

## ANGULAR ACCURACY OF STEERABLE FEATURE DETECTORS \*

ZSUZSANNA PÜSPÖKI <sup>†</sup>, JULIEN FAGEOT <sup>†</sup>, ARASH AMINI <sup>‡</sup>, JOHN PAUL WARD <sup>§</sup>, AND MICHAEL UNSER <sup>†</sup>

**Abstract.** The detection of landmarks or patterns is of interest for extracting features in biological images. Hence, algorithms for finding these keypoints have been extensively investigated in the literature, and their localization and detection properties are well known. In this paper, we study the complementary topic of local orientation estimation, which has not received similar attention. Simply stated, the problem that we address is the following: estimate the angle of rotation of a pattern with steerable filters centered at the same location, where the image is corrupted by colored isotropic Gaussian noise. For this problem, we propose an estimator formulated as linear combinations of circular harmonics with given radial profiles. We prove that the proposed estimator is unbiased. This property allows us to use a statistical framework based on the Cramér-Rao lower bound (CRLB) to study the limits on the accuracy of the corresponding class of estimators. We aim at evaluating the performance of detection methods based on steerable filters in terms of angular accuracy (as a lower bound), while considering the connection to maximum likelihood estimation. Beyond the general results, we analyze the asymptotic behaviour of the lower bound in terms of the order of steerability and propose an optimal subset of components that minimizes the bound. We define a mechanism for selecting optimal subspaces of the span of the detectors. These are characterized by the most relevant angular frequencies. Finally, we project our template to the span of circular harmonics with given radial profiles and experimentally show that the prediction accuracy achieves the predicted CRLB. As an extension, we also consider steerable wavelet detectors.

**Key words.** Cramér-Rao lower bounds, estimation of orientation, steerable filters, steerable wavelets

**AMS subject classifications.** 62H12, 62M40, 42C40

**1. Introduction.** Steerable feature detectors [39] are popular tools for the quantitative analysis of images where the information that corresponds to the notions of rotation, directionality, and orientation is a key component. Their application area is continuously growing: from nano to macro, from biomedical imaging [32] to astronomy [47], from material sciences [5] to aerial and satellite imaging [20], and so on. The analysis of local directional patterns also includes the detection of ridges and junctions of any order; applications can be found in [44, 62, 27, 37].

Typically, concerning steerable detectors, a two-step algorithm is used for identifying features of interest and determine their orientation. The first step is to find their location in an image. Then, the second step is to steer the detector centred at the given keypoints to identify their exact orientation. Our approach is in line with this practice. Steerable functions (a conventional detector or a wavelet frame of  $L_2(\mathbb{R}^2)$ ) are used to identify the center of a given pattern in an image. Then, we locally steer those filters that are centered at the same point. The paper is devoted to the analysis of this second step: studying the problem of extracting the local orientation of patterns in images. These local patterns, in particular, include junctions or crossing points with  $N$ -fold angular symmetry of any order.

In this paper, we examine the behavior of steerable filters in orientation detection, providing a lower bound on the performance of any detector that is in the span of a given

---

\*The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement N°267439 and the Hasler Foundation.

<sup>†</sup>Biomedical Imaging Group, Ecole polytechnique fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland.

<sup>‡</sup>Department of Electrical Engineering and Advanced Communications Research Institute (ACRI), Sharif University of Technology, Tehran, Iran.

<sup>§</sup>Affiliations while contributing to this work: Department of Mathematics, University of Central Florida, Orlando, FL 32816, USA and Department of Mathematics, North Carolina A&T State University, Greensboro, NC 27411, USA.

family of steerable functions. In particular, we focus on steerable filters that are linear combinations of circular harmonics with given radial profiles. We propose an estimator based on such steerable filters and show that the estimator is unbiased. This allows us to analyse the limits on the error of the estimation in a statistical framework using the Cramér-Rao lower bound (CRLB). We investigate the connection to maximum likelihood estimation and show that, under some assumptions, our measurement functions correspond to the maximum likelihood estimator. We define the best subset of harmonics as the one that minimizes the CRLB for a given class of detectors and we provide a criterion for obtaining it. We also analyze the asymptotic behaviour of the lower bound in terms of the number of harmonics. This gives a theoretical limit on the precision of the estimation one can obtain with a given class of estimators. We experimentally show that the actual performance of the given steerable detectors follows the predicted theoretical bounds.

We focus on the continuous domain setup, because it allows for compact and intuitive formulation. The discretisation of the underlying filters can be achieved using standard techniques. In practice, this step affects multiscale steerable analysis in two ways: It limits the range of angular harmonics due to angular aliasing and limits the practical range of scales. However, as shown in Figure 2, one can achieve good estimation by selecting an appropriate template with a small amount of polar-separable low order harmonics.

**1.1. State of the Art.** Classical methods to detect orientations are based on gradient information (*e.g.*, Canny edge detector [3]), on directional derivatives [62], and on the structure tensor [19]. Variations of the latter method can be found in [24] and [2]. To capture higher-order directional structures, the Hessian and higher-order derivatives can be used [12, 1]. While simple and computationally efficient, these methods have drawbacks: they only take into account one specific scale and the estimation of the orientation can be overly sensitive to noise.

Alternatively, directional pattern matching is also commonly used. It often relies on the discretization of the orientation and thus demands a tradeoff between accurate results and computational cost. An important exception to this are the steerable filters, where one may perform arbitrary (continuous) rotations and optimizations with a substantially reduced computational overhead.

The basics of steerability were formulated by Danielsson [6] in the eighties and Freeman and Adelson in the early nineties [13, 14, 15]. In [50], Simoncelli and Freeman proposed a new take on steerable filters: the steerable pyramid. The goal of their design was to combine steerability with a multiscale detection scheme. This pioneering work had many successful applications: contour detection [34], orientation analysis [48], and texture analysis and synthesis [38]. Other applications of steerable filters were presented in [23, 63, 17, 21, 46, 61, 8].

In [30], multisteerable filters were designed and adjusted to the patterns of interest to determine the precise angular distribution of coinciding branches. There, the detection and classification of polar-separable patterns (including junctions) rely on a classical structure-tensor scheme, complemented by the multisteerable filters. In [58], Unser and Chenouard proposed a unifying parametric framework for 2D steerable wavelet transforms. The goal of [58] was to combine steerability with tight wavelet frames and propose a general scheme to design such wavelets. In [40] and [41], the authors designed wavelets that can serve as a basis to detect features such as junctions and local symmetry points. An application of steerable wavelets for texture learning was presented in [9]. Here, steerable templates are represented in the Fourier domain using circular harmonics. The order of the detector is given by the number of harmonics, which also affects the performance of the detection.

An interesting instance of this direction of research is the monogenic wavelet transform that gives access to the local amplitude and the local phase of the image, additionally to the information on the local orientation [11, 16, 31, 53].

The extension of the steerable wavelet design based on the Riesz transform for higher dimensions, along with potential biomedical applications, are presented in [4]. Spherical harmonics, which are the 3D counterparts of circular harmonics, have also been used to represent and detect features and shapes in 3D [42, 51]. There, the authors also aim to identify symmetric structures; however, they search for symmetries in an entire 3D shape, not locally.

Over the last decades, the notion of steerability has been further extended and generalized to a broader group of geometric transformations. A general approach to construct scalable functions was first given by Simoncelli et al. [49]. Perona generalized this to arbitrary compact transformations without requiring the group properties [35]. Michaelis and Sommer set the basis of steerability based on Abelian Lie groups [29]. Teo, in his doctoral dissertation [55] and a series of papers with Hel-Or [56, 57], unified the existing theories and provided a solid mathematical framework for the study of transformability based on the theory of Lie groups and Lie algebras in general. A unified theory of steerable and quadrature filters based on Lie groups was provided by Krajsek and Mester [26].

The authors are not aware of any previous studies that thoroughly address the angular accuracy of steerable feature detectors.

**1.2. Roadmap.** The paper is organized as follows: In Section 2.2, we provide a description of our directional pattern model and formulate the estimation problem. In Section 2.3, we discuss our estimation strategy and propose a reference class of estimators. In Section 2.4, we identify the probability law of the measurement vector and introduce the Cramér-Rao lower bound (CRLB). In Section 2.5, we consider the connection to maximum likelihood estimation and demonstrate that our orientation estimation framework is unbiased, which is crucial for the application of the CRLB. In Section 3, we provide the CRLB for orientation estimation based on steerable filters. We consider the case when the number of angular frequencies is finite (Section 3.1) and also the asymptotic case for a fixed radial profile (Section 3.3). In Section 3.2, we examine the question of how to choose the best subset of harmonics to achieve the lowest CRLB. Section 4 contains experiments related to the CRLB. In Section 5, we extend our results to wavelets.

## 2. The Estimation Problem at a Glance.

**2.1. Notations.** We use  $f(\mathbf{x})$  with  $\mathbf{x} \in \mathbb{R}^2$  and  $f(r, \theta)$  with  $r \in \mathbb{R}^+$ ,  $\theta \in [0, 2\pi)$  to denote the Cartesian and polar representations of the same 2D function  $f$ , respectively. Corresponding notations in the Fourier domain are  $\hat{f}(\boldsymbol{\omega})$  and  $\hat{f}(\omega, \varphi)$  with  $\boldsymbol{\omega} \in \mathbb{R}^2$  and  $\omega \in \mathbb{R}^+$ ,  $\varphi \in [0, 2\pi)$ . The Fourier transform of a Lebesgue-integrable, finite-energy function  $f \in L_1(\mathbb{R}^2) \cap L_2(\mathbb{R}^2)$  is denoted by  $\mathcal{F}\{f\} = \hat{f}$  and computed according to

$$(2.1) \quad \hat{f}(\boldsymbol{\omega}) = \int_{\mathbb{R}^2} f(\mathbf{x}) e^{-j\langle \mathbf{x}, \boldsymbol{\omega} \rangle} d\mathbf{x},$$

where  $\langle \mathbf{x}, \boldsymbol{\omega} \rangle = x_1 \omega_1 + x_2 \omega_2$  is the usual scalar product on  $\mathbb{R}^2$ .

The 2D matrix of rotation by  $\theta_0$  is denoted by

$$(2.2) \quad \mathbf{R}_{\theta_0} = \begin{pmatrix} \cos(\theta_0) & -\sin(\theta_0) \\ \sin(\theta_0) & \cos(\theta_0) \end{pmatrix}.$$

TABLE 1  
Construction of measurement functions.

$q_\alpha$	radial profile	harmonics
Conventional detector $\alpha = n \in \mathbb{Z}$	$\hat{\xi}_n(\omega) = \hat{h}(\omega)$ fixed	$e^{jn\theta}$
Wavelet detector $\alpha = (n, i) \in \mathbb{Z}^2$	$\hat{\xi}_{(n,i)}(\omega) = 2^i \hat{h}(2^i \omega)$ fixed	$e^{jn\theta}$
General detector $\alpha \in \mathbb{Z}^d$	$\hat{\xi}_\alpha(\omega) = \hat{h}(\omega)$ adaptive	$e^{jn_\alpha \theta}$

**2.2. The Estimation Problem.** Our interest in this paper lies in detecting a rotated pattern  $J(\mathbf{x})$  from a noisy signal of the form

$$(2.3) \quad I(\mathbf{x}) = J(\mathbf{R}_{-\theta^*}(\mathbf{x} - \mathbf{x}_0)) + S(\mathbf{x})$$

using a steerable filter bank. Here,  $J$  denotes the general shape of the pattern of interest and  $J(\mathbf{R}_{-\theta^*}(\mathbf{x} - \mathbf{x}_0))$  its rotated version around location  $\mathbf{x}_0$  with an unknown angle  $\theta^*$ .  $S$  is the background signal, modeled as the realization of an isotropic Gaussian self-similar random field. The motivation behind this choice is that the power spectrum of many natural images is isotropic with an  $1/\|\omega\|^\gamma$  type of decay, which is consistent with long-range dependencies [36, 33]. Also, this model of background signal fits fluorescence microscopy images well [45], which is relevant to many practical applications of orientation-estimation methods. Further details on our noise model are given in Appendix A.

We shall perform the correlations between measurement filters and the image  $I$  in a sliding fashion, and we are interested in determining the performance of this detection method in terms of angular accuracy. To that end, we shall only examine the position where the detector hits the target. Thus, without loss of generality, we set  $\mathbf{x}_0 = \mathbf{0}$  in (2.3).

We analyze the image  $I$  through a family of measurement filters  $\xi_\alpha$  parameterized by a (multi)-index  $\alpha$ . The measurements at location  $\mathbf{x}_0 = \mathbf{0}$  are then given by

$$(2.4) \quad q_\alpha = \langle I, \xi_\alpha \rangle = \langle J(\mathbf{R}_{-\theta^*} \cdot), \xi_\alpha \rangle + \langle S, \xi_\alpha \rangle.$$

We also set  $u_\alpha = \langle J, \xi_\alpha \rangle$  and  $s_\alpha = \langle S, \xi_\alpha \rangle$ .

We distinguish three different cases related to the construction of  $q_\alpha$  with  $\hat{\xi}_\alpha(\omega) = \hat{h}_\alpha(\omega)e^{jn_\alpha \theta}$  (cf. Table 1). In this paper, we address the question of angular accuracy of measurement filters constructed as conventional as well as wavelet detectors.

The local orientation angle  $\theta^*$  is estimated from the vector of measurements  $\mathbf{q} = (q_\alpha)$ . Any estimator based on this framework is then a mapping  $E$  that takes the measurements  $\mathbf{q}$  and returns an estimate  $\tilde{\theta}$  of  $\theta^*$ .

**2.3. Steerable Filterbanks and Estimation Strategy.** We perform the estimation of the unknown angle  $\theta^*$  by selecting a suitable filter  $\xi$  that is a linear combination of the measurement filters  $\xi_\alpha$ , and by selecting  $\tilde{\theta}$  as the solution of

$$(2.5) \quad \tilde{\theta} = \operatorname{argmax}_{\theta_0 \in [0, 2\pi)} \langle I, \xi(\mathbf{R}_{\theta_0} \cdot) \rangle.$$

In order to define an estimator that meets the requirements of Section 2.2,  $\xi$  needs to satisfy three properties.

- The filter  $\xi$  has to be a good approximation of the pattern of interest  $J$ , with the consequence that (2.5) corresponds to the detection of this pattern at the correct orientation. We note that angular (quasi) symmetries in  $J$  might lead to mis-detection of the angle.

- The filter  $\xi$  has to be robust to the background signal  $S$ , such that the estimation  $\tilde{\theta}$  mostly depends on the pattern  $J$ .
- The estimator  $\tilde{\theta}$  in (2.5) should be computable only based on the knowledge of the measurements  $(q_\alpha)$ .

We now identify the filters  $\xi_\alpha$  and  $\xi$  that allow us to achieve these three goals. To do so, we rely on steerable filters. We briefly introduce the concepts of steerability and refer the reader to Appendix B for more details.

We consider filters  $\xi_\alpha \in L_2(\mathbb{R}^2)$  that are polar-separable and have the form

$$(2.6) \quad \xi_\alpha(r, \theta) = \eta_\alpha(r) e^{jn_\alpha \theta}$$

with  $\eta_\alpha$  the *radial profile* and  $n_\alpha \in \mathbb{Z}$  the *harmonic* of  $\xi_\alpha$ . We assume that  $\xi_\alpha$  is normalized such that  $\|\xi_\alpha\|_2 = 1$ . The polar separability of the  $\xi_\alpha$ 's implies that

$$(2.7) \quad \begin{aligned} \langle \xi_\alpha, \xi_\beta \rangle &= \int_0^\infty \eta_\alpha(r) \eta_\beta(r) r dr \int_0^{2\pi} e^{j(n_\alpha - n_\beta)\theta} d\theta \\ &= 2\pi \delta[n_\alpha - n_\beta] \int_0^\infty \eta_\alpha(r) \eta_\beta(r) r dr \end{aligned}$$

with  $\delta[\cdot]$  the Kronecker delta. This shows that the system  $(\xi_\alpha)$  is orthonormal when the  $n_\alpha$  are all distinct.

The measurement function  $\xi_\alpha$  being polar-separable, its Fourier transform is also polar-separable as

$$(2.8) \quad \hat{\xi}_\alpha(\omega, \varphi) = \hat{h}_\alpha(\omega) e^{jn_\alpha \varphi}.$$

The function  $\hat{h}_\alpha$  is related to  $\eta_\alpha$  and  $n_\alpha$  by the Hankel transform (see, e.g., [40, Proposition 2]).

We remark that the rotated version of  $\xi_\alpha$  satisfies

$$(2.9) \quad \xi_\alpha(\mathbf{R}_{\theta_0} \mathbf{x}) = \xi_\alpha(r, \theta + \theta_0) = e^{jn_\alpha \theta_0} \xi_\alpha(\mathbf{x}),$$

and is therefore steerable in the sense of Definition B.1 in Appendix B. We then select a filter of the form  $\xi = \sum_\alpha c_\alpha \xi_\alpha$ , where the vector  $\mathbf{c} = (c_\alpha)$  determines the shape of the filter. The way of selecting  $\mathbf{c}$  adequately is discussed in Proposition 2.2. Hence,  $\xi$  is the best possible filter for the pattern  $J$  in the mean-square sense, once the measurement functions are given.

As a linear combination of steerable filters (see (2.9)), the filter  $\xi$  is steerable with

$$(2.10) \quad \xi(\mathbf{R}_{\theta_0} \mathbf{x}) = \sum_\alpha c_\alpha e^{jn_\alpha \theta_0} \xi_\alpha(\mathbf{x}).$$

This means that it is sufficient to apply the filtering with  $\xi_\alpha$  only once. Then, the rotated filter for any arbitrary angle can be determined by a systematic and linear transformation of the initial basis filters.

**2.4. The Law of Measurement Vector and the Cramér-Rao Lower Bound.** We aim at evaluating the performance of the estimators of  $\theta^* \in [0, 2\pi)$  depending on the measurement vectors  $\mathbf{q} = (q_\alpha)$ .

In this section, we only consider estimators  $\tilde{\theta} = \tilde{\theta}(\mathbf{q})$  that are unbiased, a notion that needs to be carefully defined for our image model (2.3). First of all, we assume in the rest of this paper that at least one coefficients  $u_\alpha = \langle J, \xi_\alpha \rangle$  is nonzero for a frequency  $n_\alpha \neq 0$ . This excludes isotropic templates  $J$  for which the angular estimation is meaningless. We then

define  $N_{\text{sym}}$  as the greatest common divisor of the  $n_\alpha \neq 0$  such that  $u_\alpha \neq 0$ . The quantity  $N_{\text{sym}}$  reflects the angular symmetry of the template  $J$  for the radial profiles  $\hat{h}_\alpha$ . Then, the angle  $\theta^*$  is determined modulo  $2\pi/N_{\text{sym}}$ , and the unbiasedness is therefore defined as

$$(2.11) \quad \mathbb{E}[\tilde{\theta}] = \theta^* \bmod [2\pi/N_{\text{sym}}].$$

In practice, we often have that  $N_{\text{sym}} = 1$ , typically when there is no angular symmetries on the template. A case where  $N_{\text{sym}} = 3$  is illustrated in Figure 1 below for the junction  $J_2$ .

The performance of the estimator is then measured by the mean-square error

$$(2.12) \quad \mathbb{E}[(\tilde{\theta}(\mathbf{q}) - \theta^*)^2].$$

As is well-known, there is a theoretical bound for the mean-square error that cannot be surpassed by an unbiased estimator, called the Cramér-Rao lower bound (CRLB). The latter is given by (see [22])

$$(2.13) \quad \mathbb{E}[(\tilde{\theta}(\mathbf{q}) - \theta^*)^2] \geq 1/\text{FI}(\theta^*) = \text{CRLB}(\theta^*),$$

where  $\text{FI}(\theta^*)$  is the Fisher information of the measurement vector  $\mathbf{q}$ .

In Theorem 2.1, we provide the mean vector, the covariance matrix and the Gaussian measurement vector and deduce its Fisher information. We recall here that the measurement filters  $\xi_\alpha$  are polar-separable of the form (2.6).

**THEOREM 2.1.** *Consider measurement filters of the form (2.6). We assume that the  $\xi_\alpha$  are such that*

$$(2.14) \quad \int_0^\infty \omega^{1-2\gamma} |\hat{h}_\alpha(\omega)|^2 d\omega < \infty,$$

where the  $\hat{h}_\alpha$  are defined in (2.8). Then, the measurement vector  $\mathbf{q}$  is Gaussian with mean vector

$$(2.15) \quad \boldsymbol{\mu} = \boldsymbol{\mu}(\theta^*) = \left( e^{jn_\alpha\theta^*} \langle J, \xi_\alpha \rangle \right)_\alpha$$

and covariance matrix

$$(2.16) \quad \mathbf{C} = \left( \delta[n_\alpha - n_\beta] \int_0^\infty \omega^{1-2\gamma} \hat{h}_\alpha(\omega) \overline{\hat{h}_\beta(\omega)} d\omega \right)_{\alpha,\beta},$$

where the  $h_\alpha$  are given in (2.8) and  $\gamma$  is the order of the whitening operator of  $S$ . The Fisher information is then

$$(2.17) \quad \text{FI}(\theta^*) = 2\text{Re} \left( \left( \frac{d\boldsymbol{\mu}}{d\theta^*} \right)^H \mathbf{C}^{-1} \frac{d\boldsymbol{\mu}}{d\theta^*} \right),$$

where  $\mathbf{x}^H = (\mathbf{x}^*)^T$  is the Hermitian transpose of  $\mathbf{x}$ .

The proof of Theorem 2.1 is postponed to Appendix C. The condition (2.14) is necessary to make the problem well-posed, because  $\langle S, \xi_\alpha \rangle$  is well-defined under this condition. It is related to the existence of vanishing moments for  $\xi_\alpha$  (see Appendix A). It is important to notice that the mean  $\boldsymbol{\mu}(\theta^*)$  of  $\mathbf{q}$  depends on  $\theta^*$ , while the covariance  $\mathbf{C}$  does not. This is due to the assumption that the background signal is statistically isotropic.

**2.5. Maximum Likelihood Estimator and Unbiasedness.** In this paper, we consider estimators of the form (2.5), where  $\xi$  is a steerable filter given by  $\xi = \sum_{\alpha} c_{\alpha} \xi_{\alpha}$ . We show here that, under some assumptions, this corresponds to the maximum likelihood estimator for an adequate choice of the  $c_{\alpha}$ . More precisely, one selects the  $c_{\alpha}$  such that  $\xi$  corresponds to the orthogonal projection of the pattern  $J$  to the basis filters ( $\xi_{\alpha}$ ).

PROPOSITION 2.2. *We assume that the filters  $\xi_{\alpha}$  have distinct harmonics ( $n_{\alpha} \neq n_{\beta}$  for  $\alpha \neq \beta$ ) and identical Fourier radial profiles ( $\hat{h}_{\alpha} = \hat{h}$  for every  $\alpha$ ). Then, the estimator  $\tilde{\theta}$  defined in (2.5) with  $\xi = \sum_{\alpha} u_{\alpha} \xi_{\alpha}$  with  $u_{\alpha} = \langle J, \xi_{\alpha} \rangle$  corresponds to the maximum likelihood estimator*

$$(2.18) \quad \theta_{\text{MLE}} = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmax}} \mathbb{P}(\mathbf{q}|\theta_0)$$

of the image model (2.3).

*Proof.* According to Theorem 2.1, conditionally to  $\theta^* = \theta_0$ , the vector  $\mathbf{q}$  is Gaussian with mean  $\boldsymbol{\mu}(\theta_0)$  and co-variance matrix  $\mathbf{C}$ . Therefore,  $\mathbb{P}(\mathbf{q}|\theta_0)$  is proportional to  $\exp(-\frac{1}{2}(\mathbf{q} - \boldsymbol{\mu}(\theta_0))^H \mathbf{C}^{-1}(\mathbf{q} - \boldsymbol{\mu}(\theta_0)))$ . Hence,

$$(2.19) \quad \theta_{\text{MLE}} = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmin}} (\mathbf{q} - \boldsymbol{\mu}(\theta_0))^H \mathbf{C}^{-1}(\mathbf{q} - \boldsymbol{\mu}(\theta_0)).$$

The harmonics of the  $\xi_{\alpha}$  being distinct,  $\mathbf{C}$  is diagonal. Moreover,  $\mathbf{C}[n_{\alpha}, n_{\alpha}]$  depends only on  $\eta_{\alpha} = \eta$ , and is therefore independent of  $\alpha$ . This means that

$$(2.20) \quad \theta_{\text{MLE}} = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmin}} (\mathbf{q} - \boldsymbol{\mu}(\theta_0))^H \overline{(\mathbf{q} - \boldsymbol{\mu}(\theta_0))}.$$

One can develop this latter expression. We remark that  $\mathbf{q}^H \overline{\mathbf{q}}$  and  $\boldsymbol{\mu}(\theta_0)^H \overline{\boldsymbol{\mu}(\theta_0)} = \sum_{\alpha} |\langle J, \xi_{\alpha} \rangle|^2$  are independent of  $\theta_0$ . Therefore, we have finally,

$$(2.21) \quad \theta_{\text{MLE}} = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmin}} (-2\mathbf{q}^H \overline{\boldsymbol{\mu}(\theta_0)}) = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmax}} \sum_{\alpha} q_{\alpha} \overline{\langle J, \xi_{\alpha} \rangle} e^{-jn_{\alpha}\theta_0}.$$

Moreover, recalling that  $q_{\alpha} = \langle I, \xi_{\alpha} \rangle$ , we can rewrite (2.5) as

$$(2.22) \quad \tilde{\theta} = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmax}} \left\langle I, \sum_{\alpha} c_{\alpha} e^{jn_{\alpha}\theta_0} \xi_{\alpha} \right\rangle = \underset{\theta_0 \in [0, 2\pi)}{\operatorname{argmax}} \sum_{\alpha} q_{\alpha} \overline{c_{\alpha}} e^{-jn_{\alpha}\theta_0},$$

Hence,  $\tilde{\theta} = \theta_{\text{MLE}}$  as soon as  $c_{\alpha} = \langle J, \xi_{\alpha} \rangle$ , as expected.  $\square$

We now demonstrate that the orientation estimation  $\tilde{\theta}$  defined from (2.5) with  $\xi = \sum_{\alpha} u_{\alpha} \xi_{\alpha}$  is unbiased, which is crucial for the application of the Cramér-Rao bound. In particular, Proposition 2.3 implies that  $\mathbb{E}[(\tilde{\theta} - \theta^*)^2] \geq \text{CRLB}(\theta^*)$ .

PROPOSITION 2.3. *Let  $I$  be the analysed image corresponding to the model (2.3), and  $\xi_{\alpha}$  be filters corresponding to different harmonics with identical Fourier radial profiles. Let  $N_{\text{sym}}$  be the greatest common divisor of the  $n_{\alpha} \neq 0$  such that  $u_{\alpha} \neq 0$  (assumed to be nonempty, see Section 2.4). Then, the estimator (2.5) with  $\xi = \sum_{\alpha} u_{\alpha} \xi_{\alpha}$  is unbiased in the sense of (2.11).*

*Proof.* We recall here that  $u_{\alpha} = \langle J, \xi_{\alpha} \rangle$ . With Theorem 2.1, we know that  $\tilde{\mathbf{q}}$ , defined by  $\tilde{q}_{\alpha} = q_{\alpha} - e^{jn_{\alpha}\theta^*} u_{\alpha}$  is a complex Gaussian vector with 0 mean. Starting with (2.22), we therefore have that

$$(2.23) \quad \tilde{\theta} = \underset{\theta_0}{\operatorname{argmax}} \sum_{\alpha} |u_{\alpha}|^2 e^{jn_{\alpha}(\tilde{\theta} - \theta_0)} + \sum_{\alpha} \tilde{q}_{\alpha} \overline{u_{\alpha}} e^{-jn_{\alpha}\theta_0}.$$

The first part in (2.23) is a deterministic function of  $\theta_0$ ,  $2\pi/N_{\text{sym}}$  periodic, symmetric with respect to  $\theta^*$ , and reaching its maximum exactly for  $\theta_0 = \tilde{\theta} \bmod(2\pi/N_{\text{sym}})$ . The second part  $X(\theta_0) = \sum_{\alpha} \tilde{q}_{\alpha} \overline{u_{\alpha}} e^{-jn_{\alpha}\theta_0}$  in (2.23) is random and such that

$$(2.24) \quad \mathbb{P}(X(\theta_0) \in [\theta^*, \theta^* + \Delta]) = \mathbb{P}(X(\theta_0) \in [\theta^* - \Delta, \theta^*])$$

for any  $\Delta > 0$ . The symmetry of the deterministic part of (2.23) around  $\tilde{\theta}$  together with (2.24) then implies that

$$(2.25) \quad \mathbb{P}(\langle I, \sum_{\alpha} u_{\alpha} e^{jn_{\alpha}\theta_0} \rangle \in [\theta^*, \theta^* + \Delta]) = \mathbb{P}(\langle I, \sum_{\alpha} u_{\alpha} e^{jn_{\alpha}\theta_0} \rangle \in [\theta^* - \Delta, \theta^*]).$$

This means that, in average,  $\tilde{\theta}$  reaches its maximum exactly where the deterministic function  $\theta_0 \mapsto \sum_{\alpha} |u_{\alpha}|^2 e^{jn_{\alpha}(\tilde{\theta} - \theta_0)}$  does, which is a reformulation of (2.11).  $\square$

The angle  $\tilde{\theta}$  in (2.22) is numerically computed with standard optimization techniques or, simply, with a grid search. Finally,  $\hat{\theta}$  is an estimator of the form (2.5), which depends only on the measurements  $q_{\alpha}$ , and is based on the steerable filter  $\xi$  that provides the best approximation of the pattern  $J$  in span  $\{\xi_{\alpha}\}$ .

We shall now use the CRLB (2.13) together with the Fisher information (2.17) to evaluate the optimal performance achievable by an unbiased estimator.

### 3. Cramér-Rao Lower Bound for an Estimation with Distinct Harmonics.

**3.1. The Cramér-Rao Lower Bound.** In this section, we derive the CRLB in cases where the measurement functions defined in (2.6) have distinct harmonics  $n_{\alpha}$ . So that  $n_{\alpha} = n$ , we shall consequently index the measurement functions directly by their harmonics  $n \in H$ , where  $H \subset \mathbb{Z}$  is the set of all used harmonics. Specifically, we write

$$(3.1) \quad \xi_n(r, \theta) = \eta_n(r) e^{jn\theta}$$

and, in the Fourier domain,

$$(3.2) \quad \hat{\xi}_n(\omega, \varphi) = \hat{h}_n(\omega) e^{jn\varphi}.$$

We recall that for  $n \neq m$ ,  $\xi_n$  and  $\xi_m$  are orthogonal, due to the orthogonality of their angular factors (see (2.7)). Moreover, we have that, for a real image  $I$ ,

$$(3.3) \quad q_{-n} = \langle I, \xi_{-n} \rangle = \langle I, \overline{\xi_n} \rangle = \overline{\langle I, \xi_n \rangle} = \overline{q_n},$$

where we used the fact that  $\eta_n$  is real. Thus,  $q_{-n}$  and  $q_n$  essentially carry the same information, so that the CRLB based on  $q_n$ ,  $n \in H$ , is the same as the CRLB based on  $q_n$  with harmonics  $n$  in the set

$$(3.4) \quad H_+ = \{|n| : n \in H\}.$$

We further exclude  $n = 0$  from consideration since the corresponding measurement does not depend on the rotation angle  $\theta^*$ . We remark that the correlation matrix  $\mathbf{C}$  in (2.16) is diagonal due to the distinct nature of the harmonics. As a consequence, the CRLB does not depend on  $\theta^*$ , as is easily deduced from Proposition 2.1.

We now calculate the CRLB for estimating the angle  $\theta^*$  of the pattern in (2.3).

**THEOREM 3.1.** *Consider the image model (2.3). For measurements  $q = (q_n)$  using distinct harmonics  $n \in H$ , the exact form of the CRLB is*

$$(3.5) \quad \text{CRLB} = \frac{\sigma_0^2/4\pi}{\sum_{n \in H_+} \frac{n^2 |u_n|^2}{\int_0^{\infty} \omega^{1-2\gamma} |\hat{h}_n(\omega)|^2 d\omega}},$$

where we recall that  $u_n = \langle J, \xi_n \rangle$ .

The proof of Theorem 3.1 is given in Appendix D.

**3.2. Conventional Detector: Estimation From Best  $N$  Measurements.** In the case of conventional detectors, one chooses the same radial pattern  $\hat{h}$  for all measurement functions [58], which results in

$$(3.6) \quad \hat{\xi}_n(\omega, \varphi) = \hat{h}(\omega) e^{jn\varphi}.$$

For many applications, this hypothesis is relevant because the pattern of interest  $J$  is itself separable in polar coordinates [40]. This radial component is typically band-pass, given its vanishing moments and finite energy. In practice, the choice of the radial profile (band-pass filter  $\hat{h}$ ) specifies the scale of the detector.

Now we address the question of selecting the best steerable subspace for specifying our steerable matched-filter detector. Suppose we have at our disposal the finite number  $N$  of harmonics (*i.e.*, a finite number of measurements). It is then natural to ask which harmonics to consider to reduce the CRLB as much as possible. This is obtained after an immediate corollary of Theorem 3.1.

**COROLLARY 3.2.** *The CRLB for the estimation problem in the case of a single common radial profile is*

$$(3.7) \quad \text{CRLB} = \frac{1}{\sum_{n \in H_+} n^2 |u_n|^2} \frac{\sigma_0^2}{4\pi} \int_0^\infty \omega^{1-2\gamma} |\hat{h}(\omega)|^2 d\omega.$$

Corollary 3.2 simply exploits the fact that the  $\hat{h}_n$  are all equal to  $\hat{h}$  in (3.5).

Considering (3.7), choosing the best measurements is equivalent to identifying the set  $H_+ \subset \mathbb{Z}_+$  with  $N$  members, for which the sum

$$(3.8) \quad \sum_{n \in H_+} n^2 |u_n|^2$$

is maximized.

**3.3. Conventional Detector: Asymptotic Behavior.** Similarly to Section 3.2, we fix the radial part of  $\xi_n$  as being independent of  $n$  and look at the asymptotic behavior in terms of the number of harmonics, by choosing the set  $H_+$  of harmonics as  $\{1, \dots, N\}$ , and letting  $N$  tend to infinity. Based on (3.7), we conclude that the asymptotic behavior of the CRLB depends on the asymptotic (decay) properties of the coefficients  $u_n = \langle J, \xi_n \rangle$  of the directional pattern.

The question is whether the CRLB vanishes asymptotically as  $N \rightarrow \infty$ , which would suggest the theoretical possibility of perfect estimation with infinitely many measurements. We now study this question.

We start with a preliminary remark. By assumption, the pattern  $J$  has finite energy. Since the measurement functions  $\{\xi_n\}_{n \in \mathbb{Z}}$  are orthonormal, by the Bessel inequality, we have that

$$(3.9) \quad \sum_{n \in \mathbb{Z}} |u_n|^2 \leq \|J\|_2^2 < \infty,$$

hence  $(u_n)_{n \in \mathbb{Z}} \in \ell_2(\mathbb{Z})$ . We see now how to refine this latter condition to have a vanishing CRLB.

**PROPOSITION 3.3.** *In the framework of conventional detectors, for a pattern  $J \in L_2(\mathbb{R}^2)$ , the CRLB does not vanish when  $N \rightarrow \infty$  if and only if*

$$(3.10) \quad \sum_{n \in \mathbb{Z}} n^2 |u_n|_2^2 < \infty.$$

Moreover, the  $u_n$  are the Fourier coefficients of the function

$$(3.11) \quad G(\varphi) = \frac{1}{2\pi} \int_0^\infty \overline{\hat{J}(\omega, \varphi)} \hat{h}(\omega) \omega d\omega.$$

Therefore, the CRLB does not vanish if and only if the function  $G$  is differentiable with its derivative being square integrable.

*Proof.* The CRLB is inversely proportional to (3.10), implying the first equivalence. We then remark that, using Parseval relation and polar coordinates, we have

$$(3.12) \quad u_n = \langle J, \xi_n \rangle = \frac{1}{(2\pi)^2} \langle \hat{J}, \hat{\xi}_n \rangle = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^\infty \overline{\hat{J}(\omega, \varphi)} \hat{h}(\omega) e^{jn\varphi} \omega d\omega d\varphi = \frac{1}{2\pi} \int_0^{2\pi} G(\varphi) e^{jn\varphi} d\varphi.$$

The sequence  $(u_n)$  is in  $\ell_2(\mathbb{Z})$ , therefore  $G$  is square integrable. The Fourier coefficients of the (weak) derivative of  $G$  are then  $jnu_n$ . Hence,  $G$  is differentiable with a square integrable derivative if and only if the sequence  $(jnu_n)$  is in  $\ell_2(\mathbb{Z})$ , proving the second equivalence in Proposition 3.3.  $\square$

The convergence of the series in (3.7) therefore depends on the decay properties of the Fourier-series of  $G(\varphi)$ , which in turn is related to its smoothness. In particular, if  $\hat{J}$  has angular jump discontinuities that are inherited by  $G$ ,  $u_n$  will decay slowly like  $1/n$ , and the series will diverge. In this case, the CRLB will asymptotically vanish. This, for instance, happens if  $\hat{J}$  has jump discontinuities along infinite radial lines, which typically goes along with a similar discontinuity in  $J$  (see [52]). This suggests that, for such patterns, the angular error of a steerable detector can be made arbitrarily small by selecting a sufficient number of harmonics.

**4. Experiments.** In practice, the radial profil  $\hat{h}$  is typically chosen as the Laplacian of a Gaussian (LoG) or a band-pass filter (even and compactly supported in the frequency domain). For the experiments, we use the LoG filter and the first scale of the Meyer-type profile [7]

$$(4.1) \quad \hat{h}(\omega) = \begin{cases} \sin\left(\frac{\pi}{2} \nu \left(\frac{4\omega}{\pi} - 1\right)\right), & \frac{\pi}{4} < \omega \leq \frac{\pi}{2} \\ \cos\left(\frac{\pi}{2} \nu \left(\frac{2\omega}{\pi} - 1\right)\right), & \frac{\pi}{2} < \omega \leq \pi \\ 0, & \text{otherwise} \end{cases}$$

with the auxiliary function  $\nu(t) = t^4(35 - 84t + 70t^2 - 20t^3)$ ,  $\nu \in C^3([0, 1])$ .

Other typical examples include Shannon-type [59], Simoncelli [38], Papadakis [43] and Held [16] wavelets. Similar results are obtained when using these radial profiles due to the fact that all these functions approximate the indicator function  $[\pi/4, \pi/2]$ .

**4.1. CRLB for Analytical Patterns.** In this section, we compute the CRLB associated with a few directional patterns for which explicit analytical formulas are provided. Here, we are interested in the decay of the CRLB and how the choice of the harmonics effect this curve. The goal is to analyse the following quantities:  $|u_n|$ , defined in (3.12),  $n|u_n|$  (as it determines the decay rate of the CRLB), and the CRLB itself, given in (3.7), for a fixed number of harmonics. As the pattern is given in a closed form, we compute these quantities analytically, without the need of generating an actual discrete image.

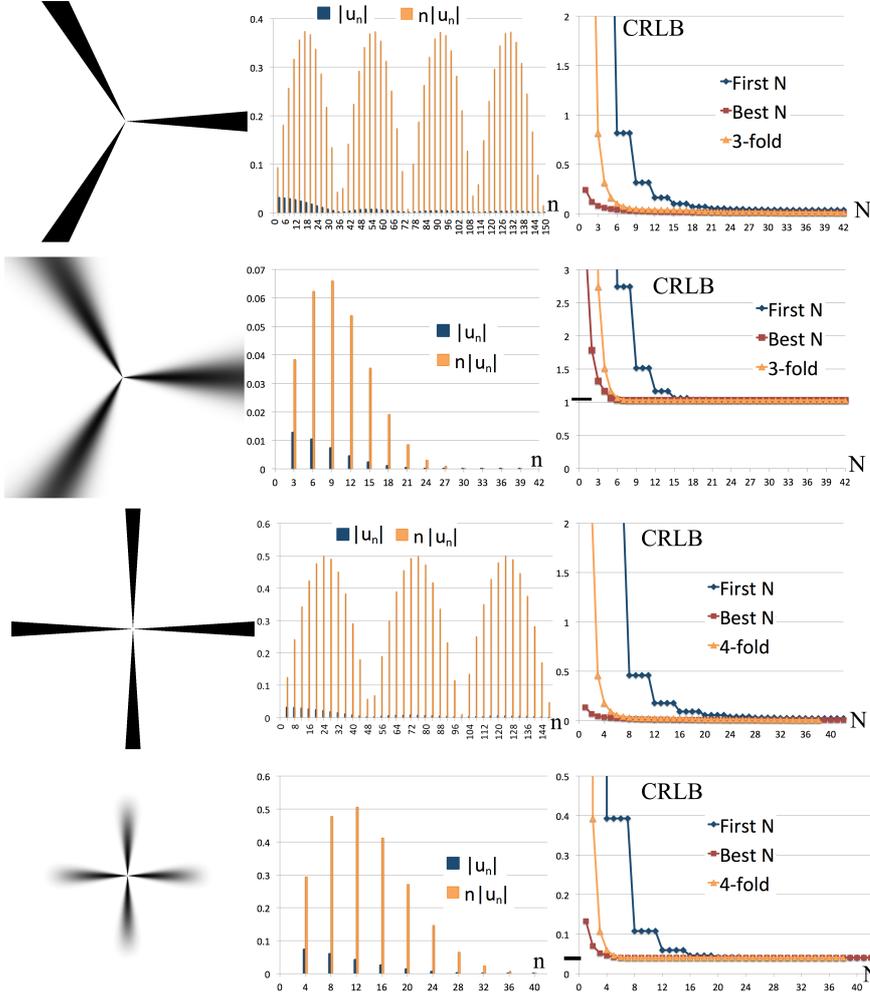


FIG. 1. First column from top to bottom: illustration of the analytically defined patterns  $\hat{J}_1$  (4.2),  $\hat{J}_2$  (4.3),  $\hat{J}_3$  (4.4) and  $\hat{J}_4$  (4.5) in the Fourier domain. Second column:  $|u_n|$  (3.12) and  $n|u_n|$  as a function of harmonics ( $n$ ). Third column: The CRLB (3.7) as a function of the number of harmonics.

We study four different types of patterns. Specifically,

$$(4.2) \quad \hat{J}_1(\omega, \varphi) = \begin{cases} 1, & \text{if } \cos(1.5\varphi)^\beta > 0.8 \\ 0, & \text{otherwise} \end{cases}$$

$$(4.3) \quad \hat{J}_2(\omega, \varphi) = \left( \frac{1}{1 + \omega^\lambda} \right) \cos(1.5\varphi)^\beta$$

$$(4.4) \quad \hat{J}_3(\omega, \varphi) = \begin{cases} 1, & \text{if } \cos(2\varphi)^\beta > 0.8 \\ 0, & \text{otherwise} \end{cases}$$

$$(4.5) \quad \hat{J}_4(\omega, \varphi) = \cos(2\varphi)^\beta e^{-\frac{1}{\alpha\omega}}$$

within the support of  $\hat{h}(\omega)$ , with  $\lambda = 2.1$ ,  $\beta = 28$  and  $\alpha = 2.5$ .

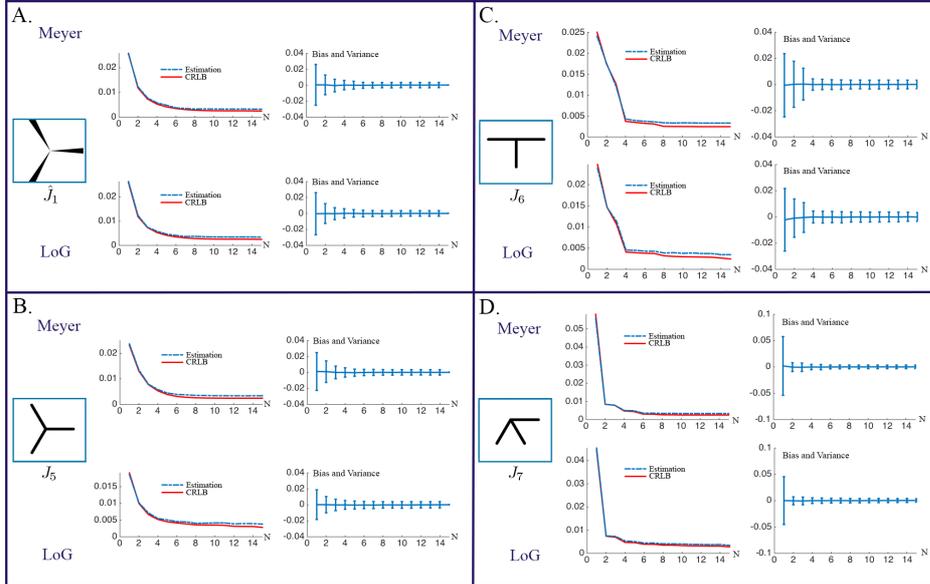


FIG. 2. Accuracy of the orientation estimation of the proposed estimator (dashed line) in the sense of (2.13) compared to the CRLB given in (3.7) (continuous line), bias and variance of the estimated angles from  $P$  measurements of the proposed estimator. All angles are calculated in radians. A.: Results on the analytically defined junction  $\hat{J}_1$  (4.2); top: with Meyer radial detector profile, bottom: with LoG detector profile. B. Results on a three-fold symmetric junction (drawn); top: with Meyer radial detector profile, bottom: with LoG detector profile. C. Results on a T-shaped junction (drawn); top: with Meyer radial detector profile, bottom: with LoG detector profile. D. Results on an arrow-shaped junction (drawn); top: with Meyer radial detector profile, bottom: with LoG detector profile.

For computing the CRLB, we apply three different strategies: “First  $N$ ”; “Best  $N$ ”; and “ $k$ -fold”. In the first case, we use the first  $N$  coefficients; in the Best  $N$  case, we select the harmonics that maximize (3.8). In the case of  $k$ -fold symmetric patterns we choose the first  $N$  multiples of  $k$  as harmonics. This latter choice accounts for the name “ $k$ -fold”.

Related to (3.5), we have chosen the variance of the continuous domain Gaussian white noise  $\sigma_0^2 = 1$  since it provides only a scaling factor and does not influence the decay of the curve. Figure 1 contains an illustration of the results.

In the case of sharp edges ( $\hat{J}_1$  (4.2) and  $\hat{J}_3$  (4.4)), the rate of decay of the circular-harmonic coefficients permits a theoretical vanishing limit for the CRLB. In the smooth angular cases ( $\hat{J}_2$  (4.3) and  $\hat{J}_4$  (4.5)), the CRLB converges to a theoretical positive value.

Moreover, as expected, for the three-fold patterns, only every third component and for the four-fold patterns every fourth element plays a significant role in the estimation of the orientation. This can be seen in the almost flat CRLB curve between multiples of three (or four, respectively) in the “First  $N$ ” strategy. We also observe a difference in performance between generically choosing the first  $N$   $k$ -fold symmetric coefficients (as a strategy for unknown  $k$ -fold patterns) and making our choice of harmonics based on maximizing (3.8).

Finally, by looking at Figure 1, we observe that, with the right choice of harmonics, the CRLB can be much reduced, even with a small number of harmonics.

**4.2. Empirical Optimality of the Proposed Estimator.** In this section, we test the accuracy of the estimation of the proposed estimator (2.22) on two symmetric and two asym-

metric patterns. Contrary to Section 4.1, here we use only one analytical (Figure 2.A) and three discrete images (Figure 2.B, 2.C and 2.D) to test the accuracy of the estimator. We are interested in the followings: 1. How does the performance of the estimator relate to the actual CRLB curve. 2. Test empirically the unbiasedness of the estimator.

The CRLB is calculated according to (3.7). By estimation, we mean the empirical error of the estimator in the sense of (2.13). The parameters of the estimator are calculated as follows. We build  $P = 1,000$  different realizations to make the experiments statistically reliable. We chose the variance of the noise such that it corresponds to an SNR of 17.22 dB. The typical SNR range of real images where the quality is still acceptable is 15-35 dB, so that our experimental conditions are representative of a noisy image. The bias of the estimator is  $\mu_P - \theta^*$ , where  $\mu_P$  is the mean of the estimated angles from  $P$  measurements  $\mu_P = \frac{1}{P} \sum_{p=1}^P \hat{\theta}_p$ . The variance of the estimator is the variance of the estimations from  $P$  measurements:  $\frac{1}{P} \sum_{p=1}^P (\hat{\theta}_p - \mu_P)^2$ . All angles are calculated in radians.

For symmetric patterns, we have chosen the analytically defined junction  $J_1$  (4.2) and a sharp, drawn three-fold junction (Figure 2,  $J_5$ ). For asymmetric patterns, we have chosen a sharp, drawn T-shape junction (Figure 2,  $J_6$ ) and a sharp, drawn arrow-type junction (Figure 2,  $J_7$ ). The coefficients  $c_\alpha$  in (2.22) are obtained in each case as the orthogonal projection of the junction of interest on the measurement functions.

For the estimator, we used every third harmonic. For the radial part of our detector, we have chosen the first scale of the Meyer wavelet profile and the Laplacian of Gaussian (LoG) filter.

The results for the patterns are illustrated in Figure 2. We first show experimentally that the estimator (2.22) is unbiased with more than one harmonic. The performances are therefore comparable to the CRLB. We then observe that the accuracy of the estimator follows closely the CRLB curve, while staying above, as expected. This proves empirically the almost optimality of the estimator (2.22) for the estimation of the angle  $\theta^*$ .

## 5. Extension to Wavelets.

**5.1. Steerable Wavelets.** In this section, we extend our results on the CRLB to wavelet detectors. Steerable wavelet frames are adapted to capture the local orientation of features, or junctions, within a multiresolution hierarchy. To simplify the notations, we consider wavelets that are centered at the origin. Moreover, we apply the multiover complex Riesz transform on a tight wavelet frame of  $L_2(\mathbb{R}^2)$ . Proposition 4.1 in [58] provides sufficient conditions on the isotropic profile that has to be defined to generate a desired wavelet system. There are various types of isotropic profiles satisfying the proposition; typical examples are given at the beginning of Section 4. For further details, we refer to [59].

Similarly to the design of conventional detectors, we take the values of  $n$  from the predefined set of harmonics  $H = \{n_0, \dots, n_N\}$ . The wavelet schemes generated in such a way are often referred to as circular harmonic wavelets [58, 18]. In an extension of (3.1), the measurement functions are indexed here by the pair  $\alpha = (n, i)$  of the harmonic  $n \in H$  and the scale  $i \in \mathbb{Z}$ . In particular, the circular harmonic wavelet  $\xi$  at scale  $i$  and harmonic channel  $n$  takes the form

$$(5.1) \quad \hat{\xi}_{n,i}(\omega, \varphi) = 2^i \hat{h}(2^i \omega) e^{jn\varphi}.$$

We note that, in the case of wavelet measurements, we have multiple measurement functions (at different scales) for the same harmonic  $n$ . We make the additional assumption that  $\hat{h}$  is real-valued, which holds true for every aforementioned radial design. Finally, given that the radial patterns  $\hat{h}$  of interest are bandpass, we assume that  $\hat{h}(\omega) = 0$  for  $\omega \notin (\pi/4, \pi]$  to set the fundamental scale.

**5.2. Multiple Scales: The Cramér-Rao Lower Bound.** We recall that the goal is to estimate  $\theta^*$  from the measurements  $q_{n,i} = \langle I, \xi_{n,i} \rangle$ . We also denote  $u_{n,i} = \langle J, \xi_{n,i} \rangle$  and  $s_{n,i} = \langle S, \xi_{n,i} \rangle$ .

To simplify future formulas, we introduce the following notations:

$$(5.2) \quad \tilde{q}_{n,i} = 2^{-i\gamma} q_{n,i},$$

$$(5.3) \quad b_z = \frac{1}{2\pi} \int_0^\infty \omega^z \hat{h}(\omega)^2 \omega d\omega,$$

$$(5.4) \quad d_z = \frac{1}{2\pi} \int_0^\infty \omega^z \hat{h}(\omega) \hat{h}(2\omega) \omega d\omega.$$

We also set the two constants

$$(5.5) \quad B = \sigma_0^2 b_{-2\gamma}, \quad D = \sigma_0^2 2^{1-\gamma} d_{-2\gamma}.$$

The angle  $\theta^*$  is estimated by steering the whole template, which can be seen as a sum of templates. The task of estimating  $\theta^*$  based on  $q_{n,i}$  is the same as its estimation based on  $\tilde{q}_{n,i}$ . In particular, the CRLB is the same. We first give the covariance matrix of the measurements  $\tilde{q}_{n,i}$ .

PROPOSITION 5.1. *The covariance matrix  $\mathbf{C}$  of the random vector  $(\tilde{q}_{n,i})_{n \in \mathbb{Z}, i \in \mathbb{Z}}$  is given by*

$$(5.6) \quad \mathbf{C}[(n, i), (m, k)] = \begin{cases} B, & \text{if } m = n, k = i \\ D, & \text{if } m = n, |k - i| = 1 \\ 0, & \text{otherwise.} \end{cases}$$

*Proof.* First,  $\mathbf{C}[(n, i), (m, k)] = 0$  as soon as  $m \neq n$  because the functions  $\xi_{n,i}$  and  $\xi_{m,k}$  are orthogonal. Then, the assumption that  $\hat{h}(\omega) = 0$  for  $\omega \notin (\pi/4, \pi]$  ensures that  $\hat{h}(2^i \omega)$  and  $\hat{h}(2^k \omega)$  do not overlap for  $|k - i| > 1$ , implying again that  $\mathbf{C}[(n, i), (m, k)] = 0$ .

We now assume that  $m = n$  and  $|k - i| \leq 1$ . When  $k = i$ , we easily recognize the quantity (5.3) with  $z = -2\gamma$  in (2.16) and therefore deduce that  $\mathbf{C}[(n, i), (n, i)] = \sigma_0^2 b_{-2\gamma} = B$ . Similarly, one shows using again (2.16) that  $\mathbf{C}[(n, i), (n, k)] = D$  when  $|k - i| = 1$ .  $\square$

The main challenge in computing the Fisher information from (2.17) for wavelet measurements is that, unlike in context of Section 3, the covariance matrix of the measurements  $\{\tilde{q}_{n,i}\}$  is not diagonal, so it is not as straightforward to invert. But this challenge is still surmountable because, as we see from (5.6), we can rearrange the measurements such that the covariance matrix is (at most) tridiagonal. The exact rearrangement of the measurements is described in details in Appendix E.

The key idea is to divide the set of measurement indices into as few disjoint sets  $G_r$  as possible ( $r = 1, \dots, g$ , where  $g$  denotes the total number of such sets). Then, each set  $G_r$  has some  $l_r$  elements of the form  $(n_r, i_r), \dots, (n_r, i_r + l_r - 1)$ , for some harmonic  $n_r \in H_+$  and minimum scale  $i_r$ . We re-index all measurements as  $\tilde{q}_{r|e}$  using the notation

$$(5.7) \quad (r|e) := (n_r, i_r + e - 1),$$

with  $r = 1, \dots, g$  and  $e = 1, \dots, l_r$ . Then, the covariance matrix of the measurements  $\{\tilde{q}_{r|e}\}$

for each fixed  $r$  is Toeplitz-tridiagonal of the form

$$(5.8) \quad \mathbf{T}_r = \begin{pmatrix} B & D & 0 & 0 & \dots & 0 \\ D & B & D & 0 & \dots & 0 \\ 0 & D & B & D & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & D & B & D \\ 0 & 0 & \dots & 0 & D & B \end{pmatrix}_{l_r \times l_r}.$$

Moreover, the overall covariance matrix is block-diagonal, with  $\mathbf{T}_r$ ,  $r = 1, \dots, g$ , as its diagonal blocks. Using this reformulation of the problem, we are able to compute the CRLB explicitly.

**THEOREM 5.2.** *For wavelet measurements  $q_{r|e}$ ,  $r = 1, \dots, g$ ,  $e = 1, \dots, l_r$ , the Fisher information is given by*

$$(5.9) \quad FI(\theta^*) = 2 \sum_{r=1}^g n_r^2 \sum_{t=1}^{l_r} \left( B + 2D \cos\left(\frac{t\pi}{l_r+1}\right) \right)^{-1} \times \left| \sum_{e=1}^{l_r} 2^{-(i_r+e-1)\gamma} u_{r|e} \sin\left(e \frac{t\pi}{l_r+1}\right) \right|^2.$$

The CRLB for the estimation problem is given by  $1/FI(\theta^*)$  and satisfies

$$(5.10) \quad \frac{B - 2|D|}{2 \sum_{n,i} n^2 4^{-i\gamma} |u_{n,i}|^2} \leq \text{CRLB} \leq \frac{B + 2|D|}{2 \sum_{n,i} n^2 4^{-i\gamma} |u_{n,i}|^2},$$

where the constant  $B - 2|D|$  is strictly positive.

The proof of Theorem 5.2 is given in Appendix F. The Fisher information is expressed as sums and inverses that are difficult to handle. However, we can give a fair estimate of the CRLB by narrowing it to a range (5.10). In the case of a conventional detector, the value of  $D$  is 0 (see (5.5)). Thus, the lower and upper bound of the range are equal and we arrive to the same exact CRLB that we computed for the case of conventional detectors (3.7). In the general case, it is crucial to remark that  $0 < B - 2|D|$  (see Appendix F); hence the lower bound in (5.10) is strictly positive. Finally, we remark that, as in the case of conventional detectors, the Fisher information, and therefore the CRLB, does not depend on  $\theta^*$ .

**5.3. Asymptotic Behavior for All Scales and Harmonics.** We now consider the case where we have access to each wavelet coefficient  $q_{n,i} = \langle I, \xi_{n,i} \rangle$  for  $n \in \mathbb{Z}$ ,  $i \in \mathbb{Z}$ . Once again, for real-valued data, we can limit ourselves to  $n \in \mathbb{Z}_+$  for computing the CRLB, due to Hermitian symmetry.

**PROPOSITION 5.3.** *Assume that the order of the whitening operator associated with the Gaussian self-similar random field  $\gamma$  is such that  $\frac{b_0}{|d_0|} \neq 2^{1-\gamma} + 2^{1+\gamma}$ . Then, by observing  $\langle I, \xi_{n,i} \rangle$ , the CRLB does not vanish if and only if*

$$(5.11) \quad \sum_{i \in \mathbb{Z}} \sum_{n \in \mathbb{Z}^+} n^2 4^{i\gamma} |u_{n,i}|^2 < \infty.$$

Moreover, if  $J$  has an expansion in terms of  $\{\xi_{n,i}\}$ , then this is also equivalent to

$$(5.12) \quad (-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \in L_2(\mathbb{R}^2).$$

The proof of Proposition 5.3 is provided in Appendix G. It is the multiscale version of Proposition 3.3. We interpret the fact that the CRLB vanishes as a possibility to perfectly

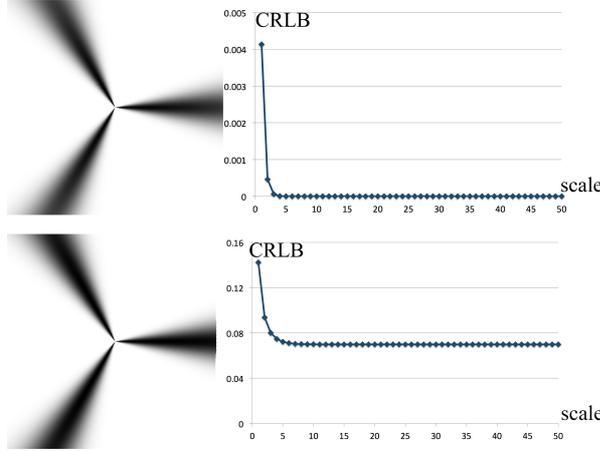


FIG. 3. First column: Illustration of the analytically defined junction  $J_2$ , where the decay rate of the pattern is  $\lambda = 2.1$  and  $4.5$ , respectively. Second column: The CRLB as a function of the number of the largest wavelet scale.

detect the correct angle from the set of measurements, assuming all the harmonics and scales are available. The assumption on  $\gamma$  is technical and simply implies that one specific value of  $\gamma$  should be avoided for a given wavelet. The values of such  $\gamma$  for different wavelets are given in Appendix G.

**5.4. Experiments with Wavelets.** In this section, we compute the CRLB of the junction  $\hat{J}_2$  (4.3) with decay rate  $\lambda = 2.1$  and  $4.5$ , respectively. These choices of the decay rate  $\lambda$  correspond to two cases: first, where the background decays faster than the junction (in the Fourier domain); and second, where it is the junction that decays faster. The value of the exponent of the Gaussian field  $\gamma$  is fixed as  $2.5$ .

For the experiments, we have chosen the Meyer-type wavelet (4.1). We again choose the variance of the continuous domain Gaussian noise  $\sigma_0^2 = 1$  since it provides a scaling factor that does not influence the rate of decay of the curve.

Based on the graphs of Figure 3, we can observe the followings. For the decay rate of  $\hat{J}_2$   $\lambda = 2.1$ , one wavelet scale was not enough for the CRLB to converge to a theoretical perfect estimation (like in Figure 1, second experiment). However, by increasing the number of scales, the 0 bound is asymptotically achievable. This illustrates the main result of Proposition 5.3, which shows that, by using finer scales, wavelets can improve the estimate, even in the case of junctions of a fixed size. For the decay rate  $\lambda = 4.5$ , as expected, the CRLB converges to a positive theoretical value.

**6. Conclusion.** In this paper, we considered the problem of estimating the orientations of features in images. In particular, we examined the orientation of patterns with rotational symmetry. Within the framework of measurement functions composed of steerable filters, we proposed a steerable angle-estimation scheme and proved that the proposed estimator is unbiased. We derived Cramér-Rao lower bounds (CRLB) on the error of estimation. We provided results on the connections to maximum likelihood estimation. Moreover, we discussed the problem of selecting the parameters in the detector functions to achieve the lowest CRLB for a given reference template. In addition to the case of conventional detectors, we also studied the bounds on the performance of steerable wavelet estimators. We proposed an estimator for identifying orientations, and we compared its estimation error to the theoretical bounds. Finally, we provided several experiments on

different realistic junctions and directional patterns that confirm the theory.

**Acknowledgments.** The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 267439. The work was also supported by the Hasler Foundation.

### Appendix A. Background Noise Model.

For our background signal  $S$  to fulfill the requirements of self-similarity and isotropy, we define it as the (non-stationary) solution of a fractional stochastic differential equation. The general theory of such models and their non-Gaussian and sparse extensions is covered in [60].

We assume that our background signal corresponds to a stochastic process  $S$  on  $\mathbb{R}^2$  that is defined as the solution of the stochastic differential equation

$$(A.1) \quad (-\Delta)^{\frac{\gamma}{2}} S = W,$$

where  $W$  is a 2D continuous-domain Gaussian white noise of variance  $\sigma_0^2$  and  $(-\Delta)^{\frac{\gamma}{2}}$  is the fractional Laplacian operator associated with the isotropic Fourier-domain multiplier  $\|\omega\|^\gamma$ , with  $\gamma \geq 0$ .

The intuitive idea here is to shape or “color” the white noise by an appropriately defined inverse fractional Laplacian, which gives to the solution an inverse-power-law spectrum. The field we obtain is the isotropic 2D generalization of the fractional Brownian motion [28].

The observation  $\langle S, f \rangle$  of the random process  $S$  through a suitable test function  $f$  is a Gaussian random variable. The random process  $S$  is Gaussian with 0 mean (because  $W$  is) and correlation form

$$(A.2) \quad \mathcal{B}_S(f, g) = \mathbb{E} \left[ \langle S, f \rangle \overline{\langle S, g \rangle} \right].$$

The random variable  $\langle W, f \rangle$  is well-defined for  $f \in L_2(\mathbb{R}^2)$  [60]. Therefore,  $\langle S, f \rangle$  is well-defined when  $(-\Delta)^{-\gamma/2} f \in L_2(\mathbb{R}^2)$ , which is equivalent, in the Fourier domain, to  $\int_0^\infty \omega^{1-2\gamma} \int_0^{2\pi} |\hat{f}(\omega, \varphi)|^2 d\varphi d\omega < \infty$ . This condition yields (2.14) when  $f = \xi_\alpha$ . Assume that  $f$  has vanishing moments until order  $\lfloor \gamma \rfloor - 1$ . This is, in particular, the case when  $f = (-\Delta)^{\gamma/2} \phi$  with  $\phi$  a smooth enough function in  $L_1(\mathbb{R}^2) \cap L_2(\mathbb{R}^2)$  [10, Proposition 5.1]. Then,  $\langle S, f \rangle$  is a well-defined Gaussian random variable with mean 0 and variance  $\sigma_0^2 \|(-\Delta)^{-\gamma/2} f\|_2^2 = \sigma_0^2 \|\phi\|_2^2$ . The conditions on  $f$  ensure that  $(-\Delta)^{-\gamma/2} f$  is square-integrable. More generally, if  $f$  and  $g$  have enough vanishing moments, then the correlation form (A.2) is evaluated as

$$\mathcal{B}_S(f, g) = \sigma_0^2 \langle (-\Delta)^{-\gamma/2} f, \overline{(-\Delta)^{-\gamma/2} g} \rangle = \frac{\sigma_0^2}{(2\pi)^2} \int_0^\infty \int_0^{2\pi} \omega^{-2\gamma} \hat{f}(\omega, \varphi) \overline{\hat{g}(\omega, \varphi)} d\varphi d\omega,$$

where we have used the Parseval relation and the polar Fourier coordinates for the last equality. We interpret (A) by saying that  $S$  is an isotropic random field with generalized power spectrum  $P_S(\omega) = P_S(\omega) = \omega^{-2\gamma}$ .

### Appendix B. Complement on Steerability.

We briefly recap some basic notions and facts on steerable functions.

**DEFINITION B.1.** *A function  $\xi$  on the plane is steerable in the finite basis  $\{\xi_\alpha\}$  if, for any rotation matrix  $\mathbf{R}_{\theta_0}$ , we can find coefficients  $\{c_\alpha(\theta_0)\}$  such that*

$$(B.1) \quad \xi(\mathbf{R}_{\theta_0} \mathbf{x}) = \sum_\alpha c_\alpha(\theta_0) \xi_\alpha(\mathbf{x}).$$

*A family of functions  $\{\xi_\alpha\}$  is steerable if its linear span is invariant under arbitrary rotations.*

The last part of Definition B.1 is equivalent to saying that, for each  $\theta_0$ , a matrix  $\mathbf{L}(\theta_0)$  exists such that

$$(B.2) \quad \begin{pmatrix} \xi_1(r, \theta + \theta_0) \\ \vdots \\ \xi_n(r, \theta + \theta_0) \end{pmatrix} = \mathbf{L}(\theta_0) \begin{pmatrix} \xi_1(r, \theta) \\ \vdots \\ \xi_n(r, \theta) \end{pmatrix}.$$

An illustrative example of such a family is  $\{\cos(\theta), \sin(\theta)\}$ , whose rotations can be written as

$$(B.3) \quad \begin{pmatrix} \cos(\theta + \theta_0) \\ \sin(\theta + \theta_0) \end{pmatrix} = \begin{pmatrix} \cos(\theta_0) & -\sin(\theta_0) \\ \sin(\theta_0) & \cos(\theta_0) \end{pmatrix} \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix},$$

which is a weighted sum of the unrotated functions.

If  $\xi_\alpha$  is polar-separable as in (2.6), then its Fourier transform is also polar-separable as

$$(B.4) \quad \hat{\xi}_\alpha(\omega, \varphi) = \hat{h}_\alpha(\omega) e^{jn_\alpha \varphi}$$

for some  $\hat{h}_\alpha$  related to  $\eta_\alpha$  and  $n_\alpha$ . We note that this formulation provides a direct connection with the  $n$ th-order complex Riesz transform and leads to a comprehensive theory for the design and analysis of steerable filters and wavelets [11, 25, 58, 40].

### Appendix C. Proof of Theorem 2.1.

The vector  $\mathbf{q}$  is Gaussian as a result of the "integration" of the background noise  $S$ . As such, its law is characterized by its mean vector and covariance matrix. In the right part of (2.5), the first term is deterministic and the second is random with mean 0. Hence, we have

$$(C.1) \quad \mathbb{E}[q_\alpha] = \langle J(\mathbf{R}_{-\theta^*} \cdot), \xi_\alpha \rangle = \langle J, \xi_\alpha(\mathbf{R}_{\theta^*} \cdot) \rangle = e^{jn_\alpha \theta^*} \langle J, \xi_\alpha \rangle,$$

where we used (2.9) for the last equality.

For the covariance matrix, we have that

$$(C.2) \quad \mathbf{C}[\alpha, \beta] = \mathbb{E} \left[ (q_\alpha - \mathbb{E}[q_\alpha]) \overline{(q_\beta - \mathbb{E}[q_\beta])} \right] = \mathbb{E} \left[ \langle S, \xi_\alpha \rangle \overline{\langle S, \xi_\beta \rangle} \right].$$

Then, applying (A) to  $f = \xi_\alpha$  and  $g = \xi_\beta$ , we deduce that

$$(C.3) \quad \mathbf{C}[\alpha, \beta] = \frac{\sigma_0^2}{(2\pi)^2} \left( \int_0^\infty \omega^{1-2\gamma} \widehat{h}_\alpha(\omega) \overline{\widehat{h}_\beta(\omega)} d\omega \right) \delta[n_\alpha - n_\beta],$$

as expected, where we have exploited the Fourier domain expression (B.4).

Finally, the Fisher information of a Gaussian model is provided by the Slepian-Bangs formula [54, B.3.3]. In our case, however, we deal with *complex* Gaussian vectors, because the measurement functions are typically complex (see (2.6)). We therefore use a generalization of the Slepian-Bangs formula adapted to complex Gaussian random vectors [54, B.3.25]. Moreover, this formula simplifies because the covariance matrix does not depend on the parameter  $\theta^*$ , giving (2.17).

### Appendix D. Proof of Theorem 3.1.

As already noted, the CRLBs based on the sets  $H$  and  $H_+$  are the same. We therefore restrict the decision to positive harmonics  $n \in H_+$ .

The functions  $\xi_n$  are assumed to have distinct harmonics; therefore, the matrix (2.16) is diagonal with

$$(D.1) \quad \mathbf{C}[n, n] = \frac{\sigma_0^2}{(2\pi)^2} \int_0^\infty \omega^{1-2\gamma} |\widehat{h}_n(\omega)|^2 d\omega.$$

Knowing that  $\mathbf{C}$  is diagonal, we easily deduce, starting from (2.17), that the Fisher information is in this case

$$(D.2) \quad \text{FI}(\theta^*) = \sum_{n \in H^+} (jn) e^{jn\theta^*} \overline{u_n} (\mathbf{C}[n, n])^{-1} (-jn) e^{-jn\theta^*} u_n = \sum_{n \in H^+} n^2 |u_n|^2 (\mathbf{C}[n, n])^{-1}.$$

Knowing  $\mathbf{C}[n, n]$ , we deduce the Fisher information, and hence the CRLB which is the inverse of the Fisher information. We finally remark that the CRLB does not depend on  $\theta^*$  in that case.

**Appendix E. Construction of the Tridiagonal Covariance Matrix.** The measurements are rearranged in the following way to create a (at most) tridiagonal covariance matrix. Our starting point is an arbitrary set of unique measurement indices of the form  $(n, i)$ , where  $n$  is the harmonic and  $i$  is the scale of the corresponding measurement. As before, since  $(n, i)$  and  $(-n, i)$  carry the same information for real-valued patterns, with  $q_{-n, i} = \overline{q_{n, i}}$  and  $u_{-n, i} = \overline{u_{n, i}}$ , we can assume that all harmonics  $n$  are positive without loss of generality. We divide the set of measurement indices into as few disjoint sets  $G_r$  as possible, subject to three conditions ( $r = 1, \dots, g$  is an index for the sets).

- Each set  $G_r$  consists of indices with a single fixed harmonic  $n_r$ .
- Each set  $G_r$  only contains indices with consecutive scales—it can also contain only a single element.
- If the two sets  $G_r, G_s$  share the same harmonic  $n_r = n_s$ , then the scales in  $G_r, G_s$  differ by a minimum of 2.

The last condition is a consequence of the first two and requires one to have a minimal number of sets. Here,  $g$  denotes the total number of such sets.

From the above conditions, it follows that each set  $G_r$  has some  $l_r$  elements of the form  $(n_r, i_r), \dots, (n_r, i_r + l_r - 1)$ . We use the notation (5.7) and re-index all measurements as  $\tilde{q}_{r|e}$ . The idea behind this re-indexing is that, following (5.6), the covariance matrix of the measurements  $\{\tilde{q}_{r|e}\}$  for each fixed  $r$  is Toeplitz-tridiagonal of the form (5.8). Moreover, the overall covariance matrix is block-diagonal, with  $\mathbf{T}_r$ ,  $r = 1, \dots, g$ , as its diagonal blocks.

**Appendix F. Proof of Theorem 5.2.** We first show that the lower bound in (5.10) is strictly positive; that is,  $B > 2|D|$ . With the notations of Section 5.2, we have that

$$(E1) \quad \begin{aligned} |d_z| &= \frac{1}{2\pi} \left| \int_{\pi/4}^{\pi/2} \omega^z \hat{h}(\omega) \hat{h}(2\omega) \omega d\omega \right| \\ &\leq \frac{1}{2\pi} \sqrt{\int_{\pi/4}^{\pi/2} \omega^z \hat{h}(\omega)^2 \omega d\omega \int_{\pi/4}^{\pi/2} \omega^z \hat{h}(2\omega)^2 \omega d\omega} \\ &= \frac{1}{2\pi} \frac{1}{2^{z/2+1}} \sqrt{\int_{\pi/4}^{\pi/2} \omega^z \hat{h}(\omega)^2 \omega d\omega \int_{\pi/2}^{\pi} \omega^z \hat{h}(\omega)^2 \omega d\omega} \\ &\leq \frac{1}{2\pi} \frac{1}{2^{z/2+2}} \left( \int_{\pi/4}^{\pi/2} \omega^z \hat{h}(\omega)^2 \omega d\omega + \int_{\pi/2}^{\pi} \omega^z \hat{h}(\omega)^2 \omega d\omega \right) \\ &= \frac{1}{2^{z/2+2}} b_z, \end{aligned}$$

where we have used the Cauchy-Schwarz inequality, and the inequality of arithmetic and geometric means. By checking the equality conditions of the two inequalities, we see that they do not happen for the cases of interest. The equality happens only if  $\hat{h}(\omega) = \hat{h}(2\omega)$  almost everywhere, which is not the case here. Thus, the inequality in (E1) is strict. Finally, by selecting  $z = -2\gamma$ , we deduce precisely that  $B > 2|D|$ .

The estimation of  $\theta^*$  from  $\{q_{n,i}\}$  is essentially the same as the estimation from  $\{\tilde{q}_{r|e}\}$ . In particular, the CRLB is the same.

The measurement vector  $\tilde{\mathbf{q}}$  is constructed by concatenating the vectors  $\tilde{\mathbf{q}}_r = (\tilde{q}_{r|1}, \dots, \tilde{q}_{r|l_r})$ ,  $r = 1, \dots, g$ . This is a normal vector, with its mean  $\tilde{\boldsymbol{\mu}}$  given by concatenating the vectors  $\tilde{\boldsymbol{\mu}}_r = (e^{jn_r\theta^*} 2^{-i_r\gamma} u_{r|1}, \dots, e^{jn_r\theta^*} 2^{-(i_r+l_r-1)\gamma} u_{r|l_r})$ ,  $r = 1, \dots, g$ .

The covariance of  $\tilde{\mathbf{q}}$ , as already noted, is block-diagonal of the form

$$(E2) \quad \tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{T}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{T}_g \end{pmatrix},$$

where the form of  $\mathbf{T}_r$  is given by (5.8). Each  $\mathbf{T}_r$  is a Toeplitz-tridiagonal matrix (or the scalar  $B$  if  $l_r = 1$ ). Consequently, its eigenvalues are given by

$$(E3) \quad \lambda_r^{(t)} = B + 2D \cos\left(\frac{t\pi}{l_r + 1}\right),$$

where  $1 \leq t \leq l_r$ . The corresponding eigenvectors are

$$(E4) \quad \mathbf{v}_r^{(t)} = \sqrt{\frac{2}{k+1}} \begin{pmatrix} \sin\left(\frac{t\pi}{l_r+1}\right) \\ \sin\left(2\frac{t\pi}{l_r+1}\right) \\ \vdots \\ \sin\left(k\frac{t\pi}{l_r+1}\right) \end{pmatrix}.$$

For positive  $n_r$  (which we can assume without loss of generality), we have, according to (2.17),

$$(E5) \quad \text{FI}(\theta^*) = 2 \operatorname{Re} \left( \left( \frac{\mathbf{d}}{\mathbf{d}\theta} \tilde{\boldsymbol{\mu}} \right)^{\text{H}} \tilde{\mathbf{C}}^{-1} \left( \frac{\mathbf{d}}{\mathbf{d}\theta} \tilde{\boldsymbol{\mu}} \right) \right) = 2 \sum_{r=1}^g n_r^2 \tilde{\boldsymbol{\mu}}_r^{\text{H}} \mathbf{T}_r^{-1} \tilde{\boldsymbol{\mu}}_r,$$

where

$$(E6) \quad \tilde{\boldsymbol{\mu}}_r = (2^{-i_r\gamma} u_{r|1}, \dots, 2^{-(i_r+l_r-1)\gamma} u_{r|l_r}).$$

Thus, the explicit formula for the Fisher information is

$$(E7) \quad \text{FI}(\theta^*) = 2 \sum_{r=1}^g n_r^2 \sum_{l=1}^{l_r} \left( B + 2D \cos\left(\frac{l\pi}{l_r+1}\right) \right)^{-1} \times \left| \sum_{e=1}^{l_r} 2^{-(i_r+e-1)\gamma} u_{r|e} \sin\left(e\frac{l\pi}{l_r+1}\right) \right|^2.$$

Also,

$$(E8) \quad \frac{2}{B+2|D|} \sum_{r=1}^g n_r^2 \|\tilde{\boldsymbol{\mu}}_r\|_2^2 \leq \text{FI}(\theta^*) \leq \frac{2}{B-2|D|} \sum_{r=1}^g n_r^2 \|\tilde{\boldsymbol{\mu}}_r\|_2^2,$$

where we have used the positivity of  $B - 2|D|$ . This shows that  $\text{FI}(\theta^*)$  is finite if and only if  $\sum_{r=1}^g n_r^2 \|\tilde{\boldsymbol{\mu}}_r\|_2^2$  is finite. Going back to the original indices  $(n, i)$ , the CRLB therefore satisfies (5.10).

**Appendix G. Proof of Proposition 5.3.** The first equivalence directly follows from Theorem 5.2. We now show that the series  $\sum_{n,i} n^2 4^{-i\gamma} |u_{n,i}|^2$  converges if and only if  $(-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \in L_2(\mathbb{R}^2)$ , implying the second equivalence. By assumption on  $J$ , we write

$$(G.1) \quad J(r, \theta) = \sum_{(n,i) \in \mathbb{Z}^2} v_{n,i} \xi_{n,i}(r, \theta).$$

We recall that, according to (53),  $\langle \partial_\theta J, \xi_n \rangle = -\langle J, \partial_\theta \xi_n \rangle$ . Therefore,

$$\begin{aligned}
\left\| (-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \right\|_2^2 &= \left\| \sum_{(n,i) \in \mathbb{Z}^2} v_{n,i} (-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} \xi_{n,i}(r, \theta) \right\|_2^2 \\
&= \left\| \sum_{(n,i) \in \mathbb{Z}^2} j n v_{n,i} (-\Delta)^{\gamma/2} \xi_{n,i}(r, \theta) \right\|_2^2 \\
\text{(G.2)} \quad &= \sum_{n_1, n_2, i_1, i_2} n_1 n_2 v_{n_1, i_1} v_{n_2, i_2}^* \langle (-\Delta)^{\gamma/2} \xi_{n_1, i_1}, (-\Delta)^{\gamma/2} \xi_{n_2, i_2} \rangle.
\end{aligned}$$

With the same polar coordinate computation as in (A), we obtain that  $\langle (-\Delta)^{\gamma/2} \xi_{n_1, i_1}, (-\Delta)^{\gamma/2} \xi_{n_2, i_2} \rangle$  is given by

$$\text{(G.3)} \quad \frac{2^{|i_1 - i_2| - 2 \min(i_1, i_2) \gamma}}{2\pi} \delta[n_1 - n_2] \int_0^\infty \omega^{2\gamma} \hat{h}(\omega) \hat{h}(2^{|i_1 - i_2|} \omega) \omega d\omega.$$

This can be simplified as

$$\text{(G.4)} \quad \langle (-\Delta)^{\gamma/2} \xi_{n_1, i_1}, (-\Delta)^{\gamma/2} \xi_{n_2, i_2} \rangle = \begin{cases} 4^{-i\gamma} b_{-2\gamma}, & \text{if } n_1 = n_2, i_1 = i_2 \\ 2 \cdot 4^{-\min(i_1, i_2) \gamma} d_{-2\gamma}, & \text{if } n_1 = n_2, |i_1 - i_2| = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$\text{(G.5)} \quad \left\| (-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \right\|_2^2 = \sum_{n \in \mathbb{Z}} n^2 \left( \sum_{i \in \mathbb{Z}} (4^{-i\gamma} |v_{n,i}|^2 b_{-2\gamma} + 2 \cdot 4^{-i\gamma} v_{n,i}^* v_{n,i+1} d_{-2\gamma} + 2 \cdot 4^{-(i-1)\gamma} v_{n,i}^* v_{n,i-1} d_{-2\gamma}) \right).$$

By employing the Cauchy-Schwarz inequality,

$$\text{(G.6)} \quad \left| \sum_i 4^{-i\gamma} v_{n,i}^* v_{n,i+1} \right| \leq \sqrt{\left( \sum_i 4^{-i\gamma} |v_{n,i}|^2 \right) \left( \sum_i 4^{-i\gamma} |v_{n,i+1}|^2 \right)} = 2^\gamma \sum_i 4^{-i\gamma} |v_{n,i}|^2.$$

Also, we deduce

$$\text{(G.7)} \quad \left| \sum_i 4^{-(i-1)\gamma} v_{n,i}^* v_{n,i-1} \right| \leq \sqrt{\left( \sum_i 4^{-(i-1)\gamma} |v_{n,i}|^2 \right) \left( \sum_i 4^{-(i-1)\gamma} |v_{n,i-1}|^2 \right)} = 2^\gamma \sum_i 4^{-i\gamma} |v_{n,i}|^2.$$

We arrive at the inequalities

$$\begin{aligned}
(b_{-2\gamma} - 2^{-\gamma+2} |d_{-2\gamma}|) \left( \sum_{(n,i) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2 \right) &\leq \left\| (-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \right\|_2^2 \\
\text{(G.8)} \quad &\leq (b_{-2\gamma} + 2^{-\gamma+2} |d_{-2\gamma}|) \left( \sum_{(n,i) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2 \right).
\end{aligned}$$

Applying (E1) with  $z = -2\gamma$ , we deduce that  $b_{-2\gamma} > 2^{-\gamma+2} |d_{-2\gamma}|$ . Hence, (G.8) means that  $(-\Delta)^{\gamma/2} \frac{\partial}{\partial \theta} J(r, \theta) \in L_2(\mathbb{R}^2)$  if and only if  $\sum_{(n,i) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2 < \infty$ .

The remaining part is to show that  $\sum_{(n,i) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2 < \infty$  if and only if  $\sum_{(n,i) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |u_{n,i}|^2 < \infty$ , from which we deduce the second part of Proposition 5.3. We therefore focus on the relationship between  $u_{n,i}$  and  $v_{n,i}$  and remark that

$$(G.9) \quad u_{n,i} = \langle J(r, \theta), \xi_{n,i} \rangle = \sum_{n',i'} v_{n',i'} \langle \xi_{n',i'}, \xi_{n,i} \rangle.$$

Again, it is straightforward to verify that

$$(G.10) \quad \langle \xi_{n',i'}, \xi_{n,i} \rangle = \begin{cases} b_0, & \text{if } n = n', i = i' \\ 2d_0, & \text{if } n = n', |i - i'| = 1 \\ 0, & \text{otherwise.} \end{cases}$$

Hence,

$$(G.11) \quad u_{n,i} = 2d_0 v_{n,i-1} + b_0 v_{n,i} + 2d_0 v_{n,i+1}$$

and

$$(G.12) \quad 2^{-i\gamma} u_{n,i} = (2^{1-\gamma} d_0) (2^{-(i-1)\gamma} v_{n,i-1}) + b_0 (2^{-i\gamma} v_{n,i}) + (2^{1+\gamma} d_0) (2^{-(i+1)\gamma} v_{n,i+1}).$$

This indicates that, for any given  $n$ , the sequence  $\{2^{-i\gamma} u_{n,i}\}_{i \in \mathbb{Z}}$  is the result of applying to the sequence  $\{2^{-i\gamma} v_{n,i}\}_i$  the FIR filter with  $z$  transform

$$(G.13) \quad F(z) = 2^{1-\gamma} d_0 z^{-1} + b_0 + 2^{1+\gamma} d_0 z.$$

This filter has no poles on the unit circle. Moreover, if

$$(G.14) \quad \frac{b_0}{|d_0|} \neq 2^{1+\gamma} + 2^{1-\gamma},$$

then none of its zeros will lie on the unit circle. Thus,

$$(G.15) \quad 0 < \min_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 \leq \max_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 < \infty.$$

The filtering relationship between the two sequences implies the following inequalities on their energies:

$$(G.16) \quad \left( \min_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 \right) \sum_{i \in \mathbb{Z}} |2^{-i\gamma} v_{n,i}|^2 \leq \sum_{i \in \mathbb{Z}} |2^{-i\gamma} u_{n,i}|^2 \leq \left( \max_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 \right) \sum_{i \in \mathbb{Z}} |2^{-i\gamma} v_{n,i}|^2.$$

Thus,

$$(G.17) \quad \left( \min_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 \right) \sum_{(i,n) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2 \leq \sum_{(i,n) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |u_{n,i}|^2 \leq \left( \max_{\omega \in [0, 2\pi)} |F(e^{j\omega})|^2 \right) \sum_{(i,n) \in \mathbb{Z}^2} n^2 4^{-i\gamma} |v_{n,i}|^2.$$

Consequently,

$$(G.18) \quad \sum_{n,i} n^2 4^{-i\gamma} |v_{n,i}|^2 < \infty \iff \sum_{n,i} n^2 4^{-i\gamma} |u_{n,i}|^2 < \infty,$$

as expected, and the proof is finished.

We finally remark that value of  $\gamma$  that is excluded in the proof satisfies (G.14). We provide the corresponding values for different radial profiles in Table 2.

TABLE 2  
Excluded  $\gamma$  values

Wavelet type	Excluded $\gamma$
Shannon [59]	$\infty$
Simoncelli [38]	0.736
Meyer [7]	1.978
Papadakis [43]	1.476
Held [16]	1.443

REFERENCES

- [1] J. BIGUN, G.H.: *Optimal orientation detection of linear symmetry*, in Proceedings of the First IEEE International Conference on Computer Vision, London, Great Britain, 1987, pp. 433–438.
- [2] J. BIGUN, T. BIGUN, AND K. NILSSON, *Recognition by symmetry derivatives and the generalized structure tensor*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (2004), pp. 1590–1605.
- [3] J. CANNY, *A computational approach to edge detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (1986), pp. 679–698.
- [4] N. CHENOARD AND M. UNSER, *3D steerable wavelets in practice*, IEEE Transactions on Image Processing, 21 (2012), pp. 4522–4533.
- [5] B. DAN, A. W. K. MA, E. H. HÁROZ, J. KONO, AND M. PASQUALI, *Nematic-like alignment in SWNT thin films from aqueous colloidal suspensions*, Industrial & Engineering Chemistry Research, 51 (2012), pp. 10232–10237.
- [6] P. E. DANIELSSON, *Rotation-invariant linear operators with directional response*, Miami, FL, 1980.
- [7] I. DAUBECHIES, *Ten lectures on wavelets*, vol. 61, SIAM, 1992.
- [8] A. DEPEURSINGE AND J. FAGEOT, *Biomedical texture operators and aggregation functions: A methodological review and user's guide*, in Biomedical Texture Analysis, Elsevier, 2018, pp. 55–94.
- [9] A. DEPEURSINGE, A. FONCUBIERTA-RODRIGUEZ, D. VAN DE VILLE, AND H. MÜLLER, *Rotation-covariant texture learning using steerable Riesz wavelets*, IEEE Transactions on Image Processing, 23 (2014), pp. 898–908.
- [10] J. FAGEOT, E. BOSTAN, AND M. UNSER, *Wavelet statistics of sparse and self-similar images*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 2951–2975.
- [11] M. FELSBERG AND G. SOMMER, *The monogenic signal*, IEEE Transactions on Signal Processing, 49 (2001), pp. 3136–3144.
- [12] W. FÖRSTNER, *A feature based correspondence algorithm for image matching*, International Archives of Photogrammetry and Remote Sensing, 26 (1986), pp. 150–166.
- [13] W. FREEMAN AND E. ADELSON, *Steerable filters for early vision, image analysis, and wavelet decomposition*, in Proceedings of the Third International Conference on Computer Vision, IEEE, 1990, pp. 406–415.
- [14] W. FREEMAN AND E. ADELSON, *The design and use of steerable filters*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13 (1991), pp. 891–906.
- [15] W. T. FREEMAN, *Steerable Filters and Local Analysis of Image Structure*, PhD thesis, Massachusetts Institute of Technology, 1992.
- [16] S. HELD, M. STORATH, P. MASSOPUST, AND B. FORSTER, *Steerable wavelet frames based on the Riesz transform*, IEEE Transactions on Image Processing, 19 (2010), pp. 653–667.
- [17] M. JACOB AND M. UNSER, *Design of steerable filters for feature detection using Canny-like criteria*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 26 (2004), pp. 1007–1019.
- [18] G. JACOVITTI AND A. NERI, *Multiresolution circular harmonic decomposition*, IEEE Transactions on Signal Processing, 48 (2000), pp. 3242–3247.
- [19] B. JAHNE, *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 4th ed., 1997.
- [20] H. JIUXIANG, A. RAZDAN, J. FEMIANI, C. MING, AND P. WONKA, *Road network extraction and intersection detection from aerial images by tracking road footprints*, IEEE Transactions on Geoscience and Remote Sensing, 45 (2007), pp. 4144–4157.
- [21] N. KARSSMEIJER AND G. TE BRAKE, *Detection of stellate distortions in mammograms*, IEEE Transactions on Medical Imaging, 15 (1996), pp. 611–619.
- [22] S. M. KAY, *Fundamentals of statistical signal processing, volume I: estimation theory*, Fundamentals of Statistical Signal Processing, Prentice Hall, 1993.
- [23] J. K. KOMINIARCZUK, K. KRAJSEK, AND R. MESTER, *Highly accurate orientation estimation using steerable filters*, in Proceedings of the International Conference on Image Processing, ICIP 2007, September

- 16-19, 2007, San Antonio, Texas, USA, 2007, pp. 277–280.
- [24] U. KÖTHER, *Edge and junction detection with an improved structure tensor*, in Pattern Recognition, vol. 2781 of Lecture Notes in Computer Science, Springer, 2003, pp. 25–32.
- [25] U. KÖTHER, *Low-level feature detection using the boundary tensor*, in Visualization and Processing of Tensor Fields, J. Weickert and H. Hagen, eds., Mathematics and Visualization, Springer Berlin Heidelberg, 2006, pp. 63–79.
- [26] K. KRAJSEK AND R. MESTER, *An unified theory for steerable and quadrature filters*, in VISAPP 2006: Proceedings of the First International Conference on Computer Vision Theory and Applications, 2 Volumes, Setúbal, Portugal, February 25-28, 2006, 2006, pp. 48–55.
- [27] M. MAIRE, P. ARBELAEZ, C. FOWLKES, AND J. MALIK, *Using contours to detect and localize junctions in natural images*, in Proceedings of the 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR'08, Anchorage, Alaska, USA, June 24-26 2008, pp. 1–8.
- [28] B. B. MANDELBROT AND J. W. VAN NESS, *Fractional brownian motions, fractional noises and applications*, SIAM review, 10 (1968), pp. 422–437.
- [29] M. MICHAELIS AND G. SOMMER, *A lie group approach to steerable filters*, Pattern Recogn. Lett., 16 (1995), pp. 1165–1174.
- [30] M. MUEHLICH, D. FRIEDRICH, AND T. AACH, *Design and implementation of multisteerable matched filters*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 34 (2012), pp. 279–291.
- [31] S. OLHEDE AND G. METIKAS, *The monogenic wavelet transform*, IEEE Transactions on Signal Processing, 57 (2009), pp. 3426–3441.
- [32] N. PATTON, T. ASLAM, T. MACGILLIVRAY, I. DEARY, B. DHILLON, R. EIKELBOOM, K. YOGESAN, AND I. CONSTABLE, *Retinal image analysis: Concepts, applications and potential*, Progress in Retinal and Eye Research, 25 (2006), pp. 99–127.
- [33] A. PENTLAND, *Fractal-based description of natural scenes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6 (1984), pp. 661–674.
- [34] P. PERONA, *Steerable-scalable kernels for edge detection and junction analysis*, in Proceedings of the Second European Conference on Computer Vision (ECCV'92), December 1992, pp. 663–672.
- [35] P. PERONA, *Deformable kernels for early vision*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 17 (1995), pp. 488–499.
- [36] B. PESQUET-POPESCU AND J. VEHEL, *Stochastic fractal models for image processing*, IEEE Signal Processing Magazine, 19 (2002), pp. 48–62.
- [37] T.-A. PHAM, M. DELALANDRE, S. BARRAT, AND J.-Y. RAMEL, *Accurate junction detection and reconstruction in line-drawing images*, in Proceedings of the 21st International Conference on Pattern Recognition ICPR'12, Tsukuba Science City, Japan, November 11-15 2012, pp. 693–696.
- [38] J. PORTILLA AND E. SIMONCELLI, *A parametric texture model based on joint statistics of complex wavelet coefficients*, International Journal of Computer Vision, 40 (2000), pp. 49–70.
- [39] Z. PÜSPÖKI, M. STORATH, D. SAGE, AND M. UNSER, *Transforms and operators for directional bioimage analysis: A survey*, in Focus on Bio-Image Informatics, W. De Vos, S. Munck, and J.-P. Timmermans, eds., vol. 219 of Advances in Anatomy, Embryology and Cell Biology, Springer International Publishing, May 21, 2016, ch. 3, pp. 69–93.
- [40] Z. PÜSPÖKI, V. UHLMANN, C. VONESCH, AND M. UNSER, *Design of steerable wavelets to detect multifold junctions*, IEEE Transactions on Image Processing, 25 (2016), pp. 643–657.
- [41] Z. PÜSPÖKI, C. VONESCH, AND M. UNSER, *Detection of symmetric junctions in biological images using 2-D steerable wavelet transforms*, in Proceedings of the Tenth IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'13), San Francisco CA, USA, April 7-11, 2013, pp. 1488–1491.
- [42] M. REISERT AND H. BURKHARDT, *Harmonic filters for generic feature detection in 3D*, in Proceedings of the Thirty-first Annual Symposium of the German Association for Pattern Recognition DAGM'09, J. Denzler, G. Notni, and H. Susse, eds., vol. 5748 of Lecture Notes in Computer Science, Springer, September 9-11 2009, pp. 131–140.
- [43] J. ROMERO, S. ALEXANDER, S. BAID, S. JAIN, AND M. PAPADAKIS, *The geometry and the analytic properties of isotropic multiresolution analysis*, Advances in Computational Mathematics, 31 (2009), pp. 283–328.
- [44] B. S. AND J.-C. OLIVO-MARIN, *Combining local filtering and multiscale analysis for edge, ridge, and curvilinear objects detection*, IEEE Transactions on Image Processing, 19 (2010), pp. 74–84.
- [45] D. SAGE, F. NEUMANN, F. HEDIGER, S. GASSER, AND M. UNSER, *Automatic tracking of individual fluorescence particles: Application to the study of chromosome dynamics*, IEEE Transactions on Image Processing, 14 (2005), pp. 1372–1383.
- [46] D. SCHMITTER, R. DELGADO-GONZALO, G. KRUEGER, AND M. UNSER, *Atlas-free brain segmentation in 3D proton-density-like MRI images*, in Proceedings of the Eleventh IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'14), Beijing, People's Republic of China, April 29-May 2, 2014, pp. 629–632.
- [47] M. SCHUH, J. BANDA, P. BERNASCONI, R. ANGRYK, AND P. MARTENS, *A comparative evaluation of automated solar filament detection*, Solar Physics, 289 (2014), pp. 2503–2524.

- [48] E. SIMONCELLI AND H. FARID, *Steerable wedge filters for local orientation analysis*, IEEE Transactions on Image Processing, 5 (1996), pp. 1377–1382.
- [49] E. SIMONCELLI, W. FREEMAN, E. ADELSON, AND D. HEEGER, *Shiftable multiscale transforms*, IEEE Transactions on Information Theory, 38 (1992), pp. 587–607.
- [50] E. SIMONCELLI AND F. W.T., *The steerable pyramid: A flexible architecture for multi-scale derivative computation*, in Proceedings of the Second IEEE International Conference on Image Processing ICIP'95, vol. 3, Washington, DC, USA, October 23-26 1995, IEEE, pp. 444–447.
- [51] H. SKIBBE, Q. WANG, O. RONNEBERGER, H. BURKHARDT, AND M. REISERT, *Fast computation of 3D spherical Fourier harmonic descriptors—A complete orthonormal basis for a rotational invariant representation of three-dimensional objects*, in Proceedings of the Twelfth IEEE International Conference on Computer Vision Workshops ICCV'09, Kyoto, Japan, September 29 - October 2 2009, IEEE, pp. 1863–1869.
- [52] P. R. SMITH, T. M. PETERS, AND R. H. T. BATES, *Image reconstruction from finite numbers of projections*, Journal of Physics A: Mathematical, Nuclear and General, 6 (1973), p. 361.
- [53] R. SOULARD, P. CARRÉ, AND C. FERNANDEZ-MALOIGNE, *Vector extension of monogenic wavelets for geometric representation of color images*, IEEE Transactions on Image Processing, 22 (2013), pp. 1070–1083.
- [54] P. STOICA AND R. MOSES, *Spectral Analysis of Signals*, Prentice-Hall, Upper Saddle River, NJ, USA, 2005.
- [55] P. TEO, *Theory and Applications of Steerable Functions*, PhD thesis, Stanford University, Stanford, CA, USA, 1998.
- [56] P. TEO AND Y. HEL-OR, *Lie generators for computing steerable functions*, Pattern Recognition Letters, 19 (1998), pp. 7–17.
- [57] P. TEO AND Y. HEL-OR, *Design of multiparameter steerable functions using cascade basis reduction*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 21 (1999), pp. 552–556.
- [58] M. UNSER AND N. CHENOUDARD, *A unifying parametric framework for 2D steerable wavelet transforms*, SIAM Journal on Imaging Sciences, 6 (2013), pp. 102–135.
- [59] M. UNSER, N. CHENOUDARD, AND D. VAN DE VILLE, *Steerable pyramids and tight wavelet frames in  $L_2(\mathbb{R}^d)$* , IEEE Transactions on Image Processing, 20 (2011), pp. 2705–2721.
- [60] M. UNSER AND P. TAFTI, *An Introduction to Sparse Stochastic Processes*, Cambridge University Press, Cambridge, UK, 2014.
- [61] J. WARD AND M. UNSER, *Harmonic singular integrals and steerable wavelets in  $L_2(\mathbb{R}^d)$* , Applied and Computational Harmonic Analysis, 36 (2014), pp. 183–197.
- [62] G.-S. XIA, J. DELON, AND Y. GOUSSEAU, *Accurate junction detection and characterization in natural images*, International Journal of Computer Vision, 106 (2014), pp. 31–56.
- [63] W. YU, K. DANILIDIS, AND G. SOMMER, *Approximate orientation steerability based on angular gaussians*, IEEE Trans. Image Processing, 10 (2001), pp. 193–205.