

## Steerable Near-Quadrature Filter Pairs in Three Dimensions\*

Tommy M. Tang<sup>†</sup> and Hemant D. Tagare<sup>‡</sup>

**Abstract.** Steerable filter pairs that are near quadrature have many image processing applications. This paper proposes a new methodology for designing such filters. The key idea is to design steerable filters by minimizing a departure-from-quadrature function. These minimizing filter pairs are almost exactly in quadrature. The polar part of the filters is nonnegative, monotonic, and highly focused around an axis, and asymptotically the filters achieve exact quadrature. These results are established by exploiting a relation between the filters and generalized Hilbert matrices. These near-quadrature filters closely approximate three dimensional Gabor filters. We experimentally verify the asymptotic mathematical results and further demonstrate the use of these filter pairs by efficient calculation of local Fourier shell correlation of cryogenic electron microscopy.

**Key words.** steerable filters, quadrature, filters, quadratic programming, Hilbert matrices

**AMS subject classifications.** 68U10, 90C20

**DOI.** 10.1137/21M143529X

**1. Introduction.** Many image and video processing applications benefit from evaluating local directional properties of images. For example, local spatio-temporal directional change is useful in understanding motion in video sequences [6]. Directional space-frequency analysis is also useful in understanding texture and other features of images [10].

Local directional properties of an image can be computationally difficult to calculate, because they have to be calculated at every pixel and for every direction. Steerable filters were introduced in [16] to simplify this calculation. Loosely speaking, a steerable filter is a directional (i.e., an anisotropic) filter with the property that it can be oriented along any direction by simply taking a linear combination of a *finite set* of its orientations (the precise definition is given in section 2). Thus, by first convolving the image with the finite set of filters, local properties of an image along *any* additional direction can be calculated simply by taking a linear combination of the finite convolutions. Computational complexity is reduced because a smaller number of convolutions are involved [16, 17, 29].

Besides steerability, other properties are also desirable in filters used for local directional analysis. Ideally, a quadrature filter pair is required for accurate representation. The filters should also be directionally focused for high orientation sensitivity; a narrower filter will yield

\*Received by the editors July 27, 2021; accepted for publication (in revised form) January 27, 2022; published electronically May 26, 2022.

<https://doi.org/10.1137/21M143529X>

**Funding:** The work of the authors was supported by NIH Grant R01GM125769.

<sup>†</sup>Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT USA ([tommytang@gmail.com](mailto:tommytang@gmail.com)).

<sup>‡</sup>Departments of Radiology and Biomedical Engineering, Electrical Engineering, Biomedical Engineering, and Statistics and Data Science, Yale University New Haven, CT, USA ([hermant.tagare@yale.edu](mailto:hermant.tagare@yale.edu), <https://noodle.med.yale.edu/hdtag/profile.html>).

a higher angular resolution analysis of orientation [15]. And the filter should have a small space-frequency uncertainty for localization [9]. As we will show, all of these properties cannot be obtained simultaneously, and this raises the question of whether some of the properties can be obtained exactly, while others are obtained approximately.

In this paper, we approach this problem by designing directional, steerable filter pairs that are near-quadrature. The main idea behind our approach is to explicitly minimize a departure-from-quadrature objective function over a set of steerable filters. The results of this minimization are surprising. It turns out that in three dimensions, there are steerable filters that are almost exactly in quadrature. Moreover, as the complexity of these filters increases, they converge to exact quadrature, while the angular width of the filters converges to zero. These filters also have low space-frequency uncertainty. Thus these filters appear to be well-behaved in terms of all four properties mentioned above.

Below, we establish several properties of these filters mathematically, using the Karush–Kuhn–Tucker (KKT) conditions of the departure-from-quadrature minimization problem, and by exploiting a close relation between the minimizers and generalized Hilbert matrices.

An important application of steerable filters is the approximation of Gabor filters. Gabor filters achieve minimum space-frequency uncertainty and are commonly used in texture and motion analysis [9]. However, Gabor filters are not steerable. In section 6.1, we show that the steerable filters we derive can closely approximate Gabor filters. And because these filters are steerable, they can be used to approximate Gabor filters oriented in any direction. Thus they may be used in many of the myriad imaging tasks in which Gabor filters are desirable, with the advantage that local properties may be calculated with decreased computational complexity.

A significant use of approximating Gabor filters by steerable filters is the analysis of local frequency content in cryogenic electron microscopy (cryo-EM) [37]. Cryo-EM is a technique for reconstructing the three dimensional shape of biological macromolecules, usually proteins or protein complexes [14]. In section 6.2, we show how our steerable filters can be used in cryo-EM. Cryo-EM and various terms associated with it are explained in detail in section 6.2.

**1.1. Organization of the paper.** The rest of the paper is organized as follows: A literature review is provided immediately below in section 1.2. Section 2 introduces steerable filters and related terminology. Section 3 formulates and solves the design of near-quadrature filters. section 4 establishes asymptotic properties of the solution. Section 5 contains results of numerical experiments confirming these properties. Section 6 contains applications of the steerable filters: In Section 6.1 we show how steerable filters can approximate Gabor filters. Section 6.2 applies this result to cryo-EM. Section 7 contains comments and a brief discussion. Section 8 concludes the paper. The proofs of the theorems in section 4 are relegated to the appendix.

**1.2. Literature review.** As mentioned above, axially symmetric steerable filters were introduced in [16], which defined steerability, established basic properties of steerable filters, and demonstrated their use in measuring orientation, angularly adaptive filtering, and contour detection. Steerable filters efficiently compute local properties of signals and functions [17, 29]. Steerable 3 dimensional (3D) wavelets are proposed in [7] based on the Reisz transform. These wavelets form a frame, but are not highly directional.

There is a close connection between steerable filters and representations of Lie groups  $SO(2)$  and  $SO(3)$  [22, 35]. Representation theory is used in [35] to describe the relationship between the steering angles and the method of producing steering coefficients with which to steer the basis. Formulas connecting the representations of the group with steering coefficients are also found in [22]. The formulas in [22] use properties of spherical harmonics.

3D steerable filters are used in a number of applications. For example, texture (such as hair, or certain kinds of biological tissue) and other image features often have directional orientation in an image. Finding the orientation of these features can be made computationally efficient with steerable filters [19]. Examples of the use of steerable filters for this purpose can be found in biomedical image analysis in 2 dimensions, e.g., [10], as well as in 3 dimensions, e.g. [1, 5].

3D steerable filters are also used for motion analysis in video sequences [6, 20, 40, 41]. Here the idea is to stack the sequence of 2 dimensions (2D) video frames as a 3D image, and then to determine the motion using the orientation of a 3D steerable filter. In another application to video analysis, 3D steerable filters are used to colorize video [28].

Steerable filters are used to analyze the quality of edges in [36] for calculating the 3D shape-from-focus from a video sequence. As an object moves close to the focal plane of a lens, its edges become sharper in the 2D image that it forms. Thus the quality of edges, and changes thereof, provide monocular clues to the 3D depth of an object. Steerable 3D filters are used in the  $2D \times \text{time}$  stack to analyze how the quality of edges changes through the video sequence [36].

More recently, steerable functions have been used in convolutional neural networks (CNN). Classical CNNs use convolutions in the first layer, and hence are translationally equivariant (translation of the data causes translation of the output of the first layer). Additional 3D rotational equivariance is obtained for the CNNs if the convolution kernels are obtained as linear combination of steerable functions [39]. Equivariant CNNs [23] are also used for texture classification [2]. CNNs that use 3D steerable filters can also be arranged to provide 3D rotational invariance [3]. As mentioned above, steerability is closely related to group representations. It has been argued that group representation methods, which make machine learning models equivariant, are critical for applications in physical sciences, where symmetry plays a key role [33].

Note that none of the 3D steerable filters mentioned above are near-quadrature; in fact, many are not even filter pairs (in the sense of quadrature).

Gabor filters achieve the lower bound of the space-frequency uncertainty described by the Heisenberg uncertainty principle and are therefore considered optimal for use in many imaging tasks [9]. However, Gabor filters cannot be steered exactly [21] and several approximations have been proposed. These include approximately steering a Gabor filter [21] or constructing approximate Gabor filters or approximate Hilbert pairs that are exactly steerable [16, 21]. These methods are mostly focused on 2, rather than 3 dimensions.

**2. Background.** This section defines the technical terms used in the paper.

**2.1. Axial symmetry.** We begin by defining radial-polar separable functions, and axially symmetric functions in  $R^3$ .

**Definition 2.1.** A function  $f : R^3 \rightarrow R$  is radial-polar separable if  $f(x, y, z) = w(r)\phi(u, v, w)$  for  $(x, y, z) \neq 0$ , where  $r = \sqrt{x^2 + y^2 + z^2}$  and  $(u, v, w) = (x/r, y/r, z/r)$ . The radial part of the function is  $w$ , and the polar part of the function is  $\phi$ .

The arguments of the polar part  $\phi$  are points  $(u, v, w)$  which lie on the unit sphere. That is, the polar part is a function from the unit sphere to the real line  $\phi : S^2 \rightarrow R$ .

Let  $\langle u, v \rangle$  denote the usual inner product of vectors  $u, v \in R^3$ .

**Definition 2.2.** A radial-polar function is axially symmetric with respect to a unit-length axis  $a = (\alpha, \beta, \gamma)$  if the polar part can be written as

$$(2.1) \quad \phi(u, v, w) = p(\langle (u, v, w), a \rangle),$$

where  $(u, v, w) \in S^2$  and  $p : [-1, 1] \rightarrow R$  is a real-valued function.

By convention, we always write the axis of an axially symmetric function as a unit vector. Below, we explicitly denote this axis as a superscript of the function. That is, we write an axially symmetric function  $f$  with the axis  $a$  as  $f^a$ .

An axially symmetric function can be rotated so that its axis of symmetry is any axis we desire. We adopt a canonical way of describing such rotated functions: we define the original function using the positive z-axis as its axis, and then replace this axis by an axis  $a$  for a rotated copy of the function. Thus the original function is specified as  $f^{(0,0,1)}$ :

$$(2.2) \quad \begin{aligned} f^{(0,0,1)}(x, y, z) &= w(r)p(\langle (x/r, y/r, z/r), (0, 0, 1) \rangle) \\ &= w(r)p(z/r). \end{aligned}$$

The rotated version of  $f^{(0,0,1)}$  is  $f^a$ . Note that the polar part  $p(z/r)$  of  $f^{(0,0,1)}$  is a function of the z-coordinate of points on a unit sphere. This coordinate take values in  $[-1, 1]$ .

## 2.2. Steerable axially symmetric functions.

**Definition 2.3.** An axially symmetric function  $f$  is steerable if there exists a finite set of axis  $a_1, \dots, a_M \in R^3$  such that for any axis  $a \in R^3$

$$(2.3) \quad f^a = \sum_{j=1}^M k_j(a) f^{a_j},$$

where  $k_j$  are real-valued functions of  $a$ .

In the above definition, the functions  $k_j$  are *steering functions*, and their values  $k_j(a)$  are *steering coefficients*. The functions  $f^{a_j}$  are the *steering basis*.

How can we find steerable axially symmetric functions? Theorem 2.4 below, established in [11, 16], gives a recipe. The theorem uses the finite set of axes  $a_1, \dots, a_M$ , where the coordinates of the axis  $a_j$  are  $a_j = (\alpha_j, \beta_j, \gamma_j)$ . The theorem requires the  $p$  function in (2.2) to be a polynomial.

**Theorem 2.4.** For  $f^{(0,0,1)}$  as defined in (2.2), let  $p : [-1, 1] \rightarrow R$  be an odd or even polynomial of  $z/r$  of order  $N$ . Then  $f^{(0,0,1)}$  is steerable, i.e.,

$$(2.4) \quad f^a = \sum_{j=1}^M k_j(a) f^{a_j}$$

for  $a = (\alpha, \beta, \gamma)$  and axes  $a_1 = (\alpha_1, \beta_1, \gamma_1), \dots, a_M = (\alpha_M, \beta_M, \gamma_M)$  if and only if  $M \geq \binom{N+2}{2}$ , and the  $k_j$  satisfy:

$$(2.5) \quad \begin{pmatrix} \alpha^N \\ \alpha^{N-1}\beta \\ \alpha^{N-1}\gamma \\ \alpha^{N-2}\beta^2 \\ \alpha^{N-2}\beta\gamma \\ \vdots \\ \gamma^N \end{pmatrix} = \begin{pmatrix} \alpha_1^N & \alpha_2^N & \dots & \alpha_M^N \\ \alpha_1^{N-1}\beta_1 & \alpha_2^{N-1}\beta_2 & \dots & \alpha_M^{N-1}\beta_M \\ \alpha_1^{N-1}\gamma_1 & \alpha_2^{N-1}\gamma_2 & \dots & \alpha_M^{N-1}\gamma_M \\ \alpha_1^{N-2}\beta_1^2 & \alpha_2^{N-2}\beta_2^2 & \dots & \alpha_M^{N-2}\beta_M^2 \\ \alpha_1^{N-2}\beta_1\gamma_1 & \alpha_2^{N-2}\beta_2\gamma_2 & \dots & \alpha_M^{N-2}\beta_M\gamma_M \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_1^N & \gamma_2^N & \dots & \gamma_M^N \end{pmatrix} \begin{pmatrix} k_1(a) \\ k_2(a) \\ k_3(a) \\ \vdots \\ k_M(a) \end{pmatrix}.$$

In other words, the recipe for creating a steerable axially symmetric function is the following: Choose an odd or even  $N$ th order polynomial  $p$  and a radial function  $w$  to define  $f^{(0,0,1)}$ . Also choose a fixed set of axes  $a_1, \dots, a_M$ . To steer  $f^{(0,0,1)}$  so that its axis is  $a$ , solve (2.5) to get the steering coefficients  $k_j(a)$ . Then, use the linear sum in (2.4).

Although the above definitions allow the function  $p$  to be a constant or a zero function, the interesting cases are where  $p$  is neither.

**2.3. From functions to filters.** In image processing, when an image is convolved with a function, the function is called a *filter*, and its Fourier transform is called the *frequency response of the filter*. To distinguish a function from its Fourier transform, we will denote the former by an ordinary font and the latter by a calligraphic font. Thus the function  $V : R^3 \rightarrow R$  has a Fourier transform  $\mathcal{V} : R^3 \rightarrow C$ .

**2.4. Quadrature.** Let  $p_e$  and  $p_o$  be odd and even functions (not necessarily polynomials) defined on  $[-1, 1]$ , and define axially symmetric even and odd functions  $f_e^{(0,0,1)}(x, y, z) = w(r)p_e(z/r)$  and  $f_o^{(0,0,1)}(x, y, z) = w(r)p_o(z/r)$  for some non zero radial function  $w(r)$ . Then, as long as  $f_e^{(0,0,1)}, f_o^{(0,0,1)}$  have finite  $L_2$  norm, we may take  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  to be frequency responses of a pair of axially symmetric filters. Referring to  $f_e^{(0,0,1)}$  or  $if_o^{(0,0,1)}$  as “frequency response of a filter” is a bit cumbersome. In the interest of simple terminology, with a slight abuse, we will refer to  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  themselves as filters.

The notion of quadrature filters is well-defined for 1 dimensional (1D) signals. Two 1D filters are in quadrature if they are Hilbert transforms of each other [27]. In other words, their Fourier transforms, say  $\mathcal{F}_1, \mathcal{F}_2$ , are related by

$$(2.6) \quad \mathcal{F}_2(\omega) = \begin{cases} i\mathcal{F}_1(\omega) & \text{for } \omega > 0, \\ 0 & \text{for } \omega = 0, \\ -i\mathcal{F}_1(\omega) & \text{for } \omega < 0. \end{cases}$$

If we assume that  $\mathcal{F}_1$  is even, then  $\mathcal{F}_2$  is odd. This notion of quadrature filters does not generalize easily to 2 or higher dimensions [12]. However, for the purposes of this paper, we do not need a complete generalization. We only need a notion of quadrature for  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  (viewed as frequency responses of filters). Following (2.6) we say that  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  (viewed as  $\mathcal{F}_1$  and  $\mathcal{F}_2$ ) are in quadrature if

$$(2.7) \quad if_o^{(0,0,1)}(\omega) = \begin{cases} if_e^{(0,0,1)}(\omega) & \text{for } \omega^T(0, 0, 1)^T > 0, \\ 0 & \text{for } \omega^T(0, 0, 1)^T = 0, \\ -if_e^{(0,0,1)}(\omega) & \text{for } \omega^T(0, 0, 1)^T < 0. \end{cases}$$

Observe that with this definition, if we restrict both of the 3D functions  $if_o$  and  $if_e$  to any line through the origin, the 1D restricted functions now satisfy (2.6), the definition of quadrature in one dimension. Now, recognizing that the radial parts  $w(r)$  are identical for both filters, and that the polar parts are functions on the sphere, we have the following.

**Definition 2.5.** *The filters  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  are in quadrature if the functions  $p_e$  and  $p_o$  satisfy  $p_e = p_o > 0$  in the north hemisphere and  $p_e = p_o = 0$  on the equator of the sphere.*

If  $w(r)$  is chosen to peak around  $\omega_0$  and close to zero away from  $\omega_0$ , then  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  can be said to measure the local frequency content at  $(0, 0, \pm\omega_0)$ . Further, if  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  can be steered, then the frequency content local to any direction, at distance  $\omega_o$ , can be measured. This is appealing and raises the question of whether the axially symmetric filters  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  can be designed to be steerable and also to be in quadrature? Unfortunately, the answer is no. To see this, recall that the polar parts of axially symmetric functions oriented along the  $z$ -axis must be functions of the  $z$ -coordinate of points on the unit sphere. And for such functions to have a finite steering basis, the polar parts must in fact be polynomials [25, 32]. That is, the polar parts of the filter pair  $f_e^{(0,0,1)}$  and  $if_o^{(0,0,1)}$  must consist of an even polynomial and an odd polynomial.

But then Definition 2.5 requires that the polar parts of our two filters agree entirely on the interval  $(0, 1]$ . Requiring an odd polynomial and even polynomial to agree on  $(0, 1]$  is impossible if both polynomials are nonzero. Thus, steerable filters designed according to Theorem 2.4 cannot be in exact quadrature.

However we can try to get these filters approximately in quadrature by minimizing a departure-from-quadrature (DQ) objective function, as shown in the next section.

**3. Departure from quadrature (DQ).** To define DQ, let  $z : S^2 \rightarrow [-1, 1]$  denote the  $z$ -axis coordinate function on the unit sphere. Then, the DQ of  $p_e$  and  $p_o$  is

$$(3.1) \quad \begin{aligned} & \int_{S^2_+} (p_e(z(a)) - p_o(z(a)))^2 da \\ &= \int_0^{\pi/2} (p_e(\cos \theta) - p_o(\cos \theta))^2 \sin \theta d\theta \\ &= \int_0^1 (p_e(z) - p_o(z))^2 dz, \end{aligned}$$

where,  $S_+^2$  is the north hemisphere,  $da$  is the infinitesimal area on the unit sphere, and  $\theta$  is the zenith angle.

Below, we minimize this DQ measure while imposing two natural constraints on  $p_e$  and  $p_o$ :

$$(3.2) \quad p_e(0) = p_o(0) = 0 \text{ and } p_e(1) = p_o(1) = 1.$$

The first constraint is the equator constraint in the definition of quadrature. The second constraint is the north hemisphere constraint of the quadrature evaluated at the north pole. Together, the constraints guarantee that the polynomials are not constant.

**3.1. Minimizing DQ.** Minimize the DQ of (3.1) subject to the constraints of (3.2) reduces to a standard quadratic program: Set  $p(z) = p_e(z) - p_o(z) = a_N z^N + a_{N-1} z^{N-1} + \dots + a_0$ . Then, the DQ objective function is  $\int_0^1 p^2(z) dz$ . In addition, the constraint  $p_e(0) = p_o(0) = 0$  translates to  $a_0 = 0$ . Setting  $a_0 = 0$  to satisfy this constraint, and letting  $a = (a_1, \dots, a_N)$  be a vector of the remaining coefficients, the DQ objective function is

$$(3.3) \quad \int_0^1 p(z)^2 dz = \sum_{i,j=1}^N a_i a_j \int_0^1 z^{i+j} dz = a^T H a,$$

where  $H_{ij} = \frac{1}{i+j+1}$ . The constraint  $p_e(1) = p_o(1) = 1$  translates to  $\sum_{i \text{ even}} a_i = 1$  and  $\sum_{i \text{ odd}} a_i = -1$ . Setting  $B^T = \begin{pmatrix} 1 & 0 & 1 & \dots & 0 \\ 0 & 1 & 0 & \dots & 1 \end{pmatrix}$ , this constraint can be written as  $B^T a = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ . Thus the minimization problem can be put into the following form.

**Problem 3.1.**

$$(3.4) \quad \hat{a} = \operatorname{argmin}_{a \in \mathbb{R}^N} \frac{1}{2} a^T H a \text{ subject to } B^T a = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

In this form, the first set of constraints in (3.2) is implicit in the objective function, while the second set of constraints is explicitly imposed. The matrix  $H$  is an instance of a generalized Hilbert matrix (with  $p = 2$ ).

**Definition 3.2.** Let  $p \geq 0$  be an integer. Then, a generalized Hilbert matrix is a matrix  $H \in M_N(\mathbb{R})$  with entries:  $H_{ij} = \frac{1}{i+j-1+p}$ .

All generalized Hilbert matrices are positive-definite [8]. Thus, Problem 3.1 is a standard quadratic program.

The Lagrangian for the quadratic program is

$$L(a, \lambda) = \frac{1}{2} a^T H a - \lambda^T \left( B^T a - \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right),$$

where  $\lambda = (\lambda_1 \ \lambda_2)^T$  is the Lagrange multiplier. According to KKT theory [18] the optimal  $\hat{a}$  is obtained by setting  $\frac{\partial L}{\partial a} = \frac{\partial L}{\partial \lambda} = 0$  which gives

$$(3.5) \quad \hat{a} = H^{-1} B \lambda, \text{ and } \lambda = K^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

where  $K = B^T H^{-1} B$ . Thus,

$$(3.6) \quad \hat{a} = H^{-1} B K^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

The  $H$  matrix is invertible because it is positive-definite. Moreover, because the columns of  $B$  (rows of  $B^T$ ) are linearly independent,  $K$  is also positive-definite, and hence invertible. Thus the solution in (3.6) is well-defined.

The optimal  $\hat{a}$  gives the *optimizing polynomial*  $p_{\hat{a}} = \hat{a}(N)x^N + \cdots + \hat{a}(1)x$  from which the optimal even and odd polynomials are available as  $p_{e,\hat{a}}(z) = \sum_{i \text{ even}} \hat{a}_i z^i$ ,  $p_{o,\hat{a}}(z) = \sum_{i \text{ odd}} -\hat{a}_i z^i$ . Note the change in the sign of the coefficients for the odd polynomial. The optimal polynomial of the higher order in the pair is either the odd or the even polynomial, depending on whether  $N$  is odd or even.

**4. Properties of the optimal solution.** The minimizing polynomial and the resulting even and odd optimal polynomials have a number of interesting properties, expressed as various theorems below. These theorems address the behavior we explained as being desirable in steerable filters: The theorems explore angular concentration, closeness to quadrature, and polar monotonicity. The proofs of these theorems are postponed to section A of the appendix. The proofs depend critically on  $H$  being a generalized Hilbert matrix.

**4.1. Positivity and shape of the polynomial.** We begin by characterizing the signs of the odd and even components of the minimizing  $\hat{a}$  of (3.6). The  $i$ th component of  $\hat{a}$  is denoted  $\hat{a}_i$ ,  $i = 1, \dots, N$ .

**Theorem 4.1.** *The even components of the optimum vector  $\hat{a}$  are positive and the odd components are negative, i.e.,  $\hat{a}_i > 0$  for  $i$  even, and  $\hat{a}_i < 0$  for  $i$  odd.*

An immediate consequence of the signs of the odd and even components is the following.

**Corollary 4.2.** *The optimal even and odd polynomials  $p_{e,\hat{a}}, p_{o,\hat{a}}$  have positive coefficients. Consequently, they are positive and have positive derivatives on  $(0, 1]$ .*

The filters  $f_e^{(0,0,1)}$  and  $f_o^{(0,0,1)}$  use these polynomials for the polar part. And the positivity of the polynomials gives a single lobe to the filters in the north and south hemispheres; there are no sidelobes. That is,  $f_e^{(0,0,1)}$  and  $f_o^{(0,0,1)}$  take high values along the axis  $(0, 0, 1)$  and their values decrease monotonically as the angle from the axis increases. That is, the filters  $f_e^{(0,0,1)}$  and  $f_o^{(0,0,1)}$  are axially “concentrated” or “focused.”

The next theorem shows that the concentration around the axis  $(0, 0, 1)$  increases with  $N$ , the order of the polynomial. Asymptotically as  $N \rightarrow \infty$ , the support of the filters shrinks to just the axis.

**Theorem 4.3.**

(a) *For any  $i$ , the  $i$ th component  $\hat{a}_i \rightarrow 0$  as  $N \rightarrow \infty$ .*

(b) *The optimal polynomials  $p_{e,\hat{a}}, p_{o,\hat{a}}$  converge pointwise to 0 in the interval  $[0, 1)$  as  $N \rightarrow \infty$ , while  $p_{e,\hat{a}}(1) = p_{o,\hat{a}}(z) = 1$ . That is, the optimal polynomials converge to the indicator function of 1 on  $[0, 1]$ .*

Thus the optimal polynomials become more and more concentrated towards the endpoint  $z = 1$  as  $N$  increases. Because the polynomials converge pointwise to 0 in  $[0, 1)$ , they converge



pointwise to 0 in  $(-1, 1)$  (the polynomials are odd and even). Hence their  $L_1$  and  $L_2$  norms converge to 0 on  $[-1, 1]$ .

Since the optimal odd and even polynomials are the polar parts of steerable filters, the steerable filters become more concentrated around their axes as  $N$  increases.

**4.2. The asymptotic behavior of the DQ.** The optimal polynomials (and hence the filters) are designed to minimize the DQ. The following theorem shows that the DQ vanishes as  $N$ , the order of the polynomial increases.

**Theorem 4.4.** *The minimum value of the DQ approaches 0 as  $N \rightarrow \infty$ , i.e.,  $\frac{1}{2}\hat{a}^T H \hat{a} \rightarrow 0$  as  $N \rightarrow \infty$ .*

In other words, exact quadrature can be approached as closely as desired by increasing the order of the polynomials. In fact, numerical calculations in section 5 show that the DQ appears to decrease exponentially with  $N$ .

The theorems above show that filters designed by using the minimizing polynomials have the desired properties: The filters are steerable. The filters are concentrated around their axis, and do not change sign in the north and south hemispheres. Moreover, the filters are close to quadrature, and the concentration around the axis and closeness to quadrature can be arbitrarily increased by increasing  $N$ .

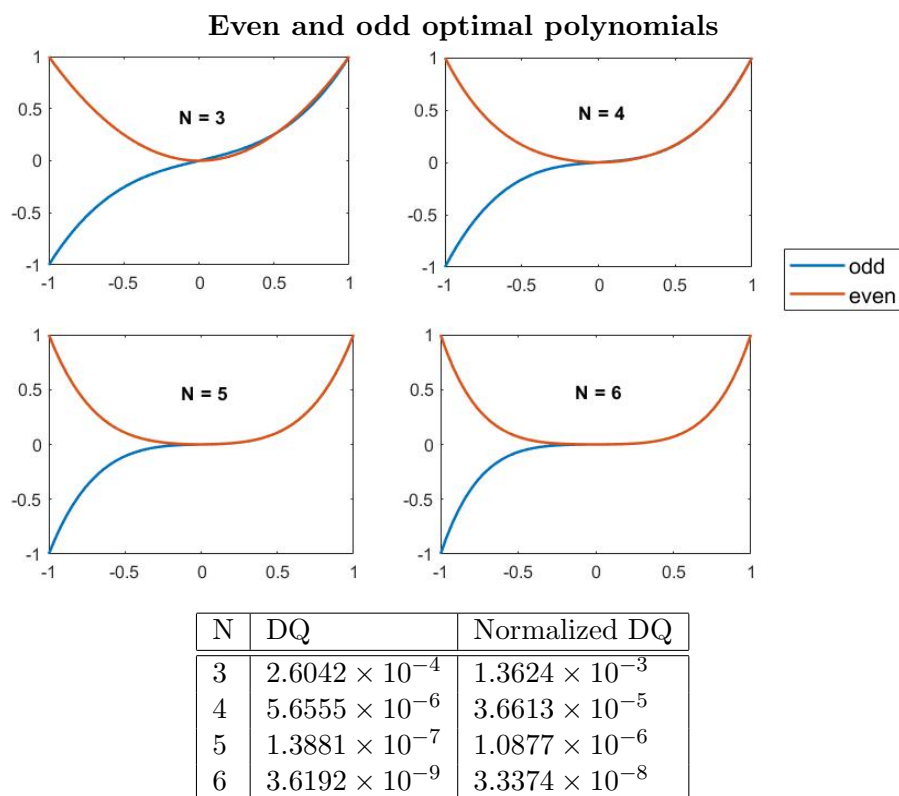
**5. Numerical results.** We now turn to present numerical results about the optimal polynomials and the resulting filters.

The theorems in section 4 gave asymptotics for several quantities. To gain additional insight, we numerically investigate the behavior of these quantities for large but finite  $N$ . We especially investigate the rate at which these quantities converge.

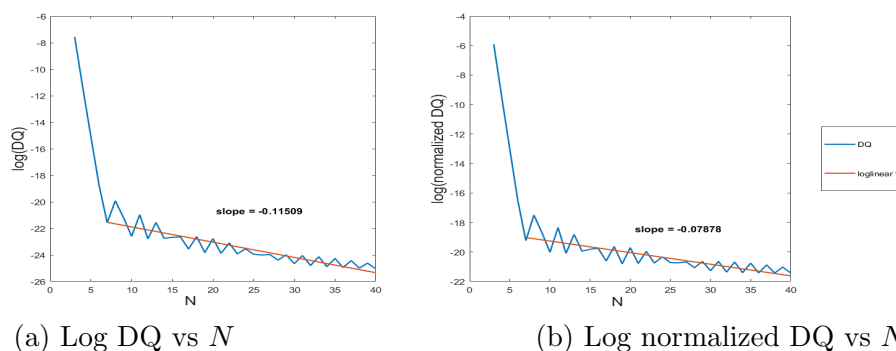
Below we plot the DQ (3.1) and other approximation errors of the filters. We report these errors as a fraction, by dividing the errors with the  $L_2$  norm of the filter, or of some other quantity. In each case, we indicate what the denominators of these *normalized errors* are.

**5.1. Optimal polynomials.** Figure 1 shows the even and odd optimal polynomials for  $N = 3, 4, 5, 6$  in the interval  $[-1, 1]$ . The positivity and monotonic behavior on  $(0, 1]$  is quite clear from the plots. The increasing concentration near 1 with increasing  $N$  is also clear. Figure 1 contains a table which shows the DQ and the normalized DQ values for these polynomials. DQ is normalized by dividing by the  $L_2$  norm of the polynomial over the interval  $[-1, 1]$ . Note the rapid decrease of the DQ and the normalized DQ with increasing  $N$ , even for these relatively low order polynomials.

Figure 2 contains the relation between the DQ, normalized DQ, and  $N$  in greater detail. Figure 2(a) shows the logarithm of DQ as a function of  $N$  for  $N = 3, \dots, 40$ . The figure also shows a linear fit to the calculated values. The linear fit (fit to the data from  $N = 7$  up to  $N = 40$ ) suggests that the DQ decreases with  $N$  as  $\exp(-0.115N)$ . Figure 2(b) shows the logarithm of the normalized DQ as a function of  $N$ . The normalized DQ is defined as the ratio of DQ to the  $L_2$  norm squared of the highest order polynomial in the optimal even odd pair. The linear fit (fit to the data from  $N = 7$  up to  $N = 40$ ) suggests that the DQ decreases with  $N$  as  $\exp(-0.079N)$ . Figure 2 provides support for Theorem 4.4.



**Figure 1.** Even and odd optimal pairs of polynomials plotted in the interval  $[-1, 1]$  for  $N = 3, 4, 5, 6$ . The even polynomial is displayed in red and the odd polynomial in blue. The table shows the DQ and normalized DQ values for the polynomial of higher order in each pair (labeled by  $N$ ).



**Figure 2.** (a) The logarithm of the DQ loss versus  $N$  with a linear fit. (b) The logarithm of the normalized DQ loss versus  $N$  with a linear fit. The normalized DQ is defined as the DQ divided by the  $L_2$  norm squared of the highest order optimal even/odd polynomial pair of order  $N$ .

Figure 3 provides support for Theorem 4.3, which says that the polynomials converge to 0 on  $[0, 1)$  (hence on  $(-1, 1)$ ). Figure 3 shows the  $L_2$  norm squared of the highest order polynomial (amongst the even/odd optimal polynomials) as a function of  $N$  for  $N = 3, \dots, 40$ .

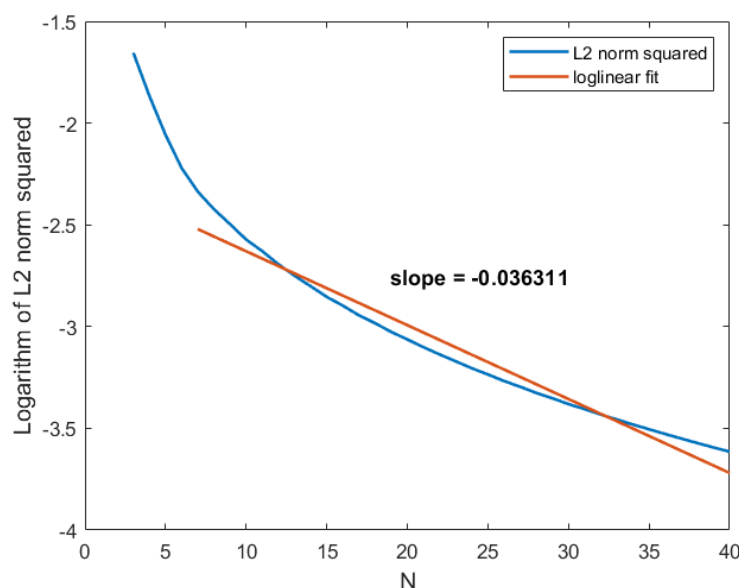


Figure 3.  $L_2$  norm squared of the optimal polynomial of higher order versus  $N$ .

Note the monotonic decrease in the  $L_2$  norm. This is consistent with Theorem 4.3(b).

**6. Applications.** In this section, we focus on two applications of steerable filters. The first is that of approximating 3D Gabor filter pairs. This is discussed below in section 6.1. The second application, presented in section 6.2, uses the 3D Gabor filter approximations to understand the local Fourier shell correlation in 3D reconstructions of the structure of protein molecules.

**6.1. Approximating 3D Gabor filter pairs.** Steerable filters using the optimal polynomials can approximate 3D Gabor filter pairs. Some notation is necessary to describe the approximation succinctly: When we fix  $N$  and obtain the optimal polynomial pair for a given maximum degree, we are simultaneously determining the order of both polynomials in the pair—the higher order polynomial will be of degree  $N$  and the other will be of degree  $N - 1$ . For example, if we set  $N = 5$ , then our higher order polynomial will be of degree 5 and therefore odd, whereas the even polynomial in the pair will have degree 4. Below, we denote the higher order optimal polynomial as  $p_N$  and the other polynomial as  $p_{N'}$ , regardless of whether  $N$  is even or odd.

Our approach to approximating 3D Gabor filter pairs is the following: We construct a steerable filter of the form

$$\hat{\mathcal{G}}(x, y, z) = e^{-a_2(r-r_0)^2 - a_1(r-r_0) - a_0} p_N(z/r),$$

where  $x, y, z$  are coordinates in the Fourier domain,  $r = \sqrt{x^2 + y^2 + z^2}$ , and  $r_0$  is the frequency at which the radial part of the function has a maximum. The coefficients  $a_0, a_1, a_2$  are parameters to be adjusted. This filter is even if  $N$  is even, else it is odd. And regardless of whether it is even or odd, it is axially symmetric about the  $z$ -axis. Note that the first

exponential term is the radial part of the filter (i.e., the term  $w(r)$ ), whereas the polynomial term is the polar part of the filter.

With our method, the polar part of the filter is determined entirely by the choice of  $N$ ; the radial part depends on  $a_2, a_1, a_0$ . Given  $N$ , we adjust  $a_2, a_1, a_0$  such that the degree  $N$  filter in our pair best approximates its counterpart of the frequency response of the Gabor filter pair. We then use the same radial function for the filter of degree  $N - 1$ . For example, if  $N = 17$ , then we choose  $a_2, a_1, a_0$  so that  $\hat{G}(x, y, z)$  best approximates the imaginary part (the odd part) of the frequency response of the Gabor filter pair, and then use those same coefficients for our even filter of degree  $N = 16$  which approximates the real part of the frequency part of the Gabor filter pair.

Let  $\mathcal{G}(x, y, z)$  be the frequency response of the chosen Gabor filter with  $r_0$  as the peak frequency. Then  $\mathcal{G}$  is approximated by  $\hat{\mathcal{G}}$  by numerically minimizing the  $L_2$  norm of  $\mathcal{G} - \hat{\mathcal{G}}$  with respect to  $a_0, a_1, a_2$ , using the MATLAB `fminsearch`. Finally, we take

$$\hat{\mathcal{G}}'(x, y, z) = e^{-a_2(r-r_0)^2 - a_1(r-r_0) - a_3} p_{N'}(z/r)$$

as an approximation of the other Gabor filter in the Gabor filter pair.

To illustrate this idea with a concrete example, we take the Fourier transform of a 3D Gabor filter pair sampled on a grid of  $128 \times 128 \times 128$  voxels. We let the frequency peak of the Fourier transform of the filter be located at 32 voxels from the zero frequency in the  $x, z$  plane, and we let the width (standard deviation) of the peak be 9.697 voxels [26]. To be precise, the exact formulas for our Gabor filter pairs are given by

$$\begin{aligned} \mathcal{G}_{odd}(x, y, z) &= \frac{1}{2} \exp \left\{ -\frac{x^2 + y^2 + (z - 32)^2}{2(9.697)^2} \right\} - \frac{1}{2} \exp \left\{ -\frac{x^2 + y^2 + (z + 32)^2}{2(9.697)^2} \right\}, \\ \mathcal{G}_{even}(x, y, z) &= \frac{1}{2} \exp \left\{ -\frac{x^2 + y^2 + (z - 32)^2}{2(9.697)^2} \right\} + \frac{1}{2} \exp \left\{ -\frac{x^2 + y^2 + (z + 32)^2}{2(9.697)^2} \right\}. \end{aligned}$$

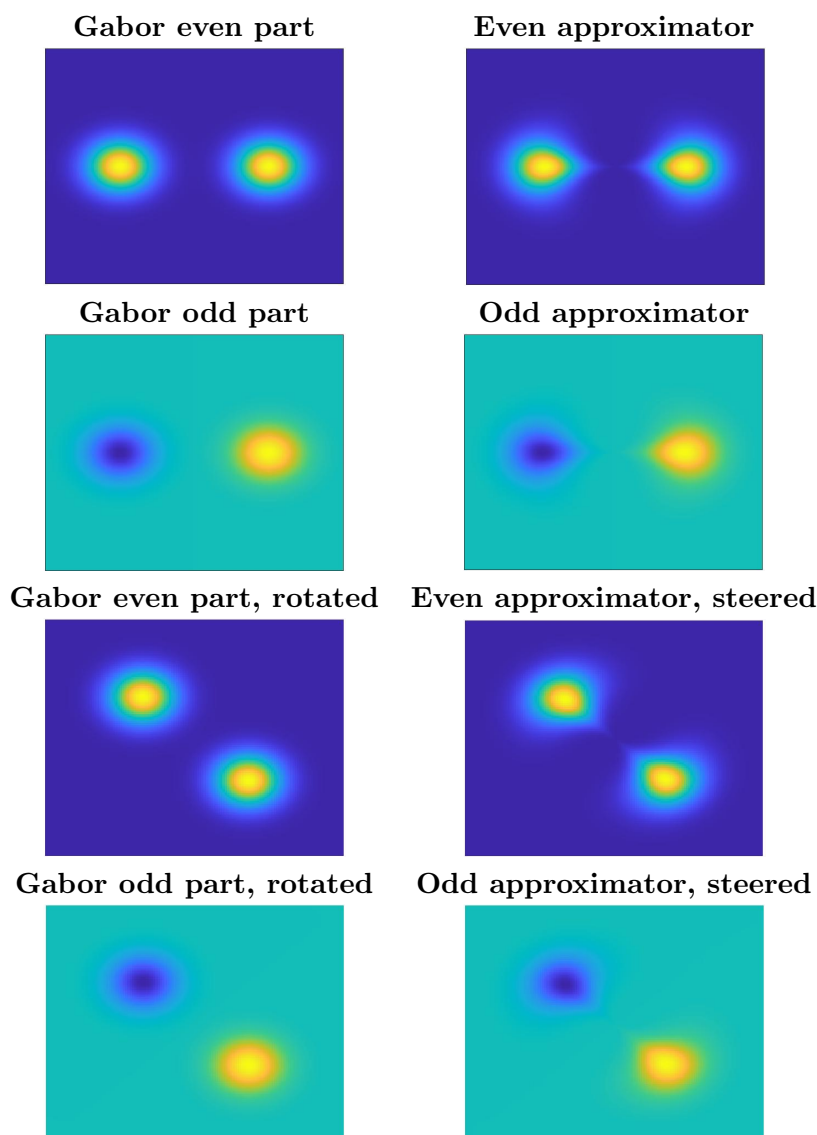
Next we perform the optimization as described above to obtain steerable approximator filters for these filters. We take the filters giving the minimum  $L_2$  norm of  $\mathcal{G} - \hat{\mathcal{G}}$  for a range of  $N$ . As shown in Figure 5 the minimum  $L_2$  norm varies with  $N$ , with the smallest minimum achieved at  $N = 23$ . For this value of  $N$ , the optimal radial part is

$$W(r) = \exp\{-0.0021(r - 32)^2 - 0.0102(r - 32) - 0.6492\}.$$

The accompanying polynomials are

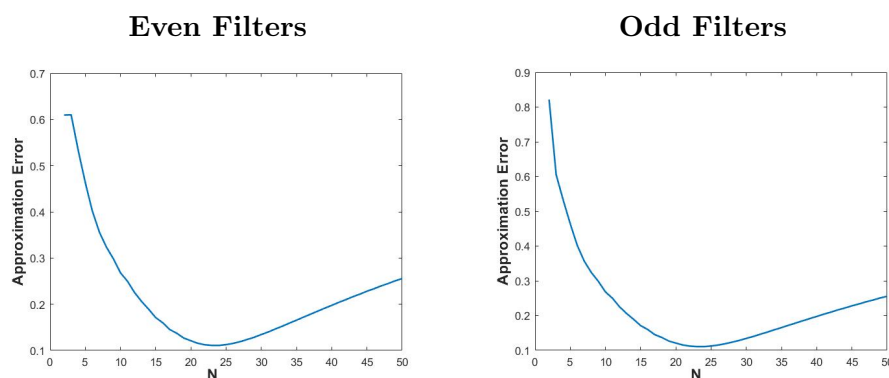
$$\begin{aligned} p_{odd}(z') &= 0.0661z'^{23} + 0.085z'^{21} + 0.098z'^{19} + 0.1042z'^{17} + 0.1038z'^{15} + \\ &\quad 0.0979z'^{13} + 0.0903z'^{11} + 0.089z'^9 + 0.1026z'^7 + 0.1101z'^5 + 0.0522z'^3 + 0.0008z', \\ p_{even}(z') &= 0.1175z'^{22} + 0.1014z'^{18} + 0.0918z'^{18} + 0.0889z'^{16} + 0.0924z'^{14} + 0.0999z'^{12} + \\ &\quad 0.1056z'^{10} + 0.0999z'^8 + 0.0823z'^6 + 0.1095z'^4 + 0.0109z'^2, \end{aligned}$$

where  $z' = z/r$ .



**Figure 4.** The real and imaginary parts of the Fourier transforms of the Gabor filters, and their corresponding optimal approximators. The original filters are shown in the first two rows; the rotated filters are shown in the bottom two rows, with the Gabor filters rotated via brute force and the approximators steered using a steering basis.

The first two rows of the left column of Figure 4 show the frequency responses of the even and odd parts of the Gabor filter pair in the  $x, z$  plane in the Fourier domain. The first two rows of the right column of Figure 4 show the frequency responses of the approximating steerable filter for  $N = 23$ . The next two rows of the figure show the steering behavior of the approximating filters. The first two columns show brute force rotation of the even and odd Gabor pairs. The next two columns show the steered filters calculated according to (2.4).



**Figure 5.** The  $L_2$  norm of  $\mathcal{G} - \hat{\mathcal{G}}$  divided by the  $L_2$  norm of  $\mathcal{G}$  for even and odd Gabor filter approximation.

Figure 5 shows the normalized  $L_2$  norm of the difference of the even and odd Gabor filters and their steering approximations, i.e.,  $\mathcal{G} - \hat{\mathcal{G}}$ , as a function of  $N$ . For normalization, the  $L_2$  norm of  $\mathcal{G} - \hat{\mathcal{G}}$  is divided by the  $L_2$  norm of  $\mathcal{G}$ . The minimum value that this normalized  $N$  takes is 0.1107 for  $N = 23$ .

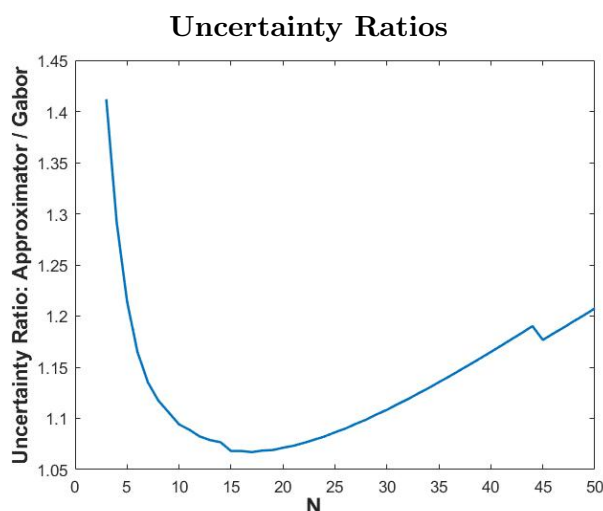
Gabor filters are useful because they have minimum space-frequency uncertainty. It is useful to compare the space-frequency uncertainty of our steerable approximations for each  $N$  with the space-frequency uncertainty of the Gabor filters that they are approximating. For any function  $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ , the space-frequency uncertainty is the product

$$\left( \int_{\mathbb{R}^3} |x|^2 |f(x)|^2 dx \right) \left( \int_{\mathbb{R}^3} |\xi|^2 |\mathcal{F}(\xi)|^2 d\xi \right),$$

where  $\mathcal{F}$  denotes the Fourier transform of function  $f$ . For all  $f$  with unit  $L_2$  norm, this product has a lower bound, which is achieved when  $\hat{f}$  is the odd or even part of the Gabor Fourier transform. We computed the space-frequency uncertainty for the even approximator filter for different  $N$  and compared it with the space-frequency uncertainty of the even part of the Gabor Fourier transform (the uncertainty of the odd part is identical to the even). Figure 6 plots the ratio of the steerable filter uncertainty to Gabor filter uncertainty as a function of  $N$ . The ratio has a minimum at  $N = 17$  and at the minimum its value is equal to 1.0747. Note that in the range  $N = 10 - 25$  the steerable filter uncertainty is within 10% of the Gabor filter uncertainty.

**6.1.1. Choice of steering directions.** To select steering directions, we used the Particle-SampleSphere from the S2 Sampling Toolbox [31]. This function gives approximately uniform  $M$  points on a sphere. For  $m$  steering directions, we let  $M = 2m$  and select the directions represented in the upper half-sphere (when  $M$  is even, the points in the lower hemisphere lie on the same directional axes as the points in the upper hemisphere).

**6.2. Local Fourier shell correlation (FSC) in Cryo-EM.** As mentioned in section 1, cryo-EM is a modern method for reconstructing the 3D molecular structure of proteins [14]. In brief, cryo-EM works as follows: Protein molecules are purified from biological material and frozen in a thin layer of vitreous ice. This sample is imaged with a transmission electron microscope.



**Figure 6.** The ratios obtained by dividing the uncertainty products of our approximators by the uncertainty product of the Gabor filter. Note that the minimum value of 1.0747 is achieved at  $N = 17$  and the low values are maintained in the range  $N = 10$  to  $N = 25$ .

The resulting image contains many tomographic images of the protein at random orientations. Several thousand such images are used to reconstruct the 3D structure of the protein. Mathematically, each reconstructed structure is a map from  $R^3$  to  $R$ . Figures 7(a)–(b) show one example; this is the reconstructed structure of the capsaicin receptor TRYPV1 obtained from the Electron Microscopy Data Bank (EMBD entry 5778, <https://www.emdataresource.org>) [24]. The background has been masked in the figure.

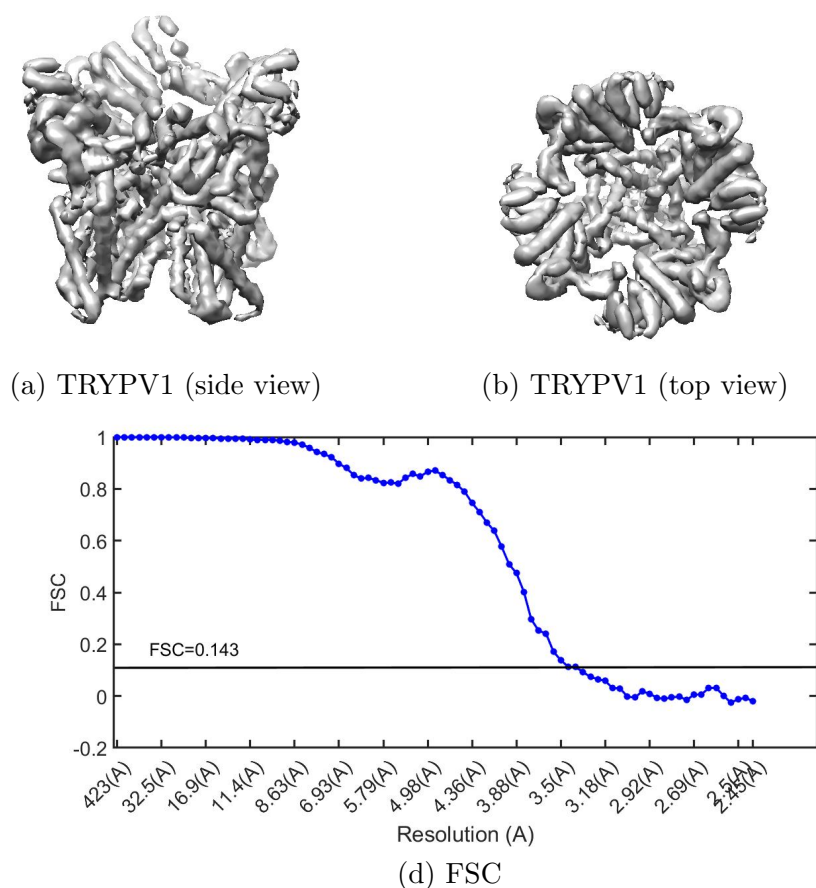
Cryo-EM images are extremely noisy, and this noise percolates into the 3D construction. This raises the possibility that artifacts due to noise may be misinterpreted as valid features of the structure. To avoid this, a form of cross validation is used. The set of all images is split into two subsets with an equal number of images, and a 3D structure is reconstructed from each subset.

Suppose that the two reconstructions are  $V_1 : R^3 \rightarrow R$  and  $V_2 : R^3 \rightarrow R$  with Fourier transforms  $\mathcal{V}_1$  and  $\mathcal{V}_2$ . Let  $S_\omega$  be the surface of a sphere of radius  $\omega$  in the Fourier domain. This surface contains all frequencies which have the same wavelength but different orientations. Define the *FSC* of the two volumes restricted to this surface as

$$(6.1) \quad \text{FSC}(\omega) = \frac{\text{Re} \int_{S_\omega} \mathcal{V}_1(u) \mathcal{V}_2^*(u) du}{\sqrt{\int_{S_\omega} \mathcal{V}_1(u) \mathcal{V}_1^*(u) du} \sqrt{\int_{S_\omega} \mathcal{V}_2(u) \mathcal{V}_2^*(u) du}},$$

where  $u$  is a point in  $S_\omega$  and  $du$  is a differential area in  $S_\omega$ . In other words,  $\text{FSC}(\omega)$  is the correlation of the Fourier transforms of the two volumes on  $S_\omega$ . High FSC values indicate a good signal-to-noise ratio for frequencies in  $S_\omega$ , and low FSC values indicate poor signal-to-noise ratio for frequencies in  $S_\omega$ .

It is standard practice in cryo-EM to plot  $\text{FSC}(\omega)$  as a function of the radius  $\omega$  (the radius is converted to a wavelength, which is expressed in angstroms). See Figure 7(d) for the



**Figure 7.** (a)–(b) Structure of the capsaicin receptor TRYPV1. (c) FSC of the structure. The structure has a resolution of  $3.275\text{Å}$  as reported in [24].

FSC of the TRYPV1 structure (the EMDB provides the two volumes to calculate the FSC with. These volumes were used to calculate the FSC in the figure. The volumes were also masked to suppress the background). Typically, the value of  $\omega$ , translated from frequency to wavelength, at which FSC equals 0.143 is taken as the *resolution* of the reconstruction [30]. It is the highest spatial frequency which can be reliably interpreted in the structure. Features of the structure at finer spatial scales are suspect. All of this is standard practice in cryo-EM [14]. A cryo-EM structure cannot be published or included in the EMDB without FSC analysis and a measurement of resolution. The resolution of the TRYPV1 receptor shown in Figure 7 is  $3.275\text{Å}$ .

One problem with the FSC is that it characterizes the volume as a whole. But often, even simple visual inspection shows that the resolution of most reconstructed volumes is not globally uniform; it can vary quite a bit locally. To deal with this, various local resolutions are defined in cryo-EM, and reviewed in [37]. The idea behind local resolution is to replace the Fourier transforms used in (6.1) by local versions of the Fourier transforms.



By analogy with (6.1), a local FSC (LFSC) at a point  $x \in R^3$  in the volume can be defined as

$$(6.2) \quad \text{LFSC}(x, \omega) = \frac{\text{Re} \int_{S_\omega} \ddot{V}_1(x, u) \ddot{V}_2^*(x, u) du}{\sqrt{\int_{S_\omega} \ddot{V}_1(x, u) \ddot{V}_1^*(x, u) du} \sqrt{\int_{S_\omega} \ddot{V}_2(x, u) \ddot{V}_2^*(x, u) du}},$$

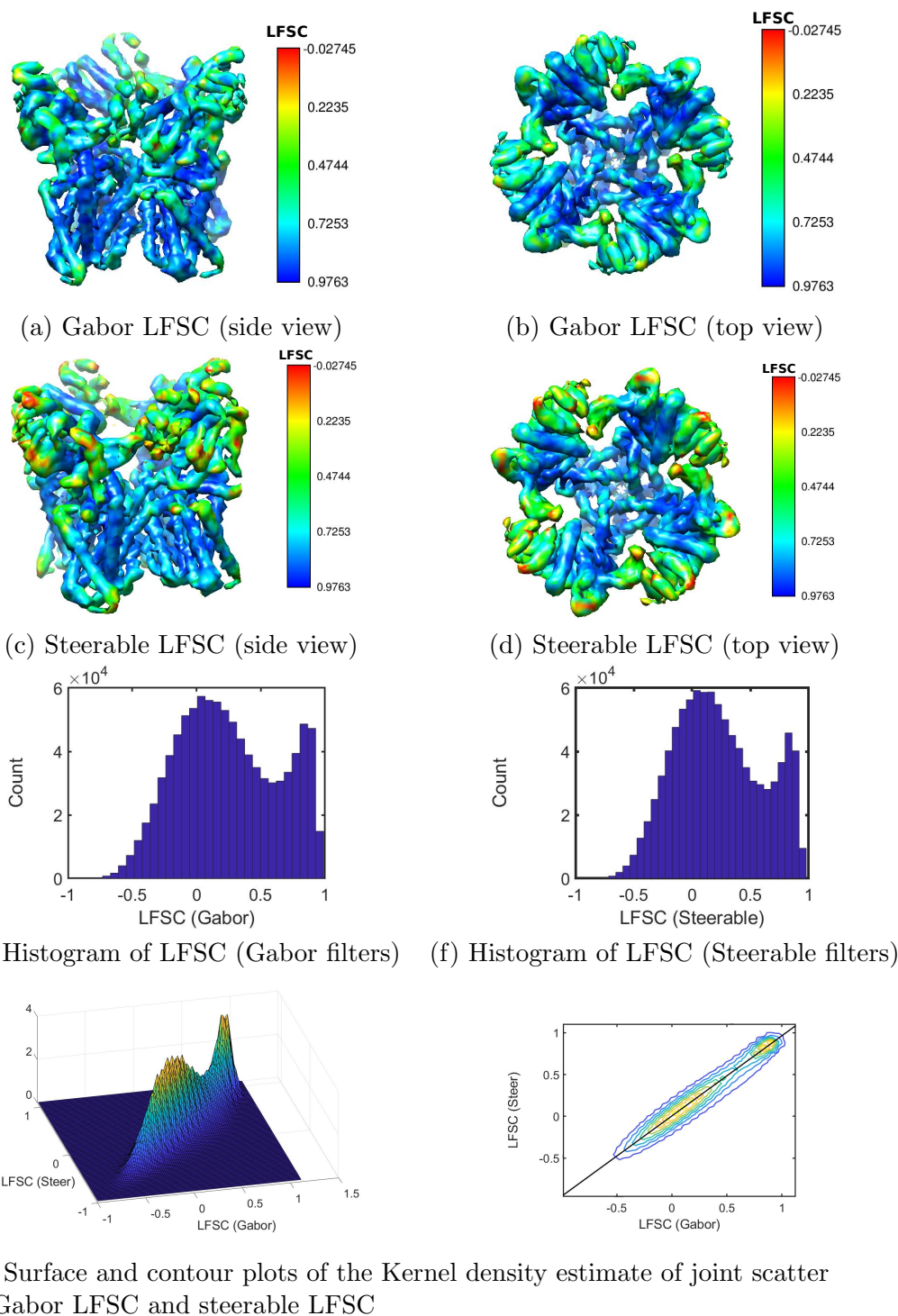
where  $\ddot{V}_1(x, u) = \mathcal{F}^{-1}(\mathcal{V}_1 \cdot (\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))(x)$  and  $\ddot{V}_2(x, u) = \mathcal{F}^{-1}(\mathcal{V}_2 \cdot (\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))(x)$ , and  $\mathcal{F}^{-1}$  is the inverse Fourier transform. In addition,  $\mathcal{G}_{\text{even}}^u$  and  $\mathcal{G}_{\text{odd}}^u$  are Fourier transforms of the even and odd Gabor filters with center frequency  $u$ . The idea is that  $\mathcal{F}^{-1}(\mathcal{V}_1 \cdot (\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))(x)$  is the convolution of  $V_1$  and  $\mathcal{F}^{-1}((\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))$  evaluated at  $x$ . Because  $\mathcal{F}^{-1}((\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))$  is a complex function whose real and imaginary parts are Gaussian enveloped cosine and sine functions,  $\mathcal{F}^{-1}((\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))(x)$  measures the local  $u$  frequency content in  $V_1$  at  $x$ . Similarly  $\mathcal{F}^{-1}(\mathcal{V}_2 \cdot (\mathcal{G}_{\text{even}}^u + i\mathcal{G}_{\text{odd}}^u))(x)$  measures the local  $u$  frequency content of  $V_2$ . Thus, (6.2) is the local FSC.

In practice, the LFSC is calculated by approximating the integrals in the numerator and denominator of (6.2) by Riemann sums [38]. That is, a finite number of vectors  $u$  are chosen to lie uniformly on the surface  $S_\omega$  and the integrals are approximated by Riemann sums involving these  $u$ 's. Each term in the Riemann sum requires a convolution with  $\mathcal{G}_{\text{even}}^u$  and  $\mathcal{G}_{\text{odd}}^u$ , which is computationally expensive.

But Gabor filters aligned along different  $u \in S_\omega$  are merely 3D rotations of each other. Thus we can use the optimal approximation of the Gabor filter pairs with the steerable filters suggested in section 6.1 to an advantage. The idea is to optimally approximate the even and odd Gabor filters  $\mathcal{G}_{\text{even}}^u$  and  $\mathcal{G}_{\text{odd}}^u$  with the corresponding 3D steerable filters  $p_{\text{even}}^u$  and  $p_{\text{odd}}^u$  aligned along  $u$ . Then, we need only evaluate the convolutions of the steerable bases of these filters. Appropriate linear sums of the convolutions with the steerable bases give convolutions with  $p_{\text{even}}^u$  and  $p_{\text{odd}}^u$  at the desired  $u$ 's. These convolutions can be taken to be  $\ddot{V}_1$  and  $\ddot{V}_2$  in LFSC. This saves computation, because there are fewer steerable bases than  $u$ 's.

As noted above, the FSC curve in Figure 7(d) reaches a value of 0.143 at 3.275 Å. We take  $S_\omega$  to be the surface of the sphere which corresponds to 3D frequencies whose wavelength is 3.275 Å. To calculate the LFSC, we approximated the integrals in the numerator and denominator with Riemann sums. The sums are evaluated at 300  $u$ 's that are uniformly spaced on  $S_\omega$ . We first calculated the LFSC using Gabor filters oriented along these  $u$ 's. The Gabor filters have a peak frequency exactly at the frequency corresponding to  $S_\omega$ . The width of the Gabor filter frequency response was set to 1.094 Å (20.7059 voxels), so that the entire Gaussian envelope of the frequency response of the Gabor filter could be contained in the FFT of the volume. Of course, calculating the LFSC using this brute force approach required  $300 \times 2$  convolutions with the Gabor filters for each volume, for a total of 1200 convolutions.

Figures 8(a)–(b) shows the LFSC calculated using Gabor filter convolutions. The TRYPV1 receptor is rendered in color in the figure. The color denotes the LFSC at every voxel. The color bar in Figures 8(a)–(b) indicates how LFSC values are mapped to color. The figure clearly shows that the LFSC values change considerably, illustrating that reconstruction has a locally variable signal-to-noise ratio, and hence has more or less structural information in different parts. This is the benefit of using the LFSC.



**Figure 8.** (a)–(b) *LFSC* calculated using Gabor filters. (c)–(d) *LFSC* calculated using steerable filters. In (a)–(d), the volume is rendered to express the *LFSC* at each voxel according to the colorbar to the right of each subfigure. (e) Histogram of *LFSC* with Gabor filters. (f) Histogram of *LFSC* with Steerable filters. (g) Surface and contour plots of the 2D kernel density estimate of the joint scatter of Gabor *LFSC* and steerable filter *LFSC*. The contour plot also shows a 45 degree line.

Next, we optimally approximated the odd and even Gabor filters with steerable 3D filters using the procedure of the previous section. The approximation used a pair of polynomials of order 17 and order 16, and their steerable basis required 171 and 153 functions, respectively. The bases were convolved with each volume for a total of  $324 \times 2 = 628$  convolutions. Thus, the results of 628 convolutions were taken and used to approximate the  $300 \times 4$  Gabor convolutions in the Riemann sums of the numerator and denominator of the LFSC.

Figures 8(c)–(d) show the LFSC calculated using the steerable approximations. The color bar in Figure 8(c)–(d) is identical to the color bar of Figure 8(a)–(b). Qualitatively, the steerable filter LFSC is quite similar to the Gabor filter LFSC. Figures 8(c)–(d) show the histograms of the LFSCs. Note that both histograms are similar and that the left peak in both histograms is close to 0.143, the value of the (global) FSC on  $S_\omega$ . To investigate relative values of the two LFSCs further, for every voxel in the volume, we took the Gabor LFSC value and the steerable filter LFSC value and made a scatter plot of these pairs of values. Figure 8(g) shows a plot of the kernel density estimate of the scatter (the scatter plot itself was too dense to convey where most of the scatter mass lies). A surface plot as well as a contour plot are shown. Note in the surface plot that the scatter is highly concentrated around the diagonal line. This is even more apparent in the contour plot, which also shows a 45 degree line. Finally, we calculated the correlation coefficient between the Gabor LFSC and the steerable filter LFSC. The correlation coefficient was 0.97, showing that the steerable filter LFSC closely approximated the Gabor LFSC. And the steerable LFSC approximation required only about half the number of the convolutions of the Gabor LFSC calculation. Of course, the number of convolutions required for the steerable approximation method does not grow with the number of Riemann sums used to compute the integral.

This example illustrates the utility of having steerable almost-conjugate 3D filters which can closely approximate Gabor filter pairs.

**7. Discussion.** The optimal filter pairs proposed in this paper have surprising properties. Even for low order polynomials (Figure 1), the filter pair is close to quadrature, monotonic, and highly directional. And it is steerable in 3D. The filter pair also approximates Gabor filters quite well.

We emphasize that our goal is not to find steerable filters per se; our interest is in finding filter pairs, especially pairs which are directional, and which are close to quadrature. Thus we exclude more complicated steerable filters (Y or star shaped filters, for example) from consideration. These filters are not directional, and quadrature is difficult to define for them.

Finally, we note that proofs of various claims in section 4 of the paper are in the appendix. The key elements of the proofs exploit properties of the inverse Hilbert matrix, especially row-sums and column-sums of the matrix. Hilbert matrices are a subclass of both the more general Cauchy and Hankel matrices [13, 4]. It is likely that this more general class of matrices may also be used to prove the same theorems.

**8. Conclusion.** 3D steerable filters can be easily designed to be near-quadrature, especially when their polar parts are polynomial functions. These filters are sharply directional, and their directional as well as quadrature properties improve with the order of the polynomial function. These properties are closely related to the properties of the generalized Hilbert matrix. These filters also provide a practically useful approximation to the even and odd 3D Gabor filter pairs.

**Appendix A. Proofs of claims.**

**A.1. Properties of generalized Hilbert matrix inverse and notation.** It is clear from our solution in (3.6) as well as our rewritten constrained objective function (3.4) that our calculations are intimately intertwined with the inverse  $H^{-1}$ . In this section we define the class of matrices to which  $H$  belongs and give important properties that we will use in subsequent proofs.

**A.1.1. Definitions and notations.** We will show that the fact that our  $H$  in (3.4) is a generalized Hilbert matrix (of order 2) is sufficient to prove the results we claim about the behavior of the solutions  $\hat{a}$  as well as the asymptotic results.

Note first that  $H^{-1}B$  is a two column vector whose values are the sum of the odd- or even-indexed entries in each row, depending on the dimension of  $H$ . Consequently, it will be useful to adopt some notation to refer to the inverse of the Hilbert matrix, its entries, and certain specialized sums, such as the sum of the entries of a row (termed row-sum) or the sum of all odd-indexed entries.

**Definition A.1.** Let  $G = H^{-1}$ , where  $H$  is a generalized Hilbert matrix of size  $N$  for constant  $p > 0$ . Let  $G_{i,j}$  denote the entries of  $G$  (we write  $G_{ij}$  when unambiguous). Furthermore, let

$$g^{oo} = \sum_{i,j \text{ odd}}^N G_{i,j}, \quad g^{ee} = \sum_{i,j \text{ even}}^N G_{i,j},$$

$$g^{oe} = \sum_{i \text{ odd}, j \text{ even}}^N G_{i,j}, \quad g^{eo} = \sum_{i \text{ even}, j \text{ odd}}^N G_{i,j}$$

be the sums of all entries of  $G$  in, for example, odd-indexed rows and odd-indexed columns. Note in particular that since  $H$  is symmetric,  $G$  is also symmetric and therefore  $g^{oe} = g^{eo}$  (when  $G$  has even dimension).

Now, let  $1 \leq i \leq N$ . Then define

$$g_i^o = \sum_{j \text{ odd}}^N G_{i,j}, \quad g_i^e = \sum_{j \text{ even}}^N G_{i,j}, \quad r_i = g_i = \sum_{j=1}^N G_{i,j}$$

as the sums of odd-indexed entries, the even-indexed entries, and all entries of a particular  $i$ th row, respectively. Finally define

$$g^o = \sum_{i \text{ odd}}^N r_i, \quad g^e = \sum_{i \text{ even}}^N r_i$$

to be the sum of all entries in odd-numbered or even-numbered rows, respectively.

**Definition A.2.** For each definition above, we adopt a notation to indicate its absolute value. We let  $s^o = |g^o|, s^e = |g^e|, s_k^e = |g_k^e|$ , etc.

In particular, note that  $s_k^e + s_k^o$  is the sum of the absolute value of the entries of a particular row. These values will prove particularly useful.

### A.1.2. Exact expression for entries and sum of entries.

**Proposition A.3.** Every generalized Hilbert matrix  $H$  is positive definite.

*Proof.* This has been proven in [8]. ■

**Proposition A.4.** Let  $G = H^{-1}$  be the inverse of a generalized Hilbert matrix of size  $N$  and constant  $p > 0$ . Then the entries of  $G$  are given by

$$(A.1) \quad G_{ij} = \frac{(-1)^{i+j}}{p+i+j-1} \left( \frac{\prod_{k=0}^{N-1} (p+i+k)(p+j+k)}{(i-1)!(N-i)!(j-1)!(N-j)!} \right).$$

*Proof.* This has been proven as well in [8]. ■

**Proposition A.5.** Let  $G = H^{-1}$  be the inverse of a generalized Hilbert matrix of size  $N$  and constant  $p > 0$ . Fix some  $i \in \{1, \dots, N\}$ . Then the sum of the entries of row  $i$  is given by

$$(A.2) \quad \sum_{j=1}^N G_{ij} = (-1)^{N+i} \frac{\prod_{k=0}^{N-1} (p+i+k)}{(i-1)!(N-i)!}$$

$$(A.3) \quad = (-1)^{N+i} i \binom{N-1+p+i}{N} \binom{N}{i}$$

*Proof.* See [34] for the first equality. The second equality is easily verified. ■

**Proposition A.6.** Let  $G = H^{-1}$  be the inverse of a generalized Hilbert matrix of size  $N$  and constant  $p > 0$ . Then the sum of all entries of the matrix is given by

$$(A.4) \quad \sum_{i,j=1}^N G_{ij} = N(p+N).$$

*Proof.* See [34] for the proof. ■

**Corollary A.7.** We obtain the following expression for products of rows:

$$r_{i_1} r_{i_2} = (p + i_1 + i_2 - 1) G_{i_1 i_2}.$$

*Proof.* Comparing the expressions of (A.1), (A.2), we see that

$$\begin{aligned} r_{i_1} r_{i_2} &= \left( \sum_{j=1}^N G_{i_1 j} \right) \left( \sum_{j=1}^N G_{i_2 j} \right) \\ &= (-1)^{2N+i_1+i_2} \frac{\prod_{k=0}^{N-1} (p+i_1+k)(p+i_2+k)}{(i_1-1)!(N-i_1)!(i_2-1)!(N-i_2)!} \\ &= (-1)^{i_1+i_2} \frac{\prod_{k=0}^{N-1} (p+i_1+k)(p+i_2+k)}{(i_1-1)!(N-i_1)!(i_2-1)!(N-i_2)!} \\ &= (p+i_1+i_2-1) G_{i_1 i_2} \end{aligned}$$

as claimed. ■

**A.1.3. Notes on magnitudes and signs of important sums.**

**Corollary A.8.** *Let  $N$  be even. Then for every  $k$ ,  $|g_k^o| < |g_k^e|$ . Furthermore  $|g^{oo}| < |g^{oe}|$  and  $|g^{eo}| < |g^{ee}|$ . If  $N$  is odd then the inequalities are reversed.*

*Proof.* This follows quickly from Propositions A.5 and A.4. ■

**Corollary A.9.** *Let  $N$  be even. Then  $g^o < 0 < g^e$ . If  $N$  is odd then the inequalities are reversed.*

**Proposition A.10.** *The quantity  $g^{oo}g^{ee} - g^{oe}g^{eo}$  is positive.*

*Proof.* Note that our quantity  $g^{oo}g^{ee} - g^{oe}g^{eo}$  is the determinant of the matrix  $K = B^T H^{-1} B = \begin{pmatrix} g^{oo} & g^{oe} \\ g^{eo} & g^{ee} \end{pmatrix}$ . Thus it suffices to show that  $K$  is positive definite. But  $H$  is positive definite, and therefore  $G = H^{-1}$  is positive definite. Furthermore, since  $B$  has full column rank, for all  $x \in \mathbb{R}^2$ , the product  $Bx = 0$  only if  $x$  is zero. Therefore for all nonzero  $x \in \mathbb{R}^2$ ,  $x^T K x = (xB)^T G(Bx) > 0$  and therefore  $K$  is positive definite. Hence  $\det K = \det B^T G B = g^{oo}g^{ee} - g^{oe}g^{eo} > 0$ . ■

**A.2. Proof of Theorem 4.1: Optimal coefficients alternate in sign.** Our goal is to prove Theorem 4.1. First we will rewrite our expressions for the solution from KKT theory. Recall from (3.6) that we have

$$\hat{\alpha} = H^{-1} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_2 \end{pmatrix}, \text{ where } \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = K^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

But using the notation we defined in Section (A.1.1), we see that we can write explicitly:

$$K = B^T H^{-1} B = \begin{pmatrix} g^{oo} & g^{oe} \\ g^{eo} & g^{ee} \end{pmatrix}$$

Thus our formula for lambda can be rewritten as

(A.5) 
$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = K^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(A.6) 
$$= \frac{1}{g^{oo}g^{ee} - g^{oe}g^{eo}} \begin{pmatrix} g^{ee} & -g^{oe} \\ -g^{eo} & g^{oo} \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

(A.7) 
$$= \frac{1}{g^{oo}g^{ee} - g^{oe}g^{eo}} \begin{pmatrix} -g^{ee} - g^{oe} \\ g^{eo} + g^{oo} \end{pmatrix}$$

(A.8) 
$$= \frac{1}{g^{oo}g^{ee} - g^{oe}g^{eo}} \begin{pmatrix} -g^e \\ g^o \end{pmatrix}.$$

The following lemma will be used to prove our theorem.

**Lemma A.11.** *Let  $s^o, s^e, s_k^o, s_k^e$  be as in Definition A.2. Furthermore, if  $N$  is even, let  $s^+ = s^o + s^e, s^- = s^e - s^o$ , and for any  $k \in \{1, \dots, N\}$ , let  $s_k^+ = s_k^o + s_k^e$  and  $s_k^- = s_k^e - s_k^o$ .*

Otherwise, let  $s^- = s^o - s^e$ ,  $s_k^- = s_k^o - s_k^e$ . Then:

$$(A.9) \quad \frac{s^+}{s^-} < \frac{s_k^+}{s_k^-}$$

Note that our  $s_k^+$  are simply defined to be the sum of the absolute values of even-indexed and odd-indexed values, and the  $s_k^-$  to be the absolute value of the sum of the even-indexed and odd-indexed values, respectively. Both  $s^-$  and  $s_k^-$  are familiar values—modulo an absolute value sign, they are simply the sum of entries in a row, and the total sum of all elements. The  $s^+$ ,  $s_k^+$  values are meant to be an absolute value analogue, achieving far larger magnitudes. The point of Lemma A.11 is to take ratios of these magnitudes and relate them directly to the coefficients of our optimal vector  $\hat{a}$ . Using our lemma, we will prove our theorem, relegating the proof of the lemma for later.

*Proof of Theorem 4.1.* We will show that inequality (A.9) implies our Theorem 4.1.

Note

$$\begin{aligned} & \frac{s^+}{s^-} < \frac{s_k^+}{s_k^-} \\ \iff & \frac{1 - s^-/s^+}{1 + s^-/s^+} < \frac{1 - s_k^-/s_k^+}{1 + s_k^-/s_k^+} \\ \iff & \frac{s^+ - s^-}{s^+ + s^-} < \frac{s_k^+ - s_k^-}{s_k^+ + s_k^-}, \end{aligned}$$

where we have taken advantage of the fact that  $0 < \frac{s^-}{s^+}, \frac{s_k^-}{s_k^+} < 1$ .

When  $N$  is even our last inequality becomes directly  $\frac{s^o}{s^e} < \frac{s_k^o}{s_k^e}$ . When  $N$  is odd, the inequality becomes  $\frac{s^e}{s^o} < \frac{s_k^e}{s_k^o}$  and taking the reciprocal of both sides, we obtain  $\frac{s^o}{s^e} > \frac{s_k^o}{s_k^e}$ . Now, let  $k < N$  be the index of any row. Then the expressions  $\frac{g^o}{g^e}, \frac{g_k^o}{g_k^e}$  must contain exactly one negative entry in each fraction, and

$$\begin{aligned} & \frac{s^o}{s^e} < \frac{s_k^o}{s_k^e} \\ \iff & \frac{|g^o|}{|g^e|} < \frac{|g_k^o|}{|g_k^e|} \\ \iff & \frac{g^o}{g^e} > \frac{g_k^o}{g_k^e}. \end{aligned}$$

(The inequality is reversed for odd  $N$ .) If  $N$  is even, then  $g^e$  is positive; otherwise  $g^o$  is negative. In either case, if  $k$  is even, our previous inequality is equivalent to

$$(A.10) \quad g^o g_k^e > g^e g_k^o$$

$$(A.11) \quad \iff (g^{oo} g^{ee} - g^{oe} g^{eo})^{-1} (-g^e g_k^o + g^o g_k^e) > 0$$

$$(A.12) \quad \iff \lambda_1 g_k^o + \lambda_2 g_k^e > 0$$

by (A.8). If  $k$  is odd, this inequality is

$$(A.13) \quad \lambda_1 g_k^o + \lambda_2 g_k^e < 0.$$

But

$$B \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_2 \end{pmatrix},$$

so the  $k$ th element of  $a$  is given by

$$(A.14) \quad \lambda_1 g_k^o + \lambda_2 g_k^e$$

so in fact our two inequalities (A.12), (A.13) together give the statement of our main Theorem 4.1. ■

*Proof of Lemma A.11.* Now onto the inequality. We can rewrite the inequality (A.9) as

$$\begin{aligned} & \frac{s^+}{s^-} < \frac{s_i^+}{s_i^-} \\ \iff & \frac{\sum_{i=1}^N |r_i|}{\sum_{i=1}^N r_i} < \frac{\sum_{j=1}^N |G_{kj}|}{\left| \sum_{j=1}^N G_{kj} \right|} \\ \iff & \frac{\sum_{i=1}^N |r_i|}{N(p+N)} < \frac{\sum_{j=1}^N |G_{kj}|}{|r_k|} \\ \iff & |r_k| \sum_{i=1}^N |r_i| < N(p+N) \sum_{j=1}^N |G_{kj}| \\ \iff & \sum_{i=1}^N (p+i+j-1) |G_{kj}| < N(p+N) \sum_{j=1}^N |G_{kj}|. \end{aligned}$$

But  $p+i+j-1 < 2N-1+p$ . When  $N \geq 2$ ,  $2N-1 < N^2$  and so  $p+i+j-1 < N^2+p < N(p+N)$ , and thus our inequality is proven. ■

**A.3. Proof of concentration and pointwise convergence in optimal polynomial pairs.**

Here we prove Theorem 4.3, relating the asymptotic decay of the polynomial coefficients and consequently their asymptotic high concentration.

Note that for a fixed degree, any polynomial with positive coefficients that add up to 1, the function is bounded from below by  $x^N$  for every  $x$ . We will exploit this fact to show the first part of Theorem 4.3, that is, that the coefficient to  $x^k$  for a fixed  $k$  for the optimal polynomial  $p_N$  vanishes as  $N \rightarrow \infty$ . We will then show that this implies that the optimal polynomial pairs derived from minimizing quadrature loss converge pointwise to an indicator function at 1—that is, the polynomials converge pointwise to 0 (except at  $x = 1$ ).

We will require the following computational lemma, whose proof will be given in section A.5.



**Lemma A.12.** *Let  $N$  be sufficiently large—say, larger than both  $p$  and 2, and fix  $j \leq N$ . Then the expression*

$$(A.15) \quad (N(p+N) - (p+i+j-1))|G_{ij}|$$

*is increasing in  $i$  on the interval  $[1, N/2]$ .*

In fact, it will be convenient to use the shorthand  $D_{ij} = (N(p+n) - (p+i+j-1))G_{ij}$ . Our shorthand does not have the entry term  $G_{ij}$  within absolute value signs but it will typically not matter. Now we prove the part of the theorem that claims that the coefficients to a fixed order vanish asymptotically.

*Proof of Theorem 4.3(a).* Recall from earlier that  $a_k$  is given by

$$(A.16) \quad a_k = \lambda_1 g_k^o + \lambda_2 g_k^e$$

$$(A.17) \quad = \lambda_G (-g^e g_k^o + g^o g_k^e)$$

$$(A.18) \quad = \frac{-g^e g_k^o + g^o g_k^e}{g^{oo} g^{ee} - (g^{oe})^2}.$$

Recall that we may write our denominator

$$g^{oo} g^{ee} - (g^{oe})^2 = N(p+N)g^{oo} - d^2,$$

where  $d$  was the absolute value of the sums of all elements in all odd rows. Using Corollary A.7 we may rewrite the denominator of A.18 as

$$\begin{aligned} & g^{oo} g^{ee} - (g^{oe})^2 \\ &= N(p+N) \sum_{i,j \text{ odd}} G_{ij} - \sum_{i,j \text{ odd}} r_i r_j \\ &= N(p+N) \sum_{i,j \text{ odd}} G_{ij} - \sum_{i,j \text{ odd}} (p+i+j-1)G_{ij} \\ &= \sum_{i,j \text{ odd}} (N(p+N) - (p+i+j-1))G_{ij} = \sum_{i,j \text{ odd}} D_{ij}. \end{aligned}$$

Now consider the numerator of A.18. Noting that  $g_k^e = -g_k^o + r_k$ , we can rewrite our numerator as

$$\begin{aligned} -g^e g_k^o + g^o g_k^e &= g^e (-g_k^o) + g^o (-g_k^o + r_k) \\ &= -g_k^o (g^e + g^o) + g^o r_k \\ &= -(N(p+N))g_k^o + g^o r_k \\ &= -(N(p+N) \sum_{j \text{ odd}} G_{kj} + \left( \sum_{j \text{ odd}} r_j \right) r_k) \\ &= \sum_{j \text{ odd}} (p+k+j-1 - N(p+N))G_{kj}. \end{aligned}$$

Now, the sign of summands is dependent entirely on  $k$ . Since  $j$  ranges over odd values,  $G_{kj}$  is negative if  $k$  is even, and positive when  $k$  is odd. But the summands have the same sign, so since we are concerned only with magnitude we may flip the sign for convenience. In other words it suffices to show that

$$(A.19) \quad (-1)^k \frac{-g^e g_k^o + g^o g_k^e}{g^{oo} g^{ee} - (g^{oe})^2} = \frac{\sum_{j \text{ odd}} D_{kj}}{\sum_{i,j \text{ odd}} D_{ij}}$$

decays to 0 as  $N \rightarrow \infty$ . Take  $N > 4k$ , and let  $S = \{i \text{ odd} : N/4 \leq i \leq N/2\}$ . First, note that this implies  $N > 4$ , and that  $|S|$  is approximately  $N/8$ . Certainly, for large  $N$  we may say that  $|S| > N/16$ . Furthermore, since  $k < N/4$ , by Lemma A.12, we have that for any  $i \in S$ ,

$$(A.20) \quad \left| \sum_{j \text{ odd}} D_{ij} \right| > \left| \sum_{j \text{ odd}} D_{kj} \right|.$$

Thus from (A.19), we may write

$$\begin{aligned} & \frac{\sum_{j \text{ odd}} D_{kj}}{\sum_{i,j \text{ odd}} D_{ij}} \\ & < \frac{\sum_{j \text{ odd}} D_{kj}}{\sum_{j \text{ odd}, i \in S} D_{ij}} \\ & < \frac{\sum_{j \text{ odd}} D_{kj}}{\frac{N}{16} \sum_{j \text{ odd}} D_{kj}} = \frac{16}{N} \end{aligned}$$

which clearly decays to 0 as  $N \rightarrow \infty$ . ■

We now prove that as we increase the degree  $N$  of our polynomials, the optimal pairs converge to a delta function at 1.

*Proof of Theorem 4.3(b).* Let  $\epsilon > 0$ . First, there is some  $M$  such that  $x^M < \frac{\epsilon}{4}$ . Recall that any  $\hat{a}^{(N)}$  corresponds to a pair of polynomials with coefficients  $\hat{a}^{(N)}(k)$ . It suffices to show that the polynomial given by  $q_N(x) = \sum_{k=1}^N |\hat{a}^{(N)}(k)| x^k$  converges pointwise to 0 at every  $x \neq 1$ , since the optimal odd/even polynomials are bounded by this polynomial. By Theorem 4.3(a), there is some  $L$  such that for every  $N > \max(M, L)$ ,  $\hat{a}^{(N)}$  satisfies  $\max(a_{N,1}, a_{N,2}, \dots, a_{N,M-1}) < \frac{\epsilon(1-x)}{2x}$ . Selecting such an  $M, N$ , let  $c = \max(a_{N,1}, a_{N,2}, \dots, a_{N,M-1})$ . Now, using the facts above and the fact that  $\sum_{k=1}^N |\hat{a}^{(N)}(k)| = 2$ ,

$$\begin{aligned} q_N(x) &= \sum_{k=1}^{M-1} a_{N,k} x^k + \sum_{k=M}^N a_{N,k} x^k \\ &< c \sum_{k=1}^{M-1} x^k + \sum_{k=M}^N a_{N,k} x^M \\ &< \frac{\epsilon(1-x)}{2} \sum_{k=1}^{\infty} x^k + x^M \sum_{k=M}^N a_{N,k} \\ &< \frac{\epsilon(1-x)}{2x} \frac{x}{1-x} + 2x^M < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$
■

as needed.

**A.4. Proof of Theorem 4.4: Asymptotic conjugacy loss vanishes.** Let us fix some  $N$  and simply write  $\hat{a}$  as the corresponding solution. We proceed by deriving an alternative expression for the objective function using our previous results and then prove some facts about the growth. Recall

$$\begin{aligned} \frac{1}{2}a^T Ha &= \frac{1}{2}\lambda^T B^T H^{-1} B \lambda \\ &= \frac{1}{2}\lambda^T K \lambda, \quad \text{where } K = \begin{pmatrix} g^{oo} & g^{oe} \\ g^{eo} & g^{ee} \end{pmatrix}. \end{aligned}$$

But  $\lambda = K^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \lambda_G \begin{pmatrix} -g^e \\ g^o \end{pmatrix}$ , where  $\lambda_G = \frac{1}{g^{oo}g^{ee} - g^{eo}g^{oe}} > 0$ , so noting the symmetry of  $K$ , we have

$$\begin{aligned} \frac{1}{2}a^T Ha &= \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix}^T K^{-1} K \lambda \\ &= \frac{\lambda_G}{2} (-1 \ 1) \begin{pmatrix} -g^e \\ g^o \end{pmatrix} \\ &= \frac{g^e + g^o}{2(g^{oo}g^{ee} - g^{eo}g^{oe})} \\ &= \frac{N(p + N)}{2(g^{oo}g^{ee} - (g^{oe})^2)}. \end{aligned}$$

Now, let  $d = |g^{oe}| - |g^{oo}|$ . Then since  $N(p + N) = g^{ee} + g^{oo} + g^{oe} + g^{eo} = g^{oo} + g^{ee} - 2|g^{oe}|$ , we may write  $g^{oo} = |g^{oe}| - d$ ,  $g^{ee} = |g^{oe}| + d + N(p + N)$ . Thus

$$\begin{aligned} g^{oo}g^{ee} - g^{eo}g^{oe} &= (|g^{oe}| + d + N(p + N))(|g^{oe}| - d) - |g^{oe}|^2 \\ &= (|g^{oe}| + d)(|g^{oe}| - d) + N(p + N)(|g^{oo}|) - |g^{oe}|^2 \\ &= N(p + N)(g^{oo}) - d^2. \end{aligned}$$

We now examine the growth of  $g^{oo}$  and  $d^2$ . The value  $d$  is simply a sign change times the sum of the odd rows. We have

$$d = (-1)^N \sum_{i \text{ odd}}^N \frac{\prod_{k=0}^{N-1} (p + i + k)}{(i-1)!(N-i)!}$$

and thus we may write

$$\begin{aligned} d^2 &= \sum_{i,j \text{ odd}}^N \frac{\prod_{k=0}^{N-1} (p + i + k)(p + j + k)}{(i-1)!(N-i)!(j-1)!(N-j)!} \\ &= \sum_{i,j \text{ odd}}^N (p + i + j - 1) |G_{ij}|. \end{aligned}$$

On the other hand, using the fact that  $g^{oo}$  is a sum of only positive entries, we see that

$$N(p + N)g^{oo} = \sum_{i,j \text{ odd}}^N N(p + N)|G_{ij}|,$$

$$g^{oo} = \sum_{i,j \text{ odd}}^N \frac{1}{p + i + j - 1} \frac{\prod_{k=0}^{N-1} (p + i + k)(p + j + k)}{(i - 1)!(N - i)!(j - 1)!(N - j)!}.$$

So we can write the denominator of our objective function value as

$$\begin{aligned} & 2(g^{oo}g^{ee} - (g^{oe})^2) \\ &= 2(N(p + N)(g^{oo}) - d^2) \\ &= 2 \sum_{i,j \text{ odd}}^N (N(p + N) - (p + i + j - 1))|G_{ij}| \\ &= \sum_{i,j \text{ odd}}^N \left( \frac{2N(p + N)}{p + i + j - 1} - 1 \right) (p + i + j - 1)|G_{ij}| \\ &\geq \sum_{i,j \text{ odd}}^N (N - 1)(p + i + j - 1)|G_{ij}| \\ &= \sum_{i,j \text{ odd}}^N (N - 1) \frac{\prod_{k=0}^{N-1} (p + i + k)(p + j + k)}{(i - 1)!(N - i)!(j - 1)!(N - j)!} \\ &\geq \sum_{i,j \text{ odd}}^N (N - 1) \frac{(N - 1)!^2}{(i - 1)!(N - i)!(j - 1)!(N - j)!} \\ &= \sum_{i,j \text{ odd}}^N (N - 1) \binom{N - 1}{i - 1} \binom{N - 1}{j - 1} \\ &= (N - 1) \left( \sum_{i' \text{ even}}^{N-1} \binom{N - 1}{i'} \right) \left( \sum_{j' \text{ even}}^{N-1} \binom{N - 1}{j'} \right) \\ &= (N - 1)(2^{N-2})(2^{N-2}) = (N - 1)2^{2N-4}. \end{aligned}$$

Since our numerator grows like  $O(N^2)$ , we are done; we see clearly that our fraction vanishes as  $N \rightarrow \infty$ .

**A.5. Proof of lemmas.** In this section we prove Lemma A.12. First we require the following auxiliary lemma.

**Lemma A.13.** *Let  $G$  be the inverse of a Hilbert matrix of any order of size  $N$ . Fix some  $0 \leq j \leq N$ . Then for  $i \in [1, N/2]$ ,*

$$(A.21) \quad \frac{|G_{i+1,j}|}{|G_{i,j}|} \geq \binom{N + p + i}{j + p + i} \left( \frac{p + i + j - 1}{p + i} \right).$$

Furthermore, both terms on the right-hand side of (A.21) are greater than or equal to 1, and for at least one of them, the numerator is larger than the denominator by at least 1.

*Proof.* First we express  $|G_{ij}|$  as a function of  $i$ . Using (A.1) and gathering together terms that do not depend on  $i$  in an absorbing coefficient  $C(N, j, p)$ , we see that

$$|G_{ij}| = C(N, j, p) \frac{1}{p+i+j-1} \left( \frac{\prod_{k=0}^{N-1} (p+i+k)}{(i-1)!(N-i)!} \right)$$

and therefore

$$\begin{aligned} & \frac{|G_{i+1,j}|}{|G_{i,j}|} \\ &= \frac{p+i+j-1}{p+i+j} \frac{\left( \frac{\prod_{k=0}^{N-1} (p+i+k)}{(i-1)!(N-i)!} \right)}{\left( \frac{(i)!(N-i-1)!}{\prod_{k=0}^{N-1} (p+i+1+k)} \right)} \\ &= \frac{(p+i+j-1)(N-i)!(i-1)! \prod_{k=0}^{N-1} (p+i+k+1)}{(p+i+j)(i)!(N-i-1)! \prod_{k=0}^{N-1} (p+i+k)} \\ &= \frac{(p+i+j)(p+i+N)(N-i)}{(p+i+j-1)(p+i)(i)} \\ &\geq \left( \frac{p+i+N}{p+i+j} \right) \left( \frac{p+i+j-1}{p+i} \right) \end{aligned}$$

since for  $i \leq N/2$ ,  $\frac{N-i}{i} \geq 1$ . Now, clearly, both fractions in our last product are at least 1. Moreover, if  $j < N$  then the numerator in the first fraction is greater than its denominator by at least 1. But if  $j = N$  then  $p+i+j-1 = p+i+N-1 > p+i$  strictly by at least 1, and so we have shown all claims in Lemma A.13. ■

Now we may prove our main lemma giving our increasing expression.

*Proof of Lemma A.12.* Let us write (A.15) as a function of  $i$ , with

$$f(i) = (N(p+N) - (p+i+j-1))|G_{ij}|.$$

We need that for  $i \in \{1, \dots, \lfloor N/2 \rfloor\}$ ,  $\frac{f(i+1)}{f(i)} > 1$ . For notational purposes, let  $A = N(p+N) - (p+i+j) = f(i+1)/|G_{i+1,j}|$ . Now, using inequality (A.21) from the result of Lemma (A.13), we may compute

$$\begin{aligned} \frac{f(i+1)}{f(i)} &= \frac{A}{A+1} \frac{|G_{i+1,j}|}{|G_{i,j}|} \\ &\geq \frac{A}{A+1} \frac{(N+p+i)(p+i+j-1)}{(j+p+i)(p+i)}. \end{aligned}$$

Lemma A.13 also states that at least one of  $\left( \frac{N+p+i}{j+p+i} \right)$ ,  $\left( \frac{p+i+j-1}{p+i} \right)$  is strictly larger than 1, that is, the numerator is greater than the denominator by at least 1. Let  $B$  be this denominator

(or if both fractions are greater than 1, pick either). Now,  $A+1 = N(p+N) - (p+i+j-1) > j+p+i \iff N^2 + Np > 2p + 2i + 2j - 1$ . But  $2i + 2j - 1 < 4N$ , so if  $N > 4$  then  $N^2 + Np > 4N + 4p > 2i + 2j - 1 + 2p$ . Similarly,  $N(p+N) - (p+i+j-1) > p+i$  under weak conditions on  $N$ , and so in all cases,  $A+1 > B$  thus we can rewrite our inequality

$$\begin{aligned} \frac{f(i+1)}{f(i)} &\geq \frac{A}{A+1} \frac{(N+p+i)(p+i+j-1)}{(j+p+i)(p+i)} \\ &\geq \frac{A}{A+1} \frac{B+1}{B} > 1 \end{aligned}$$

as needed. ■

## REFERENCES

- [1] F. AGUET, S. GEISSBUHLER, I. MARKI, T. LASSER, AND M. UNSER, *Super-resolution orientation estimation and localization of fluorescent dipoles using 3-D steerable filters*, Opt. Express, 17 (2009), pp. 6829–6848.
- [2] A. ANDREARCZYK AND A. DEPEURSINGE, *Rotational 3D Texture Classification Using Group Equivariant CNNs*, preprint, <https://arxiv.org/abs/1810.06889> (2018).
- [3] V. ANDREARCZYK, J. FAGEOT, V. OREILLER, X. MONTET, AND A. DEPEURSINGE, *Exploring local rotation invariance in 3D CNNs with steerable filters*, in Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning, M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal, and T. Vercauteren, eds., Proc. Mach. Learn. Res., (PMLR), 102 (2019), pp. 15–26, <https://proceedings.mlr.press/v102/andrearczyk19a.html>.
- [4] E. L. BASOR, Y. CHEN, AND H. WIDOM, *Determinants of Hankel matrices*, J. Funct. Anal., 179 (2001), pp. 214–234.
- [5] G. BLIN, D. SADURSKA, M. R. PORTERO, J. CHEN, NAND WATSON, AND L. S, *Nessys: A new set of tools for the automated detection of nuclei within intact tissues and dense 3D cultures*, PLoS Biol., 17 (2019), e3000388.
- [6] A. C. BOVIK, M. CLARK, AND W. S. GEISLER, *Multichannel texture analysis using localized spatial filters*, IEEE Trans. Pattern Anal. Mach. Intell., 12 (1990), pp. 55–73.
- [7] N. CHENOARD AND M. UNSER, *3D steerable wavelets in practice*, IEEE Trans. Image Process., 21 (2012), pp. 4522–4533.
- [8] A. R. COLLAR, *XIX.—On the reciprocation of certain matrices*, Proc. Roy. Soc. Edinburgh, 59 (1940), pp. 195–206.
- [9] J. G. DAUGMAN, *Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters*, J. Opt. Soc. Amer. A, 2 (1985), pp. 1160–1169.
- [10] A. DEPEURSINGE AND J. FAGEOT, *Biomedical texture operators and aggregation functions: A methodological review and user’s guide*, in Biomedical Texture Analysis, A. Depeursinge, O. Al-Kadi, and J. Mitchell, eds., Elsevier MICCAI Soc. Book Ser., Academic Press London, 2017, pp. 55–94.
- [11] K. G. DERPANIS AND J. M. GRYN, *Three-dimensional nth derivative of Gaussian separable steerable filters*, in IEEE International Conference on Image Processing 2005, Vol. 3, IEEE, Piscataway, NJ, 2005, pp. III–553.
- [12] M. FELSBERG AND G. SOMMER, *The monogenic signal*, IEEE Trans. Signal Process., 49 (2001), pp. 3136–3144, <https://doi.org/10.1109/78.969520>.
- [13] M. FIEDLER, *Notes on Hilbert and Cauchy matrices*, Linear Algebra Appl., 432 (2010), pp. 351–356.
- [14] J. FRANK, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in their Native State*, Oxford University Press, Oxford, 2006.
- [15] W. T. FREEMAN, *Steerable Filters and Local Analysis of Image Structure*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 1992.
- [16] W. T. FREEMAN AND E. H. ADELSON, *The design and use of steerable filters*, IEEE Trans. Pattern Anal. Mach. Intell., 13 (1991), pp. 891–906.

- [17] J. GIL, *Polarimetric characterization of light and media*, European Phys. J. Appl. Phys., 40 (2007), pp. 1–47.
- [18] G. GORDON AND R. TIBSHIRANI, *Karush-Kuhn-Tucker Conditions*, Class notes, Carnegie Mellon University, 2012.
- [19] D. J. HEEGER, *Optical flow using spatiotemporal filters*, Inte. J. Comput Vis, 1 (1988), pp. 279–302.
- [20] C. HUANG AND Y. CHEN, *Motion estimation method using a 3D steerable filter*, Image Vis. Comput., 13 (1995), pp. 21–32.
- [21] I. KALLIOMÄKI AND J. LAMPINEN, *On steerability of Gabor-type filters for feature detection*, Pattern Recognit. Lett., 28 (2007), pp. 904–911.
- [22] K. KRAJSEK AND R. MESTER, *A unified theory for steerable and quadrature filters*, in Advances in Computer Graphics and Computer Vision, Springer, Berlin, 2007, pp. 201–214.
- [23] B. LI, Q. WANG, AND G. LEE, *Filtra: Rethinking steerable CNN by filter transform*, in Proceedings of the 38th International Conference on Machine Learning, M. Meila and T. Zhang, eds., Proc. Mach. Learn. Res., (PMLR), 139 (2021), pp. 6515–6522.
- [24] J. D. C. Y. LIAO M, CAO E, *Structure of the TRPV1 ion channel determined by electron cryo-microscopy*, Nature, 504 (2013), pp. 107–12.
- [25] H. MCKEAN AND H. DYM, *Fourier Series and Integrals*, Academic, Boston, 1972.
- [26] J. R. MOVELLAN, *Tutorial on Gabor Filters*, manuscript.
- [27] A. PAPOULIS, *The Fourier Integral and its Applications*, McGraw-Hill, New York, 1962.
- [28] S. PAUL, S. BHATTACHARYA, AND S. GUPTA, *Spatiotemporal colorization of video using 3D steerable pyramids*, IEEE Trans. Circuits Systems Video Technol., 27 (2017), pp. 1605–1619.
- [29] A. POTAMIANOS AND P. MARAGOS, *A comparison of the energy operator and the Hilbert transform approach to signal and speech demodulation*, Signal Process., 37 (1994), pp. 95–120.
- [30] P. B. ROSENTHAL AND R. HENDERSON, *Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy*, J. Mol. Biol., 333 (2003), pp. 721–745.
- [31] A. SEMECHKO, *S2-Sampling-Toolbox*. <https://github.com/AntonSemechko/S2-Sampling-Toolbox> (2021).
- [32] F. J. SIMONS, F. DAHLEN, AND M. A. WIECZOREK, *Spatiospectral concentration on a sphere*, SIAM Rev., 48 (2006), pp. 504–536.
- [33] T. SMIDT, *Euclidean symmetry and equivariance in machine learning*, Trends Chem., 3 (2021), pp. 82–85.
- [34] R. B. SMITH, *Two theorems on inverses of finite segments of the generalized Hilbert matrix*, Math. Tables Other Aids Comput., 13 (1959), pp. 41–43.
- [35] P. C. TEO, *Theory and Applications of Steerable Functions*. Ph. D. thesis, Stanford University, Palu Alto, CA, 1998.
- [36] F. TIAN TIAN AND Y. HONGBIN, *A novel shape from focus method based on 3D steerable filters for improved performance on treating textureless region*, Opt. Commun., 410 (2018), pp. 254–261.
- [37] J. VILAS, J. HEYMANN, H. D. TAGARE, E. RAMÍREZ-APORTELA, J. M. CARAZO, AND C. SORZANO, *Local resolution estimates of cryoem reconstructions*, Curr. Opinion Struct. Biol., 64 (2020), pp. 74–78.
- [38] J. L. VILAS, H. D. TAGARE, J. VARGAS, J. M. CARAZO, AND C. O. S. SORZANO, *Measuring local-directional resolution and local anisotropy in cryo-em maps*, Nature Commun., 11 (2020), pp. 1–7.
- [39] M. WEILER, M. GEIGER, M. WELING, W. BOOMSMA, AND T. COHEN, *3D steerable CNNs: Learning rotationally equivariant features in volumetric data*, in Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS’18), Curran Associates Inc., Red Hook, NY, pp. 10402–10413.
- [40] W. YU, K. DANILIDIS, AND G. SOMMER, *A new 3D orientation steerable filter*, in Mustererkennung 2000, G. Sommer, N. Krüger, and C. Perwass, eds., Springer, Berlin, 2000, pp. 203–212.
- [41] W. YU, G. SOMMER, AND K. DANILIDIS, *3D-orientation signatures with conic kernel filtering for multiple motion analysis*, in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001, Vol. 1, IEEE Computer Society, Los Alamitos, CA, 2001, pp. I–I.