HYPERACCURATE ELLIPSE FITTING WITHOUT ITERATIONS

Kenichi Kanatani

Department of Computer Science, Okayama University, Okayama 700-8530, Japan

Prasanna Rangarajan

Department of Electrical Engineering, Southern Methodist University, Dallas, TX 75205, U.S.A.

Abstract: This paper presents a new method for fitting an ellipse to a point sequence extracted from images. It is widely known that the best fit is obtained by maximum likelihood. However, it requires iterations, which may not converge in the presence of large noise. Our approach is algebraic distance minimization; no iterations are required. Exploiting the fact that the solution depends on the way the scale is normalized, we analyze the accuracy to high order error terms with the scale normalization weight unspecified and determine it so that the bias is zero up to the second order. We demonstrate by experiments that our method is superior to the Taubin method, also algebraic and known to be highly accurate.

1 INTRODUCTION

Circular objects are projected onto camera images as ellipses, and from their 2-D shapes one can reconstruct their 3-D structure (Kanatani, 1993). For this reason, detecting ellipses in images and computing their mathematical representation are the first step of many computer vision applications including industrial robotic operations and autonomous navigation. This is done in two stages, although they are often intermingled. The first stage is to detect edges, test if a particular edge segment can be regarded as an elliptic arc, and integrate multiple arcs into ellipses (Kanatani and Ohta, 2004; Rosin Rosin and West, 1995). The second stage is to fit an equation to those edge points regarded as constituting an elliptic arc. In this paper, we concentrate on the latter.

Among many ellipse fitting algorithms presented in the past, those regarded as the most accurate are methods based on maximum likelihood (ML), and various computational schemes have been proposed including the *FNS* (*Fundamental Numerical Scheme* of Chojnacki et al. (2000), the *HEIV* (*Heteroscedastic Errors-in-Variable* of Leedan and Meer (2000) and Matei and Meer (2006), and the *projective Gauss-Newton iterations* of Kanatani and Sugaya (2007). Efforts have also been made to make the cost function more precise (Kanatani and Sugaya, 2008) and

add a posterior correction to the solution (Kanatani, 2006), but the solution of all ML-based methods already achieves the theoretical accuracy limit, called the *KCR lower bound* (Chernov and Lesort, 2004; Kanatani, 1996; Kanatani, 2008), up to high order error terms. Hence, there is practically no room for further accuracy improvement. However, all ML-based methods have one drawback: Iterations are required for nonlinear optimization, but they often fail to converge in the presence of large noise. Also, an appropriate initial guess must be provided. Therefore, accurate algebraic methods that do not require iterations are very much desired, even though the solution may not be strictly optimal.

The best known algebraic method is the least squares, also known as algebraic distance minimization or DLT (direct linear transformation) (Hartley and Zisserman, 2004), but all algebraic fitting methods have an inherent weakness: We need to impose a normalization to remove scale indeterminacy, yet the solution depends on the choice of the normalization. Al-Sharadqah and Chernov (2009) and Rangarajan and Kanatani (2009) exploited this freedom for fitting circles. Invoking the high order error analysis of Kanatani (Kanatani, 2008), they optimized the normalization so that the solution has the highest accuracy. In this paper, we apply their techniques to ellipse fitting. Doing numerical experiments, we demon-

strate that our method is superior to the method of Taubin (1991), also an algebraic method known to be very accurate (Kanatani, 2008; Kanatani and Sugaya, 2007).

2 ALGEBRAIC FITTING

An ellipse is represented by

$$Ax^2 + 2Bxy + Cy^2 + 2f_0(Dx + Ey) + f_0^2 F = 0, \quad (1)$$

where f_0 is a scale constant that has an order of x and y; without this, finite precision numerical computation would incur serious accuracy loss[1]. Our task is to compute the coefficients A, ..., F so that the ellipse of Eq. (1) passes through given points (x_α, y_α), $\alpha = 1$, ..., N, as closely as possible. The algebraic approach is to compute A, ..., F that minimize the algebraic distance

$$J = \frac{1}{N} \sum_{\alpha=1}^{N} \left(Ax_\alpha^2 + 2Bx_\alpha y_\alpha + Cy_\alpha^2 + 2f_0(Dx_\alpha + Ey_\alpha) + f_0^2 F \right)^2. \quad (2)$$

This is also known as the least squares, algebraic distance minimization, or the direct linear transformation (DLT). Evidently, Eq. (2) is minimized by $A = \cdots = F = 0$ if no scale nomalization is imposed. Frequently used normalizations include

$$F = 1, \quad (3)$$
$$A + D = 1, \quad (4)$$
$$A^2 + B^2 + C^2 + D^2 + E^2 + F^2 = 1, \quad (5)$$
$$A^2 + B^2 + C^2 + D^2 + E^2 = 1, \quad (6)$$
$$A^2 + 2B^2 + C^2 = 1, \quad (7)$$
$$AC - B^2 = 1. \quad (8)$$

Equation (3) reduces minimization of Eq. (2) to simultatneous linear equations (Albano, 1974; Cooper and Yalabik, 1976, Rosin, 1993). However, Eq. (1) with $F = 1$ cannot represent ellipses passing through the origin $(0,0)$. Equation (4) remedies this (Gander et al., 1994; Rosin, 1994). The most frequently used is[2] Eq. (5) (Paton, 1970), but some authors use Eq. (6) (Gnanadesikan, 1977). Equation (7) imposes invariace to coordinate transformations in the sense that the ellipse fitted after the coordinate system is transalated

---

[1]In our experiments, we set $f_0 = 600$, assuming images of one side less than 1000 pixels.

[2]Some authors write an ellipse as $Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$. The meaning of Eq. (5) differs for this form and for Eq. (1). In the following, we ignore such small differences; no significant consequence would result.

and rotated is the same as the originally fitted ellipse translated and rotated afterwards (Bookstein, 1979). Equation (8) prevents Eq. (1) from representing a parabola ($AC - B^2 = 0$) or a hyperbola ($AC - B^2 < 0$) (Fitzgibbon et al., 1999). Many other normalizations are conceivable, but the crucial fact is that the resulting solution depends on which normalization is imposed. The purpose of this paper is to find the "best" normalization. Write the 6-D vector of the unknown coefficients as

$$\theta = \begin{pmatrix} A & B & C & D & E & F \end{pmatrix}^\top, \quad (9)$$

and consider the class of normalizations written as

$$(\theta, \mathbf{N}\theta) = \text{constant}, \quad (10)$$

for some symmetric matrix $\mathbf{N}$, where and hereafter we denote the inner product of vectors $\mathbf{a}$ an $\mathbf{b}$ by $(\mathbf{a}, \mathbf{b})$. Equations (5), (6), and (7) can be written in this form with a positive definite or semidefinite $\mathbf{N}$, while for Eq. (8) $\mathbf{N}$ is nondefinite. In this paper, we allow nondefinite $\mathbf{N}$, so that the constant in Eq. (10) is not necessarily positive.

3 ALGEBRAIC SOLUTION

If the weight matrix $\mathbf{N}$ is given, the solution θ that minimizes Eq. (1) is immediately computed. Write

$$\xi = \begin{pmatrix} x^2 & 2xy & y^2 & 2f_0x & 2f_0y & f_0^2 \end{pmatrix}^\top. \quad (11)$$

Equation (1) is now written as

$$(\theta, \xi) = 0. \quad (12)$$

Let ξ_α be the value of ξ for (x_α, y_α). Our problem is to minimize

$$J = \frac{1}{N} \sum_{\alpha=1}^{N} (\theta, \xi_\alpha)^2 = \frac{1}{N} \sum_{\alpha=1}^{N} \theta^\top \xi_\alpha \xi_\alpha^\top \theta = (\theta, \mathbf{M}\theta), \quad (13)$$

subject to Eq. (10), where we define the 6×6 matrix $\mathbf{M}$ as follows:

$$\mathbf{M} = \frac{1}{N} \sum_{\alpha=1}^{N} \xi_\alpha \xi_\alpha^\top. \quad (14)$$

Equation (13) is a quadratic form in θ, so it is minimized subject to Eq. (10) by solving the generalized eigenvalue problem

$$\mathbf{M}\theta = \lambda \mathbf{N}\theta. \quad (15)$$

If Eq. (1) is exactly satisfied for all (x_α, y_α), i.e., $(\theta, \xi_\alpha) = 0$ for all α, Eq. (14) implies $\mathbf{M}\theta = \mathbf{0}$ and hence $\lambda = 0$. If the weight $\mathbf{N}$ is positive definite or semidefinite, the generalized eigenvalue λ is positive

in the presence of noise, so the solution is given by the generalized eigenvector θ for the smallest λ. Here, we allow $\mathbf{N}$ to be nondefinite, so λ may not be positive. In the following, we do error analysis of Eq. (15) by assuming that $\lambda \approx 0$, so the solution is given by the generalized eigenvector θ for the λ with the smallest absolute value[3]. Since the solution θ of Eq. (15) has scale indeterminacy, we hereafter adopt normalization into unit norm $\|\theta\| = 1$ rather than Eq. (10).

The resulting solution may not necessarily represent an ellipse; it may represent a parabola or hyperbola. This can be avoided by imposing Eq. (8) (Fitzgibbon et al., 1999), but here we do not exclude nonellipse solution and optimize $\mathbf{N}$ so that the resulting solution θ is as close to its true value θ as possible.

Least Squares. In the following, we call the popular method of using Eq. (5) the *least squares* for short. This is equivalent to letting $\mathbf{N}$ to be the unit matrix $\mathbf{I}$. In this case, Eq. (15) becomes an ordinary eigenvalue problem

$$\mathbf{M}\theta = \lambda\theta, \qquad (16)$$

and the solution is the unit eigenvector of $\mathbf{M}$ for the smallest eigenvalue.

Taubin Method. A well known algebraic method known to be very accurate is due to Taubin (1991), who used as $\mathbf{N}$

$$\mathbf{N}_{\mathrm{T}} = \frac{4}{N} \sum_{\alpha=1}^{N} \begin{pmatrix} x_\alpha^2 & x_\alpha y_\alpha & 0 & f_0 x_\alpha & 0 & 0 \\ x_\alpha y_\alpha & x_\alpha^2 + y_\alpha^2 & x_\alpha y_\alpha & f_0 y_\alpha & f_0 x_\alpha & 0 \\ 0 & x_\alpha y_\alpha & y_\alpha^2 & 0 & f_0 y_\alpha & 0 \\ f_0 x_\alpha & f_0 y_\alpha & 0 & f_0^2 & 0 & 0 \\ 0 & f_0 x_\alpha & f_0 y_\alpha & 0 & f_0^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \qquad (17)$$

The solution is given by the unit generalized eigenvector θ of Eq. (15) for the smallest generalized eigenvalue λ.

4 ERROR ANALYSIS

We regard each (x_α, y_α) as perturbed from its true position $(\bar{x}_\alpha, \bar{y}_\alpha)$ by $(\Delta x_\alpha, \Delta y_\alpha)$ and write

$$\xi_\alpha = \bar{\xi}_\alpha + \Delta_1 \xi_\alpha + \Delta_2 \xi_\alpha, \qquad (18)$$

where $\bar{\xi}_\alpha$ is the true value of ξ_α, and $\Delta_1 \xi_\alpha$, and $\Delta_2 \xi_\alpha$ are the noise terms of the first and the second order, respectively:

---

[3]In the presence of noise, $\mathbf{M}$ is positive definite, so $(\theta, \mathbf{M}\theta) > 0$. If $(\theta, \mathbf{N}\theta) < 0$ for the final solution, we have $\lambda < 0$. However, if we replace $\mathbf{N}$ by $-\mathbf{N}$, we obtain the same solution θ with $\lambda > 0$, so the sign of λ does not have a particular meaning.

$$\Delta_1 \xi_\alpha = \begin{pmatrix} 2\bar{x}_\alpha \Delta x_\alpha \\ 2\bar{x}_\alpha \Delta y_\alpha + 2\bar{y}_\alpha \Delta x_\alpha \\ 2\bar{y}_\alpha \Delta y_\alpha \\ 2 f_0 \Delta x_\alpha \\ 2 f_0 \Delta y_\alpha \\ 0 \end{pmatrix}, \quad \Delta_2 \xi_\alpha = \begin{pmatrix} \Delta x_\alpha^2 \\ 2\Delta x_\alpha \Delta y_\alpha \\ \Delta y_\alpha^2 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (19)$$

The second term $\Delta_2 \xi_\alpha$ is ellipse specific and was not considered in the general theory of Kanatani (2008). We define the covariance matrix of ξ_α by $V[\xi_\alpha] = E[\Delta_1 \xi_\alpha \Delta_1 \xi_\alpha^\top]$, where $E[\cdot]$ denotes expectation. If the noise terms Δx_α and Δy_α are regarded as independent random Gaussian variables of mean 0 and standard deviation σ, we obtain

$$V[\xi_\alpha] = E[\Delta_1 \xi_\alpha \Delta_1 \xi_\alpha^\top] = \sigma^2 V_0[\xi_\alpha], \qquad (20)$$

where we put

$$V_0[\xi_\alpha] = 4 \begin{pmatrix} \bar{x}_\alpha^2 & \bar{x}_\alpha \bar{y}_\alpha & 0 & f_0 \bar{x}_\alpha & 0 & 0 \\ \bar{x}_\alpha \bar{y}_\alpha & \bar{x}_\alpha^2 + \bar{y}_\alpha^2 & \bar{x}_\alpha \bar{y}_\alpha & f_0 \bar{y}_\alpha & f_0 \bar{x}_\alpha & 0 \\ 0 & \bar{x}_\alpha \bar{y}_\alpha & \bar{y}_\alpha^2 & 0 & f_0 \bar{y}_\alpha & 0 \\ f_0 \bar{x}_\alpha & f_0 \bar{y}_\alpha & 0 & f_0^2 & 0 & 0 \\ 0 & f_0 \bar{x}_\alpha & f_0 \bar{y}_\alpha & 0 & f_0^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \qquad (21)$$

Here, we have noted that $E[\Delta x_\alpha] = E[\Delta y_\alpha] = 0$, $E[\Delta x_\alpha^2] = E[\Delta y_\alpha^2] = \sigma^2$, and $E[\Delta x_\alpha \Delta y_\alpha] = 0$ according to our assumption. We call the above $V_0[\xi_\alpha]$ the *normalized covariance matrix*. Comparing Eqs. (21) and (17), we find that the Taubin method uses as $\mathbf{N}$

$$\mathbf{N}_{\mathrm{T}} = \frac{1}{N} \sum_{\alpha=1}^{N} V_0[\xi_\alpha], \qquad (22)$$

after the observations (x_α, y_α) are plugged into $(\bar{x}_\alpha, \bar{y}_\alpha)$.

5 PERTURBATION ANALYSIS

Substituting Eq. (18) into Eq. (14), we have

$$\mathbf{M} = \frac{1}{N} \sum_{\alpha=1}^{N} (\bar{\xi}_\alpha + \Delta_1 \xi_\alpha + \Delta_2 \xi_\alpha)(\bar{\xi}_\alpha + \Delta_1 \xi_\alpha + \Delta_2 \xi_\alpha)^\top$$
$$= \bar{\mathbf{M}} + \Delta_1 \mathbf{M} + \Delta_2 \mathbf{M} + \cdots, \qquad (23)$$

where $\cdots$ denotes noise terms of order three and higher, and we define $\bar{\mathbf{M}}$, $\Delta_1 \mathbf{M}$, and $\Delta_2 \mathbf{M}$ by

$$\bar{\mathbf{M}} = \frac{1}{N} \sum_{\alpha=1}^{N} \bar{\xi}_\alpha \bar{\xi}_\alpha^\top, \qquad (24)$$

$$\Delta_1 \mathbf{M} = \frac{1}{N} \sum_{\alpha=1}^{N} \left(\bar{\xi}_\alpha \Delta_1 \xi_\alpha^\top + \Delta_1 \xi_\alpha \bar{\xi}_\alpha^\top \right), \qquad (25)$$

$$\Delta_2 \mathbf{M} = \frac{1}{N} \sum_{\alpha=1}^{N} \left(\bar{\xi}_\alpha \Delta_2 \xi_\alpha^\top + \Delta_1 \xi_\alpha \Delta_1 \xi_\alpha^\top + \Delta_2 \xi_\alpha \bar{\xi}_\alpha^\top \right). \tag{26}$$

We expand the solution θ and λ of Eq. (16) in the form

$$\theta = \bar{\theta} + \Delta_1 \theta + \Delta_2 \theta + \cdots, \quad \lambda = \bar{\lambda} + \Delta_1 \lambda + \Delta_2 \lambda + \cdots, \tag{27}$$

where the barred terms are the noise-free values, and symbols Δ_1 and Δ_2 indicate the first and the second order noise terms, respectively. Substituting Eqs. (23) and (27) into Eq. (15), we obtain

$$(\bar{\mathbf{M}} + \Delta_1 \mathbf{M} + \Delta_2 \mathbf{M} + \cdots)(\bar{\theta} + \Delta_1 \theta + \Delta_2 \theta + \cdots)$$
$$= (\bar{\lambda} + \Delta_1 \lambda + \Delta_2 \lambda + \cdots) \mathbf{N}(\bar{\theta} + \Delta_1 \theta + \Delta_2 \theta + \cdots). \tag{28}$$

Expanding both sides and equating terms of the same order, we obtain

$$\bar{\mathbf{M}} \bar{\theta} = \bar{\lambda} \mathbf{N} \bar{\theta}, \tag{29}$$

$$\bar{\mathbf{M}} \Delta_1 \theta + \Delta_1 \mathbf{M} \bar{\theta} = \bar{\lambda} \mathbf{N} \Delta_1 \theta + \Delta_1 \lambda \mathbf{N} \bar{\theta}, \tag{30}$$

$$\bar{\mathbf{M}} \Delta_2 \theta + \Delta_1 \mathbf{M} \Delta_1 \theta + \Delta_2 \mathbf{M} \bar{\theta}$$
$$= \bar{\lambda} \mathbf{N} \Delta_2 \theta + \Delta_1 \lambda \mathbf{N} \Delta_1 \theta + \Delta_2 \lambda \mathbf{N} \bar{\theta}. \tag{31}$$

The noise-free values $\bar{\xi}_\alpha$ and $\bar{\theta}$ satisfy $(\bar{\xi}_\alpha, \bar{\theta}) = 0$, so Eq. (24) implies $\bar{\mathbf{M}} \bar{\theta} = \mathbf{0}$, and Eq. (29) implies $\bar{\lambda} = 0$. From Eq. (25), we have $(\bar{\theta}, \Delta_1 \mathbf{M} \bar{\theta}) = 0$. Computing the inner product of $\bar{\theta}$ and Eq. (30), we find that $\Delta_1 \lambda = 0$. Multiplying Eq. (30) by the pseudoinverse $\bar{\mathbf{M}}^-$ from left, we have

$$\Delta_1 \theta = -\bar{\mathbf{M}}^- \Delta_1 \mathbf{M} \bar{\theta}, \tag{32}$$

where we have noted that $\bar{\theta}$ is the null vector of $\bar{\mathbf{M}}$ (i.e., $\bar{\mathbf{M}} \bar{\theta} = \mathbf{0}$) and hence $\bar{\mathbf{M}}^- \bar{\mathbf{M}} (\equiv \mathbf{P}_{\bar{\theta}})$ represents orthogonal projection along $\bar{\theta}$. We have also noted that equating the first order terms in the expansion of $\| \bar{\theta} + \Delta_1 \theta + \Delta_2 \theta + \cdots \|^2 = 1$ results in $(\bar{\theta}, \Delta_1 \theta) = 0$ so $\mathbf{P}_{\bar{\theta}} \Delta_1 \theta = \Delta_1 \theta$. Substituting Eq. (32) into Eq. (31), we can express $\Delta_2 \lambda$ in the form

$$\Delta_2 \lambda = \frac{(\bar{\theta}, \Delta_2 \mathbf{M} \bar{\theta}) - (\bar{\theta}, \Delta_1 \mathbf{M} \bar{\mathbf{M}}^- \Delta_1 \mathbf{M} \bar{\theta})}{(\bar{\theta}, \mathbf{N} \bar{\theta})} = \frac{(\bar{\theta}, \mathbf{T} \bar{\theta})}{(\bar{\theta}, \mathbf{N} \bar{\theta})}, \tag{33}$$

where

$$\mathbf{T} = \Delta_2 \mathbf{M} - \Delta_1 \mathbf{M} \bar{\mathbf{M}}^- \Delta_1 \mathbf{M}. \tag{34}$$

Next, we consider the second order error $\Delta_2 \theta$. Since $\bar{\theta}$ is a unit vector and does not change its norm, we are interested in the error component orthogonal to $\bar{\theta}$. We define the orthogonal component of $\Delta_2 \theta$ by

$$\Delta_2^\perp \theta = \mathbf{P}_{\bar{\theta}} \Delta_2 \theta \ (= \bar{\mathbf{M}}^- \bar{\mathbf{M}} \Delta_2 \theta). \tag{35}$$

Multiplying Eq. (31) by $\bar{\mathbf{M}}^-$ from left and substituting Eq. (32), we obtain

$$\Delta_2^\perp \theta = \Delta_2 \lambda \bar{\mathbf{M}}^- \mathbf{N} \bar{\theta} + \bar{\mathbf{M}}^- \Delta_1 \mathbf{M} \bar{\mathbf{M}}^- \Delta_1 \mathbf{M} \bar{\theta} - \bar{\mathbf{M}}^- \Delta_2 \mathbf{M} \bar{\theta}$$
$$= \frac{(\bar{\theta}, \mathbf{T} \bar{\theta})}{(\bar{\theta}, \mathbf{N} \bar{\theta})} \bar{\mathbf{M}}^- \mathbf{N} \bar{\theta} - \bar{\mathbf{M}}^- \mathbf{T} \bar{\theta}. \tag{36}$$

6 COVARIANCE AND BIAS

6.1 General Algebraic Fitting

From Eq. (32), we see that the leading term of the covariance matrix of the solution θ is given by

$$V[\theta] = E[\Delta_1 \theta \Delta_1 \theta^\top] = \bar{\mathbf{M}}^- E[(\Delta_1 \mathbf{M} \theta)(\Delta_1 \mathbf{M} \theta)^\top] \bar{\mathbf{M}}^-$$
$$= \frac{1}{N^2} \bar{\mathbf{M}}^- E \Big[\sum_{\alpha=1}^{N} (\Delta \xi_\alpha, \theta) \bar{\xi}_\alpha \sum_{\beta=1}^{N} (\Delta \xi_\beta, \theta) \bar{\xi}_\beta^\top \Big] \bar{\mathbf{M}}^-$$
$$= \frac{1}{N^2} \bar{\mathbf{M}}^- \sum_{\alpha, \beta = 1}^{N} (\theta, E[\Delta \xi_\alpha \Delta \xi_\beta^\top] \theta) \bar{\xi}_\alpha \bar{\xi}_\beta^\top \bar{\mathbf{M}}^-$$
$$= \frac{\sigma^2}{N^2} \bar{\mathbf{M}}^- \Big(\sum_{\alpha=1}^{N} (\theta, V_0[\xi_\alpha] \theta) \bar{\xi}_\alpha \bar{\xi}_\alpha^\top \Big) \bar{\mathbf{M}}^-$$
$$= \frac{\sigma^2}{N} \bar{\mathbf{M}}^- \bar{\mathbf{M}}' \bar{\mathbf{M}}^-, \tag{37}$$

where we define

$$\bar{\mathbf{M}}' = \frac{1}{N} \sum_{\alpha=1}^{N} (\bar{\theta}, V_0[\xi_\alpha] \theta) \bar{\xi}_\alpha \bar{\xi}_\alpha^\top. \tag{38}$$

In the derivation of Eq. (37), we have noted that ξ_α is independent for different α and that $E[\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top] = \delta_{\alpha\beta} \sigma^2 V_0[\xi_\alpha]$, where $\delta_{\alpha\beta}$ is the Kronecker delta. The important observation is that the covarinace matrix $V[\theta]$ of the solution θ does not depend on the normalization weight $\mathbf{N}$. This implies that *all algebraic methods have the same the covariance matrix in the leading order*, so we are unable to reduce the covariance of θ by adjusting $\mathbf{N}$. Yet, the Taubin method is known to be far accurate than the least squares. We will show that this stems from the bias terms and that a superior method can result by reducing the bias.

Since $E[\Delta_1 \theta] = \mathbf{0}$, there is no bias in the first order: the leading bias is in the second order. In order to evaluate the second order bias $E[\Delta_2^\perp \theta]$, we evaluate the expectation of $\mathbf{T}$ in Eq. (34). We first consider the term $E[\Delta_2 \mathbf{M}]$. From Eq. (26), we see that

$$E[\Delta_2 \mathbf{M}] = \frac{1}{N} \sum_{\alpha=1}^{N} \Big(\bar{\xi}_\alpha E[\Delta_2 \xi_\alpha]^\top + E[\Delta_1 \xi_\alpha \Delta_1 \xi_\alpha^\top]$$
$$+ E[\Delta_2 \xi_\alpha] \bar{\xi}_\alpha^\top \Big)$$
$$= \frac{\sigma^2}{N} \sum_{\alpha=1}^{N} \Big(\bar{\xi}_\alpha \mathbf{e}_{13}^\top + V_0[\xi_\alpha] + \mathbf{e}_{13} \bar{\xi}_\alpha^\top \Big)$$
$$= \sigma^2 \Big(\mathbf{N}_T + 2 \mathcal{S}[\bar{\xi}_c \mathbf{e}_{13}^\top] \Big), \tag{39}$$

where we have noted the definition in Eq. (20) and used Eq. (22). The symbol $\mathcal{S}[\cdot]$ denotes symmetrization ($\mathcal{S}[\mathbf{A}] \equiv (\mathbf{A} + \mathbf{A}^\top)/2$), and the vectors $\bar{\xi}_c$ and $\mathbf{e}_{13}$ are defined by

$$\bar{\xi}_c = \frac{1}{N} \sum_{\alpha=1}^{N} \bar{\xi}_\alpha, \quad \mathbf{e}_{13} = \begin{pmatrix} 1\,0\,1\,0\,0\,0 \end{pmatrix}^\top. \quad (40)$$

We next consider the term $E[\Delta_1 \mathbf{M}\bar{\mathbf{M}}^- \Delta_1 \mathbf{M}]$. It has the form (see Appendix for the derivation)

$$E[\Delta_1 \mathbf{M}\bar{\mathbf{M}}^- \Delta_1 \mathbf{M}] = \frac{\sigma^2}{N^2} \sum_{\alpha=1}^{N} \Big(\text{tr}[\bar{\mathbf{M}}^- V_0[\xi_\alpha]] \bar{\xi}_\alpha \bar{\xi}_\alpha^\top$$

$$+ (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha] + 2\mathcal{S}[V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top] \Big), \quad (41)$$

where $\text{tr}[\cdot]$ denotes the trace. From Eqs. (39) and (41), the matrix $\mathbf{T}$ in Eq. (34) has the following expectation:

$$E[\mathbf{T}] = \sigma^2 \Big(\mathbf{N}_\text{T} + 2\mathcal{S}[\bar{\xi}_c \mathbf{e}_{13}^\top]$$

$$- \frac{1}{N^2} \sum_{\alpha=1}^{N} \Big(\text{tr}[\bar{\mathbf{M}}^- V_0[\xi_\alpha]] \bar{\xi}_\alpha \bar{\xi}_\alpha^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha]$$

$$+ 2\mathcal{S}[V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top] \Big) \Big). \quad (42)$$

Thus, the second order error $\Delta_2^\perp \theta$ in Eq. (36) has the following bias:

$$E[\Delta_2^\perp \theta] = \bar{\mathbf{M}}^- \Big(\frac{(\bar{\theta}, E[\mathbf{T}]\bar{\theta})}{(\bar{\theta}, \mathbf{N}\bar{\theta})} \mathbf{N}\bar{\theta} - E[\mathbf{T}]\bar{\theta} \Big). \quad (43)$$

6.2 Least Squares

Eq. (42) implies $(\bar{\xi}_c, \bar{\theta}) = \mathbf{0}$ and $(\bar{\xi}_\alpha, \bar{\theta}) = \mathbf{0}$. Hence, $E[\mathbf{T}]\bar{\theta}$ can be written as

$$E[\mathbf{T}]\bar{\theta} = \sigma^2 \Big(\mathbf{N}_\text{T}\bar{\theta} + (A+C)\bar{\xi}_c$$

$$- \frac{1}{N^2} \sum_{\alpha=1}^{N} \Big((\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha]\bar{\theta} + (\bar{\theta}, V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha)\bar{\xi}_\alpha \Big) \Big).$$

$$(44)$$

If we let $\mathbf{N} = \mathbf{I}$, we obtain the least squares fit. Its leading bias is

$$E[\Delta_2^\perp \theta] = \bar{\mathbf{M}}^- \Big((\bar{\theta}, E[\mathbf{T}]\bar{\theta})\bar{\theta} - E[\mathbf{T}]\bar{\theta} \Big)$$

$$= -\bar{\mathbf{M}}^- (\mathbf{I} - \bar{\theta}\bar{\theta}^\top) E[\mathbf{T}]\bar{\theta} = -\bar{\mathbf{M}}^- E[\mathbf{T}]\bar{\theta}, \quad (45)$$

where we have used the following equality:

$$\bar{\mathbf{M}}^- (\mathbf{I} - \bar{\theta}\bar{\theta}^\top) = \bar{\mathbf{M}}^- \mathbf{P}_{\bar{\theta}} = \bar{\mathbf{M}}^- \mathbf{M}\bar{\mathbf{M}}^- = \bar{\mathbf{M}}^-. \quad (46)$$

From Eqs. (44), and (45), the leading bias of the least square has the following form:

$$E[\Delta_2^\perp \theta] = -\sigma^2 \bar{\mathbf{M}}^- \Big(\mathbf{N}_\text{T}\bar{\theta} + (A+C)\bar{\xi}_c$$

$$- \frac{1}{N^2} \sum_{\alpha=1}^{N} \Big((\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha]\bar{\theta} + (\bar{\theta}, V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha)\bar{\xi}_\alpha \Big) \Big).$$

$$(47)$$

6.3 Taubin Method

Eq. (44) implies $(\bar{\xi}_c, \bar{\theta}) = \mathbf{0}$ and $(\bar{\xi}_\alpha, \bar{\theta}) = \mathbf{0}$. Hence, $(\bar{\theta}, E[\mathbf{T}]\bar{\theta})$ can be written as

$$(\bar{\theta}, E[\mathbf{T}]\bar{\theta})$$

$$= \sigma^2 \Big((\bar{\theta}, \mathbf{N}_\text{T}\bar{\theta}) - \frac{1}{N^2} \sum_{\alpha=1}^{N} (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha)(\bar{\theta}, V_0[\xi_\alpha]\bar{\theta}) \Big)$$

$$= \sigma^2 \Big((\bar{\theta}, \mathbf{N}_\text{T}\bar{\theta}) - \frac{1}{N^2} \sum_{\alpha=1}^{N} \text{tr}[\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top](\bar{\theta}, V_0[\xi_\alpha]\bar{\theta}) \Big)$$

$$= \sigma^2 \Big((\bar{\theta}, \mathbf{N}_\text{T}\bar{\theta}) - \frac{1}{N^2} \text{tr}[\bar{\mathbf{M}}^- \sum_{\alpha=1}^{N} (\bar{\theta}, V_0[\xi_\alpha]\bar{\theta})\bar{\xi}_\alpha \bar{\xi}_\alpha^\top] \Big)$$

$$= \sigma^2 (\bar{\theta}, \mathbf{N}_\text{T}\bar{\theta}) - \frac{\sigma^2}{N} \text{tr}[\bar{\mathbf{M}}^- \bar{\mathbf{M}}'], \quad (48)$$

where we have used Eq. (38). If we let $\mathbf{N} = \mathbf{N}_\text{T}$, we obtain the Taubin method. Thus, the leading bias of the Taubin fit has form

$$E[\Delta_2^\perp \theta] = -\sigma^2 \bar{\mathbf{M}}^- \Big(q\mathbf{N}_\text{T}\bar{\theta} + (A+C)\bar{\xi}_c$$

$$- \frac{1}{N^2} \sum_{\alpha=1}^{N} \Big((\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha]\bar{\theta} + (\bar{\theta}, V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha)\bar{\xi}_\alpha \Big) \Big),$$

$$(49)$$

where we put

$$q = \frac{1}{N} \frac{\text{tr}[\bar{\mathbf{M}}^- \bar{\mathbf{M}}']}{(\bar{\theta}, \mathbf{N}_\text{T}\bar{\theta})}. \quad (50)$$

Comparing Eqs. (49) and (47), we notice that the only difference is that $\mathbf{N}_\text{T}\bar{\theta}$ in Eq. (47) is replaced by $q\mathbf{N}_\text{T}\bar{\theta}$ in Eq. (49). We see from Eq. (50) that $q < 1$ when N is large. This can be regarded as one of the reasons of the high accuracy of the Taubin method, as already pointed out by Kanatani (2008).

6.4 Hyperaccurate Algebraic Fit

Now, we present our main contribution of this paper. Our proposal is to chose the weight $\mathbf{N}$ to be

$$\mathbf{N} = \mathbf{N}_\text{T} + 2\mathcal{S}[\bar{\xi}_c \mathbf{e}_{13}^\top] - \frac{1}{N^2} \sum_{\alpha=1}^{N} \Big(\text{tr}[\bar{\mathbf{M}}^- V_0[\xi_\alpha]] \bar{\xi}_\alpha \bar{\xi}_\alpha^\top$$

$$+ (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha) V_0[\xi_\alpha] + 2\mathcal{S}[V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top] \Big). \quad (51)$$

Then, we have $E[\mathbf{T}] = \sigma^2 \mathbf{N}$ from Eq. (42), and Eq. (43) becomes

$$E[\Delta_2^\perp \theta] = \sigma^2 \bar{\mathbf{M}}^- \Big(\frac{(\bar{\theta}, \mathbf{N}\bar{\theta})}{(\bar{\theta}, \mathbf{N}\bar{\theta})} \mathbf{N} - \mathbf{N} \Big)\bar{\theta} = \mathbf{0}. \quad (52)$$

Since Eq. (51) contains the true values $\bar{\xi}_\alpha$ and $\bar{\mathbf{M}}$, we evaluate them by replacing the true values $(\bar{x}_\alpha, \bar{y}_\alpha)$ in
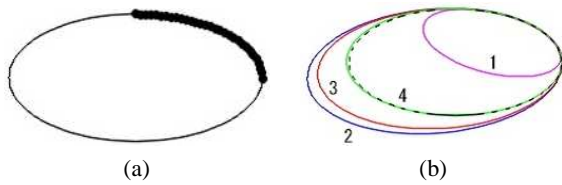
Figure 1: (a) 31 points on an ellipse. (b) Instances of fitted ellipses for $\sigma = 0.5$. 1. Least squares. 2. Taubin method. 3. Proposed method. 4. Maximum likelihood. The true shape is indicated in dashed lines.
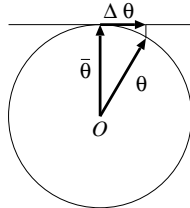

Figure 2: The true value $\bar{\theta}$, the computed value θ, and its orthogonal component $\Delta\theta$ to $\bar{\theta}$.

their definitions by the observations (x_α, y_α). This does not affect our result, because expectations of odd-order error terms vanish and hence the error in Eq. (52) is at most $O(\sigma^4)$. Thus, the second order bias is exactly 0. After the terminology used by Al-Sharadqah and Chernov (2009) for their circle fitting method, we call our method using Eq. (51) "hyperaccurate algebraic fitting".

7 NUMERICAL EXPERIMENTS

We placed 31 equidistant points in the first quadrant of the ellipse shown in Fig. 1(a). The major and the minor axis are 100 and 50 pixel long, respectively. We added to the x- and y-coordinates of each point independent Gaussian noise of mean 0 and standard deviation σ and fitted an ellipse by least squares, the Taubin method, our proposed method, and maximum likelihood[4]. Figure 1(b) shows fitted ellipses for some noise instance of $\sigma = 0.5$.

Since the computed and the true values θ and $\bar{\theta}$ are both unit vectors, we define their discrepancy $\Delta\theta$ by the orthogonal component

$$\Delta\theta = \mathbf{P}_{\bar{\theta}}\theta, \qquad (53)$$

where $\mathbf{P}_{\bar{\theta}}$ ($\equiv \mathbf{I} - \bar{\theta}\bar{\theta}$) is the orthogonal projection matrix along $\bar{\theta}$. (Fig. 2). Figures 3(a) and (b) plot for various σ the bias B and the RMS (root-mean-square)

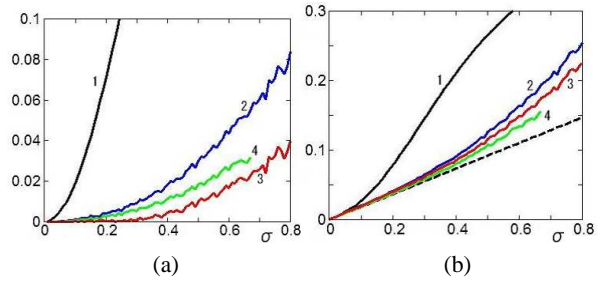[4]We used the FNS of Chojnacki et al. (2000). See Kanatani and Sugaya (2007) for the details.

Figure 3: The bias (a) and the RMS error (b) of the fitting to the data in Fig. 1(a). The horizontal axis is for the standard deviation σ of the added noise. 1. Least squares. 2. Taubin method. 3. Proposed method. 4. Maximum likelihood (interrupted due to nonconvergence). The dotted line in (b) shows the KCR lower bound.

error D defined by

$$B = \left\| \frac{1}{10000} \sum_{a=1}^{10000} \Delta\theta^{(a)} \right\|,$$

$$D = \sqrt{\frac{1}{10000} \sum_{a=1}^{10000} \|\Delta\theta^{(a)}\|^2}, \qquad (54)$$

where $\theta^{(a)}$ is the solution in the ath trial. The dotted line in Fig. 3(b) shows the KCR lower bound (Kanatani, 1996; Kanatani, 2008) given by

$$D_{\text{KCR}} = \sigma \sqrt{\text{tr}\left[\left(\sum_{\alpha=1}^N \frac{\bar{\xi}_\alpha \bar{\xi}_\alpha^\top}{(\bar{\theta}, V_0[\xi_\alpha]\bar{\theta})}\right)^-\right]}. \qquad (55)$$

Standard linear algebra routines for solving the generalized eigenvalue problem in the form of Eq. (15) assumes that the matrix $\mathbf{N}$ is positive definite. As can be seen from Eq. (17), however, the matrix $\mathbf{N}_T$ for the Taubin method is positive semidefinite having a row and a column of zeros. The matrix $\mathbf{N}$ in Eq. (51) is not positive definite, either. This causes no problem, because Eq. (15) can be written as

$$\mathbf{N}\theta = \frac{1}{\lambda}\mathbf{M}\theta. \qquad (56)$$

Since the matrix $\mathbf{M}$ in Eq. (14) is positive definite for noisy data, we can solve Eq. (56) instead of Eq. (15), using a standard routine. If the smallest eigenvalue of $\mathbf{M}$ happens to be 0, it indicates that the data are all exact; any method, e.g., LS, gives an exact solution. For noisy data, the solution θ is given by the generalized eigenvector of Eq. (56) for the generalized eigenvalue $1/\lambda$ with the largest absolute value.

As we can see from Fig. 3(a), the least square solution has a large bias, as compared to which the Taubin solution has a smaller bias, and our solution has even smaller bias. Since the least squares, the Taubin, and our solutions all have the same covariance matrix to the leading order, the bias is the decisive factor for their accuracy. This is demonstrated in

Figure 4: Left: Edge image containing a small elliptic edge segment (red). Right: Ellipses fitted to 155 edge points overlaid on the original image. From inner to outer are the ellipses computed by least squares (pink), Taubin method (blue), proposed method (red), and maximum likelihood (green). The proposed and Taubin method compute almost overlapping ellipses.

Fig. 3(b): The Taubin solution is more accurate than the least squares, and our solution is even more accurate.

On the other hand, the ML solution, which minimizes the Mahalanobis distance rather than the algebraic distance, has a larger bias than our solution, as shown in Fig. 3(a). Yet, since the covariance matrix of the ML solution is smaller than Eq. (38) (Kanatani, 2008), it achieves a higher accuracy than our solution, as shown in Fig. 3(b). However, the ML computation may not converge in the presence of large noise. Indeed, the interrupted plots of ML in Figs. 1(a) and (b) indicate that the iterations did not converge beyond that noise level. In contrast, our method, like the least squares and the Taubin method, is algebraic, so the computation can continue for however large noise.

The left of Fig. 4 is an edge image where a short elliptic arc (red) is visible. We fitted an ellipse to the 155 consecutive edge points on it by least squares, the Taubin method, our method, and ML. The right of Fig. 4 shows the resulting ellipses overlaid on the original image. We can see that the least squares solution is very poor, while the Taubin solution is close to the true shape. Our method and ML are slightly more accurate, but generally the difference is very small when the number of points is large and the noise is small as in this example.

8 CONCLUSIONS

We have presented a new algebraic method for fitting an ellipse to a point sequence extracted from images. The method known to be of the highest accuracy is maximum likelihood, but it requires iterations, which may not converge in the presence of lage noise. Also, an appropriate initial must be given. Our proposed method is algebraic and does not require iterations.

The basic principle is minimization of the alge-

braic distance. However, the solution depends on what kind of normalization is imposed. We exploited this freedom and derived a best normalization in such a way that the resulting solution has no bias up to the second order, invoking the high order error analysis of Kanatani (2008). Numerical experiments show that our method is superior to the Taubin method, also an algebraic method and known to be very accurate.

REFERENCES

Albano, R. (1974). Representation of digitized contours in terms of conics and straight-line segments, *Comput. Graphics Image Process.*, 3, 23–33.

Al-Sharadqah, A. and Chernov, N. (2009). Error analysis for circle fitting algorithms, *Elec. J. Stat.*, 3, 886-911.

Bookstein, F. J. (1979). Fitting conic sections to scattered data, *Comput. Graphics Image Process.*, 9, 56–71.

Chernov, N. and Lesort, C. (2004). Statistical efficiency of curve fitting algorithms, *Comput. Stat. Data Anal.*, 47, 713–728.

Chojnacki, W., Brooks, M. J., van den Hengel, A. and D. Gawley, D. (2000). On the fitting of surfaces to data with covariances, *IEEE Trans. Patt. Anal. Mach. Intell.*, 22, 1294–1303.

Cooper, D. B. and Yalabik, N. (1976). On the computational cost of approximating and recognizing noise-perturbed straight lines and quadratic arcs in the plane, *IEEE Trans. Computers*, 25, 1020–1032.

Fitzgibbon, A., Pilu, M. and Fisher, R. B. (1999). Direct least square fitting of ellipses, *IEEE Trans. Patt. Anal. Mach. Intell.*, 21, 476–480.

Gander, W., Golub, H. and Strebel, R. (1995). Least-squares fitting of circles and ellipses, *BIT*, 34, 558–578.

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariable Observations* (2nd ed.), Hoboken, NJ: Wiley.

Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision* (2nd ed.), Cambridge: Cambridge University Press.

Kanatani, K. (2006). Ellipse fitting with hyperaccuracy, *IEICE Trans. Inf. & Syst.*, E89-D, 2653–2660.

Kanatani, K. (1993). *Geometric Computation for Machine Vision*, Oxford: Oxford University Press.

Kanatani, K. (1996). *Statistical Optimization for Geometric Computation: Theory and Practice*, Amsterdam: Elsevier. Reprinted (2005) New York: Dover.

Kanatani, K. (2008). Statistical optimization for geometric fitting: Theoretical accuracy analysis and high order error analysis, *Int. J. Comp. Vis.* 80, 167–188.

Kanatani, K. and Ohta, N. (2004). Automatic detection of circular objects by ellipse growing, *Int. J. Image Graphics*, 4, 35–50.

Kanatani, K. and Sugaya, Y. (2008). Compact algorithm for strictly ML ellipse fitting, *Proc. 19th Int. Conf. Pattern Recognition*, Tampa, FL.

Kanatani, K. and Sugaya, Y. (2007). Performance evaluation of iterative geometric fitting algorithms, *Comp. Stat. Data Anal.*, 52, 1208–1222.

Leedan, Y. and Meer, P. (2000). Heteroscedastic regression in computer vision: Problems with bilinear constraint, *Int. J. Comput. Vision.*, 37, 127–150.

Matei, B. C. and Meer, P. (2006). Estimation of nonlinear errors-in-variables models for computer vision applications, *IEEE Trans. Patt. Anal. Mach. Intell.*, 28, 1537–1552.

Paton, K. A. (1970). Conic sections in chromosome analysis, *Patt. Recog.*, 2, 39–40.

Rangarajan, P. and Kanatani, K. (2009). Improved algebraic methods for circle fitting, *Elec. J. Stat.*, 3, 1075–1082.

Rosin, P. L. (1993). A note on the least squares fitting of ellipses, *Patt. Recog. Lett.*, 14, 799-808.

Rosin, P. L. and West, G. A. W. (1995). Nonparametric segmentation of curves into various representations, *IEEE. Trans. Patt. Anal. Mach. Intell.*, 17, 1140-1153.

Taubin, G. (1991). 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.*, 13, 1115–1138.

APPENDIX

The expectation $E[\Delta_1 \mathbf{M} \bar{\mathbf{M}}^- \Delta_1 \mathbf{M}]$ is computed as follows:

$$
E[\Delta_1 \mathbf{M} \bar{\mathbf{M}}^- \Delta_1 \mathbf{M}]
$$

$$
= E[\frac{1}{N}\sum_{\alpha=1}^{N}\left(\bar{\xi}_\alpha \Delta_1 \xi_\alpha^\top + \Delta_1 \xi_\alpha \bar{\xi}_\alpha^\top\right)\bar{\mathbf{M}}^- \frac{1}{N}\sum_{\beta=1}^{N}\left(\bar{\xi}_\beta \Delta_1 \xi_\beta^\top + \Delta_1 \xi_\beta \bar{\xi}_\beta^\top\right)]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} E[(\bar{\xi}_\alpha \Delta_1 \xi_\alpha^\top + \Delta_1 \xi_\alpha \bar{\xi}_\alpha^\top)\bar{\mathbf{M}}^-(\bar{\xi}_\beta \Delta_1 \xi_\beta^\top + \Delta_1 \xi_\beta \bar{\xi}_\beta^\top)]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} E[\bar{\xi}_\alpha \Delta_1 \xi_\alpha^\top \bar{\mathbf{M}}^- \bar{\xi}_\beta \Delta_1 \xi_\beta^\top + \bar{\xi}_\alpha \Delta_1 \xi_\alpha^\top \bar{\mathbf{M}}^- \Delta_1 \xi_\beta \bar{\xi}_\beta^\top
$$
$$
+ \Delta_1 \xi_\alpha \bar{\xi}_\alpha^\top \bar{\mathbf{M}}^- \bar{\xi}_\beta \Delta_1 \xi_\beta^\top + \Delta_1 \xi_\alpha \bar{\xi}_\alpha^\top \bar{\mathbf{M}}^- \Delta_1 \xi_\beta \bar{\xi}_\beta^\top]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} E[\bar{\xi}_\alpha (\Delta_1 \xi_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\Delta_1 \xi_\beta^\top
$$
$$
+ \bar{\xi}_\alpha(\Delta_1 \xi_\alpha, \bar{\mathbf{M}}^- \Delta_1 \xi_\beta)\bar{\xi}_\beta^\top + \Delta_1 \xi_\alpha(\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\Delta_1 \xi_\beta^\top
$$
$$
+ \Delta_1 \xi_\alpha(\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \Delta_1 \xi_\beta)\bar{\xi}_\beta^\top]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} E[(\Delta_1 \xi_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\bar{\xi}_\alpha \Delta_1 \xi_\beta^\top
$$
$$
+ (\Delta_1 \xi_\alpha, \bar{\mathbf{M}}^- \Delta_1 \xi_\beta)\bar{\xi}_\alpha \bar{\xi}_\beta^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top
$$
$$
+ \Delta_1 \xi_\alpha(\bar{\mathbf{M}}^- \Delta_1 \xi_\beta, \bar{\xi}_\alpha)\bar{\xi}_\beta^\top]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} E[\bar{\xi}_\alpha((\bar{\mathbf{M}}^- \bar{\xi}_\beta)^\top \Delta_1 \xi_\alpha)\Delta_1 \xi_\beta^\top
$$
$$
+ \mathrm{tr}[\bar{\mathbf{M}}^- \Delta_1 \xi_\beta \Delta_1 \xi_\alpha^\top]\bar{\xi}_\alpha \bar{\xi}_\beta^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top
$$
$$
+ \Delta_1 \xi_\alpha(\Delta_1 \xi_\beta^\top \bar{\mathbf{M}}^- \bar{\xi}_\alpha)\bar{\xi}_\beta^\top]
$$

$$
= \frac{1}{N^2}\sum_{\alpha,\beta=1}^{N} \Big(\bar{\xi}_\alpha \bar{\xi}_\beta^\top \bar{\mathbf{M}}^- E[\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top]
$$
$$
+ \mathrm{tr}[\bar{\mathbf{M}}^- E[\Delta_1 \xi_\beta \Delta_1 \xi_\alpha^\top]]\bar{\xi}_\alpha \bar{\xi}_\beta^\top
$$
$$
+ (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)E[\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top]
$$
$$
+ E[\Delta_1 \xi_\alpha \Delta_1 \xi_\beta^\top]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\beta^\top\Big)
$$

$$
= \frac{\sigma^2}{N^2}\sum_{\alpha,\beta=1}^{N} \Big(\bar{\xi}_\alpha \bar{\xi}_\beta^\top \bar{\mathbf{M}}^- \delta_{\alpha\beta} V_0[\xi_\alpha]
$$
$$
+ \mathrm{tr}[\bar{\mathbf{M}}^- \delta_{\alpha\beta} V_0[\xi_\alpha]]\bar{\xi}_\alpha \bar{\xi}_\beta^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\beta)\delta_{\alpha\beta} V_0[\xi_\alpha]
$$
$$
+ \delta_{\alpha\beta} V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\beta^\top\Big)
$$

$$
= \frac{\sigma^2}{N^2}\sum_{\alpha=1}^{N} \Big(\bar{\xi}_\alpha \bar{\xi}_\alpha^\top \bar{\mathbf{M}}^- V_0[\xi_\alpha]
$$
$$
+ \mathrm{tr}[\bar{\mathbf{M}}^- V_0[\xi_\alpha]]\bar{\xi}_\alpha \bar{\xi}_\alpha^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha)V_0[\xi_\alpha]
$$
$$
+ V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top\Big)
$$

$$
= \frac{\sigma^2}{N^2}\sum_{\alpha=1}^{N} \Big(\mathrm{tr}[\bar{\mathbf{M}}^- V_0[\xi_\alpha]]\bar{\xi}_\alpha \bar{\xi}_\alpha^\top + (\bar{\xi}_\alpha, \bar{\mathbf{M}}^- \bar{\xi}_\alpha)V_0[\xi_\alpha]
$$
$$
+ 2\mathcal{S}\,[V_0[\xi_\alpha]\bar{\mathbf{M}}^- \bar{\xi}_\alpha \bar{\xi}_\alpha^\top]\Big) \tag{57}
$$