

MICROGRAPH MOMENTS

1. CONTINUOUS CASE

Let $\mathcal{I} : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a micrograph, let $I_\omega : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a projection of a volume V corresponding to orientation $\omega \in \text{SO}(3)$. Suppose that \mathcal{I} contains N translated instances of I_ω for different orientations ω and is zero outside of these instances, $\mathcal{I}(\mathbf{r}) = 0$ for $\|\mathbf{r}\| > M$, and that $I_\omega(\mathbf{r}) = 0$ whenever $\|\mathbf{r}\| > L$ for all ω , and for any (fixed) norm $\|\cdot\|$. Let S_i denote the support of the instance I_{ω_i} in \mathcal{I} , and let $Z = \mathbb{R}^2 - \bigcup_i S_i$, so $\mathcal{I}(\mathbf{r}) = 0$ if $\mathbf{r} \in Z$. Also denote by \mathbf{c}_i the center of the instance of I_{ω_i} in \mathcal{I} , so $\mathcal{I}(\mathbf{r}) = I_{\omega_i}(\mathbf{r} - \mathbf{c}_i)$ whenever $\mathbf{r} \in S_i$.

We shall assume that the instances of I_ω in \mathcal{I} are well-separated. Specifically, suppose

$$d = \min_{i,j} \inf_{\mathbf{r}_i \in S_i, \mathbf{r}_j \in S_j} \|\mathbf{r}_i - \mathbf{r}_j\| > 2L,$$

i.e. the minimum separation between any two projections is at least the diameter of the support of the molecule.

For a general function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $f(\mathbf{r}) = 0$ for $\|\mathbf{r}\| > L$, define its k th moment as

$$m_k[f](\Delta \mathbf{r}_1, \dots, \Delta \mathbf{r}_k) = \frac{1}{L^n} \int_{\|\mathbf{r}\| \leq L} f(\mathbf{r}) f(\mathbf{r} + \Delta \mathbf{r}_1) \cdots f(\mathbf{r} + \Delta \mathbf{r}_k) d\mathbf{r}$$

where $m_k[f](\Delta \mathbf{r}_1, \dots, \Delta \mathbf{r}_k) = 0$ if $\|\Delta \mathbf{r}_i\| > 2L$ for any i , as then $\|\mathbf{r} + \Delta \mathbf{r}_i\| \geq \|\Delta \mathbf{r}_i\| - \|\mathbf{r}\| > 2L - L = L$, and hence $f(\mathbf{r} + \Delta \mathbf{r}_i) = 0$ for all $\|\mathbf{r}\| \leq L$. Note also that if we take the Fourier transform of the k th moment, we get [Yellott and Iverson, 2017]

$$\widehat{m_k[f]}(\mathbf{k}_1, \dots, \mathbf{k}_k) = \frac{1}{L^n} \widehat{f} \left(\sum_{i=1}^k \mathbf{k}_i \right) \prod_{i=1}^k \widehat{f}(\mathbf{k}_i).$$

Finally, for convenience we shall take the above integral to be over \mathbb{R}^2 , always possible as $f(\mathbf{r}) = 0$ whenever $\|\mathbf{r}\| > L$.

Remark 1.1. The normalization $1/L^n$ above is necessary to make the moments of the micrograph finite. Specifically, the moments of the micrograph are shown below to be *sums* of moments of the projections (because a micrograph can contain at most countably many projections - there's a point with rational coordinates in every S_i), and in the worst case $N = \Theta(M^2)$. To see this, note that N is upper-bounded by the ratio of the area of the micrograph, proportional to M^2 , by that of a projection, proportional to L^2 , (so trivially $N = O(M^2)$), and in the worst case, equals that ratio (e.g., tile a square micrograph with square projections).

We consider the moments of \mathcal{I} in the limit $N \rightarrow \infty$. To take this limit, we shall assume that $N = \Omega(M^2)$ (so by the above argument, in fact $N = \Theta(M^2)$), so $\gamma = \lim_{N \rightarrow \infty} \frac{N}{M^2} \in (0, 1)$. We also assume that the orientations of the projections

in \mathcal{I} are uniformly distributed over the sphere. For the first moment,

$$m_1[\mathcal{I}] = \frac{NL^2}{M^2} \cdot \frac{1}{N} \sum_{i=1}^N \frac{1}{L^2} \int_{\mathbb{R}^2} I_{\omega_i}(\mathbf{r}) d\mathbf{r} \rightarrow \gamma \int_{\text{SO}(3)} L^2 m_1[I_\omega] d\omega = \gamma L m_1[V],$$

since $\int_{\mathbb{R}^2} I_\omega(x, y) dx dy = \widehat{I}_\omega(\mathbf{0}) = \widehat{V}(\mathbf{0})$ as the Fourier transform of any projection corresponds to a 2D slice of the Fourier transform of the volume passing through the origin.

For the second and third moments, note that if $\mathbf{r} \in S_i$ for some i , then $\mathbf{r} + \Delta\mathbf{r} \in S_i \cup Z$ whenever $\|\Delta\mathbf{r}\| \leq 2L$ by the separation assumption $d > 2L$. Since $\mathbb{R}^2 = Z \sqcup \bigsqcup_{i=1}^N S_i$, and $\mathcal{I}(\mathbf{r}) = 0$ if $\mathbf{r} \in Z$, we have for $\|\Delta\mathbf{r}\| \leq 2L$

$$\begin{aligned} m_2[\mathcal{I}](\Delta\mathbf{r}) &= \frac{1}{M^2} \int_{\mathbb{R}^2} \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{M^2} \int_Z \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) d\mathbf{r} + \frac{1}{M^2} \sum_{i=1}^N \int_{S_i} \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) d\mathbf{r} \\ &= \frac{1}{M^2} \sum_{i=1}^N \int_{\mathbb{R}^2} I_{\omega_i}(\mathbf{r}) I_{\omega_i}(\mathbf{r} + \Delta\mathbf{r}) d\mathbf{r} \\ &= \frac{N}{M^2} \cdot \frac{1}{N} \sum_{i=1}^N L^2 m_2[I_{\omega_i}](\Delta\mathbf{r}) \\ &\rightarrow \gamma \int_{\text{SO}(3)} L^2 m_2[I_\omega](\Delta\mathbf{r}) d\omega. \end{aligned}$$

If we set $m_2(\Delta\mathbf{r}) = 0$ for $\|\Delta\mathbf{r}\| > 2L$, the above equality holds for all $\Delta\mathbf{r}$, and taking its Fourier transform (with respect to $\Delta\mathbf{r}$) and interchanging the Fourier integral with the integral over $\text{SO}(3)$ (which can *always* be done), we get

$$\widehat{m_2[\mathcal{I}]}(\mathbf{k}) \rightarrow \gamma \int_{\text{SO}(3)} L^2 \widehat{m_2[I_\omega]}(\mathbf{k}) d\omega = \gamma \langle |\widehat{I}_\omega(\mathbf{k})|^2 \rangle_\omega,$$

where $\langle \cdot \rangle_\omega$ denotes average over the orientations, so over $\text{SO}(3)$.

Similarly, for the third moment if $\mathbf{r} \in S_i$ for some i then $\mathbf{r} + \Delta\mathbf{r}_1, \mathbf{r} + \Delta\mathbf{r}_2 \in S_i \cup Z$, and hence we again get for $\|\Delta\mathbf{r}_1\| \leq 2L$ and $\|\Delta\mathbf{r}_2\| \leq 2L$ that

$$\begin{aligned} m_3[\mathcal{I}](\Delta\mathbf{r}_1, \Delta\mathbf{r}_2) &= \frac{1}{M^2} \int_{\mathbb{R}^2} \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_1) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_2) d\mathbf{r} \\ &= \frac{1}{M^2} \int_Z \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_1) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_2) d\mathbf{r} + \frac{1}{M^2} \sum_{i=1}^N \int_{S_i} \mathcal{I}(\mathbf{r}) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_1) \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}_2) d\mathbf{r} \\ &= \frac{1}{M^2} \sum_{i=1}^N \int_{\mathbb{R}^2} I_{\omega_i}(\mathbf{r}) I_{\omega_i}(\mathbf{r} + \Delta\mathbf{r}_1) I_{\omega_i}(\mathbf{r} + \Delta\mathbf{r}_2) d\mathbf{r} \\ &= \frac{N}{M^2} \cdot \frac{1}{N} \sum_{i=1}^N L^2 m_3[I_{\omega_i}](\Delta\mathbf{r}_1, \Delta\mathbf{r}_2) \\ &\rightarrow \gamma \int_{\text{SO}(3)} L^2 m_3[I_\omega](\Delta\mathbf{r}_1, \Delta\mathbf{r}_2) d\omega. \end{aligned}$$

Again, setting $m_3(\Delta \mathbf{r}_1, \Delta \mathbf{r}_2) = 0$ if either $\|\Delta \mathbf{r}_1\| > 2L$ or $\|\Delta \mathbf{r}_2\| > 2L$, we also get

$$\widehat{m_3[\mathcal{I}]}(\mathbf{k}_1, \mathbf{k}_2) = \gamma \langle \widehat{I}_\omega(\mathbf{k}_1) \widehat{I}_\omega(\mathbf{k}_2) \widehat{I}_\omega(\mathbf{k}_1 + \mathbf{k}_2) \rangle_\omega.$$

2. DISCRETE CASE

In the discrete case, assume $\mathcal{I} \in \mathbb{R}^{M \times M}$ and $I_\omega \in \mathbb{R}^{L \times L}$ (since M, L denote the support of these signals in some norm, this is valid if for instance we restrict ourselves to ℓ_p norms, since then if $\|\mathbf{r}\|_p \leq M$ then $\|\mathbf{r}\|_\infty \leq M$, while $\|(M, 0, \dots, 0)^T\|_p = \|(M, 0, \dots, 0)^T\|_\infty = 1$). Once again, we shall assume $N = \Omega(M^2)$.

In the discrete case, we define the k th moment of a signal $x \in \mathbb{R}^{L \times L}$ as

$$m_k[x][(\Delta i_1, \Delta j_1); \dots; (\Delta i_k, \Delta j_k)] = \frac{1}{L^2} \sum_{i,j=1}^L x[i, j] x[i + \Delta i_1, j + \Delta j_1] \cdots x[i + \Delta i_k, j + \Delta j_k],$$

where we set $x[i, j] = 0$ if either $i > L$ or $j > L$.

The derivation is then identical as for the continuous case, with the exception that now $\gamma = \lim_{N \rightarrow \infty} \frac{NL^2}{M^2} \in (0, 1)$ is the occupancy factor.

3. STEERING

The first two moments are cheap to compute and store - m_1 is a scalar and m_2 is radially symmetric (in the limit $N \rightarrow \infty$), so it suffices to compute a single ray of it. The third moment however is infeasible to store - it has $(2L - 1)^4$ entries, and in modern microscopes we can easily have $L \geq 300$. In addition, the above scheme does not average over in-plane rotations, which can improve estimation in the presence of noise (and accelerate convergence to the population moments without noise). We therefore propose the following steering procedure:

The third moment $m_3[\mathcal{I}](\Delta \mathbf{r}_1, \Delta \mathbf{r}_2)$ as defined above is compactly supported (at least numerically) with respect to each of its variables in both real and Fourier space, with support radii $2L$ and $1/2$ in real and Fourier space, respectively (see above formulas for m_3 in real and Fourier space). Therefore, we can expand it in a suitable product basis that is compactly supported in both real and Fourier space, e.g. Prolate Spheroidal Wave Functions (PSWFs) or Fourier-Bessel (FB):

$$m_3[\mathcal{I}](\Delta \mathbf{r}_1, \Delta \mathbf{r}_2) = \sum_{k_1, k_2 = -\infty}^{\infty} \sum_{q_1, q_2 = 1}^{\infty} m_3(k_1, q_1; k_2, q_2) \psi_{k_1, q_1}(\Delta \mathbf{r}_1) \overline{\psi_{k_2, q_2}(\Delta \mathbf{r}_2)},$$

where we conjugate the second set of basis functions for convenience, that will become apparent below (of course, the set of conjugates is also a basis, as $\overline{\psi_{k, q}} = \psi_{-k, q}$ for PSWFs and $\overline{\psi_{k, q}} = (-1)^k \psi_{-k, q}$ for FB). We shall assume that the basis $\{\psi_{k, q}(\mathbf{r})\}$ is orthonormal, so $\int_{\mathbb{R}^2} \psi_{k_1, q_1}(\mathbf{r}) \overline{\psi_{k_2, q_2}(\mathbf{r})} d\mathbf{r} = \delta_{k_1, k_2} \delta_{q_1, q_2}$, valid for both PSWFs and FB. Then, the coefficients are given by

$$\begin{aligned} m_3(k_1, q_1; k_2, q_2) &= \int_{\|\Delta \mathbf{r}_1\|, \|\Delta \mathbf{r}_2\| \leq 2L} m_3[\mathcal{I}](\Delta \mathbf{r}_1, \Delta \mathbf{r}_2) \overline{\psi_{k_1, q_1}(\Delta \mathbf{r}_1)} \psi_{k_2, q_2}(\Delta \mathbf{r}_2) d\Delta \mathbf{r}_1 d\Delta \mathbf{r}_2 \\ &= \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) \left(\int_{\|\Delta \mathbf{r}_1\| \leq 2L} \mathcal{I}(\mathbf{r} + \Delta \mathbf{r}_1) \overline{\psi_{k_1, q_1}(\Delta \mathbf{r}_1)} d\Delta \mathbf{r}_1 \right) \left(\int_{\|\Delta \mathbf{r}_2\| \leq 2L} \mathcal{I}(\mathbf{r} + \Delta \mathbf{r}_2) \psi_{k_2, q_2}(\Delta \mathbf{r}_2) d\Delta \mathbf{r}_2 \right). \end{aligned}$$

Note that if we expand $\mathcal{I}(\mathbf{r} + \Delta\mathbf{r})$ in $\{\psi_{k,q}(\Delta\mathbf{r})\}$ for fixed \mathbf{r} and $\|\Delta\mathbf{r}\| \leq 2L$, so

$$\mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) = \sum_{k,q} a_{k,q}(\mathbf{r}) \psi_{k,q}(\Delta\mathbf{r}), \quad a_{k,q}(\mathbf{r}) = \int_{\|\Delta\mathbf{r}\| \leq 2L} \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) \overline{\psi_{k,q}(\Delta\mathbf{r})},$$

and use the fact that \mathcal{I} is real, our expression becomes

$$\mathbf{m}_3(k_1, q_1; k_2, q_2) = \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) a_{k_1, q_1}(\mathbf{r}) \overline{a_{k_2, q_2}(\mathbf{r})}.$$

We further assume that all in-plane rotations of the micrograph and its reflection, or equivalently, all in-plane rotations of each such disc of radius $2L$ and its reflection, are present in our dataset. Noting that rotations and reflections commute with the Fourier transform and using the derivation of [Zhao, Landa], the expansion of the disc about \mathbf{r} by an angle α is given by

$$\mathcal{I}^{\alpha,+}(\mathbf{r} + \Delta\mathbf{r}) = \sum_{k,q} a_{k,q} e^{-ik\alpha} \psi_{k,q}(\Delta\mathbf{r}),$$

and the expansion of the reflection of that disc rotated by an angle α by

$$\mathcal{I}^{\alpha,-}(\mathbf{r} + \Delta\mathbf{r}) = \sum_{k,q} \overline{a_{k,q}} e^{-ik\alpha} \psi_{k,q}(\Delta\mathbf{r}),$$

where we used the fact that for real-valued images we have $a_{-k,q} = \overline{a_{k,q}}$ for both PSWFs and FB. We thus have

$$\begin{aligned} \mathbf{m}_3(k_1, q_1; k_2, q_2) &= \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) \left(\frac{1}{4\pi} \int_0^{2\pi} [a_{k_1, q_1}(\mathbf{r}) \overline{a_{k_2, q_2}(\mathbf{r})} + \overline{a_{k_1, q_1}(\mathbf{r})} a_{k_2, q_2}(\mathbf{r})] e^{-i(k_1 - k_2)\alpha} d\alpha \right), \\ &= \delta_{k_1, k_2} \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) \Re\{a_{k_1, q_1}(\mathbf{r}) \overline{a_{k_2, q_2}(\mathbf{r})}\} \\ &= \Re \left\{ \delta_{k_1, k_2} \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) a_{k_1, q_1}(\mathbf{r}) \overline{a_{k_2, q_2}(\mathbf{r})} \right\}, \end{aligned}$$

Thus, the bispectrum in our steerable basis $\mathbf{m}_3(k_1, q_1; k_2, q_2)$ is block-diagonal, and is effectively a 3-tensor.

Similarly, the power spectrum is compactly supported in space (with support $2L - 1$) and bandlimited as in Fourier space it is the average squared magnitude of the Fourier transform of the projections, each of which is supposedly bandlimited. Therefore, we may expand it as

$$m_2[\mathcal{I}](\Delta\mathbf{r}) = \sum_{k,q} \mathbf{m}_2(k, q) \psi_{k,q}(\Delta\mathbf{r}),$$

and obtain the expansion coefficients as

$$\begin{aligned} \mathbf{m}_2(k, q) &= \int_{\Delta\mathbf{r}} m_2[\mathcal{I}](\Delta\mathbf{r}) \overline{\psi_{k,q}(\Delta\mathbf{r})} \\ &= \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) \int_{\Delta\mathbf{r}} \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) \overline{\psi_{k,q}(\Delta\mathbf{r})} \\ &= \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) a_{k,q}(\mathbf{r}). \end{aligned}$$

Taking all rotations of \mathcal{I} and its reflection, we get

$$\begin{aligned} \mathfrak{m}_2(k, q) &= \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) \left(\frac{1}{4\pi} \int_0^{2\pi} [a_{k,q}(\mathbf{r}) + \overline{a_{k,q}(\mathbf{r})}] e^{-ik\alpha} d\alpha \right) \\ &= \delta_{k,0} \int_{\mathbf{r}} \mathcal{I}(\mathbf{r}) a_{0,q}(\mathbf{r}), \end{aligned}$$

where in the last equality we dropped the real part since both $\mathcal{I}(\mathbf{r})$ and $\psi_{0,q}$ are real valued, and hence $a_{0,q}(\mathbf{r}) = \int_{\Delta\mathbf{r}} \mathcal{I}(\mathbf{r} + \Delta\mathbf{r}) \psi_{0,q}(\Delta\mathbf{r})$ is real as well. Thus, the average power spectrum is effectively a vector.

Remark 3.1. In practice, we are going to expand many non-centered windows from the micrograph in our steerable basis. While we technically assume that these windows are zero outside, and hence compactly supported, because of the centering of PSWFs and FB this may require infinitely many terms in the summation to adequately approximate. In other words, while the above formulas are correct when expressed as integrals, we may be losing accuracy when we expand these non-centered images numerically.

Remark 3.2. The triple correlation is a 6-tensor and the bispectrum is a 4-tensor, so we lost two degrees of freedom to invariance under translations and one degree of freedom to invariance under in-plane rotations.

4. CONNECTION TO VOLUME

We derive a relation between the steered bispectrum derived in Sect. 3 to the volume, expanded in a suitable basis. Specifically, expand the Fourier-transformed volume as

$$\widehat{V}(c\mathbf{k}) = \sum_{\ell, m, s} a_{\ell, m, s} \Phi_{\ell, s}(k) Y_{\ell, m}(\mathbf{k}/k),$$

for $k \leq 1$ and zero otherwise, where c is the assumed bandlimit, $Y_{\ell, m}$ may be taken to be either real or complex, and $\Phi_{\ell, s}(r)$ is some radial basis function (in practice, either spherical bessel or the radial part of the 3D PSWFs as in [Lederman]). We work here with the complex spherical harmonics, 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.

Remark 4.1. According to Nir and Joe, optimizing over \mathbb{C} allows them to escape local minima they encountered in \mathbb{R} .

Then

$$\widehat{I}_{\omega}(ck, \theta) = \sum_{\ell, m, m', s} a_{\ell, m, s} Y_{\ell, m'}(\pi/2, 0) D_{m', m}^{\ell}(\omega) \Phi_{\ell, s}(k) e^{im'\theta},$$

whenever $k \leq 1$. Express this in 2D PSWFs so

$$\widehat{I}_{\omega}(ck, \theta) = \sum_{N, n} b_{N, n} \psi_{N, n}(k, \theta),$$

where in papers the 2D PSWFs are defined as

$$\psi_{N,n}(k, \theta) = \frac{1}{\sqrt{2\pi}} R_{N,n}(k) e^{ik\theta},$$

but in Boris' code the expansion is actually performed with respect to

$$\tilde{\psi}_{N,n}(k, \theta) = \frac{1}{2\sqrt{2\pi}} \alpha_{N,n} R_{N,n}(k) e^{ik\theta},$$

where $\alpha_{N,n}$ are the eigenvalues associated with $\psi_{N,n}$ - see Sect. 5 below.

Remark 4.2. The eigenvalues $\alpha_{N,n}$ decay with increasing radial frequency n in a step-function manner (initially constant, then super-exponential decay - see Boris' paper). Additionally, although $\alpha_{N,0}$ seems to be the same for different N , the quick decay occurs earlier (smaller n) for larger N . The above normalization of $\psi_{N,n}$ seems to ensure that the expansion coefficients decay as well. From experience trying to expand a volume in 3D prolates, this normalization seems necessary (even though it is not mentioned in any of the papers I've seen) - without it coefficients for higher N blow up.

We then get

$$\begin{aligned} b_{N,n} &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \int_0^1 \hat{I}_\omega(c k, \theta) R_{N,n}(k) e^{-iN\theta} k dk d\theta, \\ &= \sum_{\ell, m, m', s} a_{\ell, m, s} [\sqrt{2\pi} Y_{\ell, m'}(\pi/2, 0)] D_{m', m}^\ell(\omega) \left(\int_0^1 \Phi_{\ell, s}(k) R_{N,n}(k) k dk \right) \left(\frac{1}{2\pi} \int_0^{2\pi} e^{i(m'-N)\theta} d\theta \right), \\ &= \sum_{\ell \geq |N|} \sum_{m, s} a_{\ell, m, s} D_{N, m}^\ell(\omega) \beta_{\ell, s; N, n}, \end{aligned}$$

where

$$\beta_{\ell, s; N, n} = \begin{cases} \sqrt{2\pi} Y_{\ell, N}(\pi/2, 0) \int_0^1 \Phi_{\ell, s}(k) R_{N,n}(k) k dk, & \ell \geq |N|, \\ 0, & \ell < |N| \end{cases},$$

can be precomputed.

Then, back in real space,

$$\begin{aligned} I_\omega(r, \varphi) &= \sum_{N, n} \hat{\alpha}_{N, n} b_{N, n} \psi_{N, n}(r, \varphi), \\ &= \sum_{\ell=0}^L \sum_{N, m=-\ell}^{\ell} \sum_{n=0}^{n_{\max}(N)} \sum_{s=1}^{S(\ell)} a_{\ell, m, s} \hat{\beta}_{\ell, s; N, n} D_{N, m}^\ell(\omega) \psi_{N, n}(r, \varphi). \end{aligned}$$

where $\alpha_{N,n}$ is the eigenvalue corresponding to the (N, n) th PSWF, $\hat{\alpha}_{N, n} = (c/2\pi)^2 \alpha_{N, n}$, and $\hat{\beta}_{\ell, s; N, n} = \hat{\alpha}_{N, n} \beta_{\ell, s; N, n}$. In this notation, we assumed I_ω has bandlimit c and is concentrated in the unit ball in real space. We then consider the product

$$\begin{aligned} m_3(\Delta \mathbf{r}_1, \Delta \mathbf{r}_2) &= \int_{\mathbf{r}} \langle I_\omega(\mathbf{r}) I_\omega(\mathbf{r} + \Delta \mathbf{r}_1) \overline{I_\omega(\mathbf{r} + \Delta \mathbf{r}_2)} \rangle_\omega, \\ &= \sum_{\substack{N_1, n_1 \\ N_2, n_2 \\ N_3, n_3}} \langle b_{N_1, n_1} b_{N_2, n_2} \overline{b_{N_3, n_3}} \rangle_\omega \int_{\mathbf{r}} \psi_{N_1, n_1}(\mathbf{r}) \psi_{N_2, n_2}(\mathbf{r} + \Delta \mathbf{r}_1) \overline{\psi_{N_3, n_3}(\mathbf{r} + \Delta \mathbf{r}_2)}. \end{aligned}$$

Now,

$$\begin{aligned} \langle b_{N_1, n_1} b_{N_2, n_2} \overline{b_{N_3, n_3}} \rangle_\omega &= \sum_{\substack{\ell_1, m_1, s_1 \\ \ell_2, m_2, s_2 \\ \ell_3, m_3, s_3}} a_{\ell_1, m_1, s_1} a_{\ell_2, m_2, s_2} \overline{a_{\ell_3, m_3, s_3}} \langle D_{N_1, m_1}^{\ell_1}(\omega) D_{N_2, m_2}^{\ell_2} \overline{D_{N_3, m_3}^{\ell_3}} \rangle_\omega \\ &\quad \times \beta_{\ell_1, s_1; N_1, n_1} \beta_{\ell_2, s_2; N_2, n_2} \overline{\beta_{\ell_3, s_3; N_3, n_3}}, \end{aligned}$$

and [Tamir's note on Kam's bispectrum]

$$\langle D_{N_1, m_1}^{\ell_1}(\omega) D_{N_2, m_2}^{\ell_2} \overline{D_{N_3, m_3}^{\ell_3}} \rangle_\omega = (-1)^{N_3 + m_3} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ N_1 & N_2 & -N_3 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_3 \end{pmatrix},$$

and since $\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = 0$ unless $m_1 + m_2 + m_3 = 0$ and $|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2$, we conclude that

$$\begin{aligned} \langle b_{N_1, n_1} b_{N_2, n_2} \overline{b_{N_3, n_3}} \rangle_\omega &= \delta_{N_3, N_1 + N_2} \sum_{\substack{\ell_1, m_1, s_1 \\ \ell_2, m_2, s_2 \\ s_3}}^{\min(L, \ell_1 + \ell_2)} \sum_{\ell_3 = |\ell_1 - \ell_2|} a_{\ell_1, m_1, s_1} a_{\ell_2, m_2, s_2} \overline{a_{\ell_3, m_1 + m_2, s_3}} \\ &\quad \times (-1)^{N_1 + N_2 + m_1 + m_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ N_1 & N_2 & -N_1 - N_2 \end{pmatrix} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_1 - m_2 \end{pmatrix} \\ &\quad \times \beta_{\ell_1, s_1; N_1, n_1} \beta_{\ell_2, s_2; N_2, n_2} \overline{\beta_{\ell_3, s_3; N_1 + N_2, n_3}}. \end{aligned}$$

Finally, we expand

$$m_3(\Delta \mathbf{r}_1, \Delta \mathbf{r}_2) = \sum_{k, q_1, q_2} \mathbf{m}_3(k, q_1, q_2) \psi_{k, q_1}(\Delta \mathbf{r}_1) \overline{\psi_{k, q_2}(\Delta \mathbf{r}_2)},$$

where we only include the block-diagonal terms in the expansion. Defining

$$\Psi_{\ell, N, s}(\mathbf{r}) = \sum_{n=0}^{n_{\max}(N)} \beta_{\ell, s; N, n} \psi_{N, n}(\mathbf{r}),$$

the final formula reads

$$\begin{aligned} \mathbf{m}_3(k, q_1, q_2) &= \sum_{\substack{\ell_1, m_1, s_1 \\ \ell_2, m_2, s_2 \\ s_3}}^{\min(L, \ell_1 + \ell_2)} \sum_{\ell_3 = |\ell_1 - \ell_2|} a_{\ell_1, m_1, s_1} a_{\ell_2, m_2, s_2} \overline{a_{\ell_3, m_1 + m_2, s_3}} \\ &\quad \times (-1)^{m_1 + m_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_1 - m_2 \end{pmatrix} \\ &\quad \times \sum_{N_1 = -\ell_1}^{\ell_1} \sum_{N_2 = -\ell_2}^{\ell_2} (-1)^{N_1 + N_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ N_1 & N_2 & -N_1 - N_2 \end{pmatrix} \int_{\mathbf{r}} \Psi_{\ell_1, N_1, s_1}(\mathbf{r}) \rho_{\ell_2, N_2, s_2}^{(k, q_1)}(\mathbf{r}) \overline{\rho_{\ell_3, N_1 + N_2, s_3}^{(k, q_2)}(\mathbf{r})}, \end{aligned}$$

where

$$\rho_{\ell, N, s}^{(k, q)} = \int_{\Delta \mathbf{r}} \Psi_{\ell, N, s}(\mathbf{r} + \Delta \mathbf{r}) \overline{\psi_{k, q}(\Delta \mathbf{r})}.$$

In practice, the last line of the above expression for $\mathbf{m}_3(k, q_1, q_2)$ is precomputed, and both the integration over \mathbf{r} and over $\Delta \mathbf{r}$ is performed on the grid of the images in the dataset, to match the integration performed on the actual images.

4.1. Implementation. To implement the above formula, we precompute the quantities

$$W(\ell_1, \ell_2, \ell_3, m_1, m_2) = (-1)^{m_1+m_2} \begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_1 - m_2 \end{pmatrix},$$

and

$$B^{(k, q_1, q_2)}(\ell_1, \ell_2, \ell_3, s_1, s_2, s_3) = \sum_{N_1=-\ell_1}^{\ell_1} \sum_{N_2=-\ell_2}^{\ell_2} W(\ell_1, \ell_2, \ell_3, N_1, N_2) \int_{\mathbf{r}} \Psi_{\ell_1, N_1, s_1}(\mathbf{r}) \rho_{\ell_2, N_2, s_2}^{(k, q_1)}(\mathbf{r}) \overline{\rho_{\ell_3, N_1+N_2, s_3}^{(k, q_2)}(\mathbf{r})},$$

so in each iteration of the optimization we compute

$$\mathbf{m}_3(k, q_1, q_2) = \sum_{\substack{\ell_1, \ell_2, \ell_3 \\ m_1, m_2 \\ s_1, s_2, s_3}} W(\ell_1, \ell_2, \ell_3, m_1, m_2) B^{(k, q_1, q_2)}(\ell_1, \ell_2, \ell_3, s_1, s_2, s_3) \times \\ a_{\ell_1, m_1, s_1} a_{\ell_2, m_2, s_2} \overline{a_{\ell_3, m_1+m_2, s_3}}.$$

The Wigner 3j symbols are computed from the Racah formula

$$\begin{pmatrix} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} = \delta_{m_3, m_1+m_2} (-1)^{\ell_1-\ell_2+m_3} \sqrt{\Delta(\ell_1, \ell_2, \ell_3)} \\ \times \prod_{i=1}^3 \sqrt{(\ell_i - m_i)!(\ell_i + m_i)!} \times \sum_t \frac{(-1)^t}{x(t)},$$

where

$$x(t) = t!(\ell_3 - \ell_2 + t + m_1)!(\ell_3 - \ell_1 + t - m_2)!(\ell_1 + \ell_2 - \ell_3 + t)!(\ell_1 - t - m_1)!(\ell_2 - t + m_2)!,$$

the sum is over all integer t for which the arguments in the factorials in $x(t)$ are all nonnegative, and

$$\Delta(\ell_1, \ell_2, \ell_3) = \frac{(\ell_1 + \ell_2 - \ell_3)!(\ell_1 - \ell_2 + \ell_3)!(-\ell_1 + \ell_2 + \ell_3)!}{(\ell_1 + \ell_2 + \ell_3 + 1)!}.$$

In practice, the sum in the above expression is computed in infinite-precision arithmetic (in terms of `sym` objects in Matlab - compare for instance `factorial(171)` and `factorial(sym(171))`). All other terms are computed in standard double-precision.

4.2. The power spectrum. The power spectrum is easier to derive directly in Fourier space, to avoid integration of shifted prolates against centered ones. The average power spectrum in Fourier space can be derived from Kam's original formula [Kam, 1980; Eq. 10] by setting $\mathbf{k}_1 = \mathbf{k}_2$ to obtain

$$\langle |\hat{I}_\omega(k, \theta)|^2 \rangle_\omega = \frac{1}{4\pi} \sum_{\ell, m} \left| \sum_s a_{\ell, m, s} j_{\ell, s}(k) \right|^2 = \frac{1}{4\pi} \sum_{\substack{\ell, m \\ s_1, s_2}} a_{\ell, m, s_1} \overline{a_{\ell, m, s_2}} j_{\ell, s_1}(k) j_{\ell, s_2}(k),$$

where we used the fact that the normalized spherical Bessel functions $j_{\ell, s}$ are real. To expand the above in 2D PSWFs, we write

$$\langle |\hat{I}_\omega(k, \theta)|^2 \rangle_\omega = \sum_q \mathbf{m}_2(q) \psi_{0, q}(k),$$

and conclude that

$$\mathbf{m}_2(q) = \frac{\sqrt{2\pi}}{4\pi} \sum_{\substack{\ell, m \\ s_1, s_2}} a_{\ell, m, s_1} \overline{a_{\ell, m, s_2}} \int_0^1 j_{\ell, s_1}(k) j_{\ell, s_2}(k) R_{0, q}(k) k dk.$$

4.3. The mean. Since the Fourier transformed volume is given by

$$\widehat{V}(k, \theta, \varphi) = \sum_{\ell, m, s} a_{\ell, m, s} Y_{\ell, m}(\theta, \varphi) j_{\ell, s}(k),$$

since $j_{\ell, s}(0) = 0$ unless $\ell = 0$, and since $Y_{0,0}(\theta, \varphi) = \frac{1}{\sqrt{4\pi}}$, we conclude that

$$m_1[V] = \widehat{V}(\mathbf{0}) = \frac{1}{\sqrt{4\pi}} \sum_s a_{0,0,s} j_{0,s}(0).$$

5. PROPERTIES OF PSWFs

The d -dimensional PSWFs $\psi_{N,n,m}(\mathbf{r})$ are eigenfunctions of the truncated Fourier transform, so

$$\alpha_{N,n,m} \psi_{N,n,m}(\mathbf{r}) = \int_{\mathbb{D}} \psi_{N,n,m}(\boldsymbol{\omega}) e^{i c \boldsymbol{\omega} \cdot \mathbf{r}} d\boldsymbol{\omega},$$

where $\mathbb{D} = \{\mathbf{r} \in \mathbb{R}^d : \|\mathbf{r}\|_2 \leq 1\}$ is the unit ball, and c is the assumed bandlimit, and the Fourier transform is unfortunately written in terms of angular frequency $\boldsymbol{\omega}$ rather than the regular frequency denoted in previous sections by \mathbf{k} (we follow this convention here because the code does). Furthermore, $\psi_{N,n,m}$ form an orthonormal basis of both $L^2(\mathbb{D})$ and $L^2(\mathbb{R}^d)$. For a c -bandlimited function f with Fourier transform \widehat{f} , we can expand

$$\widehat{f}(c\boldsymbol{\omega}) = \sum_{N,n,m} a_{N,n,m} \psi_{N,n,m}(\boldsymbol{\omega}),$$

so

$$\begin{aligned} f(\mathbf{r}) &= \frac{1}{(2\pi)^d} \int_{c\mathbb{D}} \widehat{f}(\boldsymbol{\omega}) e^{i \boldsymbol{\omega} \cdot \mathbf{r}} d\boldsymbol{\omega} = \left(\frac{c}{2\pi}\right)^d \int_{\mathbb{D}} \widehat{f}(c\boldsymbol{\omega}) e^{i c \boldsymbol{\omega} \cdot \mathbf{r}} d\boldsymbol{\omega} \\ &= \left(\frac{c}{2\pi}\right)^d \sum_{N,n,m} a_{N,n,m} \int_{\mathbb{D}} \psi_{N,n,m}(\boldsymbol{\omega}) e^{i c \boldsymbol{\omega} \cdot \mathbf{r}} d\boldsymbol{\omega} = \left(\frac{c}{2\pi}\right)^d \sum_{N,n,m} \alpha_{N,n,m} a_{N,n,m} \psi_{N,n,m}(\mathbf{r}), \end{aligned}$$

so the expansion coefficients of $f(\mathbf{r})$ in $\{\psi_{N,n,m}\}$ are $\left(\frac{c}{2\pi}\right)^d \alpha_{N,n,m} a_{N,n,m}$. In polar form, they have a separable expression

$$\psi_{N,n,m}(\mathbf{r}) = R_{N,n}(r) S_{N,n,m}(\mathbf{r}/r),$$

where $S_{N,n,m}(\mathbf{r}/r)$ are the normalized spherical harmonics on \mathbb{S}^{d-1} , the unit sphere in \mathbb{R}^d . We have the following orthogonality relations:

$$\int_{\mathbb{S}^{d-1}} S_{N,n,m}(\mathbf{r}/r) S_{N',n',m'}(\mathbf{r}/r) = \delta_{N,N'} \delta_{m,m'}, \quad \int_0^1 R_{N,n}(r) R_{N',n'}(r) r dr = \delta_{n,n'}.$$

Note in particular that $\{R_{N,n}\}$ for different values of the angular frequency N are not necessarily orthogonal.

Remark 5.1. Off-center slices of m_3 are not centered, but are also smaller in magnitude than the centered, central slices. We therefore expect to have a bigger error between the expansion in prolates and the actual moment off-center. This may be fixed by predicting the maximum of the off-center slices (say, under the assumption of a spherical object) and translating the prolates in the expansion by that amount. However, this would introduce a dependence between the two basis functions in the expansion, making all the formulas harder to compute - the resulting product basis is likely not orthogonal.

Remark 5.2. In the code, we assume the images are sampled on the unit square with samples $\{(\frac{j}{L}, \frac{k}{L}) : j, k = -L, \dots, L\}$, so the image has size $(2L+1) \times (2L+1)$ and bandlimit $c = \pi L$. In particular, this implies that the assumed bandlimit for the micrograph patches is twice as large as that assumed for the volume and its projections. Assuming a larger bandlimit for the latter leads to bad approximation errors:

- Expanding the volume with a larger bandlimit, and using the resulting coefficients to generate projections from non-trivial directions results in nonsense.
- Using the larger bandlimit in the formula for $\mathbf{m}_3(k, q_1, q_2)$ (or equivalently, assuming such a bandlimit for the projections) results in nonsense.

6. DEBIASING THE MOMENTS

Key: The derivation is exactly the same as in the 1D MRA case, only that instead of factoring out a common signal from an average of noise as in the current draft (which we can't do here as we have "continuous heterogeneity"), we consider the viewing directions as instances of a random variable independent of the noise, and use the law of total expectation.

6.1. Stats for Dummies. Claim: Suppose $\xi_i \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$, and x_i is independently drawn from some distribution with finite first and second moments. Then

$$\frac{1}{N} \sum_{i=1}^N \xi_i^2 x_i \xrightarrow{\text{as}} \sigma^2 \mathbb{E}[x].$$

Proof: Let $X_\xi = \frac{1}{N} \sum_{i=1}^N \xi_i^2 x_i$ be the random variable in question. Then $\mathbb{E}_\xi[X_\xi] = \frac{\sigma^2}{N} \sum_{i=1}^N x_i$ trivially, and

$$\text{var}(X_\xi) = \mathbb{E}_\xi \left[\left(\frac{1}{N} \sum_{i=1}^N (\xi_i^2 - \sigma^2) x_i \right)^2 \right] = \frac{1}{N^2} \sum_{i,j=1}^N \mathbb{E}_\xi [(\xi_i^2 - \sigma^2)(\xi_j^2 - \sigma^2)] x_i x_j = \frac{2\sigma^2}{N} \left(\frac{1}{N} \sum_{i=1}^N x_i^2 \right),$$

as only terms with $i = j$ give nonzero expectation, so $\text{var}(X_\xi) \rightarrow 0$ as $N \rightarrow \infty$.

Claim: Under the above assumptions,

$$\frac{1}{N} \sum_{i=1}^N \xi_i x_i \xrightarrow{\text{as}} 0.$$

Proof: Let $Y_\xi = \frac{1}{N} \sum_{i=1}^N \xi_i x_i$. Then $\mathbb{E}_\xi[Y_\xi] = 0$ and

$$\text{var}(Y_\xi) = \frac{1}{N^2} \sum_{i,j=1}^N \mathbb{E}_\xi [\xi_i \xi_j] x_i x_j = \frac{\sigma^2}{N} \left(\frac{1}{N} \sum_{i=1}^N x_i^2 \right).$$

Claim: Under the above assumptions,

$$m_3[\xi](\Delta i_1, \Delta i_2) = \frac{1}{N} \sum_{i=1}^N \xi_i \xi_{i+\Delta i_1} \xi_{i+\Delta i_2} \rightarrow 0,$$

for all $\Delta i_1, \Delta i_2$, where it is understood that $\xi_i = 0$ for $i \notin [1, N]$.

Proof: We have

$$m_3[\xi](\Delta i_1, \Delta i_2) \rightarrow \mathbb{E}[\xi_i \xi_{i+\Delta i_1} \xi_{i+\Delta i_2}],$$

so if either $\Delta i_1 \neq \Delta i_2$ or $\Delta i_1 = \Delta i_2 \neq 0$, the expectation vanishes as ξ_i are independent, whereas if $\Delta i_1 = \Delta i_2 = 0$ then the expectation vanishes because the third moment of a Gaussian is zero.

6.2. Noisy micrograph moments. We now suppose that the micrograph is perturbed by additive white Gaussian noise, so we observe $\tilde{\mathcal{I}} = \mathcal{I} + \xi$ where $\xi \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2 I)$. We proceed to derive $\lim_{N \rightarrow \infty} m_2[\tilde{\mathcal{I}}], m_3[\tilde{\mathcal{I}}]$. For simplicity of notation, we shall use vectorized indices $\mathbf{i} = (i, j)$.

For the power spectrum:

$$\begin{aligned} m_2[\mathcal{I} + \xi](\Delta \mathbf{i}) &= \frac{1}{M^2} \sum_{i,j=1}^M \tilde{\mathcal{I}}(\mathbf{i}) \tilde{\mathcal{I}}(\mathbf{i} + \Delta \mathbf{i}) \\ &= \frac{1}{M^2} \sum_{i,j=1}^M \mathcal{I}(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}) + \frac{1}{M^2} \sum_{i,j=1}^M \mathcal{I}(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}) \\ &\quad + \frac{1}{M^2} \sum_{i,j=1}^M \xi(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}) + \frac{1}{M^2} \sum_{i,j=1}^M \xi(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}). \end{aligned}$$

Considering the terms one-by-one, the first term is independent of the noise, and as shown above converges to $\gamma \langle m_2[I_\omega] \rangle_\omega$ as $N \rightarrow \infty$. Denoting the center of the instance of I_ω in \mathcal{I} by \mathbf{s}_ω , the second term satisfies

$$\begin{aligned} \frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}) &= \frac{NL^2}{M^2} \cdot \frac{1}{NL^2} \sum_{\omega, \mathbf{i}} I_\omega(\mathbf{i}) \xi_\omega(\mathbf{i} + \Delta \mathbf{i}) \\ &\rightarrow \gamma \mathbb{E}[\xi] \mathbb{E}[I_\omega] = 0, \end{aligned}$$

where $\xi_\omega(\mathbf{i}) = \xi(\mathbf{i} + \mathbf{s}_\omega)$ and $\mathbb{E}[I_\omega]$ is proportional to the mean of the volume. A similar argument applied to the third term shows that it also vanishes as $N \rightarrow \infty$. For the fourth term, if $\Delta \mathbf{i} \neq \mathbf{0}$ then since the noise is zero mean and i.i.d. this term vanishes. If $\Delta \mathbf{i} = \mathbf{0}$ then

$$\frac{1}{m^2} \sum_{i,j=1}^m \xi(\mathbf{i})^2 \rightarrow \sigma^2.$$

Thus, we conclude

$$m_2[\mathcal{I} + \xi](\Delta \mathbf{i}) \rightarrow \gamma \langle m_2[I_\omega](\Delta \mathbf{i}) \rangle_\omega + \sigma^2 \delta(\Delta \mathbf{i}).$$

For the third moments, we get 8 terms:

$$\begin{aligned}
m_3[\mathcal{I} + \xi](\Delta \mathbf{i}_1, \Delta \mathbf{i}_2) = & \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_1) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_2)}_{(1)} + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \xi(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}_1) \xi(\mathbf{i} + \Delta \mathbf{i}_2)}_{(2)} \\
& + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}_1) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_2)}_{(3)} + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_1) \xi(\mathbf{i} + \Delta \mathbf{i}_2)}_{(4)} \\
& + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \xi(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_1) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_2)}_{(5)} + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}_1) \xi(\mathbf{i} + \Delta \mathbf{i}_2)}_{(6)} \\
& + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \xi(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}_1) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_2)}_{(7)} + \underbrace{\frac{1}{M^2} \sum_{\mathbf{i}} \xi(\mathbf{i}) \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_1) \xi(\mathbf{i} + \Delta \mathbf{i}_2)}_{(8)}.
\end{aligned}$$

We address these terms one by one:

- Term (1) is $m_3[\mathcal{I}]$, shown above to converge to $\gamma \langle m_3[I_\omega] \rangle_\omega$.
- Term (2) is $m_3[\xi]$, the bispectrum of pure noise, which vanishes as shown above.
- Terms (3)-(5) depend linearly on the noise and hence converge to zero.
- For term (6), if $\Delta \mathbf{i}_1 \neq \Delta \mathbf{i}_2$ the term vanishes as then $\xi(\mathbf{i} + \Delta \mathbf{i}_1)$ and $\xi(\mathbf{i} + \Delta \mathbf{i}_2)$ are independent. If $\Delta \mathbf{i}_1 = \Delta \mathbf{i}_2$ the term becomes

$$\frac{1}{M^2} \sum_{\mathbf{i}} \mathcal{I}(\mathbf{i}) \xi(\mathbf{i} + \Delta \mathbf{i}_1)^2 = \frac{NL^2}{M^2} \cdot \frac{1}{NL^2} \sum_{\omega, \mathbf{i}} I_\omega(\mathbf{i}) \xi_\omega(\mathbf{i} + \Delta \mathbf{i})^2 \rightarrow \gamma \sigma^2 \mathbb{E}[I_\omega] = \gamma \sigma^2 L m_1[V],$$

where $m_1[V]$ is the mean of the volume (and hence also the mean of each projection I_ω).

- For term (7), if $\Delta \mathbf{i}_1 \neq \mathbf{0}$ then once again this term vanishes, whereas if $\Delta \mathbf{i}_2 = \mathbf{0}$ then it becomes

$$\frac{1}{M^2} \sum_{\mathbf{i}} \xi(\mathbf{i})^2 \mathcal{I}(\mathbf{i} + \Delta \mathbf{i}_2) = \frac{NL^2}{M^2} \cdot \frac{1}{NL^2} \sum_{\omega, \mathbf{i}} \xi_\omega(\mathbf{i})^2 I_\omega(\mathbf{i} + \Delta \mathbf{i}_2) \rightarrow \gamma \sigma^2 L m_1[V].$$

Similarly, term (8) vanishes if $\Delta \mathbf{i}_2 \neq \mathbf{0}$ and converges to $\gamma \sigma^2 L m_1[V]$ otherwise.

Thus, we conclude that

$$m_3[\mathcal{I} + \xi](\Delta \mathbf{i}_1, \Delta \mathbf{i}_2) \rightarrow \gamma \langle m_3[I_\omega](\Delta \mathbf{i}_1, \Delta \mathbf{i}_2) \rangle_\omega + \gamma \sigma^2 L m_1[V] \left(\delta(\Delta \mathbf{i}_1 - \Delta \mathbf{i}_2) + \delta(\Delta \mathbf{i}_1) + \delta(\Delta \mathbf{i}_2) \right).$$

Note that in practice, we have $m_1[\mathcal{I} + \xi] \rightarrow m_1[\mathcal{I}] \approx \gamma L m_1[V]$ since the noise has zero mean, so we do not need prior knowledge of γ to effectively debias the bispectrum.

6.3. Expansion in PSWFs. In practice, we compute the moments of our noisy micrograph in a product basis of 2D prolates, so we need to derive the expansion coefficients of the above bias terms in these functions for debiasing.

To this end, writing (in the continuous case again)

$$\delta(\Delta \mathbf{r}_1 - \Delta \mathbf{r}_2) = \sum_{k, q_1, q_2} \mathfrak{d}(k, q_1, q_2) \psi_{k, q_1}(\Delta \mathbf{r}_1) \overline{\psi_{k, q_2}(\Delta \mathbf{r}_2)},$$

we get

$$\begin{aligned} \mathfrak{d}(k, q_1, q_2) &= \int_{\Delta \mathbf{r}_1, \Delta \mathbf{r}_2} \delta(\Delta \mathbf{r}_1 - \Delta \mathbf{r}_2) \overline{\psi_{k, q_1}(\Delta \mathbf{r}_1)} \psi_{k, q_2}(\Delta \mathbf{r}_2) \\ &= \int_{\Delta \mathbf{r}_2} \overline{\psi_{k, q_1}(\Delta \mathbf{r}_2)} \psi_{k, q_2}(\Delta \mathbf{r}_2) \\ &= \delta_{q_1, q_2}. \end{aligned}$$

Similarly, writing

$$\delta(\Delta \mathbf{r}_1) = \sum_{k, q_1, q_2} \mathfrak{d}^{(1)}(k, q_1, q_2) \psi_{k, q_1}(\Delta \mathbf{r}_1) \overline{\psi_{k, q_2}(\Delta \mathbf{r}_2)},$$

we get

$$\begin{aligned} \mathfrak{d}^{(1)}(k, q_1, q_2) &= \int_{\Delta \mathbf{r}_1, \Delta \mathbf{r}_2} \delta(\Delta \mathbf{r}_1) \overline{\psi_{k, q_1}(\Delta \mathbf{r}_1)} \psi_{k, q_2}(\Delta \mathbf{r}_2) \\ &= \overline{\psi_{k, q_1}(\mathbf{0})} \int_{\Delta \mathbf{r}_2} \psi_{k, q_2}(\Delta \mathbf{r}_2) \\ &= \delta_{k, 0} R_{0, q_1}(0) \int_0^1 R_{0, q_2}(r) r dr \\ &= \delta_{k, 0} \frac{\alpha_{0, q_2}}{2\pi} R_{0, q_1}(0) R_{0, q_2}(0), \end{aligned}$$

where we used the fact that $R_{0, q}$ satisfies

$$\frac{\alpha_{0, q}}{2\pi} R_{0, q}(r) = \int_0^1 R_{0, q}(\rho) J_0(c r \rho) \rho d\rho,$$

where J_k is the Bessel function of the first kind, and that $J_0(0) = 1$. A similar derivation applies to $\delta(\Delta \mathbf{r}_2)$.

Thus, in terms of the prolate expansion coefficients, our bias formulas become

$$\tilde{\mathfrak{m}}_3(k, q_1, q_2) = \mathfrak{m}_3(k, q_1, q_2) + \sigma^2 m_1[\tilde{\mathcal{I}}] \left[\delta_{q_1, q_2} + \delta_{k, 0} \frac{1}{2\pi} (\alpha_{0, q_1} + \alpha_{0, q_2}) R_{0, q_1}(0) R_{0, q_2}(0) \right].$$

Similarly, for the power spectrum we get

$$\tilde{\mathfrak{m}}_2(q) = \mathfrak{m}_2(q) + \sigma^2 \frac{1}{\sqrt{2\pi}} R_{0, q}(0).$$