

# Second-Order Performance Analysis and Unbiased Estimation for the Fitting of Concentric Circles

Ali Al-Sharadqah<sup>1</sup> · K. C. Ho<sup>2</sup>

Received: 8 November 2015 / Accepted: 8 August 2016 / Published online: 25 August 2016  
© Springer Science+Business Media New York 2016

**Abstract** A number of methods, both algebraic and iterative, have been developed recently for the fitting of concentric circles. Previous studies focus on first-order analysis for performance evaluation, which is appropriate only when the observation noise is small so that the bias is insignificant compared to variance. Further studies indicate that the first-order analysis does not appear sufficient in explaining and predicting the performance of an estimator for the fitting problem, especially when the noise level becomes significant. This paper extends the previous study to perform the second-order analysis and evaluate the estimation bias of several concentric circle estimators. The second-order analysis exposes important characteristics of the estimators that cannot be seen from the first-order studies. The insights gained in the theoretical study have led to the development of a new estimator that is unbiased and performs best among the algebraic solutions. An adjusted maximum likelihood estimator is also proposed that can yield an unbiased estimate while maintaining the KCR bound performance.

**Keywords** Concentric circles · Hyper-accurate method · Iterative methods · Mean squared error · Non-iterative methods · Second-order bias

## 1 Introduction

Various objects appear concentric in images. Examples include a compact disk, an iris and the cross section of a transmission pipe. Perhaps, the simplest concentric pattern is formed by circles that have a common center. A set of concentric circles are characterized by the center and the radii. The estimation of the concentric circle parameters is a challenging problem due to the nonlinear relationship between the observed data points and the unknown parameters. It has been an active research and has applications in biometrics such as iris recognition [26], commercial electronics such as camera calibration [15], and industry, such as, steel coils quality control.

Many concentric circle estimators have been developed over the years, including the preliminary solution by Dampagama [10], the suboptimum method by Benko et al. [4], the conic fitting solution by O’Leary et al. [19] and the algebraic solution by Ma and Ho [17]. Recently, Al-Sharadqah and Ho [3] have revisited previously proposed methods and developed new ones for the problem of fitting concentric circles. Both algebraic and iterative methods were investigated, and the first-order analysis was performed. The analytical results indicate that the algebraic solutions are easy to apply, but the performance is suboptimum. The iterative solutions, namely the maximum likelihood estimator (MLE), the gradient weighted algebraic fit (GRAF) and the Renormalization method (RM), can yield the KCR bound [7, 14, 20] performance. Furthermore, the MLE, using the Levenberg–Marquardt implementation, is much more computationally efficient than the GRAF and the RM. Under the first-order analysis, these estimators all have zero bias [3].

The first-order analysis is appropriate when the signal-to-noise ratio (SNR) of the measurements is high, so that the bias is negligible compared to variance. It is not uncommon

---

✉ Ali Al-Sharadqah  
ali.alsharadqah@csun.edu

K. C. Ho  
hod@missouri.edu

<sup>1</sup> Department of Mathematics, California State University–Northridge, Northridge, CA 91330-8313, USA

<sup>2</sup> Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65211, USA

in many practical scenarios where the SNR is not sufficient. In such cases, the first-order analysis will not be a reasonable indication in predicting the performance of an estimator. The first-order analysis also is not adequate to explain the behaviors of different estimators.

This paper continues the previous study [3] to analyze in the second-order performance of various algebraic and iterative concentric circle fitting methods. The second-order analysis is quite challenging and more involved than the first order. Nevertheless, it reveals very interesting results among different estimators. In particular, the concentric circle center estimate has zero essential bias for both the algebraic (except the simple fit method) and iterative solutions. Regarding the radii, the algebraic solutions have essential bias four times higher than that of the MLE and the GRAF and RM two times.

The insights gained from the second-order analysis have led us to develop a hyper-accurate estimator that is among most accurate for an algebraic solution while having zero essential bias up to second order. It is computationally efficient. Although its performance is unable to achieve the KCR bound, it can serve as a promising initial guess for the iterative methods. We also propose the adjusted MLE that has zero bias up to second order and maintains the KCR bound performance.

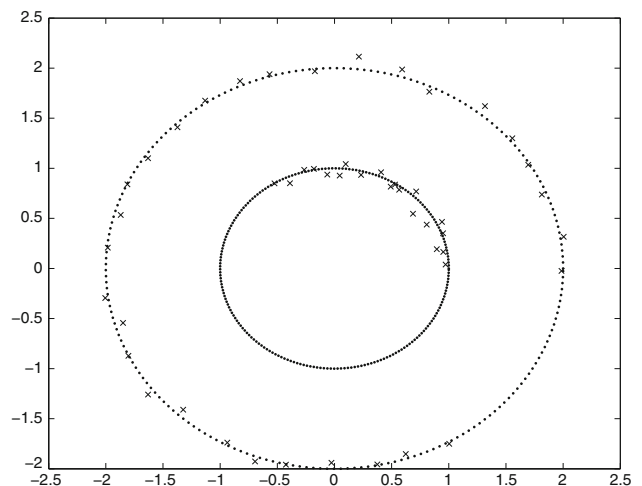
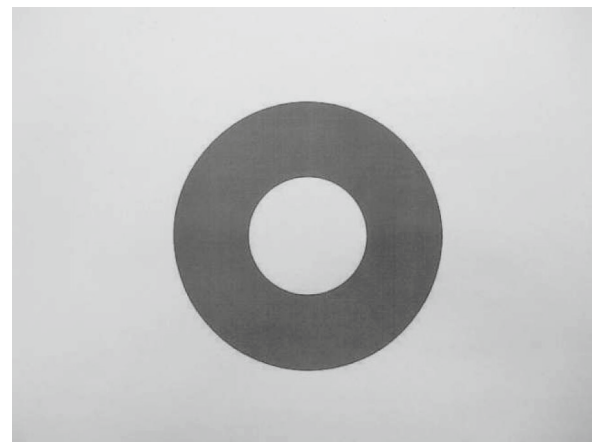
We shall process a realistic data set to support the analysis and the performance of the proposed algorithms. The data are obtained from an image of a CD/DVD, where the Canny edge detection method [5] is applied to obtain the data sample points for the fitting of two concentric circles. The results verify very well the theoretical studies.

For ease of illustration, as shown in Fig. 1, we shall consider two concentric circles. It is straightforward to extend the study to more than two concentric circles. Let us denote  $n_i$  to be the number of data points on the  $i$ th component of the concentric circles, where  $i = 1, 2$ . The data from the concentric circles are modeled as

$$s_{ij} = \tilde{s}_{ij} + \mathbf{n}_{ij}, \quad j = 1, 2, \dots, n_i, \quad i = 1, 2, \quad (1)$$

where  $\mathbf{s}_{ij} = (x_{ij}, y_{ij})^T$  represents the  $2 \times 1$  vector containing the Cartesian coordinates of the  $j$ th point observed around the  $i$ th circle. Also,  $\tilde{\mathbf{s}}_{ij} = (\tilde{x}_{ij}, \tilde{y}_{ij})^T$  is the true value, and  $\mathbf{n}_{ij} = (\delta_{ij}, \epsilon_{ij})^T$  is the observation noise of  $\mathbf{s}_{ij}$ . In this paper, we use the common notations that bold capital letters denote matrices and bold lower case letters represent vectors.  $\mathbf{1}_m$  and  $\mathbf{0}_m$  are length  $m$  vectors of ones and zeros. Also, the symbol tilde represents the true value, and  $\bullet^-$  is the Moore–Penrose pseudoinverse of  $\bullet$ . For an unknown parameter  $\bullet$  to be estimated,  $\hat{\bullet}$  is an estimate of it while  $\bullet$  itself is viewed as a variable for optimization.

Concentric circles fitting problem lies under the umbrella of Errors-in-Variables (EIV) model. Here, we shall adopt the



**Fig. 1** CD/DVD image taken by a digital camera and the extracted noisy data points (shown as cross symbols)

functional model in which the true points  $\tilde{\mathbf{s}}_{ij}$ 's are fixed, but they are unknown and satisfy the following relation

$$P_i(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{s}}_{ij}) := \|\tilde{\mathbf{s}}_{ij} - \tilde{\mathbf{c}}\|^2 - \tilde{R}_i^2 = 0, \quad (2)$$

for each  $i = 1, 2$ , and  $j = 1, \dots, n_i$ . Here, the four-dimensional vector  $\tilde{\boldsymbol{\theta}}$  is the unknown  $\tilde{\boldsymbol{\theta}} = (\tilde{\mathbf{c}}^T, \tilde{R}_1, \tilde{R}_2)^T$  to be estimated, and  $\tilde{\mathbf{c}} = (\tilde{a}, \tilde{b})^T$  is the center of the common concentric circles.  $\tilde{R}_1$  and  $\tilde{R}_2$ , with  $\tilde{R}_1 < \tilde{R}_2$ , are the two circle radii, and  $\|\cdot\|$  is the Euclidean norm.

Expanding the square in Eq. (2) and denoting  $\|\tilde{\mathbf{s}}_{ij}\|^2$  as  $\tilde{z}_{ij}$ , Eq. (2) can be expressed as

$$P_i(\tilde{\boldsymbol{\phi}}, \tilde{\mathbf{s}}_{ij}) = \tilde{A}\tilde{z}_{ij} + \tilde{B}\tilde{x}_{ij} + \tilde{C}\tilde{y}_{ij} + \tilde{D}_i = 0, \quad (3)$$

where  $\tilde{\boldsymbol{\phi}} = (\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}_1, \tilde{D}_2)^T$  is the algebraic parameter vector. Thus, one might alternatively estimate the algebraic parameter vector and then recover the natural parameters through the relationships between the two parametric spaces

$$a = \frac{-B}{2A}, \quad b = \frac{-C}{2A}, \quad R_i^2 = \frac{\zeta_i}{4A^2}, \tag{4}$$

where

$$\zeta_i = B^2 + C^2 - 4AD_i, \quad i = 1, 2. \tag{5}$$

Note here that Eq. (3) can be expressed in a compact form as  $P_i = \hat{\phi}^T \tilde{\mathbf{z}}_{ij}$ , where

$$\begin{aligned} \tilde{\mathbf{z}}_{1j} &= (\tilde{z}_{1j}, \tilde{x}_{1j}, \tilde{y}_{1j}, 1, 0)^T \text{ and} \\ \tilde{\mathbf{z}}_{2j} &= (\tilde{z}_{2j}, \tilde{x}_{2j}, \tilde{y}_{2j}, 0, 1)^T. \end{aligned} \tag{6}$$

For notation simplicity, we shall collect all  $n = n_1 + n_2$  measurement data points together as  $\mathbf{s} = \tilde{\mathbf{s}} + \mathbf{n}$  with  $\mathbf{s} = (\mathbf{s}_1^T, \mathbf{s}_2^T)^T$ ,  $\tilde{\mathbf{s}} = (\tilde{\mathbf{s}}_1^T, \tilde{\mathbf{s}}_2^T)^T$ , and  $\mathbf{n} = (\mathbf{n}_1^T, \mathbf{n}_2^T)^T$ , where  $\mathbf{s}_i = (\mathbf{s}_{i1}^T, \dots, \mathbf{s}_{in_i}^T)^T$ ,  $\mathbf{n}_i = (\mathbf{n}_{i1}^T, \dots, \mathbf{n}_{in_i}^T)^T$ . Here,  $\mathbf{s}$  and  $\mathbf{n}$  are  $2n \times 1$  vectors of the observed points and the noise vectors, respectively, while  $\tilde{\mathbf{s}}$  is the true value of  $\mathbf{s}$ .

The noise vector  $\mathbf{n}$  is modeled as a Gaussian random vector with mean zero and covariance matrix  $\mathbf{Q}$ . It is often reasonable to assume  $\mathbf{n}_{ik}$  and  $\mathbf{n}_{jl}$ , for  $k \neq l$  or  $i \neq j$ , are independent so that  $\mathbf{Q}$  is a diagonal matrix. Here, we look at the specific case that  $\text{cov}(\mathbf{n}_{ij}, \mathbf{n}_{il}) = \sigma^2 \hat{\delta}_{jl}$  for all  $l, j = 1, \dots, n_i$  and  $i = 1, 2$ , where  $\hat{\delta}_{jl}$  represents the Kronecker delta function. Finally, it is worth mentioning here that in all numerical experiments we have conducted, the condition  $\zeta_i > 0$ , for each  $i$ , has fulfilled. In our experiments, we were carefully monitoring this condition. If a single case did appear, then its estimation would not be considered in our calculations.

The paper is organized as follows. Section 2 reviews the existing methods that will be analyzed. Section 3 briefly summarizes the first-order analysis results for completeness. Section 4 focuses on the second-order analysis of the MLE and introduces the adjusted MLE that has zero essential bias. Section 5 is dedicated to the second-order analysis of the algebraic fits. It also develops the new Hyper-accurate fitting method that works best among other algebraic solutions. Section 6 performs the second-order analysis for the iterative methods. The studies in Sects. 4–6 are supported with numerical examples.

Section 7 gives three comprehensive numerical simulations and a practical experimental study with realistic and Sect. 7 summarizes our conclusion. The technical details of this paper are included in ‘‘Appendix.’’

## 2 Background

In this section, we briefly introduce the methods discussed in this paper. A full description of these methods can be found in Al-Sharadqah and Ho [3].

### 2.1 The MLE

Based on our statistical assumptions, the MLE,  $\hat{\theta}_m$ , turns out to be the minimizer of  $\mathcal{F}_1(\theta) = \sum_{i=1}^2 \sum_{j=1}^{n_i} \|\mathbf{s}_{ij} - \tilde{\mathbf{s}}_{ij}\|^2$ , subject to a system of equations given in Eq. (2). Al-Sharadqah and Ho [3] proved that, under the isotropic and homogeneity of the noisy errors that are also adopted in this paper, the MLE is equivalent to orthogonal distance regression (ODR), or the so-called the geometric fit. The geometric fit minimizes the sum of the squares of the orthogonal distances of the observed points  $\mathbf{s}_{ij}$  to the fitted circles indexed by  $(a, b, R_1)$  and  $(a, b, R_2)$ ; i.e.,

$$\mathcal{F}(\theta) = \sum_{i=1}^2 \sum_{j=1}^{n_i} (\|\mathbf{r}_{ij}\| - R_i)^2 = \sum_{i=1}^2 \sum_{j=1}^{n_i} d_{ij}^2, \tag{7}$$

where  $\mathbf{r}_{ij} = \mathbf{s}_{ij} - \mathbf{c}$ . The signed distance  $d_{ij}$  stands for the distance from  $\mathbf{s}_{ij}$  to the circle  $P_i$ . i.e., for each  $i = 1, 2$ , and  $j = 1, \dots, n_i$ ,

$$d_{ij} = r_{ij} - R_i, \quad r_{ij} = \|\mathbf{r}_{ij}\| = \sqrt{(x_{ij} - a)^2 + (y_{ij} - b)^2}. \tag{8}$$

This is a nonlinear least squares problem, and it needs an iterative procedure to obtain its solution  $\hat{\theta}_m$ . In [3], two Levenberg–Marquardt (LM)-based algorithms were proposed to compute its solution.

### 2.2 Gradient Weighted Algebraic Fit

Al-Sharadqah and Ho [3] implemented the so-called gradient weighted algebraic fit (GRAF), a well-known method in estimating the parameters for many subjects, to concentric circles fitting [8, 16, 22]. The GRAF  $\hat{\phi}_g$  is the estimator that minimizes the objective function

$$\mathcal{G}(\phi) = \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{\|\mathbf{z}_{ij}^T \phi\|^2}{\phi^T (\nabla_{\mathbf{s}_{ij}} \mathbf{z}_{ij} \text{cov}(\mathbf{n}_{ij}) \nabla_{\mathbf{s}_{ij}} \mathbf{z}_{ij}^T) \phi}, \tag{9}$$

where

$$\nabla_{\mathbf{s}_{ij}} \mathbf{z}_{ij}^T = \begin{bmatrix} 2x_{ij} & 1 & 0 & 0 & 0 \\ 2y_{ij} & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{10}$$

As an immediate consequence of our statistical assumption  $\text{cov}(\mathbf{n}_{ij}) = \sigma^2 \mathbf{I}_2$ , the objective function  $\mathcal{G}$  in Eq. (9) has the formal expression

$$\mathcal{G}(\phi) = \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{\phi^T \mathbf{M}_{ij} \phi}{\phi^T \mathbf{V}_{ij} \phi}, \tag{11}$$

where

$$\mathbf{M}_{1j} = \begin{bmatrix} z_{1j}^2 & z_{1j}x_{1j} & z_{1j}y_{1j} & z_{1j} & 0 \\ x_{1j}z_{1j} & x_{1j}^2 & x_{1j}y_{1j} & x_{1j} & 0 \\ y_{1j}z_{1j} & x_{1j}y_{1j} & y_{1j}^2 & y_{1j} & 0 \\ z_{1j} & x_{1j} & y_{1j} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{M}_{2j} = \begin{bmatrix} z_{2j}^2 & z_{2j}x_{2j} & z_{2j}y_{2j} & 0 & z_{2j} \\ x_{2j}z_{2j} & x_{2j}^2 & x_{2j}y_{2j} & 0 & x_{2j} \\ y_{2j}z_{2j} & x_{2j}y_{2j} & y_{2j}^2 & 0 & y_{2j} \\ 0 & 0 & 0 & 0 & 0 \\ z_{2j} & x_{2j} & y_{2j} & 0 & 1 \end{bmatrix}, \tag{12}$$

and

$$\mathbf{V}_{ij} = \nabla_{\mathbf{s}_{ij}} P_{ij} \nabla_{\mathbf{s}_{ij}} P_{ij}^T = \begin{bmatrix} \mathbf{V}_{ij}^{(*)} & \mathbf{0}_{3 \times 2} \\ \mathbf{0}_{2 \times 3} & \mathbf{0}_{2 \times 2} \end{bmatrix}, \tag{13}$$

where

$$\mathbf{V}_{ij}^{(*)} = \begin{bmatrix} 4z_{ij} & 2x_{ij} & 2y_{ij} \\ 2x_{ij} & 1 & 0 \\ 2y_{ij} & 0 & 1 \end{bmatrix}. \tag{14}$$

Note that in Eq. (11) the objective function  $\mathcal{G}$  represents the sum of the ratios of two homogeneous quadratic polynomials of degree 2; hence, if  $\hat{\phi}_g$  is the solution of Eq. (11), then  $c\hat{\phi}_g$  is also another solution for any nonzero constant  $c$ . To remove the indeterminacy, we restrict the domain of the solution  $\hat{\phi}_g$  on the unit ball; i.e.,  $\|\hat{\phi}_g\| = 1$ . Moreover, the solution is invariant under translation, rotation, and scaling.

Differentiating  $\mathcal{G}$  in Eq. (11) with respect to  $\phi$ , we obtain the so-called variational equation; i.e.,

$$\mathcal{M}\hat{\phi}_g = \mathcal{L}\hat{\phi}_g, \tag{15}$$

where,

$$\mathcal{M} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \gamma_{ij}^{-1} \mathbf{M}_{ij},$$

$$\mathcal{L} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \gamma_{ij}^{-2} (\hat{\phi}_g^T \mathbf{M}_{ij} \hat{\phi}_g) \mathbf{V}_{ij}, \tag{16}$$

and  $\gamma_{ij} = \hat{\phi}_g^T \mathbf{V}_{ij} \hat{\phi}_g$ . Its solution  $\hat{\phi}_g$  is the generalized eigenvector of the matrix pencil  $(\mathcal{M}, \mathcal{L})$  associated with the unity eigenvalue. Since both  $\mathcal{M}$  and  $\mathcal{L}$  depend on  $\phi$ , Eq. (15) can only be solved by iterative procedures. Two numerical schemes that appear in many computer vision literature were used to solve Eq. (15). Chojnacki et al. [8] proposed the Fundamental Numerical Scheme (FNS), while Leedan

and others [16, 18] proposed the Heteroscedastic-Error-In-Variables Scheme (HEIV).

### 2.3 Renormalization and Generalized GRAF

Kanatani developed the Renormalization method [12, 13], which took a prominent position in solving many practical problems in geometric fitting and computer vision before GRAF was developed. Al-Sharadqah and Ho extended the *Renormalization method* to the concentric circles. In our context, the Renormalization fit  $\hat{\phi}_R$  solves

$$\mathcal{M}\phi = \lambda \mathcal{N}\phi, \tag{17}$$

where  $\mathcal{N} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \gamma_{ij}^{-1} \mathbf{V}_{ij}$ . Then, one finds  $\phi$  as a generalized eigenvector for the matrix pencil  $(\mathcal{M}, \mathcal{N})$ . That is, we solve the generalized eigenvalue problem and choose the generalized eigenvector associated with the smallest positive eigenvalue  $\lambda$ . However, both of the matrices  $\mathcal{M}$  and  $\mathcal{N}$  depend on  $\phi$ ; thus, Kanatani proposed an iterative algorithm to find the solution. More details about all these numerical schemes can be found in the surveys [6, 9]. Since the Renormalization method can be considered as an approximation for GRAF and it is a special case of Eq. (17), we will call the solution of the general equation (17) by the generalized GRAF.

### 2.4 Algebraic Fits

Here, we discuss the all non-iterative algebraic methods that will be analyzed in this paper. They are called algebraic, because they minimize the sum of the squares of the algebraic distances  $\sum_{i=1}^2 \sum_{j=1}^{n_i} \|P_i(\theta, \mathbf{s}_{ij})\|^2$ . Alternatively, the problem can be rewritten in terms of the algebraic parameter  $\phi$  by expressing  $P_i(\phi, \mathbf{s}_{ij})$  as  $\phi^T \mathbf{z}_{ij}$ . Since  $\mathbf{M}_{ij} = \mathbf{z}_{ij} \mathbf{z}_{ij}^T$  and the matrix of the moments  $\mathbf{M} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \mathbf{M}_{ij}$ , the objective function we are seeking to minimize is  $\phi^T \mathbf{M} \phi$ . To avoid the trivial solution, we restrict our domain to  $\|\phi\| = 1$ , but this does not remove the indeterminacy entirely. To address this issue, a quadratic constraint in the parametric subspace of  $\mathbb{R}^5$  must be imposed. That is, we generally use  $\phi^T \mathbf{N} \phi = c$ , for some constant  $c$ . This converts our problem to the minimization problem

$$\mathcal{F}_N(\phi) = \frac{\phi^T \mathbf{M} \phi}{\phi^T \mathbf{N} \phi}. \tag{18}$$

This objective function depends on the constraint matrix  $\mathbf{N}$ , and as such its minimizer’s behavior depends upon  $\mathbf{N}$ . Different choices of  $\mathbf{N}$  lead to new estimators. For instance,  $\mathbf{N}_K = \mathbf{e}'_1 \mathbf{e}'_1^T$  with  $\mathbf{e}'_1 = (1, 0, 0, 0, 0)^T$  is the constraint matrix of the simple fit, the Taubin method has the constraint matrix

$\mathbf{N}_T = \mathbf{V}$ , while the constraint matrices  $\mathbf{N}_A = \mathbf{A}$  and  $\mathbf{N}_B = \mathbf{B}$  for Type-A and Type-B direct methods, respectively, are presented shortly. In this paper, we will investigate these four choices of  $\mathbf{N}$ . A brief description of each method is presented shortly, while details about these methods together with their properties and implementations have been thoroughly discussed in [3].

All algebraic fits share a common aspect. They minimize (18), and as such, one finds the solution by differentiating the objective function (18) with respect to  $\phi$  and equating its derivatives with zero. Therefore, the minimization problem reduces to solving the generalized eigenvalue problem

$$\mathbf{M}\phi = \lambda\mathbf{N}\phi \tag{19}$$

The matrix of the moments  $\mathbf{M}$  is positive definite in general. The only exceptional case that  $\mathbf{M}$  becomes singular when all the points for fitting lie exactly on the concentric circles, but the chance of having this situation is near zero.  $\lambda$  is the generalized eigenvalue for the pencil of matrices  $(\mathbf{M}, \mathbf{N})$ , and  $\phi$  is the corresponding generalized eigenvector. Because  $\|\phi\| = 1$ , we pick the solution as the unit generalized eigenvector associated with the eigenvalue  $\lambda$ . The choice of  $\lambda$  depends on the structure of the constraint matrix  $\mathbf{N}$ .

### 2.4.1 Simple Algebraic Fit

Perhaps, the simplest approach is to minimize the sum of the squares of the algebraic errors. That is,

$$\mathcal{F}_K(\theta) = \sum_{i=1}^2 \sum_{j=1}^{n_i} \|P_i(\theta, \mathbf{s}_{ij})\|^2 = \sum_{i=1}^2 \sum_{j=1}^{n_i} (\|\mathbf{r}_{ij}\|^2 - R_i^2)^2.$$

Equivalently, one minimizes (18) subject to  $A = 1$ . This constraint can be written as  $\phi^T \mathbf{N}_K \phi = 1$ , where  $\mathbf{N}_K = \mathbf{e}'_1 \mathbf{e}'_1{}^T$ . This method is called the simple fit (SF) and it is the fastest method. Moreover, it works well if data are distributed along long concentric circular arcs. However, this feature attribute disappears when data are sampled along short concentric circular arcs. In this paper, we will investigate this issue, and we will explain its behavior analytically.

### 2.4.2 The Taubin Method

The Taubin method is one of the most accurate algebraic method [3]. Its solution  $\hat{\phi}_T$  minimizes the objective function

$$\mathcal{F}_T(\phi) = \frac{\sum_{i=1}^2 \sum_{j=1}^{n_i} \phi^T \mathbf{M}_{ij} \phi}{\sum_{i=1}^2 \sum_{j=1}^{n_i} \phi^T \mathbf{V}_{ij} \phi} = \frac{\phi^T \mathbf{M} \phi}{\phi^T \mathbf{V} \phi}, \tag{20}$$

with

$$\mathbf{N}_T = \mathbf{V} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \mathbf{V}_{ij}. \tag{21}$$

The matrix  $\mathbf{V}_{ij}$  is defined in Eq. (13). Since  $\mathbf{N}_T$  is positive semidefinite,  $\hat{\phi}_T = \operatorname{argmin}_{\|\phi\|=1} \mathcal{F}_T(\phi)$  is the generalized eigenvector of the matrix pencil  $(\mathbf{M}, \mathbf{V})$  corresponding to the smallest positive eigenvalue  $\lambda$ . This solution is an approximation of the GRAF when the double summation is distributed over both of the numerator and the denominator of  $\mathcal{G}$  [cf. Eq. (11)]. Furthermore, regardless of its losses for the optimality criterion, the Taubin fit behaves—in some prospect—like the GRAF. This observation will be investigated analytically in our second-order error analysis.

### 2.4.3 Type-A Direct Method

Another algebraic method that was also proposed as an approximation of GRAF is the Type-A direct fit [3]. It results by presumably assuming that the observed points are close enough to the true curves. Then,  $\phi^T \mathbf{z}_{ij} \approx 0$ . Therefore,  $\phi^T \mathbf{V} \phi = n_1 \zeta_1 + n_2 \zeta_2$ , where  $\zeta_i$  is defined in Eq. (5). Thus, Eq. (11) can be approximated by

$$\mathcal{F}_A(\phi) = \frac{\phi^T \mathbf{M} \phi}{\phi^T \mathbf{N}_A \phi} \tag{22}$$

with

$$\mathbf{N}_A := \mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & -2n_1 & -2n_2 \\ 0 & n & 0 & 0 & 0 \\ 0 & 0 & n & 0 & 0 \\ -2n_1 & 0 & 0 & 0 & 0 \\ -2n_2 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{23}$$

For this method,  $\hat{\phi}_A = \operatorname{argmin}_{\|\phi\|=1} \mathcal{F}_A(\phi)$  is the generalized eigenvector that corresponds to the smallest nonnegative generalized eigenvalue  $\lambda$ .

### 2.4.4 Type-B Direct Method

This method is stemmed from the two geometric constraints of concentric circles. That is,  $\phi$  must satisfy  $\zeta_i > 0$  for each  $i = 1, 2$ . If these two constraints are combined together, we have the constraint  $\zeta_1 + \zeta_2 > 0$ . Note that this is a single homogeneous constraint of degree 2. Therefore, we can set  $2B^2 + 2C^2 - 4A(D_1 + D_2) = 1$ ; i.e.,  $\phi^T \mathbf{N}_B \phi = 1$ , where

$$\mathbf{N}_B = \mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & -2 & -2 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{24}$$

and then we minimize

$$\mathcal{F}_B(\boldsymbol{\phi}) = \frac{\boldsymbol{\phi}^T \mathbf{M} \boldsymbol{\phi}}{\boldsymbol{\phi}^T \mathbf{N}_B \boldsymbol{\phi}}. \tag{25}$$

With a very similar analysis to the Type-A direct fit, one can show that the new estimator  $\hat{\boldsymbol{\phi}}_B = \operatorname{argmin}_{\|\boldsymbol{\phi}\|=1} \mathcal{F}_B(\boldsymbol{\phi})$  is the unit generalized eigenvector of the smallest positive eigenvalue of  $(\mathbf{M}, \mathbf{B})$ . The matrix  $\mathbf{B}$  does not depend on  $n_1$  or  $n_2$ , and it coincides with the Type-A direct fit only whenever  $n_1 = n_2$ . Because of this similarity in its structure to the Type-A direct fit, this method is called the Type-B direct method (or Type-B for short).

### 3 Results of First-Order Error Analysis

To demonstrate the performance of the methods mentioned above, we turn our attention to the statistical error analysis. In general, the accuracy of an estimator  $\hat{\boldsymbol{\theta}}$  of the parameter vector  $\boldsymbol{\theta}$  is measured by its mean squared error.

$$\operatorname{MSE}(\hat{\boldsymbol{\theta}}) = \mathbb{E} \left[ \Delta \hat{\boldsymbol{\theta}} \Delta \hat{\boldsymbol{\theta}}^T \right] = \operatorname{cov}(\hat{\boldsymbol{\theta}}) + \operatorname{bias}(\hat{\boldsymbol{\theta}}) \operatorname{bias}(\hat{\boldsymbol{\theta}})^T,$$

where  $\Delta \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}$ . Evaluating the MSE, however, is not possible to our problem. However, if the noise level  $\sigma$  is not significant—as a realistic assumption in many research subjects in signal and image processing—then we can ignore the higher-order noise terms and approximate the mean squared error by

$$\operatorname{MSE}(\hat{\boldsymbol{\theta}}) = \sigma^2 \mathcal{E}_1 + \mathcal{O}(\sigma^4), \tag{26}$$

where  $\sigma^2 \mathcal{E}_1$  is the *first leading term* of the covariance matrix of  $\hat{\boldsymbol{\theta}}$  and it is the only term of order  $\sigma^2$ . In [3], the statistical analysis was restricted to the first leading term of the MSE, which is also the main term of the covariance matrix.

This leading term has a natural lower bound known by the Kanatani–Cramér–Rao (KCR) lower bound [7]. If the leading term of an estimator achieves this bound, then the estimator is optimal. The optimality issue was one of the main goals of [3]. Al-Sharadqah and Ho evaluated the leading term for each of the methods: the MLE, the GRAF, the Renormalization, and the algebraic fits and they showed that all the iterative methods are optimal in this sense. Indeed, for the MLE, it was shown in [3] that its linear approximation is  $\tilde{\boldsymbol{\theta}} + \Delta_1 \hat{\boldsymbol{\theta}}_m$ , where the first-order error term  $\Delta_1 \hat{\boldsymbol{\theta}}_m = \tilde{\mathbf{K}} \mathbf{f}_1$

and  $\mathbf{f}_1$  presented in Eq. (34) and  $\tilde{\mathbf{K}} = \tilde{\mathbf{R}} \tilde{\mathbf{W}}^T$ . Here,  $\tilde{\mathbf{R}} = (\tilde{\mathbf{W}}^T \tilde{\mathbf{W}})^{-1}$  with  $\tilde{\mathbf{W}}^T = [\tilde{\mathbf{W}}_1^T \ \tilde{\mathbf{W}}_2^T]$  and

$$\tilde{\mathbf{W}}_1 = \begin{bmatrix} \tilde{u}_{11} & \tilde{v}_{11} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{u}_{1n_1} & \tilde{v}_{1n_1} & 1 & 0 \end{bmatrix}, \tag{27}$$

$$\tilde{\mathbf{W}}_2 = \begin{bmatrix} \tilde{u}_{21} & \tilde{v}_{21} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{u}_{2n_2} & \tilde{v}_{2n_2} & 0 & 1 \end{bmatrix},$$

and

$$\tilde{u}_{ij} = \frac{\tilde{x}_{ij} - \tilde{a}}{\tilde{R}_i}, \quad \tilde{v}_{ij} = \frac{\tilde{y}_{ij} - \tilde{b}}{\tilde{R}_i}, \quad i = 1, 2, \quad j = 1, \dots, n_i. \tag{28}$$

Moreover, the linear approximation of the MLE is normally distributed with mean  $\tilde{\boldsymbol{\theta}}$  and covariance matrix  $\sigma^2 \tilde{\mathbf{R}}$ . This first leading term of the covariance of the MLE is expressed in terms of the natural parameters  $\boldsymbol{\theta} = (a, b, R_1, R_2)$  and it coincides with the KCR of  $\boldsymbol{\theta} = \boldsymbol{\theta}$ , which was derived by Ma and Ho [17]; i.e.,  $\operatorname{KCR}(\hat{\boldsymbol{\theta}}) = \sigma^2 \tilde{\mathbf{R}}$ . This shows why the MLE is optimal.

Al-Sharadqah and Ho [3] also applied the perturbation theory to the two other iterative methods and showed that the leading terms of their covariances are identical and equal to

$$\operatorname{cov}(\Delta_1 \hat{\boldsymbol{\phi}}) = \sigma^2 \tilde{\mathcal{M}}^{-1} + \mathcal{O}(\sigma^4). \tag{29}$$

where  $\tilde{\mathcal{M}} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \tilde{\mathbf{z}}_{ij} \tilde{\mathbf{z}}_{ij}^T$  and  $\tilde{\zeta}_i$  is the true value of  $\zeta_i$  [cf. Eq. (5)].

This leading term of the covariances for the GRAF and the generalized GRAF is expressed in terms of the algebraic parameters  $\boldsymbol{\theta} = \boldsymbol{\phi}$  and it coincides with the KCR lower bound [7]. That is, the KCR of  $\hat{\boldsymbol{\phi}}$  is  $\operatorname{KCR}(\hat{\boldsymbol{\phi}}) = \sigma^2 \tilde{\mathcal{M}}^{-1}$ .

It is worth mentioning here that the relationship between the lower bounds in the two parametric spaces is  $\operatorname{KCR}(\hat{\boldsymbol{\theta}}) = \tilde{\mathbf{E}} \operatorname{KCR}(\hat{\boldsymbol{\phi}}) \tilde{\mathbf{E}}^T$  where  $\mathbf{E}$  is a  $4 \times 5$  ‘‘Jacobian’’ matrix that comes from taking partial derivatives in Eq. (4), see [3] for more details. This shows that the leading term of the covariance matrices for both the Renormalization and GRAF equal to the KCR lower bound, and hence, we conclude that they are statistically optimal.

Finally, it was shown in [3] that the leading terms of the covariances of all algebraic fits are the same and equal to

$$\operatorname{cov}(\hat{\boldsymbol{\phi}}_{\text{Alg}}) = \sigma^2 \tilde{\mathbf{M}}^{-1} \left[ \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i \tilde{\mathbf{M}}_{ij} \right] \tilde{\mathbf{M}}^{-1} + \mathcal{O}(\sigma^4), \tag{30}$$

where  $\tilde{\mathbf{M}} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\mathbf{z}}_{ij} \tilde{\mathbf{z}}_{ij}^T$ . Moreover, this covariance matrix does not achieve the KCR lower bound, and this proves that the algebraic methods are suboptimal.

### 4 Second-Order Analysis of MLE

The leading term of the covariance matrix of any algebraic fit does not depend on the constraint matrix  $\mathbf{N}$ . Therefore, all algebraic concentric circle fits have the same covariance matrix (to the leading order). This is because their linear approximations, say  $\hat{\boldsymbol{\theta}}_L = \tilde{\boldsymbol{\theta}} + \Delta_1 \hat{\boldsymbol{\theta}}$ , are equal. Moreover, the leading term of the covariance matrices of algebraic fits do not achieve the KCR lower bound; i.e., they are not statistically optimal. However, the algebraic methods behave differently in practice. For instance, the Taubin method always outperforms the Type-A direct fit. The same observation appears among the iterative (optimal) methods. Some fits always perform better than others, even though they belong to the same class.

To shed more light about their behaviors, we shall trace the second most important terms in their MSEs that come from the bias. That is, we shall expand our analysis by including the second-order error term  $\Delta_2 \hat{\boldsymbol{\theta}}$  and study the quadratic approximations of  $\hat{\boldsymbol{\theta}}$

$$\hat{\boldsymbol{\theta}}_Q = \tilde{\boldsymbol{\theta}} + \Delta_1 \hat{\boldsymbol{\theta}} + \Delta_2 \hat{\boldsymbol{\theta}}.$$

We only need to evaluate  $\Delta_2 \hat{\boldsymbol{\theta}}$ , which is a quadratic form of the noisy error vector  $\mathbf{n}$ . This refines Eq. (26) to

$$\begin{aligned} \text{MSE}(\hat{\boldsymbol{\theta}}_Q) &= \text{cov}(\Delta_1 \hat{\boldsymbol{\theta}}) + \text{MSE}(\Delta_2 \hat{\boldsymbol{\theta}}) \\ &= \sigma^2 \mathcal{E}_1 + \sigma^4 \tilde{\mathbf{b}} \tilde{\mathbf{b}}^T + \sigma^4 \mathcal{E}_2 + \mathcal{O}(\sigma^6). \end{aligned} \tag{31}$$

here,  $\sigma^2 \mathcal{E}_1$  and  $\sigma^4 \mathcal{E}_2$  are the first and the second leading terms of the covariance matrix of  $\hat{\boldsymbol{\theta}}$  and their orders of magnitudes are  $\sigma^2/n$  and  $\sigma^4/n$ , respectively. The term  $\sigma^4 \tilde{\mathbf{b}} \tilde{\mathbf{b}}^T$  comes as the outer product of the second-order error bias  $\mathbb{E}(\Delta_2 \hat{\boldsymbol{\theta}}) = \sigma^2 \tilde{\mathbf{b}}$  with itself. We shall show later that  $\mathcal{O}(\sigma^2)$ -bias,  $\sigma^2 \tilde{\mathbf{b}}$ , can be decomposed into  $\tilde{\mathbf{b}} = \tilde{\mathbf{b}}_1 + \tilde{\mathbf{b}}_2$ . The most important part,  $\sigma^2 \tilde{\mathbf{b}}_1 \sim \sigma^2$ , is referred as the *essential bias* and the least important part,  $\sigma^2 \tilde{\mathbf{b}}_2 \sim \sigma^2/n$ , is called the *nonessential bias*.

It is worth mentioning that we are studying here the quadratic approximation of  $\hat{\boldsymbol{\theta}}$  and its MSE, which coincides with the MSE of  $\hat{\boldsymbol{\theta}}$  up to order of magnitude  $\sigma^4/n$ . That is, there are other  $\mathcal{O}(\sigma^4/n)$ —terms in the full MSE expression that were not included. Those terms are the expected value of the outer product of first- and the third-order errors terms, which are discarded in our analysis. However, if  $n_i$ 's are fairly large, say 20–100, then those ignored terms together with the less important part in  $\sigma^4(\tilde{\mathbf{b}}_2 \tilde{\mathbf{b}}_1^T + \tilde{\mathbf{b}}_1 \tilde{\mathbf{b}}_2^T)$  become negligible.

Accordingly, the only terms are persisted in orders of magnitudes  $\sigma^4$ . They are free of  $n$  and they come only from the essential bias.

The MLE is often considered as one of the best estimators available. We shall begin with the second-order analysis of the MLE.

#### 4.1 Second-Order Bias of MLE

If we restrict the error analysis to the first leading term, then the MLE will be an unbiased and efficient estimator of  $\hat{\boldsymbol{\theta}}_m$ . Higher-order analysis reveals the bias of the MLE persists even for very large sample size. Therefore, we address here the bias of the MLE  $\hat{\boldsymbol{\theta}}_m$ , and find the MSE of its quadratic approximation. This can be accomplished by computing its second-order error term,  $\Delta_2 \hat{\boldsymbol{\theta}}_m$ , and evaluating its expectation. Including all terms of order  $\mathcal{O}_P(\sigma^2)$  of  $d_{ij}$  shows that, for each  $i = 1, 2$  and  $j = 1, \dots, n_i$ ,

$$d_{ij} = q_{ij} - \tilde{u}_{ij} \Delta_2 \hat{a} - \tilde{v}_{ij} \Delta_2 \hat{b} - \Delta_2 \hat{R}_i + \mathcal{O}_P(\sigma^3), \tag{32}$$

where

$$q_{ij} = l_{ij} + f_{2ij}, \tag{33}$$

and

$$l_{ij} = f_{1ij} - (\tilde{u}_{ij} \Delta_1 \hat{a} + \tilde{v}_{ij} \Delta_1 \hat{b} + \Delta_1 R_i), \tag{34}$$

where  $f_{1ij} = \tilde{u}_{ij} \delta_{ij} + \tilde{v}_{ij} \epsilon_{ij}$  and

$$f_{2ij} = \frac{1}{2\tilde{R}_i} (\rho_{ij} - \tau_{ij})^2. \tag{35}$$

We also have defined

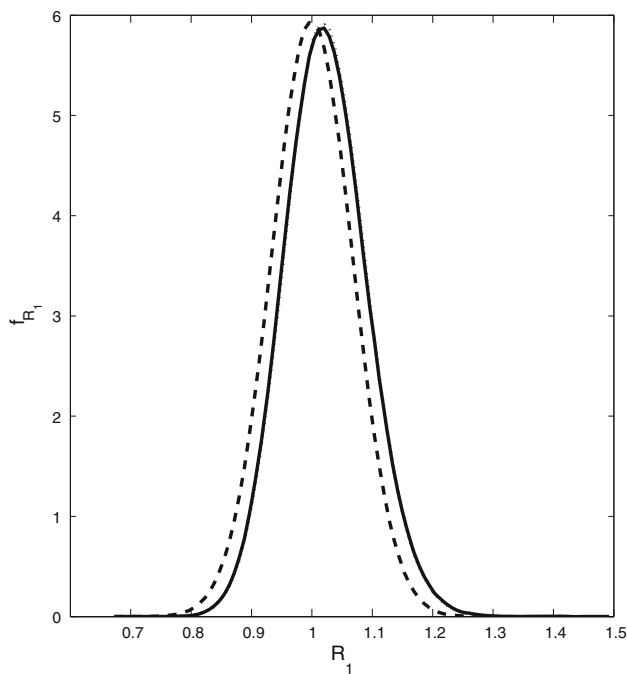
$$\rho_{ij} = \tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij}, \quad \tau_{ij} = \tilde{\mathbf{t}}_{ij}^T \Delta_1 \hat{\boldsymbol{\theta}}_m, \quad \tilde{\mathbf{t}}_{ij} = (-\tilde{v}_{ij}, \tilde{u}_{ij}, 0, 0)^T,$$

and  $\check{\mathbf{n}}_{ij} = (\delta_{ij}, \epsilon_{ij}, 0, 0)^T$ . The derivations of Eqs. (32)–(35) are long and their details can be found in “Appendix.”

The quantity  $q_{ij}$  includes all terms up to order  $\mathcal{O}_P(\sigma^2)$  and it can be expressed in the vector notations as  $\mathbf{q}_i = \mathbf{l}_i + \mathbf{f}_{2i}$ , where the vector  $\mathbf{f}_{2i} = (f_{2i1}, \dots, f_{2in_i})^T$ . Next, minimizing  $\mathcal{F}(\boldsymbol{\theta})$  to the second-order term is equivalent to minimizing

$$\begin{aligned} \mathcal{F}(\Delta_2 \hat{\boldsymbol{\theta}}_m) &= \|\tilde{\mathbf{W}}_1 \Delta_2 \hat{\boldsymbol{\theta}}_m - \mathbf{l}_1 - \mathbf{f}_{21}\|^2 \\ &\quad + \|\tilde{\mathbf{W}}_2 \Delta_2 \hat{\boldsymbol{\theta}}_m - \mathbf{l}_2 - \mathbf{f}_{22}\|^2. \end{aligned} \tag{36}$$

This least squares problem turns into  $\tilde{\mathbf{W}} \Delta_2 \hat{\boldsymbol{\theta}}_m = \mathbf{l} + \mathbf{f}_2$ , where concatenated vectors are  $\mathbf{l} = (\mathbf{l}_1^T, \mathbf{l}_2^T)^T$  and  $\mathbf{f}_2 = (\mathbf{f}_{21}^T, \mathbf{f}_{22}^T)^T$ . Thus,  $\Delta_2 \hat{\boldsymbol{\theta}}_m = \tilde{\mathbf{K}} \mathbf{l} + \tilde{\mathbf{K}} \mathbf{f}_2 = \tilde{\mathbf{K}} \mathbf{f}_2$ , and the quadratic approximation of  $\hat{\boldsymbol{\theta}}_m$  becomes  $\hat{\boldsymbol{\theta}}_Q = \tilde{\boldsymbol{\theta}} + \tilde{\mathbf{K}}(\mathbf{f}_1 + \mathbf{f}_2)$ , where  $\mathbf{f}_1 = (\mathbf{f}_{11}^T, \mathbf{f}_{12}^T)^T$ .



**Fig. 2** Kernel density functions of  $\hat{R}_1$  for the MLE (solid curve) and its linear approximation (dashed curve) and its quadratic approximation (dotted curve)

As a simple numerical experiment to validate the linear and the quadratic approximations, we positioned  $n_1 = 10$  and  $n_2 = 20$  equidistant points on the concentric circular arcs with center  $(0, 0)$  and radii  $\tilde{R}_1=1$  and  $\tilde{R}_2 = 2$ . Both arcs have the same initial angle 0. The terminal angle for the smaller circle is  $\omega_1 = 90^\circ$  and for the larger circle is  $\omega_2 = 120^\circ$ . The true points were contaminated by white Gaussian noise with standard deviation 0.2.  $N = 10^5$  ensemble runs were performed and for each run the estimate of  $\hat{R}_1$  is recorded. The linear and the quadratics approximations of  $\hat{R}_1$  are computed using  $\hat{\theta}_L$  and  $\hat{\theta}_Q$ , while the MLE is computed by the LM algorithm. The kernel densities of these estimates are shown in Fig. 2. As Fig. 2 shows, the kernel density of the linear approximation (dashed curve) significantly differs from the identical kernel densities of the MLE (solid curve) and its quadratic approximation (dotted curve). This example also demonstrates the importance of the second-order error analysis discussed in this paper.

Evaluating the expressions of the bias and the MSE of  $\Delta_2\hat{\theta}_Q$  involves lengthy calculations. To simplify our calculations, let us denote the  $n_i$ -dimensional vectors whose  $j^{\text{th}}$  entities are

$$a_{ij} = \rho_{ij}^2, \quad b_{ij} = -2\tau_{ij}\rho_{ij}, \quad c_{ij} = \tau_{ij}^2, \\ \tilde{h}_{ij} = \tilde{\mathbf{t}}_{ij}^T \tilde{\mathbf{R}} \tilde{\mathbf{t}}_{ij}, \quad j = 1, \dots, n_i$$

by  $\mathbf{a}_i$ ,  $\mathbf{b}_i$ ,  $\mathbf{c}_i$ , and  $\tilde{\mathbf{h}}_i$ , respectively. These vectors play a key role later, so some of their relevant identities are presented in the following lemma. The proof is deferred to ‘‘Appendix.’’

**Lemma 1** For each  $i, k = 1, 2$ , and  $j = 1, \dots, n_i$ , let  $\tilde{\mathbf{H}}_{ik}$  be a  $n_i \times n_k$  matrix with  $jl^{\text{th}}$  entities  $\tilde{h}_{ijkl} = \tilde{\mathbf{t}}_{ij}^T \tilde{\mathbf{R}} \tilde{\mathbf{t}}_{kl}$  and  $\tilde{\mathbf{D}}_{\tilde{\mathbf{h}}_i} = \text{diag}(\tilde{\mathbf{h}}_i)$ . Then,

$$\mathbb{E}(\mathbf{a}_i) = \sigma^2 \mathbf{1}_{n_i}, \quad \mathbb{E}(\mathbf{b}_i) = \mathbf{0}_{n_i}, \quad \mathbb{E}(\mathbf{c}_i) = \sigma^2 \tilde{\mathbf{h}}_i.$$

Moreover, for each  $i, k = 1, 2$ , one has

$$\mathbb{E}(\mathbf{a}_i \mathbf{a}_k^T) = \sigma^4 (\mathbf{1}_{n_i} \mathbf{1}_{n_k}^T + 2\hat{\delta}_{ik} \mathbf{I}_{n_i}),$$

$$\mathbb{E}(\mathbf{b}_i \mathbf{b}_k^T) = 4\hat{\delta}_{ik} \sigma^4 \tilde{\mathbf{D}}_{\tilde{\mathbf{h}}_i},$$

while  $\mathbb{E}(\mathbf{a}_i \mathbf{b}_k^T) = \mathbb{E}(\mathbf{b}_i \mathbf{c}_k^T) = \mathbf{0}_{n_i \times n_k}$  and

$$\mathbb{E}(\mathbf{c}_i \mathbf{c}_k^T) = \sigma^4 (\tilde{\mathbf{h}}_i \tilde{\mathbf{h}}_k^T + 2\tilde{\mathbf{H}}_{ik}),$$

$$\mathbb{E}(\mathbf{a}_i \mathbf{c}_k^T) = \sigma^4 \tilde{\mathbf{h}}_i \cdot \mathbf{1}_{n_k}^T. \tag{37}$$

The first advantage of these identities appears in the evaluation the bias of  $\hat{\theta}_Q$ . Since  $\mathbb{E}(\Delta_1\hat{\theta}_m) = \mathbf{0}$  we only need  $\mathbb{E}(\Delta_2\hat{\theta}_m) = \tilde{\mathbf{K}} \mathbb{E}(\mathbf{f}_2)$ , where  $\mathbf{f}_{2i} = \frac{1}{2\tilde{R}_i}(\mathbf{a}_i + \mathbf{b}_i + \mathbf{c}_i)$ , it is sufficient to find  $\mathbb{E}(\mathbf{a}_i)$ ,  $\mathbb{E}(\mathbf{b}_i)$ , and  $\mathbb{E}(\mathbf{c}_i)$ . From now on, the appearance of the notation  $\check{(\cdot)}$  over a vector means that the vector is divided by  $\frac{1}{2\tilde{R}_i}$ . For instance,  $\check{\mathbf{a}}_i$  and  $\check{\mathbf{h}}_i$  are equal to  $\frac{1}{2\tilde{R}_i} \mathbf{a}_i$  and  $\frac{1}{2\tilde{R}_i} \tilde{\mathbf{h}}_i$ , respectively; but in the case of matrices, we define the matrices  $\check{\tilde{\mathbf{D}}}_{\tilde{\mathbf{h}}_i} = \frac{1}{4\tilde{R}_i^2} \tilde{\mathbf{D}}_{\tilde{\mathbf{h}}_i}$  and  $\check{\tilde{\mathbf{H}}}_{ik} = \frac{1}{4\tilde{R}_i \tilde{R}_k} \tilde{\mathbf{H}}_{ik}$ . With these notations, we decompose  $\mathbf{f}_{2i}$  into  $\mathbf{f}_{2i} = \check{\mathbf{a}}_i + \check{\mathbf{b}}_i + \check{\mathbf{c}}_i$ , for each  $i = 1, 2$ . Based on Lemma 1,

$$\mathbb{E}(\mathbf{f}_2) = \sigma^2 \left[ \check{\mathbf{1}}_{n_1}^T \check{\mathbf{1}}_{n_2}^T \right]^T + \sigma^2 \check{\mathbf{h}}, \quad \text{where } \check{\mathbf{h}} = (\check{\mathbf{h}}_1^T, \check{\mathbf{h}}_2^T)^T. \tag{38}$$

Premultiplying  $\mathbb{E}(\mathbf{f}_2)$  by  $\tilde{\mathbf{K}}$  gives

$$\mathbb{E}(\Delta_2\hat{\theta}) = \sigma^2 \tilde{\mathbf{R}} (\check{\mathbf{w}}_1 + \check{\mathbf{w}}_2 + \check{\mathbf{W}}^T \check{\mathbf{h}}), \tag{39}$$

where

$$\check{\mathbf{w}}_1 = (\check{\mathbf{u}}_1^T \check{\mathbf{1}}_{n_1}, \check{\mathbf{v}}_1^T \check{\mathbf{1}}_{n_1}, \check{n}_1, 0)^T,$$

$$\check{\mathbf{w}}_2 = (\check{\mathbf{u}}_2^T \check{\mathbf{1}}_{n_2}, \check{\mathbf{v}}_2^T \check{\mathbf{1}}_{n_2}, 0, \check{n}_2)^T,$$

and  $\check{n}_i = \frac{n_i}{2\tilde{R}_i}$ . Since  $2\tilde{R}_i \check{\mathbf{w}}_i$  is the  $(2+i)^{\text{th}}$  column of  $\check{\mathbf{W}}^T \check{\mathbf{W}}$ , then

$$2\tilde{R}_1 (\check{\mathbf{W}}^T \check{\mathbf{W}})^{-1} \check{\mathbf{w}}_1 = (0, 0, 1, 0)^T = \mathbf{e}_3$$



and

$$2\tilde{R}_2(\tilde{\mathbf{W}}^T\tilde{\mathbf{W}})^{-1}\tilde{\mathbf{w}}_2 = (0, 0, 0, 1)^T = \mathbf{e}_4.$$

Accordingly,

$$\mathbb{E}(\Delta_2\hat{\boldsymbol{\theta}}_m) = \sigma^2(\check{\mathbf{e}}_3 + \check{\mathbf{e}}_4) + \sigma^2\tilde{\mathbf{K}}\check{\mathbf{h}} := \sigma^2\tilde{\mathbf{b}}_m, \tag{40}$$

where  $\check{\mathbf{e}}_3 = \frac{1}{2\tilde{R}_1}\mathbf{e}_3$  and  $\check{\mathbf{e}}_4 = \frac{1}{2\tilde{R}_2}\mathbf{e}_4$ . The essential bias of the MLE,  $\sigma^2(\check{\mathbf{e}}_3 + \check{\mathbf{e}}_4)$ , depends upon  $\tilde{R}_1$  and  $\tilde{R}_2$  only and it does not rely on the actual data. The second part is of order  $\sigma^2/n_i$ 's, and it represents the nonessential bias. It is worth mentioning here that the MLE of the center does not have essential bias while amounts of the essential bias of  $\hat{R}_i$ 's are  $\frac{\sigma^2}{2\tilde{R}_i}$ .

Finally, our thorough analysis shows that the MSE of  $\hat{\boldsymbol{\theta}}_Q$  is

$$\text{MSE}(\hat{\boldsymbol{\theta}}_Q) = \text{MSE}(\Delta_1\hat{\boldsymbol{\theta}}_m) + \text{MSE}(\Delta_2\hat{\boldsymbol{\theta}}_m), \tag{41}$$

where

$$\text{MSE}(\Delta_2\hat{\boldsymbol{\theta}}_m) = \sigma^4\tilde{\mathbf{b}}_m\tilde{\mathbf{b}}_m^T + \sigma^4\mathcal{E}_2, \tag{42}$$

and

$$\mathcal{E}_2 = 2\tilde{\mathbf{K}}\left(\mathbf{I}_n + \check{\mathbf{H}} + 2\check{\mathbf{D}}_{\tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2}\right)\tilde{\mathbf{K}}^T. \tag{43}$$

The complete derivations of Eqs. (41)–(43) are in ‘‘Appendix.’’

### 4.2 Adjusted MLE

The MLE has second-order bias, which can be adjusted to improve the MLE [cf. Eq. (40)]. Then one can remove the bias to improve estimating the parameters by computing the  $\hat{\boldsymbol{\theta}}_m = (\hat{a}, \hat{b}, \hat{R}_1, \hat{R}_2)$  first and then removing the realization of its bias through

$$\hat{\boldsymbol{\theta}}_{\text{adj}} = \hat{\boldsymbol{\theta}}_m - \hat{\sigma}^2\left(0, 0, \frac{1}{2\hat{R}_1}, \frac{1}{2\hat{R}_2}\right) - \hat{\sigma}^2\hat{\mathbf{K}}\hat{\mathbf{h}}, \tag{44}$$

where  $\hat{\mathbf{K}}, \hat{\mathbf{h}}$ , and  $\hat{\sigma}^2 = \mathcal{F}(\hat{\boldsymbol{\theta}}_m)/(n - 4)$  are the estimating version of  $\tilde{\mathbf{K}}, \tilde{\mathbf{h}}$ , and  $\sigma^2$ , respectively. One can show that these statistics are unbiased estimators for their true versions up to the first leading term. Accordingly, the verification that  $\hat{\boldsymbol{\theta}}_{\text{adj}}$  is unbiased estimator of  $\tilde{\boldsymbol{\theta}}$  (up to order  $\sigma^4$ ) becomes straightforward.

## 5 Second-Order Analysis of Algebraic Fits

The performance of all algebraic fits in the MSE sense depends on their variances and their biases. It is remarkable to note that the leading term of the covariance matrix of any algebraic fit does not depend on the constraint matrix  $\mathbf{N}$ . This observation implies that all algebraic circle fits have the same variance (to the leading order). However, this leading term does not attain the KCR lower bound. Consequently, all algebraic fits are suboptimal. This fact reveals a big departure from the analysis of the single circle fitting by Al-Sharadqah and Chernov [1,2].

Evaluating the bias is an essential step in parametric estimation in order to obtain a good estimator, especially that the iterative methods, such as the MLE, need a good initial guess in order to guarantee their convergence. Besides, in the cases where the noise level is large, all iterative methods, including the MLE, diverge. This makes algebraic fits the only plausible solutions at our hand. Therefore, this section is devoted to analyze the bias of the algebraic fits, such as the simple fit, the Type-A fit, the Type-B fit, and the Taubin fit. The algebraic method proposed in [19] will not be included there. Indeed, its mechanism is not only unclear, but also the quality of its performance is insufficient as demonstrated in [3].

The  $\mathbf{z}_{ij}$  is a quadratic form of the random vector  $\mathbf{s}_{ij} = (x_{ij}, y_{ij})^T$ , which can be written as

$$\mathbf{z}_{ij} = \tilde{\mathbf{z}}_{ij} + \Delta_1\mathbf{z}_{ij} + \Delta_2\mathbf{z}_{ij}. \tag{45}$$

The first-order error  $\Delta_1\mathbf{z}_{ij}$  is a linear combination of  $(\delta_{ij}, \epsilon_{ij})$ . That is,

$$\Delta_1\mathbf{z}_{ij} = \nabla\mathbf{z}_{ij}(\delta_{ij}, \epsilon_{ij})^T = \tilde{\mathbf{a}}_{ij}\delta_{ij} + \tilde{\mathbf{b}}_{ij}\epsilon_{ij},$$

where  $\tilde{\mathbf{a}}_{ij} = (2\tilde{x}_{ij}, 1, 0, 0, 0)^T$  and  $\tilde{\mathbf{b}}_{ij} = (2\tilde{y}_{ij}, 0, 1, 0, 0)^T$ . Accordingly, for each  $j = 1, \dots, n_i$ ,

$$\text{cov}(\Delta_1\mathbf{z}_{ij}, \Delta_1\mathbf{z}_{kl}) = \sigma^2\hat{\delta}_{ik}\hat{\delta}_{jl}\tilde{\mathbf{V}}_{ij},$$

where  $\tilde{\mathbf{V}}_{ij} = \tilde{\mathbf{a}}_{ij}\tilde{\mathbf{a}}_{ij}^T + \tilde{\mathbf{b}}_{ij}\tilde{\mathbf{b}}_{ij}^T$ . Moreover, the second-order error term is  $\Delta_2\mathbf{z}_{ij} = (\delta_{ij}^2 + \epsilon_{ij}^2)\mathbf{e}'_1$  and  $\mathbb{E}(\Delta_2\mathbf{z}_{ij}) = 2\mathbf{e}'_1$ , where  $\mathbf{e}'_1 = (1, 0, 0, 0, 0)^T$ .

Now, using Taylor series expansion, we can decompose  $\Delta\hat{\boldsymbol{\phi}} = \hat{\boldsymbol{\phi}} - \tilde{\boldsymbol{\phi}}$  as

$$\Delta\hat{\boldsymbol{\phi}} = \Delta_1\hat{\boldsymbol{\phi}} + \Delta_2\hat{\boldsymbol{\phi}} + \mathcal{O}_P(\sigma^3),$$

where  $\Delta_1\hat{\boldsymbol{\phi}}$  is a linear form of  $\delta_{ij}$ 's and  $\epsilon_{ij}$ 's and  $\Delta_2\hat{\boldsymbol{\phi}}$  is a quadratic form of  $\delta_{ij}$ 's and  $\epsilon_{ij}$ 's, and all other higher-order terms are represented by  $\mathcal{O}_P(\sigma^3)$ . To assess the bias, we need to evaluate  $\mathbb{E}(\Delta_2\hat{\boldsymbol{\phi}})$  because  $\mathbb{E}(\Delta_1\hat{\boldsymbol{\phi}}) = \mathbf{0}$ . Since all algebraic

fits solve the equation  $\mathbf{M}\hat{\boldsymbol{\phi}} = \lambda\mathbf{N}\hat{\boldsymbol{\phi}}$ , then if we apply matrix perturbation to  $\lambda$ ,  $\hat{\boldsymbol{\phi}}$ ,  $\mathbf{M}$  and  $\mathbf{N}$ , then  $\mathbf{M}\hat{\boldsymbol{\phi}} = \lambda\mathbf{N}\hat{\boldsymbol{\phi}}$  becomes

$$\begin{aligned} (\tilde{\mathbf{M}} + \Delta_1\mathbf{M} + \dots)(\tilde{\boldsymbol{\phi}} + \Delta_1\hat{\boldsymbol{\phi}} + \dots) &= (\tilde{\lambda} + \Delta_1\lambda + \dots) \\ (\tilde{\mathbf{N}} + \Delta_1\mathbf{N} + \dots)(\tilde{\boldsymbol{\phi}} + \Delta_1\hat{\boldsymbol{\phi}} + \dots). \end{aligned} \tag{46}$$

From the first-order analysis [3], it was shown that  $\tilde{\lambda} = \Delta_1\lambda = 0$ , and

$$\Delta_1\hat{\boldsymbol{\phi}} = -\tilde{\mathbf{M}}^{-1}\Delta_1\mathbf{M}\tilde{\boldsymbol{\phi}}, \text{ where } \Delta_1\mathbf{M} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \Delta_1\mathbf{M}_{ij}. \tag{47}$$

and

$$\Delta_1\mathbf{M}_{ij} = \tilde{\mathbf{z}}_{ij} \Delta_1\mathbf{z}_{ij}^T + \Delta_1\mathbf{z}_{ij} \tilde{\mathbf{z}}_{ij}^T.$$

Hence, equating all  $\mathcal{O}_P(\sigma^2)$ -terms in Eq. (46) yields

$$\tilde{\mathbf{M}} \Delta_2\hat{\boldsymbol{\phi}} = - \left( \Delta_1\mathbf{M}\Delta_1\hat{\boldsymbol{\phi}} + \Delta_2\mathbf{M}\tilde{\boldsymbol{\phi}} \right) + \Delta_2\lambda\tilde{\mathbf{N}}\tilde{\boldsymbol{\phi}}, \tag{48}$$

where

$$\begin{aligned} \Delta_2\mathbf{M} &= \sum_{i=1}^2 \sum_j \Delta_2\mathbf{M}_{ij}, \quad \Delta_2\mathbf{M}_{ij} \\ &= \Delta_1\mathbf{z}_{ij} \Delta_1\mathbf{z}_{ij}^T + 2\mathcal{S} \left[ \Delta_2\mathbf{z}_{ij} \tilde{\mathbf{z}}_{ij}^T \right], \end{aligned}$$

and  $\mathcal{S}$  is the symmetrization operator; i.e.,  $\mathcal{S}(\bullet) = (\bullet + \bullet^T)/2$ . Next, we find  $\Delta_2\lambda$ . Premultiplying Eq. (48) and using Eq. (47) lead to

$$\Delta_2\lambda = \frac{\tilde{\boldsymbol{\phi}}^T \mathbf{R} \tilde{\boldsymbol{\phi}}}{\tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{N}} \tilde{\boldsymbol{\phi}}} \text{ where } \mathbf{R} = \Delta_2\mathbf{M} - \Delta_1\mathbf{M}\tilde{\mathbf{M}}^{-1}\Delta_1\mathbf{M}. \tag{49}$$

Since  $\Delta_2\hat{\boldsymbol{\phi}}$  is *not* orthogonal to  $\tilde{\boldsymbol{\phi}}$ , we decompose  $\Delta_2\hat{\boldsymbol{\phi}} = \Delta_2^{\parallel}\hat{\boldsymbol{\phi}} + \Delta_2^{\perp}\hat{\boldsymbol{\phi}}$  into the components that are parallel and orthogonal to  $\tilde{\boldsymbol{\phi}}$ . Then,  $\Delta_2^{\parallel}\hat{\boldsymbol{\phi}} = -\frac{1}{2}\|\Delta_1\hat{\boldsymbol{\phi}}\|^2\tilde{\boldsymbol{\phi}}$  and  $\mathbb{E}(\Delta_2^{\parallel}\hat{\boldsymbol{\phi}}) = -\frac{1}{2}\sigma^2(\text{tr}\tilde{\mathbf{M}}^{-1})\tilde{\boldsymbol{\phi}}$ , which really accounts for the curvature of the unit sphere  $\|\hat{\boldsymbol{\phi}}\| = 1$  rather than represents the bias of  $\hat{\boldsymbol{\phi}}$ . Our goal is to evaluate  $\mathbb{E}(\Delta_2^{\perp}\hat{\boldsymbol{\phi}})$ , and we will suppress the  $\perp$  sign for brevity. Therefore, Eqs. (48)–(49) lead to

$$\Delta_2\boldsymbol{\phi} = \tilde{\mathbf{M}}^{-1} \left( \frac{\tilde{\boldsymbol{\phi}}^T \mathbf{R} \tilde{\boldsymbol{\phi}}}{\tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{N}} \tilde{\boldsymbol{\phi}}} \tilde{\mathbf{N}} \tilde{\boldsymbol{\phi}} - \mathbf{R} \tilde{\boldsymbol{\phi}} \right) = \tilde{\mathbf{M}}^{-1} \mathbf{S}_{\tilde{\mathbf{N}}} \mathbf{R} \tilde{\boldsymbol{\phi}} \tag{50}$$

with

$$\mathbf{S}_{\tilde{\mathbf{N}}} = \frac{\tilde{\mathbf{N}} \tilde{\boldsymbol{\phi}} \tilde{\boldsymbol{\phi}}^T}{\tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{N}} \tilde{\boldsymbol{\phi}}} - \mathbf{I}_5. \tag{51}$$

If  $\mathbf{R}$  is decomposed into  $\mathbf{R}_1 - \mathbf{R}_2$  with  $\mathbf{R}_1 = \Delta_2\mathbf{M}$  and  $\mathbf{R}_2 = \Delta_1\mathbf{M}\tilde{\mathbf{M}}^{-1}\Delta_1\mathbf{M}$ , then

$$\mathbb{E}(\mathbf{R}_1) = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\mathbf{V}}_{ij} + 4\mathcal{S} \left[ \mathbf{e}'_1 \tilde{\mathbf{z}}_{ij}^T \right]. \tag{52}$$

Moreover, analogous to [2,21], a thorough analysis here shows that

$$\begin{aligned} \mathbb{E}(\mathbf{R}_2) &= \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} 2\mathcal{S}[\tilde{\mathbf{V}}_{ij}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{M}}_{ij}] + \tilde{\psi}_{ij}^{(a)}\tilde{\mathbf{V}}_{ij} \\ &\quad + \text{tr}(\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{V}}_{ij})\tilde{\mathbf{M}}_{ij}, \end{aligned} \tag{53}$$

where  $\tilde{\psi}_{ij}^{(a)} = \tilde{\mathbf{z}}_{ij}^T \tilde{\mathbf{M}}^{-1} \tilde{\mathbf{z}}_{ij}$ .

It is worth mentioning here that  $\mathbb{E}(\mathbf{R}_1) \sim n\sigma^2$  and  $\mathbb{E}(\mathbf{R}_2) \sim \sigma^2$ . Since  $\tilde{\mathbf{M}}^{-1} \sim \mathcal{O}(n^{-1})$ , then the contributions of  $\mathbb{E}(\mathbf{R}_1)$  and  $\mathbb{E}(\mathbf{R}_2)$  to their biases are of orders of magnitudes  $\sigma^2$  and  $\sigma^2/n$ , respectively. Therefore, the most important term of  $\mathbf{R}$  is  $\mathbf{R}_1$ , which has a key role in our analysis, while  $\mathbf{R}_2$  has a less significant contribution to the MSE. Now, combining Eqs. (52) and (53) with Eq. (49) forms

$$\begin{aligned} \mathbb{E}(\mathbf{R}) &= \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \left( 1 - \tilde{\psi}_{ij}^{(a)} \right) \tilde{\mathbf{V}}_{ij} + 2\mathcal{S} \left[ 2\mathbf{e}'_1 \tilde{\mathbf{z}}_{ij}^T \right] \right. \\ &\quad \left. - 2\mathcal{S} \left[ \tilde{\mathbf{V}}_{ij}\tilde{\mathbf{M}}^{-1}\tilde{\mathbf{M}}_{ij} \right] - \text{tr} \left( \tilde{\mathbf{M}}^{-1}\tilde{\mathbf{V}}_{ij} \right) \tilde{\mathbf{M}}_{ij} \right]. \end{aligned} \tag{54}$$

Furthermore, since  $\tilde{\mathbf{z}}_{ij}^T \tilde{\boldsymbol{\phi}} = 0$  and  $\tilde{\boldsymbol{\phi}}^T \mathbf{e}'_1 = \tilde{A}$ , we obtain

$$2\tilde{A}\tilde{\mathbf{M}}^{-1}\mathbf{S}_{\tilde{\mathbf{N}}} \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\mathbf{z}}_{ij} = -2\tilde{A}\tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\mathbf{z}}_{ij} := I_{\tilde{\boldsymbol{\phi}}}.$$

This simplifies the general form of the algebraic fits to

$$\begin{aligned} \mathbb{E}(\Delta_2\hat{\boldsymbol{\phi}}) &= \sigma^2\tilde{\mathbf{M}}^{-1}\mathbf{S}_{\tilde{\mathbf{N}}} \sum_{i=1}^2 \sum_{j=1}^{n_i} 2\tilde{A}\tilde{\mathbf{z}}_{ij} \\ &\quad + \left( \left( 1 - \tilde{\psi}_{ij}^{(a)} \right) \mathbf{I}_5 - \tilde{\mathbf{M}}_{ij}\tilde{\mathbf{M}}^{-1} \right) \tilde{\mathbf{V}}_{ij}\tilde{\boldsymbol{\phi}} \\ &= \sigma^2 I_{\tilde{\boldsymbol{\phi}}} + \sigma^2\tilde{\mathbf{M}}^{-1}\mathbf{S}_{\tilde{\mathbf{N}}} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \left( 1 - \tilde{\psi}_{ij}^{(a)} \right) \mathbf{I}_5 \right. \\ &\quad \left. - \tilde{\mathbf{M}}_{ij}\tilde{\mathbf{M}}^{-1} \right] \tilde{\mathbf{V}}_{ij}\tilde{\boldsymbol{\phi}}. \end{aligned} \tag{55}$$

Equation (55) provides the general expression of the bias for any algebraic fit. Each algebraic fit has its specific matrix  $\mathbf{N}$ , and as such, each method has its own bias.

**Bias of the Taubin fit** In this case, recall that  $\tilde{\mathbf{N}} = \tilde{\mathbf{V}} = \sum_i^2 \sum_j^{n_i} \tilde{\mathbf{V}}_{ij}$  [cf. Eq. (13)] and  $\tilde{y}_{ij} = \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} = \tilde{\zeta}_i$ . Thus, a thorough analysis shows that

$$\mathbb{E}(\Delta_2 \hat{\phi}_T) = \sigma^2 I_{\tilde{\phi}} + \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \tilde{\psi}_{ij}^{(a)} \left( 1 - \frac{\tilde{\zeta}_i}{\tilde{\zeta}} \right) \mathbf{I}_5 + \tilde{\mathbf{M}}_{ij} \tilde{\mathbf{M}}^{-1} \right] \tilde{\mathbf{V}}_{ij} \tilde{\phi}, \tag{56}$$

where

$$\tilde{\zeta} = \tilde{\phi}^T \tilde{\mathbf{V}} \tilde{\phi} = n_1 \tilde{\zeta}_1 + n_2 \tilde{\zeta}_2.$$

**Bias of Type-A direct fit**  $\hat{\phi}_A$  The matrix  $\mathbf{N} = \mathbf{A}$  is free of data. Thus,  $\tilde{\mathbf{N}} = \mathbf{A}$  and

$$\mathbb{E}(\Delta_2 \hat{\phi}_A) = \sigma^2 I_{\tilde{\phi}} - \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \left( 1 - \tilde{\psi}_{ij}^{(a)} \right) \mathbf{I}_5 - \tilde{\mathbf{M}}_{ij} \tilde{\mathbf{M}}^{-1} \right) \tilde{\mathbf{V}}_{ij} \tilde{\phi} + \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^n \tilde{\zeta}^{-1} \tilde{\zeta}_i \left( 1 - \tilde{\psi}_{ij}^{(a)} \right) \mathbf{A} \tilde{\phi}. \tag{57}$$

Now, the very helpful relation

$$\mathbf{A} \tilde{\phi} = \tilde{\mathbf{V}} \tilde{\phi} - 2\tilde{A} \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{z}_{ij} \tag{58}$$

reduces the expression of  $\mathbb{E}(\Delta_2 \hat{\phi}_A)$  to

$$\mathbb{E}(\Delta_2 \hat{\phi}_A) = 2\sigma^2 I_{\tilde{\phi}} - \alpha \sigma^2 \tilde{\mathbf{M}}^{-1} \mathbf{A} \tilde{\phi} + \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \tilde{\mathbf{M}}_{ij} \tilde{\mathbf{M}}^{-1} + \tilde{\psi}_{ij}^{(a)} \mathbf{I}_5 \right) \tilde{\mathbf{V}}_{ij} \tilde{\phi}, \tag{59}$$

where  $\alpha = \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}^{-1} \tilde{\zeta}_i \tilde{\psi}_{ij}^{(a)}$ . The only terms in Eqs. (56) and (59) with order of magnitude  $\sigma^2$  are  $\sigma^2 I_{\tilde{\phi}}$  and  $2\sigma^2 I_{\tilde{\phi}}$ .

**Bias of Type-B direct fit**  $\hat{\phi}_B$  The matrix  $\mathbf{N} = \mathbf{B}$  is free of data. In this case,  $\tilde{\mathbf{N}} = \mathbf{B}$  and

$$\mathbb{E}(\Delta_2 \hat{\phi}_B) = 2\sigma^2 I_{\tilde{\phi}} + \sigma^2 \tilde{\mathbf{M}}^{-1} \left( \frac{n_1 \tilde{\zeta}_1 (1 - \tilde{\psi}_1) + n_2 \tilde{\zeta}_2 (1 - \tilde{\psi}_2)}{\tilde{\zeta}_1 + \tilde{\zeta}_2} \mathbf{B} - \mathbf{A} \right) \tilde{\phi} + \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \tilde{\mathbf{M}}_{ij} \tilde{\mathbf{M}}^{-1} + \tilde{\psi}_{ij}^{(a)} \mathbf{I}_5 \right) \tilde{\mathbf{V}}_{ij} \tilde{\phi},$$

where  $\tilde{\psi}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \tilde{\psi}_{ij}^{(a)}$ .

**Bias of simple algebraic fit** The same approach applies here to derive the bias the simple fit, in which its constraint matrix

$\mathbf{N} = \mathbf{e}'_1 \mathbf{e}_1^T$ . Then, after direct algebraic manipulations, we obtain

$$\mathbb{E}(\Delta_2 \phi_K) = 2\sigma^2 I_{\tilde{\phi}} + \sigma^2 \tilde{\mathbf{M}}^{-1} \left( \tilde{A}^{-2} \tilde{\zeta} \mathbf{e}'_1 \mathbf{e}_1^T - \tilde{\mathbf{A}} \right) \tilde{\phi} + \sigma^2 \tilde{\mathbf{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ -\tilde{A}^{-1} \tilde{\psi}_{ij}^{(a)} \tilde{\zeta}_i \mathbf{e}'_1 + \left( \tilde{\mathbf{M}}_{ij} \tilde{\mathbf{M}}^{-1} + \tilde{\psi}_{ij}^{(a)} \mathbf{I}_5 \right) \tilde{\mathbf{V}}_{ij} \tilde{\phi} \right].$$

### 5.1 Comparisons Between Algebraic Fits

From the error analysis, the second-order biases of the Taubin fit, the Type-A direct fit, the Type-B direct fit, and the simple algebraic fit have been derived. For comparison purpose, it is more appropriate to convert their bias expressions to the natural geometric parameters  $\theta$ . They are related by  $\mathbb{E}(\Delta_2 \hat{\theta}) = \tilde{\mathbf{E}} \mathbb{E}(\Delta_2 \hat{\phi}) + \mathcal{O}(\sigma^3)$ , where  $\tilde{\mathbf{E}}$  is the Jacobian matrix. To simplify the comparisons, we shall decide that the smaller the essential bias, the better the estimator. For fairly large  $n$ , only the essential bias in each of Eqs. (56) and (59) persist.

The essential bias of the Taubin fit in term of its natural geometric parameter  $\theta$  is

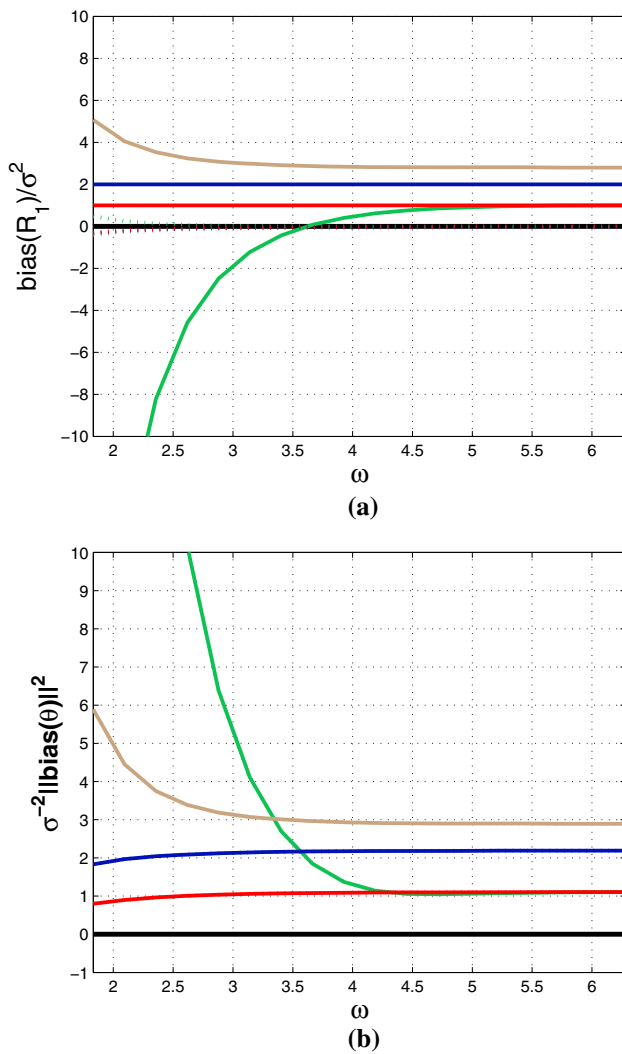
$$\sigma^2 \tilde{\mathbf{E}} I_{\tilde{\phi}} = \sigma^2 \left( 0, 0, \tilde{R}_1^{-1}, \tilde{R}_2^{-1} \right)^T = \sigma^2 I_{\tilde{\theta}},$$

implying its center has a zero essential bias while the radii have data-free essential biases. Consequently, we can see now that the Type-A fit has essential bias twice as much as the bias of the Taubin fit. This answers why and by how much the Taubin fit outperforms the Type-A direct fit. The essential bias for the two other fits is listed in Table 1, and they depend on the true values of the parameters and the observations.

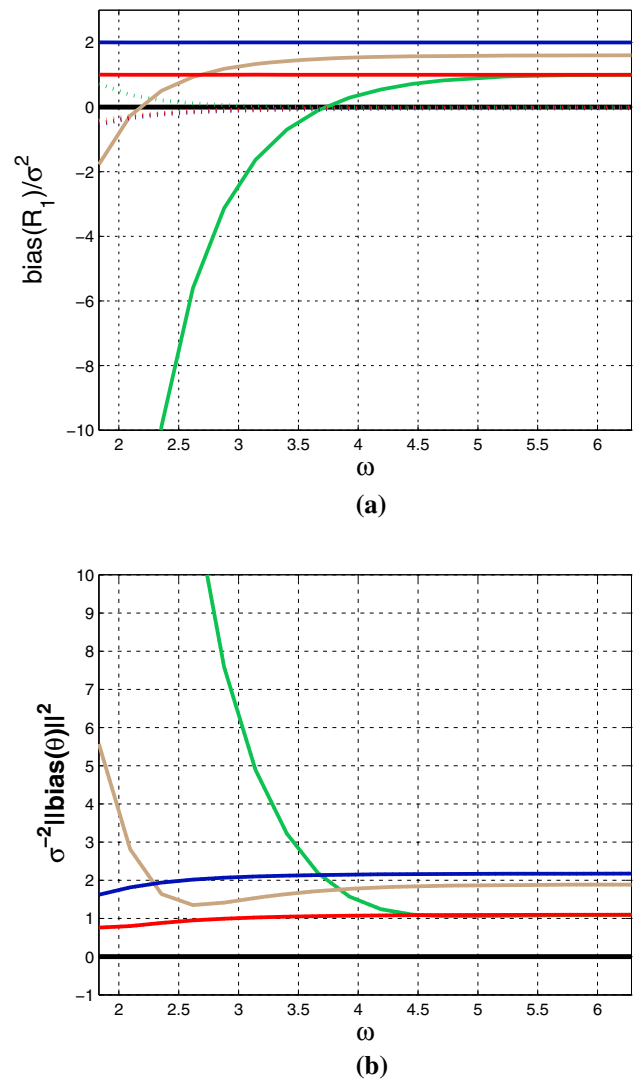
To analytically investigate their performances, we plotted both parts of the normalized bias (NB) of  $\hat{R}_1$ ; i.e.,  $\text{NB}(\hat{R}_1) = \frac{\mathbb{E}(\hat{R}_1)}{\sigma^2}$  and the normalized bias of  $\hat{\theta}$ ; i.e.,  $\text{NB}(\hat{\theta}) = \frac{\|\mathbb{E}(\Delta_2 \hat{\theta})\|}{\sigma^2}$ . Firstly, we positioned  $n_1 = 50$  and  $n_2 = 100$  equidistant points on concentric circular arcs with central angles initiated at  $0^\circ$  and terminated at  $\omega^\circ$  (i.e.,  $\frac{\pi}{180^\circ} \omega^\circ$  radian). The center

**Table 1** Essential biases of the algebraic fits in terms the geometric parameter vector  $\theta$

	Essential bias of $\hat{\theta}$
Taubin fit	$\sigma^2 I_{\tilde{\theta}}$
Type-A	$2\sigma^2 I_{\tilde{\theta}}$
Type-B	$2\sigma^2 I_{\tilde{\theta}} + \sigma^2 \tilde{\mathbf{E}} \tilde{\mathbf{M}}^{-1} \left( \frac{n_1 \tilde{\zeta}_1 + n_2 \tilde{\zeta}_2}{\tilde{\zeta}_1 + \tilde{\zeta}_2} \mathbf{B} - \tilde{\mathbf{A}} \right) \tilde{\phi}$
Simple fit	$2\sigma^2 I_{\tilde{\theta}} + \sigma^2 \tilde{\mathbf{E}} \tilde{\mathbf{M}}^{-1} \left( \tilde{A}^{-2} \tilde{\zeta} \mathbf{e}'_1 \mathbf{e}_1^T - \tilde{\mathbf{A}} \right) \tilde{\phi}$



**Fig. 3** Normalized bias of the radius  $\hat{R}_1$ , i.e.,  $NB(\hat{R}_1) = \sigma^{-2}\mathbb{E}(\hat{R}_1)$ , and the normalized bias of  $\hat{\theta}$ , i.e.,  $NB(\hat{\theta}) = \sigma^{-2}\|\mathbb{E}(\Delta_2\hat{\theta})\|$ , for the four algebraic fits whenever  $n_1 = 50$  and  $n_2 = 100$ . The tested methods are the Type-A direct fit (blue), the Type-B direct fit (gold), the simple fit (green), and the Taubin fit (red). **a** Normalized bias of  $\hat{R}_1$ ,  $NB(\hat{R}_1)$ , whenever  $n_1 = 50$  and  $n_2 = 100$ . **b** Normalized bias of  $\hat{\theta}$ ,  $NB(\hat{\theta})$ , whenever  $n_1 = 50$  and  $n_2 = 100$  (Color figure online)



**Fig. 4** Normalized bias of the radius  $\hat{R}_1$ , i.e.,  $NB(\hat{R}_1) = \sigma^{-2}\mathbb{E}(\hat{R}_1)$ , and the normalized bias of  $\hat{\theta}$ , i.e.,  $NB(\hat{\theta}) = \sigma^{-2}\|\mathbb{E}(\Delta_2\hat{\theta})\|$ , for the four algebraic fits whenever  $n_1 = 100$  and  $n_2 = 50$ . The tested methods are the Type-A direct fit (blue), the Type-B direct fit (gold), the simple fit (green), and the Taubin fit (red). **a** Normalized bias of  $\hat{R}_1$ ,  $NB(\hat{R}_1)$ , whenever  $n_1 = 100$  and  $n_2 = 50$ . **b** Normalized bias of  $\hat{\theta}$ ,  $NB(\hat{\theta})$ , whenever  $n_1 = 100$  and  $n_2 = 50$ . (Color figure online)

of the concentric circles is  $(0, 0)$ , and the radii are  $\tilde{R}_1 = 1$  and  $\tilde{R}_2 = 2$ , see Fig. 3. Then, we repeated the process with  $n_1 = 100$  and  $n_2 = 50$  and the results are shown in Fig. 4. The  $x$ -axis represents the length of the arcs  $\omega$ , which varies from  $\pi/2$  to  $2\pi$ .

Figures 3 and 4 reveal the behaviors of the normalized essential biases (represented by the solid lines) and the normalized nonessential biases (represented by the dotted lines). The first subfigure in each figure represents the normalized bias of  $\hat{R}_1$ ,  $NB(\hat{R}_1)$ , and the second subfigure represents the total normalized bias of  $\hat{\theta}$ ,  $NB(\hat{\theta})$ . Our pack includes the Type-A direct fit (blue), the Type-B direct fit (gold), the simple fit (green), and the Taubin fit (red). The figures show how

the essential biases dominate the nonessential biases that are quite close to zero. This observation remains valid, even if  $n_i$ 's were chosen much smaller than our choice, provided that  $\omega$  is relatively large, say  $\omega = \pi$ .

As Figs. 3 and 4 show, the normalized essential biases of  $\hat{R}_1$  and  $\hat{\theta}$  for each of the Type-A fit and the Taubin fit are always 2 and 1, respectively. On the other hand, the two other fits have much more involved situations.

The simple fit has the same essential bias as the Taubin fit for  $\omega \geq 4.5$  radian, but its tendency of giving smaller radii appears for short circular arcs when  $\omega \leq 3.5$  radian. The bias in radius estimate becomes the smallest among all

other methods whenever  $\omega \in (3.5, 4.5)$ , but this does not mean that the simple fit returns the best estimator over this zone. The reason behind this is that the bias of center estimate becomes very large. One can verify this phenomenon by plotting the analytic expression of the magnitude of the normalized bias of the center or by investigating the behavior of the normalized bias of  $\hat{\theta}$  as shown in Figs. 3b and 4b. Accordingly, we conclude that the simple fit coincides with the Taubin fit whenever observations are sampled along large concentric circular arcs. This resemblance behavior disappears whenever  $\omega$  becomes smaller and satisfies  $\omega < 4.5$ . Meanwhile, the simple fit continues to behave better than the Type-A direct fit over the zone  $\omega \in (3.5, 4.5)$ . As  $\omega$  reduces, the simple fit increases the bias considerably and its performance becomes the worst.

On the other hand, the Type-B direct fit has three scenarios depending on the difference  $n_1 - n_2$ .

1. The simplest case appears whenever  $n_1 = n_2$ . In this case, the two direct fits coincide.
2. The most practical scenario that Fig. 3 shows the case of  $n_1 \ll n_2$ . The normalized essential biases of  $\hat{R}_1$  and  $\hat{\theta}$  obtained by the Type-B direct fit reach their horizontal asymptotes that equal 2.8 whenever  $\omega \geq 3$ , but they blow out as  $\omega$  becomes smaller. This makes the Type-B fit the second worst fit, after the simple fit, for short arc and the worst fit in case of long circular arcs.
3. Figure 4 demonstrates the third scenario that appears whenever  $n_1 \gg n_2$ . In this case, the horizontal asymptotes of the normalized essential bias of  $\hat{R}_1$  and  $\hat{\theta}$  for the Type-B direct fit equal to 1.8. Moreover, on the zone  $\omega \in (2, 2.6)$  the normalized bias of  $\hat{R}_1$  becomes the smallest among all other fits but the magnitude of the normalized bias of the center increases, which makes the normalized total bias of  $\hat{\theta}$  larger as shown in Fig. 4b. Lastly, for  $\omega \in (0, 2)$ , both of the center and the bias have heavy biases.

### 5.2 Hyper-Accurate Fit

A pertinent approach to the analysis of Al-Sharadqah and Chernov, this rigorous analysis allows us to propose a new algebraic method that its estimator has zero essential bias and it possesses translation and rotation invariance properties. Since the essential bias of the Type-A direct fit is two times the essential bias of the Taubin fit, one can design a new constraint matrix  $\mathbf{H}$  such that its noiseless version is  $\tilde{\mathbf{H}} = 2\tilde{\mathbf{V}} - \tilde{\mathbf{A}}$ . That is, if  $\tilde{\mathbf{H}}$  is substituted in the general formula of the bias in Eq. (55), then the new estimator, say  $\hat{\phi}_H$  will have zero essential bias. The constraint matrix  $\mathbf{H}$  can be defined as

$$\mathbf{H} = 2\mathbf{V} - \mathbf{A} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \begin{bmatrix} 8z_{ij} & 4x_{ij} & 4y_{ij} & 2\hat{\delta}_{1i} & 2\hat{\delta}_{2i} \\ 4x_{ij} & 1 & 0 & 0 & 0 \\ 4y_{ij} & 0 & 1 & 0 & 0 \\ 2\hat{\delta}_{1i} & 0 & 0 & 0 & 0 \\ 2\hat{\delta}_{2i} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The generalized eigenvalue problem  $\mathbf{M}\phi = \lambda\mathbf{H}\phi$  has 5 pairs of solutions  $(\lambda, \hat{\phi})$  that also give zero essential biases, but the solution must also minimize the sum of the squares of the algebraic distances  $\phi^T\mathbf{M}\phi$ , subject to a quadratic constraint in  $\phi$ . Since  $\mathbf{A}\tilde{\phi} = \tilde{\mathbf{V}}\tilde{\phi} - 2\tilde{\mathbf{A}}\sum_{i=1}^2\sum_{j=1}^{n_i}\tilde{z}_{ij}$ , then

$$\tilde{\phi}^T\tilde{\mathbf{V}}\tilde{\phi} = \tilde{\phi}^T\tilde{\mathbf{A}}\tilde{\phi} = \tilde{\zeta}$$

and  $\tilde{\phi}^T\tilde{\mathbf{H}}\tilde{\phi} = \tilde{\zeta}$ . The quantity  $\tilde{\zeta} = n_1\tilde{\zeta}_1 + n_2\tilde{\zeta}_2$  is positive because  $\tilde{\zeta}_i = \tilde{B}^2 + \tilde{C}^2 - 4\tilde{A}\tilde{D}_i > 0$  for each  $i$ . This means that the new estimator

$$\hat{\phi}_H = \operatorname{argmin}_{\phi \neq 0} \phi^T\mathbf{M}\phi = \operatorname{argmin}_{\|\phi\|=1} \phi^T\mathbf{M}\phi \tag{60}$$

must possess the property  $\phi^T\mathbf{H}\phi > 0$ . Therefore, one can set the constraint  $\phi^T\mathbf{H}\phi = 1$ .

The constraint matrix  $\mathbf{H}$  is a symmetric but not a positive definite. It has one zero eigenvalue, three positive eigenvalues, and one negative eigenvalue. Therefore, its solution is the unit generalized eigenvector of the pencil  $(\mathbf{M}, \mathbf{H})$  that associates with the smallest positive eigenvalue  $\lambda$ . However, the matrix  $\mathbf{H}$  is singular; thus, it is a numerically stable strategy to alternatively solve the problem by choosing the unit generalized eigenvector of the pencil  $(\mathbf{H}, \mathbf{M})$  that associates with the largest positive eigenvalue.

The solution of this problem can be found by some built-in functions in many software packages, such as MATLAB. Alternatively one can use the following trick. Since the matrix of the moments  $\mathbf{M}$  can be written in terms of the so-called design matrix  $\mathbf{Z}$ , i.e.,  $\mathbf{M} = \mathbf{Z}^T\mathbf{Z}$ , then we can compute the “economic size” SVD of  $\mathbf{Z}$  by  $\mathbf{Z} = \mathbf{U}_0\mathbf{D}_0\mathbf{V}_0^T$ . Now we define  $\phi' = \mathbf{V}_0^T\phi$  and choose the eigenvector  $\hat{\phi}'$  of the largest positive eigenvalue  $\eta$  that solves  $(\mathbf{D}_0^{-2}\mathbf{V}_0^T\mathbf{H}\mathbf{V}_0)\hat{\phi}' = \eta\hat{\phi}'$ , and finally replicate the solution back to  $\hat{\phi}_H = \mathbf{V}_0\hat{\phi}'$ .

We note that the new estimator is invariant under translations and rotations because the constraint matrix  $\mathbf{H}$  is a linear combination of two others,  $\mathbf{V}$  and  $\mathbf{A}$ , that satisfy the invariance requirements. Finally, we call this method hyper-accurate algebraic method (or Hyper method for short).

*Remark 1* In the case that the observation belongs to the more general Gaussian assumptions, such as  $\operatorname{cov}(\mathbf{n}_{ij}, \mathbf{n}_{kl}) =$

**Table 2** Normalized bias (NB) of the elements of  $\hat{\theta}$  and its normalized MSE (NMSE),  $n_1 = 50$  and  $n_2 = 100$  and  $\omega = 240^\circ$

$\omega = 240^\circ$	S	B	A	Taub	MLE	Hyp
NB( $\hat{a}$ )	-0.47	0.03	-0.01	-0.01	-0.00	-0.01
NB( $\hat{b}$ )	0.82	-0.04	0.02	-0.02	0.01	0.02
NB( $\hat{R}_1$ )	0.64	2.75	1.97	1.01	0.52	0.04
NB( $\hat{R}_2$ )	0.12	0.95	0.97	0.50	0.25	0.01
NB( $\hat{\theta}$ )	1.15	2.92	2.19	1.12	0.58	0.04
NMSE( $\hat{\theta}$ )	0.10	0.14	0.12	0.10	0.08	0.09

**Table 3** Normalized bias (NB) of the elements of  $\hat{\theta}$  and its normalized MSE (NMSE),  $n_1 = 50$  and  $n_2 = 100$  and  $\omega = 120^\circ$

$\omega = 120^\circ$	S	B	A	Tau	MLE	Hyp
NB( $\hat{a}$ )	8.67	-0.73	0.00	-0.02	-0.01	0.00
NB( $\hat{b}$ )	15.01	-1.26	0.01	-0.03	0.09	0.01
NB( $\hat{R}_1$ )	12.78	3.94	2.02	1.08	0.65	0.06
NB( $\hat{R}_2$ )	-13.60	2.15	1.00	0.54	0.36	0.03
NB( $\hat{\theta}$ )	25.47	4.72	2.25	1.21	0.75	0.07
NMSE( $\hat{\theta}$ )	4.35	1.04	0.94	0.90	0.74	0.88

$\sigma^2 \hat{\delta}_{ik} \hat{\delta}_{jl} \mathbf{Q}_{ij}$ , then the Hyper method still has the constraint matrix  $\mathbf{H} = 2\mathbf{V} - \mathbf{A}$ , where  $\mathbf{V} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \mathbf{V}_{ij}$  and  $\mathbf{V}_{ij}$  in Eq. (13) will be replaced by

$$\mathbf{V}_{ij} = \nabla_{s_{ij}} P_i \text{cov}(\mathbf{n}_{ij}, \mathbf{n}_{ij}) \nabla_{s_{ij}} P_i.$$

### 5.3 Numerical Experiments

Next we turn our attention to run some numerical experiments and validate our second-order biases. We shall adopt the same models discussed in Sect. 5.1, and we shall examine the performances of each of the MLE and the algebraic methods. We examine firstly long concentric arcs with  $\omega = 240^\circ$  whenever  $n_1 = 50$  and  $n_2 = 100$  (Table 2) and  $n_1 = 100$  and  $n_2 = 50$  (Table 4). Then, we repeat this process again for short circular arcs with  $\omega = 120^\circ$  as listed in Tables 3 and 5.

The true points were contaminated by zero-mean white Gaussian noise with standard deviation  $\sigma = .05$ . A total of  $N = 10^5$  ensemble trials were performed. Tables 2, 3 and 5 summarize the normalized values of the biases of each component of  $\hat{\theta} = (\hat{a}, \hat{b}, \hat{R}_1, \hat{R}_2)^T$ . It also exhibits the normalized total bias of  $\hat{\theta}$ , i.e., NB( $\hat{\theta}$ ), as well as the normalized mean squared error (NMSE) of  $\hat{\theta}$ , which were estimated by

**Table 4** Normalized bias (NB) of the elements of  $\hat{\theta}$  and its normalized MSE (NMSE),  $n_1 = 100$  and  $n_2 = 50$  and  $\omega = 240^\circ$

$\omega = 240^\circ$	S	B	A	Tau	MLE	Hyp
NB( $\hat{a}$ )	-0.54	-0.07	-0.01	-0.01	0.00	-0.01
NB( $\hat{b}$ )	0.94	0.12	0.03	0.02	-0.01	0.03
NB( $\hat{R}_1$ )	0.58	1.72	1.93	0.99	0.52	0.04
NB( $\hat{R}_2$ )	0.06	0.91	0.95	0.48	0.26	0.00
NB( $\hat{\theta}$ )	1.23	1.95	2.15	1.11	0.58	0.051
NMSE( $\hat{\theta}$ )	0.12	0.14	0.15	0.12	0.08	0.11

**Table 5** Normalized bias (NB) of the elements of  $\hat{\theta}$  and its normalized MSE (NMSE),  $n_1 = 100$  and  $n_2 = 50$  and  $\omega = 120^\circ$

$\omega = 120^\circ$	S	B	A	Tau	MLE	Hyp
NB( $\hat{a}$ )	9.80	1.62	0.07	0.04	-0.04	0.07
NB( $\hat{b}$ )	16.99	1.96	0.14	0.09	-0.07	0.14
NB( $\hat{R}_1$ )	-14.61	-0.20	1.90	1.00	0.62	-0.01
NB( $\hat{R}_2$ )	-15.36	-0.74	0.88	0.45	0.35	-0.07
NB( $\hat{\theta}$ )	28.89	2.66	2.10	1.10	0.72	0.17
NMSE( $\hat{\theta}$ )	5.48	1.10	1.06	1.09	0.75	1.08

$$\begin{aligned} \text{NMSE}(\hat{\theta}) &= \sqrt{\sum_{k=1}^N \frac{\|\hat{\theta}^{(k)} - \tilde{\theta}\|^2}{N\sigma^2}} \\ \text{NB}(\hat{\theta}) &= \left\| \frac{1}{\sigma^2 N} \sum_{k=1}^N (\hat{\theta}^{(k)} - \tilde{\theta}) \right\|, \end{aligned} \tag{61}$$

where  $\hat{\theta}^{(k)}$  is the estimate of  $\theta$  in the  $k$ th trial. The tables include all of the four algebraic fits and the MLE. Note that the latter is the only method that its first-order variance achieves the KCR lower bound.

First, we observe that the normalized bias of radius estimates ( $\hat{R}_1, \hat{R}_2$ ) for the MLE is close to the theoretical normalized essential biases  $(\frac{1}{2R_1}, \frac{1}{2R_2}) = (0.5, 0.25)$  as listed in Tables 2 and 4. The tables also show that this bias is half of the bias of the Taubin method, which is half of the bias from the Type-A fit. Since these algebraic methods have the same first-order covariance matrix, the MSE of the Taubin fit is smaller than the Type-A method. On the other hand, it is clear that the Hyper fit has the smallest bias among all experimented fits and its NMSE is smaller than all other algebraic fits. Note that for short-arc cases these conclusions remain valid even though their results are slightly different. These differences comes from the contributions of the nonessential biases ( $\approx 10\text{--}15\%$  of the essential biases).

Second, the following observations regarding the simple fit validate our theoretical analysis. Tables 2 and 4 show how well behaved the simple fit is for long circular arcs. Its biases

for  $\hat{R}_1$  and  $\hat{R}_2$  are the smallest compared with other estimators (MLE, Type-A, Type-B, Taubin). However, its center estimate in this case is the worst, and as such, this makes the simple fit have similar performance as the Taubin fit. On the other hand, if the observations are clustered along short circular arcs, then the simple method will return a much smaller concentric circles. This attenuation behavior for the radii is clearly listed in Tables 3 and 5.

Third, in the regard of the behavior of the normalized bias of Type-B fit, Tables 2, 3, 4 and 5 show an excellent agreement with our conclusions presented in Sect. 5.1. Tables 2 and 4 show that the normalized bias of  $\hat{R}_1$  for the Type-B fit approaches its theoretical limit 2.8 and 1.8 (respectively)  $\omega$  becomes large. Whenever  $n_2 = 100$  and  $n_1 = 50$ , Table 4 shows that the total normalized bias of Type-B method is smaller than the Type-A method, and as such, the NMSE of  $\hat{\theta}_B$  is smaller than the NMSE of  $\hat{\theta}_A$ . Note that this situation is reversed whenever  $n_1 = 50$  and  $n_2 = 100$  as listed in Table 2. In the latter case,  $NB(\hat{R}_1)$  for the Type-B method approaches its theoretical limit 2.8 as  $\omega$  increases, and its  $NB(\hat{\theta}_B)$  becomes larger than all other algebraic methods. Consequently, it becomes the worst in this case (recall that the simple fit behaves like the Taubin fit for large concentric circular arcs).

Another excellent agreement between the numerical experiments and the theoretical results occurs again in our last observation regarding the behavior of the Type-B fit in case of short concentric circular arcs (Tables 3 and 5). Here, the bias of the center estimate is relatively large for the Type-B if compared with the other methods (excluding the simple fit) but its radius depends on the difference  $n_1 - n_2$ . For  $n_1 = 50$  and  $n_2 = 100$ , its  $NB(\hat{R}_i)$  is also much larger than the others, and as a result, its has the highest bias among other fits (except the simple fit). This demonstrates why it falls behind the Type-A fit but it is still better than the simple fit in the MSE sense. Finally, for  $n_1 = 100$  and  $n_2 = 50$ ,  $NB(\hat{R}_i)$ 's are negative, but with small magnitudes if compared with the other fits (excluding the simple fit). This makes the Type-B method the second worst method after the simple method, when data are sampled along short circular arcs with  $n_1 \gg n_2$ .

### 6 Second-Order Analysis of Iterative Methods

In this section, we will evaluate the bias of the GRAF and the generalized GRAF. We will show why the GRAF has a smaller bias than the Renormalization method. We first start with the GRAF. To apply the perturbation analysis to GRAF  $\hat{\phi}_g$ , we express  $\hat{\phi}_g - \tilde{\phi} = \Delta_1 \hat{\phi}_g + \Delta_2 \hat{\phi}_g + \mathcal{O}_P(\sigma^3)$ , where  $\Delta_2 \hat{\phi}_g$  is a quadratic form of  $\delta_{ij}$ 's and  $\epsilon_{ij}$ 's and all other higher-order terms are represented by  $\mathcal{O}_P(\sigma^3)$ . Here  $\Delta_1 \hat{\phi}_g$

and  $\Delta_2 \hat{\phi}_g$  can be obtained by applying the Taylor expansion of  $\gamma_{ij}^{-1} = (\hat{\phi}_g^T \mathbf{V}_{ij} \hat{\phi}_g)^{-1}$  around its true value  $\tilde{\gamma}_{ij}^{-1}$ , where  $\tilde{\gamma}_{ij} = \tilde{\phi}^T \tilde{\mathbf{V}}_{ij} \tilde{\phi}$  can be simplified further—after some straightforward algebraic manipulations—to

$$\tilde{\gamma}_{ij} = 4\tilde{A} \left( \tilde{A}\tilde{z}_{ij} + \tilde{B}\tilde{x}_{ij} + \tilde{C}\tilde{y}_{ij} \right) + \tilde{B}^2 + \tilde{C}^2 = \tilde{\zeta}_i, \tag{62}$$

where  $\tilde{\zeta}_i$  is the true value of  $\zeta_i$  [cf. Eq. (5)]. Therefore, if we use the Taylor series expansion of  $\gamma_{ij}^{-1} = (\hat{\phi}_g^T \mathbf{V}_{ij} \hat{\phi}_g)^{-1}$  around  $\tilde{\gamma}_{ij}^{-1}$ , then we obtain

$$\gamma_{ij}^{-1} = \tilde{\zeta}_i^{-1} \left( 1 - \tilde{\zeta}_i^{-1} (\Delta_1 \gamma_{ij} + \Delta_2 \gamma_{ij}) + \tilde{\zeta}_i^{-2} (\Delta_1 \gamma_{ij})^2 \right) + \mathcal{O}_P(\sigma^3).$$

Accordingly,

$$\mathcal{M} = \tilde{\mathcal{M}} + \Delta_1^1 \mathcal{M} + \Delta_1^2 \mathcal{M} + \Delta_2^1 \mathcal{M} + \Delta_2^2 \mathcal{M} + \Delta_2^3 \mathcal{M} + \mathcal{O}_P(\sigma^3),$$

where  $\Delta_1^j \mathcal{M}$  and  $\Delta_2^k \mathcal{M}$ , for  $j = 1, 2$  and  $k = 1, 2, 3$  denote the various “parts” of  $\Delta_1 \mathcal{M}$  and  $\Delta_2 \mathcal{M}$ , respectively. Their formal definitions are listed below.

$$\begin{aligned} \Delta_1^1 \mathcal{M} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{1}{\tilde{\zeta}_i} \Delta_1 \mathbf{M}_{ij}, \\ \Delta_1^2 \mathcal{M} &= - \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{\Delta_1 \gamma_{ij}}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij}, \\ \Delta_2^1 \mathcal{M} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{1}{\tilde{\zeta}_i} \Delta_2 \mathbf{M}_{ij}, \\ \Delta_2^2 \mathcal{M} &= - \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{\Delta_1 \gamma_{ij}}{\tilde{\zeta}_i^2} \Delta_1 \mathbf{M}_{ij}, \\ \Delta_2^3 \mathcal{M} &= \sum_{i=1}^n \frac{-\Delta_2 \gamma_{ij} + \tilde{\zeta}_i^{-1} (\Delta_1 \gamma_{ij})^2}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij}. \end{aligned}$$

here,  $\Delta_2 \gamma_{ij}$  is irrelevant because  $\Delta_2^3 \mathcal{M} \tilde{\phi} = \mathbf{0}$  and its expression is omitted, while

$$\Delta_1 \gamma_{ij} = 2 \left( \Delta_1 \hat{\phi}_g^T \tilde{\mathbf{V}}_{ij} \tilde{\phi} \right) + \tilde{\phi}^T \Delta_1 \mathbf{V}_{ij} \tilde{\phi}, \tag{63}$$

where  $\Delta_1 \mathbf{V}_{ij} = \tilde{\mathbf{T}}_{ij} \delta_{ij} + \tilde{\mathbf{S}}_{ij} \epsilon_{ij}$ , with

$$\begin{aligned} \tilde{\mathbf{T}}_{ij} &= \begin{bmatrix} 8\tilde{x}_{ij} & 2 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \\ \tilde{\mathbf{S}}_{ij} &= \begin{bmatrix} 8\tilde{y}_{ij} & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \tag{64}$$

Now, perturbing Eq. (15) leads to

$$\begin{aligned} (\tilde{\mathcal{M}} + \Delta_1 \mathcal{M} + \dots) (\tilde{\boldsymbol{\phi}} + \Delta_1 \hat{\boldsymbol{\phi}}_g + \dots) &= \dots \\ (\tilde{\mathcal{L}} + \Delta_1 \mathcal{L} + \dots) (\tilde{\boldsymbol{\phi}} + \Delta_1 \hat{\boldsymbol{\phi}}_g + \dots). \end{aligned} \tag{65}$$

Since  $\hat{\boldsymbol{\phi}}_g^T \mathbf{M}_{ij} \hat{\boldsymbol{\phi}}_g = (\mathbf{z}_{ij}^T \hat{\boldsymbol{\phi}}_g)^2$  and  $\tilde{\mathbf{z}}_{ij}^T \tilde{\boldsymbol{\phi}} = 0$ , we have  $\hat{\boldsymbol{\phi}}_g^T \mathbf{M}_{ij} \hat{\boldsymbol{\phi}}_g \sim \mathcal{O}_P(\sigma^2)$ , and hence  $\tilde{\mathcal{L}} = \Delta_1 \mathcal{L} = \mathbf{0}$ , thus  $\mathcal{L} = \Delta_2 \mathcal{L} + \mathcal{O}_P(\sigma^3)$ , where

$$\begin{aligned} \Delta_2 \mathcal{L} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \left[ (\Delta_1 \mathbf{z}_{ij}^T \tilde{\boldsymbol{\phi}})^2 + (\tilde{\mathbf{z}}_{ij}^T \Delta_1 \hat{\boldsymbol{\phi}}_g)^2 \right. \\ &\quad \left. + 2 (\tilde{\mathbf{z}}_{ij}^T \Delta_1 \hat{\boldsymbol{\phi}}_g) (\Delta_1 \mathbf{z}_{ij}^T \tilde{\boldsymbol{\phi}}) \right] \tilde{\mathbf{V}}_{ij}. \end{aligned} \tag{66}$$

Thus, equating all  $\mathcal{O}(\sigma)$ -terms in Eq. (65) yields  $\Delta_1 \hat{\boldsymbol{\phi}}_g = -\tilde{\mathcal{M}}^{-1} \Delta_1 \mathcal{M} \tilde{\boldsymbol{\phi}}$ , and further,

$$\Delta_1 \hat{\boldsymbol{\phi}}_g = -\tilde{\mathcal{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} (\tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{z}_{ij}) \tilde{\mathbf{z}}_{ij}. \tag{67}$$

Note here that  $\|\hat{\boldsymbol{\phi}}_g\| = \|\tilde{\boldsymbol{\phi}}\| = 1$  implies  $\Delta_1 \hat{\boldsymbol{\phi}}_g$  is orthogonal to  $\tilde{\boldsymbol{\phi}}$ . Also, it is worth mentioning here that  $\tilde{\boldsymbol{\phi}}$  is in the null space of  $\Delta_1^2 \mathcal{M}$ .

To find the second-order error bias, we first equating all  $\mathcal{O}(\sigma^2)$ - terms and obtain

$$\tilde{\mathcal{M}} \Delta_2 \hat{\boldsymbol{\phi}}_g + \Delta_1 \mathcal{M} \Delta_1 \hat{\boldsymbol{\phi}}_g + \Delta_2 \mathcal{M} \tilde{\boldsymbol{\phi}} = \Delta_2 \mathcal{L} \tilde{\boldsymbol{\phi}}.$$

Since  $\Delta_1^2 \mathcal{M} \tilde{\boldsymbol{\phi}} = \Delta_2^3 \mathcal{M} \tilde{\boldsymbol{\phi}} = \mathbf{0}$ , we have

$$\begin{aligned} \Delta_2 \hat{\boldsymbol{\phi}}_g &= \tilde{\mathcal{M}}^{-1} (\Delta_2 \mathcal{L} \tilde{\boldsymbol{\phi}} - \Delta_1 \mathcal{M} \Delta_1 \hat{\boldsymbol{\phi}}_g - \Delta_2 \mathcal{M} \tilde{\boldsymbol{\phi}}_g) \\ &= \tilde{\mathcal{M}}^{-1} (\Delta_2 \mathcal{L} - \mathbf{R}^*) \tilde{\boldsymbol{\phi}}, \end{aligned} \tag{68}$$

where we denote for brevity

$$\mathbf{R}^* = \mathbf{R}_1^* - \mathbf{R}_2^* - \mathbf{R}_3^*, \tag{69}$$

and

$$\mathbf{R}_1^* = \Delta_1^2 \mathcal{M} + \Delta_2^2 \mathcal{M}, \quad \mathbf{R}_2^* = (\Delta_1 \mathcal{M} + \Delta_1^2 \mathcal{M}) \tilde{\mathcal{M}}^{-1} \Delta_1 \mathcal{M},$$

and  $\mathbf{R}_3^* = (\Delta_1 \mathcal{M} + \Delta_1^2 \mathcal{M}) \tilde{\mathcal{M}}^{-1} \Delta_1^2 \mathcal{M}$ . Note that  $\mathbb{E}(\mathbf{R}_3^*) \tilde{\boldsymbol{\phi}} = \mathbf{0}$ , so it will be discarded in our analysis. Accordingly, we present the following identities and we moved its derivation to ‘‘Appendix’’; i.e.,

$$\mathbb{E}(\mathbf{R}^*) \tilde{\boldsymbol{\phi}} = \sigma^2 (\tilde{\mathbf{G}}_1 + \tilde{\mathbf{G}}_2) \tilde{\boldsymbol{\phi}} = \sigma^2 (\tilde{\mathbf{G}}_1 + \tilde{\mathbf{G}}_2 + \tilde{\mathbf{G}}_2^T) \tilde{\boldsymbol{\phi}}, \tag{70}$$

where

$$\tilde{\mathbf{G}}_1 = \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \frac{1}{\tilde{\zeta}_i} - \frac{\tilde{\psi}_{ij}}{\tilde{\zeta}_i^2} \right) \tilde{\mathbf{V}}_{ij}, \tag{71}$$

and  $\tilde{\psi}_i = \tilde{\mathbf{z}}_{ij}^T \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{z}}_{ij}$ . Also,

$$\begin{aligned} \tilde{\mathbf{G}}_2 &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \frac{2}{\tilde{\zeta}_i} \tilde{\mathbf{z}}_{ij} \mathbf{e}_1^T - \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{V}}_{ij} \right. \\ &\quad \left. + \frac{1}{\tilde{\zeta}_i^3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{r}}_{ij} - \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{r}}_{ij} \right]. \end{aligned} \tag{72}$$

Here we have defined

$$\tilde{\mathbf{r}}_{ij} = (\tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{T}}_{ij} \tilde{\boldsymbol{\phi}}) \tilde{\mathbf{z}}_{ij} \tilde{\mathbf{a}}_{ij}^T + (\tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{S}}_{ij} \tilde{\boldsymbol{\phi}}) \tilde{\mathbf{z}}_{ij} \tilde{\mathbf{b}}_{ij}^T, \tag{73}$$

where  $\tilde{\mathbf{a}}_{ij}$  and  $\tilde{\mathbf{b}}_{ij}$  are the first and the second columns of  $\nabla \mathbf{z}_{ij}$ .

Furthermore, the interesting identity  $\tilde{\mathbf{r}}_{ij} \tilde{\boldsymbol{\phi}} = 4\tilde{A} \tilde{\zeta}_i \tilde{\mathbf{z}}_{ij}$  simplifies  $\tilde{\mathbf{G}}_2 \tilde{\boldsymbol{\phi}}$ ; i.e.,

$$\begin{aligned} \tilde{\mathbf{G}}_2 \tilde{\boldsymbol{\phi}} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \frac{-2\tilde{A}}{\tilde{\zeta}_i} \tilde{\mathbf{z}}_{ij} - \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} + \frac{4\tilde{A}}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{z}}_{ij} \right] \\ &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \left[ \frac{2\tilde{A}}{\tilde{\zeta}_i} \left( -1 + \frac{2\tilde{\psi}_{ij}}{\tilde{\zeta}_i} \right) \tilde{\mathbf{z}}_{ij} - \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \right]. \end{aligned} \tag{74}$$

To find the bias of the GRAF, then we need to compute  $\mathbb{E}(\Delta_2 \mathcal{L})$ , which was already computed by Kanatani [20, pp. 181–182] and Al-Sharadqah and Chernov [2] for conic fitting; thus, after simple algebraic manipulations, we obtain

$$\mathbb{E}(\Delta_2 \mathcal{L}) \tilde{\boldsymbol{\phi}} = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \frac{1}{\tilde{\zeta}_i} - \frac{\tilde{\psi}_{ij}}{\tilde{\zeta}_i^2} \right) \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} = \sigma^2 \tilde{\mathbf{G}}_1 \tilde{\boldsymbol{\phi}}.$$



Hence,  $\mathbb{E}(\Delta_2 \hat{\phi}_g) = -\sigma^2 \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{G}}_2 \tilde{\phi}$ . Thus

$$\begin{aligned} \mathbb{E}(\Delta_2 \hat{\phi}_g) &= \sigma^2 \tilde{\mathcal{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{2\tilde{A}}{\tilde{\zeta}_i} \left( 1 - \frac{2\tilde{\psi}_{ij}}{\tilde{\zeta}_i} \right) \tilde{\mathbf{z}}_{ij} \\ &\quad + \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{V}}_{ij} \tilde{\phi}. \end{aligned} \tag{75}$$

Since  $\sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \tilde{\mathbf{z}}_{ij}$  is the  $(3+i)$ th column of  $\tilde{\mathcal{M}}$  for each  $i = 1, 2$ , then the bias of GRAF is reduced to

$$\begin{aligned} \mathbb{E}(\Delta_2 \hat{\phi}_g) &= -\sigma^2 \tilde{\mathbf{I}}_{\tilde{\phi}} + \sigma^2 \tilde{\mathcal{M}}^{-1} \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{-4\tilde{A}\tilde{\psi}_{ij}}{\tilde{\zeta}_i^2} \tilde{\mathbf{z}}_{ij} \\ &\quad + \frac{1}{\tilde{\zeta}_i^2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-1} \tilde{\mathbf{V}}_{ij} \tilde{\phi}. \end{aligned} \tag{76}$$

For large  $n = n_1 + n_2$ , we have  $\tilde{\mathcal{M}} = \mathcal{O}(n)$ ; hence both of  $\tilde{\mathcal{M}}^{-1}$  and  $\tilde{\psi}_{ij}$  are of order of magnitude  $\mathcal{O}(1/n)$ . Therefore, the essential bias of the GRAF is  $-\sigma^2 \tilde{\mathbf{I}}_{\tilde{\phi}}$ , which is the negative of the essential bias of the Taubin fit as revealed by Table 1. Now, premultiplying Eq. (76) by  $\tilde{\mathbf{E}}$  expresses the bias of the GRAF in terms of its geometric parameters  $\theta$ . Most importantly, we notice that the center has only nonessential bias while its radii have both parts, from which the essential part is  $\sigma^2 \left( \frac{-1}{\tilde{R}_1}, \frac{-1}{\tilde{R}_1} \right)$ . This essential bias is the negative of the essential bias of the Taubin fit. Hence, the GRAF and the Taubin fit—in absolute value—have essential biases for the two radii twice as much as the MLE have. In other words, the GRAF tends to return concentric circles with smaller radii, while the MLE and the Taubin method have the opposite tendency.

### 6.1 Analysis of Generalized GRAF

Here, we analyze the bias of the generalized GRAF that solves the generalized eigenvalue problem in Eq. (17). This can be done by applying matrix perturbation to  $\mathcal{M}$ ,  $\mathcal{N}$ ,  $\hat{\phi}$ , and  $\lambda$  in Eq. (17) and using the same approach implemented in the second-order analysis of the algebraic fits. That is, perturbing

$$\begin{aligned} &(\tilde{\mathcal{M}} + \Delta_1 \mathcal{M} + \Delta_2 \mathcal{M} + \dots)(\tilde{\phi} + \Delta_1 \hat{\phi} + \Delta_2 \hat{\phi} + \dots) \\ &= (\tilde{\lambda} + \Delta_1 \lambda + \Delta_2 \lambda + \dots)(\tilde{\mathcal{N}} + \Delta_1 \mathcal{N} + \dots) \\ &\quad (\tilde{\phi} + \Delta_1 \hat{\phi} + \dots). \end{aligned} \tag{77}$$

Then, one can show that  $\tilde{\lambda} = \Delta_1 \lambda = 0$ . Therefore,  $\Delta_1 \hat{\phi}$  is the same as Eq. (67) and its covariance also equals to  $\sigma^2 \tilde{\mathcal{M}}^{-1}$  [cf. Eq. (29)].

To evaluate the bias, we first equate all  $\mathcal{O}_P(\sigma^2)$  in Eq. (46) and obtain

$$\tilde{\mathcal{M}} \Delta_2 \hat{\phi} = - \left( \Delta_2 \mathcal{M} \tilde{\phi} + \Delta_1 \mathcal{M} \Delta_1 \hat{\phi} \right) + \Delta_2 \lambda \tilde{\mathcal{N}} \tilde{\phi}. \tag{78}$$

From Eq. (67), we conclude  $\Delta_2 \mathcal{M} \tilde{\phi} + \Delta_1 \mathcal{M} \Delta_1 \hat{\phi}$

$$\begin{aligned} &= \left( \Delta_2^1 \mathcal{M} + \Delta_2^2 \mathcal{M} \right) \tilde{\phi} - \left( \Delta_1^1 \mathcal{M} + \Delta_1^2 \mathcal{M} \right) \tilde{\mathcal{M}}^{-1} \Delta_1^1 \mathcal{M} \tilde{\phi} \\ &= \mathbf{R}_1^* \tilde{\phi} - \mathbf{R}_2^* \tilde{\phi} = \mathbf{R}^* \tilde{\phi}. \end{aligned}$$

To find  $\Delta_2 \lambda$ , we premultiply Eq. (78) with  $\tilde{\phi}^T$  and obtain  $\tilde{\phi}^T \tilde{\mathcal{M}} = \mathbf{0}^T$ . Hence,  $\Delta_2 \lambda = \frac{\tilde{\phi}^T \mathbf{R}^* \tilde{\phi}}{\tilde{\phi}^T \tilde{\mathcal{N}} \tilde{\phi}}$ , and  $\Delta_2 \hat{\phi}$  can be written as

$$\begin{aligned} \Delta_2 \hat{\phi} &= -\tilde{\mathcal{M}}^{-1} \mathbf{R}^* \tilde{\phi} + \tilde{\mathcal{M}}^{-1} \frac{\tilde{\mathcal{N}} \tilde{\phi} \tilde{\mathbf{A}}^T}{\tilde{\phi}^T \tilde{\mathcal{N}} \tilde{\mathbf{A}}} \mathbf{R}^* \tilde{\phi} \\ &= \tilde{\mathcal{M}}^{-1} \mathbf{S}_{\tilde{\mathcal{N}}} \mathbb{E}(\mathbf{R}^*) \tilde{\phi}, \end{aligned} \tag{79}$$

where the symmetrization operator  $\mathbf{S}_{\bullet}$  is defined in Eq. (51). Now, premultiplying Eq. (70) by  $\tilde{\phi}^T$  and recalling that  $\tilde{\zeta}_i = \tilde{\phi}^T \tilde{\mathbf{V}}_{ij} \tilde{\phi}$  yield  $\tilde{\phi}^T \mathbb{E}(\mathbf{R}^*) \tilde{\phi} = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} (1 - \tilde{\zeta}_i^{-1} \tilde{\psi}_{ij})$ . Also, since  $\text{kernel}(\tilde{\mathcal{M}}) = \text{span}(\tilde{\mathbf{A}})$ , which is a one-dimensional subspace of  $\mathbb{R}^5$ , we have the identity

$$\begin{aligned} \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \tilde{\psi}_{ij} &= \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \text{tr}(\tilde{\mathcal{M}}^{-1} \tilde{\mathbf{M}}_{ij}) \\ &= \text{tr}(\tilde{\mathcal{M}}^{-1} \tilde{\mathcal{M}}) = 4. \end{aligned}$$

Thus,

$$\mathbb{E}(\Delta_2 \hat{\phi}) = -\sigma^2 \tilde{\mathcal{M}}^{-1} (\tilde{\mathbf{G}}_1 + \tilde{\mathbf{G}}_2) \tilde{\phi} + \sigma^2 \tilde{\mathcal{M}}^{-1} \frac{n-4}{\tilde{\phi}^T \tilde{\mathcal{N}} \tilde{\phi}} \tilde{\mathcal{N}} \tilde{\phi}.$$

The Renormalization scheme [13, 14] is a special case of the generalized GRAF  $\mathcal{N} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \gamma_{ij}^{-1} \mathbf{V}_{ij}$ . In this case,  $\tilde{\phi}^T \tilde{\mathcal{N}} \tilde{\phi} = n$ , hence

$$\mathbb{E}(\Delta_2 \hat{\phi}_R) = -\sigma^2 \tilde{\mathcal{M}}^{-1} (\tilde{\mathbf{G}}_1^* + \tilde{\mathbf{G}}_2) \tilde{\phi}, \tag{80}$$

where

$$\tilde{\mathbf{G}}_1^* = \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \frac{4}{n\tilde{\zeta}_i} - \frac{\tilde{\psi}_{ij}}{\tilde{\zeta}_i^2} \right) \tilde{\mathbf{V}}_{ij}. \tag{81}$$

Note here that  $\tilde{\mathcal{M}}^{-1} \tilde{\mathbf{G}}_1^* = \mathcal{O}(1/n)$  and  $\tilde{\mathcal{M}}^{-1} \tilde{\mathbf{G}}_1 = \mathcal{O}(1)$ . Thus, the Renormalization scheme suppresses  $\tilde{\mathbf{G}}_1$  to a matrix of a smaller order, while GRAF eliminates it completely. Finally, combining Eqs. (74) and (76) with (80) and (81)

**Table 6** Normalized bias (NB) of the elements of  $\hat{\theta}$  and its normalized MSE (NMSE),  $n_1 = 50$  and  $n_2 = 100$  and  $\omega = 180^\circ$

$\omega = 180^\circ$	$\hat{\theta}_{\text{adj}}$	$\hat{\theta}_m$	GRAF	RM	Taubin
NB( $\hat{a}$ )	-0.001	-0.001	-0.001	-0.001	-0.002
NB( $\hat{b}$ )	-0.018	-0.018	-0.052	-0.030	0.007
NB( $\hat{R}_1$ )	0.030	0.530	-0.903	-0.918	1.011
NB( $\hat{R}_2$ )	0.021	0.278	-0.434	-0.448	0.506
NB( $\hat{\theta}$ )	0.041	0.598	1.01	1.022	1.131
NMSE( $\hat{\theta}$ )	0.162	0.165	0.168	0.168	0.197

gives the final expression of the second-order bias of the Renormalization method.

To demonstrate our results on these two iterative methods and their connection with the MLE and the Taubin method, we experimented the same settings as before that  $\hat{\theta} = (0, 0, 1, 2)$  with  $n_1 = 50, n_2 = 100$ , and  $\omega = 180^\circ$ . Table 6 shows the NMSE and the normalized bias (NB) of  $10^5$  estimates of  $\hat{\theta}$  computed based on Eq. (61).

As listed in Table 6, NB( $\hat{a}$ ) and NB( $\hat{b}$ ) are negligible. Also, for each  $i = 1, 2$ , we clearly see that NB( $\hat{R}_i$ ) is close to  $\tilde{R}_i^{-1}$  for the Taubin fit and it is near  $-\tilde{R}_i^{-1}$  for each of the GRAF and the Renormalization methods. It is about  $(2\tilde{R}_i)^{-1}$  for the MLE. In total, we notice that the normalized biases of  $\hat{\theta}$  for the GRAF, the Renormalization fit, and the Taubin fit are equal, while the normalized bias of  $\hat{\theta}_m$  is approximately equal to  $\frac{1}{\sqrt{2}}$  of the normalized biased of the GRAF and the others. Moreover, the adjusted MLE has the smallest normalized MSE, as being a refinement for the MLE, whose performance is superior to all other methods.

### 7 Numerical Experiments

We shall examine our theoretical results through some numerical experiments using synthetic data and real image data.

#### 7.1 Experiments on Synthetic Data

In this section, we present results of experiments obtained from synthetic data. We considered the estimators obtained from the five algebraic fits: The simple fit (S), Taubin fit (T), Type-A direct fit (A), Type-B direct fit (B), and Hyper fit (H). Also, we included the iterative methods: GRAF based on the FNS scheme (F), the Renormalization (R) method, the MLE (M), and its refinement (AM), after supplying them with the solution of the Hyper fit as the initial guess. We define  $n_1 = 50$  and  $n_2 = 100$  equidistant points on the concentric circular arcs with center  $(0, 0)$  and radii  $\tilde{R}_1=1$  and  $\tilde{R}_2 = 2$ . Both arcs have the same initial angle 0 and terminal

angle  $\omega$ . Three cases were considered here with  $\omega = 240^\circ, \omega = 120^\circ$ , and  $\omega = 60^\circ$ . The true points were contaminated by zero-mean white Gaussian noise with standard deviation  $\sigma$ . For each  $\sigma, N = 10^4$  ensemble trials were performed and for each trial the estimates are computed. We plotted (a) the NMSE of  $\hat{\theta}$  and (b) the NB of  $\hat{\theta}$ , as the noise level  $\sigma$  increases from 0. In Figs. 5, 6 and 7a, the horizontal dashed line represents the normalized KCR, which is computed by taking the sum of the diagonal elements of  $\tilde{\mathbf{R}} = (\tilde{\mathbf{W}}^T \tilde{\mathbf{W}})^{-1}$ .

Figures 5, 6 and 7 exhibit the NMSE and the normalized bias NB for each of the fitting methods in the three cases. As a general observation, whenever  $\sigma$  is approaching 0, the NMSE of all iterative methods approach the  $\sigma^{-2}$ KCR lower bound while all non-iterative methods do not. This observation validates the theoretical results in [3]. Figures 5, 6 and 7 also reveal why the problem of fitting concentric circles differs from the single circle fitting where algebraic fits and the MLE are optimal in the sense of reaching the KCR lower bound. The Hyper method and the Adjusted MLE, followed by the MLE, have the smallest biases among other method. This supports, based on the MSE criterion, why the MLE and the adjusted MLE are the best estimators among all other fits, and why Hyper fit outperforms all algebraic methods.

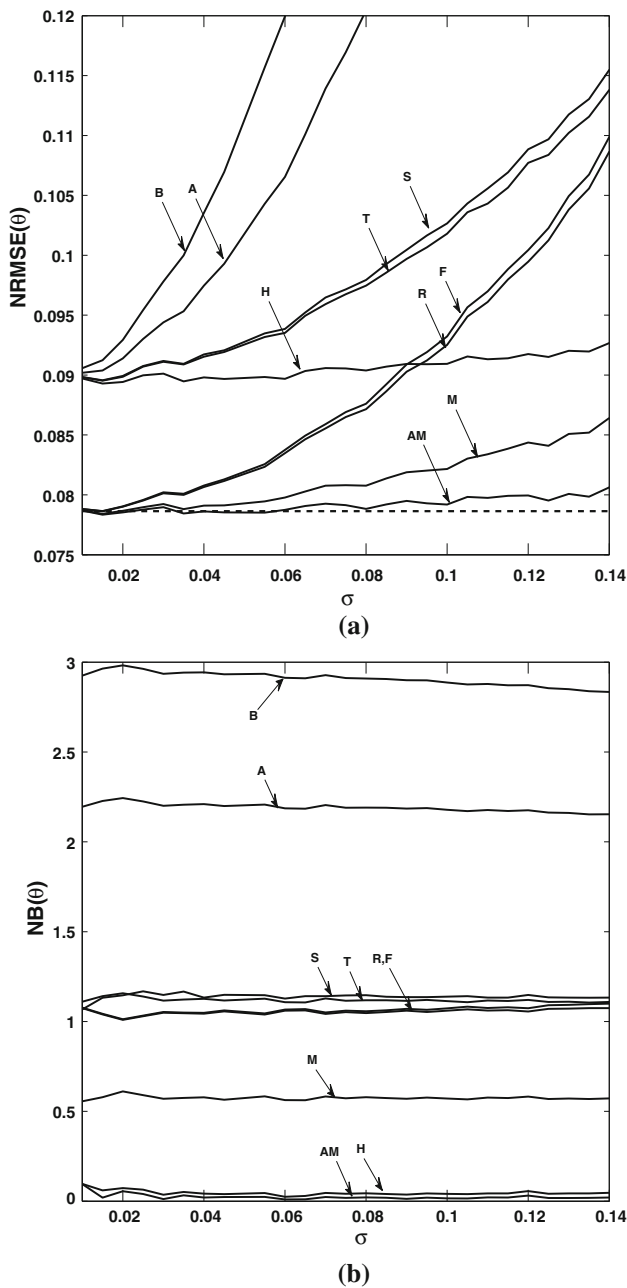
The behavior of the simple fit depends heavily on the length of the circular arcs. Its bias and MSE are close to the bias and the MSE of the Taubin method in the case of long circular arcs, but its bias blows out, and as such, its MSE does whenever the arc lengths reduce. For the  $\omega = 60^\circ$  case, its NB is even close to 500, while the second largest bias does not exceed 53. Therefore, we omitted its bias curve from the figure.

The significance of the essential bias can be clearly seen in the case of long circular arcs, in which the following relations NB( $\hat{\theta}_A$ ) > NB( $\hat{\theta}_T$ ) > NB( $\hat{\theta}_H$ ) hold, and as such, these relations also hold for their NMSEs. However, this significant difference between the Type-A and the Taubin method decreases whenever  $\omega$  is small. In Fig. 7b, the contribution of the nonessential biases becomes large, by which the Type-A and Taubin methods have nearly identical performance, but they still fall behind the Hyper method.

Finally, the numerical experiments show how the Renormalization method slightly outperforms the GRAF for long concentric arcs but it becomes unstable in the case of very short arcs (see Fig. 7a). In addition, its bias is the largest among the other fits (A, T, F, M, AM, H).

#### 7.2 Experiments on Real Data

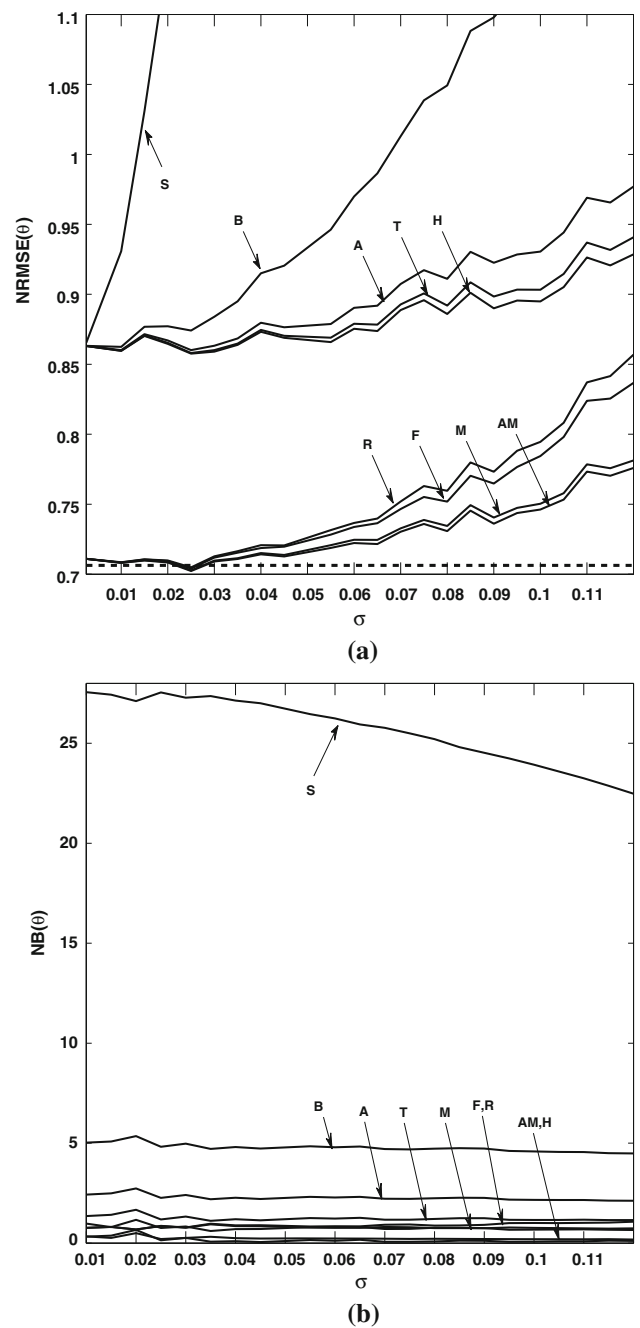
We now turn our attention by corroborating our analytical studies and findings with a set of practical data points obtained from an image of a CD/DVD. The image in binary form is shown in Fig. 1, where Canny's edge detection method [5] is applied to extract the data points in the inner and outer



**Fig. 5** NMSE( $\hat{\theta}$ ) in (a) and the NB( $\hat{\theta}$ ) in (b) of the parameters of the fitted concentric circles at additive observation noise with standard deviation  $\sigma$ . The settings are  $\hat{\theta} = (0, 0, 1, 2)^T$ ,  $n_1 = 50$ ,  $n_2 = 100$ ,  $\omega = 240^\circ$ , and  $N = 10^4$  independent trials. The tested fits are: the simple fit (S), the Taubin fit (T), the Type-A direct fit (A), the Type-B direct fit (B), the Hyper fit (H), the GRAF (F), the Renormalization (R) method, the MLE (M), and the adjusted MLE (AM). The horizontal dashed line represents the normalized KCR lower bound. **a** Normalized NMSE of  $\hat{\theta}$ , NMSE( $\hat{\theta}$ ), in the case  $\omega = 240^\circ$ . **b** Normalized bias of  $\hat{\theta}$ , NB( $\hat{\theta}$ ), in the case  $\omega = 240^\circ$

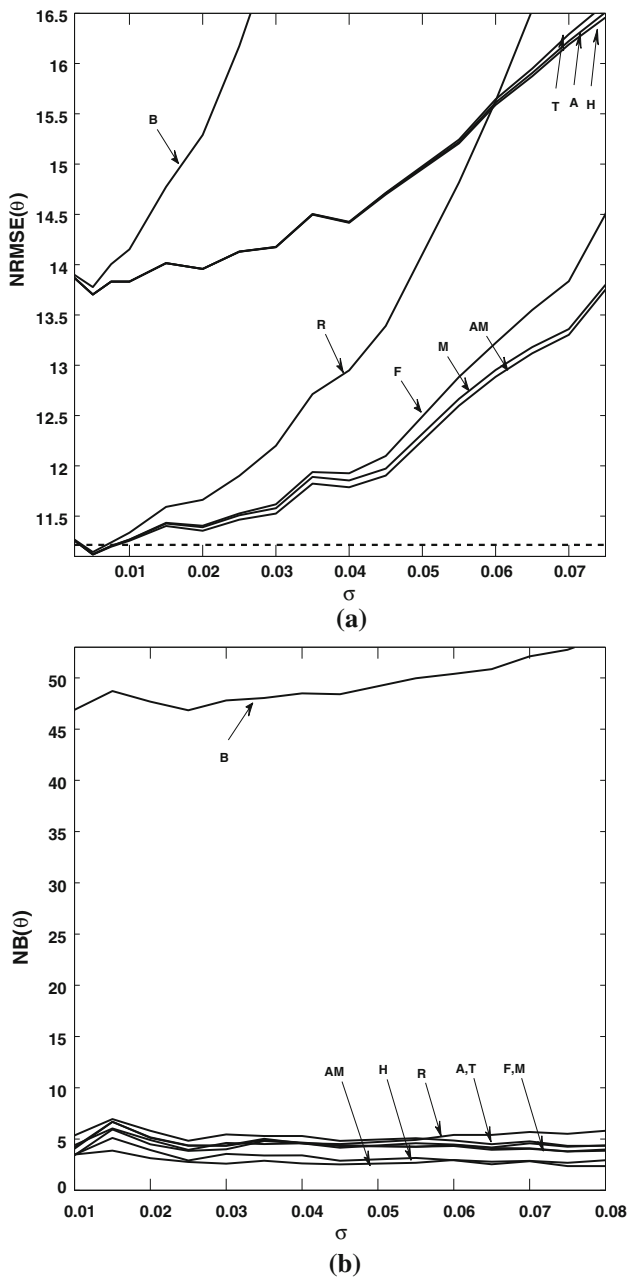
circles. We have a total of  $n_1 = 50$  and  $n_2 = 100$  points from them respectively.

In the experiment, we directly apply various algorithms to the data, where we added uncorrelated Gaussian noise of



**Fig. 6** The NMSE( $\hat{\theta}$ ) in (a) and the NB( $\hat{\theta}$ ) in (b) of the parameters of the fitted concentric circles at additive observation noise with standard deviation  $\sigma$ . The settings are  $\hat{\theta} = (0, 0, 1, 2)^T$ ,  $n_1 = 50$ ,  $n_2 = 100$ ,  $\omega = 120^\circ$  and  $N = 10^4$  independent trials. The tested fits are: the simple fit (S), the Taubin fit (T), the Type-A direct fit (A), the Type-B direct fit (B), the Hyper fit (H), the GRAF (F), the Renormalization (R) method, the MLE (M), and the adjusted MLE (AM). The horizontal dashed line represents the normalized KCR lower bound. **a** The normalized NMSE of  $\hat{\theta}$ , NMSE( $\hat{\theta}$ ), in the case  $\omega = 120^\circ$ . **b** The normalized bias of  $\hat{\theta}$ , NB( $\hat{\theta}$ ), in the case  $\omega = 120^\circ$

standard deviation  $\sigma = 0.05$  to the data points of the inner and outer circles. Only a segment of the data from 0 to 120 degrees is used, giving  $n_1 = 50$  and  $n_2 = 100$  points from



**Fig. 7** The  $NMSE(\hat{\theta})$  in (a) and the  $NB(\hat{\theta})$  in (b) of the parameters of the fitted concentric circles at additive observation noise with standard deviation  $\sigma$ . The settings are  $\theta = (0, 0, 1, 2)^T$ ,  $n_1 = 50$ ,  $n_2 = 100$ ,  $\omega = 60^\circ$  and  $N = 10^4$  independent trials. The tested fits are: the simple fit (S), the Taubin fit (T), the Type-A direct fit (A), the Type-B direct fit (B), the Hyper fit (H), the GRAF (F), the Renormalization (R) method, the MLE (M), and the adjusted MLE (AM). The horizontal dashed line represents the normalized KCR lower bound. The curves for the simple fit (S) were removed in this experiment. **a** The normalized NMSE of  $\hat{\theta}$ ,  $NMSE(\hat{\theta})$ , in the case  $\omega = 60^\circ$ . **b** The normalized bias of  $\hat{\theta}$ ,  $NB(\hat{\theta})$ , in the case  $\omega = 60^\circ$

the two concentric circles for estimation. We took a total of 100,000 ensemble runs. The averages of the estimation parameters are obtained for various fitting methods to examine the estimation bias, and they are listed in Table 7.

**Table 7** The averages of the estimation parameters that are obtained from various fitting methods

	$\bar{a}$	$\bar{b}$	$\bar{R}_1$	$\bar{R}_2$
S	35.64	-218.33	41.45	121.14
B	326.17	-257.40	74.24	153.74
A	327.24	-252.7	68.65	149.80
T	327.13	-253.16	66.98	149.27
H	327.23	-252.70	64.46	147.95
M	327.21	-252.92	65.70	148.58
AM	327.15	-252.78	64.66	148.05
R	327.48	-251.48	61.38	146.20
F	327.61	-251.26	68.69	150.95

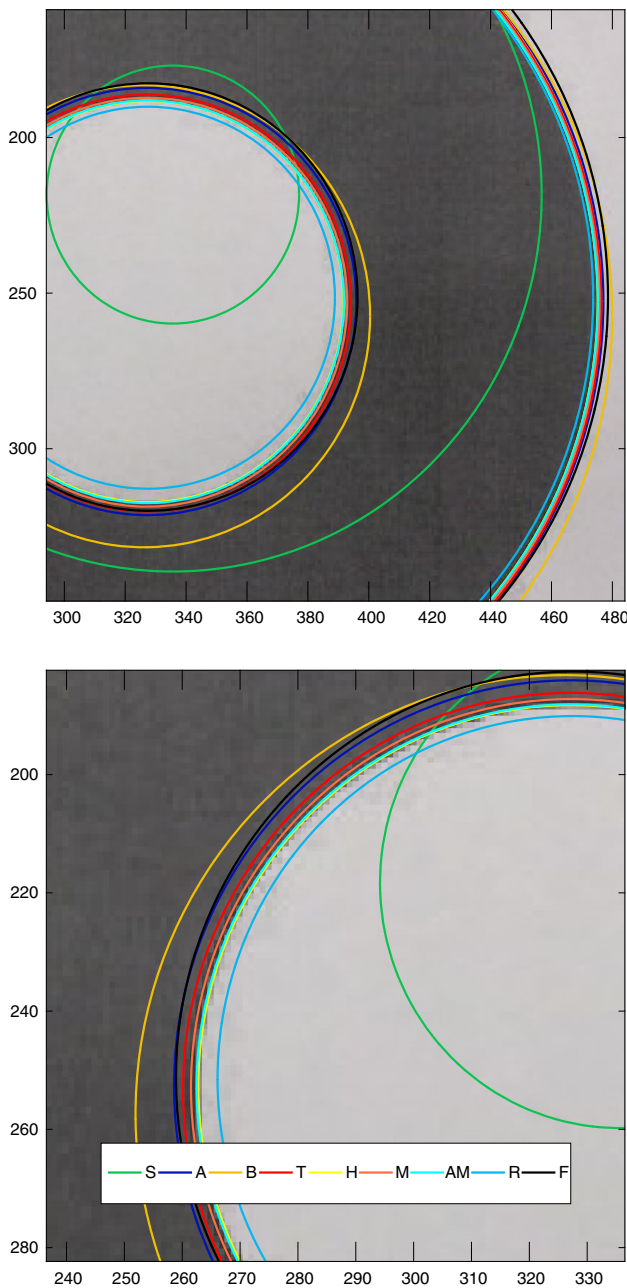
The corresponding fitted circles are shown in Fig. 8, where we have zoomed in to better observe the fitting performance. Both the Hyperfit (H) and the Adjusted MLE (AM) yield the best results as anticipated. Next best result is from the MLE (M), which tends to overestimate a little the parameters. The Renormalization method (R) under estimates the radii as indicated by our theoretical study. Furthermore, the Taubin (T) and Type-A (A), Type-B (B) fits overestimate the radii as the analytical investigation tells.

The experimental results from the real data confirm very well the theoretical studies and the improvements from the proposed H and AM methods. However, instead of underestimating the radii as the analytic results show, the GRAF (F), overestimates the radii. This problem might be more related to numerical issues and it requires further investigations.

### 8 Conclusions

In this paper, we applied the second-order error analysis to the problem of fitting concentric circles to a number of coordinate points. We derived the biases of the MLE, the algebraic fits, the GRAF, and the Renormalization method. We showed that the bias determines the performance of each method in its category (non-iterative and iterative). The smaller the bias, the smaller the MSE, and as such, the better the performance among each group. The most important part of the bias is the essential bias, and its order of magnitude  $\sigma^4$  dominates the value of the second-order bias as long as the arcs are not too short, such as  $\omega < 90^\circ$ .

Accordingly, we rigorously showed that the essential bias of the center estimate is zero for each of the MLE, the GRAF, the Renormalization, the Taubin, and the Type-A direct fit. However, their essential biases for the radii persist and they, after normalized by the noise power, depend only on the values of  $\tilde{R}_1$  and  $\tilde{R}_2$ . Indeed, we showed that the ratio of the essential biases of  $\hat{R}_i$  obtained from the MLE to the Taubin



**Fig. 8** The averages of estimators for  $N = 10^5$  independent trials that were generated after adding  $\sigma = .05$  to the real data. The settings are  $n_1 = 50$  and  $n_2 = 100$  around a short concentric arcs with  $\omega \approx 120^\circ$ . The tested fits are: the simple fit (S), the Taubin fit (T), the Type-A direct fit (A), the Type-B direct fit (B), the Hyper fit (H), the GRAF (F), the Renormalization (R) method, the MLE (M), and the adjusted MLE (AM)

fit and to the Type-A fit is 1:2:4. Also, the essential biases for the radii from the GRAF and the Renormalization method are equal, and they are identical to the negatives of those of the Taubin fit.

We also illustrated why the simple fit works well as the Taubin method in the case of long circular arcs but this excel-

lent performance diminishes as the length of arcs becomes smaller. Furthermore, we gained understanding of how the performance of the two types of direct fits were related to each other, and how the relationship of their performance depends on the sign and the magnitude of  $n_1 - n_2$ .

Additionally, our analysis allowed us to develop a new non-iterative algebraic fit, called the Hyper-accurate method that has zero essential bias. We addressed its implementation and demonstrated its superiority over all existing algebraic fits. The Hyper-accurate method was the most accurate algebraic fit among those we examined, and hence, it is considered as the best initial guess to feed the LM-based MLE and the other iterative methods. Finally, we modified the MLE to reduce its bias. This modified version of the MLE, called the Adjusted MLE, showed significant improvement over the MLE in our numerical experiments.

We validated our theoretical results with a series of numerical experiments using synthetic data. The agreement between them was excellent. We also applied the methods to a practical data set taken from an image of a CD/DVD. The results show an excellent match of the analytic results in practical applications.

The investigation of various methods discussed in this paper is based on the mean squared error of the geometric parameters  $(a, b, R_1, R_2)$  and their statistical properties. Accordingly, the numerical experiments are designed to maintain the consistency of the study. To have a full understanding of these methods, further work needs to be conducted using the geometric error [23–25] and confirm if the new estimators are indeed as good as they appear to be. The use of this measure is outside what we would like to cover in the current investigation. Depending on applications, the geometric distance measure could provide another dimension for evaluating algorithms. We plan to examine the performance of different algorithms in terms of the geometric distance measure and will report the results elsewhere when the study is complete.

**Acknowledgements** The authors would like to thank the reviewers and the editor for providing many suggestions that resulted in improving this paper.

### Appendix

**Derivation of Eq. (32)** First, we define

$$\alpha_{ij} = \delta_{ij} - \Delta_1 \hat{a} - \Delta_2 \hat{a}, \quad \beta_{ij} = \epsilon_{ij} - \Delta_1 \hat{b} - \Delta_2 \hat{b},$$

and

$$\zeta_{ij} = \frac{2}{\tilde{R}_i} \tilde{u}_{ij} \alpha_{ij} + \frac{2}{\tilde{R}_i} \tilde{v}_{ij} \beta_{ij} + \frac{\alpha_{ij}^2}{\tilde{R}_i^2} + \frac{\beta_{ij}^2}{\tilde{R}_i^2}.$$

Now, for each  $i = 1, 2$  and  $j = 1, \dots, n_i$ , expanding the distance  $d_{ij}$  to the second-order terms gives

$$d_{ij} = \tilde{R}_i \sqrt{1 + \zeta_{ij}} - \tilde{R}_i - \Delta_1 \hat{R}_i - \Delta_2 \hat{R}_i.$$

This expression can be simplified further using the approximation

$$\begin{aligned} \sqrt{1 + \zeta_{ij}} &\approx 1 + \frac{1}{2} \zeta_{ij} - \frac{1}{8} \zeta_{ij}^2 \\ &\approx 1 + \frac{2}{2\tilde{R}_i} \tilde{u}_{ij} \alpha_{ij} + \frac{2}{2\tilde{R}_i} \tilde{v}_{ij} \beta_{ij} + \frac{\alpha_{ij}^2}{2\tilde{R}_i^2} + \frac{\beta_{ij}^2}{2\tilde{R}_i^2} \\ &\quad - \frac{1}{8} \left( \frac{2}{\tilde{R}_i} \tilde{u}_{ij} \alpha_{ij} + \frac{2}{\tilde{R}_i} \tilde{v}_{ij} \beta_{ij} \right)^2. \end{aligned}$$

Keeping all terms of order  $\mathcal{O}_P(\sigma^2)$  and discarding the less significant terms in  $\sqrt{1 + \zeta_{ij}}$ , the above expression becomes

$$\begin{aligned} &\approx 1 + \frac{2}{2\tilde{R}_i} \tilde{u}_{ij} \alpha_{ij} + \frac{2}{2\tilde{R}_i} \tilde{v}_{ij} \beta_{ij} + \frac{(\delta_{ij} - \Delta_1 \hat{a})^2}{2\tilde{R}_i^2} \\ &\quad + \frac{(\epsilon_{ij} - \Delta_1 \hat{b})^2}{2\tilde{R}_i^2} \\ &\quad - \frac{1}{2\tilde{R}_i^2} \left( \tilde{u}_{ij} (\delta_{ij} - \Delta_1 \hat{a}) + \tilde{v}_{ij} (\epsilon_{ij} - \Delta_1 \hat{b}) \right)^2 \\ &\approx 1 + \frac{1}{\tilde{R}_i} \tilde{u}_{ij} \alpha_{ij} + \frac{1}{\tilde{R}_i} \tilde{v}_{ij} \beta_{ij} + \frac{(1 - \tilde{u}_{ij}^2) (\delta_{ij} - \Delta_1 \hat{a})^2}{2\tilde{R}_i^2} \\ &\quad + \frac{(1 - \tilde{v}_{ij}^2) (\epsilon_{ij} - \Delta_1 \hat{b})^2}{2\tilde{R}_i^2} \\ &\quad - \frac{2}{2\tilde{R}_i^2} \tilde{u}_{ij} \tilde{v}_{ij} (\delta_{ij} - \Delta_1 \hat{a}) (\epsilon_{ij} - \Delta_1 \hat{b}). \end{aligned}$$

Since  $1 - \tilde{u}_{ij}^2 = \tilde{v}_{ij}^2$ , we conclude

$$\begin{aligned} d_{ij} &= \tilde{R}_i \sqrt{1 + \zeta_{ij}} - \hat{R}_i \\ &\approx \tilde{u}_{ij} \alpha_{ij} + \tilde{v}_{ij} \beta_{ij} + \frac{\tilde{v}_{ij}^2 (\delta_{ij} - \Delta_1 \hat{a})^2}{2\tilde{R}_i} \\ &\quad + \frac{\tilde{u}_{ij}^2 (\epsilon_{ij} - \Delta_1 \hat{b})^2}{2\tilde{R}_i} \\ &\quad - \frac{2}{2\tilde{R}_i} \tilde{u}_{ij} \tilde{v}_{ij} (\delta_{ij} - \Delta_1 \hat{a}) (\epsilon_{ij} - \Delta_1 \hat{b}) \\ &\approx q_{ij} - \tilde{u}_{ij} \Delta_2 \hat{a} - \tilde{v}_{ij} \Delta_2 \hat{b} - \Delta_2 \hat{R}_i. \end{aligned}$$

*Proof of Lemma 1* The first three assertions will be proven if we show

$$\mathbb{E}(\rho_{ij}^2) = \sigma^2, \quad \mathbb{E}(\tau_{ij}^2) = \sigma^2 \tilde{h}_{ij}, \quad \mathbb{E}(\tau_{ij} \rho_{ij}) = 0.$$

Firstly, observe that for any  $j = 1, \dots, n_i$

$$\mathbb{E}(\rho_{ij}^2) = \tilde{\mathbf{t}}_{ij}^T \mathbb{E}(\check{\mathbf{n}}_{ij} \check{\mathbf{n}}_{ij}^T) \tilde{\mathbf{t}}_{ij} = \hat{\delta}_{ij} \sigma^2 \tilde{\mathbf{t}}_{ij}^T \tilde{\mathbf{t}}_{ij} = \sigma^2,$$

which comes from the identity  $\|\tilde{\mathbf{t}}_{ij}\|_2^2 = \tilde{u}_{ij}^2 + \tilde{v}_{ij}^2 = 1$ . Next, we compute  $\mathbb{E}(\tau_{ij}^2)$ . Since  $\text{cov}(\Delta_1 \hat{\boldsymbol{\theta}}_m) = \sigma^2 \tilde{\mathbf{R}}$ , then, for any  $j = 1, \dots, n_i$ , one has

$$\mathbb{E}(\tau_{ij}^2) = \tilde{\mathbf{t}}_{ij}^T \mathbb{E}(\Delta_1 \hat{\boldsymbol{\theta}}_m \Delta_1 \hat{\boldsymbol{\theta}}_m^T) \tilde{\mathbf{t}}_{ij} = \sigma^2 \tilde{h}_{ij}.$$

Finally, we compute  $\mathbb{E}(\tau_{ij} \rho_{ij})$ . Recall the definition of  $\Delta_1 \hat{\boldsymbol{\theta}}_m = \tilde{\mathbf{K}} \mathbf{f}_1 = \tilde{\mathbf{R}} \tilde{\mathbf{W}}^T \mathbf{f}_1$ , which can be rewritten as  $\tilde{\mathbf{R}} \mathbf{g}$  with

$$\mathbf{g} = \left[ \sum_{i=1}^2 \sum_{t=1}^{n_i} \tilde{u}_{it} f_{1it}, \sum_{i=1}^2 \sum_{t=1}^{n_i} \tilde{v}_{it} f_{1it}, \sum_{n=1}^{n_1} f_{11n}, \sum_{t=1}^{n_2} f_{12t} \right]^T,$$

then for each  $t = 1, \dots, n_i$ ,  $\mathbb{E}(f_{1it} \rho_{ij}) = \mathbb{E}(f_{1it} (\check{\mathbf{n}}_{ij}^T \tilde{\mathbf{t}}_{ij})) = 0$  for all  $j = 1, \dots, n_i$ . Thus,  $\mathbb{E}(\tau_{ij} \rho_{ij}) = 0$ .

Next, we compute the expectations of the outer products of  $\mathbf{a}_i$ ,  $\mathbf{b}_i$ , and  $\mathbf{c}_i$  starting with  $\mathbb{E}(\mathbf{a}_i \mathbf{a}_i^T)$ . Since

$$\mathbb{E}(a_{ij} a_{kl}) = \mathbb{E}(\rho_{ij}^2 \rho_{kl}^2) = \mathbb{E} \left( \left( \tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij} \right)^2 \left( \check{\mathbf{n}}_{kl}^T \tilde{\mathbf{t}}_{kl} \right)^2 \right).$$

Using the Isserlis' Theorem [11] gives

$$\begin{aligned} \mathbb{E}(a_{ij} a_{kl}) &= \mathbb{E} \left( \left[ \tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij} \right]^2 \right) \mathbb{E} \left( \left[ \check{\mathbf{n}}_{kl}^T \tilde{\mathbf{t}}_{kl} \right]^2 \right) \\ &\quad + 2 \mathbb{E} \left( \left( \tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij} \right) \left( \check{\mathbf{n}}_{kl}^T \tilde{\mathbf{t}}_{kl} \right) \right). \end{aligned}$$

But it is easy to show that  $\mathbb{E}((\tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij})^2) = \sigma^2$  and

$$\mathbb{E} \left( \left( \tilde{\mathbf{t}}_{ij}^T \check{\mathbf{n}}_{ij} \right) \left( \check{\mathbf{n}}_{kl}^T \tilde{\mathbf{t}}_{kl} \right) \right) = \hat{\delta}_{ik} \hat{\delta}_{jl} \sigma^2.$$

Therefore,

$$\mathbb{E}(a_{ij} a_{kl}) = \sigma^4 + 2 \hat{\delta}_{ik} \hat{\delta}_{jl} \sigma^4.$$

Next, we compute  $\mathbb{E}(b_{ij} b_{kl}) = 4 \mathbb{E}(\tau_{ij} \rho_{ij} \tau_{kl} \rho_{kl})$ . The pair  $(\tau_{ij}, \rho_{kl})$  are uncorrelated for any choice of  $i$  and  $k$ . Besides,

$$\mathbb{E}(\rho_{ij} \rho_{kl}) = \hat{\delta}_{ik} \hat{\delta}_{jl} \sigma^2, \quad \text{and} \quad \mathbb{E}(\tau_{ij} \tau_{kl}) = \sigma^2 \tilde{\mathbf{t}}_{ij}^T \tilde{\mathbf{R}} \tilde{\mathbf{t}}_{kl} = \sigma^2 \tilde{h}_{ijkl}.$$

Therefore,

$$\mathbb{E}(b_{ij} b_{kl}) = 4 \mathbb{E}(\tau_{ij} \tau_{kl}) \mathbb{E}(\rho_{ij} \rho_{kl}) = 4 \sigma^4 \tilde{h}_{ijkl} \hat{\delta}_{ik} \hat{\delta}_{jl}.$$

Next we evaluate  $\mathbb{E}(c_{ij} c_{kl}) = \mathbb{E}(\tau_{ij}^2 \tau_{kl}^2)$ . If  $\tilde{h}_{ijij}$  is expressed by  $\tilde{h}_{ij}$ , then using the general formulas of the expected value of the product of two quadratic forms of random variables [11], we obtain

$$\begin{aligned} \mathbb{E}(c_{ij}c_{kl}) &= \mathbb{E}(\tau_{ij}^2) \mathbb{E}(\tau_{kl}^2) + 2[\mathbb{E}(\tau_{ij}\tau_{kl})]^2 \\ &= \sigma^4 (\tilde{h}_{ij}\tilde{h}_{kl} + 2\tilde{h}_{ijkl}). \end{aligned}$$

In the same analog, we find

$$\mathbb{E}(a_{ij}c_{kl}) = \mathbb{E}(\tau_{ij}^2\rho_{kl}^2) = \mathbb{E}(\tau_{ij}^2) \mathbb{E}(\rho_{kl}^2) = \sigma^4\tilde{h}_{ij},$$

because  $\tau_{ij}$  and  $\rho_{kl}$  are independent random variables. Finally, following the same reason, we observe that  $\mathbb{E}(a_{ij}b_{kl}) = -2\mathbb{E}(\rho_{ij}^2\tau_{kl}\rho_{kl}) = 0$  and also  $\mathbb{E}(b_{ij}c_{kl}) = 0$ . From these results, we find the desired identities. This completes the proof of the lemma.  $\square$

**Derivation of Eqs. (41)–(43)** The MSE of  $\Delta_1\hat{\theta}_m$  is equal to  $\sigma^2\tilde{\mathbf{R}}$ . We shall next evaluate  $\text{MSE}(\Delta_2\hat{\theta}_m)$ . Let us define  $\check{\mathbf{a}} = (\check{\mathbf{a}}_1^T, \check{\mathbf{a}}_2^T)^T$  and other variables in the same manner. Then

$$\begin{aligned} \mathbb{E}(\check{\mathbf{a}}\check{\mathbf{a}}^T) &= \begin{bmatrix} \mathbb{E}(\check{\mathbf{a}}_1\check{\mathbf{a}}_1^T) & \mathbb{E}(\check{\mathbf{a}}_1\check{\mathbf{a}}_2^T) \\ \mathbb{E}(\check{\mathbf{a}}_2\check{\mathbf{a}}_1^T) & \mathbb{E}(\check{\mathbf{a}}_2\check{\mathbf{a}}_2^T) \end{bmatrix} \\ &= \sigma^4 \begin{bmatrix} \check{\mathbf{I}}_{n_1}\check{\mathbf{I}}_{n_1}^T & \check{\mathbf{I}}_{n_1}\check{\mathbf{I}}_{n_2}^T \\ \check{\mathbf{I}}_{n_2}\check{\mathbf{I}}_{n_1}^T & \check{\mathbf{I}}_{n_2}\check{\mathbf{I}}_{n_2}^T \end{bmatrix} + 2\sigma^4\check{\mathbf{I}}_n, \end{aligned}$$

where

$$\check{\mathbf{I}}_n = \begin{bmatrix} \frac{1}{4R_1^2}\mathbf{I}_{n_1} & \mathbf{0}_{n_1 \times n_2} \\ \mathbf{0}_{n_2 \times n_1} & \frac{1}{4R_2^2}\mathbf{I}_{n_2} \end{bmatrix} = \begin{bmatrix} \check{\mathbf{I}}_{n_1} & \mathbf{0}_{n_1 \times n_2} \\ \mathbf{0}_{n_2 \times n_1} & \check{\mathbf{I}}_{n_2} \end{bmatrix}.$$

Premultiplying and then postmultiplying  $\mathbb{E}(\check{\mathbf{a}}\check{\mathbf{a}}^T)$  by  $\tilde{\mathbf{K}}$  and  $\tilde{\mathbf{K}}^T$ , respectively, lead to

$$\tilde{\mathbf{K}}\mathbb{E}(\check{\mathbf{a}}\check{\mathbf{a}}^T)\tilde{\mathbf{K}}^T = \sigma^4\tilde{\mathbf{b}}_1\tilde{\mathbf{b}}_1^T + 2\sigma^4\tilde{\mathbf{K}}\check{\mathbf{I}}_n\tilde{\mathbf{K}}^T. \tag{82}$$

Now we compute  $\mathbb{E}(\check{\mathbf{a}}\check{\mathbf{c}}^T)$ . It is easy to show that

$$\mathbb{E}(\check{\mathbf{a}}\check{\mathbf{c}}^T) = \begin{bmatrix} \mathbb{E}(\check{\mathbf{a}}_1\check{\mathbf{c}}_1^T) & \mathbb{E}(\check{\mathbf{a}}_1\check{\mathbf{c}}_2^T) \\ \mathbb{E}(\check{\mathbf{a}}_2\check{\mathbf{c}}_1^T) & \mathbb{E}(\check{\mathbf{a}}_2\check{\mathbf{c}}_2^T) \end{bmatrix} = \sigma^4 \begin{bmatrix} \check{\mathbf{h}}_1\check{\mathbf{I}}_{n_1}^T & \check{\mathbf{h}}_1\check{\mathbf{I}}_{n_2}^T \\ \check{\mathbf{h}}_2\check{\mathbf{I}}_{n_1}^T & \check{\mathbf{h}}_2\check{\mathbf{I}}_{n_2}^T \end{bmatrix},$$

and as such, if it is premultiplied and then postmultiplied by  $\tilde{\mathbf{K}}$  and  $\tilde{\mathbf{K}}^T$ , respectively, we obtain

$$\begin{aligned} 2\mathcal{S} \left[ \tilde{\mathbf{K}}\mathbb{E}(\check{\mathbf{a}}\check{\mathbf{c}}^T)\tilde{\mathbf{K}}^T \right] &= 2\mathcal{S} \left[ \sigma^4\tilde{\mathbf{K}}\check{\mathbf{h}}(\check{\mathbf{e}}_3 + \check{\mathbf{e}}_4)^T \right] \\ &= 2\sigma^4\mathcal{S} \left[ \tilde{\mathbf{b}}_2\tilde{\mathbf{b}}_2^T \right]. \end{aligned}$$

Finally, the less important but the most complicated expressions in the MSE of  $\Delta_2\hat{\theta}_m$  come from  $\mathbb{E}(\tilde{\mathbf{b}}\tilde{\mathbf{b}}^T)$  and  $\mathbb{E}(\tilde{\mathbf{c}}\tilde{\mathbf{c}}^T)$ . After lengthy but direct calculations, we have

$$\mathbb{E} \left[ \left( (\check{\mathbf{b}} + \check{\mathbf{c}})(\check{\mathbf{b}} + \check{\mathbf{c}}) \right)^T \right] = \sigma^4 \left( \check{\mathbf{h}}\check{\mathbf{h}}^T + 2\check{\mathbf{H}} + 4\check{\mathbf{D}}_{\check{\mathbf{h}}_1, \check{\mathbf{h}}_2} \right),$$

where

$$\check{\mathbf{H}} = \begin{bmatrix} \check{\mathbf{H}}_{11} & \check{\mathbf{H}}_{12} \\ \check{\mathbf{H}}_{21} & \check{\mathbf{H}}_{22} \end{bmatrix},$$

and  $\check{\mathbf{D}}_{\check{\mathbf{h}}_1, \check{\mathbf{h}}_2} = \text{diag}(\check{\mathbf{h}}_1, \check{\mathbf{h}}_2)$ . Premultiplying and postmultiplying this expression by  $\tilde{\mathbf{K}}$  and  $\tilde{\mathbf{K}}^T$  give

$$\begin{aligned} \tilde{\mathbf{K}}\mathbb{E} \left[ \left( (\check{\mathbf{b}} + \check{\mathbf{c}})(\check{\mathbf{b}} + \check{\mathbf{c}}) \right)^T \right] \tilde{\mathbf{K}}^T \\ = \sigma^4\tilde{\mathbf{b}}_2\tilde{\mathbf{b}}_2^T + 2\sigma^4\tilde{\mathbf{K}} \left( \check{\mathbf{H}} + 2\check{\mathbf{D}}_{\check{\mathbf{h}}_1, \check{\mathbf{h}}_2} \right) \tilde{\mathbf{K}}^T. \end{aligned}$$

Finally, combining all expressions together gives us the MSE of  $\Delta_2\hat{\theta}_m$ .

*Proof of Eq. (70)* Let  $\mathbf{A}_{ij} = (\tilde{\phi}^T\tilde{\mathbf{V}}_{ij}\tilde{\mathcal{M}}^{-1}\Delta_1^1\mathcal{M}\tilde{\phi})\Delta_1^1\mathcal{M}$ . Then

$$\begin{aligned} \mathbb{E}(\mathbf{A}_{ij}) &= \sum_{k=1}^2 \sum_{l=1}^{n_k} \frac{2}{\zeta_k^2} \mathbb{E} \left[ \left( \tilde{\phi}^T\tilde{\mathbf{V}}_{ij}\tilde{\mathcal{M}}^{-1}(\Delta_1\mathbf{z}_{kl}\tilde{\phi})\tilde{\mathbf{z}}_{kl} \right) \mathcal{S} \left[ \Delta_1\mathbf{z}_{kl}\tilde{\mathbf{z}}_{kl}^T \right] \right] \\ &= 2\sigma^2 \sum_{k=1}^2 \sum_{l=1}^{n_k} \frac{1}{\zeta_k^2} \left( \tilde{\phi}^T\tilde{\mathbf{V}}_{ij}\tilde{\mathcal{M}}^{-1}\tilde{\mathbf{z}}_{kl} \right) \mathcal{S} \left[ \tilde{\mathbf{V}}_{kl}\tilde{\phi}\tilde{\mathbf{z}}_{kl}^T \right]. \end{aligned} \tag{83}$$

Here again  $\mathcal{S}(\bullet) = (\bullet + \bullet^T)/2$  is the symmetrization operator. Recall that  $\gamma_{ij} = \hat{\phi}^T\mathbf{V}_{ij}\hat{\phi}$  and  $\Delta_1\gamma_{ij}$  is given in Eq. (63). We shall evaluate Eq. (69) term by term. First, consider

$$\begin{aligned} \mathbf{I}_1 &:= \mathbb{E} \left( \Delta_1^1\mathcal{M}\tilde{\mathcal{M}}^{-1}\Delta_1^1\mathcal{M} \right) \\ &= 4 \sum_{i,k=1}^2 \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \zeta_i^{-1}\zeta_k^{-1} \mathbb{E} \left( \mathcal{S} \left[ \tilde{\mathbf{z}}_{ij}\Delta_1\mathbf{z}_{ij}^T \right] \right) \tilde{\mathcal{M}}^{-1} \\ &\quad \mathcal{S} \left[ \Delta_1\mathbf{z}_{kl}\tilde{\mathbf{z}}_{kl}^T \right] \\ &= 4 \sum_{i=1}^2 \sum_{j=1}^{n_i} \zeta_i^{-2} \times \mathbb{E} \left( \mathcal{S} \left[ \tilde{\mathbf{z}}_{ij}\Delta_1\mathbf{z}_{ij}^T \right] \right) \tilde{\mathcal{M}}^{-1} \\ &\quad \left( \mathcal{S} \left[ \tilde{\mathbf{z}}_{ij}\Delta_1\mathbf{z}_{ij}^T \right] \right). \end{aligned}$$

By direct inspection,  $\mathbb{E}(\Delta_1\mathbf{z}_{ij}^T\tilde{\mathcal{M}}^{-1}\Delta_1\mathbf{z}_{ij}) = \sigma^2\text{tr}(\tilde{\mathcal{M}}^{-1}\tilde{\mathbf{V}}_{ij})$ , hence if we using the definition of  $\tilde{\psi}_{ij} = \tilde{\mathbf{z}}_{ij}^T\tilde{\mathcal{M}}^{-1}\tilde{\mathbf{z}}_{ij}$ , then

$$\begin{aligned} \mathbf{I}_1 &= \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \zeta_i^{-2} \left[ 2\mathcal{S} \left[ \tilde{\mathbf{V}}_{ij}\tilde{\mathcal{M}}^{-1}\tilde{\mathbf{M}}_{ij} \right] \right. \\ &\quad \left. + \text{tr} \left( \tilde{\mathcal{M}}^{-1}\tilde{\mathbf{V}}_{ij} \right) \tilde{\mathbf{M}}_{ij} + \tilde{\psi}_{ij}\tilde{\mathbf{V}}_{ij} \right]. \end{aligned} \tag{84}$$

We shall make use of an auxiliary formula below that follows from Eq. (67):

$$\mathbb{E} \left( \Delta_1 \mathbf{A} \Delta_1 \mathbf{z}_{ij}^T \right) = -\sigma^2 \tilde{\zeta}_i^{-1} \tilde{\mathcal{M}}^{-\tilde{\mathbf{z}}_{ij}} \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tag{85}$$

and hence  $\mathbb{E}(\Delta_1 \mathbf{z}_{ij} \Delta_1 \hat{\boldsymbol{\phi}}^T) = -\sigma^2 \tilde{\zeta}_i^{-1} \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \tilde{\mathbf{z}}_{ij}^T \tilde{\mathcal{M}}^{-}$ . Next we compute

$$\begin{aligned} \mathbf{I}_2 &:= \mathbb{E} \left( \Delta_1^2 \mathcal{M} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right) \\ &= -\sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E} \left( \Delta_1 \gamma_{ij} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right). \end{aligned} \tag{86}$$

Using  $\Delta_1^1 \mathcal{M} = \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \Delta_1 \mathbf{M}_{ij}$ , and then Eqs. (63) and (85) give

$$\begin{aligned} \mathbf{I}_2 &= -\sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E} \left[ \left( 2 \left( \Delta_1 \hat{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \right) \right. \right. \\ &\quad \left. \left. + \tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right] \\ &= -\sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E} \left( 2 \left( \Delta_1 \hat{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right) \\ &\quad - \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E} \left( \left( \tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right) \\ &= -2 \sum_{i,k=1}^2 \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \tilde{\zeta}_i^{-2} \tilde{\zeta}_k^{-1} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \mathbb{E} \\ &\quad \left( \Delta_1 \mathbf{M}_{kl} \Delta_1 \hat{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \right) \\ &\quad - \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E} \left( \left( \tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \Delta_1^1 \mathcal{M} \right) \\ &= 4\sigma^2 \sum_{i,k=1}^2 \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \tilde{\zeta}_i^{-2} \tilde{\zeta}_k^{-2} \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{z}}_{kl} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \\ &\quad \mathcal{S} \left[ \tilde{\mathbf{V}}_{kl} \tilde{\boldsymbol{\phi}} \tilde{\mathbf{z}}_{kl}^T \right] - \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \mathbb{E} \\ &\quad \left( \left( \tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} \right) \Delta_1 \mathbf{M}_{ij} \right). \end{aligned}$$

Now recall that  $\tilde{\mathbf{V}}_{ij} = \tilde{\mathbf{a}}_{ij} \tilde{\mathbf{a}}_{ij}^T + \tilde{\mathbf{b}}_{ij} \tilde{\mathbf{b}}_{ij}^T$ ; thus, the noisy version of  $\tilde{\mathbf{V}}_{ij}$  has the first-order error term expressed as

$$\Delta_1 \mathbf{V}_{ij} = 2\mathcal{S} \left[ \tilde{\mathbf{a}}_{ij} \Delta_1 \mathbf{a}_{ij}^T + \tilde{\mathbf{b}}_{ij} \Delta_1 \mathbf{b}_{ij}^T \right].$$

Thus,  $\tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} = \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{T}}_{ij} \tilde{\boldsymbol{\phi}} \right) \delta_{ij} + \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{S}}_{ij} \tilde{\boldsymbol{\phi}} \right) \epsilon_{ij}$ . Also  $\mathbb{E}(\delta_{ij} \Delta_1 \mathbf{M}_{ij}) = 2\sigma^2 \mathcal{S}[\tilde{\mathbf{a}}_{ij} \tilde{\mathbf{z}}_{ij}^T]$  and  $\mathbb{E}(\epsilon_{ij} \Delta_1 \mathbf{M}_{ij}) = 2\sigma^2$

$\mathcal{S}[\tilde{\mathbf{b}}_{ij} \tilde{\mathbf{z}}_{ij}^T]$ , where  $\tilde{\mathbf{a}}_{ij}$  and  $\tilde{\mathbf{b}}_{ij}$  denote the first and second columns of  $\nabla \mathbf{z}_{ij}$ . Therefore,  $\mathbb{E}[(\tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}}) \Delta_1 \mathbf{M}_{ij}]$

$$\begin{aligned} &= 2\sigma^2 \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{T}}_{ij} \tilde{\boldsymbol{\phi}} \right) \mathcal{S} \left[ \tilde{\mathbf{a}}_{ij} \tilde{\mathbf{z}}_{ij}^T \right] + 2\sigma^2 \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{S}}_{ij} \tilde{\boldsymbol{\phi}} \right) \mathcal{S} \left[ \tilde{\mathbf{b}}_{ij} \tilde{\mathbf{z}}_{ij}^T \right] \\ &= 2\sigma^2 \mathcal{S}[\tilde{\boldsymbol{\Gamma}}_{ij}], \end{aligned} \tag{87}$$

where  $\tilde{\boldsymbol{\Gamma}}_{ij}$  is defined in Eq. (73). They imply that

$$\begin{aligned} \mathbf{I}_2 &= 4\sigma^2 \sum_{i,k=1}^2 \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \frac{1}{\tilde{\zeta}_i^2 \tilde{\zeta}_k} \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{z}}_{kl} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \\ &\quad \mathcal{S} \left[ \tilde{\mathbf{V}}_{kl} \tilde{\boldsymbol{\phi}} \tilde{\mathbf{z}}_{kl}^T \right] \\ &\quad - 2\sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{1}{\tilde{\zeta}_i^3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \mathcal{S}[\tilde{\boldsymbol{\Gamma}}_{ij}]. \end{aligned} \tag{88}$$

Next, using  $\Delta_2 \mathbf{M}_{ij} = 2\mathcal{S}[\Delta_2 \mathbf{z}_{ij} \tilde{\mathbf{z}}_{ij}^T] + \Delta_1 \mathbf{z}_{ij} \Delta_1 \mathbf{z}_{ij}^T$  and  $\Delta_2 \mathbf{z}_{ij} = (\delta_{ij}^2 + \epsilon_{ij}^2) \mathbf{e}'_i$  give

$$\begin{aligned} \mathbb{E}(\Delta_2^1 \mathcal{M}) &= \sum_{i=1}^2 \sum_{i=1}^{n_i} \tilde{\zeta}_i^{-1} \mathbb{E}(\Delta_2 \mathbf{M}_{ij}) \\ &= \sum_{i=1}^2 \sum_{i=1}^{n_i} \tilde{\zeta}_i^{-1} \mathbb{E} \left( \Delta_1 \mathbf{z}_{ij} \Delta_1 \mathbf{z}_{ij}^T + 2\mathcal{S} \left[ \tilde{\mathbf{z}}_{ij} \Delta_2 \mathbf{z}_{ij}^T \right] \right), \end{aligned} \tag{89}$$

which simply becomes

$$\mathbb{E} \left( \Delta_2^1 \mathcal{M} \right) = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{1}{\tilde{\zeta}_i} \left( \tilde{\mathbf{V}}_{ij} + 4\mathcal{S} \left[ \mathbf{e}'_i \tilde{\mathbf{z}}_{ij}^T \right] \right). \tag{90}$$

Lastly, we evaluate

$$\begin{aligned} \mathbb{E} \left( \Delta_2^2 \mathcal{M} \right) &= -\sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \mathbb{E}(\Delta_1 \gamma_{ij} \Delta_1 \mathbf{M}_{ij}) \\ &= -\sum_{i=1}^2 \sum_{j=1}^{n_i} \frac{1}{\tilde{\zeta}_i^2} \mathbb{E} \left( \left( 2 \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \Delta_1 \hat{\boldsymbol{\phi}} \right) \right. \right. \\ &\quad \left. \left. + \tilde{\boldsymbol{\phi}}^T \Delta_1 \mathbf{V}_{ij} \tilde{\boldsymbol{\phi}} \right) \Delta_1 \mathbf{M}_{ij} \right). \end{aligned}$$

The first term comes from Eq. (85), and the second comes from Eq. (87). Thus,

$$\begin{aligned} \mathbb{E} \left( \Delta_2^2 \mathcal{M} \right) &= 2\sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( \frac{2}{\tilde{\zeta}_i^3} \left( \tilde{\boldsymbol{\phi}}^T \tilde{\mathbf{V}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{z}}_{ij} \right) \right. \\ &\quad \left. \mathcal{S} \left[ \tilde{\mathbf{V}}_{ij} \tilde{\boldsymbol{\phi}} \tilde{\mathbf{z}}_{ij}^T \right] - \frac{1}{\tilde{\zeta}_i^2} \mathcal{S}[\tilde{\boldsymbol{\Gamma}}_{ij}] \right) \end{aligned} \tag{91}$$

This completes the integration of Eq. (69).



To prove Eq. (70), we multiply the identities in Eqs. (84), (88), (90), and (91) by  $\tilde{\phi}$ . First,

$$\mathbf{I}_1 \tilde{\phi} = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} (\tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} + \tilde{\psi}_{ij} \mathbf{I}_5) \tilde{\mathbf{V}}_{ij} \tilde{\phi}. \tag{92}$$

Using the following useful identities

$$2\mathcal{S} \left[ \tilde{\mathbf{V}}_{ij} \tilde{\phi} \tilde{\mathbf{z}}_{ij}^T \right] \tilde{\phi} = \tilde{\mathbf{z}}_{ij} \left( \tilde{\phi}^T \tilde{\mathbf{V}}_{ij} \tilde{\phi} \right) = \zeta_i \tilde{\mathbf{z}}_{ij}, \quad \text{and} \quad \tilde{\mathbf{F}}_i^T \tilde{\phi} = \mathbf{0}.$$

We have  $\mathbf{I}_2 \tilde{\phi}$

$$\begin{aligned} \mathbf{I}_2 \tilde{\phi} &= 2\sigma^2 \sum_{i,k=1}^2 \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \tilde{\zeta}_i^{-2} \tilde{\zeta}_k^{-1} \left( \tilde{\phi}^T \tilde{\mathbf{V}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{z}}_{kl} \right) \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{z}}_{kl} \\ &\quad - \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{F}}_{ij} \tilde{\phi}, \end{aligned}$$

and further,

$$\begin{aligned} \mathbf{I}_2 \tilde{\phi} &= 2\sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \left[ \sum_{k=1}^2 \sum_{l=1}^{n_k} \tilde{\zeta}_k^{-1} \tilde{\mathbf{M}}_{kl} \right] \tilde{\mathcal{M}}^{-} \tilde{\mathbf{V}}_{ij} \tilde{\phi} \\ &\quad - \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{F}}_{ij} \tilde{\phi}. \end{aligned}$$

Applying the definition of  $\tilde{\mathcal{M}}$  and the identity  $\tilde{\mathcal{M}}^{-} \tilde{\mathcal{M}} \tilde{\mathcal{M}}^{-} = \tilde{\mathcal{M}}^{-}$  yields

$$\begin{aligned} \mathbf{I}_2 \tilde{\phi} &= 2\sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{V}}_{ij} \tilde{\phi} \\ &\quad - \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-3} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{F}}_{ij} \tilde{\phi}. \tag{93} \end{aligned}$$

Next

$$\mathbb{E} \left( \Delta_2^1 \mathcal{M} \right) \tilde{\phi} = \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \tilde{\zeta}_i^{-1} \left( \tilde{\mathbf{V}}_{ij} + 2\tilde{\mathbf{z}}_{ij} \mathbf{e}_1^T \right) \tilde{\phi}. \tag{94}$$

Lastly, we apply the above identities again to evaluate

$$\begin{aligned} \mathbb{E} \left( \Delta_2^2 \mathcal{M} \right) \tilde{\phi} &= \sigma^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \\ &\quad \left( 2\tilde{\zeta}_i^{-2} \tilde{\mathbf{M}}_{ij} \tilde{\mathcal{M}}^{-} \tilde{\mathbf{V}}_{ij} \tilde{\phi} - \tilde{\zeta}_i^{-2} \tilde{\mathbf{F}}_{ij} \tilde{\mathbf{A}} \right). \tag{95} \end{aligned}$$

Now, observing that  $\tilde{\mathbf{G}}_2^T \tilde{\phi} = \mathbf{0}$  and combining Eqs. (92), (93), and (94), (95) complete the derivation of Eq. (70). Note that some terms in Eq. (93) and Eq. (95) cancel each other.  $\square$

## References

- Al-Sharadqah, A.: Further statistical analysis for circle fitting algorithms. *Electron. J. Stat.* **8**(2), 2741–2778 (2014)
- Al-Sharadqah, A., Chernov, N.: Error analysis for circle fitting algorithms. *Electron. J. Stat.* **3**(1), 886–911 (2009)
- Al-Sharadqah, A., Ho, K.C.: Solutions and evaluations for fitting of concentric circles. *Signal Process.* **120**, 468–479 (2016)
- Benko, P., Kós, G., Várady, T., Andor, L., Martin, R.: Constrained fitting in reverse engineering. *Comput. Aided Geom. Des.* **19**(3), 173–205 (2002)
- Canny, J.: A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986)
- Chernov, N.: On the convergence of fitting algorithms in computer vision. *J. Math. Imaging Vis.* **27**, 231–239 (2007)
- Chernov, N., Lesort, C.: Statistical efficiency of curve fitting algorithms. *Comput. Stat. Data Anal.* **47**, 713–728 (2004)
- Chojnacki, W., Brooks, M.J., van den Hengel, A., Gawley, D.: On the fitting of surfaces to data with covariances. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 1294–1303 (2000)
- Chojnacki, W., Brooks, M.J., van den Hengel, A.: Rationalising the renormalisation method of Kanatani. *J. Math. Imaging Vis.* **14**, 21–38 (2001)
- Dampegama, S.D.: Determination of original size of the abhayagiriya stupa. In: 21st Asian Conference on Remote Sensing (2000)
- Isserlis, L.: On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika* **12**(1/2), 134–139 (1918)
- Kanatani, K.: Renormalization for unbiased estimation. In: Proceedings 4th International Conference Computer Vision, Berlin, Germany, pp. 599–606 (1993)
- Kanatani, K.: Statistical bias of conic fitting and renormalization. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 320–326 (1994)
- Kanatani, K.: *Statistical Optimization for Geometric Computation: Theory and Practice*. Elsevier, Amsterdam (1996)
- Kweon, I.S., Kim, J.S., Kim, H.-W.: A camera calibration method using concentric circles of vision applications. In: Asian Conference on Computer Vision, pp. 512–520 (2002)
- Leedan, Y., Meer, P.: Heteroscedastic regression in computer vision: problems with bilinear constraint. *Int. J. Comput. Vis.* **37**, 127–150 (2000)
- Ma, Z., Ho, K.C.: Asymptotically efficient estimators for the fittings of coupled circles and ellipses. *Dig. Signal Process.* **25**(2), 28–40 (2014)
- Matei, B., Meer, P.: Reduction of bias in maximum likelihood ellipse fitting. In: 15th International Conference on Computer Vision and Pattern Recognition, vol. 3, Barcelona, Spain, pp. 802–806 (2000)
- O’Leary, P., Harker, M., Zsombor Murray, P.: Direct and least square fitting of coupled geometric objects for metric vision. *IEEE Proc. Vis. Image Signal Process.* **152**, 687–694 (2005)
- Kanatani, K.: Statistical optimization for geometric fitting: theoretical accuracy bound and high order error analysis. *Int. J. Comput. Vis.* **80**, 167–188 (2008)
- Kanatani, K., Rangarajan, P.: Hyper least squares fitting of circles and ellipses. *Comput. Stat. Data Anal.* **55**, 2197–2208 (2011)
- Sampson, P.D.: Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm. *Comput. Graph. Image Process.* **18**, 97–108 (1982)
- Szpak, Z., Chojnacki, W., van den Hengel, A.: A comparison of ellipse fitting methods and implications for multiple-view geometry estimation. In: 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA), pp. 1–8 (2012)
- Szpak, Z., Chojnacki, W., van den Hengel, A.: Guaranteed ellipse fitting with the sampson distance. In: Fitzgibbon, A., Lazebnik,

- S., Perona, P., Sato, Y., Schmid, C. (eds.) *Computer Vision ECCV 2012. Lecture Notes in Computer Science*, vol. 7576, pp. 87–100. Springer, Berlin (2012)
25. Szpak, Z., Chojnacki, W., van den Hengel, A.: Guaranteed ellipse fitting with a confidence region and an uncertainty measure for centre, axes, and orientation. *J. Math. Imaging Vis.* **52**(2), 173–199 (2015)
  26. Yahya, A.E., Nordin, M.J.: A new technique for iris localization in iris recognition systems. *J. Inf. Technol.* **7**, 924–929 (2008)



**Ali Al-Sharadqah** received his B.Sc. and M.Sc. in Mathematics from Jordan University of Science and Technology in 2001 and 2004 respectively, and his Ph.D. in applied mathematics with a concentration in statistics from the University of Alabama at Birmingham in 2011. He is an assistant professor of statistics in the Department of Mathematics at California State University (CSUN) since 2016. Before that he spent three years as assistant professor of statistics at East Carolina

University (2013–2016) and two years as a visiting assistant professor at the University of Mississippi (2011–2013). His research interests extend from statistical analysis of geometric problems to computer vision and signal processing.



**K. C. Ho** was born in Hong Kong. He received the B.Sc. degree with First Class Honors in Electronics in 1988 and the Ph.D. degree in Electronic Engineering in 1991. He was a research associate in the Royal Military College of Canada from 1991 to 1994. He joined the Bell-Northern Research, Montreal, Canada, in 1995 as a member of scientific staff. He was a faculty in the Department of Electrical Engineering at the University of Saskatchewan,

Saskatoon, Canada, from September 1996 to August 1997. Since September 1997, he has been with the University of Missouri and

is currently a Professor in the Electrical and Computer Engineering Department. His research interests are in sensor array processing, source localization, subsurface object detection, wireless communications, and adaptive processing. He was very active in the development of the ITU-T Standard Recommendations from 1995 to 2012. He was the Rapporteur of ITU-T Q15/SG16: Voice Gateway Signal Processing Functions and Circuit Multiplication Equipment/Systems from 2009 to 2012, and the Associate Rapporteur of ITU-T Q16/SG16: Speech Enhancement Functions in Signal Processing Network Equipment in 2012. He was the Editor of the ITU-T Recommendations G.160: Voice Enhancement Devices from 2006 to 2012, G.168: Digital Network Echo Cancellers from 2000 to 2012, and G.799.2: Mechanism for Dynamic Coordination of Signal Processing Network Equipment from 2004 to 2009. Dr. Ho has served as a Technical Co-Chair of the IEEE International Conference on Acoustics, Speech and Signal Processing 2016 (ICASSP2016). Dr. Ho was the Past Chair (2015), the Chair (2013–2014) and Vice-Chair (2011–2012) of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. He was in the organizing committees of the IEEE SAM 2008 Workshop and the IEEE CAMSAP 2011. He was an Associate Editor of the IEEE Transactions on Signal Processing (2009–2013, 2003–2006) and the IEEE Signal Processing Letters (2004–2008). He received the Senior Faculty Research Award in 2014 and 2009, the Junior Faculty Research Award in 2003, and the Teaching Award in 2006 from the College of Engineering at the University of Missouri. He is an inventor of 21 patents in the USA, Canada, Europe, and Asia on geolocation and signal processing for wireless communications.