AUTOCORRELATION ANALYSIS FOR CRYO-EM WITH SPARSITY CONSTRAINTS: IMPROVED SAMPLE COMPLEXITY AND PROJECTION-BASED ALGORITHMS

TAMIR BENDORY, YUEHAW KHOO, JOE KILEEL, OSCAR MICKELIN, AND AMIT SINGER

ABSTRACT. The number of noisy images required for molecular reconstruction in single-particle cryo-electron microscopy (cryo-EM) is governed by the autocorrelations of the observed, randomly-oriented, noisy projection images. In this work, we consider the effect of imposing sparsity priors on the molecule. We use techniques from signal processing, optimization, and applied algebraic geometry to obtain new theoretical and computational contributions for this challenging non-linear inverse problem with sparsity constraints. We prove that molecular structures modeled as sums of idealized point masses are uniquely determined by the second-order autocorrelation of their projection images, implying that the sample complexity is proportional to the square of the variance of the noise. This theory improves upon the non-sparse case, where the third-order autocorrelation is required for uniformly-oriented particle images and the sample complexity scales with the cube of the noise variance. Furthermore, we build a computational framework to reconstruct molecular structures which are sparse in the wavelet basis. This method combines the sparse representation for the molecule with projection-based techniques used for phase retrieval in X-ray crystallography.

1. INTRODUCTION

Sparsity is a ubiquitous prior in many linear inverse problems, including regression [85, 41], compressed sensing [25, 17, 28], and various image processing applications [27], to name a few. While sparse priors are also used for non-linear inverse problems, their applicability and theoretical foundations are limited to a few specific (usually linear and quadratic) models, e.g., [15, 21, 31, 88]. Motivated by single-particle cryo-EM—an imaging technology for determining the 3-D structure of biological molecules—this paper uses modern techniques from signal processing, optimization, and applied algebraic geometry to provide new theoretical analysis and computational methods for a challenging non-linear inverse problem with sparsity constraints.

Cryo-EM has garnered increasing interest in the past decade due to a series of technological and algorithmic breakthroughs, driving a striking improvement in
the obtainable resolution, up to the level where individual atoms can be distin-
guished. This has in turn opened new scientific horizons and led to many biological
discoveries, e.g., [55, 3, 16].

In a cryo-EM experiment, a solution containing molecules to be imaged is rapidly
frozen into a thin ice layer, which is then placed in an electron microscope. Next,
the microscope acquires an image, called micrograph, which contains multiple 2-D
tomographic projection images of the molecules. The 3-D orientations of individual
projection images are unknown and random. To avoid damaging the samples, the
electron dose must be kept low, resulting in a low signal-to-noise ratio (SNR). The
cryo-EM computational problem is reconstructing the 3-D molecular structure from
these projection images [36].

The renewed interest in cryo-EM led to a thorough investigation of its mathe-
matical and statistical foundations [80, 6]. In particular, a crucial challenge from
a statistical perspective is understanding the sample complexity of cryo-EM, i.e.,
the number of images that are required to obtain accurate reconstructions. A re-
markable result revealed an intimate connection between the sample complexity of
cryo-EM (and related statistical models) and the method of moments in the low
SNR regime. If the distribution of the 3-D rotations is uniform, the method of
moments reduces to autocorrelation analysis. In particular, it was shown that if \( d \)
is the lowest degree moment of the observations (i.e., the randomly oriented tomo-
graphic projections) that determines the molecular structure uniquely, a necessary
condition for recovery is \( n = \omega(\sigma^2 d) \), where \( \sigma^2 \) is the variance of the noise [67, 1].
Specifically, if the distribution of rotations is uniform, then the third-order autocor-
relation is the lowest order autocorrelation that determines a generic 3-D structure,
implying a sample complexity of \( n = \omega(\sigma^6) \) [4]. This agrees with long-standing em-
pirical evidence [79].

Autocorrelation analysis was first introduced to cryo-EM by Zvi Kam more than
40 years ago [50]. Kam showed that the second-order autocorrelation of the pro-
jection images (which can be estimated with \( n = \omega(\sigma^2) \) observations) determines
the 3-D structure up to a set of orthogonal matrices, under the assumption that
the rotations are drawn from a uniform distribution. A few methods have been
proposed to resolve the missing orthogonal matrices based on typically unavail-
able side information, such as homologous models with known structure, or a few
clean projections images, in order to construct ab initio models [12, 59, 46]. These
ab initio models can then be refined using expectation-maximization: the preva-
ient algorithmic framework for the cryo-EM reconstruction problem that aims to
maximize the non-convex posterior distribution [72, 74, 69]. In addition, it was
recently shown that if the distribution of rotations is non-uniform, then there is
at most a finite list of structures that agree with the second moment of the ob-
servations [76]. Techniques that are inspired by Kam’s method were also proposed
as a solution to the molecular reconstruction problem in X-ray free-electron lasers
(XFEL), which, akin to the reconstruction problem in cryo-EM, involves recovering
a 3-D structure from its randomly oriented diffraction images [71, 56, 24, 57]. In
contrast to cryo-EM, in XFEL the rotations are always uniformly distributed and
the reconstruction problem is more involved since the measurements consist of the
magnitudes of Fourier coefficients without their phases.

The first contribution of this paper is proving that if the sought-for molecular
structure can be described as a sparse combination of ideal point masses, then the
structure can be determined uniquely from the second-order autocorrelation of the observations, even if the rotations are distributed uniformly. This eliminates the orthogonal matrix ambiguity in Kam’s original paper [50]. This result is the first theoretical guarantee for unique recovery from the second moment. It combines the sparsity assumption with proof techniques from real algebraic geometry, substantially reducing the sample complexity of the cryo-EM reconstruction problem from \( n = \omega(\sigma^6) \) to \( n = \omega(\sigma^4) \). The argument is constructive in the sense that it provides a polynomial-time recovery algorithm. However, said algorithm is tailored to the specific model of point mass functions. It is not well-suited to data discretized into pixels or voxels because it hinges on the ability to cluster points in the support of the second moment into certain distinct components, see Remark 11. (Accurate clustering becomes difficult when data is discretized and the components are close to each other.)

The second contribution of this paper is a practical algorithm, fusing the second-order autocorrelation of the projections with a sparse representation of the molecular structure, and requiring only \( n = \omega(\sigma^4) \) observations. The algorithm builds on the realization that a typical 3-D structure can be represented by only a few coefficients in a suitable basis. Similar sparsity assumptions have been leveraged in a wide variety of scientific and engineering applications, including compressed sensing [25, 17, 28], image processing [27], and phase retrieval [15, 21, 47, 31, 88], to name a few. In cryo-EM, there have been several attempts to represent the 3-D structure as either a sparse mixture of Gaussians [87, 49, 48, 52, 89, 19, 70] or using alternative bases [86]. Yet, the sparsity property has still not been fully harnessed to represent and recover 3-D molecular structures, and it is not part of the standard computational pipeline of cryo-EM.

The technique is based on a new connection between Kam’s theory for cryo-EM and the crystallographic phase retrieval problem—recovering a sparse signal from its Fourier magnitudes. In particular, we adapt projection-based algorithms that were designed for the crystallographic phase retrieval problem to the cryo-EM setting. These algorithms were extensively validated on experimental X-ray crystallography datasets by prior researchers, see for example [34, 29, 60, 32, 31]. Here, we demonstrate on simulated data that they are also useful in constructing ab initio models in cryo-EM. They can be used to mitigate computational and model bias issues associated with the non-convexity of the cryo-EM reconstruction problem [82, 69, 43]. This new computational approach opens the door to merging more aspects of the phase retrieval and cryo-EM fields in future work.

The rest of the paper is organized as follows. In Section 2, we provide background on the reconstruction problem in cryo-EM, the method of moments, Kam’s theory, and the crystallographic phase retrieval problem. In Section 3, we prove that a structure composed of an ensemble of ideal point masses subject to uniform rotations can be recovered from the second order autocorrelation, implying a sample complexity of \( n = \omega(\sigma^4) \). Section 4 outlines the practical computational framework and presents numerical results. Section 5 concludes the paper and discusses potential theoretical and computational extensions.

2. Preliminaries

2.1. The cryo-EM problem. Cryo-EM reconstruction seeks to determine a 3-D molecular structure \( \Phi \) from its 2-D noisy tomographic projections, taken at random
viewing angles. In this work, we focus on the case of uniformly random rotations. Uniformity is often taken as a baseline model, was the setting in Kam’s paper [50], and is harder than the case of a non-uniform distribution of rotations in the sense that it requires asymptotically more images (without sparsity priors) [4, 76]. Formally, let \( \mu \) be the Haar probability measure on the compact group \( \text{SO}(3) \) of 3-D rotations, representing the uniform distribution. Assuming that we observe i.i.d. 2-D images of \( \Phi \) after it has been randomly rotated according to \( \mu \) and then tomographically projected to the plane, each projection image is modeled as:

\[
I_R(x, y) = \int_{z = -\infty}^{\infty} (R \cdot \Phi)(x, y, z) dz + \varepsilon(x, y), \quad R \sim \mu,
\]

where \( \varepsilon(x, y) \) is a white Gaussian noise with known variance \( \sigma^2 \), and \( R \cdot \Phi \) denotes the action of the rotation \( R \) on \( \Phi \). Here, typically, the variance of the noise \( \sigma^2 \) is much greater than the magnitude of the clean projection.

The cryo-EM problem is to estimate the molecular structure \( \Phi \) from \( n \) realizations of (1), i.e., from the 2-D observations \( I_{R_1}, I_{R_2}, \ldots, I_{R_n} \). In Section 5, we discuss how the proposed framework can be extended to account for additional aspects in the generative model for cryo-EM images.

### 2.2. The method of moments.

Our theoretical and computational contributions are based on the method of moments—a basic statistical inference technique tracing back to the seminal paper of Karl Pearson in the end of the 19th century. Specifically, we use the second moment of the observations (1), and relate it to the sought-for 3-D structure.

The (debiased) second observable moment is given by

\[
\mathcal{M}_2((x_1, y_1), (x_2, y_2)) = \frac{1}{n} \sum_{i=1}^{n} I_{R_i}(x_1, y_1)I_{R_i}(x_2, y_2) - B(\sigma^2),
\]

where \( B(\sigma^2) \) is a bias term that depends only on the noise variance. For large enough \( n \), we have

\[
\mathcal{M}_2((x_1, y_1), (x_2, y_2)) \approx M_2((x_1, y_1), (x_2, y_2)),
\]

where

\[
M_2((x_1, y_1), (x_2, y_2)) := \int_{\text{SO}(3)} I_R(x_1, y_1)I_R(x_2, y_2)d\mu(R) - B(\sigma^2),
\]

denotes the (debiased) population second moment, which is a function of \( \Phi \) through \( \text{SO}(3) \). More precisely, for \( n = \omega(\sigma^4) \) it holds \( \mathcal{M}_2 = M_2 + o(1) \) with high probability.

The idea of the method of moments is to find a structure \( \Phi \) which matches the observable moments. It is an alternative to other standard statistical estimation methods, e.g., maximum likelihood estimation (MLE). That said, a recent paper suggests that in the low SNR regime, the method of moments approximates the MLE [51].

### 2.3. Kam’s method.

We detail a specific approach to autocorrelation analysis in cryo-EM, introduced by Kam [50]. To this end, we need to introduce a convenient basis for representing a 3-D structure \( \Phi \), the spherical Bessel basis [2]. An expansion
of maximum degree $L$ is defined by first expanding the Fourier transform $\mathcal{F}(\Phi)$ of $\Phi$ in spherical harmonics as

$$\mathcal{F}(\Phi)(k, \theta, \varphi) \approx \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} A_{\ell m}(k) Y_{\ell m}^{\ell}(\theta, \varphi),$$

where $k$ denotes the radial frequency and $Y_{\ell m}^{\ell}(\theta, \varphi)$ are the spherical harmonics basis functions. In addition, the spherical harmonics coefficients $A_{\ell m}(k)$ are expanded by spherical Bessel functions, up to degree $S_{\ell}$, as

$$A_{\ell m}(k) \approx \sum_{s=1}^{S_{\ell}} a_{\ell m s} j_{\ell s}(k).$$

The functions $j_{\ell s}(k)$ are the normalized spherical Bessel functions. By allowing $L$ and $S_{\ell}$ to grow unboundedly, any structure of a biological molecule can be represented in this basis. However, when expanding 3-D molecular structures from discretized projection images, the Nyquist criterion applied to the projection images limits the amount of extractable information. This determines bounds on the maximally allowable truncation parameters $L$ and $S_{\ell}$; see Appendix C for a detailed description.

We aim to recover the coefficients $a_{\ell m s}$ (up to rotation and reflection in $\mathbb{R}^3$). As will be shown next, it is convenient to gather the coefficients into matrices $A_{\ell}$, of size $S_{\ell} \times (2\ell + 1)$, via $A_{\ell}(s, m) := a_{\ell m s}$, for each $\ell = 0, \ldots, L$.

Provided that the distribution of viewing angles is uniform, Kam \[50\] showed that the second moment of the Fourier transform of the projection images provides estimates for the following $S_{\ell} \times S_{\ell}$ matrices:

$$C_{\ell}(s_1, s_2) := \sum_{m=-\ell}^{\ell} A_{\ell}(s_1, m) \overline{A_{\ell}(s_2, m)} = A_{\ell} A_{\ell}^\dagger.$$

Applying the Cholesky decomposition to each $C_{\ell}$ in (7) and imposing $\Phi$ to be real-valued, knowledge of (7) identifies each matrix of coefficient $A_{\ell}$ up to an unknown real, orthogonal transformation, provided $S_{\ell} \geq 2\ell + 1$. That is, we can compute $A_{\ell} O_{\ell}$ for some unknown orthogonal matrix $O_{\ell}$ in the group $O(2\ell + 1)$. Therefore, the second moment determines $\Phi$ up to a set of orthogonal matrices. To recover these matrices, and thus the 3-D structure, additional information is required. In this paper, we suggest using a sparsity assumption.

2.4. Crystallographic phase retrieval. One of the contributions of this paper is to relate the cryo-EM reconstruction problem to crystallographic phase retrieval. Phase retrieval is the main computational challenge in X-ray crystallography, which is still a leading method for elucidating the atomic structure of molecules. The prevalence of crystallography is witnessed by the remarkable fact that 25 Nobel Prizes have been awarded for work directly or indirectly involving crystallography \[38\]. Although there exist additional important phase retrieval applications (see for example \[77, 24, 7, 5\]), X-ray crystallography is by far the most widely investigated application.

The crystallographic phase retrieval problem entails recovering a sparse signal $x$ from its periodic autocorrelation (or, equivalently, from its Fourier transform magnitudes, namely, its power spectrum). While simply stated, and despite its importance, the theoretical foundations of this problem have been investigated only
recently. In particular, it was conjectured that a generic sparse signal can be recovered from its periodic autocorrelation if the number of non-zero entries is smaller than half the signal’s length [9]. This conjecture was verified for a few cases. The relation of the crystallographic phase retrieval problem with the beltway problem from combinatorial optimization is explored in [39]. Our theoretical reconstruction guarantees in the following section can be viewed as analogous results in the setting of cryo-EM.

The standard algorithms for crystallographic phase retrieval build on two projection operators: one onto the measured data (the power spectrum) and the second onto the space of sparse signals. While simple algorithms that alternate between these two projections tend to quickly stagnate, a more sophisticated family of algorithms, based on reflections, shows excellent performance, though their running time is exponential in the sparsity level [31, 11]. These algorithms are tightly related to splitting methods, such as Douglas-Rachford and the alternating direction method of multipliers (ADMM), and have been applied to a wide variety of problems [32]. A main contribution of this paper is a modification of these algorithms to autocorrelation analysis for cryo-EM. In particular, we focus on one such algorithm, called relaxed-reflect-reflect (RRR), but alternative algorithms, such as Fienup’s hybrid input-output algorithm [34], the difference map algorithm [29], and the relaxed averaged alternating reflections algorithm [60], can be adapted to cryo-EM by the same strategy. Importantly, if the model is accurate (e.g., no noise) RRR iterations halt only when they find a solution that satisfies both constraints (defined by the projection operators). Thus, RRR does not suffer from local minima as gradient-based algorithms do.

3. Superior sample complexity: The second moment suffices for sums of point masses

This section presents our main theoretical result: the second moment suffices to recover an idealized sparse volume, i.e., a volume given as a weighted sum of point masses. We deduce that the second moment also suffices for a pixelated and blurred variant of the model. Our theorems imply an associated sample complexity of \( n = \omega(\sigma^4) \). This stands in contrast to previous results which do not assume sparsity. There, the third moment is required for recovery, and the associated sample complexity is \( n = \omega(\sigma^6) \) [4, 33].

3.1. Models and main theoretical results. We use an atomistic representation of a molecule. In our first idealized model, an atom is specified by a weighted Dirac delta function, and a molecule is a sum of such point masses. Let \( a_1, \ldots, a_p \in \mathbb{R}^3 \) be the 3-D points representing atom locations, and \( w_1, \ldots, w_p \) be positive weights corresponding to the scattering potentials of the individual atoms. Then

\[
\Phi := \sum_{i=1}^{p} w_i \delta_{a_i}
\]

is the molecule composed of the atoms \((a_i, w_i)\). Relabeling if necessary, we assume that the \( \ell^2 \)-norms \( \|a_i\| \) are in descending order.
We model each projection image $I_R$ as a mixture of $p$ Dirac delta functions on $\mathbb{R}^2$ plus noise:

\begin{equation}
I_R(x, y) := \sum_{i=1}^{p} w_i \delta_{\pi R a_i}(x, y) + \varepsilon(x, y).
\end{equation}

Here $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ denotes coordinate projection onto the first two coordinates, and $\varepsilon$ is a white Gaussian noise with (known) variance $\sigma^2$. Given $n$ 2-D images as in \(9\), the reconstruction problem is to recover the atoms $(a_i, w_i)$ in \(8\) up to a global rotation and reflection. (The reflection ambiguity exists because a molecule and its reflection in the microscope’s image plane are indistinguishable given cryo-EM data, see e.g. [36].) See Figure 1 for an illustration of the setup.

Under this model the (debiased) second population moment, obtained by substituting \(9\) into \(4\), reads:

\begin{equation}
M_2((x_1, y_1), (x_2, y_2)) = \sum_{i=1}^{p} \sum_{j=1}^{p} w_i w_j \times \int_{SO(3)} \delta_{\pi R a_i}(x_1, y_1) \delta_{\pi R a_j}(x_2, y_2) d\mu(R),
\end{equation}

where $\mu$ is the uniform distribution. Note that $M_2$ is a measure on $\mathbb{R}^2 \times \mathbb{R}^2$.

We introduce certain assumptions on the atom locations:

**A1.** The vectors $a_i$ are pairwise linearly independent;

**A2.** The norms $\|a_i\|$ are distinct.

We remark that conditions **A1** and **A2** are quite restrictive, e.g., ruling out molecules with nontrivial point-group symmetries.

Our first main theoretical result is stated as follows.
Theorem 1. Consider the model given by \( \text{Eq. (8)} \). Assume that conditions A1-A2 hold. Then the support of the second moment \( M_2 \) uniquely determines the set \( \{ (||a_i||^2, ||a_j||^2, (a_i, a_j)) : i, j = 1, \ldots, p \} \). Therefore, \( M_2 \) (Eq. \( \text{10} \)) uniquely determines the set of atom locations \( \{ a_i : i = 1, \ldots, p \} \) up to a rotation and reflection in \( \mathbb{R}^3 \).

Theorem 1 is proven in Subsection 3.3 after auxiliary results are given in Subsection 3.2.

Building on the uniqueness in Theorem 1, we obtain the following constructive result as well.

Theorem 2. Consider the model given by \( \text{Eq. (8)} \). Assume that \( p \geq 3 \) and conditions A1-A2 hold. Then Algorithm 2 (described in Subsection 3.4) recovers the set of atom locations \( \{ a_i, w_i : i = 1, \ldots, p \} \) from the second moment \( M_2 \) (Eq. (10)) in \( \mathcal{O}(p^2) \) flops.

Theorem 2 is proven in Subsection 3.4.

Prior works on the sample complexity of cryo-EM [4, 66, 1] do not directly apply to the model above. The principal reason is that the measurement formation defined by \( \text{Eq. (9)} \) is not finite-dimensional. Therefore, we consider a pixelated and blurred variant of the model. The molecule is still specified by a collection of atoms \( \{ (a_i, w_i) : i = 1, \ldots, p \} \). However, each projection image now consists of \( 2^m \times 2^m \) pixels:

\[
M_2^{[m]}((j_1, j_2), (j_3, j_4)) = \int_{j_3 \tau}^{(j_3+1) \tau} \int_{j_4 \tau}^{(j_4+1) \tau} \int_{j_2 \tau}^{(j_2+1) \tau} \int_{j_1 \tau}^{(j_1+1) \tau} f_{j_1, j_2}(x, y) \cdot k(x, y) dx dy + \varepsilon(j_1, j_2).
\]

Here, we discretized \([-1, 1] \) into equi-sized squares, where \( j_1, j_2 \in \{ -2^{m-1}, -2^{m-1}+1, \ldots, 2^{m-1} - 1 \} \) and \( \tau = 1/2^{m-1} \). Also, \( * \) denotes convolution and \( k(x, y) \) is the isotropic Gaussian kernel with fixed variance \( \kappa^2 \), i.e., \( k(x, y) = e^{-x^2+y^2}/\pi\kappa^2 \). Last, the noise satisfies \( \varepsilon(j_1, j_2) \overset{i.i.d.}{\sim} N(0, \sigma^2) \).

In the pixelated model, the (debiased) second population moment equals:

\[
M_2^{[m]}((j_1, j_2), (j_3, j_4)) = \int_{j_3 \tau}^{(j_3+1) \tau} \int_{j_4 \tau}^{(j_4+1) \tau} \int_{j_2 \tau}^{(j_2+1) \tau} \int_{j_1 \tau}^{(j_1+1) \tau} M_2((x_1, y_1), (x_2, y_2)) \cdot k(x_1, y_1)k(x_2, y_2) \cdot dx_1 dy_1 dx_2 dy_2.
\]

We now state our main result for the pixelated model. Its proof relies on Theorems 1 and 2.

Theorem 3. Consider the model given by \( \text{Eq. (11)} \). Fix an integer \( p \geq 3 \) and real numbers \( r > 0 \) and \( w_+ > w_- > 0 \). Assume that A1-A2 hold, and for each \( i = 1, \ldots, p \) we have \( ||a_i|| \leq r \) and \( w_- \leq w_i \leq w_+ \). Then there exists \( m' = m'(p, r, w_+, w_-) \) with the following property. Whenever \( m \geq m' \) and \( 2^m \times 2^m \) pixels are used in \( \text{Eq. (11)} \), then the second moment \( M_2^{[m]} \) (Eq. (12)) uniquely determines the set of atoms \( \{ (a_i, w_i) : i = 1, \ldots, p \} \) up to a rotation and reflection in \( \mathbb{R}^3 \).

As the details are technical, we prove Theorem 3 in the appendices. We only use two properties of the Gaussian kernel \( k \): that it is real-analytic and that its Fourier transform does not vanish.
Corollary 4. Assume the setting of Theorem 3 with $m \geq m'$. Then, the sample complexity for generic unique recovery (in the sense of [1]) is $n = \omega(\sigma^4)$ as $\sigma \to \infty$.

The rest of the section provides the proofs of Theorems 1 and 2, with Theorem 3, Corollary 4 and supporting results shown in the appendices. We emphasize that Algorithm 1 is a theoretical algorithm, not intended for use in practice due to its noise sensitivity as explained in Remark 11. By contrast, Algorithm 2 in the subsequent section is built for practical situations.

3.2. Support of $M_2$. To recover the atoms from $M_2$, the main information that we use is actually qualitative. Specifically, we rely on the particular structure of the support of the second moment $M_2$ in $\mathbb{R}^2 \times \mathbb{R}^2$. To describe this, we need to first understand the possible images of one pair of atoms.

Definition 5. For $i, j = 1, \ldots, p$, let $\theta_{ij} : SO(3) \to \mathbb{R}^2 \times \mathbb{R}^2$ be the map given by $\theta_{ij}(R) = (\pi R a_i, \pi R a_j)$.

Definition 6. For $i, j = 1, \ldots, p$, let $S_{ij} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ be the image of $\theta_{ij}$, i.e., $S_{ij} = \{(x_1, x_2) \in \mathbb{R}^2 \times \mathbb{R}^2 : \exists R \in SO(3) \text{ s.t. } \pi R a_i = x_1, \pi R a_j = x_2\}$.

The next lemma characterizes $S_{ij} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ as the solution set to a system of polynomial equations and inequalities. This will enable proof techniques from real algebraic geometry.

Lemma 7. $S_{ij}$ is connected, compact and semialgebraic. Letting $(x_1, y_1, x_2, y_2)$ be variables on $\mathbb{R}^2 \times \mathbb{R}^2$, then $S_{ij}$ is cut out by one quartic equation and two quadratic inequalities:

$$\left(\|a_i\|^2 - x_1^2 - y_1^2\right)\left(\|a_j\|^2 - x_2^2 - y_2^2\right) = (\langle a_i, a_j \rangle - x_1 x_2 - y_1 y_2)^2,$$

$$x_1^2 + y_1^2 \leq \|a_i\|^2 \quad \text{and} \quad x_2^2 + y_2^2 \leq \|a_j\|^2.$$

It has dimension 3 as a semialgebraic set if condition A1 holds.

A few different examples of the sets $S_{ij}$ are illustrated in Figure 2 for varying values of $\langle a_i, a_j \rangle$ and $\|a_i\|, \|a_j\|$. There we show the projection of $S_{ij}$ to $\mathbb{R}^3$ when $y_2$ is dropped.

The next result is immediate from Definitions 5 and 6.

Lemma 8. The second moment is

$$M_2 = \sum_{i,j=1}^k w_i w_j (\theta_{ij})_* (\mu),$$

where the subscripts indicate the pushforward measure defined by $(\theta_{ij})_* (\mu)(\cdot) = \mu(\theta_{ij}^{-1}(\cdot))$. In particular, the support of $M_2$ is $\cup_{i,j=1}^p S_{ij}$.

3.3. Information-theoretic uniqueness: Proof of Theorem 1. We begin by proving Theorem 1. The key is a converse to Lemma 7. While Lemma 7 implies the quartic equation in (13) (plus the quadratic inequalities there) determine the set $S_{ij}$, we need that $S_{ij}$ determines the quartic. The proof of this converse uses results from real algebraic geometry [14].
Lemma 9. Assume that condition A1 holds. Let $i \neq j$. Then the ideal of the real Zariski closure of $S_{ij}$ in $\mathbb{R}^2 \times \mathbb{R}^2$ is principal and generated by the quartic polynomial
\begin{equation}
q_{ij} = (\|a_i\|^2 - x_1^2 - y_1^2)(\|a_j\|^2 - x_2^2 - y_2^2) - (\langle a_i, a_j \rangle - x_1x_2 - y_1y_2)^2
\end{equation}
Further, $q_{ij}$ is irreducible over $\mathbb{R}$.

Corollary 10. Assume that conditions A1-A2 hold. Then, the irredundant irreducible decomposition of the Zariski closure of the support of $M_2$ is
\begin{equation}
\{ (x_1, x_2) : x_1 = x_2 \} \cup \bigcup_{i \neq j} S_{ij}.
\end{equation}
We can now prove the information-theoretic uniqueness.

Proof of Theorem 7. The support of $M_2$ determines the real radical prime ideal of each of top-dimensional irreducible component of its Zariski closure. By A1-A2, Corollary 10 and Lemma 9, these ideals are $\langle q_{ij} \rangle = \{ q_{ij}h : h \in \mathbb{R}[x_1, y_1, x_2, y_2] \}$ for $i \neq j$. The ideal $\langle q_{ij} \rangle$ uniquely determines $q_{ij}$, since the coefficient of $x_1^2y_2^2$ in (15) is 1. Extracting the coefficients of $x_1^2, x_1^2, x_1x_2$ in (15), $q_{ij}$ determines the triple $(\|a_i\|^2, \|a_j\|^2, \langle a_i, a_j \rangle)$. Ranging over $i, j$, we have proven that the support of $M_2$ fixes the set:
\begin{equation}
\{ (\|a_i\|^2, \|a_j\|^2, \langle a_i, a_j \rangle) : i \neq j \}.
\end{equation}
By A2 and our assumption the norms $\|a_i\|$ are descending, knowledge of (17) lets us fill in the Gram matrix:

\[
G = A^\top A \in \mathbb{R}^{p \times p},
\]

where

\[
A = (a_1 \ldots a_p) \in \mathbb{R}^{3 \times p}.
\]

However $G$ determines $A$ up to left multiplication by a $3 \times 3$ orthogonal matrix.

Indeed considering a truncated rank-3 eigendecomposition, we write

\[
G = QDQ^\top,
\]

where $Q \in \mathbb{R}^{p \times 3}$ has orthonormal columns and $D \in \mathbb{R}^{3 \times 3}$ is diagonal and positive-semidefinite. Then,

\[
A = OD^{1/2}Q^\top,
\]

for some $O \in O(3)$. Therefore, the atoms’ locations $a_i$ are determined up to a global rotation and reflection. □

3.4. Recovery algorithm: Proof of Theorem 2. Now, we move forward and prove Theorem 2. We present Algorithm 1 for efficiently recovering the atoms $(a_i, w_i)$ from $M_2$. The algorithm is theoretical in that it relies on oracle access to the following information.

**Assumption 1.** We assume oracle access to:

- O1. \{sample$(S_{ij}) : i \neq j\}$, where sample$(S_{ij})$ consists of four or more Zariski-generic points on $S_{ij}$;
- O2. the value of the measure $M_2$ on the set $S_{ij}$ for all $i \neq j$.

**Remark 11.** In principle, O1 and O2 can be estimated from the sample moment $M_2$ if $n = \omega(\sigma^4)$. It would require the ability to identify points in the support of $M_2$ and cluster them according to the components $S_{ij}$. However this encounters difficulty when dealing with noisy moments discretized in pixels. The next section is dedicated to a different computational framework, better-suited for practical settings.

Proceeding, Algorithm 1 interpolates sample$(S_{ij})$ to recover $q_{ij}$ from Lemma 12.

**Lemma 12.** Assume that condition A1 holds, and O1 is known. Let sample$(S_{ij}) = \{((x_{1k}, y_{1k}), (x_{2k}, y_{2k})) : k = 1, \ldots, |\text{sample}(S_{ij})|\}$. Consider the matrix

\[
\begin{pmatrix}
1 & x_{1k}^2 + y_{1k}^2 & x_{2k}^2 + y_{2k}^2 & \ldots \\
\vdots & x_{1k}x_{2k} + y_{1k}y_{2k} & x_{1k}y_{2k} + y_{1k}x_{2k} & \ldots \\
\end{pmatrix}^\top
\]

Then it has rank 4, with kernel spanned by

\[
\begin{pmatrix}
\|a_i\|^2\|a_j\|^2 - \langle a_i, a_j \rangle^2 \\
-\|a_i\|^2 \\
-\|a_j\|^2 \\
2\langle a_i, a_j \rangle \\
1
\end{pmatrix}.
\]
By Lemma 12, we compute the triples \( \{ (\|a_i\|^2, \|a_j\|^2, \langle a_i, a_j \rangle) : i \neq j \} \) in \( O(p^2) \) time, by forming the matrices (22) and computing their kernels. We then fill in the Gram matrix (18) as in the proof of Theorem 1. The atoms’ locations \( a_i \) are recovered from the truncated eigendecomposition (21).

The calculation of the weights \( w_i \) is based on the following.

**Lemma 13.** Assume that conditions A1–A2 hold. Then, for each \( i \neq j \), the measure of \( S_{ij} \) with respect to \( M_2 \) is

\[
M_2(S_{ij}) = w_i w_j.
\]

Therefore, \( O2 \) tells us all off-diagonal entries of \( ww^\top \in \mathbb{R}^{p \times p} \). We complete this uniquely to a rank-1 matrix by using

\[
(w w^\top)_{ii} = \frac{(w w^\top)_{ij}(w w^\top)_{ji}}{(w w^\top)_{jj}},
\]

where \( j, j' \) are any indices such that \( i, j, j' \) are all distinct. (This step requires \( p \geq 3 \).) The weights \( w_i \) are lastly recovered either by computing the leading eigenvector/eigenvalue pair of \( ww^\top \) or as the square root of the diagonal of \( ww^\top \), using the fact that the \( w \) are non-negative.

**Remark 14.** We note that (25) is a particular case of the problem of recovering a low-rank matrix with corrupted diagonal entries; see, e.g., [65, 72, 73] for more on that problem.

We summarize the procedure of this section in Algorithm 1

---

**Algorithm 1** Recovering a sparse structure from its second moment

**Input:** Second population moment \( M_2 \) as in (10)

**Output:** Atoms \( \{(a_i, w_i) : i = 1, \ldots, p\} \) up to a rotation and reflection in \( \mathbb{R}^3 \)

1. Access \( O1 \) in Assumption 1
2. Recover the unordered set \( \{ (\|a_i\|^2, \|a_j\|^2, \langle a_i, a_j \rangle) : i \neq j \} \) using Lemma 12
3. Fill in the Gram matrix \( G = A^\top A \) with \( A \) from (19)
4. Recover the \( a_i \) up to orthogonal transformation by computing a truncated eigendecomposition of \( G \) as in (21)
5. Access \( O2 \) in Assumption 1
6. Fill in the off-diagonal entries of \( ww^\top \) using Lemma 13
7. Complete \( ww^\top \) using (25)
8. Recover the \( w_i \) from \( ww^\top \)

**return** \( \{(a_i, w_i) : i = 1, \ldots, k\} \)

---

**Proof of Theorem 3** The considerations above show Algorithm 1 correctly recovers the set of atoms \( \{(a_i, w_i) : i = 1, \ldots, p\} \) from \( M_2 \) (up to a rotation and reflection in \( \mathbb{R}^3 \)). It costs \( O(p^2) \) in flops and storage once \( O1 \) and \( O2 \) are available if we use a randomized algorithm [64] to compute the truncated decomposition in (20).

**4. Kam’s method with sparsity constraints**

This section introduces a computational framework to leverage sparsity in recovering the underlying molecular structure. The goal is to devise a principled way to compute ab initio approximations of the underlying structures, that can
Figure 3. Reconstruction results when applying the proposed algorithm to three example structures. (A) Resolution of the reconstructed volumes as a function of the number of iterations. The resolution is determined by the 0.5 cutoff of the Fourier Shell Correlation (FSC). (B) Example FSC curves for reconstructed volumes as function of non-dimensionalized frequency. (C,D,E) Visualization of the reconstructed volumes for the three example structures: EMD-0409 [44], EMD-25892 [61] and the Shepp-Logan phantom [78], respectively. (Top) Ground-truth structure. (Middle) Truncation of ground-truth structure into the spherical Bessel basis using $L = 8, 12, 12$, respectively. (Bottom) Reconstructed volume returned by Algorithm 2 using $L = 8, 12, 12$, respectively. The visualizations were rendered by UCSF Chimera [68].
Analogously to (10), the moments of the projection images furnish information about the underlying 3-D structure. Unlike our theoretical results, however, we consider a general 3-D structure \( \Phi \) expanded in a spherical Bessel basis as in Section 2.3, and assume it can be represented by only a few wavelet coefficients. The next section introduces wavelet bases, and later we provide the details of the projection-based algorithm.

4.1. Wavelet bases. We encode sparsity of a 3-D molecular structure \( \Phi \) by a sparse expansion in wavelets [22]—a popular choice of sparsifying, localized bases in a wide range of applications [63]. Our algorithm can easily be adapted to any specific wavelet basis, and, more generally, to any choice of basis, for instance, sparsifying bases learned through data.

We denote the multilevel wavelet basis by \( f_{m,n} \), where \( m = 1, \ldots, m_{\text{max}} \) denotes the level of the wavelet and \( n = 1, \ldots, n_{\text{max}}(m) \) the index of the function within the level. As a shorthand, we define \( W : \mathbb{R}^{M \times M \times M} \rightarrow \mathbb{R}^{m_{\text{max}}n_{\text{max}}(m)} \) as the map sending a 3-D structure to its vector of coefficients when expanded in the wavelet basis, i.e.,

\[
W(\Phi) = (\langle \Phi, f_{m,n} \rangle)_{m,n=1}^{m_{\text{max}},n_{\text{max}}(m)}.
\]

Likewise, \( W^{-1} : \mathbb{R}^{m_{\text{max}}n_{\text{max}}(m)} \rightarrow \mathbb{R}^{M \times M \times M} \) then maps a wavelet coefficient vector into its 3-D expansion by

\[
W^{-1}(c_{m,n})_{m,n=1}^{m_{\text{max}},n_{\text{max}}(m)} = \sum_{m,n} c_{m,n} f_{m,n}.
\]

An additional advantage of using wavelet bases is that \( W \) and \( W^{-1} \) can then be applied in linear time \( O(M^3) \) using a fast wavelet transform [62].

4.2. Projection-based algorithm. For a discretized 3-D structure \( \Phi \) of size \( M \times M \times M \), we define the mapping \( SB : \mathbb{R}^{M \times M \times M} \rightarrow \prod_{\ell=0}^{L} \mathbb{C}^{S_{\ell} \times (2\ell+1)} \) of the structure into its coordinates in the spherical Bessel basis by

\[
SB(\Phi) = (A_0, A_1, \ldots, A_L).
\]

The inverse mapping \(SB^{-1} : \prod_{\ell=0}^{L} \mathbb{C}^{S_{\ell} \times (2\ell+1)} \rightarrow \mathbb{R}^{M \times M \times M} \) then expands a set of coefficients in the spherical Bessel basis into its corresponding 3-D structure:

\[
SB^{-1}(A_0, \ldots, A_L) = F^{-1} \left( \sum_{\ell,m,s} a_{\ell,m,s} j_{\ell s}(k) Y_{\ell m}(\theta, \varphi) \right).
\]

As discussed in Section 2, Kam’s method identifies matrices \( A_\ell O_\ell \), for \( O_\ell \) an unknown orthogonal matrix, for each \( \ell \) with \( S_\ell \geq 2\ell + 1 \). By possibly reducing the value of \( L \) to the largest index with this property, we will for ease of notation assume that this property holds for \( \ell = 0, \ldots, L \). Therefore, at the onset of the algorithm, we have access to a set of coefficient matrices \( B = (B_0, \ldots, B_L) \) satisfying

\[
B_\ell = A_\ell O_\ell, \quad O_\ell \in O(2\ell + 1),
\]

for unknown orthogonal matrices \( O_\ell \). Our algorithm aims to recover an approximation of these unknown orthogonal matrices, which leads to an approximation of \( \Phi \). This orthogonal matrix retrieval problem is an analogue to the problem of the missing phases in the phase retrieval problem [10]. We therefore adapt a popular
algorithm from the phase retrieval literature into the problem of cryo-EM. The algorithm repeatedly utilizes two projections onto the set of structures with a given sparsity level and a set determined by the projection images. These two projections are the main pillars of the algorithm and can be used in different ways, as explained next. But first, we define the two projection operators.

4.2.1. **First projection: Moment constraint.** We begin by defining the first projection operator, denoted by \( \rho_1 \), as the projection onto the set defined by the \( C_\ell \) matrices in (7). Let \( SB(\Phi) = (A_0, \ldots, A_L) \) in \( \prod_{\ell=0}^L \mathbb{C}^{S_\ell \times (2\ell+1)} \) be the ordered collection of matrices of coefficients in the spherical Bessel basis. Define \( \rho_1(\Phi) \) as the projection

\[
\rho_1(\Phi) = SB^{-1}(D_0, \ldots, D_L),
\]

where the matrices \( D_\ell \) are defined by

\[
(D_0, \ldots, D_L) = \operatorname{argmin}_{(D_0, \ldots, D_L)} \{ \|A_\ell - D_\ell\|_F : D_\ell D_\ell^* = C_\ell \},
\]

with \( C_\ell \) from (7). (32) is an instance of the Orthogonal Procrustes problem. Although it is a non-convex optimization problem, it can be solved in closed form in terms of the singular value decomposition of \( B_\ell^T A_\ell \), see e.g., [75]. In the implementation, the matrices defining the operations \( SB \) and \( SB^{-1} \) are precomputed. The computational complexity of subsequently solving an instance of (31) is then \( O(L^4 + \sum_{\ell=0}^L S_\ell \ell^2 + M^3 \log M + M^3 \sum_{\ell} (S_\ell)\) = \( O(L^4 + M^3 \log M + M^3 L^3) \), since typically \( S_\ell = O(L) \).

4.2.2. **Second projection: Sparsity constraint.** The projection \( \rho_2 \) promotes sparsity in a given local wavelet basis. For a structure \( \Phi \) and an integer \( K \), define \( \rho_2(\Phi, K) \) as the structure with wavelet coefficients obtained by retaining the \( K \) largest components of \( W(\Phi) \) and replacing the remaining elements by zero, i.e.,

\[
\rho_2(\Phi, K) = W^{-1}\left((\alpha_{m,n} c_{m,n})_{m,n=1}^{m_{\max}, n_{\max} (m)}\right),
\]

where the coefficients are defined by \( W(\Phi) = (c_{m,n})_{m,n=1}^{m_{\max}, n_{\max} (m)} \) and \( \alpha_{m,n} = 1 \) if \( c_{m,n} \) has magnitude among the \( K \) largest magnitudes of the \( c_{m,n} \), and zero otherwise. The computational complexity of this step is \( O(M^3) \). We again emphasize that, generally, any localized basis or frame can be used to define \( \rho_2 \), and we fix a wavelet basis for the sake of definiteness.

4.2.3. **Algorithm.** A straightforward algorithm to attempt to recover \( \Phi \) is through alternating projections. This procedure is described by fixing a sparsity level \( K \) and iterating the two projections \( \rho_1 \) and \( \rho_2 \) in turn. The use of the two projections in an alternating fashion is intended to promote convergence to an intersection point of the two sets. In the case of projecting onto convex sets, convergence results are known [20], but convergence is not guaranteed for the non-convex projections in (31) and (33). Indeed, for non-convex sets, alternating projection schemes frequently suffer from convergence to local minima, and a method to escape the local minima is required. To achieve this, the phase retrieval literature details different iteration schemes combining the two projections \( \rho_1 \) and \( \rho_2 \) in different ways, for
instance using the Relaxed-Reflect-Reflect (RRR) algorithm \[30, 31\]. In terms of the projection operators, this iterative scheme can be written out as

\[
\begin{align*}
\Phi^{(n+1/3)} &= \rho_1 \left( \Phi^{(n)} \right), \\
\Phi^{(n+2/3)} &= \rho_2 \left( 2\Phi^{(n+1/3)} - \Phi^{(n)}, K \right), \\
\Phi^{(n+1)} &= \Phi^{(n)} + \beta \left( \Phi^{(n+2/3)} - \Phi^{(n+1/3)} \right),
\end{align*}
\]

where \( \beta \in (0, 2) \) is a scalar hyperparameter. The algorithm is summarized in Algorithm 2. As aforementioned, other phase retrieval algorithms which are based on two projection operators, such as the difference map algorithm and the relaxed averaged alternating reflections algorithm, can be adapted to cryo-EM similarly.

**Algorithm 2** Recovering \( \Phi \)

**Input:** Projection images, sparsity level \( K \), maximum number of iterations \( N \), hyperparameter \( \beta \).

**Output:** Estimated structure \( \Phi \)

1. Form the matrices \( C_\ell \) from (7).
2. Compute the Cholesky factorizations of the \( C_\ell \) to produce matrices \( B_\ell \) in (30).
3. \( \Phi^{(0)} = SB^{-1}(B_0, \ldots, B_L) \)
4. For \( n = 0, \ldots, N - 1 \) do
   - \( \Phi^{(n+1/3)} = \rho_1 \left( \Phi^{(n)} \right) \)
   - \( \Phi^{(n+2/3)} = \rho_2 \left( 2\Phi^{(n+1/3)} - \Phi^{(n)}, K \right) \)
   - \( \Phi^{(n+1)} = \Phi^{(n)} + \beta \left( \Phi^{(n+2/3)} - \Phi^{(n+1/3)} \right) \)

Return \( \Phi^{(N-2/3)} \)

**4.3. Simulation results.** We apply Algorithm 2 to two structures from the online EM data bank \[58\], EMD-0409 \[44\] and EMD-25892 \[61\], as well as the Shepp-Logan phantom \[78\]. For each structure, we run Algorithm 2 for a given value of \( L \) and \( K \) to obtain a reconstruction. To measure the reconstruction quality, we follow the standard procedure in the cryo-EM community, and compute the Fourier shell correlation (FSC) between the estimated structure and the ground truth structure.
truth. Specifically, the FSC of two structures $\Phi_1$ and $\Phi_2$ is defined by
\begin{equation}
FSC(k) = \frac{\sum_{r_i : \|r_i\| = k} \mathcal{F}(\Phi_1)(r_i) \mathcal{F}(\Phi_2)(r_i)}{\sqrt{\sum_{\|r_i\| = k} |\mathcal{F}(\Phi_1)(r_i)|^2 \sum_{\|r_i\| = k} |\mathcal{F}(\Phi_2)(r_i)|^2}}.
\end{equation}

where one structure is the estimated structure, and the second is the ground truth. The FSC is real-valued because of symmetry of the summation. The resolution is determined when the FSC curve drops below 0.5.

EMD-0409 has dimensions $128 \times 128 \times 128$, with each voxel having physical length of $1.117$ Å. EMD-25892 has dimensions $320 \times 320 \times 320$, and voxel size $1.68$ Å. The volumes were downsampled by a factor of 2 and 5, respectively, to give structures of size $64 \times 64 \times 64$. The ground truth matrices were generated exactly and the matrices $B_\ell$ in (30) were generated using $O_\ell$ chosen uniformly at random. To fix the units for the Shepp-Logan phantom, we assume the voxels to have side length 1 Å. The simulations used Haar wavelets to define $W$.

The result of applying Algorithm 2 to each structure is shown in Figure 3. For all three example structures, during a run of the algorithm, the resolution initially rapidly improves. Afterwards, the improvement slows down and exhibits an exploratory and oscillating behavior. This is typical for RRR-based algorithms [31].

During the run of the algorithm, the optimal resolutions obtained for the three structures were 17.2 Å, 20.52 Å, and 4.59 Å, respectively, and the resolutions at the initialization of the algorithm were 21.3 Å, 329.8 Å, and 8.4 Å, respectively.

As a comparison, the resolutions between the ground truth structures and their truncation into the spherical Bessel bases with the chosen values of $L$ are 6.0 Å, 18.50 Å, and 2.20 Å, respectively; these resolutions are bounds on the optimally obtainable resolutions.

Figure 4 also shows a comparison of EMD-0409 with its reconstruction, truncated to different values of $L$. Visually, the reconstructed element captures the relevant features of the ground truth for each value of $L$. The resolution increases with $L$.

Movie 1 visualizes the reconstructed volume as a function of the iteration number. Note that the reconstruction at each step is visually similar to the ground truth, although the computed resolution noticeably improves during the run of the algorithm. This implies that even knowledge of the coefficients $A_\ell O_\ell$ with the wrong rotation matrices $O_\ell$ provides some information about the ground truth.

5. Discussion

The contribution of this paper is twofold. As the first contribution, our theoretical results imply that a sparse mixture of point masses can be uniquely recovered from the second order moment, even in the case of a uniform distribution of viewing angles, whereas previous work has only proven recovery using the third order moment. Thus, fewer images are required for reconstruction. This has a number of potential experimental implications. Firstly, since microscope time is expensive, this may greatly reduce the cost of the experimental part of the cryo-EM pipeline. This might be especially important for XFEL, where throughput is a major bottleneck and viewing directions are uniformly distributed [31, 54]. Secondly, it may enable reconstruction of structures where a limited number of projection images can be captured. This might be the case, for example, when the molecule may appear in several conformational states, and a limited number of images will be available for each conformation.
The second contribution is a new algorithm for ab initio modeling. The computational framework introduced in this article opens the door to incorporating a number of promising techniques from crystallographic phase retrieval into cryo-EM algorithms. There is, for instance, flexibility in choosing the projection operators ρ₁ and ρ₂. It may include biologically-oriented priors, such as minimum atom-atom distance or Wilson statistics [51, 40], or data-driven priors based on previously resolved structures [33]. A systematic study of adapting these techniques will be initiated in coming work. Additional future work includes extending the use of sparsifying priors in other parts of the cryo-EM reconstruction pipeline, for instance in existing approaches to iterative refinement [74] or in autocorrelation analysis using micrographs without particle picking [8]. Yet another important direction is to incorporate the sparsity prior into reconstruction by the method of moments when there is a non-uniform distribution of viewing directions [76].

Appendix A. Proofs of auxiliary results for Theorems 1 and 2

A.1. Proof of Lemma 7. Note that $S_{ij}$ is connected and compact, since $S_{ij} = \theta_{ij}(SO(3))$ and $SO(3)$ is compact and connected, while $\theta_{ij}$ is continuous. It also is semialgebraic, as $SO(3)$ is a real algebraic variety and $\theta_{ij}$ is a polynomial map (see the Tarski-Seidenberg theorem [14]).

Define $T_{ij} \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ as the set cut out by the three constraints in (13). Assume $((x_1, y_1), (x_2, y_2)) \in S_{ij}$. By definition of $S_{ij}$, there exist $R \in SO(3)$ and $z_1, z_2 \in \mathbb{R}$ such that $Ra_i = (x_1, y_1, z_1)^T$ and $Ra_j = (x_2, y_2, z_2)^T$. Then

$$\|a_i\|^2 - x_1^2 - y_1^2 = \|Ra_i\|^2 - x_1^2 - y_1^2 = z_1^2,$$

likewise $\|a_j\|^2 - x_2^2 - y_2^2 = z_2^2$, and

$$\langle a_i, a_j \rangle - x_1 x_2 - y_1 y_2 = (\langle Ra_i, Ra_j \rangle - x_1 x_2 - y_1 y_2)^2 = (z_1 z_2)^2.$$

Also, $x_1^2 + y_1^2 \leq \|Ra_i\|^2 = \|a_i\|^2$, and similarly $x_2^2 + y_2^2 \leq \|a_j\|^2$. These show that $((x_1, y_1), (x_2, y_2)) \in T_{ij}$, whence $S_{ij} \subseteq T_{ij}$.

For the converse, take $((x_1, y_1), (x_2, y_2)) \in T_{ij}$. Let

$$z_1 = \sqrt{\|a_i\|^2 - x_1^2 - y_1^2} \quad \text{and} \quad z_2 = \varepsilon \sqrt{\|a_j\|^2 - x_2^2 - y_2^2},$$

where $\varepsilon = \text{sign}(\langle a_i, a_j \rangle - x_1 x_2 - y_1 y_2)$. Put $b_i = (x_1, y_1, z_1)^T$ and $b_j = (x_2, y_2, z_2)^T$ in $\mathbb{R}^3$. By the choice of $z_1$ and $z_2$,

$$\|b_i\| = \|a_i\| \quad \text{and} \quad \|b_j\| = \|a_j\|.$$

Also, from the equality constraint in (13), it holds $z_1^2 z_2^2 = (\langle a_i, a_j \rangle - x_1 x_2 - y_1 y_2)^2$. This with the choice of $\varepsilon$ implies

$$\langle b_i, b_j \rangle = \langle a_i, a_j \rangle.$$

From (36) and (37), there exists $R \in SO(3)$ such $b_i = Ra_i$ and $b_j = Ra_j$. Hence $((x_1, y_1), (x_2, y_2)) \in S_{ij}$, whence $T_{ij} \subseteq S_{ij}$. We conclude $T_{ij} = S_{ij}$.

The dimension of $S_{ij}$ as a semialgebraic set is the maximal dimension of a cell in any cylindrical algebraic decomposition of it [14, Cor. 2.8.9]. This agrees with the maximal rank attained by the differential of $\theta_{ij}$:

$$\dim(S_{ij}) = \max_{R} \text{rank}(D\theta_{ij} : T_R(SO(3)) \to T_{\theta_{ij}}(\mathbb{R}^2 \times \mathbb{R}^2)).$$
where $T$ denotes tangent space. We recall that the tangent space to rotation matrices is parameterized by skew-symmetric matrices. Specifically, $T_R(\text{SO}(3)) = \{ [s] \times R : s \in \mathbb{R}^3 \}$, where

$$[s] \times := \begin{pmatrix} 0 & s_3 & -s_2 \\ -s_3 & 0 & s_1 \\ s_2 & -s_1 & 0 \end{pmatrix}.$$ 

Then, $D \theta_{ij}([s] \times R) = ([s] \times R_a, [s] \times R_a_j) \in \mathbb{R}^2 \times \mathbb{R}^2 = T_{\theta_{ij}(R)}(\mathbb{R}^2 \times \mathbb{R}^2)$. Putting $(x_1, y_1, z_1)^T := R_a$ and $(x_2, y_2, z_2)^T := R_a_j$, we rewrite

$$D \theta_{ij}([s] \times R) = W(R)s, \quad \text{where } W(R) := \begin{pmatrix} 0 & -z_1 & y_1 \\ z_1 & 0 & -x_1 \\ 0 & -z_2 & y_2 \\ z_2 & 0 & -x_2 \end{pmatrix}.$$ 

Thus, $\dim(S_{ij}) = 3$, unless $W(R)$ is rank-deficient for all $R \in \text{SO}(3)$. We claim it is rank-deficient for specific $R$ if and only if $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ are linearly dependent or $z_1 = z_2 = 0$. This is proven using a computer algebra system, e.g. [42]. Indeed, if $I$ is ideal in the ring $\mathbb{Q}[x_1, y_1, z_1, x_2, y_2, z_2]$ generated by the $3 \times 3$ minors of $W(R)$, the claim follows from calculating the primary decomposition [26]:

$$I = \langle z_1 y_2 - y_1 z_2, z_1 x_2 - x_1 z_2, y_1 x_2 - x_1 y_2 \rangle \cap \langle z_2^2, z_1 z_2, z_2^2 \rangle.$$ 

Given the claim, rank($W(R)$) < 3 for all $R \in \text{SO}(3)$ if and only if $a_i$ and $a_j$ are linearly dependent. In other words, $S_{ij}$ has dimension 3 if and only if $a_i$ and $a_j$ are linearly independent. The proof of Lemma 7 is complete. \hfill $\square$

A.2. Proof of Lemma 8. As mentioned in the main text, this is immediate from Definitions 5 and 6 and (10). \hfill $\square$

A.3. Proof of Lemma 9. If $X$ is a subset of a real Euclidean space $\mathbb{R}^k$, we write $\overline{X}$ for the Zariski closure in $\mathbb{R}^k$, and $\mathcal{I}(\overline{X})$ for the real radical ideal of $\overline{X}$.

First, note that $\overline{S_{ij}}$ is irreducible. This is because $S_{ij} = \theta_{ij}(\text{SO}(3))$, $\theta_{ij}$ is polynomial and $\text{SO}(3)$ is an irreducible algebraic variety. So, $\mathcal{I}(\overline{S_{ij}})$ is prime [14 Thm. 2.8.3(ii)]. Also, $\overline{S_{ij}}$ has dimension 3 as an algebraic variety. This is by [14 Prop. 2.8.2], and Lemma 7 which states $S_{ij}$ has dimension 3 as a semialgebraic set. So, $\mathcal{I}(\overline{S_{ij}})$ has height 1 [14 Def. 2.8.1]. Here, every prime ideal with height 1 is principal, as $\mathbb{R}[x_1, x_2, y_2]$ is a unique factorization domain. It follows that

$$\mathcal{I}(\overline{S_{ij}}) = \langle f \rangle,$$

for some irreducible polynomial $f \in \mathbb{R}[x_1, y_1, x_2, y_2]$, where angle brackets indicate ideal generation.

By Lemma 3 we know $S_{ij} \subseteq \mathcal{Z}(q_{ij})$, where $\mathcal{Z}$ denotes the zero set in $\mathbb{R}^2 \times \mathbb{R}^2$. Taking closures, $\overline{S_{ij}} \subseteq \mathcal{Z}(q_{ij})$. Equivalently, $q_{ij} \in \mathcal{I}(\overline{S_{ij}})$. By (38), this means $f$ evenly divides $q_{ij}$, say,

$$q_{ij} = fg,$$

for some $g \in \mathbb{R}[x_1, y_1, x_2, y_2]$. To conclude the proof, it suffices to prove that $g$ is a nonzero scalar. Then $\langle q_{ij} \rangle = \langle f \rangle = \mathcal{I}(\overline{S_{ij}})$, and $q_{ij}$ is irreducible because $f$ is.

For a contradiction, assume that $f$ has positive degree. Then, (39) implies

$$\langle q_{ij} \rangle_{\text{top}} = f_{\text{top}}g_{\text{top}},$$

where the subscript indicates the top total degree part of the polynomial. Here,

\[
q_{ij} = \left( \|a_i\|^2 \|a_j\|^2 - \langle a_i, a_j \rangle^2 \right) - \|a_j\|^2 x_1^2 - \|a_i\|^2 y_1^2 - \|a_i\|^2 y_2^2 + 2 \langle a_i, a_j \rangle x_1 x_2 + 2 \langle a_i, a_j \rangle y_1 y_2 + x_1^2 y_2^2 + y_1^2 x_2^2 - 2 x_1 y_1 x_2 y_2.
\]

(41)

Thus,

\[
(q_{ij})_{\text{top}} = x_2 y_2 + y_1 x_2^2 - 2 x_1 y_1 x_2 y_2 = (x_1 y_2 - y_1 x_2)^2.
\]

From (40), the assumption that \(f\) has positive degree and unique factorization, we deduce that (possibly after multiplying by nonzero scalars)

\[
f_{\text{top}} = g_{\text{top}} = x_1 y_2 - y_1 x_2.
\]

Therefore,

\[
f = x_1 y_2 - x_2 y_1 + \alpha x_1 + \beta y_1 + \gamma x_2 + \delta y_2 + \varepsilon,
\]

\[
g = x_1 y_2 - x_2 y_1 + \zeta x_1 + \eta y_1 + \theta x_2 + \iota y_2 + \kappa,
\]

for some \(\alpha, \beta, \gamma, \delta, \varepsilon, \zeta, \eta, \theta, \iota, \kappa \in \mathbb{R}\). Now we insert (42) and (41) into (39). Equating the constants and the coefficients of \(x_1^2, y_1^2, x_2^2, y_2^2\) gives

\[
\begin{pmatrix}
\|a_i\|^2 \|a_j\|^2 - \langle a_i, a_j \rangle^2 \\
-\|a_j\|^2 = \alpha \zeta \\
-\|a_j\|^2 = \beta \eta \\
-\|a_i\|^2 = \gamma \theta \\
-\|a_i\|^2 = \delta \iota.
\end{pmatrix}
\]

(43)

All the left-hand sides in (43) are nonzero because \(a_i\) and \(a_j\) are linearly independent. Therefore, \(\alpha, \ldots, \kappa\) are all nonzero. Next, we equate the coefficients of \(x_1 y_1, x_2 y_2\) in (39). The result is that

\[
\begin{pmatrix}
\alpha \\
\beta \\
\gamma \\
\delta
\end{pmatrix}
\times
\begin{pmatrix}
\zeta \\
\eta \\
\theta \\
\iota
\end{pmatrix} = \begin{pmatrix}
\kappa \\
\varepsilon
\end{pmatrix},
\]

whence the \(4 \times 2\) matrix has rank 1. So all its \(2 \times 2\) minors vanish, in particular

\[
\alpha \eta - \beta \zeta = 0.
\]

(44)

Finally, we equate the coefficients of \(x_1 y_1\) in (39):

\[
0 = \alpha \eta + \beta \zeta.
\]

(45)

Eq. (44) and (45) imply

\[
\alpha \eta = \beta \zeta = 0.
\]

(46)

But this contradicts the earlier finding that \(\alpha, \ldots, \kappa\) are all nonzero. So the assumption that \(f\) has positive degree is false, and \(f\) is a nonzero scalar. This proves Lemma 9. \(\square\)
A.4. Proof of Corollary 10 By Lemma 8 the support of $M_2$ is $\cup_{i,j=1}^p S_{ij}$. This has Zariski closure $\cup_{i,j=1}^p \overline{S_{ij}}$. We claim its irredundant irreducible decomposition is

$$\{(x_1, x_2) : x_1 = x_2\} \cup \bigcup_{i \neq j} \overline{S_{ij}}.$$  

The claim follows from several facts. First, for all $i \neq j$, $\overline{S_{ij}}$ is irreducible. It has dimension 3 and defining equation $q_{ij}$ [41]. By A1 and Lemma 8 When $i \neq j$, $i' \neq j'$ and $(i,j) \neq (i',j')$, then $q_{ij}$ and $q_{i'j'}$ are not scalar multiples of each other by A2 (cf. their coefficients on $x_1^2, x_2^2, x_1 y_1^2$). Hence $\overline{S_{ij}} \neq \overline{S_{i'j'}}$. Next, $\overline{S_{ii}} = \{(x_1, x_2) : x_1 = x_2\}$ (all i). Further, $\{(x_1, x_2) : x_1 = x_2\} \not\subseteq \overline{S_{ij}}$ (all $i \neq j$), because when we substitute $x_1 = x_2$ and $y_1 = y_2$ into [41], we get a nonzero result as the constant term does not vanish. All together, [47] is the claimed irredundant irreducible decomposition as wanted.

A.5. Proof of Lemma 12 This follows from $I(S_{ij}) = \langle q_{ij} \rangle$ (Lemma 9), the expression [41] for $q_{ij}$, and the proof of Thm. 3.

A.6. Proof of Lemma 13 From [14], we have

$$M_2(S_{ij}) = \sum_{i',j'=1}^p w_{i'} w_{j'} \mu(\theta_{i'j'}^{-1}(S_{ij})).$$

Here $\mu(\theta_{ij}^{-1}(S_{ij})) = \mu(\text{SO}(3)) = 1$ by definition of $S_{ij}$. On the other hand, for all $i',j' = 1, \ldots, p$ with $(i',j') \neq (i,j)$, $S_{ij} \cap S_{i'j'}$ is a semialgebraic set with positive codimension in $S_{i'j'}$ by the fact that [47] is an irredundant irreducible decomposition. Then $\theta_{i'j'}^{-1}(S_{ij}) = \theta_{i'j'}^{-1}(S_{ij} \cap S_{i'j'})$ is a semialgebraic set with positive codimension in $\text{SO}(3)$. Since $\mu$ is absolutely continuous, it implies $\mu(\theta_{i'j'}^{-1}(S_{ij})) = 0$. Eq. [24] follows.

APPENDIX B. SAMPLE COMPLEXITY: PROOFS OF THEOREMS 3 AND COROLLARY 4

B.1. Proof of Theorem 3 The proof is divided into several steps.

Step 0: We state a general fact about real analytic functions that we will use: Let $H(y, z) : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ be real analytic jointly in $(y, z)$. Let $\nu$ be a compactly supported and absolutely continuous measure on $\mathbb{R}^n$. Then $\int H(y, z) d\nu(z)$ is real-analytic in $y$. This can be justified by appropriately differentiating under the integral sign, see [35].

Step 1: We introduce notation.

First, define

$$X = (B(0, r) \times [w_-, w_+])^\times p \subseteq \mathbb{R}^{4p},$$

where $B(0, r)$ is the $\ell^2$-ball of radius $r$ in $\mathbb{R}^3$ centered at 0. $X$ is the space of possible molecules.

Next, slightly modifying the notation in the main text, write

$$M^m_2 : X \to \mathbb{R}^{2m \times 2m} \otimes \mathbb{R}^{2m \times 2m}$$
for the map associating a molecule to its pixelated second moment (12) when \(2^m \times 2^m\) pixels are used. Explicitly,

\[
M_2^{[m]} \{ (a_i, w_i) \} = \left\{ \sum_{i,j=1}^p w_i w_j \delta_{\pi R^a_i} (x_1, y_1) \delta_{\pi R^a_j} (x_2, y_2) \ast (k(x_1, y_1) k(x_2, y_2)) d\mu(R) dx_1 dy_1 dx_2 dy_2 \right\}
\]

for \(j_1, j_2, j_3, j_4 \in \{-2^{m-1}, \ldots, 2^{m-1} - 1\}\). Then, \(M_2^{[m]}\) is a real analytic function by Step 0. Indeed, the Gaussian kernel \(k(x, y)\) is real analytic, so the above integrand is real analytic in all variables. (Also, integration over \(\text{SO}(3)\) is replaced by integration against a compactly supported absolutely continuous measure on \(\mathbb{R}^3\) if we parameterize \(\text{SO}(3)\) with Euler angles.)

Thirdly, we put

\[
L^{[m]} : \mathbb{R}^{2m+1} \times \mathbb{R}^{2m+1} \to \mathbb{R}^{2m} \times \mathbb{R}^{2m}
\]

for the obvious linear map which lowers the resolution of the second moment by a factor of two, i.e. \(L^{[m]}(t) = s\) where

\[
s_{j_1, j_2, j_3, j_4} = \sum_{\gamma_1 \in \{0, 1\}} \sum_{\gamma_2 \in \{0, 1\}} \sum_{\gamma_3 \in \{0, 1\}} t_{2j_1 + \gamma_1, 2j_2 + \gamma_2, 2j_3 + \gamma_3, 2j_4 + \gamma_4}
\]

for \(j_1, j_2, j_3, j_4 \in \{-2^{m-1}, \ldots, 2^{m-1} - 1\}\). For all \(m\), it holds

\[
M_2^{[m]} = L^{[m]} \circ M_2^{[m+1]}
\]

Step 2: We prove that a certain stabilization occurs as \(m \to \infty\).

Consider \(X^{\times 2} = (B(0, r) \times [w_-, w_+])^\times p \times (B(0, r) \times [w_-, w_+])^\times p \subseteq \mathbb{R}^{8p}\), where the variables on \(\mathbb{R}^{8p}\) are \(\{a_i, w_i\}, \{b_i, v_i\}\). Regard \(X^{\times 2}\) as a semianalytic set (i.e., a subset of Euclidean space locally defined by real analytic equations and inequalities). Let \(\mathcal{O}(X^{\times 2})\) denote the ring of real analytic functions on \(X\). Then \(\mathcal{O}(X^{\times 2})\) is a Noetherian ring, because \(X\) is compact [47] Théorème I. 9. (Note that \(X^{\times 2}\) automatically satisfies the Stein hypothesis in loc. cit. since we are in the real case, see [83].)

We define

\[
\mathcal{I}^{[m]} = \text{ideal in } \mathcal{O}(X^{\times 2}) \text{ generated by the } 2^{4m}\text{ coordinate functions of } M_2^{[m]}(\{a_i, w_i\}) - M_2^{[m]}(\{b_i, v_i\}).
\]

For all \(m\), we have

\[
\mathcal{I}^{[m+1]} \supseteq \mathcal{I}^{[m]}
\]

by (49). From Noetherianity, there exists \(m' = m'(p, r, w_+, w_-)\) such that

\[
\mathcal{I}^{[m]} = \mathcal{I}^{[m']} \quad \forall m \geq m'.
\]

Thus the corresponding zero sets in \(X^{\times 2}\) stabilize too:

\[
\left\{ \{a_i, w_i\}, \{b_i, v_i\} : M_2^{[m]}(\{a_i, w_i\}) = M_2^{[m]}(\{b_i, v_i\}) \right\} \subseteq X^{\times 2}
\]

is constant in \(m\) if \(m \geq m'\). Equivalently, for all \(\{a_i, w_i\} \in X\) we have:

\[
\text{the fiber } (M_2^{[m]})^{-1}(M_2^{[m]}(\{a_i, w_i\})) \subseteq X \quad \text{is constant in } m \text{ if } m \geq m'.
\]

Step 3: We deduce an equality in which there is no pixelation.
Specifically, fix \{(a_i, w_i)\}, \{(b_i, v_i)\} ∈ X such that \{(a_i, w_i)\} satisfies A1-A2 in the main text and

\[ M_2^{[m']}((a_i, w_i)) = M_2^{[m']}((b_i, v_i)) \]

holds. We claim there is an equality between unipixelated (but still blurred) second moments:

\[ M_2((a_i, w_i)) * (k ⊗ k) = M_2((b_i, v_i)) * (k ⊗ k). \]

To see this, note by Step 0 that both sides of (52) are real-valued real analytic functions on \(\mathbb{R}^2 \times \mathbb{R}^2\). From continuity, if they differ on \([-1, 1]^2\) there must exist a product of sufficiently small pixels where their integrals differ, i.e. \(m ≥ m'\) and \(j_1, j_2, j_3, j_4 \in \{-2^{m-1}, \ldots, 2^{m-1} - 1\}\) such that

\[ M_2^{[m]}((a_i, w_i)) ((j_1, j_2), (j_3, j_4)) ≠ M_2^{[m']}((b_i, v_i)) ((j_1, j_2), (j_3, j_4)). \]

However, (53) contradicts (51) and (50). Thus, (52) holds on \([-1, 1]^2 \times [-1, 1]^2\). By real analyticity, (52) then holds on all of \(\mathbb{R}^2 \times \mathbb{R}^2\) as wanted.

**Step 4:** We undo the Gaussian blurring.

Continue with (52). Because \(M_2(\cdot)\) is compactly supported, it identifies with a tempered distribution. The Fourier transform is thus applicable to (52). By the convolution theorem \([43\text{ Thm. 7.1.15}]\), it gives

\[ \hat{M}_2((a_i, w_i)) \hat{k} ⊗ \hat{k} = \hat{M}_2((b_i, v_i)) \hat{k} ⊗ \hat{k}. \]

Note that the Paley-Wiener theorem \([43\text{ Thm. 7.1.14}]\) implies \(\hat{M}_2(\cdot)\) is a function rather just a distribution (moreover it is extendable to an entire function). Also, \(\hat{k} ⊗ \hat{k}\) is a Gaussian function. Therefore, (54) can be regarded as an equality of functions rather than just distributions. Since \(\hat{k} ⊗ \hat{k} ≠ 0\) everywhere, it implies

\[ \hat{M}_2((a_i, w_i)) = \hat{M}_2((b_i, v_i)), \]

whence

\[ M_2((a_i, w_i)) = M_2((b_i, v_i)), \]

using the fact that the Fourier transform is an automorphism on tempered distributions.

**Step 5:** We use Theorem 2 to conclude.

The above steps have shown: there exists \(m' = m'(p, r, w_+, w_-)\) such that if \(m ≥ m'\) then \(\{(a_i, w_i)\}, \{(b_i, v_i)\} ∈ X\) and \(M_2^{[m]}((a_i, w_i)) = M_2^{[m']}((b_i, v_i))\) imply \(M_2((a_i, w_i)) = M_2((b_i, v_i))\). However by Theorem 2 \(M_2((a_i, w_i)) = M_2((b_i, v_i))\) implies \(\{(a_i, w_i)\}\) and \(\{(b_i, v_i)\}\) are equal up to a rotation/reflection in \(\mathbb{R}^3\), provided \(\{(a_i, w_i)\}\) satisfies A1-A2.

The proof of Theorem 3 is complete. □

B.2. **Proof of Corollary 4.** This now follows immediately from [41 Sec. 3] or [66 Sec. 2], because by Theorem 3 the second moment \(M_2^{[m]}((a_i, w_i))\) uniquely determines the signal \(\{(a_i, w_i)\}\) up to the group action of \(O(3)\), provided \(\{(a_i, w_i)\}\) satisfies the Zariski-open conditions A1-A2. □
Appendix C. Normalized Bessel functions and the Nyquist criterion

The spherical Bessel basis defined in the main text uses the normalized spherical Bessel functions \( j_{\ell s}(k) \) defined by

\[
j_{\ell s}(k) = \frac{1}{c\sqrt{\pi j_{\ell+1}(R_{\ell,s})}} j_{\ell}(R_{\ell,s} c),
\]

where \( j_{\ell} \) is the \( \ell \)th Bessel function of the first kind \cite[Eq. 10.2.1]{23}, \( c \) the bandlimit of the projection images and \( R_{\ell,s} \) the \( s \)th positive solution to \( j_{\ell} = 0 \). The Nyquist criterion determines the maximally allowable value of the truncation parameter \( S_{\ell} \) by defining \( S_{\ell} \) as the largest integer \( s \) satisfying,

\[
R_{\ell,s+1} \leq 2\pi c R,
\]

assuming the projection images are supported on a disk of radius \( R \) \cite[Eq. 10.2.1]{13}. Our numerical experiments used \( c = 0.5 \) and a radius \( R \) of 32 voxels.

References


Email address: bendory@tauex.tau.ac.il

Email address: ykhoo@uchicago.edu

Email address: jkileel@math.utexas.edu

Email address: hm6655@princeton.edu

Email address: amits@math.princeton.edu