

SO(3)-INVARIANT PCA WITH APPLICATION TO MOLECULAR DATA

Michael Fraiman^{*} Paulina Hoyos[†] Tamir Bendory[¶] Joe Kileel[‡]
Oscar Mickelin[‡] Nir Sharon^{*} Amit Singer[§]

^{*} School of Mathematical Sciences, Tel Aviv University, Israel

[†] Department of Mathematics, UT Austin, USA

[¶] School of Electrical and Computer Engineering, Tel Aviv University, Israel

[‡] Yau Mathematical Sciences Center, Tsinghua University, China

[§] Program in Applied and Computational Mathematics
and Department of Mathematics, Princeton University, USA

ABSTRACT

Principal component analysis (PCA) is a fundamental technique for dimensionality reduction and denoising; however, its application to three-dimensional data with arbitrary orientations—common in structural biology—presents significant challenges. A naive approach requires augmenting the dataset with many rotated copies of each sample, incurring prohibitive computational costs. In this paper, we extend PCA to 3D volumetric datasets with unknown orientations by developing an efficient and principled framework for SO(3)-invariant PCA that implicitly accounts for all rotations without explicit data augmentation. By exploiting underlying algebraic structure, we demonstrate that the computation involves only the square root of the total number of covariance entries, resulting in a substantial reduction in complexity. We validate the method on real-world molecular datasets, demonstrating its effectiveness and opening up new possibilities for large-scale, high-dimensional reconstruction problems.

Index Terms— steerable PCA, group invariants, 3D volumes, cryo-EM, spherical Bessel, ball harmonics

1. INTRODUCTION

Principal component analysis (PCA) is a fundamental technique in data science and statistics, especially when dealing with high-dimensional datasets. By extracting the leading eigenvectors of the covariance matrix, PCA identifies directions of maximum variability and projects the data onto a lower-dimensional space. This simple and interpretable method is widely used for denoising, dimensionality reduction, visualization, feature extraction, and compression, while often preserving the essential structure of the data [1, 2].

In some scenarios, the orientation of the data samples is arbitrary or random. A notable example, and the focus of this

paper, arises when the samples are three-dimensional functions, each observed up to an unknown rotation. As in [3, 4], rather than using the costly naive fix—augmenting each sample with many rotated copies to handle random orientations—we exploit the underlying algebraic structure and group symmetries. This results in a numerically accurate method that implicitly accounts for all rotations, while greatly reducing computation. In particular, we prove that, under a suitable representation, the SO(3)-invariant covariance matrix is low-dimensional and its eigenvectors can be computed efficiently. The problem is formally stated in Section 2, and the derivation of the SO(3)-invariant PCA is presented in Section 3.

The primary motivation of this paper arises from molecular datasets in structural biology applications, including single-particle cryo-electron microscopy and cryo-electron tomography [5, 6, 7]. These problems are computationally demanding, involving the reconstruction of 3D structures from massive, high-dimensional, noisy datasets [8]. Although PCA has been highly effective in various stages of cryo-EM image analysis, its use has largely been limited to 2D raw images, leaving a gap in methodologies for processing 3D volumes [3, 9, 10]. We extend PCA to 3D molecular data, enabling a principled approach for dimensionality reduction and denoising in volumetric reconstructions. Section 4 demonstrates the effectiveness of our method on available molecular datasets, and Section 5 outlines potential future applications.

2. PROBLEM FORMULATION

Let $\phi: \mathbb{R}^3 \rightarrow \mathbb{C}$ be a function supported on the unit ball (e.g., the electro-static potential of a molecule, or its Fourier transform). We assume the function is well-approximated by the following finite-term expansion in spherical coordinates:

$$\phi(r, \theta, \varphi) = \sum_{\ell=0}^L \sum_{m=-\ell}^{\ell} \sum_{s=1}^{S(\ell)} f_{\ell m s} j_{\ell s}(r) Y_{\ell}^m(\theta, \varphi), \quad (2.1)$$

M.F. and P.H. contributed equally to this work.

where $f_{\ell ms} \in \mathbb{C}$ are expansion coefficients. The basis functions $j_{\ell s}(r)Y_{\ell}^m(\theta, \varphi)$ are called *ball harmonics* [11]. Here $j_{\ell s}$ is the normalized spherical Bessel functions given by $j_{\ell s}(r) := \frac{4}{|j_{\ell+1}(u_{\ell s})|} j_{\ell}(u_{\ell s}r)$, where j_{ℓ} is the ℓ -th spherical Bessel function of the first kind, and $u_{\ell s}$ is the s -th positive zero of j_{ℓ} . The complex spherical harmonics Y_{ℓ}^m are given by $Y_{\ell}^m(\theta, \varphi) := \sqrt{\frac{2\ell+1}{4\pi} \cdot \frac{(\ell-m)!}{(\ell+m)!}} P_{\ell}^m(\cos \theta) e^{im\varphi}$, where P_{ℓ}^m are the associated Legendre polynomials with the Condon–Shortley phase. Here $S(\ell)$ denotes the number of radial components retained at the angular degree ℓ , which is determined by the Nyquist sampling rate of the discretized volume and decreases with ℓ . Under Nyquist sampling on an $N \times N \times N$ grid, we have $S = \Theta(N)$ and $L = \mathcal{O}(N)$. Any square-integrable function on the unit ball can be represented in the form (2.1) as $L \rightarrow \infty$, and the finite expansion is common and useful to represent smooth, 3D volumes. See [11] for a fast algorithm to compute the expansion numerically.

A 3D rotation $R \in \text{SO}(3)$ acts on the function $\phi: \mathbb{R}^3 \rightarrow \mathbb{C}$ via $(R \cdot \phi)(x) := \phi(R^T x)$ for $x \in \mathbb{R}^3$. In spherical coordinates, the action is given by

$$\begin{aligned} (R \cdot \phi)(r, \theta, \varphi) \\ = \sum_{\ell=0}^L \sum_{m=-\ell}^{\ell} \sum_{s=1}^{S(\ell)} \left(\sum_{k=-\ell}^{\ell} D_{mk}^{\ell}(R) f_{\ell ks} \right) j_{\ell s}(r) Y_{\ell}^m(\theta, \varphi), \end{aligned}$$

where $D_{mk}^{\ell}(R)$ are the Wigner D-matrices. This amounts to transforming the expansion coefficients of ϕ via

$$R \cdot f_{\ell ms} = \sum_{k=-\ell}^{\ell} D_{mk}^{\ell}(R) f_{\ell ks}. \quad (2.2)$$

We assume access to n volumes of the form (2.1), each subjected to an arbitrary 3D rotation as described above and potentially corrupted by noise. Each volume resides in a high-dimensional space $\mathbb{R}^D$, where $D = \sum_{\ell=0}^L (2\ell+1)S(\ell)$. The goal of PCA is to identify a lower-dimensional subspace isomorphic to $\mathbb{R}^d$ with $d \ll D$, such that the volumes can be well-approximated by their projections onto this subspace. To this end, we compute the $\text{SO}(3)$ -invariant covariance matrix and extract its leading eigenvectors, as explained below.

3. $\text{SO}(3)$ -INVARIANT PRINCIPAL COMPONENTS

Suppose we collect n volumes $\{\phi^{(i)}\}_{i=1}^n$ of the form (2.1), and consider all of their possible 3D rotations. The sample mean of this $\text{SO}(3)$ -augmented collection is given by

$$\begin{aligned} \phi_{\text{mean}}(r, \theta, \varphi) &= \frac{1}{n} \sum_{i=1}^n \int_{\text{SO}(3)} (R \cdot \phi^{(i)})(r, \theta, \varphi) dR \\ &= \sum_{s=1}^{S(\ell)} \left(\frac{1}{n} \sum_{i=1}^n f_{00s}^{(i)} \right) j_{0s}(r) Y_0^0(\theta, \varphi), \end{aligned}$$

where the last equality follows from Equation (2.2) and orthogonality relations of the Wigner D-matrices. Since $Y_0^0(\theta, \varphi)$ is constant, ϕ_{mean} is a radially symmetric function.

The full $\text{SO}(3)$ -invariant covariance is defined as

$$\mathcal{C} = \frac{1}{n} \sum_{i=1}^n \int_{\text{SO}(3)} (R \cdot \phi^{(i)} - \phi_{\text{mean}}) \overline{(R \cdot \phi^{(i)} - \phi_{\text{mean}})} dR.$$

To express $\mathcal{C}$ in the ball harmonics basis using the expansion coefficients $f_{\ell ms}^{(i)}$, we first note that centering the volumes amounts to updating the zero-frequency coefficients $f_{00s}^{(i)}$ to $f_{00s}^{(i)} - \frac{1}{n} \sum_{i=1}^n f_{00s}^{(i)}$. We then compute

$$\begin{aligned} \mathcal{C}_{(\ell, m, s), (\ell', m', s')} &= \frac{1}{n} \sum_{i=1}^n \int_{\text{SO}(3)} (R \cdot f_{\ell ms}^{(i)}) \overline{(R \cdot f_{\ell' m' s'}^{(i)})} dR \\ &= \frac{1}{n} \sum_{i=1}^n \sum_{k=-\ell}^{\ell} \sum_{k'=-\ell'}^{\ell'} f_{\ell ks}^{(i)} \overline{f_{\ell' k' s'}^{(i)}} \int_{\text{SO}(3)} D_{mk}^{\ell}(R) \overline{D_{m' k'}^{\ell'}(R)} dR \\ &= \delta_{mm'} \cdot \delta_{\ell\ell'} \cdot \frac{1}{n} \cdot \frac{1}{2\ell+1} \sum_{i=1}^n \sum_{k=-\ell}^{\ell} f_{\ell ks}^{(i)} \overline{f_{\ell ks'}^{(i)}}, \end{aligned} \quad (3.1)$$

where the last equality follows from orthogonality relations of the Wigner D-matrices. Thus, if we set $C_{\ell} \in \mathbb{C}^{S(\ell) \times S(\ell)}$ as

$$C_{\ell}(s, s') := \frac{1}{n} \cdot \frac{1}{2\ell+1} \sum_{i=1}^n \sum_{m=-\ell}^{\ell} f_{\ell ms}^{(i)} \overline{f_{\ell ms'}^{(i)}}, \quad (3.2)$$

then

$$\mathcal{C} = \bigoplus_{\ell=0}^L (I_{2\ell+1} \otimes C_{\ell}). \quad (3.3)$$

In other words, the matrix representation of $\mathcal{C}$ in (3.1) with respect to the ball harmonics basis is block-diagonal, with each block $C_{\ell} \in \mathbb{C}^{S(\ell) \times S(\ell)}$ occurring $(2\ell+1)$ times. This reduction is possible due to the averaging over $\text{SO}(3)$. We note that formulas (3.2) and (3.3) extend results of [3] from 2D to 3D.

Let $\{(v_{\ell s}, \lambda_{\ell s})\}_{s=1}^{S(\ell)}$ be the eigenvector/eigenvalue pairs of the matrix C_{ℓ} in (3.2). Each $v_{\ell s}$ gives rise to a set $V_{\ell s} := \{u_{\ell sm}\}_{m=-\ell}^{\ell}$, where $u_{\ell sm} := e_m \otimes v_{\ell s}$ and e_m is the m -th standard basis vector in $\mathbb{R}^{2\ell+1}$. Each element of $V_{\ell s}$ is an eigenvector of $I_{2\ell+1} \otimes C_{\ell}$ with the eigenvalue $\lambda_{\ell s}$. By (3.3), all such vectors, padded with zeros, are the eigenvectors of $\mathcal{C}$, and they constitute a basis of $\mathbb{R}^D$. This lets us obtain the eigenvectors of $\mathcal{C}$ from those of $\{C_{\ell}\}_{\ell=0}^L$ at no extra cost.

We order the eigenvectors $\{u_{\ell_j s_j m_j}\}_{j=1}^D$ as follows. First, we order the sets $\{V_{\ell_j s_j}\}_{j=1}^{D'}$, where $D' = \sum_{\ell=0}^L S(\ell)$, so the corresponding eigenvalues decrease. Then, inside each $V_{\ell_j s_j}$ we order the vectors by m . The eigenvolumes $\mathbf{V} = \{\psi_j\}_{j=1}^D$ are given by $\psi_j(r, \theta, \varphi) = \sum_{t=1}^{S(\ell_j)} (v_{\ell_j s_j})_t j_{\ell_j t}(r) Y_{\ell_j}^{m_j}(\theta, \varphi)$, where $(\cdot)_t$ is the t -th entry of a vector. Each volume ϕ decomposes as $\phi = \sum_{j=1}^D \alpha_j \psi_j$ with $\alpha_j := \sum_{t=1}^{S(\ell_j)} f_{\ell_j m_j t} \overline{(v_{\ell_j s_j})_t}$.

Finally, the $\text{SO}(3)$ -invariant PCA approximation of ϕ is $\sum_{j=1}^d \alpha_j \psi_j$ for $d \ll D$, with d chosen by a rank criterion.

Computational Complexity. Suppose each volume in the dataset $\{\phi^{(i)}\}_{i=1}^n$ is discretized on a cubic grid of side length N , so that each volume contains N^3 voxels. Recall $S := \max_{\ell} S(\ell)$. Computation of the ball harmonics expansion coefficients takes $\mathcal{O}(N^3 \log^2(N))$ operations [11]. Computing the ℓ -th block C_{ℓ} of the covariance matrix requires $\mathcal{O}(S^2 \ell)$ operations; summing over ℓ , there are $\mathcal{O}(S^2 L^2)$ operations. The eigendecomposition takes $\mathcal{O}(S^3)$ operations for each block. Overall, the complexity of computing the (full) $\text{SO}(3)$ -invariant PCA basis is $\mathcal{O}(S^2 L^2 + S^3 L)$. For comparison, a computation of the usual dense covariance of $\{\phi^{(i)}\}_{i=1}^n$ in the ball harmonics basis would require $\mathcal{O}(S^2 L^4)$ operations, and its eigendecomposition would cost $\mathcal{O}(S^3 L^6)$. Under Nyquist sampling, our algorithm requires $\mathcal{O}(N^4)$ operations in total compared to $\mathcal{O}(N^9)$ for the usual computation. Thus, our formulation of the $\text{SO}(3)$ -invariant covariance simultaneously achieves two objectives: it implicitly incorporates all rotated versions of the data, and it reduces the computational complexity of PCA by a factor of about N^5 .

O(3)-Invariant PCA. To construct an $\text{O}(3)$ -invariant covariance, we could augment the dataset with reflections of the volumes in the xy -plane, and then perform the $\text{SO}(3)$ -invariant procedure. This reflection transforms the coefficients by replacing $f_{\ell m s}$ with $(-1)^{\ell+m} f_{\ell m s}$. It turns out that the $\text{O}(3)$ -invariant covariance matrix is identical to (3.3).

4. NUMERICAL RESULTS

We test our framework on a dataset of 1,419 volumes from the Protein Data Bank [12] like in [13]. Each volume is expanded in ball harmonics, by the implementation of [11] with a bandlimit $L = 20$ and resolutions $N = 64, 128$, and 256 (i.e., each volume consists of N^3 voxels). Our code is at https://github.com/MichaelFraiman/PCA_SO-3.

First we compute a collection of eigenvolumes for varying parameters for volumes of size $N = 64$. See Figure 1 for the results. Figure 2 shows two molecular structures and their approximation based on the d leading principal volumes.

Comparison Against a Fixed Basis. Next, we aim to show that the data-driven PCA approximation outperforms the fixed ball harmonics basis when it is truncated. We quantify the accuracy of an approximation of a structure ϕ by

$$w_{\phi}^V(d) := \frac{\sum_{j=0}^d |\alpha_j|^2}{\sum_{j=0}^D |\alpha_j|^2}, \quad (4.1)$$

where $\{\alpha_j\}_{j=1}^D$ are the expansion coefficients of ϕ in an orthonormal basis V . Figure 3 compares $w_{\phi}^V(d)$ for a representative volume contained in the dataset under three choices of basis V : the PCA basis $\mathbf{V}$, the ball harmonics basis with coefficients ordered by decreasing absolute value (this ordering is specific to each volume), and the ball harmonics basis

with coefficients ordered by increasing u_{ls} (the s -th positive root of j_l) with near-zero coefficients filtered out. Two conclusions are evident. First, adaptive PCA offers a clear advantage over a fixed basis. Second, for volumes sampled on an $N \times N \times N$ grid, the accuracy of the PCA approximation remains nearly unchanged as N increases. In contrast, that of the ball harmonics basis declines sharply, even when the expansion coefficients are reordered by decreasing magnitude.

Synthesizing Proteins. PCA can be thought of as giving a generative model. Generative models are useful for data augmentation in modern machine learning because they can create new, realistic examples that mimic the true data distribution. This may be especially valuable for structural biology, where data is expensive to collect. With PCA, approximations of the molecular structures lie in a linear space $\mathbf{L} \cong \mathbb{R}^d \subseteq \mathbb{R}^D$. We suggest synthesizing other elements of $\mathbf{L}$ (that is, “fake proteins”) in the following way. Let Φ be the dataset of proteins and let $\{\psi_j\}_{j=1}^d$ be its low-rank basis obtained via the $\text{SO}(3)$ -invariant PCA procedure. For each $\phi^{(i)} \in \Phi$, we consider its expansion $\phi^{(i)} = \sum_{j=1}^d \alpha_j^{(i)} \psi_j$. The mean μ_j and the variance σ_j are estimated so that the samples $\{\alpha_j^{(i)}\}_{\phi^{(i)} \in \Phi}$ are modeled as $\mathcal{N}(\mu_j, \sigma_j^2)$. Now, a random element of $\mathbf{L}$ can be sampled as $\sum_{j=1}^d \beta_j \psi_j$, where $\beta_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$ is a random sequence of coefficients. We present a few examples of synthesized proteins in Figure 4, compared to some real ones.

5. PATHWAYS TO APPLICATIONS IN CRYO-EM

This paper derived an efficient method for performing $\text{SO}(3)$ -invariant PCA on 3D data samples with random orientations. To conclude, we outline potential applications of the framework, with emphasis on 3D reconstruction in cryo-EM.

The main computational workhorse in cryo-EM is the expectation-maximization (EM) algorithm [14]. Although it is successful in recovering a wide range of molecular structures, there is strong evidence that its performance deteriorates as the noise level increases [15]. One possible remedy is to apply EM directly on a subspace where the molecular structure is expected to lie, a strategy known as subspace EM [16]. Using the computational framework developed in this work, this subspace can be learned from molecular databases and incorporated into existing EM pipelines. We note that implementing this will present challenges, e.g., due to CTF effects [17].

As an alternative to EM, the method of moments (MoM) has been proposed for cryo-EM, particularly for scenarios with extremely high noise. This approach is auspicious for small molecular structures, which typically induce high noise levels and remain beyond the reach of current cryo-EM technology [18]. Recent work has shown that MoM can be improved when the structure is constrained to a subspace [19]. The subspace could be learned using the framework developed in this paper, after which moments would be expressed in terms of expansion coefficients in the resulting basis.

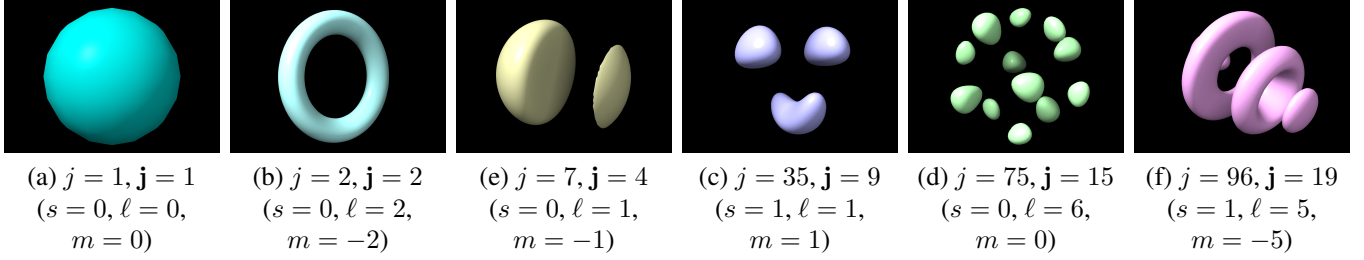


Fig. 1. Eigenvolumes for $N = 64$. The notation is retained from Section 3: j stands for the rank of the eigenvolume ψ_j in the ordering of $u_{\ell_j s_j m_j}$, while $\mathbf{j}$ stands for the rank of the eigenvalue $\lambda_{\ell_j s_j}$.

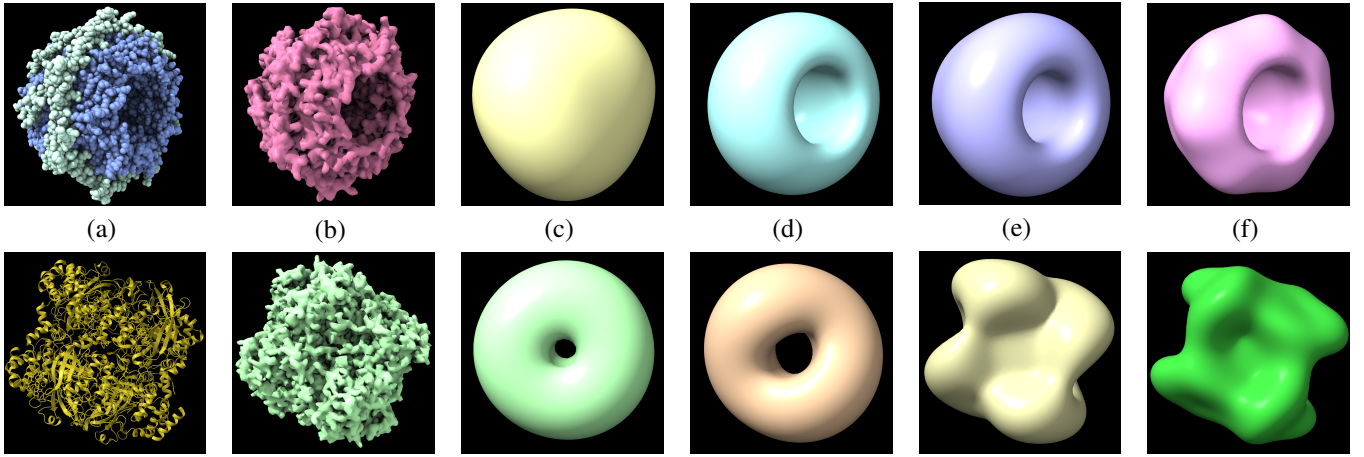


Fig. 2. Volume reconstructions with $N = 128$. In each row: (a) reference volume; (b) ball harmonics expansion (bandlimit $L = 20$); (c–f) reconstructions using the top d eigenvectors with $d = 10, 20, 100, 200$, respectively. Top row: volume with PDB index 1avo; bottom row with 1dgb.

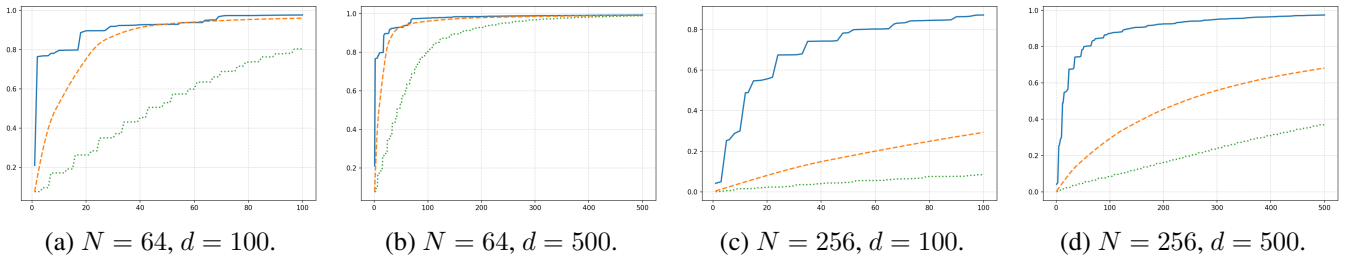


Fig. 3. Comparison of approximations with $\text{SO}(3)$ -invariant PCA and the ball harmonics basis (BH) for the sample volume 1fzf. The plots show $w_\phi^V(k)$ for $k = 1, \dots, d$ under three choices of the basis V : PCA (solid blue), sorted BH (dashed orange), BH sorted by u_{ls} (dotted green).

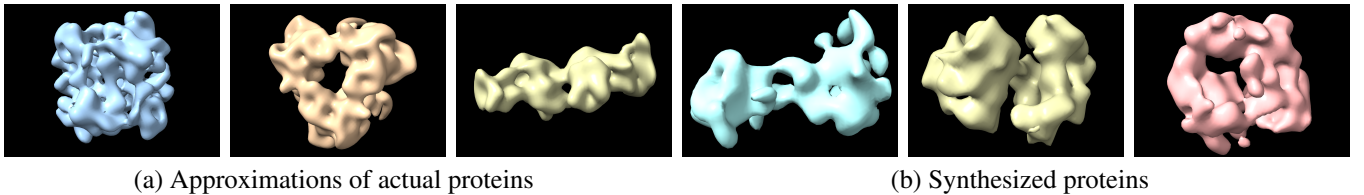


Fig. 4. Examples of synthesis of random proteins using $d = 200$ eigenvectors with resolution $N = 64$ and bandlimit $L = 20$. The PDB indexes of the real proteins are 1dgb, 1cb5, 1fzf, respectively.

Acknowledgments

T.B. and N.S. are supported in part by the NSF-BSF under Grant 2024791. T.B. is also supported in part by BSF under Grant 2020159, in part by ISF under Grant 1924/21, and in part by a grant from The Center for AI and Data Science at Tel Aviv University. N.S. is also partially supported by the BSF award 2024266 and the DFG award 514588180. J.K. is supported in part by NSF DMS 2309782, NSF DMS 2436499, NSF CISE-IIS 2312746, DE SC0025312, and the Sloan Foundation. A.S. is supported in part by AFOSR FA9550-23-1-0249, the Simons Foundation Math+X Investigator Award, NSF DMS 2510039, and NIH/NIGMS R01GM13678001.

6. REFERENCES

- [1] Ian T Jolliffe and Jorge Cadima, “Principal component analysis: A review and recent developments,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp. 20150202, 2016.
- [2] Ian Jolliffe, “Principal component analysis,” in *International Encyclopedia of Statistical Science*, pp. 1094–1096. Springer, 2011.
- [3] Zhizhen Zhao, Yoel Shkolnisky, and Amit Singer, “Fast steerable principal component analysis,” *IEEE Transactions on Computational Imaging*, vol. 2, no. 1, pp. 1–12, 2016.
- [4] Eitan Rosen, Paulina Hoyos, Xiuyuan Cheng, Joe Kileel, and Yoel Shkolnisky, “The G -invariant graph Laplacian part I: Convergence rate and eigendecomposition,” *Applied and Computational Harmonic Analysis*, vol. 71, pp. 101637, February 2024.
- [5] Eva Nogales, “The development of cryo-EM into a mainstream structural biology technique,” *Nature Methods*, vol. 13, no. 1, pp. 24–27, 2016.
- [6] Yifan Cheng, “Single-particle cryo-EM—How did it get here and where will it go,” *Science*, vol. 361, no. 6405, pp. 876–880, 2018.
- [7] Abigail JI Watson and Alberto Bartesaghi, “Advances in cryo-ET data processing: Meeting the demands of visual proteomics,” *Current Opinion in Structural Biology*, vol. 87, pp. 102861, 2024.
- [8] Tamir Bendory, Alberto Bartesaghi, and Amit Singer, “Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities,” *IEEE Signal Processing Magazine*, vol. 37, no. 2, pp. 58–76, 2020.
- [9] Tejal Bhamre, Teng Zhang, and Amit Singer, “Denoising and covariance estimation of single particle cryo-EM images,” *Journal of Structural Biology*, vol. 195, no. 1, pp. 72–81, 2016.
- [10] Marin Van Heel and Joachim Frank, “Use of multivariate statistics in analysing the images of biological macromolecules,” *Ultramicroscopy*, vol. 6, no. 1, pp. 187–194, 1981.
- [11] Joe Kileel, Nicholas F Marshall, Oscar Mickelin, and Amit Singer, “Fast expansion into harmonics on the ball,” *SIAM Journal on Scientific Computing*, vol. 47, no. 2, pp. A1117–A1144, 2025.
- [12] EMBL-EBI/wwPDB, “EMDB archive,” <https://ftp.ebi.ac.uk/pub/databases/emdb/structures/>, Subset of 1,419 entries used; accessed 8 June 2025.
- [13] Andy Zhang, Oscar Mickelin, Joe Kileel, Eric J. Verbeke, Nicholas F. Marshall, Marc Aurèle Gilles, and Amit Singer, “Moment-based metrics for molecules computable from cryogenic electron microscopy images,” *Biological Imaging*, vol. 4, pp. e3, 2024.
- [14] Sjors HW Scheres, “RELION: Implementation of a Bayesian approach to cryo-EM structure determination,” *Journal of Structural Biology*, vol. 180, no. 3, pp. 519–530, 2012.
- [15] Amnon Balanov, Wasim Huleihel, and Tamir Bendory, “Expectation-maximization for multi-reference alignment: Two pitfalls and one remedy,” *arXiv preprint arXiv:2505.21435*, 2025.
- [16] Nicha C Dvornek, Fred J Sigworth, and Hemant D Tagare, “SubspaceEM: A fast maximum-a-posteriori algorithm for cryo-EM single particle reconstruction,” *Journal of Structural Biology*, vol. 190, no. 2, pp. 200–214, 2015.
- [17] Nicholas F Marshall, Oscar Mickelin, Yunpeng Shi, and Amit Singer, “Fast principal component analysis for cryo-electron microscopy images,” *Biological Imaging*, vol. 3, pp. e2, 2023.
- [18] Tamir Bendory, Nicolas Boumal, William Leeb, Eitan Levin, and Amit Singer, “Toward single particle reconstruction without particle picking: Breaking the detection limit,” *SIAM Journal on Imaging Sciences*, vol. 16, no. 2, pp. 886–910, 2023.
- [19] Jeremy Hoskins, Yuehaw Khoo, Oscar Mickelin, Amit Singer, and Yuguan Wang, “Subspace method of moments for ab initio 3-D single-particle cryo-EM reconstruction,” *arXiv preprint arXiv:2410.06889*, 2024.