A} - \mathbf{E}^* \mathbf{A} = \mathbf{V} \boldsymbol{\Sigma}^2 \mathbf{V}^* - \mathbf{E}^* \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^* = \mathbf{V} [\boldsymbol{\Sigma}^2 - \mathbf{E}_0^* \boldsymbol{\Sigma}] \mathbf{V}^*$$

and similarly,  $\mathbf{A}^*(\mathbf{A} - \mathbf{E}) = \{\mathbf{V}[\boldsymbol{\Sigma}^2 - \mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^*\}^* = \mathbf{V}[\boldsymbol{\Sigma}^2 - \boldsymbol{\Sigma}\mathbf{E}_0]\mathbf{V}^*$ . Then

$$\begin{aligned} \{(\mathbf{A} - \mathbf{E})^*\mathbf{A}\mathbf{A}^*(\mathbf{A} - \mathbf{E})\}^{0.5} &= \{\mathbf{V}[\boldsymbol{\Sigma}^2 - \mathbf{E}_0^*\boldsymbol{\Sigma}][\boldsymbol{\Sigma}^2 - \boldsymbol{\Sigma}\mathbf{E}_0]\mathbf{V}^*\}^{0.5} \\ &= \mathbf{V}\{[\boldsymbol{\Sigma}^2 - \mathbf{E}_0^*\boldsymbol{\Sigma}][\boldsymbol{\Sigma}^2 - \boldsymbol{\Sigma}\mathbf{E}_0]\}^{0.5}\mathbf{V}^* \\ &= \mathbf{V}[\boldsymbol{\Sigma}^4 - \mathbf{E}_0^*\boldsymbol{\Sigma}^3 - \boldsymbol{\Sigma}^3\mathbf{E}_0 + O(\|\mathbf{E}\|_F^2)]^{0.5}\mathbf{V}^* \end{aligned} \quad (29)$$

Applying Lemma 9.1, we have that

$$[\boldsymbol{\Sigma}^4 - \mathbf{E}_0^*\boldsymbol{\Sigma}^3 - \boldsymbol{\Sigma}^3\mathbf{E}_0 + o(\|\mathbf{E}\|_F)]^{0.5} = \boldsymbol{\Sigma}^2 + \mathbf{Z} + o(\|\mathbf{E}\|_F), \quad (30)$$

where the  $ij$ -th entry of  $\mathbf{Z}$  is given by

$$\mathbf{Z}_{ij} = -\frac{\mathbf{E}_{0,ji}^*\sigma_j^3 + \mathbf{E}_{0,ij}\sigma_i^3}{\sigma_i^2 + \sigma_j^2}.$$

Combining (27)-(30), we have

$$\begin{aligned} \hat{\mathbf{A}}_{\text{LS}} &= [\mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{V}^* + \mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}^{-1}\mathbf{V}^*]\mathbf{V}[\boldsymbol{\Sigma}^2 + \mathbf{Z}]\mathbf{V}^* + o(\|\mathbf{E}\|_F) \\ &= \mathbf{A} + [\mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{V}^*]\mathbf{V}\mathbf{Z}\mathbf{V}^* + \mathbf{U}\boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}^{-1}\mathbf{V}^*\mathbf{V}\boldsymbol{\Sigma}^2\mathbf{V}^* + o(\|\mathbf{E}\|_F) \\ &= \mathbf{A} + \mathbf{U}[\boldsymbol{\Sigma}^{-1}\mathbf{Z} + \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^* + o(\|\mathbf{E}\|_F), \end{aligned} \quad (31)$$

and the  $ij$ -th entry of  $[\boldsymbol{\Sigma}^{-1}\mathbf{Z} + \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]$  can be explicitly written down by

$$\begin{aligned} \sigma_i^{-1}\mathbf{Z}_{ij} + \sigma_i^{-1}\mathbf{E}_{0,ji}^*\sigma_j &= -\frac{\mathbf{E}_{0,ji}^*\sigma_j^3 + \mathbf{E}_{0,ij}\sigma_i^3}{\sigma_i(\sigma_i^2 + \sigma_j^2)} + \mathbf{E}_{0,ji}^*\frac{\sigma_j}{\sigma_i} \\ &= \mathbf{E}_{0,ji}^*\frac{\sigma_i\sigma_j}{\sigma_i^2 + \sigma_j^2} - \mathbf{E}_{0,ij}\frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2}. \end{aligned}$$

### 9.2. Expectation when $\mathbf{V}$ is uniformly distributed

From the analysis in the previous section, we have

$$\mathbb{E}(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{A}) = \mathbf{U} \mathbb{E}_{\mathbf{V}}[\boldsymbol{\Sigma}^{-1}\mathbf{Z} + \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^* + o(\|\mathbf{E}\|_F),$$

when  $\mathbf{V}$  is uniformly distributed on the set of all orthogonal matrices or all unitary matrices.

Now let us check the  $ik$ -th entry of  $\mathbb{E}_{\mathbf{V}}[\boldsymbol{\Sigma}^{-1}\mathbf{Z} + \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^*$ , which is

$$\begin{aligned} & \sum_j \left[ \mathbf{E}_{0,ji}^* \frac{\sigma_i \sigma_j}{\sigma_i^2 + \sigma_j^2} - \mathbf{E}_{0,ij} \frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2} \right] \mathbf{V}_{kj}^* \\ &= \sum_j \left[ \left( \sum_{m,n} \mathbf{U}_{mj}^* \mathbf{E}_{mn} \mathbf{V}_{ni} \right)^* \frac{\sigma_i \sigma_j}{\sigma_i^2 + \sigma_j^2} - \sum_{m,n} \mathbf{U}_{mi}^* \mathbf{E}_{mn} \mathbf{V}_{nj} \frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2} \right] \mathbf{V}_{kj}^*. \end{aligned}$$

Applying the facts that for real-valued matrices  $\mathbf{V}$  we have

$$\mathbb{E}_{\mathbf{V}} \mathbf{V}_{ij} \mathbf{V}_{mn} = \begin{cases} \frac{1}{D}, & \text{if } i = m \text{ and } j = n \\ 0, & \text{otherwise,} \end{cases}$$

and for complex-valued matrices  $\mathbf{V}$ ,  $\mathbb{E}_{\mathbf{V}} \mathbf{V}_{ij} \mathbf{V}_{mn} = 0$  for all  $(i, j, m, n)$ , and

$$\mathbb{E}_{\mathbf{V}} \mathbf{V}_{ij} \mathbf{V}_{mn}^* = \begin{cases} \frac{1}{D}, & \text{if } i = m \text{ and } j = n \\ 0, & \text{otherwise,} \end{cases}$$

its expectation is given by

$$\begin{aligned} & \frac{1}{D} \left\{ \sum_m (\mathbf{U}_{mi}^* \mathbf{E}_{mk})^* \frac{1}{2} - \sum_{m,j} \mathbf{U}_{mi}^* \mathbf{E}_{mk} \frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2} \right\} \\ &= \begin{cases} \frac{1}{D} \left\{ \frac{1}{2} [\mathbf{U}^* \mathbf{E}]_{ik} - [\mathbf{U}^* \mathbf{E}]_{ik} \sum_j \frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2} \right\}, & \text{when } \mathbf{V} \text{ is real-valued,} \\ -\frac{1}{D} [\mathbf{U}^* \mathbf{E}]_{ik} \sum_j \frac{\sigma_i^2}{\sigma_i^2 + \sigma_j^2}, & \text{when } \mathbf{V} \text{ is complex-valued.} \end{cases} \end{aligned}$$

Combining these elementwise expectations into a matrix,  $\mathbb{E}_{\mathbf{V}}[\boldsymbol{\Sigma}^{-1}\mathbf{Z} - \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^* = -\mathbf{T}\mathbf{U}^*\mathbf{E}$ . Therefore, we have

$$\mathbb{E}(\hat{\mathbf{A}}_{\text{LS}} - \mathbf{A}) = \mathbf{U} \mathbb{E}_{\mathbf{V}}[\boldsymbol{\Sigma}^{-1}\mathbf{Z} - \boldsymbol{\Sigma}^{-1}\mathbf{E}_0^*\boldsymbol{\Sigma}]\mathbf{V}^* = -\mathbf{U}\mathbf{T}\mathbf{U}^*\mathbf{E} + o(\|\mathbf{E}\|_F). \quad (32)$$

### 9.3. Lemmas

**Lemma 9.1.** *For a diagonal matrix  $\mathbf{X} = \text{diag}(x_1, x_2, \dots, x_D)$ , the  $ij$ -th entry of  $(\mathbf{X} + \mathbf{E})^{0.5} - \mathbf{X}^{0.5}$  is given by*

$$[(\mathbf{X} + \mathbf{E})^{0.5} - \mathbf{X}^{0.5}]_{ij} = [\mathbf{E}_{ij} \cdot \frac{1}{x_i^{0.5} + x_j^{0.5}}] + o(\|\mathbf{E}\|_F).$$

*Proof.* The proof is based on the following observation: if  $\mathbf{Y}$  is diagonal and  $\mathbf{C}$  is small,

$$(\mathbf{Y} + \mathbf{C})^2 - \mathbf{Y}^2 = \mathbf{Y}\mathbf{C} + \mathbf{C}\mathbf{Y} + \mathbf{C}^2 = [\mathbf{C}_{ij} \cdot (y_i + y_j)] + o(\|\mathbf{C}\|_F),$$

where  $[C_{ij} \cdot (y_i + y_j)]$  denotes a matrix of  $D \times D$ , with  $ij$ -th entry given by  $C_{ij} \cdot (y_i + y_j)$ .

Then the lemma is proved by applying this observation to  $\mathbf{Y} = \mathbf{X}^{0.5}$  and  $(\mathbf{Y} + \mathbf{C})^2 = \mathbf{X} + \mathbf{E}$ :

$$\mathbf{E} = [[(\mathbf{X} + \mathbf{E})^{0.5} - \mathbf{X}^{0.5}]_{ij} \cdot (x_i^{0.5} + x_j^{0.5})] + o(\|(\mathbf{X} + \mathbf{E})^{0.5} - \mathbf{X}^{0.5}\|_F).$$

□