Department of Electrical
and
Computer Systems Engineering

Technical Report
MECSE-21-2006

# A Bilinear Approach to the Parameter Estimation of a general Heteroscedastic Linear System with Application to Conic Fitting

P. Chen and D. Suter

MONASH
UNIVERSITY

# A bilinear approach to the parameter estimation of a general heteroscedastic linear system, with application to conic fitting

Pei Chen* and David Suter
Dept ECSE, P. O. Box 35, Monash University, Australia, 3800

**Abstract**: In this paper, we study the parameter estimation problem in a general heteroscedastic linear system, by putting the problem in the framework of the bilinear approach to low-rank matrix approximation. The ellipse fitting problem is studied as a specific example of the general theory. Despite the impression given in the literature, the ellipse fitting problem is still unsolved when the data comes from a small section of the ellipse. Although there are already some good approaches to the problem of conic fitting, such as FNS and HEIV, convergence in these iterative approaches is not ensured, as pointed out in the literature. Another limitation of these approaches is that they can't model the correlations among different rows of the "general measurement matrix". Our method, of employing the bilinear approach to solve the general heteroscedastic parameter estimation problem, overcomes these limitations: it is convergent and can cope with a general heteroscedastic problem. Experiments show that the proposed bilinear approach performs slightly better than other competing approaches.

**Keywords**: Parameter estimation, heteroscedastic uncertainty, bilinear approach, low-rank matrix approximation, least squares estimate, Mahalanobis distance, conic fitting.

* Corresponding author: chenpei75@yahoo.com

# 1 Introduction

Parameter estimation in a heteroscedastic system has become an active subject, in order to overcome the difficulties of the total least squares (TLS) method [13], as can be found in [6-8, 22, 23, 25, 26]. Another active research topic is to employ the bilinear approach to calculate the low-rank approximation of a large matrix in some challenging environments [11, 24, 27, 30, 31], where the traditional SVD [10] does not work or its solution is not optimal. Here, in this paper, we apply the bilinear approach to solve the parameter estimation problem in a general heteroscedastic environment. First, we review the work on these two research topics.

## 1.1 Parameter estimation in a heteroscedastic system

Many parameter estimation problems can be reduced to the following linear form:

$$\mathbf{w}^T(\mathbf{x})\boldsymbol{\theta} = 0 \tag{1}$$

$\mathbf{w}(\mathbf{x})$ are $n \times 1$ carriers of the observed quantity $\mathbf{x}$, for example, a prominent problem in computer vision: conic fitting. We will study the conic fitting problem in section 4.

Suppose $m$ different quantities $\mathbf{x}_i$ $(i = 1,2,\cdots,m)$ are observed. We arrange the carriers as a general "measurement matrix" $\mathbf{W} \in R^{m,n}$:

$$\mathbf{W} = \begin{bmatrix} \mathbf{w}^T(\mathbf{x}_1) \\ \mathbf{w}^T(\mathbf{x}_2) \\ \vdots \\ \mathbf{w}^T(\mathbf{x}_m) \end{bmatrix} \tag{2}$$

Without loss of generality, suppose $m \geq n$. If not, (2) is an underdetermined system. If $\mathbf{W}$ is noise free, it is rank deficient, with a rank of $n$-1 at most. However, it quickly becomes full rank, due to noise. Many optimization approaches and their associated objective functions have been proposed to solve this parameter estimation problem, as can be found in a comprehensive survey [33]. Among them, a straightforward solution to (1) is the right singular vector of $\mathbf{W}$, associated with the least singular value. Such a solution is usually called as the TLS estimate [13], because it minimizes the following objective function:

$$\sum \boldsymbol{\theta}^T \mathbf{w}(\mathbf{x}_i)\mathbf{w}^T(\mathbf{x}_i)\boldsymbol{\theta} \big/ \| \boldsymbol{\theta} \|^2 \tag{3}$$

It is also the maximum likelihood (ML) estimate if the noise/uncertainty in the carriers **w** (not the observed quantities **x**) is i.i.d. Gaussian.

However, the assumption of i.i.d. Gaussianality usually does not hold, especially in the system (1), because the carriers are transformed quantities of the observed data. Even if the noise in **x** can reasonably be assumed to be i.i.d. Gaussian, the uncertainty in the carriers **w** often loses this property. The violation of the i.i.d. Gaussianality makes the problem challenging to the TLS method. For example, a biased estimate is obtained by the TLS method, if the noisy points come from a segment of the conic, as testified experimentally [22, 23] and proved theoretically [20, 21].

In order to overcome the difficulties, introduced by the non-i.i.d. Gaussianality, Kanatani analyzed this problem from a geometric statistics view and devised the renormalization method [19-21]. The idea behind this is to approximately equalize the noise in all carriers. Other general approaches to this heteroscedastic problem include HEIV [22, 23, 25, 26] and FNS [6, 8]. In the HEIV model, the covariance matrix $\mathbf{C}_i$ between the carriers in $\mathbf{w}_i$ is first obtained from a linearization process, then, the parameters $\boldsymbol{\theta}$ are estimated by minimizing the Mahalanobis distance:

$$\sum_{i=1}^{m} (\mathbf{w}_i - \mathbf{w}_{io})^T \mathbf{C}_i^{-} (\mathbf{w}_i - \mathbf{w}_{io}) \tag{4}$$

where $\mathbf{C}^{-}$ is the pseudo inverse of **C** and $\mathbf{w}_{io}$ is the underlying ground truth of $\mathbf{w}_i$. This minimization problem is reduced to a generalized eigenproblem, where the generalized eigenvector, associated with the least eigenvalue, is needed. In the FNS method, an approximated maximum-likelihood (AML) objective function is employed. It is also reduced to a generalized eigenproblem. In [8], it has been proved that these two approaches, HEIV and FNS, are intimately related.

Another interesting problem in conic fitting, or in general second-order curve fitting, is to analyze the bias of the estimates. Ideally, the estimates are unbiased, like Kanatani's renormalization method [20] and Werman and Geyzel's method [32], which have been explicitly proved as unbiased. Note that Werman and Geyzel's method [32] is for

general second-order curve fitting, and that we do not imply other competing approaches, like HEIV and FNS, are biased. And, in this paper, we do not intend to analyze the biases of the estimates, and we will only concentrate on the whole performance of the estimates.

## 1.2 Bilinear approach to the low-rank matrix approximation

The SVD is the basic tool for calculating the low-rank matrix approximation. The principle behind the SVD [10] states that any matrix, $\mathbf{W} \in R^{m,n}$, can be decomposed into

$$\mathbf{W} = \mathbf{U}\mathbf{\Sigma}V^T \tag{5}$$

where $\mathbf{U} \in O^{m,m}$, $\mathbf{V} \in O^{n,n}$ and $\mathbf{\Sigma} = diag\{\sigma_1, \sigma_2, \cdots, \sigma_p\} \in R^{m,n}$, with $p = \min(m,n)$ and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$. An important fact [10] is that one can easily construct $\mathbf{W}^r$, the closest rank *r* approximation of **W**, measured by *2*-norm or Frobenius-norm, as:

$$\mathbf{W}^r = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{6}$$

Specifically,

$$\left\| \mathbf{W} - \mathbf{W}^r \right\|_2 = \sigma_{r+1} \tag{7}$$

$$\left\| \mathbf{W} - \mathbf{W}^r \right\|_F = \sqrt{\sum_{j=r+1}^{p} \sigma_j^2} \tag{8}$$

From the optimality measured by the Frobenius-norm, the estimate by (6) is also the ML estimate [12, 28, 29], if the noise in the matrix **W** is i.i.d. Gaussian.

However, the SVD method does not work on an incomplete matrix (with missing data). Moreover, the solution by (6) is not optimal if the noise in **W** does not obey the i.i.d. Gaussian model. Efforts have been devoted to the missing data problem [5, 11, 15-18] and the heteroscedastic noise problem [1-4, 14]. Except these efforts, another

promising approach to these problems is the bilinear approach [11, 30, 31][1], where one tries to fit **W** as the product:

$$\mathbf{RS} \tag{9}$$

with $\mathbf{R} \in R^{m,r}$ and $\mathbf{S} \in R^{r,n}$. To do so, one iteratively updates **R** and **S**, by alternately holding **S** and **R** constant, respectively:

$$\mathbf{R}_{new} = \min_{\mathbf{R}} \| \mathbf{W} - \mathbf{RS} \|_F \tag{10}$$

$$\mathbf{S}_{new} = \min_{\mathbf{S}} \| \mathbf{W} - \mathbf{RS} \|_F \tag{11}$$

The objective functions in (10) and (11) can be reformulated, respectively, as:

$$\| \mathbf{W} - \mathbf{RS} \|_F^2 = \sum_i \| \mathbf{w}'_i - \mathbf{r}'_i \mathbf{S} \|_F^2 = \sum_i \| \mathbf{S}^T (\mathbf{r}'_i)^T - (\mathbf{w}'_i)^T \|_F^2 \tag{12}$$

$$\| \mathbf{W} - \mathbf{RS} \|_F^2 = \sum_i \| \mathbf{R}\mathbf{s}_i - \mathbf{w}_i \|_F^2 \tag{13}$$

where $\mathbf{w}'_i$ is the $i^{th}$ row of **W** and $\mathbf{r}'_i$ is the $i^{th}$ row of **R**, and $\mathbf{s}_i$ and $\mathbf{w}_i$ are the $i^{th}$ columns of **S** and **W**, respectively[2]. If the noise in **W** is i.i.d. Gaussian, $\mathbf{r}'_i$ in (12), or $\mathbf{s}_i$ in (13), can be separately calculated as the least squares (LS) solution, which minimizes

$$\hat{\mathbf{r}}'_i = \min_{\mathbf{r}'_i} \| \mathbf{S}^T (\mathbf{r}'_i)^T - (\mathbf{w}'_i)^T \|_F^2 \tag{14}$$

$$\hat{\mathbf{s}}_i = \min_{\mathbf{s}_i} \| \mathbf{R}\mathbf{s}_i - \mathbf{w}_i \|_F^2 \tag{15}$$

Note the similarity between (12) and (13), or between (14) and (15). In (12) or (14), each row of **R**, $\mathbf{r}'_i$, needs to be computed; and similarly, each column of **S**, $\mathbf{s}_i$, needs to be computed in (13) or (15). Intrinsically, these two sub-problems are same: to solve a linear system. This way, each sub-step of the iteration is reduced to solving a linear system:

$$\mathbf{Ax}=\mathbf{b} \tag{16}$$

with the LS solution as:

$$\hat{\mathbf{x}} = \mathbf{A}^{-}\mathbf{b} \tag{17}$$

---

[1] In [31], the bilinear approach is called the PowerFactorization method.

[2] In the following, a matrix is usually denoted by a bold capital letter, eg **W**. Its $i^{th}$ column is denoted by $\mathbf{w}_i$ and its $i^{th}$ row is denoted by $\mathbf{w}'_i$.

However, the updatings of **R** and **S** in a heteroscedastic linear system, which is the subject of our paper, are no so straightforward as (17), as can be found in section 3.

***Algorithm 1*** (Bilinear algorithm):
Given an input matrix **W**, factor **W** as the product of **RS** as (9).

1. Initialize the factor **R** in (9). For example, **R** can be randomly generated as an *m-by-r* matrix, as in [30].
2. While keeping **R** constant, update each column of **S**, $\mathbf{s}_i$, according to the LS solution (15).
3. While keeping **S** constant, update each row of **R**, $\mathbf{r}_i'$, according to the LS solution (14).
4. Calculate the product of $\hat{\mathbf{W}} = \hat{\mathbf{R}}\hat{\mathbf{S}}$. If this is the first iteration, go into step 2; else if $\| \hat{\mathbf{W}} - \mathbf{W}_{old} \| < \varepsilon$, where $\varepsilon$ is a small positive number, end the iteration and output $\hat{\mathbf{R}}$ and $\hat{\mathbf{S}}$, else store $\hat{\mathbf{W}}$ as $\mathbf{W}_{old}$ and go to step 2.

In the limit, as $\varepsilon \to 0$, the product of **RS** approaches the *r*-rank approximation matrix by SVD, i.e. that from (6). Thus, the rank *r* approximation of matrix W can also be solved by iteratively updating its factors row-by-row (column-by-column), as in algorithm 1. More details about the bilinear approach can be found in [30], and we will revisit this point in section 3. In this bilinear approach to the low-rank approximation, the missing data problem can be naturally coped with, and a scalar-weighted uncertainty can also be incorporated [30]. Moreover, this bilinear approach can be further developed to incorporate directional uncertainty [27] (although the measurement matrix was assumed to be complete in [27], the method can be naturally extended to the missing data problem, with directional uncertainty.)

However, the bilinear algorithm, algorithm 1, also suffers from the heteroscedastic noise problem. The reason for this lies in the fact that the central idea of the bilinear algorithm is its LS-based updating rules, as step 2 and step 3 in algorithm 1. In (14) and (15), the objective function is to minimize the sum of the square difference, without considering any statistical properties among the data in **W**.

### 1.3 The issues to be studied and the organization of this paper

In HEIV and FNS, only the correlation among the carriers in each $\mathbf{w}(\mathbf{x}_i)$ can be dealt with, although this is the most common case in practice. (Note that there are *n* transformed quantities in each $\mathbf{w}(\mathbf{x}_i)$.) In this paper, we will consider the general case, where the uncertainties in different carriers $\{\mathbf{w}(\mathbf{x}_i)\}$ are correlated. To do so, we rephrase the general heteroscedastic parameter estimation problem into the framework of the bilinear approach. Then, to make our theory concrete, we consider a specific computer vision task: conic fitting.

In section 2, we formulate the parameter estimation problem with an objective function which is subtly different from (4), and then we rephrase this problem in the framework of the low-rank matrix approximation. In section 3, we present our bilinear approach to the problem of the low-rank approximation in the heteroscedastic system. In section 4, we study the specific computer vision task: conic fitting, including the issue of noise level estimation. In section 5, our results, with comparison with other competing approaches, are presented.

# 2 The parameter estimation problem

## 2.1 Objective function to be minimized

Temporarily, we suppose that the noise model for the carriers $\{\mathbf{w}(\mathbf{x}_i)\}$ is known. More precisely, the correlated Gaussian model, with covariance matrix $\mathbf{C} \in R^{mn,mn}$, is employed to characterize the uncertainties among the vectorized carriers $vec1\{\mathbf{W}\}$, where

$$vec1\{\mathbf{W}\} = \begin{bmatrix} \mathbf{w}(\mathbf{x}_1) \\ \mathbf{w}(\mathbf{x}_2) \\ \vdots \\ \mathbf{w}(\mathbf{x}_m) \end{bmatrix} \in R^{mn,1} \tag{18}$$

Please note that the covariance matrix $\mathbf{C}$ is symmetric and positive semi-definite, and can be factorized into $\mathbf{C} = \sum_{i=1}^{mn} \sigma_i \mathbf{u}_i \mathbf{u}_i^T$, with $\sigma_i \geq 0$. The characterization of the uncertainty using correlated Gaussian models is application dependent. For example, the uncertainties in the applications of conic fitting and fundamental matrix estimation have been studied in [22, 23]. In section 4.1, we will include the conic fitting as an example of how to obtain the correlated Gaussian model.

We start by defining the following modified Mahalanobis distance as the objective function to be minimized:

$$\min_{\mathbf{l}} (vec1\{\mathbf{W}\} - \mathbf{l})^T \mathbf{C}^+ (vec1\{\mathbf{W}\} - \mathbf{l}) \tag{19}$$

where $\mathbf{C}^+ = \sum_{i=1}^{mn} \frac{\mathbf{u}_i \mathbf{u}_i^T}{\sigma_i}$. The vector $\mathbf{l} = \begin{bmatrix} \mathbf{l}_1 \\ \mathbf{l}_2 \\ \vdots \\ \mathbf{l}_m \end{bmatrix} \in R^{mn,1}$ in (19), with $\mathbf{l}_i \in R^n$, is associated with a rank *n*-1 matrix $\mathbf{L} = \begin{bmatrix} \mathbf{l}_1^T \\ \mathbf{l}_2^T \\ \vdots \\ \mathbf{l}_m^T \end{bmatrix} \in R^{m,n}$.

In plain language, the minimization of the objection function (19) is to obtain a rank *n*-1 approximation matrix, which has the minimal modified Mahalanobis distance to the general measurement matrix. However, because the (modified) Mahalanobis distance is defined for vector and is not applicable to matrix, we have to vectorize the *m*-by-*n* matrix to an *mn* vector, as in (18). If the uncertainties in the general measurement matrix are Gaussian, i.i.d. or correlated, the minimizer of the (19) is the ML estimate, as will be shown in section 3.

Assume that $\hat{\mathbf{l}}$ is the solution of the system of (19), and it has an associated rank *n*-1 matrix $\hat{\mathbf{L}}$. ***The solution of the system of (1) is taken as the right singular vector of*** $\hat{\mathbf{L}}$***, associated with the least singular value.***

If the uncertainties in different carriers $\mathbf{w}(\mathbf{x}_i)$ and $\mathbf{w}(\mathbf{x}_j)$ for $i \neq j$ are independent, the objective function (19) can be formulated as

$$\sum_{i=1}^{m} (\mathbf{w}_i - \mathbf{w}_{io})^T \mathbf{C}_i^{+} (\mathbf{w}_i - \mathbf{w}_{io}) \tag{20}$$

where $\mathbf{C}_i$ is the covariance matrix for $\mathbf{w}(\mathbf{x}_i)$ and the matrix, $\begin{bmatrix} \mathbf{w}_{1o}^T \\ \mathbf{w}_{2o}^T \\ \vdots \\ \mathbf{w}_{mo}^T \end{bmatrix}$, has a rank of *n*-1.

Despite the similarity between (20) and (4), which is the objective function of the HEIV method, please note the difference between them. First, $\mathbf{w}_{io}$ is the assumed underlying ground truth, in (4). In contrast, in (20), $\mathbf{w}_{io}$ can be characterized by the property that its associated matrix has a rank of *n*-1. We are deliberately projecting onto the "nearest" rank *n*-1 matrix as the *starting* point of our bilinear approach to the heteroscedastic problem. Second, the modified Mahalanobis distance is employed in (20). In contrast, the Mahalanobis distance is employed in (4). They are identical if the covariance matrix is non-singular. However, there is a difference in cases, where the covariance matrix is singular, i.e., some singular values of the covariance matrix are zeroes. Obviously, if $\sigma_i$ is zero, and (19) or (20) are not mathematically meaningful. It

will become clear in section 3.1.1, that, in such cases, this can be reduced to an equality constrained LS problem [10]. In contrast, (4) can be reduced to a LS problem.

# 3 The bilinear approach to the heteroscedastic parameter estimation

As stated as (10-15), the central idea of the bilinear approach is to iteratively update **R** or **S**, by holding **S** or **R** constant, respectively. Although, in section 1.2, we have reformulated each updating step of the bilinear approach in a simple mathematical language, as the linear system (16); the case becomes complicated if the uncertainty model in **W** is not i.i.d. Gaussian. Our approach to this problem is to extend the bilinear approach to cope with heteroscedastic noise. As in algorithm 1, we also iteratively update the factors **R** and **S**, while keeping **S** and **R** constant, respectively. However, we can't simply update **R** and **S** row-by-row/column-by-column, using the LS based rule.

In order to simplify the development of the solution to the low-rank approximation in a general heteroscedastic system, we first consider the case of (20), where the uncertainties between different carriers $\mathbf{w}(\mathbf{x}_i)$ and $\mathbf{w}(\mathbf{x}_j)$ for $i \neq j$ are independent and the uncertainty in $\mathbf{w}(\mathbf{x}_i)$ is assumed to be a Gaussian density with a covariance matrix $\mathbf{C}_i$. This particular case will be developed in sections 3.1 and 3.2, and the algorithm will be given in section 3.2.

## 3.1 Update of R

In the case of (20), the uncertainties between different rows of **W** are assumed to be independent, so we can *separately* update each row of **R**, i.e., row by row. The updating of each row of **R** equals to solving the linear system (16). Note that, **A** is $\mathbf{S}^T$, and **b** is the transform of the $i^{th}$ row of **W**, $(\mathbf{w}'_i)^T$; and the estimated $\hat{\mathbf{x}}$ would be the transform of the $i^{th}$ row of **R**, $(\mathbf{r}'_i)^T$, which is to be updated. Because the uncertainties in **b** are modeled as correlated Gaussian noise, with a presumably known covariance matrix of **C**, and (16) becomes a *heteroscedastic linear* system, we can't use the LS-based solution (17).

Instead, we use the following minimization objective function for a linear heteroscedastic system (16):

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} (\mathbf{Ax} - \mathbf{b})^T \mathbf{C}^+ (\mathbf{Ax} - \mathbf{b}) \tag{21}$$

Suppose $\mathbf{C} = \mathbf{U} diag(d_1, d_2, \cdots, d_n)\mathbf{U}^T$ . Define $\mathbf{Q} = diag(1/\sqrt{d_1}, 1/\sqrt{d_2}, \cdots, 1/\sqrt{d_n})\mathbf{U}^T$ . The solution to (21) is:

$$\hat{\mathbf{x}} = (\mathbf{QA})^+ \mathbf{Qb} \tag{22}$$

*Proof*: We arrange the minimization objective function in (21) as:

$$(\mathbf{Ax} - \mathbf{b})^T \mathbf{Q}^T \mathbf{Q} (\mathbf{Ax} - \mathbf{b}) = (\mathbf{QAx} - \mathbf{Qb})^T (\mathbf{QAx} - \mathbf{Qb})$$

Obviously, (22) is the solution of minimizing the above objective function, and consequently, is the solution of (21).

It will become clear in section 3.1.1, that the uncertainties in $\mathbf{Qb}$ are i.i.d. Gaussian. And, (22) is the LS solution the equation of $(\mathbf{QA})\mathbf{x} = \mathbf{Qb}$ . By the transformation of $\mathbf{Q}$, the heteroscedastic noise property in each row of $\mathbf{W}$ is considered.

### 3.1.1 Case with zero singular values in the covariance matrix C

As we note in section 2.1, the modified pseudo inverse of the covariance matrix $\mathbf{C}$ does not make sense if $\mathbf{C}$ has some zero singular values. However, there is usually a constant carrier in (1), i.e., this component is noise free. Consequently, $\mathbf{C}$ has, at least, a zero singular value. Here, we study this case and present our solution to this problem. *This analysis also presents the central idea of the updating in a heteroscedastic system*: *we transform the quantities so that the transformed uncertainties become i.i.d. Gaussian and the LS solution can be applied to the transformed system.*

First, we study the covariance matrix of the transformed $\mathbf{b}$, $\mathbf{U}^T\mathbf{b}$.

$$\text{cov}(\mathbf{U}^T\mathbf{b}) = \mathbf{U}^T \text{cov}(\mathbf{b})\mathbf{U} = \mathbf{U}^T\mathbf{C}\mathbf{U} = diag(d_1, d_2, \cdots, d_n) \tag{23}$$

(23) means that the coupled uncertainties in $\mathbf{b}$ have been decoupled in the transformed $\mathbf{U}^T\mathbf{b}$. If all $d_i \neq 0$ and the coupled uncertainties in $\mathbf{b}$ are Gaussian, the uncertainties in $\mathbf{Qb}$ are i.i.d. Gaussian.

Any zero $d_i$ in (23) means that the $\mathbf{u}_i$-direction component of $\mathbf{b}$, $\mathbf{u}_i^T\mathbf{b}$, has no uncertainties or noise. Without loss of generality, we suppose the last $k$ $d_i$ for $i = n-k+1,\cdots,n$ are zeroes.

In the following, we transform **b** into two parts: one part is with i.i.d. Gaussian noise, and the other is noise free. Define $\mathbf{A}_1 = diag(1/\sqrt{d_1},1/\sqrt{d_2},\cdots,1/\sqrt{d_{n-k}})\begin{bmatrix}\mathbf{u}_1^T\\ \mathbf{u}_2^T\\ \vdots\\ \mathbf{u}_{n-k}^T\end{bmatrix}\mathbf{A}$,

$\mathbf{A}_2 = \begin{bmatrix}\mathbf{u}_{n-k+1}^T\\ \mathbf{u}_{n-k+2}^T\\ \vdots\\ \mathbf{u}_n^T\end{bmatrix}\mathbf{A}$ , $\mathbf{b}_1 = diag(1/\sqrt{d_1},1/\sqrt{d_2},\cdots,1/\sqrt{d_{n-k}})\begin{bmatrix}\mathbf{u}_1^T\\ \mathbf{u}_2^T\\ \vdots\\ \mathbf{u}_{n-k}^T\end{bmatrix}\mathbf{b}$ and

$\mathbf{b}_2 = \begin{bmatrix}\mathbf{u}_{n-k+1}^T\\ \mathbf{u}_{n-k+2}^T\\ \vdots\\ \mathbf{u}_n^T\end{bmatrix}\mathbf{b}$.

Now, it is clear that the uncertainties in $\mathbf{b}_1$ are i.i.d. Gaussian, and that $\mathbf{b}_2$ is noise free. Thus, the optimum estimate of (21) should be the solution of the following constrained minimization problem:

$$\min_{\mathbf{A}_2\mathbf{x}=\mathbf{b}_2} \mathbf{A}_1\mathbf{x} = \mathbf{b}_1 \qquad (24)$$

(24) is an equality constrained least squares problem, and its solution can be found in [10] (See the appendix, too).

Now, it is clear that our objective function in (19) or (20) makes sense if we adopt the interpretation of $0/0 = 0$. More importantly, the solution of (24) is the minimizer of (21), under this assumption.

In contrast, if we employ (4) as the objective function in zero-singular-value cases, as in the HEIV method, (21) will be reduced to a simple LS problem: $\min \mathbf{A}_1\mathbf{x} = \mathbf{b}_1$, without the equality constrained $\mathbf{A}_2\mathbf{x} = \mathbf{b}_2$. Obviously, the objective function of (4) can't be

employed in such cases. In order to overcome this difficulty in [22, 23], the constant component is not included in (4), and has been separately considered from the other columns.

*As can be seen above, the updates of **R** can be computed, row by row, because the noises in **W** are row-independent. And, because the noises in each row of **W** are correlated, a non-Frobenius-norm objective function (21) is employed. The central idea in the updating of each row of* **R** *is that, we transform the quantities so that the transformed uncertainties become i.i.d. Gaussian and the LS solution can be applied to the transformed system. This principle also applies to the analysis below.*

## 3.2 Update of S

The scene changes, as the updating of **S** is concerned. Because the uncertainties in different columns of **W**, $\mathbf{w}_i$ and $\mathbf{w}_j$ for $(i \neq j)$ in (13), are not independent, the updating of **S** can't be dealt with, column by column, as in the updating of **S** above. We have to jointly solve a matrix equation: **AX=B**. Note that, **A** is **R**, and **B** is **W**; and that **X** is the **S**, which is to be updated.

Fundamentally, we abstract (13) as the following minimization problem.

Suppose $\mathbf{A} = \begin{bmatrix} \mathbf{a}'_1 \\ \mathbf{a}'_2 \\ \vdots \\ \mathbf{a}'_m \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} \mathbf{b}'_1 \\ \mathbf{b}'_2 \\ \vdots \\ \mathbf{b}'_m \end{bmatrix}$, with $\mathbf{a}'_i \in R^{1,n}$ and $\mathbf{b}'_i \in R^{1,r}$. Suppose $\mathbf{b}'^T_i$ is corrupted by correlated Gaussian noise with $\mathbf{C}_i$ covariance matrix, which can be factorized into $\mathbf{C}_i = (\mathbf{U}^i) diag(d_1^i, d_2^i, \cdots, d_r^i)(\mathbf{U}^i)^T$. And, the uncertainties in $\mathbf{b}'^T_i$ are independent of those in $\mathbf{b}'^T_j$ for $j \neq i$. Similar as (21), we employ the following objective function:

$$\mathbf{X} = \min_{\mathbf{X} \in R^{n,r}} \sum_{i=1}^{n} (\mathbf{a}'_i \mathbf{X} - \mathbf{b}'_i) \mathbf{C}_i^{+} (\mathbf{a}'_i \mathbf{X} - \mathbf{b}'_i)^T \tag{25}$$

We suppose that $d_i^j \neq 0$. If not, we can convert the problem to an equality constrained least squares problem (24), as in section 3.1.1. Define

$\mathbf{\Omega}_i = diag(1/\sqrt{d_1^i}, 1/\sqrt{d_2^i}, \cdots, 1/\sqrt{d_r^i})(\mathbf{U}^i)^T$. The correlated uncertainties in **B** can be decoupled by:

$$\begin{bmatrix} \mathbf{a}_1'\mathbf{X}\mathbf{Q}_1^T \\ \mathbf{a}_2'\mathbf{X}\mathbf{Q}_2^T \\ \vdots \\ \mathbf{a}_m'\mathbf{X}\mathbf{Q}_m^T \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1'\mathbf{Q}_1^T \\ \mathbf{b}_2'\mathbf{Q}_2^T \\ \vdots \\ \mathbf{b}_m'\mathbf{Q}_m^T \end{bmatrix} \tag{26}$$

(26) equals to the following linear system:

$$\begin{bmatrix} \mathbf{Q}_1 \otimes \mathbf{a}_1' \\ \mathbf{Q}_2 \otimes \mathbf{a}_2' \\ \vdots \\ \mathbf{Q}_m \otimes \mathbf{a}_m' \end{bmatrix} vec(\mathbf{X}) = \begin{bmatrix} vec(\mathbf{b}_1'\mathbf{Q}_1^T) \\ vec(\mathbf{b}_2'\mathbf{Q}_2^T) \\ \vdots \\ vec(\mathbf{b}_m'\mathbf{Q}_m^T) \end{bmatrix} \tag{27}$$

where $\otimes$ denotes the Kronecker product of two matrices, and, for a matrix **X** with *r* columns, $vec(\mathbf{X}) = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_r \end{bmatrix}$. (27) comes from the property of the Kronecker product: $vec(\mathbf{AXB}) = (\mathbf{B}^T \otimes \mathbf{A})vec(\mathbf{X})$. The uncertainties in the right side of (27) have already been made i.i.d. Gaussian. Thus, the minimizer to *vec*(**X**), and consequently **X**, can be obtained by the LS estimation.

Note the solution of *vec*(**X**) from (27), and consequently **X**, minimizes the objective function in (25).

*Proof*: we arrange the minimization objective function in (25) as:

$$\begin{aligned} &\sum_{i=1}^{n} (a_i'\mathbf{X} - \mathbf{b}_i')\mathbf{Q}_i^T\mathbf{Q}_i(a_i'\mathbf{X} - \mathbf{b}_i')^T \\ &= \sum_{i=1}^{n} (\mathbf{a}_i'\mathbf{X}\mathbf{Q}_i^T - \mathbf{b}_i'\mathbf{Q}_i^T)(\mathbf{a}_i'\mathbf{X}\mathbf{Q}_i^T - \mathbf{b}_i'\mathbf{Q}_i^T)^T \\ &= \sum_{i=1}^{n} [(\mathbf{Q}_i \otimes \mathbf{a}_i')vec(\mathbf{X}) - vec(\mathbf{b}_i'\mathbf{Q}_i^T)][(\mathbf{Q}_i \otimes \mathbf{a}_i')vec(\mathbf{X}) - vec(\mathbf{b}_i'\mathbf{Q}_i^T)]^T \end{aligned}$$

The uncertainties in $vec(\mathbf{b}_i'\mathbf{Q}_i^T)$ are i.i.d. Gaussian, so the LS estimate can be applied to (27). The LS estimate of *vec*(**X**) in (27) minimizes the above objective function, and consequently, the related **X** minimizes the objective function in (25).

### 3.2.1 Constant column in the measurement matrix

Assume that there is a constant column in the general measurement matrix, as will be found in the conic fitting, i.e. $\mathbf{M} = [\mathbf{M}', \mathbf{1}]$. In such cases, we can single out the constant column in the above updating of **S**. The last column of **S**, $\mathbf{s}_r$, can be calculated as:

$$\hat{\mathbf{s}}_r = \mathbf{R}^{-}\mathbf{1} \tag{28}$$

Note, in the updating of each row of **R**, we convert the minimization problem of (21) as an equality constrained LS estimation problem of (24). Consequently, $\hat{\mathbf{r}}_i'\mathbf{S} = [\hat{\mathbf{m}}', 1]$. It is clear that the approximated measurement matrix $\hat{\mathbf{R}}\mathbf{S}$, after each updating of **R**, has an exact constant column **l** (with all ones). This means that $\mathbf{l} \in span(\hat{\mathbf{R}})$. So, (28) holds without any error, i.e. $\mathbf{R}\hat{\mathbf{s}}_r = \mathbf{1}$.

### 3.2.2 Discussion of the convergence of the bilinear approach

In sections 3.1 and 3.1.1, we studied the updating of **R**, where the objective function is (20). Because of the assumed independence among the uncertainties in different rows of **W** in (2), we can separately update each row of **R**, minimizing the associated part in the sum of (20). In section 3.2, we jointly updated **S** in order to incorporate the correlation among different columns of **W**. This way, the objective function in (20) decreases after each updating step of **R** or **S**. From these observations, we can see that the bilinear approach converges, in contrast to the lack of proof of convergence of the HEIV or FNS methods.

### 3.2.3 The algorithm for the case with independent rows in W

Here, according to the analysis in sections 3.1 and 3.2, we outline our bilinear algorithm, which deals with a special class of matrix **W**: the uncertainties in different rows of **W** are assumed to be independent.

***Algorithm 2***:

Given an input matrix **W**, factor **W** as the product of **RS** as (9).

1. Initialize the factor **R** in (9).
2. While keeping **R** constant, update the factor **S** as a whole, according to the analysis in section 3.2.

3. While keeping **S** constant, update each row of **R**, $\mathbf{r}_i'$, according to the analysis in section 3.13.2. Note that this step can be done row-by-row.
4. Calculate the product of $\hat{\mathbf{W}} = \hat{\mathbf{R}}\hat{\mathbf{S}}$. If this is the first iteration, go into step 2; else if $\| \hat{\mathbf{W}} - \mathbf{W}_{old} \| < \varepsilon$, where $\varepsilon$ is a small positive number, end the iteration and output $\hat{\mathbf{R}}$ and $\hat{\mathbf{S}}$, else store $\hat{\mathbf{W}}$ as $\mathbf{W}_{old}$ and go to step 2.

## 3.3 A more general update

In sections 3.1 and 3.2, we have presented how to update **R** and **S**, in a linear system, where row-independent noises exist. Because the uncertainties in **W** are row-independent, we can update **R**, row by row; while the update of **S** can't be done column by column, and has to be jointly considered.

In a general heteroscedastic system, the uncertainties are not row-independent or row-independent. Then, the updates of both **R** and **S** have to be reduced to solving a matrix equation:

$$\mathbf{AX=B} \tag{29}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_m' \end{bmatrix} \in R^{m,r}$, $\mathbf{B} = \begin{bmatrix} \mathbf{b}_1' \\ \mathbf{b}_2' \\ \vdots \\ \mathbf{b}_m' \end{bmatrix} \in R^{m,n}$, $\mathbf{X} \in R^{r,n}$. In a general heteroscedastic case, the uncertainties in **B** are characterized by the covariance matrix **C** for the vectorized $vec1(\mathbf{B}) = \begin{bmatrix} \mathbf{b}_1' & \mathbf{b}_2' & \cdots & \mathbf{b}_m' \end{bmatrix}^T \in R^{mn,1}$. The objective function to be minimized is:

$$\hat{\mathbf{X}} = \min_{\mathbf{X}} (vec1(\mathbf{AX} - \mathbf{B}))\mathbf{C}^{+}(vec1(\mathbf{AX} - \mathbf{B}))^T \tag{30}$$

Similarly, **C** can be factorized as $\mathbf{C} = \mathbf{U} diag(d_1, d_2, \cdots, d_{mn})\mathbf{U}^T$, and define $\mathbf{Q} = diag(1/\sqrt{d_1}, 1/\sqrt{d_2}, \cdots, 1/\sqrt{d_{mn}})\mathbf{U}^T$.

First, we convert the equation of $\mathbf{a}_i'\mathbf{X} = \mathbf{b}_i'$ to $(I_n \otimes \mathbf{a}_i')vec(\mathbf{X}) = \mathbf{b}_i'^T$. Then, **AX=B** can be rewritten as:

$$\begin{bmatrix} I_n \otimes \mathbf{a}'_1 \\ I_n \otimes \mathbf{a}'_2 \\ \vdots \\ I_n \otimes \mathbf{a}'_m \end{bmatrix} vec(\mathbf{X}) = \begin{bmatrix} \mathbf{b}'^T_1 \\ \mathbf{b}'^T_2 \\ \vdots \\ \mathbf{b}'^T_m \end{bmatrix} \tag{31}$$

and further as

$$\mathbf{Q}\begin{bmatrix} I_n \otimes \mathbf{a}'_1 \\ I_n \otimes \mathbf{a}'_2 \\ \vdots \\ I_n \otimes \mathbf{a}'_m \end{bmatrix} vec(\mathbf{X}) = \mathbf{Q}\begin{bmatrix} \mathbf{b}'^T_1 \\ \mathbf{b}'^T_2 \\ \vdots \\ \mathbf{b}'^T_m \end{bmatrix} \tag{32}$$

The uncertainties in the right side of (32) have been i.i.d. Gaussian if the uncertainties in **B** are Gaussian. Thus, (32) can be solved by the LS method, and the associated **X** can be obtained.

### 3.4 Disussion of the optimality

From the above sections, we can see that the optimal solution, at least a local optimal solution, is iteratively obtained, if we evaluate the estimate using the objective functions in (19) or (20). However, it is not the ML estimate if the uncertainties in **W** are not Gaussian. Because of this, we assumed the uncertainties in **W** are Gaussian when we referred to the ML estimate above.

# 4 Application in conic fitting

We can reasonably assume that the observed data **x** is corrupted with i.i.d. Gaussian noise. However, the uncertainty within the measurement matrix **W** often loses this i.i.d. Gaussianality. Furthermore, as can be observed in section 3 and in the objective function in (19) or (20), the crux of our bilinear approach to the heteroscedastic low-rank approximation, and consequently of the heteroscedastic parameter estimation problem, is to obtain the covariance matrix of the noise in the carriers **w** (one row of **W**). In this section, we will study this important issue of modeling the heteroscedastic characteristic of the carriers **w**, by taking the conic fitting problem as a specific example. This issue has been analyzed in [22, 23], where the covariance matrix of the carriers was obtained.

As in [6, 8, 22, 23], we also assume that each component of the observed **x** is corrupted with i.i.d. and $\sigma^2$-variance Gaussian noise, and consequently, that the uncertainties in different rows of **W** are independent.

## 4.1 Covariance matrix in the conic fitting

A conic is characterized by the following constraint:

$$ax^2 + bxy + cy^2 + dx + ey + f = 0 \tag{33}$$

The carriers in (33) are $x^2$, $xy$, $y^2$, $x$, $y$, and 1. By the linearization, we reformulate (33) in the form of (1):

$$[x_i, y_i, x_i y_i, x_i^2, y_i^2, 1][d, e, b, a, c, f]^T = 0 \tag{34}$$

The conic fitting problem is to estimate the parameters, $a$, $b$, $c$, $d$, $e$ and $f$, from a few (at least 6), noisy points.

We can neglect the constant component, by using the techniques suggested in section 3.1.1 and 3.2.1. So, we only need study the uncertainty model for the first five carriers $[x, y, xy, x^2, y^2]$.

Suppose we observe $x$, $y$, with noise $\varepsilon_x$ and $\varepsilon_y$ in them, respectively. The uncertainties in the carriers $[x, y, xy, x^2, y^2]$, introduced by $\varepsilon_x$ and $\varepsilon_y$, are:

$$\boldsymbol{\varepsilon} = \begin{bmatrix} x_o + \varepsilon_x \\ y_o + \varepsilon_y \\ (x_o + \varepsilon_x)(y_o + \varepsilon_y) \\ (x_o + \varepsilon_x)^2 \\ (y_o + \varepsilon_y)^2 \end{bmatrix} - \begin{bmatrix} x_o \\ y_o \\ x_o y_o \\ x_o^2 \\ y_o^2 \end{bmatrix} = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ y_o\varepsilon_x + x_o\varepsilon_y + \varepsilon_x\varepsilon_y \\ 2x_o\varepsilon_x + \varepsilon_x^2 \\ 2y_o\varepsilon_y + \varepsilon_y^2 \end{bmatrix} \tag{35}$$

And, $\boldsymbol{\varepsilon}$ can be expressed as:

$$\boldsymbol{\varepsilon} = \mathbf{D}[\varepsilon_x, \varepsilon_y]^T + [0,0,\varepsilon_x\varepsilon_y, \varepsilon_x^2, \varepsilon_y^2]^T \tag{36}$$

where

$$\mathbf{D} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ y_o & x_o \\ 2x_o & 0 \\ 0 & 2y_o \end{bmatrix} \tag{37}$$

and the subscript "*o*" denotes the underlying ground truth of the associated quantity. In practice, the ground truth is unknown and we use the observed quantities in place of the ground truth. So, in the following, we do not use the symbol "*o*" in the subscript.

The covariance matrix of $\boldsymbol{\varepsilon}$ can be expressed in a matrix form. As in [22, 23], we employ the following covariance matrix to characterize the uncertainties in $[x, y, xy, x^2, y^2]$:

$$\sigma^2 \begin{bmatrix} 1 & 0 & y & 2x & 0 \\ 0 & 1 & x & 0 & 2y \\ y & x & x^2 + y^2 + \sigma^2 & 2xy & 2xy \\ 2x & 0 & 2xy & 4x^2 + 2\sigma^2 & 0 \\ 0 & 2y & 2xy & 0 & 4y^2 + 2\sigma^2 \end{bmatrix} = \sigma^2 \widetilde{\mathbf{C}} \tag{38}$$

If we drop the $\sigma^2$ in $\widetilde{\mathbf{C}}$ in (38), we have:

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & y & 2x & 0 \\ 0 & 1 & x & 0 & 2y \\ y & x & x^2+y^2 & 2xy & 2xy \\ 2x & 0 & 2xy & 4x^2 & 0 \\ 0 & 2y & 2xy & 0 & 4y^2 \end{bmatrix} = \mathbf{D}\mathbf{D}^T \tag{39}$$

From (39), **C** has a rank of 2. Then, **C** can be factored as: $\mathbf{C} = \mathbf{U}diag(d_1, d_2, 0, 0, 0)\mathbf{U}^T$, and we have

$$span(\mathbf{u}_1, \mathbf{u}_2) = span(\mathbf{D}) \tag{40}$$

If the $x$ and $y$ coordinates are much larger than the noise in them (this is true in most points), it would hold that

$$span(\mathbf{u}_1, \mathbf{u}_2) \approx span(\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2) \tag{41}$$

where $\tilde{\mathbf{u}}_1, \tilde{\mathbf{u}}_2$ are the singular vectors of $\tilde{\mathbf{C}}$, associated with the two largest singular values. This can be obtained from the matrix perturbation theory, by regarding the terms of $\sigma^2$ in $\tilde{\mathbf{C}}$ as some perturbation.

From (35), (36), (39), (40) and (41), the first order uncertainties are modeled by the covariance matrix **C**, and consequently approximately by the first two singular vectors of $\tilde{\mathbf{C}}$, associated with the two largest singular values. This property will be used in the noise level estimation.

## 4.2 Noise level estimation

As can be observed in (38), the noise level, $\sigma$, in the observed quantities is needed in obtaining the covariance matrix of the carriers. Because the second order terms in (36) are not Gaussian, we only use their first order uncertainties in estimating the noise level in the observed data. Taking the conic fitting as an example, the first order uncertainties are $\mathbf{D}[\varepsilon_x, \varepsilon_y]^T$.

First, we have the following fact

$$[\varepsilon_x, \varepsilon_y]\mathbf{D}^T(\mathbf{D}\mathbf{D}^T)^{-}\mathbf{D}[\varepsilon_x, \varepsilon_y]^T = \varepsilon_x^2 + \varepsilon_y^2 \tag{42}$$

where $\mathbf{D}^T(\mathbf{D}\mathbf{D}^T)^{-}\mathbf{D} = diag(1,1)$.

From (35), (36), (39), (40) and (41), the rank 2 approximation of the covariance matrix $\widetilde{\mathbf{C}}$, is approximately $\mathbf{C}$ if the $x$ and $y$ coordinates are much larger than the noise level: $\widetilde{\mathbf{C}}^2 \approx \mathbf{C} = \mathbf{D}\mathbf{D}^T$. Moreover, the uncertainties captured by the 2 largest singular vectors of $\widetilde{\mathbf{C}}$, are approximately $\mathbf{D}[\varepsilon_x, \varepsilon_y]^T$. Combining these observations and (42), we employ the following estimate for the noise level:

$$\sqrt{\frac{1}{2m}\sum_{i=1}^{m}\mathbf{e}_i'\widetilde{\mathbf{C}}_i^{-2}\mathbf{e}_i'^T} \tag{43}$$

where $\widetilde{\mathbf{C}}_i^{-2}$ is the pseudo inverse of the rank-2 approximation matrix of $\widetilde{\mathbf{C}}_i$ in (38), and $\mathbf{e}'$ is the $i^{th}$ row of the error matrix $\mathbf{E}$, $\mathbf{E} = \mathbf{W} - \hat{\mathbf{R}}\hat{\mathbf{S}}$, with $\hat{\mathbf{R}}$ and $\hat{\mathbf{S}}$ as the current estimates in the bilinear approach (9). Note, in the calculation of the error matrix $\mathbf{E}$, the constant column in $\mathbf{W}$ is not included.

# 5 Experimental results

In this section, we conduct experiments on the conic fitting, to validate the correctness of our general theory in section 3. With this aim, we mainly compare our approach with other competing approaches to this problem: including FNS [6], HEIV [22, 23], KAN [20, 21] and the *constrained* TLS method [9]. The method in [9] is a specific implementation of the TLS method [13], for the conic fitting problem, as pointed out in [23], **in particular it enforces that the solution is an ellipse**.

It has been established in [8], that the HEIV and the FNS are intimately related, with only different numerical solution; and it has also experimentally proved that both of them have almost same performance, where the AML objective function is employed as a criterion. The following experiments suggest that HEIV performs better than FNS in the more challenging problems, for example, where the points distribute in a small portion (e.g., a quarter) of the ellipse; although they have almost same performance in other mildly challenging settings, for example, where the points are from an half ellipse. We do not know the reason for this difference in performance.

In all the experiments, we use the following setting: the true ellipse has a major axis of size 100 and a minor axis of size 50. Two factors have much influence on the estimates of, almost all, the methods mentioned above: the noise level and the span of the points. All the methods produce good estimates, indeed estimates which are almost same, if the points span the whole ellipse. Because of this, we do not run experiments on the whole ellipse.

## 5.1 Noise level=2 over a half ellipse

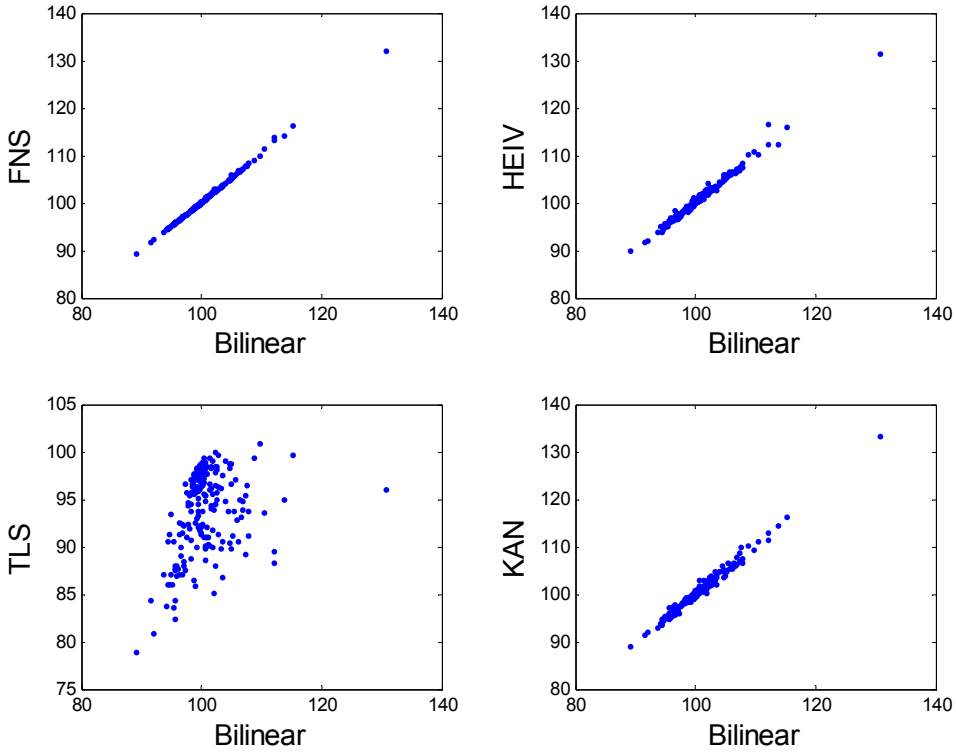
First, we conduct the experiments in the following setting. On the ellipse $r^2\cos^2\theta/a^2 + r^2\sin^2\theta/b^2 = 1$, a point is determined by the angle $\theta$. In this experiment, 100 points are randomly generated on a half ellipse: $\theta$, with a uniform distribution, is randomly distributed within $(\theta_0, \theta_0 + \pi)$; and $\theta_0$ is also uniformly located within $[0, 2\pi)$. Then, i.i.d. Gaussian noise, with noise level of 2, is added to the 100 points. The experiment is repeated 200 times (with different random samples).

Surprisingly, our bilinear approach performs almost identically to HEIV, FNS and KAN, all of which perform better than TLS.

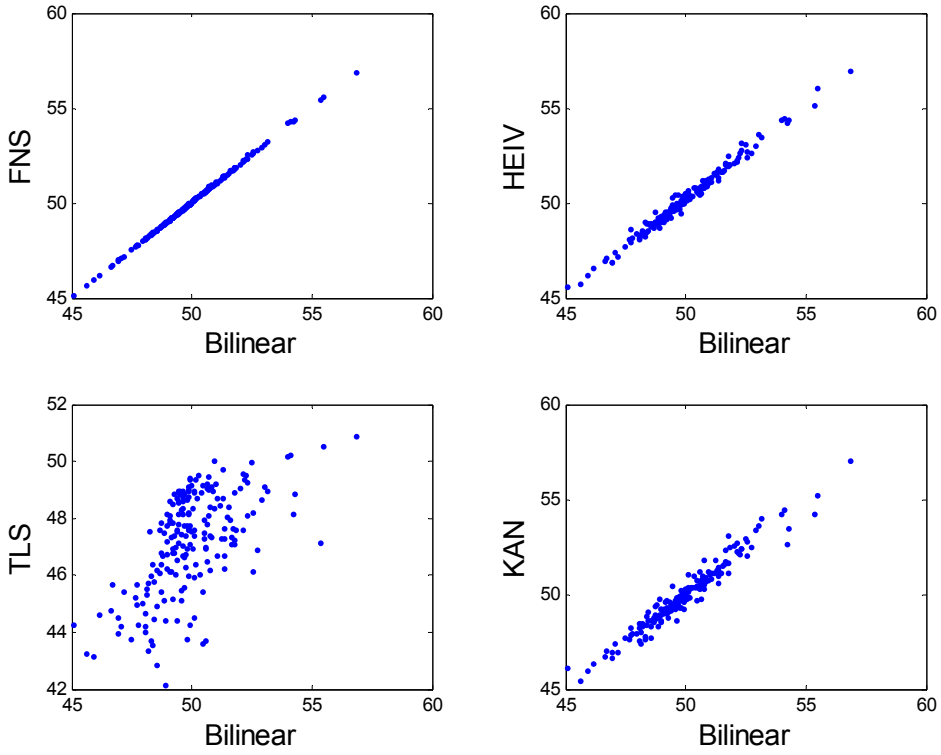
We experimentally find that the error in the five parameters above is not independent. The error in the coordinates of the center and the orientation angle are strongly dependent on the estimates of the major length and the minor length. If both the major length and the minor length are correctly estimated, the estimates of the other three parameters probably have small errors. Because of this observation, we mainly resort to the major length and the minor length in the evaluation of the methods, in the following. In figure 1, we show the performance of the methods, in contrast with that of the Bilinear. We can observe a strong linear correlation between the four approaches: HEIV, FNS, KAN and Bilinear.

Table 1: The statistics of the estimated major length, minor length, x and y coordinates of the center, and the angle between the major axis and the horizontal axis. The ground truth is listed in the first row. For every method, its mean, with its standard deviation in the brackets, is listed in each row. ***Noise =2 over ½ ellipse***

| | Major(100) | Minor(50) | Center_X(0) | Center_Y(0) | Angle(0) |
|---|---|---|---|---|---|
| Bilinear | 100.6332<br>(4.3768) | 49.9533<br>(1.6987) | -0.1382<br>(4.3744) | 0.1164<br>(1.7272) | 0.0218<br>(1.2630) |
| FNS | 100.7461<br>(4.5155) | 49.9794<br>(1.7131) | -0.1540<br>(4.5255) | 0.1194<br>(1.7441) | 0.0176<br>(1.2746) |
| HEIV | 100.8303<br>(4.4745) | 50.1564<br>(1.6956) | -0.1783<br>(4.4852) | 0.1182<br>(1.7454) | 0.0235<br>(1.2878) |
| KAN | 100.6214<br>(4.5938) | 49.9326<br>(1.7042) | -0.1670<br>(4.5990) | 0.0903<br>(1.7457) | 0.0328<br>(1.2966) |
| TLS | 93.6993<br>(4.3251) | 47.1748<br>(1.7675) | 0.2508<br>(7.5858) | -0.3264<br>(3.3746) | 0.4267<br>(2.6456) |

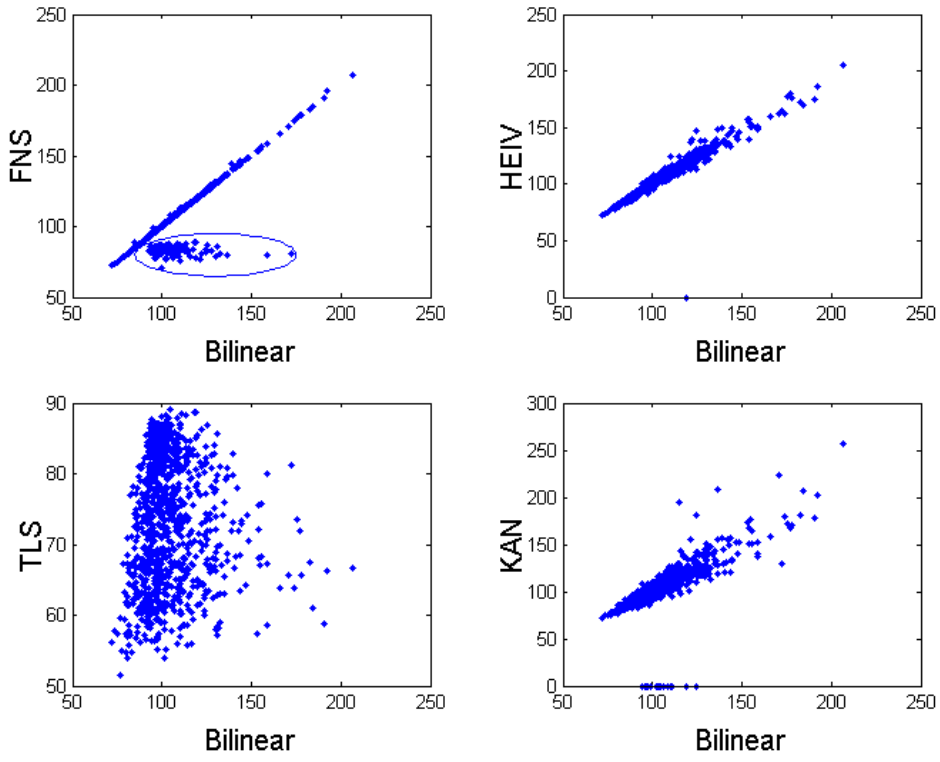



(a)

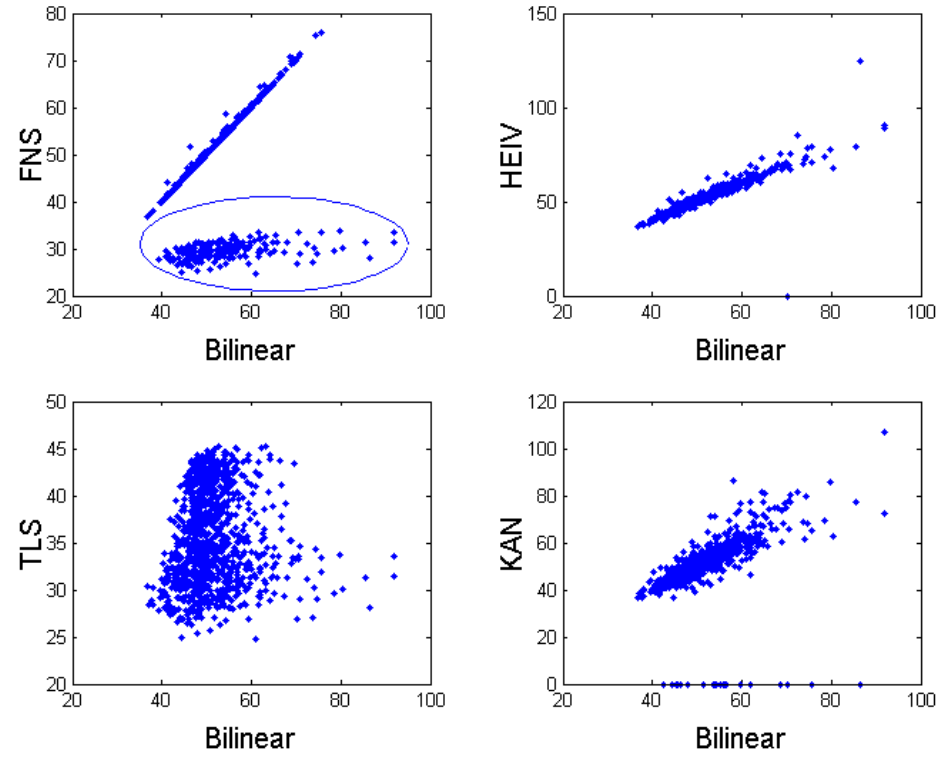



(b)

Figure 1: The comparison of the bilinear approach, with the FNS, HEIV and KAN, TLS. (a) is the estimated major length and (b) is the estimated minor length.




(a)




(b)

Figure 2: See the caption of figure 1. ***Noise =2 over 3/8 ellipse***. In two of the graphs, there are a significant number of "outlier" results that we have highlighted by drawing an enclosing boundary around them.

## 5.2 Noise level=2 over 3/8 ellipse

The next experiment differs from the above experiment in that now the points are randomly distributed on a random 3/8 ellipse. The noise level is still 2. In order to present a better comparison, we repeated the experiment 1000 times to obtain the statistics (listed in table 2).

Table 2: See the caption of table 1. ***noise=2 over 3/8 ellipse***

| | Major(100) | Minor(50) | Center_X(0) | Center_Y(0) | Angel(0) |
|---|---|---|---|---|---|
| Bilinear | 103.6820<br>(16.0952) | 51.0717<br>(6.6344) | 0.4578<br>(16.0868) | 0.2702<br>(6.9479) | -0.0461<br>(3.2988) |
| FNS | 100.3527<br>(17.4383) | 46.9251<br>(9.4859) | 0.4511<br>(16.4996) | 0.2104<br>(9.7672) | 0.0434<br>(3.6045) |
| HEIV | 103.7216<br>(15.5261) | 51.4157<br>(6.8837) | 0.3804<br>(15.4628) | 0.3439<br>(7.2849) | -0.0313<br>(3.3207) |
| KAN | 104.4603<br>(18.5451) | 51.2461<br>(7.3436) | 0.3138<br>(18.5808) | 0.0375<br>(7.8171) | -0.0042<br>(3.5893) |
| TLS | 73.8776<br>(9.0869) | 35.7235<br>(5.0653) | 0.1016<br>(28.0569) | -0.2730<br>(13.6420) | -0.2007<br>(10.5892) |

However, taken alone, the statistics in table 2 do not adequately reflect the performance of the methods. Consider also figure 2 and table 3. We find that the FNS method performs much worse than the HEIV, KAN and Bilinear approaches in some cases, as can be observed in figure 2. (Note, although there are a few cases in the circles in figure 2, where the Bilinear, HEIV and KAN also produce "bad" estimates; in many cases, the Bilinear, HEIV and KAN produce "good" estimates, as can observed in table 3). The problem with the FNS method is that there is no guarantee of convergence. Because of the lack of convergence, in some cases, the FNS stops after the first iteration step, and consequently, its estimate is the same as the initial estimate, which we chose as the TLS estimate for initializing FNS. (This also accounts for the fact that the FNS method produces almost 100% ellipses in the following experiments, which are even more challenging. Note, the TLS always produces an ellipse because the constraint $4de - c^2 = 1$ is enforced in the TLS method.)

To summarize: it is difficult to evaluate the approaches in this setting. This is because one approach scores better in a few cases, while another approach scores better in other different cases. Moreover, as we will see in the more challenging experiments below,

some estimates are wildly wrong, (they may not even be ellipses - except when the special TLS method is employed, enforcing the elliptic constraint). For these reasons, the statistics above, by themselves, can not reliably reflect the performance of the approaches. Worse, the wildly wrong estimates make the statistics, like mean, misleading in assessing the performance of the methods. For example, from the mean of the major length, the FNS is the best method. However, if we examine the figures in figure 2 in detail, we find that FNS actually is worse than the HEIV, KAN and Bilinear methods.

In order to present a meaningful comparison, we mainly resort to the following statistics: for a method, how often does it produce good estimates? As in figure 2, we only use the estimated major length and the estimated minor length in evaluating the performance, because the accuracy of other parameters is strongly dependent on the accuracy of the lengths. More precisely, we regard an estimate as "good" if the error of the estimated major, and minor, lengths fall short of 10% or 20% of the true lengths. In this example, we regard the estimate good if its major length lies in [90,110] or [80,120] and if its minor length lies in [45,55] or [40,60]. From table 3, we can observe that the KAN and FNS methods perform a little worse than HEIV and the proposed Bilinear method.

To provide an indication of our measure, two examples of "good" estimates are shown in figure 3. The good estimates by these four methods are shown in figure 4.

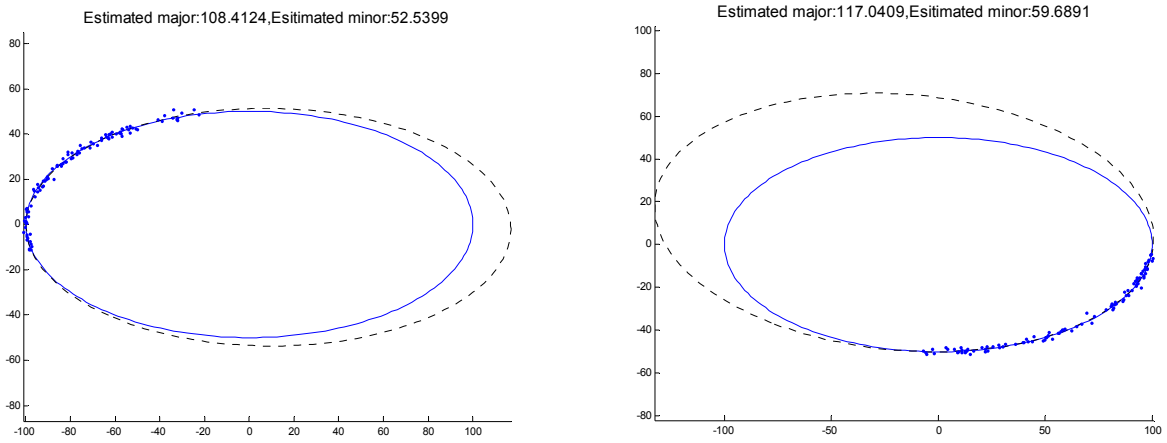



Figure 3: Two example of "good" estimates, falling in the 10% and 20% range.

Table 3: The "good" estimates for ***noise=2 over 3/8 ellipse***. See the definition of "good" estimate in the text.

| | 10% | 20% | Other ellipses | Non-ellipse |
|---|---|---|---|---|
| Bilinear | 532 | 825 | 175 | 0 |
| HEIV | 526 | 820 | 179 | 1 |
| KAN | 470 | 773 | 205 | 22 |
| FNS | 444 | 686 | 314 | 0 |
| LS | 0 | 1 | 999 | 0 |

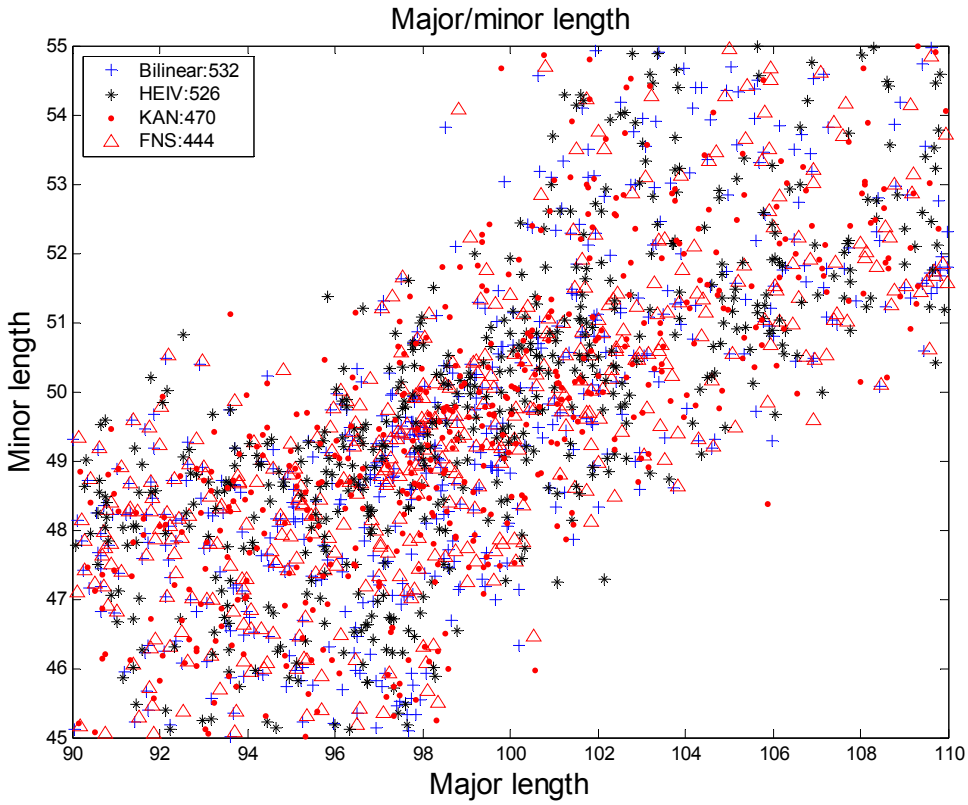



Figure 4: The "good" estimates in 1000 trials of the Bilinear, FNS, HEIV and KAN approaches for ***noise=2 over 3/8 ellipse***. The number after the approaches in the legend is how often the associated approach produces "good" estimates in 1000 trials.

## 5.3 Noise level=1 over a quarter ellipse

In this experiment the noise level is 1 and the points are from a quarter of the ellipse. We also run 1000 trials for this setting. As we discussed above, the statistic of mean and standard deviation are not good indexes for comparing. We only list how often the approaches succeed in producing "good" estimates in 1000 trials.

Table 4: The "good" estimates for ***noise=1 over 1/4 ellipse***. See the definition of "good" estimate in the text.

| | 10% | 20% | Other ellipses | Non-ellipse |
|---|---|---|---|---|
| Bilinear | 229 | 421 | 569 | 10 |
| HEIV | 222 | 425 | 553 | 22 |
| KAN | 188 | 383 | 481 | 136 |
| FNS | 125 | 246 | 752 | 2 |
| TLS | 0 | 0 | 1000 | 0 |

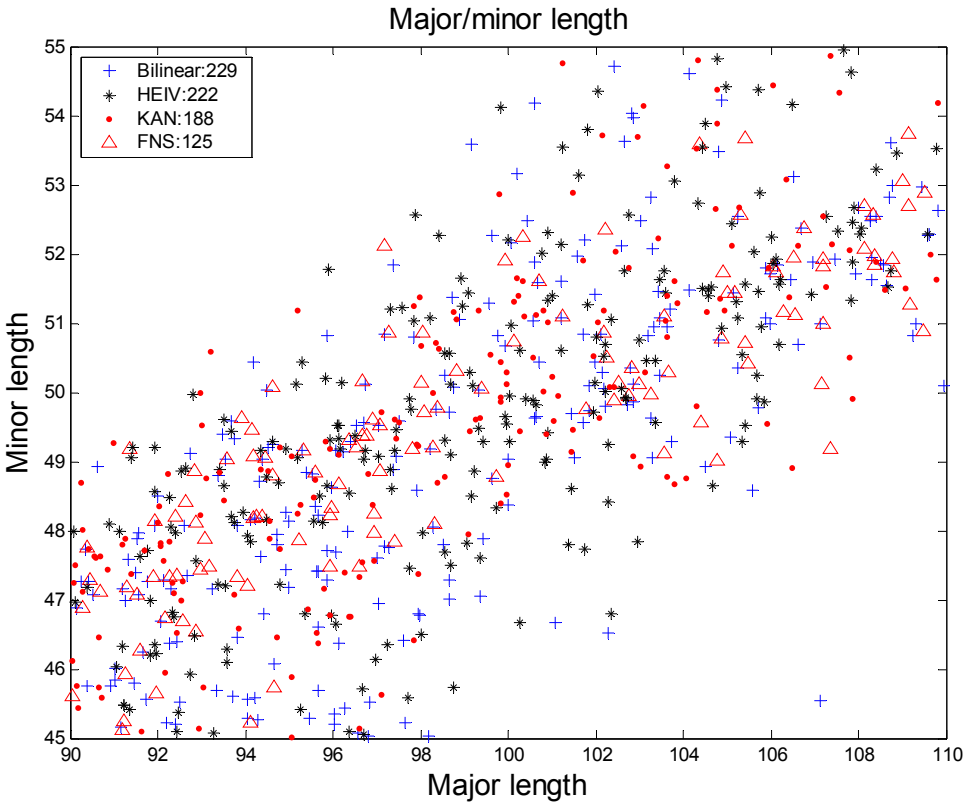



Figure 5: The "good" estimates in 1000 trials of the Bilinear, FNS, HEIV and KAN approaches for ***noise=1 over 1/4 ellipse***. The number after the approaches in the legend is how often the associated approach produces "good" estimates in 1000 trials. The "good" estimates are defined by the 10% range.

Although there is a strong linear correlation between the results produced by the HEIV, FNS, KAN and Bilinear methods in some settings, as can be observed in figure 1 and figure 2; they actually have quite different performance in this more challenging environment.

Note: even though the HEIV and the bilinear methods seem to have a similar performance, in terms of the statistics in table 3 and table 4, they actually have different outputs in many cases. For example, although the Bilinear approach and the HEIV approach produce a similar result, in terms of how often they produce "good" estimates; there are only 142 cases, where both approaches simultaneously produce "good" estimates, falling in the 10% range. This means that, in 80 cases, while the HEIV result falls in the range of 10%, the Bilinear does not. On the other hand, the Bilinear approach produces good estimates in 87 cases, where the HEIV approach does not.

We also comment that, due to the moderately high failure rate, none of these approaches can't be regarded as a solution to the conic fitting problem in the most challenging forms (data over a small arc of the ellipse only). .

## 5.4 Noise level=2 over a quarter ellipse

In this last experiment the noise level is 2 and the points are from a quarter of the ellipse. As in section 5.3, we only list how often the approaches succeed in producing "good" estimates in 1000 trials.

Table 5: The "good" estimates for ***noise=2 over 1/4 ellipse*** . See the definition of "good" estimate in the text.

| | 10% | 20% | Other ellipses | Non-ellipse |
|---|---|---|---|---|
| Bilinear | 88 | 211 | 663 | 126 |
| HEIV | 75 | 179 | 577 | 244 |
| KAN | 29 | 75 | 313 | 612 |
| FNS | 10 | 26 | 965 | 9 |
| TLS | 0 | 0 | 1000 | 0 |

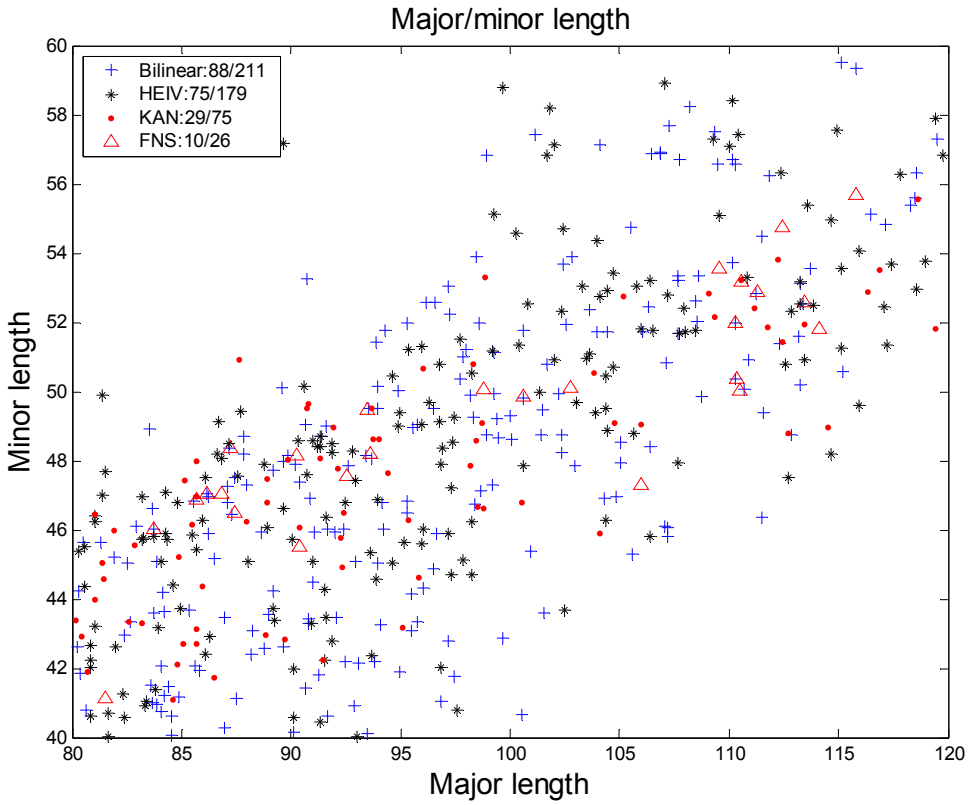



Figure 6: The "good" estimates in 1000 trials of the Bilinear, FNS, HEIV and KAN approaches for ***noise=2 over 1/4 ellipse***. The numbers after the approaches in the legend are how often the associated approach produces "good" estimates in 1000 trials. The "good" estimates are defined by the 10% and 20% range, respectively.

We remark that, only on 19 or 92 cases, out of the 1000 trials, both HEIV and Bilinear produce "good" estimates, in terms of the 10% or 20% ranges, respectively.

## 5.5 Comments on the experimental results

Although convergence to the ML estimate, at least a local optimum estimate, can be ensured in the proposed bilinear approach, as discussed in section 3.4, the results are not so good as expected. From the table 4 and table 5, the bilinear approach can't be regarded as a good solution to the problem, where the points only span a quarter of the ellipse; because it has only about 10-20% success rate of "good" estimates.

There are two possible reasons for this. First, as suggested in section 3.4, only a local optimal solution can be ensured in the iteration process. If the initial estimate deviates

far from the global optimum estimate, the iteration is possibly trapped in other local minimum. In our experiments, we took the TLS result [9] as the initial estimate for the bilinear approach. However, it has been suggested in [23] that the TLS result is not "adequate to be used as an initial solution." Also note that we took the TLS result as the initial estimate for FNS approach. This possibly accounts for the fact that FNS performs worse than HEIV in the challenging settings, as shown in table 4 and table 5.

The second reason, possibly, is due to the specific nature of the conic fitting problem. Also as discussed in section 3.4, the optimal solution, measured by (19) or (20), does not imply the ML estimate, because the ML optimality applies only when the uncertainties in the general measurement matrix **W** are Gaussian. As we analysed in section 4.1, the second order uncertainties are not Gaussian. Strictly, even if we obtain the optimal solution, measured by (19) or (20), it is not the ML estimate.

As pointed out in the experiments, although the results, by FNS, HEIV, and KAN show a strong correlation with the result by the proposed bilinear approach, there are many situations where one method succeeds and other methods fail. There is no clear "safe-bet" in this regards. Because of this, we also have to remark that, although that the bilinear approach outperforms other competing approaches, by different margins, as can be observed in experiments above; we do not claim that the proposed bilinear approach can replace the other approaches.

## 5.6 Question raised

We have stated that the approaches, including FNS, HEIV, KAN and the proposed bilinear approach, can't be regarded as a good solution to the conic fitting problem, when the points only span a small arc of the ellipse. Here, we highlight aspects of the problem from another point of view. Figure 7.a shows a "bad" estimate, whose estimated major and minor lengths are 230 and 80, respectively. In terms of the estimated parameters, this estimated ellipse is wildly wrong, because it is far from the truth. However, from figure 7.b, we can find that the estimated ellipse fits the points very well. If we separate the fitted ellipse and the underlying ground truth, it is very difficult to decide which one fits the noisy points better, as shown in figure 7.c and figure 7.d. This suggests that in many cases it may actually be unreasonable to expect

the true ellipse to be recovered in such extreme cases of only a small fraction of the ellipse containing data.

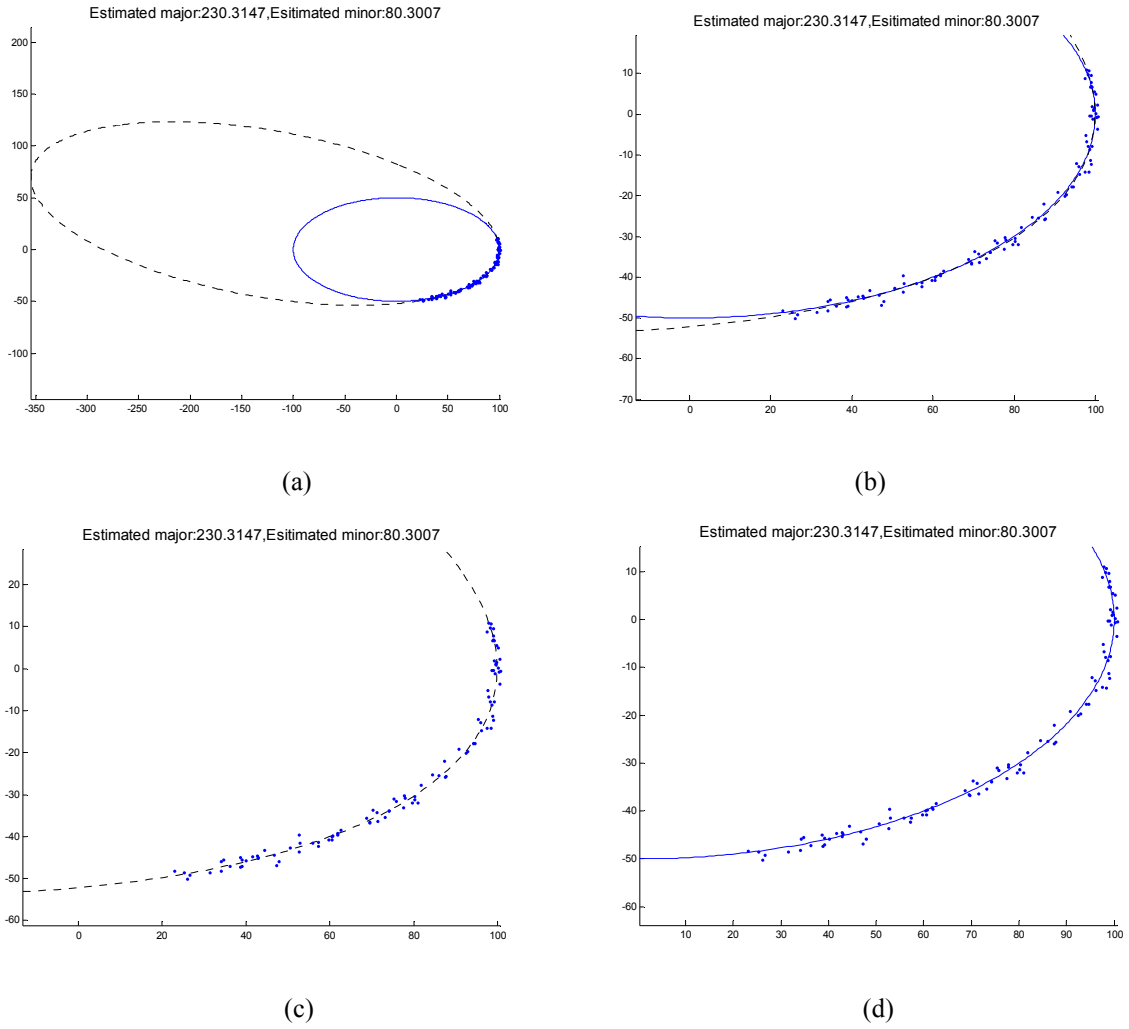



Figure 7: A bad estimate, with details. The solid ellipse is the ground truth, and the dotted ellipse is the estimated ellipse. The dots, '.', are the noisy feature points.

# 6 Conclusion and future work

In this paper, we present a general theory of the parameter estimation problem in a heteroscedastic linear system. This theory suggests a bilinear solution method which we implemented and tested. The method was shown to perform relatively well, and, for ellipse fitting where the data covers a large fraction of the ellipse, the results are good. However, *none* of the methods investigated, including ours, can be considered adequate for fitting data from a small arc of the ellipse. As we illustrated in our concluding section, it is perhaps true that in at least some of the cases where the methods fail, it is unreasonable to expect any method to produce the "true" solution. However, we have no way of making such a notion precise at this stage; and testing the "reasonableness" of the task will be our future work.

# Appendix: Equality Constrained Least Squares

The equality constrained least squares problem is as:

$$\min_{\mathbf{A}_2\mathbf{x}=\mathbf{b}_2} \| \mathbf{A}_1\mathbf{x}-\mathbf{b}_1 \| \tag{A.1}$$

where $\mathbf{A}_1 \in R^{m,n}$, $\mathbf{A}_2 \in R^{p,n}$, $\mathbf{b}_1 \in R^{m,1}$, and $\mathbf{b}_2 \in R^{p,1}$.

Without loss of generality, assume $rank(\mathbf{A}_2)=p$ and $p<n$. Let $\mathbf{Q}^T\mathbf{A}_2^T=\begin{bmatrix}\mathbf{R}\\ \mathbf{0}\end{bmatrix}$ be the QR factorization of $\mathbf{A}_2^T$, where $\mathbf{R}$ is a $p\times p$ upper triangle matrix. Set $\mathbf{A}_1\mathbf{Q}=[\mathbf{P}_1,\mathbf{P}_2]$ and $\mathbf{Q}^T\mathbf{x}=\begin{bmatrix}\mathbf{y}\\ \mathbf{z}\end{bmatrix}$, where $\mathbf{y}\in R^{p,1}$, $\mathbf{z}\in R^{n-p,1}$. With these transformations, (A.1) becomes

$$\min_{\mathbf{R}^T\mathbf{y}=\mathbf{b}_2} \| \mathbf{P}_1\mathbf{y}+\mathbf{P}_2\mathbf{z}-\mathbf{b}_1 \| \tag{A.2}$$

where the vector $\hat{\mathbf{y}}$ can be determined from the constraint $\mathbf{R}^T\mathbf{y}=\mathbf{b}_2$. (A.2) becomes

$$\min_{\mathbf{z}} \| \mathbf{P}_2\mathbf{z}-(\mathbf{b}_1-\mathbf{P}_1\hat{\mathbf{y}}) \| \tag{A.3}$$

which is an unconstrained LS problem. The solution to the equality constrained LS problem (A.1) is:

$$\hat{\mathbf{x}}=\mathbf{Q}\begin{bmatrix}\hat{\mathbf{y}}\\ \hat{\mathbf{z}}\end{bmatrix} \tag{A.4}$$

# References:


1. Aguiar, P.M.Q. and J.M.F. Moura. *Factorization as a rank 1 problem*. in *Proc. Conf. Computer Vision and Pattern Recognition*. 1999.
2. Aguiar, P.M.Q. and J.M.F. Moura. *Weighted factorization*. in *ICIP*. 2000.
3. Aguiar, P.M.Q. and J.M.F. Moura, *Rank 1 weighted factorization for 3D structure recovery: algorithms and performance analysis.* IEEE Trans Pattern Analysis and Machine Intelligence, 2003. **25**(9): p. 1134-1149.
4. Anandan, P. and M. Irani, *Factorization with uncertainty.* Int'l J. Computer vision, 2002. **49**(2/3): p. 101-116.
5. Chen, P. and D. Suter, *Recovering the missing components in a large noisy low-rank matrix: Application to SFM.* IEEE Trans Pattern Analysis and Machine Intelligence, 2004. **26**(8): p. 1051-1063.
6. Chojnacki, W., et al., *On the fitting of surfaces to data with covariances.* IEEE Trans Pattern Analysis and Machine Intelligence, 2000. **22**(11): p. 1294-1303.
7. Chojnacki, W., et al., *Revisiting Hartley's normalized eight-point algorithm.* IEEE Trans Pattern Analysis and Machine Intelligence, 2003. **25**(9): p. 1172 - 1177.
8. Chojnacki, W., et al., *From fns to heiv: a link between two vision parameter estimation methods.* IEEE Trans Pattern Analysis and Machine Intelligence, 2004. **26**(2): p. 264-268.
9. Fitzgibbon, A., M. Pilu, and R.B. Fisher, *Direct least square fitting of ellipses.* IEEE Trans Pattern Analysis and Machine Intelligence, 1999. **21**(5): p. 476-480.
10. Golub, G.H. and C.F.V. Loan, *Matrix Computations*. 3nd ed. 1996, Baltimore: Johns Hopkins University Press.
11. Guerreiro, R.F.C. and P.M.Q. Aguiar. *Estimation of Rank Deficient Matrices from Partial Observations: Two-Step Iterative Algorithms*. in *Energy Minimization Methods in Computer Vision and Pattern Recognition*. 2003.
12. Hartley, R. and A. Zisserman, *Multiple view geometry in computer vision*. 2000: Cambridge University Press.
13. Huffel, S.V. and J. Vandewalle, *The total least squares problem: computational aspects and analysis*. 1991: SIAM.
14. Irani, M. and P. Anandan. *Factorization with uncertainty*. in *Proc. European Conf. Computer Vision*. 2000.
15. Jacobs, D. *Linear fitting with missing data: Applications to structure-from-motion and to characterizing intensity images*. in *Proc. Conf. Computer Vision and Pattern Recognition*. 1997.
16. Jacobs, D., *Linear fitting with missing data for structure-from-motion.* Computer Vision and Image Understanding, 2001. **82**: p. 57-81.
17. Kahl, F. and A. Heyden. *Structure and motion from points, lines and conics with affine cameras*. in *Proc. European Conf. Computer Vision*. 1998.
18. Kahl, F. and A. Heyden, *Affine structure and motion from points, lines and conics.* Int'l J. Computer vision, 1999. **33**(3): p. 163-180.
19. Kanatani, K., *Unbiased estimation and statistical analysis of 3-D rigid motion from two views.* IEEE Trans Pattern Analysis and Machine Intelligence, 1993. **15**(1): p. 37 - 50.
20. Kanatani, K., *Statistical bias of conic fitting and renormalization.* IEEE Trans Pattern Analysis and Machine Intelligence, 1994. **16**(3): p. 320 - 326.

21. Kanatani, K., *Statistical Optimi zation for Geometric Computation: Theory and Practice*. 1996: Amsterdam: Elsevier.
22. Leedan, Y. and P. Meer. *Estimation with bilinear constraints in computer vision*. in *Proc. Int'l Conf. Computer Vision*. 1999.
23. Leedan, Y. and P. Meer, *Heteroscedastic Regression in Computer Vision: Problems with Bilinear Constraint.* Int'l J. Computer vision, 2000. **37**(2): p. 127-150.
24. Mahamud, S., et al. *Provably-Convergent Iterative Methods for Projective Structure from Motion*. in *Proc. Conf. Computer Vision and Pattern Recognition*. 2001.
25. Matei, B. and P. Meer. *Reduction of bias in maximum likelihood ellipse fitting*. in *ICPR*. 2000.
26. Matei, B. and P. Meer. *A general method for Errors-in-Variables problems in computer vision*. in *Proc. Conf. Computer Vision and Pattern Recognition*. 2000.
27. Morris, D.D. and T. Kanade. *A unified factorization algorithm for points, line segments and planes with uncertainty models*. in *Proc. Int'l Conf. Computer Vision*. 1998.
28. Press, W.H., et al., *Numerical recipes in C*. 2nd ed. 1992: Cambridge University Press.
29. Reid, I.D. and D.W. Murray, *Active tracking of foveated feature clusters using affine structure.* Int'l J. Computer vision, 1996. **18**(1): p. 41-60.
30. Shum, H., K. Ikeuchi, and R. Reddy, *Principal component analysis with missing data and its applications to polyhedral object modeling.* IEEE Trans Pattern Analysis and Machine Intelligence, 1995. **17**(9): p. 854-867.
31. Vidal, R. and R. Hartley. *Motion segmentation with missing data using PowerFactorization and GPCA*. in *Proc. Conf. Computer Vision and Pattern Recognition*. 2004.
32. Werman, M. and Z. Geyzel, *Fitting a Second Degree Curve in the Presence of Error.* IEEE Trans Pattern Analysis and Machine Intelligence, 1995. **17**(2): p. 207-211.
33. Zhang, Z., *parameter estimation techniques: A tutorial with application to conic fitting.* Image and Vision Computing, 1997. **15**: p. 59-76.