Weighted Total Least Squares for Quadratic Errors-in-Variables Regression

Peng Liu, Kailai Li, Gustaf Hendeby, Fredrik Gustafsson

***Abstract*—In this paper, we present a study on using weighted total least squares method for parameter estimation of errors-in-variables models with quadratic regressors. The statistics of error is analyzed to fill in the gap between basic assumptions in weighted total least squares and our case. A modified Cramér-Rao lower bound is introduced for error quantification in the proposed method. We perform evaluations based on simulations with comparisons to standard least squares and generalized total least squares. Numerical results show that the proposed method outperforms the others in terms of estimation accuracy.**

I. Introduction

Parameter estimation is a fundamental problem in statistical signal processing and has been frequently investigated in ubiquitous application scenarios, such as localization, sensor calibration and epidemic modeling [1]–[3]. The basic structure of the problem is established upon linear formulations with fixed regressor matrix and measurements under noise. In this case, the *least squares* (LS) method is the off-the-shelf solution to parameter estimation and corresponding statistical analysis [4]. In practice, however, the regressor matrix is obtained via observations that are corrupted with noise. This leads to the so-called *errors-in-variables* (EIV) problem, and estimates delivered by the standard least square are in general biased, and statistical analysis is often unreliable.

To mitigate this issue, the *total least squares* (TLS) can be deployed as an alternative, with estimates given in closed form via singular value decomposition [5], [6]. One limitation of the approach is that noise terms in the regressor matrix and measurements are required to be zero-mean and independent and identically distributed, which is hardly feasible in reality. In the cases where regressor matrices are of certain special structures, e.g., with diagonal-constant form, the *structured total least squares* (STLS) method can be applied [7], [8]. Another common case for simplification appears when some columns in the regressor matrix are free of error, EIV problems can then be solved through the *generalized total least squares* (GTLS) [9].

Towards EIV regression of general setup, one promising solution is the *weighted total least squares* (WTLS). In [10], the WTLS problem was applied with Lagrange multipliers to handle noises of Kronecker product covariance structure. Another variant was proposed in [11] to cope with fully correlated noise covariance, where both constrained and unconstrained formulations for optimization were investigated. It is shown in [12] that the structure of the regressor matrix can be preserved in its error term provided that the covariance matrix satisfies certain basic conditions. This insight was further generalized in [13], where the constraints on the regressor matrices were further relaxed.

The aforementioned WTLS variants have been widely exploited, in particular, in geodesic-related application scenarios [14]–[17]. However, existing works have only referred to EIV problems with regressor matrices of linear dependencies. In terms of nonlinear EIV regressions, a typical measure is to linearize the nonlinear elements in the regressor matrix for adaptation to a linear setup, such that WTLS can be applied in its standard formulation.

In order to quantify the uncertainty of EIV estimates, the *Cramér-Rao lower bound* (CRLB) has been investigated to provide a theoretical reference for evaluation. One study can be found in [18], where the CRLB is provided for linear EIV models under Gaussian-distributed observation noise. As for nonlinear EIV regressions, however, the Gaussian assumption no longer holds, and a closed-form expression of CRLB is in general infeasible. By marginalizing over the observed signal, methods have been proposed in this regard based on numerical integration or approximation [19], [20].

Contribution

In this work, we investigate the general solution to nonlinear errors-in-variables regression problem with regressor matrices of quadratic structure. For that, we adopt weighted total least squares method with moment analysis on the error term of the quadratic component. For evaluation of the WTLS estimates in nonlinear EIV regression, a principled study on CRLB is further provided based on numerical integration. The proposed WTLS-based approach is compared to ordinary least squares and generalized total least squares methods based on simulation. Numerical results show that the WTLS-based method delivers superior accuracy over LS and GTLS.

The remainder of the paper is structured as follows. Problem formulation of general EIV regression is introduced in Sec. II. The WTLS method is adapted to quadratic EIV models in Sec. III with a dedicated discussion on corresponding statistical properties. Further, a modified version of the CRLB is provided in Sec. IV. The proposed techniques are validated based on simulations in Sec. V, and the work is concluded in Sec. VI.

The authors are with the Division of Automatic Control, Department of Electrical Engineering, Linköping University, Linköping, Sweden. The work is partially funded by the Swedish Research Council under grant Scalable Kalman Filters. Email: {peng.liu, kailai.li, gustaf.hendeby, fredrik.gustafsson}@liu.se

II. Problem Formulation

Throughout the paper, we consider the nonlinear EIV regression model of the following form

$$y = \underline{A}(x^*)\underline{\theta} + w_y \,, \tag{1}$$

where $\underline{A}(x^*)$ is a $1 \times d$ row vector. $y \in \mathbb{R}$ is measured variable with noise. In general, $\underline{A}(x^*)$ denotes a nonlinear vector-valued function of a scalar x^*, while $\underline{\theta}$ contains the unknown parameters to be estimated. Further, w_y is *additive white Gaussian noise* (AWGN) with variance of $\sigma^2_{w_y}$, and x^* denotes noise-free data for the regressor with distribution $x^* \sim \mathcal{N}(\mu_{x^*}, \sigma^2_{x^*})$. In practice, only an observation x of x^* is available following

$$x = x^* + w_x \,, \tag{2}$$

where w_x is AWGN of variance $\sigma^2_{w_x}$. This configuration gives us an EIV model, because both the measurement y and the regressor $\underline{A}(x)$ are affected by noise. To simplify this problem, we assume that w_y and w_x are independent to each other, and $\sigma^2_{w_y}$ and $\sigma^2_{w_x}$ are known, but not necessarily to be equal. To simplify the discussion, we can assume that x^* and w_x are independent. The ordinary least squares method delivers biased estimate in this case due to the measurement noise in x. Total least squares and its variants shall be used instead. Consider N samples, then

$$\begin{aligned} \underline{y} &= [\, y_1, y_2, \ldots, y_N \,]^\top \in \mathbb{R}^N \,, \\ \underline{x} &= [\, x_1, x_2, \ldots, x_N \,]^\top \in \mathbb{R}^N \,, \\ \mathbf{A}(x) &= [\, \underline{A}^\top(x_1), \underline{A}^\top(x_2), \ldots, \underline{A}^\top(x_N) \,]^\top \in \mathbb{R}^{N\times d} \,. \end{aligned} \tag{3}$$

The subscript indicates samples. The model parameter $\underline{\theta}$ should be estimated given the noise-corrupted measurements $\underline{y}$ and $\mathbf{A}(x)$.

III. Proposed Method

In this section, we discuss the details about how to use WTLS to solve the regression problem in (1) and (2). We introduce the principle of WTLS in the first subsection, and adapt the WTLS to the quadratic EIV problem afterward.

A. Weighted Total Least Squares

The EIV problem based on (1) and (2) with N samples is given by

$$\underline{y} + \underline{e}_y = (\mathbf{A} + \mathcal{E}_\mathbf{A})\underline{\theta} \,, \tag{4}$$

where $\mathcal{E}_\mathbf{A}$ is the correction term for $\mathbf{A}$. $\underline{e}_y$ is the correction term for $\underline{y}$. It can be verified that (4) is equivalent to

$$\mathbf{B}\underline{e} + \mathbf{B}\underline{Y} = \underline{0}_N \,, \tag{5}$$

where $\underline{e} = \mathbf{vec}([\, \mathcal{E}_\mathbf{A}, \underline{e}_y \,]) \in \mathbb{R}^{N(d+1)}$ of covariance $\mathbf{Q}$ (concretized in Sec. III-B). $\mathbf{vec}(\cdot)$ denotes the vectorization of a matrix, and $\underline{0}_N$ is a N-dimensional zero vector. $\underline{Y} = \mathbf{vec}([\, \mathbf{A}, \underline{y} \,])$, which composes all observations. $\mathbf{B} = [\, \underline{\theta}^\top \otimes \mathbf{I}, -\mathbf{I} \,]$, with $\otimes$ denoting the Kronecker product and $\mathbf{I}$ the identity matrix with proper dimension. If $\underline{e}$ is zero-mean, the WTLS problem is to solve the following constrained optimization problem

$$\min_{\{\underline{\theta}, \underline{e}\}} \{ \underline{e}^\top \mathbf{P} \underline{e} \} \,, \quad \text{s.t.} \quad \mathbf{B}(\underline{e} + \underline{Y}) = \underline{0}_N \,, \tag{6}$$

where matrix $\mathbf{P} = \mathbf{Q}^{-1}$. Solving (6) cannot be done in closed form. The method of the Lagrange multipliers can be applied, with both the parameter $\underline{\theta}$ and the correction vector $\underline{e}$ estimated iteratively. The Lagrange function of (6) is given as

$$L(\underline{e}, \underline{\theta}, \underline{\lambda}) = \underline{e}^\top \mathbf{P} \underline{e} + 2\underline{\lambda}^\top (\mathbf{B}\underline{e} + \mathbf{B}\underline{Y}) \,, \tag{7}$$

where $\underline{\lambda}$ contains the Lagrange multipliers. By setting the partial derivative of $L(\underline{e}, \underline{\theta}, \underline{\lambda})$ with respect to $\underline{e}$, $\underline{\theta}$ and $\underline{\lambda}$ to zero, we get the following equations

$$\begin{aligned} \frac{\partial L}{\partial e}\Big|_{\hat{\underline{e}}, \hat{\underline{\lambda}}, \hat{\underline{\theta}}} &= 2\mathbf{P}\hat{\underline{e}} + 2\hat{\mathbf{B}}^\top \hat{\underline{\lambda}} = \underline{0}_{N(d+1)} \,, \\ \frac{\partial L}{\partial \lambda}\Big|_{\hat{\underline{e}}, \hat{\underline{\lambda}}, \hat{\underline{\theta}}} &= 2\hat{\mathbf{B}}\hat{\underline{e}} + 2\hat{\mathbf{B}}\underline{Y} = \underline{0}_N \,, \\ \frac{\partial L}{\partial \underline{\theta}}\Big|_{\hat{\underline{e}}, \hat{\underline{\lambda}}, \hat{\underline{\theta}}} &= -2\mathbf{A}^\top \hat{\underline{\lambda}} - 2\hat{\mathcal{E}}_\mathbf{A}^\top \hat{\underline{\lambda}} = \underline{0}_d \,, \end{aligned} \tag{8}$$

respectively. Here, $\hat{(\cdot)}$ denotes an estimate. An iterative solution to (6) by solving (8) follows

$$\begin{aligned} \hat{\mathbf{B}}^{i+1} &= [\, (\hat{\underline{\theta}}^i)^\top \otimes \mathbf{I}, -\mathbf{I} \,] \,, \\ \hat{\underline{\lambda}}^{i+1} &= \big(\hat{\mathbf{B}}^{i+1}\mathbf{Q}(\hat{\mathbf{B}}^{i+1})^\top\big)^{-1}(\mathbf{A}\hat{\underline{\theta}}^i - \underline{y}) \,, \\ \hat{\underline{e}}^{i+1} &= -\mathbf{Q}(\hat{\mathbf{B}}^{i+1})^\top \hat{\underline{\lambda}}^{i+1} \,, \\ \hat{\underline{\theta}}^{i+1} &= (\mathbf{C}^{i+1})^{-1}\mathbf{D}^{i+1}(\underline{y} + \hat{\mathcal{E}}_A^{i+1}\hat{\underline{\theta}}^i) \,, \end{aligned} \tag{9}$$

where the matrices $\mathbf{C}^{i+1}$ and $\mathbf{D}^{i+1}$ are given by

$$\begin{aligned} \mathbf{C}^{i+1} &= (\mathbf{A} + \hat{\mathcal{E}}_\mathbf{A}^{i+1})^\top \big(\hat{\mathbf{B}}^{i+1}\mathbf{Q}(\hat{\mathbf{B}}^{i+1})^\top\big)^{-1}(\mathbf{A} + \hat{\mathcal{E}}_\mathbf{A}^{i+1}) \,, \\ \mathbf{D}^{i+1} &= (\mathbf{A} + \hat{\mathcal{E}}_\mathbf{A}^{i+1})^\top \big(\hat{\mathbf{B}}^{i+1}\mathbf{Q}(\hat{\mathbf{B}}^{i+1})^\top\big)^{-1} \,. \end{aligned}$$

Here, $\hat{\mathcal{E}}_\mathbf{A}^{i+1}$ can be extracted from $\hat{\underline{e}}^{i+1}$ in the corresponding position and reshaped to matrix form. Given that $\underline{e}$ is zero-mean with a known covariance matrix, the WTLS problem of (6) can be solved iteratively using equations in (9). We provide the pseudo-code summarizing the derivation above in Alg.1. The initial value $\hat{\underline{\theta}}^0$ of the parameter can be obtained via standard least squares. Note that the algorithm works as long as the matrix $\hat{\mathbf{B}}\mathbf{Q}\hat{\mathbf{B}}^\top$ is non-singular, even if we have a singular covariance matrix $\mathbf{Q}$. This is a quite common case that can occur in practice, and an example follows.

B. Statistical Analysis of Error in Regressor Matrix

In general, the regressor $\underline{A}(x)$ in (1) involves nonlinear functions of the observation x. In this case, analyzing statistical property for the error term is nontrivial. In order to showcase the general structure of analysis, we now assume that elements in $\underline{A}(x)$ follow a quadratic form of x. Extensions to other polynomial cases are straightforward based on the analytical moment computation of Gaussian random variables. The regressor matrix of a quadratic EIV model follows

$$\underline{A}(x) = [\, 1, \, x, \, x^2 \,] \,. \tag{10}$$

Algorithm 1: WTLS-Based EIV regression

Input : $\hat{\underline{\theta}}^0$, $\mathbf{Q}$, $\mathbf{A}$, $\underline{y}$, threshold ϵ
Output: $\hat{\underline{\theta}}$
$i \leftarrow 1$;
while $\eta < \epsilon$ **do**
  $\hat{\mathbf{B}}^i \leftarrow [\,(\hat{\underline{\theta}}^{i-1})^\top \otimes \mathbf{I}, -\mathbf{I}\,]$;
  $\hat{\underline{\lambda}}^i \leftarrow (\hat{\mathbf{B}}^i \mathbf{Q} (\hat{\mathbf{B}}^i)^\top)^{-1} (\mathbf{A} \hat{\underline{\theta}}^{i-1} - \underline{y})$;
  $\hat{\underline{e}}^i \leftarrow -\mathbf{Q} (\hat{\mathbf{B}}^i)^\top \hat{\underline{\lambda}}$;
  $\mathbf{C}^i \leftarrow (\mathbf{A} + \hat{\mathcal{E}}^i_{\mathbf{A}})^\top (\hat{\mathbf{B}}^i \mathbf{Q} (\hat{\mathbf{B}}^i)^\top)^{-1} (\mathbf{A} + \hat{\mathcal{E}}^i_{\mathbf{A}})$;
  $\mathbf{D}^i \leftarrow (\mathbf{A} + \hat{\mathcal{E}}^i_{\mathbf{A}})^\top (\hat{\mathbf{B}}^i \mathbf{Q} (\hat{\mathbf{B}}^i)^\top)^{-1}$;
  $\hat{\underline{\theta}}^i \leftarrow (\mathbf{C}^i)^{-1} \mathbf{D}^i (\hat{\mathbf{B}}^i \mathbf{Q} (\hat{\mathbf{B}}^i)^\top)^{-1} (\mathbf{y} + \hat{\mathcal{E}}^i_{\mathbf{A}} \hat{\underline{\theta}}^{i-1})$;
  $\eta \leftarrow \| \underline{\theta}^i - \underline{\theta}^{i-1} \|$;
end
$\hat{\underline{\theta}} \leftarrow \hat{\underline{\theta}}^i$;

According to (5), the measurement correction term $\underline{e}_y$ is assumed to be zero-mean with variance $\sigma^2_{w_y}$. For brevity, we also assume that the first column in $\mathbf{A}$ is composed of constants that are noise-free. Thus, the first column of the regressor correction matrix $\mathcal{E}_{\mathbf{A}}$ is zero-valued. We now concatenate the correction terms for regressor and measurement as

$$[\,\mathcal{E}_{\mathbf{A}}, e_y\,] = [\,\underline{0}_N\,, \underline{e}_x, \underline{e}_{x^2}, \underline{e}_y\,]\,. \tag{11}$$

The non-conventional part in (11) is the third column, which refers to the error for the quadratic term. To simplify the notation for analysis, we now derive the error term w.r.t. one observation sample, where

$$x^2 = (x^* + w_x)^2 = (x^*)^2 + w_x^2 + 2x^* w_x\,.$$

Thus, the whole error term for the quadratic terms is given by

$$w_{x^2} = x^2 - (x^*)^2 = w_x^2 + 2x^*\, w_x\,. \tag{12}$$

Consequently, the mean and covariance of w_{x^2} follow

$$\begin{aligned} \mu_{w_{x^2}} &= \mathbf{E}(w_x^2 + 2x^* w_x) = \sigma^2_{w_x} \quad \text{and} \\ \sigma^2_{w_{x^2}} &= \mathbf{var}(w_x^2 + 2x^* w_x) = 2\sigma^4_{w_x} + 4\sigma^2_{x^*}\sigma^2_{w_x}\,, \end{aligned} \tag{13}$$

respectively. After computing the variance for each sample, it is straightforward to extend it to vector form given that the error for each sample is independent from each other.

On the other hand, the mean value of the correction matrix $\left[\mathcal{E}_{\mathbf{A}}, \underline{e}_y\right]$ can be computed as

$$\mathbf{E}\big([\,\mathcal{E}_{\mathbf{A}}, \underline{e}_y\,]\big) = [\,\underline{0}_N, \underline{0}_N, \sigma^2_{w_x}\underline{1}_N, \underline{0}_N\,]\,, \tag{14}$$

with $\underline{1}_N \in \mathbb{R}^N$ denoting a vector of ones. The next step is to compute the covariance matrix of $\underline{e} = \mathbf{vec}([\mathcal{E}_{\mathbf{A}}, \underline{e}_y])$. As we discussed in (11), there are four columns in $[\mathcal{E}_{\mathbf{A}}, \underline{e}_y]$ and the first column is zero-valued. Thus, we only need to compute the covariance matrix for the last three columns. Computing covariance matrices for $\underline{e}_x$ and $\underline{e}_y$ is straightforward, and their cross-covariance matrix is $\mathbf{0}_N$. As for the covariance matrix of $\underline{e}_{x^2}$, we exploit the covariance of the observation noise term w_{x^2} based on the statistical property given in (13). This follows $\mathbf{cov}(\underline{e}_{x^2}) = \sigma^2_{w_{x^2}}\mathbf{I}_N$, where $\mathbf{I}_N$ is an $N \times N$ identity matrix. The cross-covariance between $\underline{e}_{x^2}$ and $\underline{e}_y$ is also $\mathbf{0}_N$. The last part is the cross-covariance matrix between the second and the third column in (11). To simplify the notation, we consider the scalar case in the cross-covariance calculation that can be expressed as follows

$$\mathbf{cov}(e_x, e_{x^2}) = \mathbf{cov}(w_x, w_{x^2}) = 2\mathbf{E}(x^* w_x^2) = 2\mu_{x^*}\sigma^2_{w_x}\,.$$

Based on the derivations above, covariance of the correction matrix in (11) can be obtained in the following form $\mathbf{Q} =$

$$\begin{bmatrix} \mathbf{0}_N & \mathbf{0}_N & \mathbf{0}_N & \mathbf{0}_N \\ \mathbf{0}_N & \sigma^2_{w_x}\mathbf{I}_N & 2\sigma^2_{w_x}\mu_{x^*}\mathbf{I}_N, & \mathbf{0}_N \\ \mathbf{0}_N & 2\sigma^2_{w_x}\mu_{x^*}\mathbf{I}_N & 2\sigma^4_{w_x}\mathbf{I}_N + 4\sigma^2_{w_x}\sigma^2_{x^*}\mathbf{I}_N & \mathbf{0}_N \\ \mathbf{0}_N & \mathbf{0}_N & \mathbf{0}_N & \sigma^2_{w_y}\mathbf{I}_N \end{bmatrix} \tag{15}$$

which is deployed to the WTLS method summarized in Alg.1.

C. Implementation

So far, we have computed the mean and covariance matrix for the complete correction matrix in (11). The next step is to compensate the bias term and construct the WTLS problem in accordance with Section III-A. Because of the non-zero mean in (14), we first compensate the mean for the measurement data $\underline{Y}$ in (5) to get the unbiased measurement $\underline{Y}^c$ as follows

$$\underline{Y}^c = \underline{Y} - \mathbf{vec}\big(\mathbf{E}\big([\,\mathcal{E}_{\mathbf{A}}, \underline{e}_y\,]\big)\big)\,. \tag{16}$$

The bias compensated measurement $\underline{Y}^c$ is corrupted by zero-mean noise, which satisfies the setup required by WTLS. Once bias compensation is done, we can exploit the WTLS algorithm in Alg. 1 for quadratic EIV regression by concretizing the elements in covariance matrix $\mathbf{Q}$ in (15).

IV. Cramér-Rao Lower Bound

Computing the theoretical lower bound in terms of the *root mean square error* (RMSE) is necessary for most regression problems for the sake of evaluation. Therefore, we investigate the CRLB of (1) considering (2) with the quadratic example in (10) in this section. In EIV configuration, we need to estimate both the model parameter and the correction in the regressor matrix. It means that both $\underline{\theta}$ and $\underline{x}^*$ need to be estimated given the noise corrupted measurement $\underline{y}$ and $\underline{x}$. As shown in (2), x^* is assumed to be a Gaussian random variable that can be treated as a nuisance parameter besides the deterministic parameter $\underline{\theta}$. In this regard, the so-called *modified Cramér-Rao lower bound* proposed in [19] is applicable as an alternative to conventional CRLB. Given the random parameter $\underline{x}^*$, we denote the conventional CRLB of the estimate as $\mathbf{J}(\underline{\theta}|\underline{x}^*)$. Based thereon, the modified CRLB is given as

$$\mathbf{MJ}(\underline{\theta}) = \mathbf{E}_{x^*}(\mathbf{J}(\underline{\theta}|\underline{x}^*))\,, \tag{17}$$

where the conventional CRLB follows

$$\mathbf{J}(\underline{\theta}|\underline{x}^*) = -\left[\mathbf{E}_{\underline{y},\underline{x}}\left(\frac{\partial^2 \mathcal{L}(\underline{\theta}, \underline{x}^*)}{\partial\underline{\theta}\partial\underline{\theta}^\top}\right)\right]^{-1}\,, \tag{18}$$

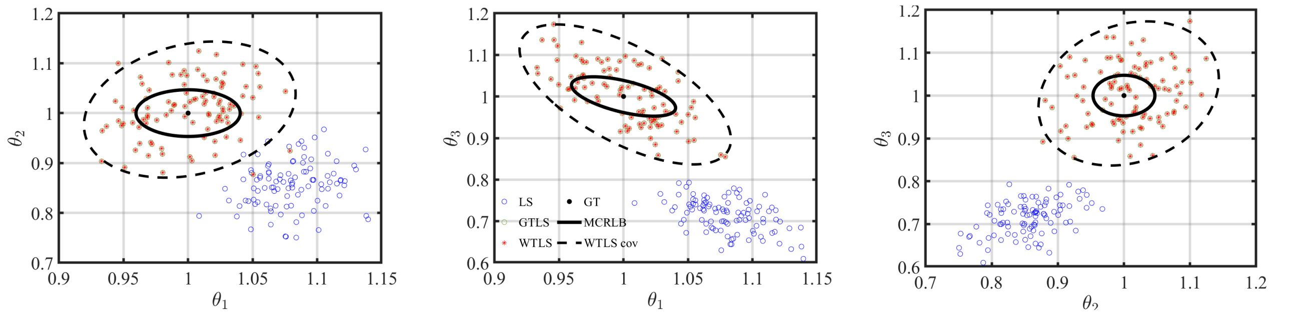


Fig. 1: Quadratic EIV regression in scenario 1 of Sec. V with deterministic constant term in regressor matrix. Compared with the ground truth (GT), the proposed WTLS resembles the GTLS with better accuracy than LS. The dashed and solid black ellipses indicate the uncertainty estimate (95% confidence interval) given by WTLS and modified CRLB, respectively.

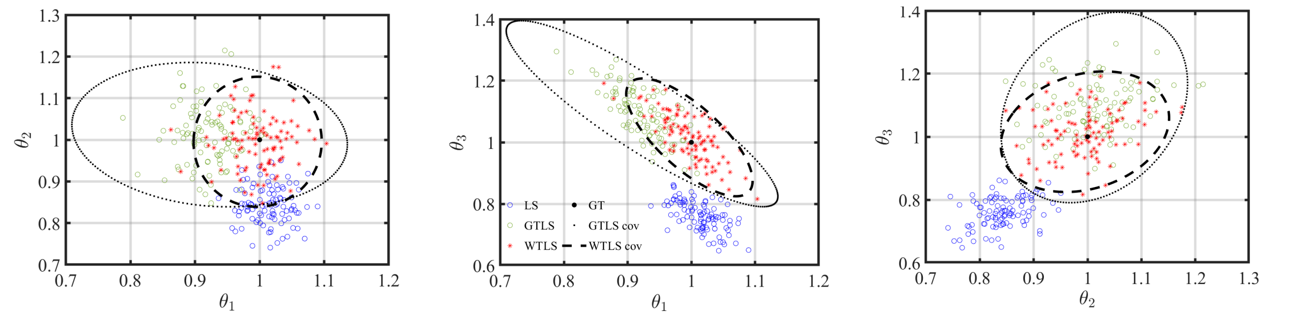


Fig. 2: Quadratic EIV regression in scenario 1 of Sec. V with uncertain constants in regressor matrix. The proposed WTLS outperforms both GTLS and LS.

with $\mathcal{L}(\underline{\theta}, \underline{x}^*) = \log P(\underline{y}, \underline{x}|\underline{\theta}, \underline{x}^*)$. Based on the assumption of independence among different observations, the log-likelihood function $\mathcal{L}(\underline{\theta}, \underline{x}^*)$ can be derived as

$$\begin{aligned} \mathcal{L}(\underline{\theta}, \underline{x}^*) &= \log P(\underline{y}|\underline{\theta}, \underline{x}^*) + \log P(\underline{x}|\underline{x}^*) \\ &= \sum_{i=1}^{N} \log P(y_i|\underline{\theta}, x_i^*) + \sum_{i=1}^{N} \log P(x_i|x_i^*)\,. \end{aligned} \tag{19}$$

To obtain the value of the conventional CRLB as shown in (18), we need to calculate the second-order derivative of $\mathcal{L}(\underline{\theta}, \underline{x}^*)$ with respect to $\underline{\theta}$. This only refers to the first term in (19). According to (1), we have $P(y_i|\underline{\theta}, x_i^*) \sim \mathcal{N}(\underline{A}(x_i^*)\underline{\theta}, \sigma_{w_y}^2)$. Thus, the second order derivative can be computed analytically as follows

$$\begin{aligned} \frac{\partial^2 \mathcal{L}(\underline{\theta}, \underline{x}^*)}{\partial \underline{\theta}\, \underline{\theta}^\top} &= -\frac{1}{\sigma_{w_y}^2} \sum_{i=1}^{N} \underline{A}(x_i^*)^\top \underline{A}(x_i^*) \\ &= -\frac{N}{\sigma_{w_y}^2} \begin{bmatrix} 1 & \overline{x^*} & \overline{(x^*)^2} \\ \overline{x^*} & \overline{(x^*)^2} & \overline{(x^*)^3} \\ \overline{(x^*)^2} & \overline{(x^*)^3} & \overline{(x^*)^4} \end{bmatrix}, \end{aligned} \tag{20}$$

with $\overline{(x^*)^k} = \frac{1}{N}\sum_{i=1}^{N} (x_i^*)^k$. (20) is independent of $\underline{x}$ and $\underline{y}$, which results in

$$\mathbf{J}(\underline{\theta}|\underline{x}^*) = \frac{\sigma_{w_y}^2}{N} \begin{bmatrix} 1 & \overline{x^*} & \overline{(x^*)^2} \\ \overline{x^*} & \overline{(x^*)^2} & \overline{(x^*)^3} \\ \overline{(x^*)^2} & \overline{(x^*)^3} & \overline{(x^*)^4} \end{bmatrix}^{-1}. \tag{21}$$

Computing the modified CRLB in (18) follows the integration

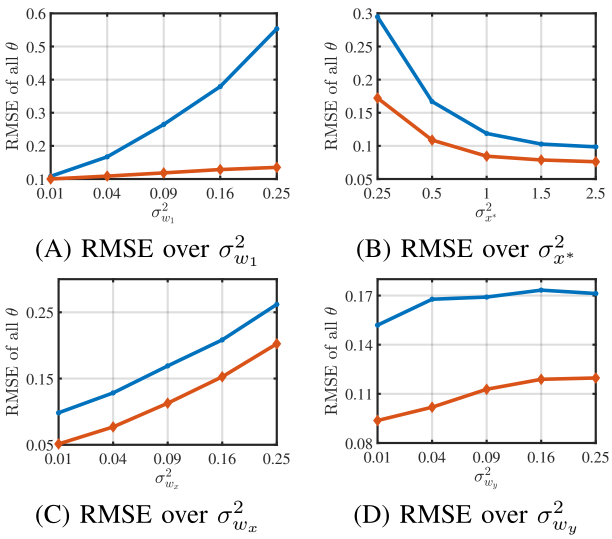
$$\mathbf{MJ}(\underline{\theta}) = \int_{-\infty}^{\infty} P(\underline{x}^*) \mathbf{J}(\underline{\theta}|\underline{x}^*) d\underline{x}^*\,. \tag{22}$$

The distribution in the integral (22) is given by $P(\underline{x}^*) \sim \mathcal{N}(\mu_{x^*}\underline{1}_N, \sigma_{x^*}^2 \mathbf{I}_{N\times N})$. Because of the nonlinearity in $\mathbf{J}(\underline{\theta}|\underline{x}^*)$, there is no analytical solution to (22). Therefore, we approximate it through *Monte Carlo* (MC) method.

V. Numerical Simulation

In this section, we investigate the performance of the proposed WTLS method in the quadratic EIV regression introduced in Sec. II. The following two simulation scenarios are discussed with results shown based on 100 MC runs.

Scenario 1: We evaluate the WTLS and compare it with LS and GTLS in solving EIV problem configured according to TABLE I. We first set up the regressor matrix $\mathbf{A}$, with its first column being noise-free, which corresponds to the standard setup of the GTLS method. In Fig. 1, it can be validated that the proposed WTLS delivers statistically the same result as GTLS. In order to verify the advantages of WTLS regarding solving the general case of nonlinear EIV regressions, an AWGN of variance $\sigma_{w_1}^2 = 0.04$ is further added to the regressor constants (first column of $\mathbf{A}$). Fig. 2 shows that the WTLS method delivers the best result in terms of estimation accuracy thanks to the proposed full covariance matrix and bias compensation. The setup of adding noise to constant terms in regressor can be useful in practice. One potential use case



(A) RMSE over $\sigma^2_{w_1}$ (B) RMSE over $\sigma^2_{x^*}$

(C) RMSE over $\sigma^2_{w_x}$ (D) RMSE over $\sigma^2_{w_y}$

Fig. 3: RMSE over various noise configurations. Red and blue curves denote results of WTLS and GTLS, respectively.

is temporal calibration of sensor networks, where uncertain clock offsets of sensor nodes are estimated [21].

Scenario 2: We now keep the setup of uncertain constants in the regressor matrix and vary its noise level of the remaining parameters following the configuration in TABLE I. The standard LS is not applicable as shown in the former scenario, thus it is not considered here. As shown in Fig. 3-(A), the proposed WTLS makes more improvement on estimation accuracy as $\sigma^2_{w_1}$ increases. Further, we keep $\sigma^2_{w_1} = 0.04$, and vary $\sigma^2_{x^*}$ and $\sigma^2_{w_x}$ to perform evaluation over different signal-to-noise ratios. Shown in Fig 3-(B) and (C), WTLS shows consistently better performance over GTLS. We also perform evaluation over various measurement noise levels as shown in Fig. 3-(D), and WTLS outperforms GTLS with a considerable margin.

TABLE I: Parameter configuration in evaluation.

parameter	μ_{x^*}	$\sigma^2_{x^*}$	$\sigma^2_{w_x}$	$\sigma^2_{w_y}$	$\underline{\theta}$	N
value	0	0.5	0.09	0.09	$[1,1,1]^\top$	500

VI. Conclusion

In this paper, we presented a novel study on exploiting weighted total least squares for solving errors-in-variables regression of quadratic structure. A general description of applying WTLS to nonlinear EIV models was provided, and the fundamental procedure of establishing the weighting matrix was showcased by investigating statistical properties of the regressor. In accordance with the proposed estimation method, we introduced the modified Cramér-Rao lower bound for evaluating the estimates. It serves as an alternative to the conventional CRLB in EIV regression containing random noise-free signal. Numerical results from simulations show superior performance of WTLS in quadratic EIV regression over standard LS and GTLS methods. Based on our current results, it is possible to improve the estimation accuracy. One promising direction is to exploit the expectation maximization algorithm in maximum likelihood framework. Besides the quadratic EIV regression problem considered in our current work, it is also promising to investigate other types of nonlinear structures in the regressor matrix. Furthermore, the proposed approach is to be applied to real-world problems for providing more extensive insights in practice. One possibility is the usage in the field of epidemic modelling [3].

References

[1] D. Tedaldi, A. Pretto, and E. Menegatti, "A Robust and Easy to Implement Method for IMU Calibration without External Equipments," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 3042–3049.

[2] F. Gustafsson, *Statistical Sensor Fusion*. Studentlitteratur, 2010.

[3] P. Liu, G. Hendeby, and F. Gustafsson, "Joint Estimation of States and Parameters in Stochastic SIR Model," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2022, pp. 1–6.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., 1993.

[5] G. H. Golub and C. F. Van Loan, "An Analysis of the Total Least Squares Problem," *SIAM journal on numerical analysis*, vol. 17, no. 6, pp. 883–893, 1980.

[6] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal processing*, vol. 87, no. 10, pp. 2283–2302, 2007.

[7] B. De Moor, "Structured Total Least Squares and L2 Approximation Problems," *Linear Algebra and Its Applications*, vol. 188, pp. 163–205, 1993.

[8] P. Lemmerling and S. Van Huffel, "Structured Total Least Squares: Analysis, Algorithms and Applications," *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*, pp. 79–91, 2002.

[9] S. Van Huffel and J. Vandewalle, "Analysis and Properties of the Generalized Total Least Squares Problem $AX \approx B$ When Some or All Columns in A are Subject to Error," *SIAM Journal on Matrix Analysis and Applications*, vol. 10, no. 3, p. 294, 1989.

[10] B. Schaffrin and A. Wieser, "On Weighted Total Least-Squares Adjustment for Linear Regression," *Journal of geodesy*, vol. 82, no. 7, pp. 415–421, 2008.

[11] X. Fang, "Weighted Total Least Squares Solutions for Applications in Geodesy," Ph.D. dissertation, Hannover: Gottfried Wilhelm Leibniz Universität Hannover, 2011.

[12] V. Mahboub, "On Weighted Total Least-Squares for Geodetic Transformations," *Journal of Geodesy*, vol. 86, no. 5, pp. 359–367, 2012.

[13] X. Fang, "A Structured and Constrained Total Least-Squares Solution with Cross-Covariances," *Studia Geophysica et Geodaetica*, vol. 58, pp. 1–16, 2014.

[14] R. M. Vaghefi, M. R. Gholami, and E. G. Ström, "Bearing-Only Target Localization with Uncertainties in Observer Position," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications Workshops*, 2010, pp. 238–242.

[15] K. Doğançay, "Bearings-Only Target Localization Using Total Least Squares," *Signal Processing*, vol. 85, no. 9, pp. 1695–1710, 2005.

[16] D. Wang, L. Zhang, and Y. Wu, "Constrained Total Least Squares Algorithm for Passive Location Based on Bearing-Only Measurements," *Science in China Series F: Information Sciences*, vol. 50, pp. 576–586, 2007.

[17] S. Jazaeri and A. Amiri-Simkooei, "Weighted Total Least Squares for Solving Non-linear Problem: GNSS Point Positioning," *Survey Review*, vol. 47, no. 343, pp. 265–271, 2015.

[18] A. Wiesel, Y. C. Eldar, and A. Beck, "Maximum Likelihood Estimation in Linear Models with a Gaussian Model Matrix," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 292–295, 2006.

[19] R. Miller and C. Chang, "A modified Cramér-Rao Bound and Its Applications," *IEEE Transactions on Information theory*, vol. 24, no. 3, pp. 398–400, 1978.

[20] O. Nestares, D. J. Fleet, and D. J. Heeger, "Likelihood functions and confidence bounds for total-least-squares problems," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, 2000, pp. 523–530.

[21] M. Kok, J. D. Hol, and T. B. Schön, "Indoor Positioning Using Ultrawideband and Inertial Measurements," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 4, pp. 1293–1303, 2015.