IPSJ Transactions on Computer Vision and Applications Vol.5 19–29 (Apr. 2013)

[DOI: 10.2197/ipsjtcva.5.19]

Regular Paper

Hyperaccurate Correction of Maximum Likelihood for Geometric Estimation

Kenichi Kanatani[1,a)] Yasuyuki Sugaya[2,b)]

Received: August 7, 2012, Accepted: February 1, 2012, Released: April 11, 2013

Abstract: The best known method for optimally computing parameters from noisy data based on geometric constraints is maximum likelihood (ML). This paper reinvestigates "hyperaccurate correction" for further improving the accuracy of ML. In the past, only the case of a single scalar constraint was studied. In this paper, we extend it to multiple constraints given in the form of vector equations. By detailed error analysis, we illuminate the existence of a term that has been ignored in the past. Doing simulation experiments of ellipse fitting, fundamental matrix, and homography computation, we show that the new term does not effectively affect the final solution. However, we show that our hyperaccurate correction is even superior to hyper-renormalization, the latest method regarded as the best fitting method, but that the iterations of ML computation do not necessarily converge in the presence of large noise.

Keywords: geometric estimation, maximum likelihood, bias estimation, hyperaccurate correction

1. Introduction

One of the most fundamental tasks of computer vision is to compute the 2-D and 3-D shapes of objects based on *geometric constraints*, by which we mean properties that can be described by relatively simple equations such as the objects being lines or planes, their being parallel or orthogonal, and the camera imaging geometry being perspective projection. We call the inference based on such geometric constraints *geometric estimation*. Techniques for optimal geometric estimation in the presence of noise has been extensively studied since 1980s by many researchers including the authors [3], [6] .

Currently, it widely is recognized that the highest accuracy is achieved by methods based on *maximum likelihood* (*ML*) and those based on *renormalization* [9]. For ML, we minimize the *Mahalanobis distance*, a special case of which is the *reprojection error* [3]. One of the problem of ML is that it is not a *convex problem* [1], for which a global optimum is easily obtained. Kahl and Hartley [4] showed that if the L_2-norm used in ML optimization is replaced by the L_∞-norm, the problem can be converted to a *quasiconvex problem*, for which a global optimum can be obtained by iteratively using linear programming (LP) and second order conic programming (SCOP) and that the accuracy is comparable to ML, although ML is theoretically desirable if it can be computed. In this paper, we concentrate on problem for which the ML solution can be obtained. Kanatani [7], [8] showed that the accuracy of ML is further improved by analyzing the statistical bias of the solution and subtracting it, which he called *hyperaccurate correction*.

On the other hand, renormalization [5], [6] does not minimize any cost function; it directly estimates a biasless solution. Recently, Kanatani et al. [10] did higher order error analysis and derived an improved version, called *hyper-renormalization*. They showed that it can compute a solution without bias up to second order noise terms and demonstrated by experiments that it is superior to ML [10]. According to comparative experiments, it is observed that ML with hyperaccurate correction slightly surpasses hyper-renormalization. However, the iterations for ML computation, such as the FNS of Chojnacki et al. [2], do not necessarily converge in the presence of large noise. Hyper-renormalization, on the other hand, is very robust to noise and converges after a few iterations, because it is an iterative improvement of *HyperLS* [13], [14], [22], an algebraic method with very high accuracy. Thus, ML with hyperaccurate correction and hyper-renormalization both have strength and weakness, as reviewed by Kanatani [9].

The bias of ML has also been studied in the domain of traditional statistical estimation, where observations are explicitly expressed in terms of noise (such expressions are called the *statistical model*) and the estimation performance is evaluated by asymptotic analysis in the limit $N \to \infty$ of the number N of observations. Okatani and Deguchi [19] adopted this approach to computer vision problems by introducing auxiliary variables, reducing the problem to the form of nonlinear regression, and employing semiparametric modeling. They also attempted to remove bias by analyzing the curvature of a hypersurface defined by the statistical model [20] and using a bias removal scheme based on projected scores [21]. The main difference of the geometric estimation we consider here from traditional statistical estimation is that the constraints are treated as implicit functions and estimation performance is evaluated in the perturbation limit $\sigma \to 0$

1 Department of Computer Science, Okayama University, Japan
2 Department of Computer Science and Engineering, Toyohashi University of Technology, Japan
a) kanatani@suri.cs.okayama-u.ac.jp
b) sugaya@iim.cs.tut.ac.jp

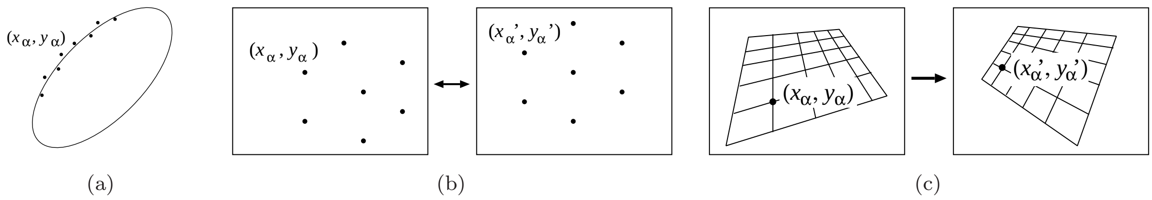

Fig. 1 (a) Fitting an ellipse to a point sequence. (b) Computing the fundamental matrix from corresponding points between two images. (c) Computing a homography between two images.

of the noise level σ [8].

The purpose of this paper is to reexamine the hyperaccurate correction of ML, for which only the case of a single constraint was analyzed in the past [7], [8]. Here, we extend it to multiple constraints given as a vector equation. Doing a detailed error analysis, we point out the existence of a term that has been ignored in the past. We also do numerical experiments to see how that term affects the final solution.

In Sec. 2, we give a mathematical formulation of our geometric estimation. We introduce our noise model in Sec. 3, and describe the ML optimization procedure in Sec. 4. We do error analysis of ML in the multiple constraint case in Sec. 5 and explicitly evaluate the bias of the resulting solution in Sec. 6. Our extended hyperaccurate correction scheme is described in Sec. 7. In Sec. 8, we do numerical experiments to compare the accuracy of hyperaccurate correction with existing methods including hyper-renormalization. We also examine the effect of the newly introduced term. In Sec. 9, we conclude.

2. Geometric Estimation

The geometric estimation problem we consider here is defined as follows. Suppose we observe N noisy observations $\boldsymbol{x}_1$, ..., $\boldsymbol{x}_N$, which are n-D vectors. We assume that their noiseless values $\bar{\boldsymbol{x}}_1$, ..., $\bar{\boldsymbol{x}}_N$ should satisfy L equations or "constraints"

$$F^{(k)}(\boldsymbol{x};\boldsymbol{\theta}) = 0, \qquad k = 1, ..., L, \tag{1}$$

where $\boldsymbol{\theta}$ is an unknown parameter vector that specifies the 2-D/3-D shapes of the objects we are viewing or their 2-D/3-D motions. The function $F^{(k)}(\boldsymbol{x};\boldsymbol{\theta})$ in Eq. (1) has generally a nonlinear form, but in many practical applications it is linear in unknown parameters or can be reparameterized so that it is linear in them. Then, Eq. (1) can be written in the form

$$(\boldsymbol{\xi}^{(k)}(\boldsymbol{x}), \boldsymbol{\theta}) = 0, \qquad k = 1, ..., L, \tag{2}$$

where and hereafter we denote the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by $(\boldsymbol{a}, \boldsymbol{b})$. In Eq. (2), $\boldsymbol{\xi}^{(k)}(\boldsymbol{x})$ is some nonlinear mapping of $\boldsymbol{x}$ from $\mathcal{R}^m$ to $\mathcal{R}^n$, where m and n are the dimensions of the data $\boldsymbol{x}_\alpha$ and the parameter $\boldsymbol{\theta}$, respectively: the ith component $\xi_i^{(k)}(\boldsymbol{x})$ of $\boldsymbol{\xi}^{(k)}(\boldsymbol{x})$ are those terms of (1) that are multiplied by the ith component θ_i of $\boldsymbol{\theta}$, and those terms that do not involve $\boldsymbol{\theta}$ are regarded as multiplied by a constant, which we also regard as one component of $\boldsymbol{\theta}$ (see the examples below). We further assume that the L vectors $\boldsymbol{\xi}^{(k)}(\boldsymbol{x})$ need not be linearly independent. We call the number r of independent ones the *rank* of the constraint. Since the vector $\boldsymbol{\theta}$ in (2) has scale indeterminacy, we normalize it to unit norm: $\|\boldsymbol{\theta}\| = 1$.

Example 1 (Ellipse fitting). Given a point sequence (x_α, y_α), $\alpha = 1$, ..., N, we wish to fit an ellipse of the form

$$Ax^2 + 2Bxy + Cy^2 + 2f_0(Dx + Ey) + f_0^2 F = 0. \tag{3}$$

(Fig. 1(a)). If we let

$$\boldsymbol{\xi} = (x^2, 2xy, y^2, 2f_0x, 2f_0y, f_0^2)^\top,$$
$$\boldsymbol{\theta} = (A, B, C, D, E, F)^\top, \tag{4}$$

the ellipse equation (3) has the form of Eq. (2) with $L = 1$. Here, f_0 is a scale constant to stabilize finite length numerical computation so that the components of the vector $\boldsymbol{\xi}$ have the same order of magnitude. We choose it to be of an approximate magnitude of the data x_α and y_α.

Example 2 (Fundamental matrix computation). Corresponding points (x, y) and (x', y') in two images of the same scene taken from different positions satisfy the *epipolar equation* [3]

$$(\begin{pmatrix} x \\ y \\ f_0 \end{pmatrix}, \boldsymbol{F}\begin{pmatrix} x' \\ y' \\ f_0 \end{pmatrix}) = 0, \tag{5}$$

where $\boldsymbol{F}$ is a matrix of rank 2 called the *fundamental matrix*, from which we can compute the camera positions and the 3-D structure of the scene (Fig. 1(b)). As in the case of ellipse fitting, f_0 is a scale constant to stabilize finite length computation. If we let

$$\boldsymbol{\xi} = (xx', xy', f_0x, yx', yy', f_0y, f_0x', f_0y', f_0^2)^\top,$$
$$\boldsymbol{\theta} = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33})^\top, \tag{6}$$

the eipolar equation (5) has the form of Eq. (2) with $L = 1$.

Example 3 (Homography computation). Two images of a planar surface or infinitely far away scene (Fig. 1(c)) are related by a *homography* of the form

$$x' = f_0\frac{h_{11}x + h_{12}y + h_{13}f_0}{h_{31}x + h_{32}y + h_{33}f_0},$$
$$y' = f_0\frac{h_{21}x + h_{22}y + h_{23}f_0}{h_{31}x + h_{32}y + h_{33}f_0}, \tag{7}$$

where f_0 is a scale constant of the order of the data. In matrix form, Eq. (7) is equivalently rewritten as

$$\begin{pmatrix} x' \\ y' \\ f_0 \end{pmatrix} \simeq \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ f_0 \end{pmatrix}, \tag{8}$$

where $\simeq$ denotes equality up to a nonzero multiplier. This equation means that the vectors on both sides are parallel to each other, so we can alternatively write this as

$$\begin{pmatrix} x' \\ y' \\ f_0 \end{pmatrix} \times \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ f_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{9}$$

The three components of this vector equation have the form of Eq. (2) with $L = 3$ as follows [11]:

$$(\boldsymbol{\xi}^{(1)}, \boldsymbol{\theta}) = 0, \quad (\boldsymbol{\xi}^{(2)}, \boldsymbol{\theta}) = 0, \quad (\boldsymbol{\xi}^{(3)}, \boldsymbol{\theta}) = 0. \tag{10}$$

Here, we define

$$\begin{aligned} \boldsymbol{\theta} &= (h_{11}\, h_{12}\, h_{13}\, h_{21}\, h_{22}\, h_{23}\, h_{31}\, h_{32}\, h_{33})^\top, \\ \boldsymbol{\xi}^{(1)} &= (0, 0, 0, -f_0 x, -f_0 y, -f_0^2, xy', yy', f_0 y')^\top, \\ \boldsymbol{\xi}^{(2)} &= (f_0 x, f_0 y, f_0^2, 0, 0, 0, -xx', -yx', -f_0 x')^\top, \\ \boldsymbol{\xi}^{(3)} &= (-xy', -yy', -f_0 y', xx', yx', f_0 x', 0, 0, 0)^\top. \end{aligned} \tag{11}$$

The three components of (9) have the form of (2) with $L = 3$. Note that $\boldsymbol{\xi}^{(1)}$, $\boldsymbol{\xi}^{(2)}$, and $\boldsymbol{\xi}^{(3)}$ are *linearly dependent*; only two of them are independent, so the rank is $r = 2$.

3. Noise modeling

We regard each observation $\boldsymbol{x}_\alpha$ as perturbed from its true value $\bar{\boldsymbol{x}}_\alpha$ by independent Gaussian noise $\Delta\boldsymbol{x}_\alpha$ of mean $\boldsymbol{0}$ and covariance matrix $\sigma^2 V_0[\boldsymbol{x}_\alpha]$, where σ is an unknown constant, which we call the *noise level*, that describes the magnitude of noise, while $V_0[\boldsymbol{x}_\alpha]$ is a matrix, which we call the *normalized covariance matrix*, that specifies the orientation dependence of the noise distribution. We assume that the normalized covariance matrix $V_0[\boldsymbol{x}_\alpha]$ is known. The separation of $V[\boldsymbol{x}_\alpha]$ into σ^2 and $V_0[\boldsymbol{x}_\alpha]$ is merely a matter of convenience; there is no fixed rule. This convention is motivated by the fact that estimation of the absolute magnitude of data uncertainty is very difficult in practice, while optimal estimation can be done only from the knowledge of $V_0[\boldsymbol{x}_\alpha]$.

If the noise distribution is homogeneous, i.e., the same for all $\boldsymbol{x}_\alpha$, and isotropic, i.e. the same for all directions, we can let $V_0[\boldsymbol{x}_\alpha] = \boldsymbol{I}$ (the identity). It has been observed that for feature point detection in 2-D images, it is sufficient to assume homogeneous and isotopic noise with $V_0[\boldsymbol{x}_\alpha] = \boldsymbol{I}$ for most applications [18], while accurate estimation of $V_0[\boldsymbol{x}_\alpha]$ is crucial for 3-D data [12] because 3-D data are obtained by 3-D sensors, such as stereo vision and laser sensing, which have strong orientation dependence with different accuracy in the depth direction and the directions orthogonal to it.

Let us write $\boldsymbol{\xi}^{(k)}(\boldsymbol{x}_\alpha)$ simply as $\boldsymbol{\xi}_\alpha^{(k)}$. It can be expanded in the form

$$\boldsymbol{\xi}_\alpha^{(k)} = \bar{\boldsymbol{\xi}}_\alpha^{(k)} + \Delta_1\boldsymbol{\xi}_\alpha^{(k)} + \Delta_2\boldsymbol{\xi}_\alpha^{(k)} + \cdots, \tag{12}$$

where and hereafter the bar denotes the noiseless value and Δ_m denotes mth order terms in the noise level σ. The first order noise term $\Delta_1\boldsymbol{\xi}_\alpha^{(k)}$ is expressed in terms of the original noise term $\Delta\boldsymbol{x}_\alpha$ in $\boldsymbol{x}_\alpha$ and the Jacobian matrices of the mapping $\boldsymbol{\xi}^{(k)}(\boldsymbol{x})$ in the following form:

$$\Delta_1\boldsymbol{\xi}_\alpha^{(k)} = \boldsymbol{T}_\alpha^{(k)}\Delta\boldsymbol{x}_\alpha, \quad \boldsymbol{T}_\alpha^{(k)} \equiv \left.\frac{\partial\boldsymbol{\xi}^{(k)}(\boldsymbol{x})}{\partial\boldsymbol{x}}\right|_{\boldsymbol{x}=\bar{\boldsymbol{x}}_\alpha}. \tag{13}$$

We define the covariance matrix $V^{(kl)}[\boldsymbol{\xi}_\alpha]$ between $\boldsymbol{\xi}_\alpha^{(k)}$ and $\boldsymbol{\xi}_\alpha^{(l)}$ by

$$V^{(kl)}[\boldsymbol{\xi}_\alpha] = E[\Delta\boldsymbol{\xi}_\alpha^{(k)}\Delta\boldsymbol{\xi}_\alpha^{(l)\top}] = V_0^{(kl)}[\boldsymbol{\xi}_\alpha], \tag{14}$$

where $E[\,\cdot\,]$ denotes the expectation over data uncertainty. The following relationship holds:

$$V^{(kl)}[\boldsymbol{\xi}_\alpha] = \boldsymbol{T}_\alpha^{(k)} V[\boldsymbol{x}_\alpha] \boldsymbol{T}_\alpha^{(l)\top}. \tag{15}$$

Example 4 (Ellipse fitting). The first order noise term $\Delta_1\boldsymbol{\xi}_\alpha$ is

$$\Delta_1\boldsymbol{\xi}_\alpha = \boldsymbol{T}_\alpha \begin{pmatrix} \Delta x_\alpha \\ \Delta y_\alpha \end{pmatrix}, \quad \boldsymbol{T}_\alpha = 2\begin{pmatrix} \bar{x}_\alpha & \bar{y}_\alpha & 0 & f_0 & 0 & 0 \\ 0 & \bar{x}_\alpha & \bar{y}_\alpha & 0 & f_0 & 0 \end{pmatrix}^\top, \tag{16}$$

and the second order noise term $\Delta_2\boldsymbol{\xi}_\alpha$ is

$$\Delta_2\boldsymbol{\xi}_\alpha = (\Delta x_\alpha^2, 2\Delta x_\alpha \Delta y_\alpha, \Delta y_\alpha^2, 0, 0, 0)^\top. \tag{17}$$

Example 5 (Fundamental matrix computation). The first order noise term $\Delta_1\boldsymbol{\xi}_\alpha$ is

$$\Delta_1\boldsymbol{\xi}_\alpha = \boldsymbol{T}_\alpha(\Delta x_\alpha, \Delta y_\alpha, \Delta x'_\alpha, \Delta y'_\alpha)^\top, \quad \boldsymbol{T}_\alpha = \begin{pmatrix} \bar{x}'_\alpha & \bar{y}'_\alpha & f_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{x}'_\alpha & \bar{y}'_\alpha & f_0 & 0 & 0 & 0 \\ \bar{x}_\alpha & 0 & 0 & \bar{y}_\alpha & 0 & 0 & 1 & 0 & 0 \\ 0 & \bar{x}_\alpha & 0 & 0 & \bar{y}_\alpha & 0 & 0 & f_0 & 0 \end{pmatrix}^\top, \tag{18}$$

and the second order noise term $\Delta_2\boldsymbol{\xi}_\alpha$ is

$$\Delta_2\boldsymbol{\xi}_\alpha = (\Delta x_\alpha \Delta x'_\alpha, \Delta x_\alpha \Delta y'_\alpha, 0, \Delta y_\alpha \Delta x'_\alpha, \Delta y_\alpha \Delta y'_\alpha, 0, 0, 0, 0)^\top. \tag{19}$$

Example 6 (Homography computation). The first order noise term $\Delta_1\boldsymbol{\xi}_\alpha^{(k)}$ is

$$\begin{aligned} \Delta_1\boldsymbol{\xi}_\alpha^{(k)} &= \boldsymbol{T}_\alpha^{(k)}(\Delta x_\alpha, \Delta y_\alpha, \Delta x'_\alpha, \Delta y'_\alpha)^\top, \\ \boldsymbol{T}_\alpha^{(1)} &= \begin{pmatrix} 0 & 0 & 0 & -f_0 & 0 & 0 & \bar{y}'_\alpha & 0 & 0 \\ 0 & 0 & 0 & 0 & -f_0 & 0 & 0 & \bar{y}'_\alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bar{x}_\alpha & \bar{y}_\alpha & f_0 \end{pmatrix}^\top, \\ \boldsymbol{T}_\alpha^{(2)} &= \begin{pmatrix} f_0 & 0 & 0 & 0 & 0 & 0 & -\bar{x}'_\alpha & 0 & 0 \\ 0 & f_0 & 0 & 0 & 0 & 0 & 0 & -\bar{x}'_\alpha & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\bar{x}_\alpha & -\bar{y}_\alpha & -f_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top, \\ \boldsymbol{T}_\alpha^{(3)} &= \begin{pmatrix} -\bar{y}'_\alpha & 0 & 0 & \bar{x}'_\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & -\bar{y}'_\alpha & 0 & 0 & \bar{x}'_\alpha & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{x}_\alpha & \bar{y}_\alpha & f_0 & 0 & 0 & 0 \\ -\bar{x}_\alpha & -\bar{y}_\alpha & -f_0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^\top, \end{aligned} \tag{20}$$

and the second order noise term $\Delta_2\boldsymbol{\xi}_\alpha$ is

$$\begin{aligned} \Delta_2\boldsymbol{\xi}_\alpha^{(1)} &= (0, 0, 0, 0, 0, 0, \Delta x_\alpha \Delta y'_\alpha, \Delta y_\alpha \Delta y'_\alpha, 0)^\top, \\ \Delta_2\boldsymbol{\xi}_\alpha^{(2)} &= (0, 0, 0, 0, 0, 0, -\Delta x'_\alpha \Delta x_\alpha, -\Delta x'_\alpha \Delta y_\alpha, 0)^\top, \\ \Delta_2\boldsymbol{\xi}_\alpha^{(3)} &= (-\Delta y'_\alpha \Delta x_\alpha, -\Delta y'_\alpha \Delta y_\alpha, 0, \Delta x'_\alpha \Delta x_\alpha, \Delta x'_\alpha \Delta y_\alpha, 0, 0, 0, 0)^\top. \end{aligned} \tag{21}$$

The Jacobian matrices $\boldsymbol{T}_\alpha^{(k)}$ contain the true values of the observations, which are replaced by observed values. It has been confirmed by many experiments that this replacement does not affect the final results. Here, the covariance matrices among $\boldsymbol{\xi}_\alpha^{(k)}$ are defined in terms of the first order derivatives of $\boldsymbol{\xi}^{(k)}(\boldsymbol{x}_\alpha)$ in terms of the Jacobian matrices $\boldsymbol{T}_\alpha^{(k)}$, but it has also been confirmed that inclusion of higher order derivatives does not affect the final results.

4. Maximum Likelihood

In our setting, *maximum likelihood* (*ML*) is to minimize the *Mahalanobis distance*

$$J = \frac{1}{N}\sum_{\alpha=1}^{N}(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha, V_0[\boldsymbol{x}_\alpha]^{-1}(\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha)), \quad (22)$$

subject to

$$(\boldsymbol{\xi}(\bar{\boldsymbol{x}}_\alpha), \boldsymbol{\theta}) = 0. \quad (23)$$

If the noise is homogeneous and isotropic, we can let $V_0[\boldsymbol{x}_\alpha] = \boldsymbol{I}$, so the right side of Eq. (22) is $(1/N)\sum_{\alpha=1}^{N}\|\boldsymbol{x}_\alpha - \bar{\boldsymbol{x}}_\alpha\|^2$, which is commonly referred to as the *reprojection error* [3].

Minimizing Eq. (22) subject to Eq. (23) is generally a complicated nonlinear optimization, but the computation is simplified if the transformed variable $\boldsymbol{\xi}_\alpha^{(k)}$ is regarded as subject to independent Gaussian noise of mean $\boldsymbol{0}$ and covariance matrices $V^{(kl)}[\boldsymbol{\xi}_\alpha] = \sigma^2 V_0^{(kl)}[\boldsymbol{\xi}_\alpha]$, although this is not strictly true. Under this Gaussian noise approximation, the constraint in Eq. (23) can be eliminated using Lagrange multipliers [6]. Then, the Mahalanobis distance in Eq. (22) reduces to

$$J = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3} W_\alpha^{(kl)}\boldsymbol{\xi}_\alpha^{(k)}\boldsymbol{\xi}_\alpha^{(l)\top}, \quad (24)$$

where $W_\alpha^{(kl)}$ is the (kl) element of the pseudoinvers of truncated rank r of of the matrix whose (kl) element is $(\boldsymbol{\theta}, V_0^{(kl)}[\boldsymbol{\xi}_\alpha]\boldsymbol{\theta})$. We write this symbolically as follows:

$$\left(W_\alpha^{(kl)}\right) = \left((\boldsymbol{\theta}, V_0^{(kl)}[\boldsymbol{\xi}_\alpha]\boldsymbol{\theta})\right)_r^-. \quad (25)$$

By "truncated rank r", we mean that the eigenvalues except the r largest ones are replaced by 0 in the spectral decomposition. Today, Eq. (24) is known as the *Sampson error* [3] after the pioneering ellipse fitting scheme of P. D. Sampson [23]. In the single constraint case ($L = 1$), Eq. (24) is easily minimized by the FNS of Chojnacki et al. [2], which can be straightforwardly extended to the multiple constraint case ($L > 1$) [11]. The minimizer of the Sampson error (24), or the "Sampson solution" for short, is not exactly the ML solution that minimizes Eq. (22), but we can modify Eq. (24) by using the computed Sampson solution, minimize the resulting modified Sampson error, and iterate this process. It can be shown that in the end the modified Sampson error coincides with the Mahalanobis distance (22), meaning that we obtain the exact ML solution [15]. It has been observed that Sampson error modification iterations converge after a few rounds but the solution does not change except a few of the least significant digits [11], [16], [17]. Hence, we can practically identify the Sampson solution with the exact ML solution. We now do detailed error analysis of the solution that minimizes Eq. (24).

5. Error Analysis

The derivative of Eq. (4) with respect to $\boldsymbol{\theta}$ has the following form [11]:

$$\nabla_{\mathbf{u}} J = 2(\boldsymbol{M} - \boldsymbol{L})\boldsymbol{\theta}, \quad (26)$$

$$\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3} W_\alpha^{(kl)}\boldsymbol{\xi}_\alpha^{(k)}\boldsymbol{\xi}_\alpha^{(l)\top},$$
$$\boldsymbol{L} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3} W_\alpha^{(km)}W_\alpha^{(ln)}(\boldsymbol{\xi}_\alpha^{(m)}, \boldsymbol{\theta})(\boldsymbol{\xi}_\alpha^{(n)}, \boldsymbol{\theta})V_0^{(kl)}[\boldsymbol{\xi}_\alpha]. \quad (27)$$

If Eq. (12) is substituted, the matrix $\boldsymbol{M}$ is expanded in the form

$$\boldsymbol{M} = \bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_2\boldsymbol{M} + \cdots, \quad (28)$$

where $\cdots$ denotes terms of order 3 or higher in σ. The terms $\Delta_1\boldsymbol{M}$ and $\Delta_2\boldsymbol{M}$ have the following expressions:

$$\Delta_1\boldsymbol{M} = \Delta_1^0\boldsymbol{M} + \Delta_1^*\boldsymbol{M}, \quad (29)$$

$$\Delta_2\boldsymbol{M} = \Delta_2^0\boldsymbol{M} + \Delta_2^*\boldsymbol{M} + \Delta_2^\dagger\boldsymbol{M}, \quad (30)$$

$$\Delta_1^0\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(kl)}(\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top} + \bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(l)\top}), \quad (31)$$

$$\Delta_1^*\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\Delta_1 W_\alpha^{(kl)}\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}, \quad (32)$$

$$\Delta_2^0\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(kl)}(\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(l)\top} + \Delta_2\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}$$
$$+ \bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_2\boldsymbol{\xi}_\alpha^{(l)\top}), \quad (33)$$

$$\Delta_2^*\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\Delta_1 W_\alpha^{(kl)}(\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top} + \bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(l)\top}), \quad (34)$$

$$\Delta_2^\dagger\boldsymbol{M} \equiv \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\Delta_2 W_\alpha^{(kl)}\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}. \quad (35)$$

Here, $\Delta_1 W_\alpha^{(kl)}$ and $\Delta_2 W_\alpha^{(kl)}$ are written as follows (see Appendix A.1):

$$\Delta_1 W_\alpha^{(kl)} = -2\sum_{m,n=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\Delta_1\boldsymbol{\theta}, V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}}),$$
$$\Delta_2 W_\alpha^{(kl)} = -\sum_{m,n=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}\Big((\Delta_1\boldsymbol{\theta}, V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\Delta_1\boldsymbol{\theta}),$$
$$+2(\Delta_2\boldsymbol{\theta}, V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\Big). \quad (36)$$

For the matrix $\boldsymbol{L}$ in Eq. (27), we obtain from $(\bar{\boldsymbol{\xi}}_\alpha^{(k)}, \bar{\boldsymbol{\theta}}) = 0$

$$\boldsymbol{L} = \bar{\boldsymbol{L}} + \Delta_1\boldsymbol{L} + \Delta_2\boldsymbol{L} + \cdots, \quad \bar{\boldsymbol{L}} = \Delta_1\boldsymbol{L} = \boldsymbol{O}, \quad (37)$$

$$\Delta_2\boldsymbol{L} = \frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}\Big((\bar{\boldsymbol{\xi}}_\alpha^{(m)}, \Delta_1\boldsymbol{\theta})(\bar{\boldsymbol{\xi}}_\alpha^{(n)}, \Delta_1\boldsymbol{\theta})$$
$$+(\bar{\boldsymbol{\xi}}_\alpha^{(m)}, \Delta_1\boldsymbol{\theta})(\Delta_1\boldsymbol{\xi}_\alpha^{(n)}, \bar{\boldsymbol{\theta}}) + (\Delta_1\boldsymbol{\xi}_\alpha^{(m)}, \bar{\boldsymbol{\theta}})(\bar{\boldsymbol{\xi}}_\alpha^{(n)}, \Delta_1\boldsymbol{\theta})$$
$$+(\Delta_1\boldsymbol{\xi}_\alpha^{(m)}, \bar{\boldsymbol{\theta}})(\Delta_1\boldsymbol{\xi}_\alpha^{(n)}, \bar{\boldsymbol{\theta}})\Big)V_0^{(kl)}[\boldsymbol{\xi}_\alpha]. \quad (38)$$

Substituting this into $\boldsymbol{M}\boldsymbol{\theta} = \boldsymbol{L}\boldsymbol{\theta}$, which is obtained by letting Eq. (26) be $\boldsymbol{0}$, we have

$$(\bar{\boldsymbol{M}} + \Delta_1\boldsymbol{M} + \Delta_2\boldsymbol{M} + \cdots)(\bar{\boldsymbol{\theta}} + \Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{\theta} + \cdots)$$
$$= \Delta_2\boldsymbol{L}(\bar{\boldsymbol{\theta}} + \Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{\theta} + \cdots). \quad (39)$$

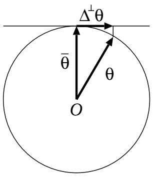

Fig. 2 The true value $\bar{\boldsymbol{\theta}}$, the computed value $\boldsymbol{\theta}$, and its orthogonal component $\Delta^{\perp}\boldsymbol{\theta}$ of $\bar{\boldsymbol{\theta}}$.

Equating terms of the same order on both sides, we obtain

$$\bar{\boldsymbol{M}}\Delta_1\boldsymbol{\theta} + \Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}} = \boldsymbol{0}, \tag{40}$$

$$\bar{\boldsymbol{M}}\Delta_2\boldsymbol{\theta} + \Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta} + \Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}} = \Delta_2\boldsymbol{L}\bar{\boldsymbol{\theta}}. \tag{41}$$

Multiplying Eq. (40) by the pseudoinverse $\bar{\boldsymbol{M}}^-$ on both sides and noting that $\bar{\boldsymbol{M}}^-\bar{\boldsymbol{M}} = \boldsymbol{P}_{\bar{\boldsymbol{\theta}}}$ (the projection matrix along $\bar{\boldsymbol{\theta}}$) and that $\Delta_1\boldsymbol{\theta}$ is orthogonal to $\bar{\boldsymbol{\theta}}$, we can write the first order error term $\Delta_1\boldsymbol{\theta}$ as follows:

$$\Delta_1\boldsymbol{\theta} = -\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}} = -\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}. \tag{42}$$

Here, we have noted that $(\bar{\boldsymbol{\xi}}_\alpha, \bar{\boldsymbol{\theta}}) = 0$ implies $\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{\theta}} = \boldsymbol{0}$. Multiplying Eq. (41) by $\bar{\boldsymbol{M}}^-$ on both sides, we obtain

$$\Delta_2^{\perp}\boldsymbol{\theta} = -\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta} - \bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}} + \bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{L}\bar{\boldsymbol{\theta}}, \tag{43}$$

where we defined $\Delta_2^{\perp}\boldsymbol{\theta} = \boldsymbol{P}_{\bar{\boldsymbol{\theta}}}\Delta_2\boldsymbol{\theta}$ to be the error component of $\Delta_2\boldsymbol{\theta}$ orthogonal to $\bar{\boldsymbol{\theta}}$ (Fig. 2).

6. Bias Analysis

Since the expectation of odd-order error terms is zero, we have $E[\Delta_1\boldsymbol{\theta}] = \boldsymbol{0}$. This means that the first order bias is $\boldsymbol{0}$, so we focus on the second order bias $E[\Delta_2^{\top}\boldsymbol{\theta}]$. From Eq. (43), we obtain

$$E[\Delta_2^{\perp}\boldsymbol{\theta}] = -E[\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta}] - E[\bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}}] + E[\bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{L}\bar{\boldsymbol{\theta}}]. \tag{44}$$

We now evaluated each term separately. The basic strategy is to eliminate the noise terms $\Delta_1\boldsymbol{\xi}_\alpha^{(k)}$ in the expectation expression, using the identity

$$E[\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\beta^{(l)\top}] = \sigma^2\delta_{\alpha\beta}V_0^{(kl)}[\boldsymbol{\xi}_\alpha], \tag{45}$$

obtained from our assumption of independent noise, where $\delta_{\alpha\beta}$ is the Kronecker delta, taking 1 for $\alpha = \beta$ and 0 otherwise. We also eliminate $\Delta_2\boldsymbol{\xi}_\alpha^{(k)}$ in the expectation expression by defining a new quantity $\boldsymbol{e}_\alpha^{(k)}$ by

$$E[\Delta_2\boldsymbol{\xi}_\alpha^{(k)}] = \sigma^2\boldsymbol{e}_\alpha^{(k)}. \tag{46}$$

6.1 The first term

The first term of Eq. (44) is written as

$$-E[\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta}] = E[\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}] + E[\bar{\boldsymbol{M}}^-\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}]. \tag{47}$$

The first term on the left side is written as follows:

$$\begin{aligned} &E[\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}] \\ &= E\Big[\frac{1}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha,\beta=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(kl)}\bar{W}_\beta^{(mn)}\Big(\Delta_1\boldsymbol{\xi}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top} \\ &\quad + \bar{\boldsymbol{\xi}}_\alpha^{(k)}\Delta_1\boldsymbol{\xi}_\alpha^{(l)\top}\Big)\bar{\boldsymbol{M}}^-(\Delta_1\boldsymbol{\xi}_\beta^{(m)}, \bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\beta^{(n)}\Big] \\ &= \frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(kl)}\bar{W}_\alpha^{(mn)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \\ &\quad \bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(n)})V_0^{(km)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}} \\ &\quad + \frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(kl)}\bar{W}_\alpha^{(mn)}(\bar{\boldsymbol{\theta}}, \\ &\quad V_0^{(ml)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(n)})\bar{\boldsymbol{\xi}}_\alpha^{(k)}. \end{aligned} \tag{48}$$

From Eq. (34) and the second of Eq. (36), we obtain

$$\begin{aligned} \Delta_1^*\boldsymbol{M} &= \frac{2}{N}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}, \\ &\quad V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(l)\top}. \end{aligned} \tag{49}$$

Hence the second term on the right side of Eq. (47) is

$$\begin{aligned} &E[\bar{\boldsymbol{M}}^-\Delta_1^*\boldsymbol{M}\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}] \\ &= E\Big[\frac{2}{N}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}, \\ &\quad V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \bar{\boldsymbol{M}}^-\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\Big] \\ &= \frac{2}{N}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \\ &\quad E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top]V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)} \\ &= \frac{2\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \bar{\boldsymbol{M}}^-V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}, \end{aligned} \tag{50}$$

where we have used the fact that $E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top]$ has the expression (see Appendix A.2)

$$E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top] = \frac{\sigma^2}{N}\bar{\boldsymbol{M}}^-, \tag{51}$$

which is the leading term of the covariance matrix of the computed $\boldsymbol{\theta}$. Equation (51) it coincides with the theoretical accuracy limit called the *KCR* (*Kanatani-Cramer-Rao*) *lower bound* [6], [8], [14]. Thus, Eq. (47) is written as follows:

$$\begin{aligned} &E[\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\Delta_1\boldsymbol{\theta}] \\ &= \frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_\alpha^{(kl)}\bar{W}_\alpha^{(mn)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \bar{\boldsymbol{M}}^-\bar{\boldsymbol{\xi}}_\alpha^{(n)}) \\ &\quad V_0^{(km)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}} + \frac{3\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)}, \\ &\quad \bar{\boldsymbol{M}}^-V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}. \end{aligned} \tag{52}$$

6.2 The second term

The second term of Eq. (44) is written as follows:

$$-E[\bar{\boldsymbol{M}}^-\Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}}] = -E[\bar{\boldsymbol{M}}^-\Delta_2^0\boldsymbol{M}\bar{\boldsymbol{\theta}}] - E[\bar{\boldsymbol{M}}^-\Delta_2^*\boldsymbol{M}\bar{\boldsymbol{\theta}}]. \tag{53}$$

The first term on the right side is written as

$$-E[\bar{\boldsymbol{M}}^{-}\Delta_2^0\boldsymbol{M}\bar{\boldsymbol{\theta}}] = -\bar{\boldsymbol{M}}^{-}\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_{\alpha}^{(kl)}$$
$$\Big(E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(k)}\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)\top}] + \bar{\boldsymbol{\xi}}_{\alpha}^{(k)}E[\Delta_2\boldsymbol{\xi}_{\alpha}^{(l)\top}]\Big)\bar{\boldsymbol{\theta}}$$
$$= -\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_{\alpha}^{(kl)}\Big(V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}} + (\boldsymbol{e}_{\alpha}^{(k)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(l)}\Big), \tag{54}$$

where Eq. (46) is used. The second term on the right side of Eq. (53) is written as

$$-E[\bar{\boldsymbol{M}}^{-}\Delta_2^*\boldsymbol{M}\bar{\boldsymbol{\theta}}]$$
$$= -E[\bar{\boldsymbol{M}}^{-}\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\Delta_1 W_{\alpha}^{(kl)}(\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(k)}]$$
$$= -2\bar{\boldsymbol{M}}^{-}\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\theta}},$$
$$E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)}(\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}})^{\top}]\bar{\boldsymbol{M}}^{-}V_0^{(mn)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(k)}. \tag{55}$$

The expression $E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)}(\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}})^{\top}]$ in the above equation is evaluated as follows:

$$E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)}(\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}})^{\top}] = E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)}\Big(\frac{1}{N}\sum_{\beta=1}^{N}\sum_{p,q=1}^{3}\bar{W}_{\beta}^{(pq)}$$
$$\Big(\Delta_1\boldsymbol{\xi}_{\beta}^{(p)}\bar{\boldsymbol{\xi}}_{\beta}^{(q)\top} + \bar{\boldsymbol{\xi}}_{\beta}^{(p)}\Delta_1\boldsymbol{\xi}_{\beta}^{(q)\top}\Big)\bar{\boldsymbol{\theta}}\Big)^{\top}]$$
$$= \frac{1}{N}\sum_{\beta=1}^{N}\sum_{p,q=1}^{3}\bar{W}_{\beta}^{(pq)}E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(l)}\Delta_1\boldsymbol{\xi}_{\beta}^{(q)\top}]\bar{\boldsymbol{\theta}}\bar{\boldsymbol{\xi}}_{\beta}^{(p)\top}$$
$$= \frac{\sigma^2}{N}\sum_{p,q=1}^{3}\bar{W}_{\alpha}^{(pq)}V_0^{(lq)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}\bar{\boldsymbol{\xi}}_{\alpha}^{(p)\top}. \tag{56}$$

Hence, Eq. (55) has the following form:

$$-E[\bar{\boldsymbol{M}}^{-}\Delta_2^*\boldsymbol{M}\bar{\boldsymbol{\theta}}]$$
$$= -\frac{2\sigma^2}{N^2}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n,p,q=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}\bar{W}_{\alpha}^{(pq)}(\bar{\boldsymbol{\theta}},$$
$$V_0^{(lq)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})(\bar{\boldsymbol{\xi}}_{\alpha}^{(p)},\bar{\boldsymbol{M}}^{-}V_0^{(mn)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(k)}$$
$$= -\frac{2\sigma^2}{N^2}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(kl)}\bar{W}_{\alpha}^{(mn)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(k)},$$
$$\bar{\boldsymbol{M}}^{-}V_0^{(lm)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(n)}. \tag{57}$$

In the above derivation, we have used the identity

$$\sum_{m,n=1}^{3}\bar{W}_{\alpha}^{(km)}(\bar{\boldsymbol{\theta}},V_0^{(mn)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})\bar{W}_{\alpha}^{(nl)} = \bar{W}_{\alpha}^{(kl)}, \tag{58}$$

which is a consequence of the identity $\bar{\boldsymbol{W}}_{\alpha}\bar{\boldsymbol{W}}_{\alpha}^{-}\bar{\boldsymbol{W}}_{\alpha} = \bar{\boldsymbol{W}}_{\alpha}$, where $\bar{\boldsymbol{W}}_{\alpha}$ is the matrix whose (kl) element is $\bar{W}_{\alpha}^{(kl)}$. Note that the (kl) element of the pseudoinverse of the matrix $\bar{\boldsymbol{W}}_{\alpha}^{-}$ is $(\bar{\boldsymbol{\theta}},V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})$ from the definition of $\bar{W}_{\alpha}^{(kl)}$ in Eq. (25). Thus, Eq. (53) can be written as follows:

$$-E[\bar{\boldsymbol{M}}^{-}\Delta_2\boldsymbol{M}\bar{\boldsymbol{\theta}}]$$
$$= -\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_{\alpha}^{(kl)}\Big(V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}} + (\boldsymbol{e}_{\alpha}^{(k)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(l)}\Big)$$
$$-\frac{2\sigma^2}{N^2}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(kl)}\bar{W}_{\alpha}^{(mn)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(k)},$$
$$\bar{\boldsymbol{M}}^{-}V_0^{(lm)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_{\alpha}^{(n)}. \tag{59}$$

6.3 The third term

The third term of Eq. (44) is rewritten as

$$E[\bar{\boldsymbol{M}}^{-}\Delta_2\boldsymbol{L}\bar{\boldsymbol{\theta}}]$$
$$= E\Big[\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(m)},\Delta_1\boldsymbol{\theta})(\bar{\boldsymbol{\xi}}_{\alpha}^{(n)},$$
$$\Delta_1\boldsymbol{\theta})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}\Big]$$
$$+E\Big[\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(m)},\Delta_1\boldsymbol{\theta})$$
$$(\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)},\bar{\boldsymbol{\theta}})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}\Big]$$
$$+E\Big[\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\Delta_1\boldsymbol{\xi}_{\alpha}^{(m)},\bar{\boldsymbol{\theta}})$$
$$(\bar{\boldsymbol{\xi}}_{\alpha}^{(n)},\Delta_1\boldsymbol{\theta})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}\Big]$$
$$+E\Big[\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\Delta_1\boldsymbol{\xi}_{\alpha}^{(m)},\bar{\boldsymbol{\theta}})$$
$$(\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)},\bar{\boldsymbol{\theta}})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}\Big]$$
$$= \frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(m)},\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_{\alpha}^{(n)})$$
$$V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}$$
$$+\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(m)},$$
$$E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)\top}]\bar{\boldsymbol{\theta}})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}$$
$$+\frac{1}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\theta}},$$
$$E[\Delta_1\boldsymbol{\xi}_{\alpha}^{(m)}\Delta_1\boldsymbol{\theta}^{\top}]\bar{\boldsymbol{\xi}}_{\alpha}^{(n)})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}$$
$$+\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}^{(kl)}V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}, \tag{60}$$

where we have used Eqs. (51) and (58). The expression $E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)\top}]$ in the above equation can be evaluated as follows:

$$E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)\top}] = -E[\bar{\boldsymbol{M}}^{-}\Delta_1^0\boldsymbol{M}\bar{\boldsymbol{\theta}}\Delta_1\boldsymbol{\xi}_{\alpha}^{(n)\top}]$$
$$= -\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^{-}\sum_{p,q=1}^{3}\bar{W}_{\alpha}^{(pq)}\bar{\boldsymbol{\xi}}_{\alpha}^{(p)}\bar{\boldsymbol{\theta}}^{\top}V_0^{(qn)}[\boldsymbol{\xi}_{\alpha}]. \tag{61}$$

Hence, Eq. (60) has the following form:

$$E[\bar{\boldsymbol{M}}^{-}\Delta_2\boldsymbol{L}\bar{\boldsymbol{\theta}}]$$
$$= -\frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l,m,n=1}^{3}\bar{W}_{\alpha}^{(km)}\bar{W}_{\alpha}^{(ln)}(\bar{\boldsymbol{\xi}}_{\alpha}^{(m)},$$
$$\bar{\boldsymbol{M}}^{-}\bar{\boldsymbol{\xi}}_{\alpha}^{(n)})V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}$$
$$+\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^{-}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}^{(kl)}V_0^{(kl)}[\boldsymbol{\xi}_{\alpha}]\bar{\boldsymbol{\theta}}. \tag{62}$$

6.4 Second order bias

From the above results, we conclude that Eq. (44) has the following form:

$$E[\Delta_2^{\perp}\boldsymbol{\theta}] = -\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^- \sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(kl)}(\boldsymbol{e}_\alpha^{(k)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(l)}$$
$$+\frac{\sigma^2}{N^2}\bar{\boldsymbol{M}}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\bar{\boldsymbol{\xi}}_\alpha^{(l)},\bar{\boldsymbol{M}}^-V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}. \quad (63)$$

7. Hyperaccurate Correction

In order to correct the solution $\boldsymbol{\theta}$ by Eq. (63), we need to estimate the unknown σ^2. We also need to approximate the noiseless values used in Eq. (63) by observed values. The resulting procedure is as follows:

(1) Using the ML solution $\boldsymbol{\theta}$ and the matrix $\boldsymbol{M}$ computed from it, estimate σ^2 by

$$\hat{\sigma}^2 = \frac{(\boldsymbol{\theta},\boldsymbol{M}\boldsymbol{\theta})}{r-(n-1)/N}, \quad (64)$$

where n is the dimension of $\boldsymbol{\theta}$.

(2) Compute the correction term by

$$\Delta_c\boldsymbol{\theta} = -\frac{\hat{\sigma}^2}{N}\boldsymbol{M}_{n-1}^- \sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}W_\alpha^{(kl)}(\boldsymbol{e}_\alpha^{(k)},\boldsymbol{\theta})\boldsymbol{\xi}_\alpha^{(l)}$$
$$+\frac{\hat{\sigma}^2}{N^2}\boldsymbol{M}_{n-1}^-\sum_{\alpha=1}^{N}\sum_{k,l=1}^{3}W_\alpha^{(km)}W_\alpha^{(ln)}(\boldsymbol{\xi}_\alpha^{(l)},$$
$$\boldsymbol{M}_{n-1}^-V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\boldsymbol{\theta})\boldsymbol{\xi}_\alpha^{(k)}, \quad (65)$$

where $\boldsymbol{M}_{n-1}^-$ is the pseudoinverse of $\boldsymbol{M}$ with truncated rank $n-1$.

(3) Correct the ML solution $\boldsymbol{\theta}$ to

$$\boldsymbol{\theta} \leftarrow \mathcal{N}[\boldsymbol{\theta}-\Delta_c\boldsymbol{\theta}], \quad (66)$$

where $\mathcal{N}[\,\cdot\,]$ designates normalization to unit norm ($\mathcal{N}[\boldsymbol{a}] \equiv \boldsymbol{a}/\|\boldsymbol{a}\|$).

The estimation formula of Eq. (64) is obtained by noting that if the minimum value of Eq. (22) is $\hat{J}$, then $N\hat{J}/\sigma^2$ is subject to a χ^2 distribution with $Nr-(n-1)$ degrees of freedom [6] and that the expectation of a χ^2 variable is equal to its degrees of freedom. Replacing the true values by their observations introduces errors of $O(\sigma)$, but since Eq. (65) is $O(\sigma^2)$ and the expectation of odd order noise terms is zero, the resulting error of Eq. (65) is $O(\sigma^4)$. Hence, the bias of the corrected $\boldsymbol{\theta}$ is still $\boldsymbol{0}$ except $O(\sigma^4)$.

Note that Eq. (65) is an analytical expression, so it can be immediately evaluated without any iterations. Of course, the truncated pseudoinverse $\boldsymbol{M}_{n-1}^-$ need to be evaluated, but this is no significant cost (recall that $n=6$ for ellipse fitting and $n=9$ for fundamental matrix and homography computation). Here, we are assuming that the ML solution $\boldsymbol{\theta}$ is already computed by some means. For this, any available method can be used, but if the FNS of Chojnacki et al. [2] or its extension [11] is used, all the quantities that appear in Eq. (65), i.e., $\boldsymbol{\xi}_\alpha^{(k)}$, $V_0^{(kl)}[\boldsymbol{\xi}_\alpha]$, $\boldsymbol{M}$, and $W_\alpha^{(kl)}$, are *already evaluated in the course of the FNS computation*. Hence, there is practically no additional computational cost for evaluating Eq. (65).

From Eq. (17), we see that the vector $\boldsymbol{e}^{(k)}$ in Eq. (65) is

$$\boldsymbol{e} = (1,0,1,0,0,0)^\top \quad (67)$$

for ellipse fitting. However, we see from Eqs. (19) and (21) that $\boldsymbol{e}^{(k)} = \boldsymbol{0}$ for the fundamental matrix and homography computation. In general, $\boldsymbol{e}^{(k)}$ is $\boldsymbol{0}$ for typical "multiview" constraints for computer vision, because noise in different images is assumed to be uncorrelated.

In the past study [7], [8], the terms that involve $\boldsymbol{e}_\alpha^{(k)}$ are ignored. We now show by simulation that omission of $\boldsymbol{e}_\alpha^{(k)}$ does not effectively affect the results.

8. Experiments

8.1 Evaluation of accuracy

Since the computed $\boldsymbol{\theta}$ and its true value $\bar{\boldsymbol{\theta}}$ are both unit vectors, we measure the discrepancy $\Delta\boldsymbol{\theta}$ between them by the orthogonal component to $\bar{\boldsymbol{\theta}}$ (Fig. 2),

$$\Delta^{\perp}\boldsymbol{\theta} = \boldsymbol{P}_{\bar{\boldsymbol{\theta}}}\boldsymbol{\theta}, \qquad \boldsymbol{P}_{\bar{\boldsymbol{\theta}}} \equiv \boldsymbol{I}-\bar{\boldsymbol{\theta}}\bar{\boldsymbol{\theta}}^\top, \quad (68)$$

where $\boldsymbol{P}_{\bar{\boldsymbol{\theta}}}$ is the projection matrix along $\bar{\boldsymbol{\theta}}$. We generate M independent noise instances and evaluate the bias B and the RMS (root-mean-square) error D defined by

$$B = \Big\|\frac{1}{M}\sum_{a=1}^{M}\Delta^{\perp}\boldsymbol{\theta}^{(a)}\Big\|, \quad (69)$$

$$D = \sqrt{\frac{1}{M}\sum_{a=1}^{M}\|\Delta^{\perp}\boldsymbol{\theta}^{(a)}\|^2}, \quad (70)$$

where $\boldsymbol{\theta}^{(a)}$ is the solution in the ath trial. The KCR lower bound (see Eq. (29)) on D is given by

$$D \geq \frac{\sigma}{\sqrt{N}}\sqrt{\mathrm{tr}\bar{\boldsymbol{M}}^-}, \quad (71)$$

where tr denotes the matrix trace. For comparison, we tested the following eight methods:

(1) Least squares (LS) [8].
(2) Iterative reweight [8].
(3) The Taubin method [24] and its extension to multiple constraints (we omit the details).
(4) Renormalization [5], [6].
(5) HyperLS [13], [14], [22].
(6) Hyper-renormalization [10] and its extension to multiple constraints (we omit the details).
(7) ML [15].
(8) ML with hyperaccurate correction.

8.2 Ellipse fitting

We define 30 equidistant points on the ellipse shown in Fig. 3. The major and minor axis are set to 100 and 50 pixels, respectively. We add independent Gaussian noise of mean 0 and standard deviation σ (pixels) to the x and y coordinates of each point and fit an ellipse.

Figures 4(a), (b) plot the bias B and the RMS error D, respectively, defined in (69) and (70) over 10000 independent trials for

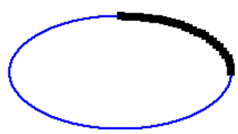

Fig. 3 Thirty points on an ellipse.

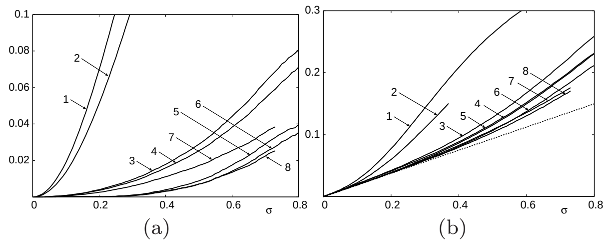

Fig. 4 The bias (a) and the RMS error (b) of the fitted ellipse for the standard deviation σ of the noise added to the data in Fig. 3 over 10000 independent trials. 1) LS, 2) iterative reweight, 3) Taubin, 4) renormalization, 5) HyperLS, 6) hyper-renormalization, 7) ML, 8) ML with hyperaccurate correction. The dotted line in (b) indicates the KCR lower bound. The interrupted plots indicate that iterations do not always converge beyond that noise level.

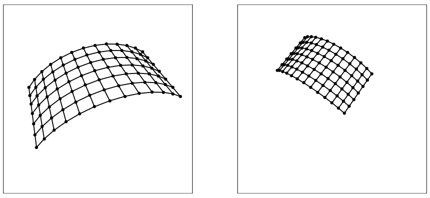

Fig. 5 Simulated images of a curved grid surface viewed from two directions.

each σ. The dotted line in Fig. 4(b) is the KCR lower bound of (71). We can see that LS and iterative reweight have very large bias and RMS error, while hyper-renormalization and ML with hyperaccurate correction both have very small bias and small RMS error. We can also see that although the difference is very small, the bias and the RMS error of ML with hypercorrection are even smaller than those of hyper-renormalization. However, as the interrupted plots in Fig. 4 show, the iterations of ML computation (we used the FNS of Chojnacki et al. [2]) do not converge in the presence of large noise. We also compared our solution with and without using the $\boldsymbol{e}^{(k)}$ term in Eq. (65) and found that the plots in Fig. 4 are unchanged.

8.3 Fundamental matrix computation

Figure 5 shows simulated images of a curved grid surface viewed from two directions. The image size is 600×600 pixels, and the focal length is 600 pixels. We add Gaussian noise of mean 0 and standard deviation σ (pixels) to the x and y coordinates of each grid point independently and compute the fundamental matrix $\boldsymbol{F}$. The fundamental matrix $\boldsymbol{F}$ has rank 2, so it is constrained to be $\det \boldsymbol{F} = 0$ [3]. Basically, the following three approaches exist for imposing this rank constraint [17]:

(1) A posteriori correction: The matrix $\boldsymbol{F}$ is optimally computed without considering the rank constraint and then optimally corrected so that it is satisfied.

(2) Internal access: The matrix $\boldsymbol{F}$ is parameterized so that the rank constraint is identically satisfied and then optimized within the resulting smaller parameter space.

(3) External access: Iterations are done in the space of unconstrained $\boldsymbol{F}$ in such a way the rank constraint is automatically satisfied at the time of convergence.

Here, we adopt the a posteriori correction approach and compare the accuracy of various methods without considering the rank constraint.

Figures 6(a), (b) plot the bias B and the RMS error D, respectively, defined in (69) and (70) over 10000 independent trials for each σ. The dotted line in Fig. 6(b) is the KCR lower bound of (71). We can see that LS and iterative reweight have very large bias and RMS error and that other hyper-renormalization and ML with hyperaccurate correction have very small bias. However, the RMS error is almost the same for all methods other than LS and iterative reweight. Yet, a close examination shows that ML with hypercorrection exhibits the highest accuracy.

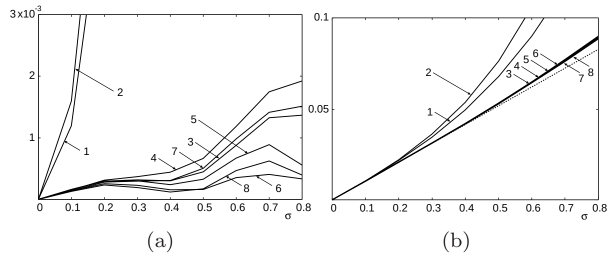

Fig. 6 The bias (a) and the RMS error (b) of the computed fundamental matrix for the standard deviation σ of the noise added to the data in Fig. 5 over 10000 independent trials. 1) LS, 2) iterative reweight, 3) Taubin, 4) renormalization, 5) HyperLS, 6) hyper-renormalization, 7) ML, 8) ML with hyperaccurate correction. The dotted line in (b) indicates the KCR lower bound.

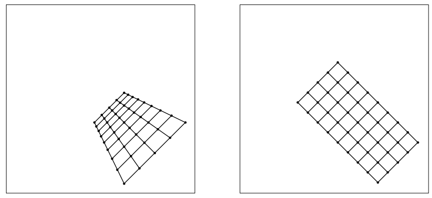

Fig. 7 Simulated images of a planar grid surface viewed from two directions.

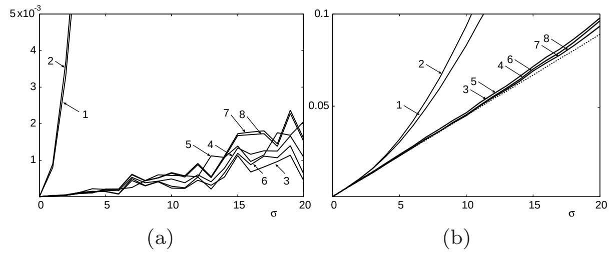

Fig. 8 The bias (a) and the RMS error (b) of the computed homography for the standard deviation σ of the noise added to the data in Fig. 7 over 10000 independent trials. 1) LS, 2) iterative reweight, 3) Taubin, 4) renormalization, 5) HyperLS, 6) hyper-renormalization, 7) ML, 8) ML with hyperaccurate correction. The dotted line in (b) indicates the KCR lower bound.

8.4 Homography computation

Figure 7 shows simulated images of a planar grid surface viewed from two directions. The image size is 800×800 pixels, and the focal length is 600 pixels. We add Gaussian noise of mean 0 and standard deviation σ (pixels) to the x and y coordinates of each grid point independently and compute the homography between the two images.

Figures 8(a), (b) plot the bias B and the RMS error D, respectively, defined in (69) and (70) over 10000 independent trials for each σ. The dotted line in Fig. 8(b) is the KCR lower bound of (71). As in the case of ellipse fitting and fundamental matrix computation, LS and iterative reweight have very large bias, resulting in large RMS error. However, the bias and RMS error of all other methods are almost the same, and the KCR lower bound is almost achieved. Yet, a close examination shows that ML with hypercorrection exhibits the highest accuracy.

9. Concluding Remarks

We reexamined the scheme of hyperaccurate correction for improving the accuracy of ML of geometric estimation based on geometric constraints. So far, this was done only in the case of a single scalar constraint. In this paper, we extended it to the case of multiple constraints given as a vector equation and pointed out the existence of a new correction term which was ignored in the past.

The correction is done by evaluating a single analytical expression without iterations with almost no additional computational cost. Moreover, if the FNS of Chojnacki et al. [2] or its extension [11] is used for computing the ML solution, all the quantities necessary for the correction are already evaluated in the course of the FNS computation, and hence practically no additional cost is required.

We compared our hyperaccurate correction with the hyper-renormalization of Kanatani et al. [10], the latest method regarded as the best fitting method, by do numerical simulation of ellipse fitting and fundamental matrix and homography computation. We observed the following:

(1) Inclusion of the new correction term does not effectively affect the final solution.
(2) The combination of ML and hyperaccurate correction can achieve the highest accuracy among all existing methods.
(3) For hyperaccurate correction, we first need to compute the ML solution, but the iterations for it do not necessarily converge in the presence of large noise.

We conclude that ML with hyperaccurate correction and hyper-renormalization [10], which achieves the next highest accuracy, are currently the best methods of all, each having its own advantage and disadvantage.

Acknowledgments. The authors thank Ali Al-Shadraqah of the University of Mississippi for helpful discussions. This work was supported in part by JSPS Grant-in-Aid for Challenging Exploratory Research (24650086).

References

[1] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, U.K. (2004).
[2] Chojnacki, W., Brooks, M.J., van den Hengel, A. and Gawley, D.: On the fitting of surfaces to data with covariances, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.22, No.11, pp.1294–1303 (2000).
[3] Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, U.K. (2004).
[4] Kahl, F. and Hartley, R.: Multiple-view geometry under the L_∞-norm, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.30, No.9, pp.1603–1617 (2008).
[5] Kanatani, K.: Renormalization for unbiased estimation, *Proc. 4th Int. Conf. Comput. Vis.*, Berlin, Germany, pp.599–606 (1993).
[6] Kanatani, K.: *Statistical Optimization for Geometric Computation: Theory and Practice* Elsevier, Amsterdam, The Netherlands (1996); reprinted, Dover, York, NY, U.S.A. (2005).
[7] Kanatani, K.: Ellipse fitting with hyperaccuracy, *IEICE Trans. Inf. & Syst.*, Vol.E89-D, No.10, pp.2653–2660 (2006).
[8] Kanatani, K.: Statistical optimization for geometric fitting: Theoretical accuracy bound and high order error analysis, *Int. J. Comput. Vis.*, Vol.80, No.2, pp.167–188 (2008).
[9] Kanatani, K.: Optimization techniques for geometric estimation: Beyond minimization, *Proc. Joint IAPR Int. Workshop Structural and Syntactic Pattern Recognition and Statistical Techniques in Pattern Recognition*, Hiroshima, Japan, pp.11–30 (2012).
[10] Kanatani, K., Al-Sharadqah, A., Chernov, N. and Sugaya, Y.: Renormalization returns: Hyper-renormalization and its applications, *Proc. 12th Euro. Conf. Comput. Vis.*, Firenze, Italy, Vol.3, pp.385–398 (2012).
[11] Kanatani, K. and Niitsuma, H.: Optimal two-view planar triangulation, *IPSJ Trans. Comput. Vis. Appl.*, Vol.3, pp.67–79 (2011).
[12] Kanatani, K. and Niitsuma, H.: Optimal computation of 3-D similarity: Gauss-Newton vs. Gauss-Helmert, *Comput. Stat. Data Anal.*, Vol.56, No.12 , pp.4470-4483 (2012).
[13] Kanatani, K. and Rangarajan, P.: Hyper least squares fitting of circles and ellipses, *Comput. Stat. Data Anal.*, Vol.55, No.6, pp.2197–2208 (2011).
[14] Kanatani, K., Rangarajan, P., Sugaya, Y. and Niitsuma, H.: HyperLS for parameter estimation in geometric fitting, *IPSJ Trans. Comput. Vis. Appl.*, Vol.3, pp.80–94 (2011).
[15] Kanatani, K. and Sugaya, Y.: Unified computation of strict maximum likelihood for geometric fitting, *J. Math. Imaging Vis.*, Vol.38, No.1, pp.1–13 (2010).
[16] Kanatani, K. and Sugaya, Y.: Compact algorithm for strictly ML ellipse fitting, *Proc. 19th Int. Conf. Patt. Recog.*, Tampa, FL, U.S.A. (2008).
[17] Kanatani, K. and Sugaya, Y.: Compact fundamental matrix computation, *IPSJ Trans. Comput. Vis. Appl.*, Vol.2, pp.59–70 (2010).
[18] Kanazawa, Y. and Kanatani, K.: Do we really have to consider covariance matrices for image feature points? *Proc. Int. Conf. Compute. Vis.*, Vancouver, Canada, Vol.2, pp.5586–591 (2001).
[19] Okatani, T. and Deguchi, K.: Toward a statistically optimal method for estimating geometric relations from noisy data: Cases of linear relations, *Proc. IEEE Conf. Comput. Vis. Patt. Recog.*, Madison, WI, U.S.A., Vol.1, pp.432–439 (2003).
[20] Okatani, T. and Deguchi, K.: On bias correction for geometric parameter estimation in computer vision, *Proc. IEEE Conf. Comput. Vis. Patt. Recog.*, Miami Beach, FL, U.S.A., pp.959–966 (2009).
[21] Okatani, T. and Deguchi, K.: Improving accuracy of geometric parameter estimation using projected score method, *Proc. Int. Conf. Comput. Vis.*, Kyoto, Japan, pp.1733–1740 (2009).
[22] Rangarajan, P. and Kanatani, K.: Improved algebraic methods for circle fitting, *Elec. J. Stat.*, Vol.3, pp.1075–1082 (2009).
[23] Sampson, P.D: Fitting conic sections to "very scattered" data: An iterative refinement of the Bookstein algorithm. *Comp. Graphics Image Process.*, Vol.18, No.1, pp.97–108 (1982).
[24] Taubin, G.: Estimation of planar curves, surfaces, and non-planar space curves defined by implicit equations with applications to edge and range image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.*, Vol.13, No.11, pp.1115–1138 (1991).

Appendix

A.1 Derivation of Eq. (36)

Let $\boldsymbol{W}_\alpha$ be the matrix whose (kl) element is $W_\alpha^{(kl)}$ (kl), and $\boldsymbol{V}_\alpha$ be the matrix whose (kl) element is $(\bar{\boldsymbol{\theta}}, V_0^{(kl)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})$. By the definition of $W_\alpha^{(kl)}$ (kl), we have $\boldsymbol{W}_\alpha = (\boldsymbol{V}_\alpha)_r^-$ and hence the identity $\boldsymbol{V}_\alpha\boldsymbol{W}_\alpha\boldsymbol{V}_\alpha = \boldsymbol{V}_\alpha$. Its expansion is

$$(\bar{\boldsymbol{V}}_\alpha + \Delta_1\boldsymbol{V}_\alpha + \Delta_2\boldsymbol{V}_\alpha + \cdots)(\bar{\boldsymbol{W}}_\alpha + \Delta_1\boldsymbol{W}_\alpha + \Delta_2\boldsymbol{W}_\alpha + \cdots)(\bar{\boldsymbol{V}}_\alpha + \Delta_1\boldsymbol{V}_\alpha + \Delta_2\boldsymbol{V}_\alpha + \cdots) = (\bar{\boldsymbol{V}}_\alpha + \Delta_1\boldsymbol{V}_\alpha + \Delta_2\boldsymbol{V}_\alpha + \cdots). \tag{A.1}$$

We derive Eq. (36) by equating the terms of the same order on both sides and using the identities $\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha = \bar{\boldsymbol{W}}_\alpha$ and $\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha = \bar{\boldsymbol{V}}_\alpha$. We also note that $\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha = \bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha$ is the projection matrix onto the common domain of $\bar{\boldsymbol{V}}_\alpha$ and $\bar{\boldsymbol{W}}_\alpha$ and that the errors $\Delta_1\boldsymbol{V}_\alpha$ and $\Delta_1\boldsymbol{W}_\alpha$ arise within that domain. Equating the first order error terms, we obtain

$$\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha + \bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha + \bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha = \Delta_1\boldsymbol{V}_\alpha. \tag{A.2}$$

Multiplying both sides by $\bar{\boldsymbol{W}}_\alpha$ from left and right, we obtain

$$\begin{aligned}&\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha+\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha=\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha.\end{aligned}\tag{A.3}$$

Hence, we have

$$\begin{aligned}&\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha+\Delta_1\boldsymbol{W}_\alpha+\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&=\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha,\end{aligned}\tag{A.4}$$

from which $\Delta_1\boldsymbol{W}_\alpha$ is expressed in the form

$$\Delta_1\boldsymbol{W}_\alpha=-\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha.\tag{A.5}$$

Its (kl) element is

$$\begin{aligned}\Delta_1W_\alpha^{(kl)}&=-\sum_{m,n=1}^{L}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{ln}\Delta_1V_\alpha^{(mn)}\\&=-2\sum_{m,n=1}^{L}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}(\Delta_1\boldsymbol{\theta},V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}}).\end{aligned}\tag{A.6}$$

Thus, we obtain the first of Eq. (36). Equating the second order error terms on both sides of Eq. (A.1), we obtain

$$\begin{aligned}&\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha+\bar{\boldsymbol{V}}_\alpha\Delta_2\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha+\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\\&+\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha\Delta_1\boldsymbol{V}_\alpha+\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\\&+\Delta_1\boldsymbol{V}_\alpha\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha=\Delta_2\boldsymbol{V}_\alpha.\end{aligned}\tag{A.7}$$

Multiplying both sides by $\bar{\boldsymbol{W}}_\alpha$ from left and right, we obtain

$$\begin{aligned}&\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha+\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_2\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha+\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha=\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha,\end{aligned}\tag{A.8}$$

which is rewritten as

$$\begin{aligned}&\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha+\Delta_2\boldsymbol{W}_\alpha+\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\Delta_1\boldsymbol{W}_\alpha(\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha)\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha(\bar{\boldsymbol{V}}_\alpha\bar{\boldsymbol{W}}_\alpha)\Delta_1\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&+\bar{\boldsymbol{W}}_\alpha\Delta_1\boldsymbol{V}_\alpha(\bar{\boldsymbol{W}}_\alpha\bar{\boldsymbol{V}}_\alpha)\Delta_1\boldsymbol{W}_\alpha=\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha.\end{aligned}\tag{A.9}$$

Substituting Eq. (A.5) into this, we obtain

$$\begin{aligned}&\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha+\Delta_2\boldsymbol{W}_\alpha+\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha\\&-\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha+\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha\\&-\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha=\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha,\end{aligned}\tag{A.10}$$

from which $\Delta_2\boldsymbol{W}_\alpha$ is expressed in the form

$$\Delta_2\boldsymbol{W}_\alpha=\Delta_1\boldsymbol{W}_\alpha\bar{\boldsymbol{V}}_\alpha\Delta_1\boldsymbol{W}_\alpha-\bar{\boldsymbol{W}}_\alpha\Delta_2\boldsymbol{V}_\alpha\bar{\boldsymbol{W}}_\alpha.\tag{A.11}$$

Its (kl) element is

$$\begin{aligned}\Delta_2W_\alpha^{(kl)}&=\sum_{m,n=1}^{L}\Delta_1W_\alpha^{(km)}\bar{V}_\alpha^{(mn)}\Delta_1W_\alpha^{(nl)}\\&\quad-\sum_{m,n=1}^{L}\bar{W}_\alpha^{(km)}\Delta_2V_\alpha^{(mn)}\bar{W}_\alpha^{(nl)}\\&=\sum_{m,n=1}^{L}\Delta_1W_\alpha^{(km)}\Delta_1W_\alpha^{(ln)}(\bar{\boldsymbol{\theta}},V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\\&\quad-\sum_{m,n=1}^{L}\bar{W}_\alpha^{(km)}\bar{W}_\alpha^{(ln)}\Big((\Delta_1\boldsymbol{\theta},V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\Delta_1\boldsymbol{\theta})\\&\quad+2(\Delta_2\boldsymbol{\theta},V_0^{(mn)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\Big).\end{aligned}\tag{A.12}$$

Thus, we obtain the second of Eq. (36).

A.2 Derivation of Eq. (51)

Substituting Eqs. (31) and into Eq. (42), and noting that $\boldsymbol{\xi}_\alpha^{(k)\top}\boldsymbol{\theta} = 0$, we can write $\Delta_1\boldsymbol{\theta}$ as follows:

$$\begin{aligned}\Delta_1\boldsymbol{\theta}&=-\bar{\boldsymbol{M}}^-\Delta_1\boldsymbol{M}\bar{\boldsymbol{\theta}}\\&=-\bar{\boldsymbol{M}}^-\Big(\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{L}\bar{W}_\alpha^{(kl)}(\Delta_1\boldsymbol{\xi}_\alpha^{(l)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\Big).\end{aligned}\tag{A.13}$$

We evaluate $E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top]$ by eliminating the noise terms $\boldsymbol{\xi}_\alpha^{(k)}$, using the identities of Eqs. (45) and (58).

$$\begin{aligned}&E[\Delta_1\boldsymbol{\theta}\Delta_1\boldsymbol{\theta}^\top]\\&=E[\bar{\boldsymbol{M}}^-\Big(\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,l=1}^{L}\bar{W}_\alpha^{(kl)}(\Delta_1\boldsymbol{\xi}_\alpha^{(l)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\Big)\\&\quad\frac{1}{N}\sum_{\beta=1}^{N}\sum_{m,n=1}^{L}\bar{W}_\beta^{(mn)}(\Delta_1\boldsymbol{\xi}_\beta^{(n)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\beta^{(m)\top}\Big)\bar{\boldsymbol{M}}^-]\\&=E[\bar{\boldsymbol{M}}^-\Big(\frac{1}{N^2}\sum_{\alpha,\beta=1}^{N}\sum_{k,l,m,n=1}^{L}\bar{W}_\alpha^{(kl)}\bar{W}_\beta^{(mn)}\\&\quad(\bar{\boldsymbol{\theta}},\Delta_1\boldsymbol{\xi}_\alpha^{(l)})(\Delta_1\boldsymbol{\xi}_\beta^{(n)},\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\beta^{(m)\top}\Big)\bar{\boldsymbol{M}}^-]\\&=\bar{\boldsymbol{M}}^-\Big(\frac{1}{N^2}\sum_{\alpha,\beta=1}^{N}\sum_{k,l,m,n=1}^{L}\bar{W}_\alpha^{(kl)}\bar{W}_\beta^{(mn)}\\&\quad(\bar{\boldsymbol{\theta}},E[\Delta_1\boldsymbol{\xi}_\alpha^{(l)}\Delta_1\boldsymbol{\xi}_\beta^{\top(n)}]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\beta^{(m)\top}\Big)\bar{\boldsymbol{M}}^-\\&=\bar{\boldsymbol{M}}^-\Big(\frac{1}{N^2}\sum_{\alpha,\beta=1}^{N}\sum_{k,l,m,n=1}^{L}\bar{W}_\alpha^{(kl)}\bar{W}_\beta^{(mn)}\\&\quad(\bar{\boldsymbol{\theta}},\sigma^2\delta_{\alpha\beta}V_0^{(ln)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\beta^{(m)\top}\Big)\bar{\boldsymbol{M}}^-\\&=\bar{\boldsymbol{M}}^-\Big(\frac{\sigma^2}{N^2}\sum_{\alpha=1}^{N}\sum_{k,m=1}^{L}\Big(\sum_{l,n=1}^{L}\bar{W}_\alpha^{(kl)}(\bar{\boldsymbol{\theta}},V_0^{(ln)}[\boldsymbol{\xi}_\alpha]\bar{\boldsymbol{\theta}})\\&\quad\bar{W}_\alpha^{(mn)}\Big)\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(m)\top}\Big)\bar{\boldsymbol{M}}^-\\&=\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^-\Big(\frac{1}{N}\sum_{\alpha=1}^{N}\sum_{k,m=1}^{L}\bar{W}_\alpha^{(km)}\bar{\boldsymbol{\xi}}_\alpha^{(k)}\bar{\boldsymbol{\xi}}_\alpha^{(m)\top}\Big)\bar{\boldsymbol{M}}^-\\&=\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^-\bar{\boldsymbol{M}}\bar{\boldsymbol{M}}^-=\frac{\sigma^2}{N}\bar{\boldsymbol{M}}^-.\end{aligned}\tag{A.14}$$

Kenichi Kanatani received his B.E., M.S., and Ph.D. in applied mathematics from the University of Tokyo in 1972, 1974 and 1979, respectively. After serving as Professor of computer science at Gunma University, Gunma, Japan, he is currently Professor of computer science at Okayama University, Okayama, Japan. He is the author of many books on computer vision and received many awards including the best paper awards from IPSJ (1987) and IEICE (2005). He is an IEEE Fellow.

Yasuyuki Sugaya received his B.E., M.S., and Ph.D. in computer science from the University of Tsukuba, Ibaraki, Japan, in 1996, 1998, and 2001, respectively. From 2001 to 2006, he was Assistant Professor of computer science at Okayama University, Okayama, Japan. Currently, he is Associate Professor of computer science and engineering at Toyohashi University of Technology, Toyohashi, Aichi, Japan. His research interests include image processing and computer vision. He received the IEICE best paper award in 2005.

(Communicated by *Chu-Song Chen*)