IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 24 February 2024, accepted 7 March 2024, date of publication 11 March 2024, date of current version 19 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375869

RESEARCH ARTICLE

TDoA Localization in Wireless Sensor Networks Using Constrained Total Least Squares and Newton's Methods

BAMRUNG TAUSIESAKUL, (Member, IEEE), AND KRISSADA ASAVASKULKIET, (Member, IEEE)
Department of Electrical Engineering, Faculty of Engineering, Mahidol University, Phutthamonthon, Nakhon Pathom 73170, Thailand
Corresponding author: Krissada Asavaskulkiet (krissada.asa@mahidol.edu)

ABSTRACT An important service in the wireless systems for the human daily life is the information of a mobile user location. Wireless sensor network is a structure that can be used to determine the mobile user position. The time-difference-of-arrival (TDoA) technique is often considered for wireless localization due to the low cost of the sensor network. In this work, the error covariance matrices are derived for predicting the error performance of the conventional closed-form constrained total least squares (CTLS) estimator. More importantly, three new Newton's methods are proposed for computing the CTLS solution in the TDoA localization. In addition, theoretical performance of both new Newton-based approaches is provided in closed forms. Numerical simulation is conducted to compare the derived theoretical prediction with the corresponding actual estimation error. It is illustrated that the two new Newton-based techniques provide better performance, in terms of lower bias and root mean square error, less computational time, and more reliability, than the former Newton-based algorithms. Furthermore, the derived expressions for the theoretical error covariance well coincide with the actual random estimation results.

INDEX TERMS First-order Taylor's series expansion, error covariance, localization, time difference of arrival, Newton's method.

I. INTRODUCTION

The use of electronic technology for the control of ground transportation systems including traffic aid systems, traffic control systems, automatic vehicle identification, location, and monitoring systems receives much attention nowadays. The information of a mobile user location is an important service in the wireless systems for the human daily life. Wireless sensor network (WSN) is a fundamental wireless structure that still receives much attention [1], [2], [3], [4], [5], especially for localization [6], [7], [8], [9], [10]. The measured quantity that can be used for finding a user location in a WSN includes received signal strength (RSS) indicator [11], [12], time of arrival (ToA) [13], angle of arrival (AoA), time difference of arrival (TDoA) [14], [15], [16], [17], [18], [19], [20] etc. Orthogonal frequency division multiplexing (OFDM) is taken into account in distance estimation for combating the fading effect, which is caused by multipath propagation. As benefitting from perfect periodic autocorrelation, Zadoff-Chu sequences are transmitted in the preambles of the OFDM in [21]. A pair of right-hand circularly polarized helical antennas result in the accuracy on the order of centimeter under occupying 20 MHz of bandwidth. The implementation of AoA localization can be found, e.g., in [22]. Phase interferometric approach is adopted therein to deal with multiple sources. The phase interferometric approach is validated in an experimental setup, where multiple transmitting units are separated from each other and the receiver is a 4-element microstrip uniform linear array antenna connected to four individual custom designed radio frequency frontends.

Various aspects of wireless localization are discussed in [23] and [24]. RSS approach is a simple and cost-efficient technique, yet prone to fading effect or multipath propagation and path loss [25]. The AoA method can achieve high localization accuracy, but incurs directional antennas, complex algorithms, and complex hardware. The ToA technique can provide high localization accuracy only when

The associate editor coordinating the review of this manuscript and approving it for publication was Huan Zhou.

the time synchronization between the transmitter and the receiver is perfect. The TDoA technique plays a prominent role in wireless localization due to the low cost of sensor network [26], [27], [28]. While the ToA localization requires a precise synchronization between the wireless node and the mobile station, the TDoA needs the strict synchronization among several wireless sensors [23]. In other words, the TDoA gets rid of the stringent synchronization between the wireless node and the mobile user [24].

The TDoA localization problem can be cast into a second-order polynomial equation [29]. Due to its nonlinear nature, the least squares fit is adopted to the TDoA localization in [30] and [31]. TDoA measurement and sensor position are considered to impose the errors in [32]. The mean square error (MSE) caused by the receiver location subject to an error is investigated in [33] by means of frequency difference of arrival and the TDoA. In [34], the sensor positions are subject to random errors and each signal emitter does not share the same time and frequency. Under both conditions, the solutions to the weighted least squares criteria are derived in closed forms. Constrained least squares criterion is considered for determining the mobile position in [35]. In [36] and [37], a constrained weighted least squares criterion is proposed for the TDoA localization.

Constrained total least squares (CTLS) criterion is introduced in [38]. It is a special case of total least squares (TLS) method [39], [40], [41], when the perturbations in the modeling matrix $\boldsymbol{A}$ and the received measurement vector $\boldsymbol{b}$ are known to be algebraically related. This TLS variant is applied in [38] to the problems of sinusoidal frequency estimation and angle-of-arrival estimation. In [42], the CTLS criterion is addressed when the constraint is unitary. It is pointed out in [43] that the CTLS objective in [38] is equivalent to the structured TLS idea in [44], where the matrices have a special structure.

Apart from the eigenvalue problems [45], the CTLS can be solved by the Newton's method (NM) [46], [47], [48], [49], [50], [51], [52], [53]. The CTLS criterion solved by the NM is adopted for sensor localization by means of the bearing angles [46], the TDoA [47], [48], [51], differential received signal strength [53], joint TDoA and angle of arrival [50], joint TDoA and wave velocity [52], and a general framework [49]. In this work, we pay attention to only the TDoA localization.

A. SHORTCOMINGS OF [47]

When the objective function in an optimization can be differentiated up to the second order, the NM often attracts much attention to solving an unconstrained problem. This is because it can offer the second-order or quadratic speed of convergence, provided that an initialized point is sufficiently close to the true value. The NM is adopted in [47] for the TDOA localization. This work is also considered in [54], [55], and [56]. The disadvantage of the NM in [47] is that the gradient and Hessian involve cumbersome expressions. Their complicated forms lead to inefficient computation and incur ill conditions for the inverse of the Hessian matrix. In addition, the theoretical error performance of the proposed method is not validated therein. In this work, the gradient and Hessian are expressed in concise forms, which lead to an efficient computation and alleviates the ill condition for the inverse of the Hessian matrix.

B. SHORTCOMINGS OF [48]

Total least squares approach still spans its application to the TDoA localization, e.g., in [48], [52], and [57]. The NM in [48] follows the framework in [47]. The key idea is that the second-order term of the noise is incorporated into the algorithm design at the beginning. Although the performance of the NM-CTLS in [48] outperforms that in [47] in terms of localization error, the shortcoming of the CTLS in [48] is that exploiting the quadratic term of the TDoA measurement noise gives rise to more complexity due to larger size of matrices. Theoretical performance analysis in terms of error covariance is also derived therein, based on the results in [47].

C. SHORTCOMINGS OF [51]

Alternating direction method of multipliers (ADMM) is an algorithm that can solve several convex optimization problems by breaking them into smaller pieces, each of which are easier to solve. Due to this advantage, it is applied to a number of areas. The ADMM method for the TDoA localization is introduced in [51] and also discussed in [58], [59], and [60]. The error expressions therein are concentrated on the relative mobile user position vector, yet not the mobile user position vector as same as in [47] and this work. The first downside of the work in [51] is that the unknown parameter vector includes a redundant quantity, which is the norm of the relative mobile user position. The redundant parameterization can deteriorate localization accuracy, because an additional search dimension introduces more local minima for the optimization and increases the computational effort in each iteration. The second drawback in [51] is that the gradient and Hessian also appear to be complicated, leading to computational demand. Third, the ADMM brings about an additional parameter, i.e. the regularization parameter, that needs to be chosen. This issue remains open and undiscussed, if one would like to obtain the minimal localization error from choosing an optimal value for this parameter. Furthermore, the benefit of convex optimization is not explored for the ADMM [51].

D. CONTRIBUTION OF THIS WORK

In this work, we derive the error covariances of the previous CTLS estimate that is expressed in a closed form [61]. The derivations are based on the first-order Taylor's series expansion (FOTSE). All expressions we drive herein are represented in closed forms and there is no need for further computing any expectation term. More importantly,

we propose an implementable version of the typical CTLS using a fixed-point iteration and two improved versions of the CTLS based on the NM.

Numerical simulation is conducted to

- investigate the closeness between the predicted theoretical errors and the actual CTLS estimation errors, and
- demonstrate better error performance of the proposed CTLS methods.

Simulation reveals that the new derived expressions of the theoretical error covariance are accurate, i.e. can well capture the simulated MSEs at a high signal-to-noise ratio, a large number of sensors. In addition, the two new CTLS methods involving Newton's method provide less mobile position estimation error and less computational time, which can fulfill the requirement of compact computing power in [27].

Contribution of this work thus includes

- the expressions of the error covariance given by the CTLS estimator
  and
- the proposed CTLS algorithms based on the fixed-point iteration and Newton's method for sensor network localization using the TDoA.
- The error expressions concentrating on the mobile user position for both proposed CTLS approaches are also derived.
- Simulation is provided to validate the accuracy of the proposed theoretical prediction.

Merit of this work lies in

- the accurate theoretical prediction of the CTLS estimation error performance
  and
- two new accurate methods in determining the mobile user position.

E. PAPER ORGANIZATION

The structure of this paper is arranged as follows.

- Introduction
  - Shortcomings of [47]
  - Shortcomings of [48]
  - Shortcomings of [51]
  - Contribution of This Work
  - Paper organization
- Localization via Time Difference of Arrival
  - Time Difference of Arrival
  - Linear Model of TDoA Perturbation
  - Noise Approximation
  - Noise Assumptions
  - Model Limitation
- Previous Methods
  - Constrained Total Least Squares
  - Constrained Total Least Squares Using Alternating Direction Method of Multipliers
  - Constrained Total Least Squares Using Newton's Method

TABLE 1. Mathematical notations and their meanings.

Notation	Meaning
$\|\cdot\|_2$	the ℓ_2 norm or the Euclidean norm
$\hat{\boldsymbol{x}}$	the estimated value of an unknown variable $\boldsymbol{x}$
$(\cdot)^\mathsf{T}$	the transpose of a matrix or a vector
$(\cdot)^{-1}$	the inverse of a square matrix
$\mathcal{E}_{\mathbf{n}}\{\cdot\}$	the expectation of a quantity with respect to the random variable $\mathbf{n}$
$\boldsymbol{I}_M$	the identity matrix of a size $M \times M$
$\boldsymbol{x}_0$	the true value of the unknown variable $\boldsymbol{x}$
$z \sim \mathcal{N}_{\mathbb{R}}(\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_{zz})$	the random variable z has a real-valued Gaussian distribution with mean $\boldsymbol{\mu}_z$ and covariance $\boldsymbol{\Sigma}_{zz}$
$\mathrm{tr}(\cdot)$	the trace of a square matrix
$\mathrm{rank}(\cdot)$	the rank of a matrix
$\boldsymbol{D}(\cdot)$	the diagonal matrix whose diagonal is taken from the vector $\cdot$
$\odot$	the Hadamard product
$\boldsymbol{D}^2(\cdot)$	the product given by $\boldsymbol{D}(\cdot)\boldsymbol{D}(\cdot)$
$\boldsymbol{D}^{-2}(\cdot)$	the product given by $\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)$
$\boldsymbol{D}^{-3}(\cdot)$	the product given by $\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)$
$\boldsymbol{D}^{-4}(\cdot)$	the product given by $\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)\boldsymbol{D}^{-1}(\cdot)$
$\mathbf{1}_N$	the column vector containing only ones of size N
$\mathbf{0}_N$	the column vector containing only zeros of size N

- Proposed Methods
  - Constrained Total Least Squares Using Fixed-Point Iteration
  - Constrained Total Least Squares Using Newton's Method with Two Steps
  - Constrained Total Least Squares Using Newton's Method with One Step
- Numerical Examples
  - Algorithmic Comparison
  - Simulation Results
    - Effect of the standard deviation of the TDoA error
    - Effect of the number of sensors
    - Effect of the imperfect synchronization
- Conclusion
- Appendices
  - Preliminary Results
  - Proof of Lemma 1
  - Proof of Lemma 2
  - Proof of Lemma 3
  - Proof of Corollary 1

All Mathematical notations in this paper are listed in Table 1 and all abbreviations are provided in Table 2. Sec. II addresses the localization model that adopts the TDoA as the measured signal. In Sec. III, we review previous methods and derive the theoretical error performance of the CTLS in the closed form. Sec. IV presents three new methods based on the CTLS for the TDoA localization. In Sec. V, numerical examples are conducted to examine the closeness of the theoretical error prediction to the actual CTLS estimation error performance and to investigate the superior performance of the proposed algorithms. In Sec. VI, concluding remark is given to this work.

II. LOCALIZATION VIA TIME DIFFERENCE OF ARRIVAL

Let $\boldsymbol{p} \in \mathbb{R}^{2\times 1}$ be an unknown position of a mobile user in a two-dimensional Cartesian coordinates, i.e.

$$\boldsymbol{p} = \begin{bmatrix} x \\ y \end{bmatrix}. \tag{1}$$

TABLE 2. Acronyms and their full names.

Acronym	Full name
ADMM	alternating direction method of multipliers
CTLS	constrained total least squares
FOTSE	first-order Taylor's series expansion
FP	fixed point
MSE	mean-squared error
NM	Newton's method
OSNM	one-step Newton's method
RMSE	root-mean-squared error
SNR	signal-to-noise ratio
TDoA	time difference of arrival
TSNM	two-step Newton's method

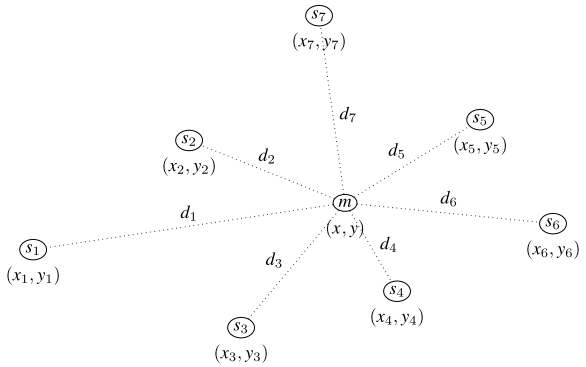

FIGURE 1. Wireless sensors, namely $s_1, s_2, \ldots, s_K$, for $K = 7$ locate in the vicinity of the mobile user m.

Wireless sensor network is a system that has a mobile user and K fixed sensors with different positions

$$\boldsymbol{p}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}, \tag{2}$$

for $k \in \{1, 2, \ldots, K\}$. An example of the geometrical illustration is in Fig. 1. The model in Fig. 1 can be perceived as the use of mobile radio on land with applications to dispatch and control vehicles or status monitoring and reporting.

A. TIME DIFFERENCE OF ARRIVAL

Let d_k be the distance from the k-th sensor to the mobile user, given by

$$\begin{aligned} d_k &= \|\boldsymbol{p} - \boldsymbol{p}_k\|_2 \\ &= \sqrt{(x_k - x)^2 + (y_k - y)^2}, \end{aligned} \tag{3}$$

where $\|\cdot\|_2$ is the ℓ_2 norm of a vector $\cdot$. Assume that

- all sensor positions, $\boldsymbol{p}_k$ for $k \in \{1, 2, \ldots, K\}$, are perfectly known and
- the mobile user position $\boldsymbol{p}$ is unknown and needs to be determined.

The TDoA strategy assumes that the first sensor is designated as the reference node. The distance from the k-th sensor to the first sensor can be expressed as [51, eq. (1)]

$$\begin{aligned} \delta_{k,1} &= d_k - d_1 \\ &= \sqrt{(x_k - x)^2 + (y_k - y)^2} - \sqrt{(x_1 - x)^2 + (y_1 - y)^2}. \end{aligned} \tag{4}$$

In the absence of noise, there exists an equality

$$\boldsymbol{A}_0 \boldsymbol{\theta}_0 = \boldsymbol{b}_0, \tag{5}$$

where $\boldsymbol{A}_0 \in \mathbb{R}^{(K-1)\times 3}$ is the matrix, given by

$$\boldsymbol{A}_0 = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & \delta_{2,1} \\ x_3 - x_1 & y_3 - y_1 & \delta_{3,1} \\ \vdots & \vdots & \vdots \\ x_K - x_1 & y_K - y_1 & \delta_{K,1} \end{bmatrix}, \tag{6}$$

$\boldsymbol{\theta}_0 \in \mathbb{R}^{3\times 1}$ is the unknown true value vector, given by

$$\boldsymbol{\theta}_0 = \begin{bmatrix} x - x_1 \\ y - y_1 \\ d_1 \end{bmatrix}, \tag{7}$$

and $\boldsymbol{b}_0 \in \mathbb{R}^{(K-1)\times 1}$ is the observation vector, given by

$$\boldsymbol{b}_0 = \frac{1}{2} \begin{bmatrix} (x_2 - x_1)^2 + (y_2 - y_1)^2 - \delta_{2,1}^2 \\ (x_3 - x_1)^2 + (y_3 - y_1)^2 - \delta_{3,1}^2 \\ \vdots \\ (x_K - x_1)^2 + (y_K - y_1)^2 - \delta_{K,1}^2 \end{bmatrix}. \tag{8}$$

Under perfect time synchronization among K sensors, the TDoA from the first sensor to the k-th sensor can be observed without any noise. It has a relationship to the distance with respect to the first sensor according to

$$\tau_{k,1} = \frac{1}{c} \delta_{k,1}, \tag{9}$$

where $c = 299,792,458 \approx 3 \times 10^8$ [m/s] is the speed of light.

B. LINEAR MODEL OF TDOA PERTURBATION

In realistic measurement, it is difficult to keep the synchronization among K sensors perfect. Let $\eta_{k,1}$ be a TDoA measurement error, which is given by

$$\eta_{k,1} = \frac{1}{c} n_{k,1}, \tag{10}$$

where $n_{k,1}$ is the corresponding relative distance error. Therefore, the measured TDoA $\tilde{\tau}_{k,1}$ encounters the additive noise according to

$$\begin{aligned} \tilde{\tau}_{k,1} &= \underbrace{\tau_{k,1}}_{=\frac{1}{c}\delta_{k,1};\ (9)} + \underbrace{\eta_{k,1}}_{=\frac{1}{c}n_{k,1};\ (10)} \\ &= \frac{1}{c}(\delta_{k,1} + n_{k,1}). \end{aligned} \tag{11}$$

The data we can receive is given by

$$\boldsymbol{A}\boldsymbol{\theta} = \boldsymbol{b}, \tag{12}$$

where $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$ is the noisy value of the true value $\boldsymbol{A}_0$ according to

$$\begin{aligned} \boldsymbol{A} &= \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & \delta_{2,1} + n_{2,1} \\ x_3 - x_1 & y_3 - y_1 & \delta_{3,1} + n_{3,1} \\ \vdots & \vdots & \vdots \\ x_K - x_1 & y_K - y_1 & \delta_{K,1} + n_{K,1} \end{bmatrix} \\ &= \boldsymbol{A}_0 + \boldsymbol{\Delta}_{\mathrm{A}}, \end{aligned} \tag{13}$$

with $\boldsymbol{\Delta}_{\mathrm{A}} \in \mathbb{R}^{(K-1)\times 3}$ that represents of the perturbation of $\boldsymbol{A}$, given by

$$\boldsymbol{\Delta}_{\mathrm{A}} = \begin{bmatrix} 0 & 0 & n_{2,1} \\ 0 & 0 & n_{3,1} \\ \vdots & \vdots & \vdots \\ 0 & 0 & n_{K,1} \end{bmatrix}, \tag{14}$$

and $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$ is the noisy value of the true value $\boldsymbol{b}_0$ corresponding to

$$\begin{aligned} \boldsymbol{b} &= \frac{1}{2}\begin{bmatrix} (x_2-x_1)^2+(y_2-y_1)^2-(\delta_{2,1}+n_{2,1})^2 \\ (x_3-x_1)^2+(y_3-y_1)^2-(\delta_{3,1}+n_{3,1})^2 \\ \vdots \\ (x_K-x_1)^2+(y_K-y_1)^2-(\delta_{K,1}+n_{K,1})^2 \end{bmatrix} \\ &= \boldsymbol{b}_0 + \boldsymbol{\delta}_{\mathrm{b}}, \end{aligned} \tag{15}$$

with $\boldsymbol{\delta}_{\mathrm{b}} \in \mathbb{R}^{(K-1)\times 1}$ that represents of the perturbation of $\boldsymbol{b}$. The perturbation of $\boldsymbol{A}$ can be expressed as [47, eq. (6)]

$$\boldsymbol{\Delta}_{\mathrm{A}} = \begin{bmatrix} 0 & 0 & \boldsymbol{n} \end{bmatrix}, \tag{16}$$

where $\boldsymbol{n}$ is the noise caused by the imperfect TDoA measurement, written by

$$\boldsymbol{n} = \begin{bmatrix} n_{2,1} \\ n_{3,1} \\ \vdots \\ n_{K,1} \end{bmatrix}. \tag{17}$$

C. NOISE APPROXIMATION

The perturbation of $\boldsymbol{b}$ can be approximated as [47, eq. (6)]

$$\begin{aligned} \boldsymbol{\delta}_{\mathrm{b}} &= -\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n} - \frac{1}{2}\boldsymbol{D}(\boldsymbol{n})\boldsymbol{n} \\ &\approx -\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}, \end{aligned} \tag{18}$$

where $\boldsymbol{\delta} \in \mathbb{R}^{(K-1)\times 1}$ is the relative distance vector, written by

$$\boldsymbol{\delta} = \begin{bmatrix} \delta_{2,1} \\ \delta_{3,1} \\ \vdots \\ \delta_{k,1} \end{bmatrix}, \tag{19}$$

and $\boldsymbol{D}(\cdot)$ is the diagonal matrix whose diagonal is taken from the vector $\cdot$. In [48], the approximation in (18) is not taken into account. However, the quadratic term of the additive noise brings about more complexity. From (16) and (18), one can see that there are errors at both sides of (12). The TDoA localization can be considered as a structured TLS problem, when the perturbations have a special structure as shown in (16) and (18). Model description of Fig. 1 can be summarized in Tbl. 3. From (6) and (19), there exists a relation

$$\boldsymbol{A}_0 = \begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta} \end{bmatrix}, \tag{20}$$

where $\boldsymbol{\delta}_{\mathrm{x}} \in \mathbb{R}^{(K-1)\times 1}$ is the vector, given by

$$\boldsymbol{\delta}_{\mathrm{x}} = \begin{bmatrix} x_2 - x_1 \\ x_3 - x_1 \\ \vdots \\ x_K - x_1 \end{bmatrix}, \tag{21}$$

and $\boldsymbol{\delta}_{\mathrm{y}} \in \mathbb{R}^{(K-1)\times 1}$ is the vector, given by

$$\boldsymbol{\delta}_{\mathrm{y}} = \begin{bmatrix} y_2 - y_1 \\ y_3 - y_1 \\ \vdots \\ y_K - y_1 \end{bmatrix}. \tag{22}$$

D. NOISE ASSUMPTIONS

The TDoA measurement error is assumed to obey $\eta_{k,1} \sim \mathcal{N}_{\mathbb{R}}(0, \sigma_\eta^2)$ for $k \in \{2, 3, \ldots, K\}$, where σ_η is the standard deviation of the TDoA measurement error, e.g., $\sigma_\eta = 10$[ns] [51, Sec. IV.A]. If the elements of the noise vector $\boldsymbol{n}$ have a zero mean and an identical noise variance σ_{n}^2 [47], we can see that

$$\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\} = \boldsymbol{0}, \tag{23a}$$

$$\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\} = \sigma_{\mathrm{n}}^2 \boldsymbol{I}_{K-1}, \tag{23b}$$

where $\mathcal{E}_{\mathbf{n}}\{\cdot\}$ is the expectation of $\cdot$ with respect to the random variable $\mathbf{n}$, $\cdot^{\mathsf{T}}$ is the transpose of a vector $\cdot$ or a matrix $\cdot$, and $\boldsymbol{I}_N$ is the identity matrix of size $N \times N$.

E. MODEL LIMITATION

The limitation of the TDoA localization in this model is that the approximation in (18) causes the loss of the accuracy for all methods designed to handle the TDoA localization problem. To overcome this issue, one should avoid the approximation in (18) and have to deal with the second-order term of the noise. We shall leave this idea for future work.

III. PREVIOUS METHODS

A classical method in solving (12) adopts the LS criterion [31, eq. (7)], i.e.

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\ell_2} &= \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}\|_2^2 \\ &= \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{b}, \end{aligned} \tag{24}$$

where $\cdot^{-1}$ is the inverse of a square matrix $\cdot$. In what follows, we review a few sophisticated approaches for achieving higher localization accuracy.

A. CONSTRAINED TOTAL LEAST SQUARES

It is fruitful to note that

$$\operatorname{rank}\left(\boldsymbol{D}(\boldsymbol{d})\right) = K - 1, \tag{25}$$

for

$$d_2 \neq d_3 \neq \ldots \neq d_K, \tag{26}$$

TABLE 3. Model of Fig. 1.

Problem	Figure	Perturbation	True Values	Known	Unknown
TDoA localization	Fig. 1	$\boldsymbol{\Delta}_{\mathrm{A}} = \begin{bmatrix}\mathbf{0} & \mathbf{0} & \boldsymbol{n}\end{bmatrix}$, $\boldsymbol{\delta}_{\mathrm{b}} \approx -\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}$	$\boldsymbol{A}_0, \boldsymbol{b}_0$	$\boldsymbol{A}, \boldsymbol{b}$	$\boldsymbol{p}, \boldsymbol{n}, \boldsymbol{A}_0, \boldsymbol{b}_0$

where rank$(\cdot)$ is the rank of a matrix $\cdot$ and $\boldsymbol{d} \in \mathbb{R}^{(K-1)\times 1}$ is the vector given by

$$\boldsymbol{d} = \begin{bmatrix} d_2 \\ d_3 \\ \vdots \\ d_K \end{bmatrix}. \tag{27}$$

The observation model in (12) results in [47, eq. (8)]

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b} &= (\boldsymbol{A}_0 + \boldsymbol{\Delta}_{\mathrm{A}})\boldsymbol{\theta} - (\boldsymbol{b}_0 + \boldsymbol{\delta}_{\mathrm{b}}) \\
&= \underbrace{\boldsymbol{A}_0\boldsymbol{\theta}}_{=\boldsymbol{b}_0;\ (5)} + \underbrace{\boldsymbol{\Delta}_{\mathrm{A}}}_{=[\mathbf{0}\ \mathbf{0}\ \boldsymbol{n}];\ (16)} \underbrace{\boldsymbol{\theta}}_{=\begin{bmatrix} x-x_1 \\ y-y_1 \\ d_1 \end{bmatrix};\ (7)} - \boldsymbol{b}_0 - \underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{\approx -\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n};\ (18)} \\
&\approx \boldsymbol{b}_0 - \boldsymbol{b}_0 + \mathbf{0}(x - x_1) + \mathbf{0}(y - y_1) \\
&\quad + \boldsymbol{n}d_1 - (-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}) \\
&= d_1\boldsymbol{n} + \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n} \\
&= (\boldsymbol{D}(\boldsymbol{\delta}) + d_1\boldsymbol{I}_{K-1})\,\boldsymbol{n} \\
&= (\boldsymbol{D}(\boldsymbol{\delta} + d_1\mathbf{1}))\,\boldsymbol{n} \\
&= \boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}.
\end{aligned} \tag{28}$$

Solving (12) for $\boldsymbol{\theta}$ in a minimal ℓ_2-norm sense of $\boldsymbol{n}$ subject to the condition $\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b} = \boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}$ in (28) gives [38]

$$\begin{bmatrix} \hat{\boldsymbol{\theta}}_{\ell_2} \\ \hat{\boldsymbol{n}}_{\ell_2} \end{bmatrix} = \arg\min_{\begin{bmatrix}\boldsymbol{\theta} \\ \boldsymbol{n}\end{bmatrix}} \|\boldsymbol{n}\|_2^2 \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b} = \boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}. \tag{29}$$

The above problem is known as constrained total least squares (CTLS) [51], [61]. To reduce the computational effort, the joint optimization problem in (29) can be decoupled into two separate optimization problems.

Lemma 1 (CTLS solution for $\boldsymbol{\theta}$): The CTLS solution of $\boldsymbol{\theta}$ in (29) is given by

$$\hat{\boldsymbol{\theta}}_{\ell_2} = \arg\min_{\boldsymbol{\theta}} (\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b})^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}), \tag{30}$$

which becomes an unconstrained problem. Let the CTLS cost function in (30) be

$$f_{\mathrm{CTLS}}(\boldsymbol{\theta}) = (\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b})^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}). \tag{31}$$

The gradient of the CTLS can be written as

$$\boldsymbol{g}_{\theta}(\boldsymbol{\theta}) = 2\boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}). \tag{32}$$

The closed-form solution to (30) can be shown as

$$\hat{\boldsymbol{\theta}}_{\ell_2} = \left(\boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{b}. \tag{33}$$

Proof: The derivation is based on typical algebra. See Appendix B. ■

The result in (33) corresponds to [61, eq. (31)]. The CTLS estimation can be summarized in Algorithm 4. The CTLS in Algorithm 4 requires a single calculation of its closed-form solution. It finally converts the first two elements resided in $\boldsymbol{\theta}$ to the mobile position $\boldsymbol{p}$.

Algorithm 1 Constrained Total Least Squares (CTLS) [61]

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$

Output: $\hat{\boldsymbol{p}}_{\mathrm{CTLS}} \in \mathbb{R}^{N\times 1}$

$$\boldsymbol{d} \leftarrow \begin{bmatrix} \|\boldsymbol{p} - \boldsymbol{p}_2\|_2 \\ \|\boldsymbol{p} - \boldsymbol{p}_3\|_2 \\ \vdots \\ \|\boldsymbol{p} - \boldsymbol{p}_K\|_2 \end{bmatrix}$$

$$\hat{\boldsymbol{\theta}}_{\ell_2} \leftarrow \left(\boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{b}$$

$$[\hat{\boldsymbol{p}}_{\ell_2}]_1 \leftarrow [\hat{\boldsymbol{\theta}}_{\ell_2}]_1 + [\boldsymbol{p}_1]_1$$

$$[\hat{\boldsymbol{p}}_{\ell_2}]_2 \leftarrow [\hat{\boldsymbol{\theta}}_{\ell_2}]_2 + [\boldsymbol{p}_1]_2$$

return $\hat{\boldsymbol{p}}_{\ell_2}$

Lemma 2 (The Hessian of the CTLS with respect to $\boldsymbol{\theta}$): The Hessian of the CTLS with respect to $\boldsymbol{\theta}$ can be expressed as

$$\begin{aligned}
\boldsymbol{H}_{\theta\theta}(\boldsymbol{\theta}) &= \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathsf{T}}} f_{\mathrm{CTLS}}(\boldsymbol{\theta}) \\
&= 2\boldsymbol{A}^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}.
\end{aligned} \tag{34}$$

The expectation of (34) can be shown as

$$\begin{aligned}
\bar{\boldsymbol{H}}_{\theta\theta}(\boldsymbol{\theta}) &= \mathcal{E}_{\mathbf{n}}\{\boldsymbol{H}_{\theta\theta}(\boldsymbol{\theta})\} \\
&= 2\left(\boldsymbol{A}_0^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{n}}^2 \operatorname{tr}(\boldsymbol{D}^{-2}(\boldsymbol{d})) \end{bmatrix}\right),
\end{aligned} \tag{35}$$

where $\operatorname{tr}(\cdot)$ is the trace of a square matrix $\cdot$.

Proof: The derivation is based on typical algebra. See Appendix C. ■

Based on the FOTSE, the CTLS estimate can be approximated by [62, p. 281]

$$\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}} - \boldsymbol{\theta}_0 \simeq -\bar{\boldsymbol{H}}_{\theta\theta}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{g}_{\theta}(\boldsymbol{\theta}_0), \tag{36}$$

where $\bar{\boldsymbol{H}}_{\theta\theta}(\boldsymbol{\theta}_0)$ is given by (35) and $\boldsymbol{g}_{\theta}(\boldsymbol{\theta}_0)$ is given by (32).

Lemma 3 (The error covariance matrix of the CTLS estimate): The error covariance matrix of the CTLS estimate is given by

$$\begin{aligned}
\boldsymbol{\Sigma}_{\theta\theta}(\hat{\boldsymbol{\theta}}_{CTLS}) &= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{\theta}}_{CTLS} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_{CTLS} - \boldsymbol{\theta}_0)^{\mathsf{T}}\} \\
&\approx \sigma_{\mathrm{n}}^2 \boldsymbol{Q}_1^{-1} \boldsymbol{Q}_2 \boldsymbol{Q}_1^{-1},
\end{aligned} \tag{37}$$

where $\boldsymbol{Q}_1 \in \mathbb{R}^{3\times 3}$ is given by

$$\boldsymbol{Q}_1 = \boldsymbol{A}_0^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \operatorname{tr}(\boldsymbol{D}^{-2}(\boldsymbol{d})) \end{bmatrix}, \tag{38}$$

and $\boldsymbol{Q}_2 \in \mathbb{R}^{3\times 3}$ *is given by*

$$\boldsymbol{Q}_2 = \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}, \tag{39}$$

with α_3 *given by*

$$\alpha_3 = 2\sum_{k=2}^{K}\frac{1}{d_k^2} + \sum_{k_1=2}^{K}\sum_{k_2=2}^{K}\frac{1}{d_{k_1}d_{k_2}}. \tag{40}$$

Proof: The derivation is based on the FOTSE. See Appendix D. ∎

Corollary 1 (The error covariance matrix of the CTLS estimate for small noise variance): For a small σ_{n}^2*, the error covariance in (37) can be simplified into*

$$\lim_{\sigma_{\mathrm{n}}^2\to 0}\boldsymbol{\Sigma}_{\theta\theta}(\hat{\boldsymbol{\theta}}_{CTLS}) \approx \sigma_{\mathrm{n}}^2\left(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\right)^{-1}. \tag{41}$$

Proof: The derivation is straightforward from (37). See Appendix E. ∎

To our best knowledge, the error covariance matrix in (37) and its simplification in (41) are not available yet in any previous work. The error covariance matrix of the mobile position that is estimated by the CTLS in (33) is given by

$$\begin{aligned}
&\boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\mathrm{CTLS}}) \\
&= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_0)(\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_0)^{\mathsf{T}}\} \\
&= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_1 - \boldsymbol{p}_0 + \boldsymbol{p}_1)(\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_1 - \boldsymbol{p}_0 + \boldsymbol{p}_1)^{\mathsf{T}}\} \\
&= \mathcal{E}_{\mathbf{n}}\{(\underbrace{\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_1}_{=[\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}}]_{1:2,1}} - \underbrace{(\boldsymbol{p}_0 - \boldsymbol{p}_1)}_{=[\boldsymbol{\theta}_0]_{1:2,1}})(\underbrace{\hat{\boldsymbol{p}}_{\mathrm{CTLS}} - \boldsymbol{p}_1}_{=[\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}}]_{1:2,1}} - \underbrace{(\boldsymbol{p}_0 - \boldsymbol{p}_1)}_{=[\boldsymbol{\theta}_0]_{1:2,1}})^{\mathsf{T}}\} \\
&= [\boldsymbol{\Sigma}_{\theta\theta}(\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}})]_{1:2,1:2},
\end{aligned} \tag{42}$$

which yields the first 2×2 sub-block matrix of $\boldsymbol{\Sigma}_{\theta\theta}(\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}}) \in \mathbb{R}^{3\times 3}$ in (37).

B. CONSTRAINED TOTAL LEAST SQUARES USING ALTERNATING DIRECTION METHOD OF MULTIPLIERS

In [51], the ADMM is applied to the CTLS function in (31). A constraint of

$$\|\boldsymbol{v}\|_2 = d_1, \tag{43}$$

is introduced, where $\boldsymbol{v} \in \mathbb{R}^{2\times 1}$ is the vector that contains the first two elements of $\boldsymbol{\theta}$, i.e.

$$\boldsymbol{v} = \boldsymbol{p} - \boldsymbol{p}_1. \tag{44}$$

It is worth noting that

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b} &= \begin{bmatrix}\boldsymbol{A}_2 & \boldsymbol{a}_3\end{bmatrix}\begin{bmatrix}\boldsymbol{v} \\ d_1\end{bmatrix} - \boldsymbol{b} \\
&= \boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b},
\end{aligned} \tag{45}$$

where $\boldsymbol{A}_2 \in \mathbb{R}^{(K-1)\times 2}$ is the matrix that contains the first two columns of $\boldsymbol{A}$ and $\boldsymbol{a}_3 \in \mathbb{R}^{(K-1)\times 1}$ is the vector that contains the last column of $\boldsymbol{A}$. Extended from [47, eq. (13)], the objective function of the CTLS can be formulated as the augmented Lagrangian according to

$$\begin{aligned}
L(\boldsymbol{v}, d_1, \rho) &= (\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad + \lambda(\|\boldsymbol{v}\|_2 - d_1) + \frac{1}{2}\rho(\|\boldsymbol{v}\|_2 - d_1)^2,
\end{aligned} \tag{46}$$

where $\lambda \in \mathbb{R}^{1\times 1}$ is the Lagrange multiplier and $\rho \in \mathbb{R}_+^{1\times 1}$ is the regularization parameter. The partial derivatives of the augmented Lagrangian in (46) with respect to $\boldsymbol{v}$ and d_1 can be written as [51, eq. (21)-(26)]

$$\begin{aligned}
\frac{\partial}{\partial \boldsymbol{v}}L &= 2\boldsymbol{A}_2^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad + \frac{1}{\|\boldsymbol{v}\|_2}\rho\left(\|\boldsymbol{v}\|_2 - d_1 + \frac{\lambda}{\rho}\right)\boldsymbol{v},
\end{aligned} \tag{47}$$

$$\begin{aligned}
\frac{\partial^2}{\partial \boldsymbol{v}\partial \boldsymbol{v}^{\mathsf{T}}}L &= 2\boldsymbol{A}_2^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_2 + \frac{1}{\|\boldsymbol{v}\|_2}\rho\Bigg(\frac{1}{\|\boldsymbol{v}\|_2}\boldsymbol{v}\boldsymbol{v}^{\mathsf{T}} \\
&\quad + \left(\|\boldsymbol{v}\|_2 - d_1 + \frac{\lambda}{\rho}\right)\left(\boldsymbol{I}_2 - \frac{1}{\|\boldsymbol{v}\|_2^2}\boldsymbol{v}\boldsymbol{v}^{\mathsf{T}}\right)\Bigg),
\end{aligned} \tag{48}$$

$$\begin{aligned}
\frac{\partial}{\partial d_1}L &= 2\boldsymbol{a}_3^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad - 2(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-3}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad - \rho\left(\|\boldsymbol{v}\|_2 - d_1 + \frac{\lambda}{\rho}\right),
\end{aligned} \tag{49}$$

and

$$\begin{aligned}
\frac{\partial^2}{\partial d_1^2}L &= 2\boldsymbol{a}_3^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{a}_3 - 8\boldsymbol{a}_3^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad - 2(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-4}(\boldsymbol{d})(\boldsymbol{A}_2\boldsymbol{v} + \boldsymbol{a}_3 d_1 - \boldsymbol{b}) \\
&\quad + \rho.
\end{aligned} \tag{50}$$

The iterative computation of the CTLS using the ADMM can be summarized as follows. The computation in Algorithm 2 is an iterative computation. Inside each loop, there are four variables that need to be updated, such as two variables in $\hat{\boldsymbol{v}}[n]$, one variable in $\hat{d}_1[n]$, and one variable in $\hat{\lambda}[n]$. Therefore, the Newton's iterations for $\hat{\boldsymbol{v}}[n]$ and $\hat{d}_1[n]$ are intensive in Algorithm 2.

Based on (42), we can find that

$$\begin{aligned}
&\boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}}) \\
&= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_0)(\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_0)^{\mathsf{T}}\} \\
&= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_1 - \boldsymbol{p}_0 + \boldsymbol{p}_1) \\
&\quad (\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_1 - \boldsymbol{p}_0 + \boldsymbol{p}_1)^{\mathsf{T}}\} \\
&= \mathcal{E}_{\mathbf{n}}\{(\underbrace{\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_1}_{=\hat{\boldsymbol{v}}_{\text{ADMM-CTLS}}} - \underbrace{(\boldsymbol{p}_0 - \boldsymbol{p}_1)}_{=\boldsymbol{v}_0}) \\
&\quad (\underbrace{\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} - \boldsymbol{p}_1}_{=\hat{\boldsymbol{v}}_{\text{ADMM-CTLS}}} - \underbrace{(\boldsymbol{p}_0 - \boldsymbol{p}_1)}_{=\boldsymbol{v}_0})^{\mathsf{T}}\} \\
&= \boldsymbol{\Sigma}_{vv}(\hat{\boldsymbol{v}}_{\text{ADMM-CTLS}}).
\end{aligned} \tag{51}$$

The result in (51) implies that the error covariance matrix of $\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}}$ is the same as that of $\hat{\boldsymbol{v}}_{\text{ADMM-CTLS}}$.

Algorithm 2 Constrained Total Least Squares Using Alternating Direction Method of Multipliers (ADMM-CTLS) [51]

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$, $\epsilon_{\min} \in (0, 1)$, $N_{\max} \in \mathbb{Z}_+^{1\times 1}$, $\rho \in (0, 1)$, $\hat{\boldsymbol{p}}_{\text{CTLS}}$

Output: $\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}} \in \mathbb{R}^{N\times 1}$

$n \leftarrow 0$

$\hat{\boldsymbol{v}}[0] \leftarrow \hat{\boldsymbol{p}}_{\text{CTLS}} - \boldsymbol{p}_1$

while $\epsilon_{\hat{\boldsymbol{v}}} > \epsilon_{\min} \wedge n \leq N_{\max}$ **do**

$n \leftarrow n + 1$

$$\hat{\boldsymbol{v}}[n] \leftarrow \hat{\boldsymbol{v}}[n-1] - \underbrace{\left(\frac{\partial^2}{\partial \boldsymbol{v}\partial \boldsymbol{v}^{\mathsf{T}}} L\Big|_{\substack{v=\hat{v}[n-1]\\ d_1=\hat{d}_1[n-1]\\ \lambda=\hat{\lambda}[n-1]}}\right)^{-1}}_{(48)} \underbrace{\frac{\partial}{\partial \boldsymbol{v}} L\Big|_{\substack{v=\hat{v}[n-1]\\ d_1=\hat{d}_1[n-1]\\ \lambda=\hat{\lambda}[n-1]}}}_{(47)}$$

$$\hat{d}_1[n] \leftarrow \hat{d}_1[n-1] - \frac{\underbrace{\frac{\partial}{\partial d_1} L\Big|_{\substack{v=\hat{v}[n]\\ d_1=\hat{d}_1[n-1]\\ \lambda=\hat{\lambda}[n-1]}}}_{(49)}}{\underbrace{\frac{\partial^2}{\partial d_1^2} L\Big|_{\substack{v=\hat{v}[n]\\ d_1=\hat{d}_1[n-1]\\ \lambda=\hat{\lambda}[n-1]}}}_{(50)}}$$

$\hat{\lambda}[n] \leftarrow \hat{\lambda}[n-1] + \rho(\|\hat{\boldsymbol{v}}[n]\|_2 - \hat{d}_1[n])$

$\epsilon_{\hat{\boldsymbol{v}}} \leftarrow \frac{\|\hat{\boldsymbol{v}}[n]-\hat{\boldsymbol{v}}[n-1]\|_2}{\|\hat{\boldsymbol{v}}[n-1]\|_2}$

end while

$\hat{\boldsymbol{p}}_{\ell_2} \leftarrow \hat{\boldsymbol{v}}[n] + \boldsymbol{p}_1$

return $\hat{\boldsymbol{p}}_{\ell_2}$

C. CONSTRAINED TOTAL LEAST SQUARES USING NEWTON'S METHOD

Let $\boldsymbol{v}_0 \in \mathbb{R}^{2\times 1}$ be the true value of $\boldsymbol{v}$, i.e.

$$\boldsymbol{v}_0 = \boldsymbol{p}_0 - \boldsymbol{p}_1. \tag{52}$$

One can see that

$$\boldsymbol{\theta}_0 = \begin{bmatrix} \boldsymbol{v}_0 \\ \|\boldsymbol{v}_0\|_2 \end{bmatrix}. \tag{53}$$

We can derive [47, eq. (A.4)]

$$\begin{aligned} \boldsymbol{g}_{\text{v}}(\boldsymbol{v}) &= \frac{\partial}{\partial \boldsymbol{v}} f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= 2\left(\boldsymbol{\Theta}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\beta} - \frac{\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{D}^{-3}(\boldsymbol{d})\boldsymbol{\beta}}{\|\boldsymbol{v}\|_2}\boldsymbol{v}\right), \end{aligned} \tag{54}$$

and [47, eq. (A.5)]

$$\begin{aligned} \boldsymbol{H}_{\text{vv}}(\boldsymbol{v}) &= \frac{\partial^2}{\partial \boldsymbol{v}\partial \boldsymbol{v}^{\mathsf{T}}} f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \frac{2}{\|\boldsymbol{v}\|_2}\left(\boldsymbol{a}_3^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\beta} - \boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{D}^{-3}(\boldsymbol{d})\boldsymbol{\beta}\right) \\ &\quad \left(\boldsymbol{I}_2 - \frac{1}{\|\boldsymbol{v}\|_2^2}\boldsymbol{v}\boldsymbol{v}^{\mathsf{T}}\right) - \frac{4}{\|\boldsymbol{v}\|_2}\boldsymbol{\Theta}\boldsymbol{D}^{-3}(\boldsymbol{d})\boldsymbol{\beta}\boldsymbol{v}^{\mathsf{T}} \\ &\quad + \frac{6\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{D}^{-4}(\boldsymbol{d})\boldsymbol{\beta}}{\|\boldsymbol{v}\|_2^2}\boldsymbol{v}\boldsymbol{v}^{\mathsf{T}} - \frac{4}{\|\boldsymbol{v}\|_2}\boldsymbol{v}\boldsymbol{\beta}^{\mathsf{T}}\boldsymbol{D}^{-3}(\boldsymbol{d})\boldsymbol{\Theta}^{\mathsf{T}} \\ &\quad + 2\boldsymbol{\Theta}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\Theta}^{\mathsf{T}}, \end{aligned} \tag{55}$$

Algorithm 3 Constrained Total Least Squares Using Newton's Method (NM-CTLS) [47]

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$, $\epsilon_{\min} \in (0, 1)$, $N_{\max} \in \mathbb{Z}_+^{1\times 1}$, $\hat{\boldsymbol{p}}_{\text{CTLS}}$

Output: $\hat{\boldsymbol{p}}_{\text{NM-CTLS}} \in \mathbb{R}^{N\times 1}$

$n \leftarrow 0$

$\hat{\boldsymbol{v}}[0] \leftarrow \hat{\boldsymbol{p}}_{\text{CTLS}} - \boldsymbol{p}_1$

while $\epsilon_{\hat{\boldsymbol{v}}} > \epsilon_{\min} \wedge n \leq N_{\max}$ **do**

$n \leftarrow n + 1$

$$\hat{\boldsymbol{v}}[n] \leftarrow \hat{\boldsymbol{v}}[n-1] - \big(\underbrace{\boldsymbol{H}_{\text{vv}}(\hat{\boldsymbol{v}}[n-1])}_{(55)}\big)^{-1} \underbrace{\boldsymbol{g}_{\text{v}}(\hat{\boldsymbol{v}}[n-1])}_{(54)}$$

$\epsilon_{\hat{\boldsymbol{v}}} \leftarrow \frac{\|\hat{\boldsymbol{v}}[n]-\hat{\boldsymbol{v}}[n-1]\|_2}{\|\hat{\boldsymbol{v}}[n-1]\|_2}$

end while

$\hat{\boldsymbol{p}}_{\ell_2} \leftarrow \hat{\boldsymbol{v}}[n] + \boldsymbol{p}_1$

return $\hat{\boldsymbol{p}}_{\ell_2}$

where $\boldsymbol{\Theta} \in \mathbb{R}^{2\times(K-1)}$ is given by

$$\boldsymbol{\Theta} = \boldsymbol{A}_2^{\mathsf{T}} + \frac{1}{\|\boldsymbol{v}\|_2}\boldsymbol{v}\boldsymbol{a}_3^{\mathsf{T}}, \tag{56}$$

$\boldsymbol{\beta} \in \mathbb{R}^{(K-1)\times 1}$ is given by

$$\boldsymbol{\beta} = \boldsymbol{A}_2\boldsymbol{v} + \|\boldsymbol{v}\|_2\boldsymbol{a}_3 - \boldsymbol{b}. \tag{57}$$

An iterative computation using the gradient in (54) and the Hessian in (55) is as follows. At the end of Algorithm 3, one needs to convert $\hat{\boldsymbol{v}}[n]$ to $\hat{\boldsymbol{p}}_{\ell_2}$, which is deemed the second step. In Algorithm 3, there is only a single update for the Newton's iteration, rather than three updates needed in Algorithm 2. Furthermore, it does not involve the regularization parameter ρ as required in Algorithm 2. Based on the derivation in (51), the error covariance matrix of the NM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{NM-CTLS}} = \hat{\boldsymbol{p}}_{\ell_2}$ in Algorithm 3 is given by [47, eq. (20)]

$$\begin{aligned} \boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{NM-CTLS}}) &= \boldsymbol{\Sigma}_{vv}(\hat{\boldsymbol{v}}_{\text{NM-CTLS}}) \\ &= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{v}}_{\text{NM-CTLS}} - \boldsymbol{v}_0)(\hat{\boldsymbol{v}}_{\text{NM-CTLS}} - \boldsymbol{v}_0)^{\mathsf{T}}\} \\ &\approx \sigma_{\text{n}}^2\left(\boldsymbol{B}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{B}_0\right)^{-1}. \end{aligned} \tag{58}$$

where $\boldsymbol{B}_0 \in \mathbb{R}^{(K-1)\times 2}$ is given by

$$\boldsymbol{B}_0 = \boldsymbol{A}_2 + \frac{1}{\|\boldsymbol{v}_0\|_2}\boldsymbol{\delta}\boldsymbol{v}_0^{\mathsf{T}}. \tag{59}$$

The NM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{NM-CTLS}}$ in Algorithm 3 is proposed earlier than the ADMM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}}$ in Algorithm 2. Note that in [51], the NM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{NM-CTLS}}$ in Algorithm 3 is not compared to the ADMM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}}$ in Algorithm 2.

IV. PROPOSED METHODS

In this section, three new methods based on the CTLS are presented for the TDoA localization.

Algorithm 4 Fixed-Point Constrained Total Least Squares (FP-CTLS)

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$
Output: $\hat{\boldsymbol{p}}_{\text{FP-CTLS}} \in \mathbb{R}^{N\times 1}$
$n \leftarrow 0$
$\hat{\boldsymbol{p}}[0] \leftarrow \hat{\boldsymbol{p}}_{\text{LS}}$
while $\epsilon_{\hat{p}} > \epsilon_{\min} \wedge n \leq N_{\max}$ **do**
$n \leftarrow n+1$
$\hat{\boldsymbol{d}} \leftarrow \begin{bmatrix} \|\hat{\boldsymbol{p}}[n-1]-\boldsymbol{p}_2\|_2 \\ \|\hat{\boldsymbol{p}}[n-1]-\boldsymbol{p}_3\|_2 \\ \vdots \\ \|\hat{\boldsymbol{p}}[n-1]-\boldsymbol{p}_K\|_2 \end{bmatrix}$
$\hat{\boldsymbol{\theta}}_{\ell_2} \leftarrow \left(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\hat{\boldsymbol{d}})\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\hat{\boldsymbol{d}})\boldsymbol{b}$
$[\hat{\boldsymbol{p}}[n]]_1 \leftarrow [\hat{\boldsymbol{\theta}}_{\ell_2}]_1 + [\boldsymbol{p}_1]_1$
$[\hat{\boldsymbol{p}}[n]]_2 \leftarrow [\hat{\boldsymbol{\theta}}_{\ell_2}]_2 + [\boldsymbol{p}_1]_2$
$\epsilon_{\hat{p}} \leftarrow \frac{\|\hat{\boldsymbol{p}}[n]-\hat{\boldsymbol{p}}[n-1]\|_2}{\|\hat{\boldsymbol{p}}[n-1]\|_2}$
end while
return $\hat{\boldsymbol{p}}_{\ell_2}$

A. CONSTRAINED TOTAL LEAST SQUARES USING FIXED-POINT ITERATION

The fixed-point CTLS (FP-CTLS) estimation can be summarized in Algorithm 4. The FP-CTLS in Algorithm 4 entails an iterative calculation of the implicit form in (33). In each loop, it converts the first two elements residing in $\boldsymbol{\theta}$ to the mobile position $\boldsymbol{p}$.

B. CONSTRAINED TOTAL LEAST SQUARES USING NEWTON'S METHOD WITH TWO STEPS

The gradient in (54) and the Hessian in (55) appear to be complicated, leading to computational demands. In what follows, we derive a new simpler expression of both quantities. Substituting (53) into (31), we obtain

$$f_{\text{CTLS}}(\boldsymbol{v}) = \left(\boldsymbol{A}\begin{bmatrix}\boldsymbol{v}\\ \|\boldsymbol{v}\|_2\end{bmatrix} - \boldsymbol{b}\right)^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d}) \left(\boldsymbol{A}\begin{bmatrix}\boldsymbol{v}\\ \|\boldsymbol{v}\|_2\end{bmatrix} - \boldsymbol{b}\right). \tag{60}$$

By using the chain rule, one may find that

$$\begin{aligned}\boldsymbol{g}_v(\boldsymbol{v}) &= \frac{\partial}{\partial \boldsymbol{p}} f_{\text{CTLS}}(\boldsymbol{v}) \\ &= \boldsymbol{J}_{v\theta}(\boldsymbol{v}) \underbrace{\boldsymbol{g}_\theta(\boldsymbol{\theta})}_{=2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b});\ (32)} \\ &= 2\boldsymbol{J}_{v\theta}(\boldsymbol{p})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}),\end{aligned} \tag{61}$$

and

$$\begin{aligned}\boldsymbol{H}_{vv}(\boldsymbol{v}) &= \frac{\partial^2}{\partial \boldsymbol{v}\partial \boldsymbol{v}^{\mathsf{T}}} f_{\text{CTLS}}(\boldsymbol{v}) \\ &= \boldsymbol{J}_{v\theta}(\boldsymbol{p}) \underbrace{\boldsymbol{H}_{\theta\theta}(\boldsymbol{\theta})}_{=2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A};\ (34)} \boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{p}) \\ &= 2\boldsymbol{J}_{v\theta}(\boldsymbol{p})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{p}),\end{aligned} \tag{62}$$

where $\boldsymbol{J}_{v\theta} \in \mathbb{R}^{2\times 3}$ is the Jacobian of $\boldsymbol{\theta}$ with respect to $\boldsymbol{v}$, which can be shown as

$$\begin{aligned}\boldsymbol{J}_{v\theta}(\boldsymbol{v}) &= \frac{\partial}{\partial \boldsymbol{v}}\boldsymbol{\theta}^{\mathsf{T}} \\ &= \left[\tfrac{\partial}{\partial \boldsymbol{v}}\boldsymbol{v}^{\mathsf{T}} \quad \tfrac{\partial}{\partial \boldsymbol{v}}\|\boldsymbol{v}\|_2\right] \\ &= \left[\boldsymbol{I}_2 \quad \tfrac{1}{d_1}\boldsymbol{v}\right].\end{aligned} \tag{63}$$

The incorporation of (61) and (62) into Algorithm 3 leads to a new algorithm, which can be referred to as Constrained Total Least Squares Using Newton's Method with Two Steps (TSNM-CTLS).

Algorithm 5 Constrained Total Least Squares Using Newton's Method With Two Steps (TSNM-CTLS)

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$, $\epsilon_{\min} \in (0, 1)$, $N_{\max} \in \mathbb{Z}_+^{1\times 1}$, $\hat{\boldsymbol{p}}_{\text{CTLS}}$
Output: $\hat{\boldsymbol{p}}_{\text{NM-CTLS}} \in \mathbb{R}^{N\times 1}$
$n \leftarrow 0$
$\hat{\boldsymbol{v}}[0] \leftarrow \hat{\boldsymbol{p}}_{\text{CTLS}} - \boldsymbol{p}_1$
while $\epsilon_{\hat{v}} > \epsilon_{\min} \wedge n \leq N_{\max}$ **do**
$n \leftarrow n+1$
$\hat{\boldsymbol{v}}[n] \leftarrow \hat{\boldsymbol{v}}[n-1] - \left(\boldsymbol{J}_{v\theta}(\boldsymbol{v})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{v})\right)^{-1}$
$\boldsymbol{J}_{v\theta}(\boldsymbol{v})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\left(\boldsymbol{A}\begin{bmatrix}\boldsymbol{v}\\ \|\boldsymbol{v}\|_2\end{bmatrix} - \boldsymbol{b}\right)\Big|_{\boldsymbol{v}=\hat{\boldsymbol{v}}[n-1]}$
$\epsilon_{\hat{v}} \leftarrow \frac{\|\hat{\boldsymbol{v}}[n]-\hat{\boldsymbol{v}}[n-1]\|_2}{\|\hat{\boldsymbol{v}}[n-1]\|_2}$
end while
$\hat{\boldsymbol{p}}_{\ell_2} \leftarrow \hat{\boldsymbol{v}}[n] + \boldsymbol{p}_1$
return $\hat{\boldsymbol{p}}_{\ell_2}$

The Jacobian notion is introduced in Algorithm 5 in order to differentiate Algorithm 5 from Algorithm 3. Based on the derivation in (51), the error covariance matrix of the TSNM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{TSNM-CTLS}} = \hat{\boldsymbol{p}}_{\ell_2}$ in Algorithm 5 can borrow the results in (37), yielding

$$\begin{aligned}&\boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{TSNM-CTLS}}) \\ &= \boldsymbol{\Sigma}_{vv}(\hat{\boldsymbol{v}}_{\text{TSNM-CTLS}}) \\ &= \mathcal{E}_{\mathbf{n}}\left\{(\hat{\boldsymbol{v}}_{\text{TSNM-CTLS}} - \boldsymbol{v}_0)(\hat{\boldsymbol{v}}_{\text{TSNM-CTLS}} - \boldsymbol{v}_0)^{\mathsf{T}}\right\} \\ &\approx \sigma_{\text{n}}^2 \left(\boldsymbol{J}_{v\theta}(\boldsymbol{v}_0)\boldsymbol{Q}_1\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{v}_0)\right)^{-1} \boldsymbol{J}_{v\theta}(\boldsymbol{v}_0)\boldsymbol{Q}_2\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{v}_0) \\ &\quad \left(\boldsymbol{J}_{v\theta}(\boldsymbol{v}_0)\boldsymbol{Q}_1\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{v}_0)\right)^{-1}.\end{aligned} \tag{64}$$

As same as (41), one can derive

$$\begin{aligned}&\lim_{\sigma_{\text{n}}^2\to 0} \boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{TSNM-CTLS}}) \\ &\quad \approx \sigma_{\text{n}}^2 \left(\boldsymbol{J}_{v\theta}(\boldsymbol{v}_0)\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\boldsymbol{J}_{v\theta}^{\mathsf{T}}(\boldsymbol{v}_0)\right)^{-1}.\end{aligned} \tag{65}$$

In the next subsection, a new approach is presented without the conversion of $\hat{\boldsymbol{v}}$ to $\hat{\boldsymbol{p}}_{\ell_2}$ as performed in both Algorithm 3 and Algorithm 5.

C. CONSTRAINED TOTAL LEAST SQUARES USING NEWTON'S METHOD WITH ONE STEP

As $d_1 = \|\boldsymbol{p} - \boldsymbol{p}_1\|_2$, another form of (7) can be written as

$$\boldsymbol{\theta}_0 = \begin{bmatrix} \boldsymbol{p}_0 - \boldsymbol{p}_1 \\ \|\boldsymbol{p}_0 - \boldsymbol{p}_1\|_2 \end{bmatrix}. \tag{66}$$

The Jacobian $\boldsymbol{J}_{p\theta} \in \mathbb{R}^{2\times 3}$ can be shown as

$$\begin{aligned} \boldsymbol{J}_{p\theta}(\boldsymbol{p}) &= \frac{\partial}{\partial \boldsymbol{p}} \boldsymbol{\theta}^{\mathsf{T}} \\ &= \begin{bmatrix} \frac{\partial}{\partial \boldsymbol{p}} (\boldsymbol{p} - \boldsymbol{p}_1)^{\mathsf{T}} & \frac{\partial}{\partial \boldsymbol{p}} \|\boldsymbol{p} - \boldsymbol{p}_1\|_2 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{I}_2 & \frac{1}{d_1}(\boldsymbol{p} - \boldsymbol{p}_1) \end{bmatrix}. \end{aligned} \tag{67}$$

The optimization search in (30) involves three variables in $\boldsymbol{\theta}$. Note that this search is redundant, because it also includes the additional parameter d_1, which actually depends on $\boldsymbol{p}$. We propose an alternative estimation that entails only two variables without the dedicated d_1, i.e.

$$\begin{aligned} \hat{\boldsymbol{p}}_{\ell_2} = \arg\min_{\boldsymbol{p}} &\left(\boldsymbol{A} \begin{bmatrix} \boldsymbol{p} - \boldsymbol{p}_1 \\ \|\boldsymbol{p} - \boldsymbol{p}_1\|_2 \end{bmatrix} - \boldsymbol{b} \right)^{\mathsf{T}} \boldsymbol{D}^{-2}(\boldsymbol{d}) \\ &\times \left(\boldsymbol{A} \begin{bmatrix} \boldsymbol{p} - \boldsymbol{p}_1 \\ \|\boldsymbol{p} - \boldsymbol{p}_1\|_2 \end{bmatrix} - \boldsymbol{b} \right). \end{aligned} \tag{68}$$

The problem in (68) is nonconvex for the variable $\boldsymbol{p}$ and involves several local minima of the variable $\boldsymbol{p}$. It is hard to find a closed-form solution of (68) in a similar way to (33). Approximately, given a good initialization $\hat{\boldsymbol{p}}[0]$, we would rather compute the solution $\hat{\boldsymbol{p}}_{\ell_2}$ in an iterative manner according to the Newton's method. By using the chain rule, one can find that

$$\begin{aligned} \boldsymbol{g}_p(\boldsymbol{p}) &= \frac{\partial}{\partial \boldsymbol{p}} f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \boldsymbol{J}_{p\theta}(\boldsymbol{p}) \underbrace{\qquad \boldsymbol{g}_\theta(\boldsymbol{\theta}) \qquad}_{=2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b});\ (32)} \\ &= 2\boldsymbol{J}_{p\theta}(\boldsymbol{p})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{b}), \end{aligned} \tag{69}$$

and

$$\begin{aligned} \boldsymbol{H}_{pp}(\boldsymbol{p}) &= \frac{\partial^2}{\partial \boldsymbol{p} \partial \boldsymbol{p}^{\mathsf{T}}} f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \boldsymbol{J}_{p\theta}(\boldsymbol{p}) \underbrace{\quad \boldsymbol{H}_{\theta\theta}(\boldsymbol{\theta}) \quad}_{=2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A};\ (34)} \boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}) \\ &= 2\boldsymbol{J}_{p\theta}(\boldsymbol{p})\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}). \end{aligned} \tag{70}$$

For a maximum number of iterations, denoted by $N_{\max}$, and for $n \in \{1, 2, \ldots, N_{\max}\}$, the Newton's iteration reads as

$$\begin{aligned} \hat{\boldsymbol{p}}[n+1] = \hat{\boldsymbol{p}}[n] - \Big(& \underbrace{\qquad \boldsymbol{H}_{pp}(\hat{\boldsymbol{p}}[n]) \qquad}_{=2\boldsymbol{J}_{p\theta}(\hat{\boldsymbol{p}}[n])\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\hat{\boldsymbol{p}}[n]);\ (70)} \Big)^{-1} \\ & \underbrace{\boldsymbol{g}_p(\hat{\boldsymbol{p}}[n])}_{=2\boldsymbol{J}_{p\theta}(\hat{\boldsymbol{p}}[n])\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b});\ (69)} \\ = \hat{\boldsymbol{p}}[n] - & \left(\boldsymbol{J}_{p\theta}(\hat{\boldsymbol{p}}[n])\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\hat{\boldsymbol{p}}[n]) \right)^{-1} \end{aligned}$$

Algorithm 6 Constrained Total Least Squares Using Newton's Method With One Step (OSNM-CTLS)

Input: $\boldsymbol{A} \in \mathbb{R}^{(K-1)\times 3}$, $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$, $\boldsymbol{p}_k \in \mathbb{R}^{2\times 1}$ for $k \in \{1, 2, \ldots, K\}$, $\epsilon_{\min} \in (0, 1)$, $N_{\max} \in \mathbb{Z}_+^{1\times 1}$, $\hat{\boldsymbol{p}}_{\text{CTLS}}$
Output: $\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}} \in \mathbb{R}^{N\times 1}$

```
n ← 0
p̂[0] ← p̂_CTLS
while ε_p̂ > ε_min ∧ n ≤ N_max do
    n ← n + 1
    p̂[n] ← p̂[n − 1] − ( J_pθ(p) A^T D^{-2}(d) A J_pθ^T(p) )^{-1}
                J_pθ(p) A^T D^{-2}(d) ( A [ (p − p_1)/‖p − p_1‖_2 ] − b ) |_{p=p̂[n−1]}
    ε_p̂ ← ‖p̂[n] − p̂[n−1]‖_2 / ‖p̂[n−1]‖_2
end while
p̂_{ℓ_2} ← p̂[n]
return p̂_{ℓ_2}
```

$$\times \boldsymbol{J}_{p\theta}(\hat{\boldsymbol{p}}[n])\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d}) \left(\boldsymbol{A} \begin{bmatrix} \hat{\boldsymbol{p}}[n] - \boldsymbol{p}_1 \\ \|\hat{\boldsymbol{p}}[n] - \boldsymbol{p}_1\|_2 \end{bmatrix} - \boldsymbol{b} \right). \tag{71}$$

This iterative computation can be proposed as follows. Similar to (37), the error covariance matrix of the OSNM-CTLS estimate $\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}} = \hat{\boldsymbol{p}}_{\ell_2}$ in (68) is given by

$$\begin{aligned} &\boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}}) \\ &= \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}} - \boldsymbol{p}_0)(\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}} - \boldsymbol{p}_0)^{\mathsf{T}}\} \\ &\approx \sigma_{\text{n}}^2 \left(\boldsymbol{J}_{p\theta}(\boldsymbol{p}_0)\boldsymbol{Q}_1\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}_0) \right)^{-1} \boldsymbol{J}_{p\theta}(\boldsymbol{p}_0)\boldsymbol{Q}_2\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}_0) \\ &\quad \times \left(\boldsymbol{J}_{p\theta}(\boldsymbol{p}_0)\boldsymbol{Q}_1\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}_0) \right)^{-1}. \end{aligned} \tag{72}$$

As same as (41), one can derive

$$\begin{aligned} &\lim_{\sigma_{\text{n}}^2 \to 0} \boldsymbol{\Sigma}_{pp}(\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}}) \\ &\approx \sigma_{\text{n}}^2 \left(\boldsymbol{J}_{p\theta}(\boldsymbol{p}_0)\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\boldsymbol{J}_{p\theta}^{\mathsf{T}}(\boldsymbol{p}_0) \right)^{-1}. \end{aligned} \tag{73}$$

At the true value $\boldsymbol{\theta}_0$, the Jacobian $\boldsymbol{J}_{v\theta}(\boldsymbol{v}_0)$ in (63) is equal to the Jacobian $\boldsymbol{J}_{p\theta}(\boldsymbol{p}_0)$ in (67). Therefore, the error covariance matrix in (64) is exactly the same as that in (72), as well as the simplified expression in (65) is exactly the same as that in (73). Hence, it is sufficient to investigate only each of these coincident covariance matrices.

The design of the TSNM-CTLS and the OSNM-CTLS methods comes from investigating the derivations of the gradient and the Hessian of the CTLS objective function in a previous work, i.e. the NM-CTLS in [47]. After an exploration, we end up using the Jacobian matrix. This technique leads to a smoother iteration for Newton's method. The smoothness is in the sense that there is no ill condition of the Hessian matrix that can cause the deviation of the search from the desired solution. The nature of Newton's method is that if Newton's iteration can be well updated, it will more quickly and more precisely reach the desired solution.

TABLE 4. The novelty of the related works and the proposed methods and their motivations.

Algorithms	References	Novelty	Motivations
LS	[31]	closed-form solution	simplicity
NM-CTLS	[47]	CTLS using NM	fast convergence
ADMM-CTLS	[51]	CTLS using ADMM	constraint manipulation
FP-CTLS	proposed herein	CTLS using fixed point	straightforward iteration
TSNM-CTLS	proposed herein	two-dimensional search	straightforward derivatives
OSNM-CTLS	proposed herein	two-dimensional search	straightforward derivatives

TABLE 5. Number of realizations for each random variable.

Notation	Number of realizations for
N_{R}	the random variable $\boldsymbol{n}$

The quick iterative computation can be resulted from less number of iterations and/or less computational burden in each iteration. This is the reason why one will see that the TSNM-CTLS and the OSNM-CTLS methods consume less computational time than the NM-CTLS approach. More precision to the solution of both methods can be observed from the simulation results in terms of lower bias and RMSE than the NM-CTLS approach. Hence, our objective is to improve Newton's iteration method, thus allowing for both more mobile position estimation accuracy and lower computational cost.

Complexity analysis of Newton-based methods, such as Algorithm 3, Algorithm 5, and Algorithm 6 is as follows. Let $O(\cdot)$ be the big 'o' notation of Bachman-Landau symbols [63], [64]. For any Newton's method, the n-th iteration requires $\frac{1}{6}N_p^3 + O(N_p^2)$ multiplications per iteration [65, p. 44], where N_p is the size of the unknown position vector $\boldsymbol{p}$ in Algorithm 6 or $\boldsymbol{v}$ in Algorithm 3 and Algorithm 5. One can see that for a two-dimensional localization problem, there are two unknown variables, such as x and y, i.e. $N_p = 2$.

The convergence of Newton-based methods, such as Algorithm 3, Algorithm 5, and Algorithm 6 can be observed as follows. If the initialized values for either $\boldsymbol{p}$ or $\boldsymbol{v}$ are sufficiently close to their true values for some iteration indices n and if the true values of the Hessian matrices are positive definite, the Newton's method converges at the second order [65, p. 46].

V. NUMERICAL EXAMPLES

A. ALGORITHMIC COMPARISON

For any value of the first sensor position $\boldsymbol{p}_1$, the first 2×2 sub-block matrix of $\boldsymbol{\Sigma}_{\theta\theta}$ represents the error covariance of the mobile position estimate given by the CTLS in (33). We refer to

- the square root of the trace of the first 2×2 sub-block matrix of $\boldsymbol{\Sigma}_{\theta\theta}$ in (37) as CTLS approximation,
- the square root of the trace of the first 2×2 sub-block matrix of $\boldsymbol{\Sigma}_{\theta\theta}$ in (41) as simplified CTLS approximation,
- the square root of the trace of $\boldsymbol{\Sigma}_{vv}$ in (58) as NM-CTLS approximation,
- the square root of the trace of $\boldsymbol{\Sigma}_{pp}$ in (72) as OSNM-CTLS approximation, and
- the square root of the trace of $\boldsymbol{\Sigma}_{pp}$ in (73) as simplified OSNM-CTLS approximation.

Note that except for (58), the above four quantities are, to our best knowledge, unavailable in the former works. All these four amounts are provided herein in order to verify whether our estimation results coincide with their theoretical approximations or not. Regarding the estimation methods, we refer to

- the output $\hat{\boldsymbol{p}}_{\text{FP-CTLS}}$ of Algorithm 4 as FP-CTLS estimate,
- the output $\hat{\boldsymbol{p}}_{\text{ADMM-CTLS}}$ of Algorithm 2 as ADMM-CTLS estimate,
- the output $\hat{\boldsymbol{p}}_{\text{NM-CTLS}}$ of Algorithm 3 as NM-CTLS estimate,
- the output $\hat{\boldsymbol{p}}_{\text{TSNM-CTLS}}$ of Algorithm 5 as TSNM-CTLS estimate, and
- the output $\hat{\boldsymbol{p}}_{\text{OSNM-CTLS}}$ of Algorithm 6 as OSNM-CTLS estimate.

The initialization values, such as $\hat{\boldsymbol{v}}[0]$ required in Algorithm 2, Algorithm 3, and Algorithm 5, as well as $\hat{\boldsymbol{p}}[0]$ needed in Algorithm 4 and Algorithm 6, are taken from the LS estimate $\hat{\boldsymbol{\theta}}_{\text{LS}}$ in (24). We use $N_{\max} = 100$ and $\epsilon_{\min} = 10^{-6}$ in all iterative algorithms and $\rho = 10^{-6}$ in Algorithm 2.

B. SIMULATION RESULTS

Similar to [51, Sec. IV], we divide the simulation into two scenarios, such as the case in which the standard deviation of the TDoA error changes and the case in which the number of sensors increases.

1) EFFECT OF THE STANDARD DEVIATION OF THE TDOA ERROR

We consider the TDoA localization that has $K = 7$ sensors. For the (x, y)-coordinates in meter, the sensor location entails [51, Sec. IV.B]

$$\begin{aligned}\boldsymbol{P} &= \begin{bmatrix}\boldsymbol{p}_1 & \boldsymbol{p}_2 & \cdots & \boldsymbol{p}_K\end{bmatrix} \\ &= \begin{bmatrix}0 & -260 & -80 & 160 & 360 & 700 & -200 \\ 0 & 300 & 400 & 180 & 60 & -120 & 730\end{bmatrix},\end{aligned} \tag{74}$$

and the mobile user is at

$$\boldsymbol{p} = \begin{bmatrix}1{,}200 \\ 800\end{bmatrix}. \tag{75}$$

In Fig. 2, the bias norm given by each algorithm is shown as a function of TDoA error standard deviation. The greater

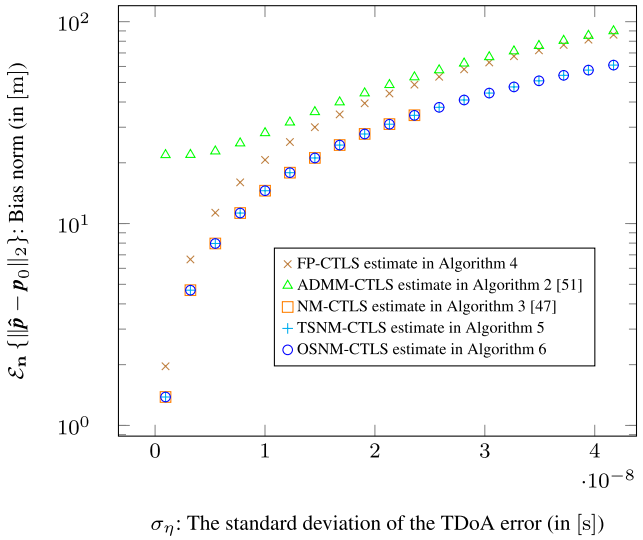

FIGURE 2. Bias norm as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

the TDoA error standard deviation, the larger the bias norm we can get from each method. In terms of the first-order error, i.e. bias, one can see that the proposed TSNM-CTLS and OSNM-CTLS approaches provide the lowest error, followed by the FP-CTLS and the ADMM-CTLS method, respectively. The NM-CTLS algorithm is the same as the proposed TSNM-CTLS and OSNM-CTLS approaches for a small TDoA error variance. It heavily deviates from other algorithms for a large TDoA error variance. This unexpected result is caused by the ill condition of the Hessian matrix, leading to a large perturbation during its inverse.

In Fig. 3, the error performance on the second order in terms of the RMSE is shown as a function of the TDoA error standard deviation. One can see that the simplified CTLS approximation in (41) can well capture the RMSE by the FP-CTLS estimate in Algorithm 4, while the simplified OSNM-CTLS approximation in (73) well coincides with the proposed TSNM-CTLS estimate in Algorithm 5 and the proposed OSNM-CTLS estimate in Algorithm 6. On the other hands, the theoretical prediction of the NM-CTLS in (58) has a significant mismatch with respect to the actual estimation by Algorithm 3.

In Fig. 4, one can clearly see that the proposed TSNM-CTLS estimation in Algorithm 5 and the proposed OSNM-CTLS estimation in Algorithm 6 provide lower RMSE than the ADMM-CTLS estimate in Algorithm 2 presented in [51], which has higher RMSE than the FP-CTLS estimate in Algorithm 4. The RMSE predictions directly given by the FOTSE in (37) and (72) do not exactly match the actual estimation errors. This mismatch is obvious when the TDoA error standard deviation is large.

In Fig. 5, the computational time consumed by the ADMM-CTLS is greatest, followed by the NM-CTLS, the TSNM-CTLS, the OSNM-CTLS, and the FP-CTLS, respectively. This is because

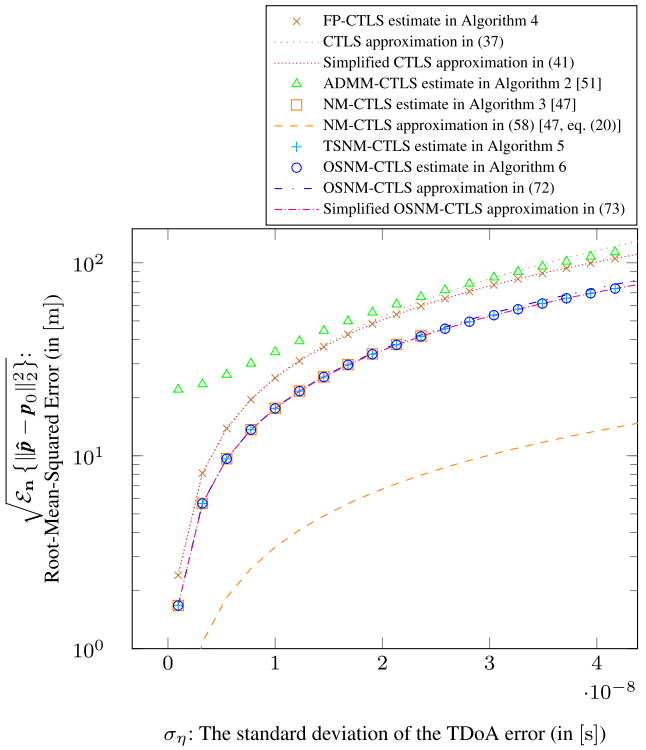

FIGURE 3. RMSE as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

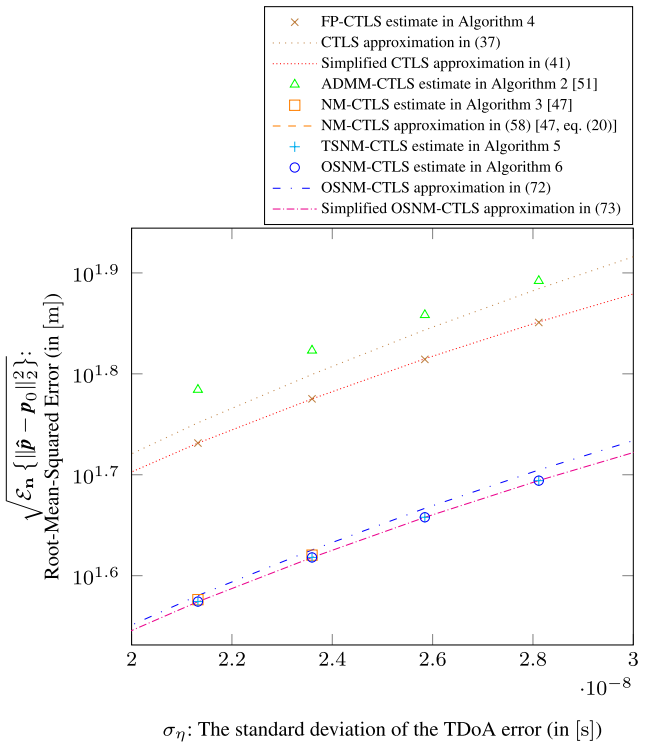

FIGURE 4. A zoom of Fig. 3, which represents the RMSE as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

- the computational time taken by all iterative methods includes the closed-form LS estimation time due to the use of their initializing values, such as $\hat{\boldsymbol{v}}[0]$ in

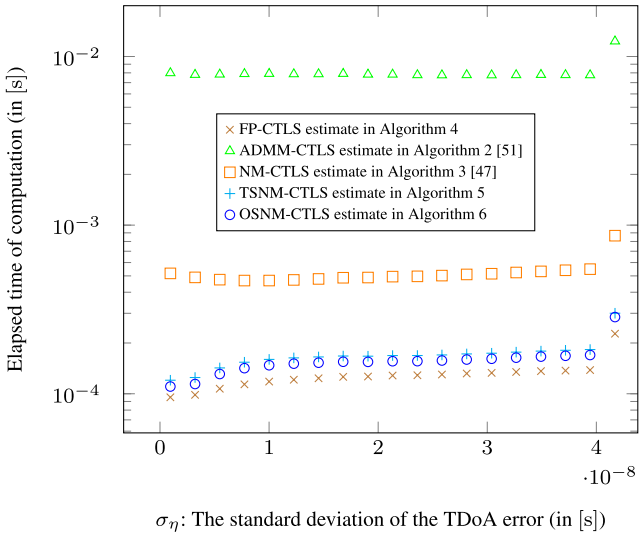

FIGURE 5. Computational time as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

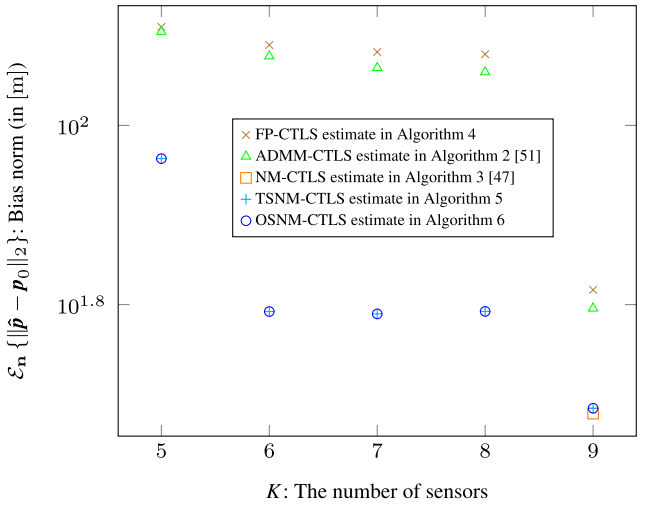

FIGURE 6. Bias norm as a function of the number of sensors from $N_R = 1{,}000{,}000$ independent runs.

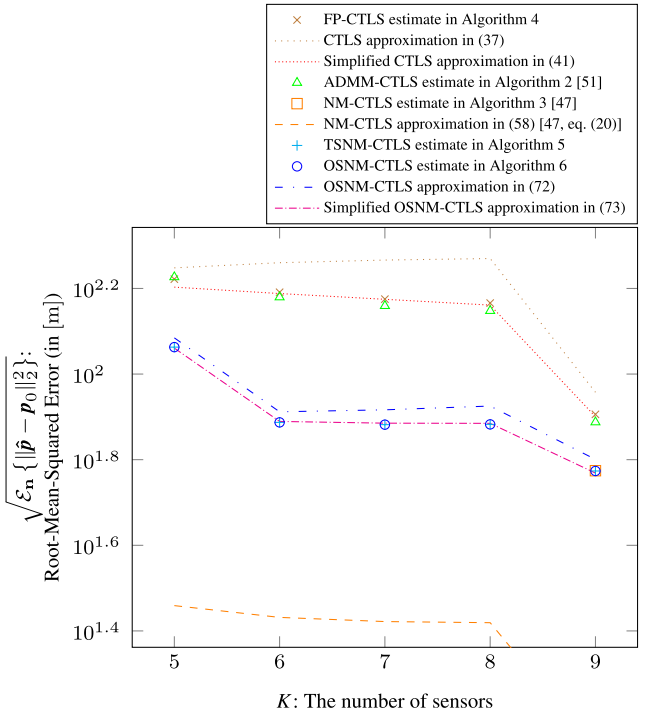

FIGURE 7. RMSE as a function of the number of sensors from $N_R = 1{,}000{,}000$ independent runs.

Algorithm 2, Algorithm 3, and Algorithm 5 as well as $\hat{\boldsymbol{p}}[0]$ in Algorithm 4 and Algorithm 6, respectively, and
- both the ADMM-CTLS and the NM-CTLS incur heavy Newton's iterations.

The NM-CTLS estimate in Algorithm 3, the FP-CTLS estimate in Algorithm 4, the TSNM-CTLS estimate in Algorithm 5, and the OSNM-CTLS estimate in Algorithm 6 take less time in computation than the ADMM-CTLS estimate in Algorithm 2, because they have only an update, compared to three updates required by the ADMM-CTLS estimate in Algorithm 2. Especially, the two updates in Algorithm 2 are computationally intensive due to the nature of Newton's iteration. It is worth noting that the OSNM-CTLS estimate in Algorithm 6 consumes less computational time than the TSNM-CTLS estimate in Algorithm 5.

2) EFFECT OF THE NUMBER OF SENSORS

We consider the TDoA localization that varies the number of participating sensors from $K = 5$ to $K = 9$ sensors. For the (x, y)-coordinates in meter, the sensor location entails [51, Sec. IV.C]

$$\boldsymbol{P} = \begin{bmatrix} 0 & -260 & -80 & 160 & 360 & 640 & 320 & 220 & 370 \\ 0 & 300 & 400 & 180 & 60 & 280 & 150 & 120 & -30 \end{bmatrix}, \tag{76}$$

and the true position of the mobile user is the same as that in (75).

In Fig. 6, the norm of the biases caused by each method is shown as a function of the number of participating sensors. One can see that when more information of the TDoA from the sensors is available, the localization system can gain more accuracy in terms of bias.

In Fig. 7, the RMSE is shown as a function of the number of sensors K. Similar to Fig. 6, one can see that when the wireless sensor network receives more knowledge of the distances, its localization performance in terms of the RMSE improves.

Again, in Fig. 8, the ADMM-CTLS encounters intensive updates in the Newton's iteration, thus consuming the greatest amount of the computational time. It is followed by the NM-CTLS, which also involves Newton's iteration yet with less computational burden. Note that when the number of sensors is sufficiently large, the inverse of the Hessian matrix inside the Newton's iteration is not stringent. In this situation, the TSNM-CTLS method takes less computational time than the

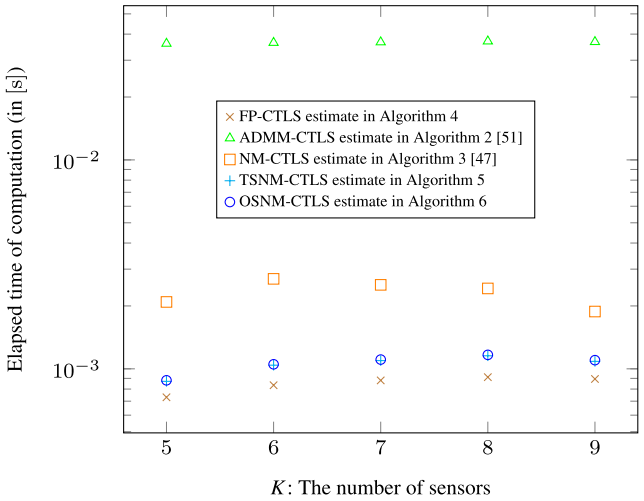

FIGURE 8. Computational time as a function of the number of sensors from $N_R = 1{,}000{,}000$ independent runs.

OSNM-CTLS approach. The FP-CTLS consumes the least amount of computational time.

3) EFFECT OF THE IMPERFECT SYNCHRONIZATION

In reality, multipath and shadowing may arise in electromagnetic wave propagation. Furthermore, the first sensor and the k-th sensor may not share the same exact time in their clocks [66]. These effects can be realized as an additional term to the right-hand side of (11) according to [67, eq. (5)]

$$\begin{aligned}\tilde{\tau}_{k,1} &= \underbrace{\tau_{k,1}}_{=\frac{1}{c}\delta_{k,1};\ (9)} + \underbrace{\eta_{k,1}}_{=\frac{1}{c}n_{k,1};\ (10)} + \delta_{\tau_{k,1}} \\ &= \frac{1}{c}(\delta_{k,1} + n_{k,1}) + \delta_{\tau_{k,1}}, \end{aligned} \tag{77}$$

where $\delta_{\tau_{k,1}} \in \mathrm{R}^{1\times 1}$ is the offset time between the first sensor and the k-th sensor caused by the synchronization error and the propagation fluctuation, such as multipath and shadowing. We shall see the characterization and the deterioration of all algorithms considered in the former comparison, when the offset time $\delta_{\tau_{k,1}}$ occurs in the previous simulation setups and is not compensated or removed. The estimation and the compensation of the unknown offset time $\delta_{\tau_{k,1}}$ are beyond the scope of this paper and can be performed in terms of, e.g., multi-hop synchronization, etc. [68].

Following from (13), we find that the noisy value of the true value $\boldsymbol{A}_0$ now becomes

$$\begin{aligned}\boldsymbol{A} &= \begin{bmatrix} x_2 - x_1 & y_2 - y_1 & \delta_{2,1} + n_{2,1} + c\delta_{\tau_{2,1}} \\ x_3 - x_1 & y_3 - y_1 & \delta_{3,1} + n_{3,1} + c\delta_{\tau_{3,1}} \\ \vdots & \vdots & \vdots \\ x_K - x_1 & y_K - y_1 & \delta_{K,1} + n_{K,1} + c\delta_{\tau_{K,1}} \end{bmatrix} \\ &= \boldsymbol{A}_0 + \boldsymbol{\Delta}_{\mathrm{A}}, \end{aligned} \tag{78}$$

where the perturbation matrix $\boldsymbol{\Delta}_{\mathrm{A}} \in \mathbb{R}^{(K-1)\times 3}$ is given by

$$\begin{aligned}\boldsymbol{\Delta}_{\mathrm{A}} &= \begin{bmatrix} 0 & 0 & n_{2,1} + c\delta_{\tau_{2,1}} \\ 0 & 0 & n_{3,1} + c\delta_{\tau_{3,1}} \\ \vdots & \vdots & \vdots \\ 0 & 0 & n_{K,1} + c\delta_{\tau_{K,1}} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} + c\boldsymbol{\delta}_{\tau} \end{bmatrix}, \end{aligned} \tag{79}$$

with $\boldsymbol{\delta}_{\tau} \in \mathbb{R}^{(K-1)\times 1}$ is the vector that contains $K-1$ time offsets, i.e.

$$\boldsymbol{\delta}_{\tau} = \begin{bmatrix} \delta_{\tau_{2,1}} \\ \delta_{\tau_{3,1}} \\ \vdots \\ \delta_{\tau_{K,1}} \end{bmatrix}. \tag{80}$$

The vector $\boldsymbol{b} \in \mathbb{R}^{(K-1)\times 1}$ can be expressed with time offsets as

$$\begin{aligned}\boldsymbol{b} &= \frac{1}{2}\begin{bmatrix} (x_2 - x_1)^2 + (y_2 - y_1)^2 - (\delta_{2,1} + n_{2,1} + c\delta_{\tau_{2,1}})^2 \\ (x_3 - x_1)^2 + (y_3 - y_1)^2 - (\delta_{3,1} + n_{3,1} + c\delta_{\tau_{3,1}})^2 \\ \vdots \\ (x_K - x_1)^2 + (y_K - y_1)^2 - (\delta_{K,1} + n_{K,1} + c\delta_{\tau_{K,1}})^2 \end{bmatrix} \\ &= \boldsymbol{b}_0 + \boldsymbol{\delta}_{\mathrm{b}}, \end{aligned} \tag{81}$$

where the perturbation $\boldsymbol{\delta}_{\mathrm{b}}$ can be approximated as

$$\begin{aligned}\boldsymbol{\delta}_{\mathrm{b}} &= -\boldsymbol{D}(\boldsymbol{\delta})(\boldsymbol{n} + c\boldsymbol{\delta}_{\tau}) - \frac{1}{2}\boldsymbol{D}(\boldsymbol{n} + c\boldsymbol{\delta}_{\tau})(\boldsymbol{n} + c\boldsymbol{\delta}_{\tau}) \\ &\approx -\boldsymbol{D}(\boldsymbol{\delta})(\boldsymbol{n} + c\boldsymbol{\delta}_{\tau}). \end{aligned} \tag{82}$$

For simplicity, we assume that the synchronization and propagation errors in $\boldsymbol{\delta}_{\tau}$ are uniform with

$$\delta_{\tau_{2,1}} = \delta_{\tau_{3,1}} = \ldots = \delta_{\tau_{K,1}} = \delta_{\tau}, \tag{83}$$

thus resulting in

$$\boldsymbol{\delta}_{\tau} = \delta_{\tau}\boldsymbol{1}_{K-1}, \tag{84}$$

where $\delta_{\tau} \in \mathbb{R}^{1\times 1}$ is a common time offset and $\boldsymbol{1}_N$ is the column vector containing only ones of size N. In the mismatch simulation caused by the asynchronization, multipath, and shadowing, let the offset time be $\delta_{\tau} = 20$ [ns].

Figs. 9 and 10 replicate the simulation framework of Figs. 2 - 5, except that the timing offset in (77) is assumed to occur instead of adopting the perfect synchronization in (11).

In Fig. 9, for σ_{η} less than 30 [ns], the FP-CTLS bias is lower than those by the other methods and the typical NM-CTLS bias is slightly lower than the proposed TSNM-CTLS and OSNM-CTLS algorithms. However, for σ_{η} greater than 30 [ns], the proposed TSNM-CTLS and OSNM-CTLS algorithms outperform the other methods. It is worth noting that the proposed FP-CTLS method and the ADMM-CTLS approach do not significantly suffer from the existence of the time offset.

In Fig. 10, the RMSE tendency is similar to the bias trend in Fig. 9. However, most derived theoretical predictions of the RMSE fail to capture the actual RMSE from the estimation,

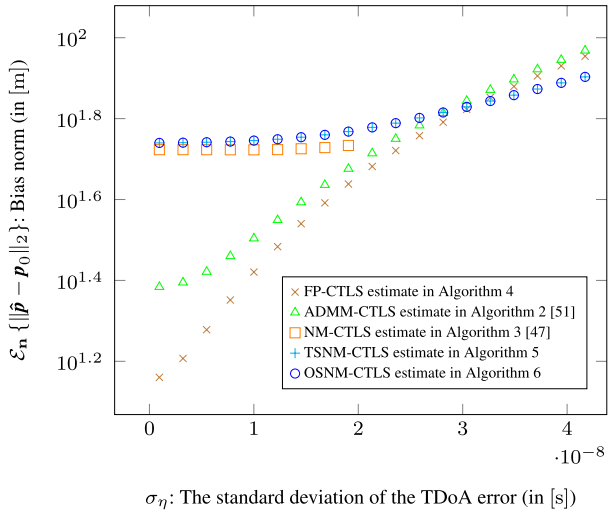

FIGURE 9. Bias norm as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

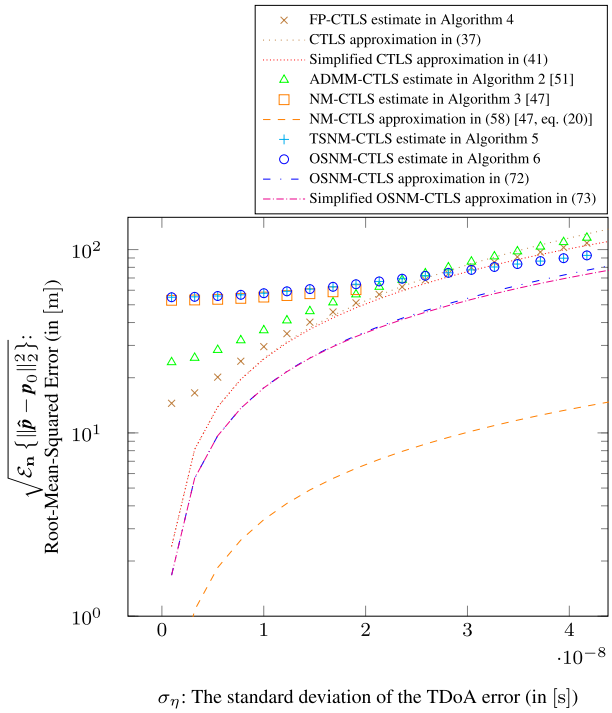

FIGURE 10. RMSE as a function of TDoA error standard deviation σ_η from $N_R = 1{,}000{,}000$ independent runs.

except for the coincidence between the FP-CTLS RMSE and the simplified CTLS approximation in (41) for σ_η greater than 20 [ns]. Note that in this region, the conventional NM-CTLS method is unreliable to provide the acceptable RMSE.

Figs. 11 and 12 represent the error performance of all localization algorithms in Figs. 6, 7, and 8, where the number of sensors changes according to (76). Their simulation setup is similar to that in Figs. 6 - 8, except that the TDoA offset in (77) is taken into account.

For $\sigma_\tau = 4$ [ns], in Fig. 11, the proposed methods, such as the TSNM-CTLS and the OSNM-CTLS approaches

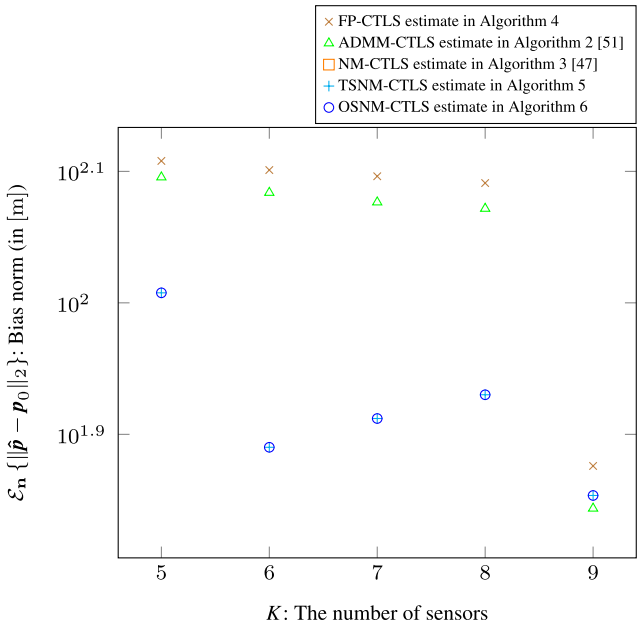

FIGURE 11. Bias norm as a function of the number of sensors from $N_R = 1{,}000{,}000$ independent runs.

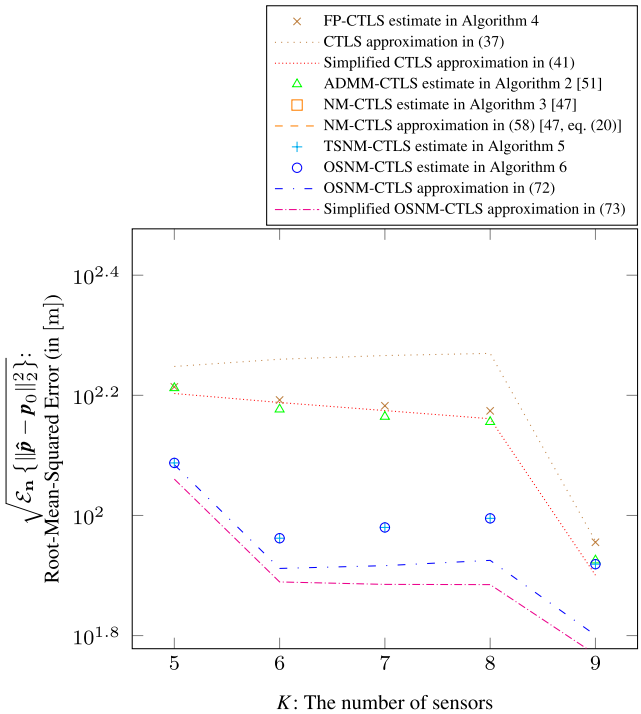

FIGURE 12. RMSE as a function of the number of sensors from $N_R = 1{,}000{,}000$ independent runs.

provide the lowest bias and are followed by the ADMM-CTLS and the FP-CTLS algorithms, respectively.

In Fig. 12, when the time offset of 4 [ns] exists, the sophisticated TSNM-CTLS and OSCTLS algorithms still work well but they deviate from their theoretical RMSE prediction. The ADMM-CTLS method coincides with the above two algorithms when the number of involved sensors is 9.

VI. CONCLUSION

We propose three alternative techniques for the CTLS solved by the fixed-point iteration and the NM. By using the chain rules, the new derived gradient vectors and Hessian matrices can be elaborated in compact forms, thus leading to i) less computational burden than the former approaches and ii) more computational reliability, in terms of no degenerate matrix inverse caused by the ill condition of the Hessian matrix, than the previous methods. Theoretical error performance prediction is also provided for the conventional CTLS estimate and each proposed algorithm. It is validated through numerical simulation that the simplified RMSE expressions well comply with the corresponding actual estimation results. Future work would be the design of an algorithm that can handle the imperfect synchronization raised in (77). This idea can be done by keeping the time offset at the beginning, i.e. the formulation of the CTLS objective function.

APPENDIX A
PRELIMINARY RESULTS

Based on (136), the expectations of some scalars for $\boldsymbol{u} \in \mathbb{R}^{(K-1)\times 1}$ and $\boldsymbol{v} \in \mathbb{R}^{(K-1)\times 1}$ can be determined by

$$\begin{aligned}
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{v}\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\ \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2\boldsymbol{I}_{K-1};\ (23b)}\ \boldsymbol{D}(\boldsymbol{d})\boldsymbol{v}\\
&= \sigma_{\mathrm{n}}^2\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^2(\boldsymbol{d})\boldsymbol{v}, \qquad (85)\\
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\mathcal{E}_{\mathbf{n}}\Big\{\ \underbrace{\boldsymbol{n}}_{=\begin{bmatrix} n_{2,1}\\ n_{3,1}\\ \vdots\\ n_{K,1}\end{bmatrix};\ (17)}\ \underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K} d_k n_{k,1}^2;\ (17),\ (27)}\ \Big\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\sum_{k=2}^{K} d_k \underbrace{\begin{bmatrix} \vdots\\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{\tilde{k},1}n_{k,1}^2\}}_{=\begin{cases}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k,1}^3\}}_{=0;\ (103)}, & \tilde{k}=k,\\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{\tilde{k},1}\}}_{=0;\ (23a)}\mathcal{E}_{\mathbf{n}}\{n_{k,1}^2\}, & \tilde{k}\neq k,\end{cases}}\\ \vdots\end{bmatrix}}_{\tilde{k}\in\{2,3,\ldots,K\}}\\
&= 0, \qquad (86)\\
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{v}\}\\
&= \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}}_{=\boldsymbol{0}^{\mathsf{T}};\ (86)}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{v}\\
&= 0, \qquad (87)
\end{aligned}$$

and

$$\begin{aligned}
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\}\\
&= \mathcal{E}_{\mathbf{n}}\Big\{\ \underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K} d_k n_{k,1}^2;\ (17),\ (27)}\ \underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K} d_k n_{k,1}^2;\ (17),\ (27)}\ \Big\}\\
&= \sum_{k_1=2}^{K}\sum_{k_2=2}^{K} d_{k_1}d_{k_2}\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_1,1}^2 n_{k_2,1}^2\}}_{=\begin{cases}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k,1}^4\}}_{=3\sigma_{\mathrm{n}}^4;\ (104)}, & k_1=k_2=k,\\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_1,1}^2\}}_{=\sigma_{\mathrm{n}}^2;\ (23b)}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_2,1}^2\}}_{=\sigma_{\mathrm{n}}^2;\ (23b)}, & k_1\neq k_2\end{cases}}\\
&= \sum_{k_1=2}^{K}\left(d_{k_1}^2 3\sigma_{\mathrm{n}}^4 + \sum_{k_2=2\neq k_1}^{K} d_{k_1}d_{k_2}\sigma_{\mathrm{n}}^4\right)\\
&= \sigma_{\mathrm{n}}^4\underbrace{\left(2\sum_{k=2}^{K} d_k^2 + \sum_{k_1=2}^{K}\sum_{k_2=2}^{K} d_{k_1}d_{k_2}\right)}_{=\alpha_1}\\
&= \sigma_{\mathrm{n}}^4\alpha_1, \qquad (88)
\end{aligned}$$

where α_1 is given by

$$\alpha_1 = 2\sum_{k=2}^{K} d_k^2 + \sum_{k_1=2}^{K}\sum_{k_2=2}^{K} d_{k_1}d_{k_2}. \qquad (89)$$

The expectations of (16) and (18) are given by

$$\begin{aligned}
\mathcal{E}_{\mathbf{n}}\{\boldsymbol{\Delta}_{\mathrm{A}}\} &= \begin{bmatrix}0 & 0 & \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\}}_{=\boldsymbol{0};\ (23a)}\end{bmatrix}\\
&= \boldsymbol{O}_{(K-1)\times 3}, \qquad (90)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{E}_{\mathbf{n}}\{\boldsymbol{\delta}_{\mathrm{b}}\} &= -\boldsymbol{D}(\boldsymbol{\delta})\ \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\}}_{=\boldsymbol{0};\ (23a)}\\
&= \boldsymbol{0}_{(K-1)\times 1}. \qquad (91)
\end{aligned}$$

Let us consider the products

$$\begin{aligned}
\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}_0 &= \begin{bmatrix}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}\end{bmatrix}^{\mathsf{T}}\boldsymbol{\Phi}\begin{bmatrix}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\\ \boldsymbol{\delta}^{\mathsf{T}}\end{bmatrix}\begin{bmatrix}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\Phi}\boldsymbol{\delta}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}\\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}\\ \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}\end{bmatrix}, \qquad (92)\\
\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathrm{A}} &= \begin{bmatrix}0 & 0 & \boldsymbol{n}\end{bmatrix}^{\mathsf{T}}\boldsymbol{\Phi}\begin{bmatrix}0 & 0 & \boldsymbol{n}\end{bmatrix}\\
&= \begin{bmatrix}\boldsymbol{0}^{\mathsf{T}}\\ \boldsymbol{0}^{\mathsf{T}}\\ \boldsymbol{n}^{\mathsf{T}}\end{bmatrix}\begin{bmatrix}0 & 0 & \boldsymbol{\Phi}\boldsymbol{n}\end{bmatrix}
\end{aligned}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n} \end{bmatrix}, \tag{93}$$

$$\begin{aligned} \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathrm{A}} &= \begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta} \end{bmatrix}^{\mathsf{T}} \boldsymbol{\Phi} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}} \\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}} \\ \boldsymbol{\delta}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\Phi}\boldsymbol{n} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n} \\ 0 & 0 & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n} \\ 0 & 0 & \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n} \end{bmatrix}, \end{aligned} \tag{94}$$

$$\begin{aligned} \boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}_0 &= \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} \end{bmatrix}^{\mathsf{T}} \boldsymbol{\Phi} \begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{n}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\Phi}\boldsymbol{\delta} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta} \end{bmatrix}, \end{aligned} \tag{95}$$

$$\begin{aligned} \boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{b}} &= \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} \end{bmatrix}^{\mathsf{T}} \boldsymbol{\Phi}\big(-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \\ &= -\begin{bmatrix} \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{n}^{\mathsf{T}} \end{bmatrix} \boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n} \\ &= -\begin{bmatrix} 0 \\ 0 \\ \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n} \end{bmatrix}. \end{aligned} \tag{96}$$

Let us consider the expectations

$$\begin{aligned} \mathcal{E}_{\mathbf{n}}\{\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{b}}\} &= \mathcal{E}_{\mathbf{n}}\{\underbrace{\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} \end{bmatrix}^{\mathsf{T}} \boldsymbol{\Phi}\big(-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)}_{(96)}\} \\ &= -\begin{bmatrix} 0 \\ 0 \\ \mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\} \end{bmatrix} \\ &= -\begin{bmatrix} 0 \\ 0 \\ \mathrm{tr}(\boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta}) \quad \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2\boldsymbol{I}_{K-1};\ (23b)} \quad) \end{bmatrix} \\ &= -\sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 \\ 0 \\ \mathrm{tr}(\boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta})) \end{bmatrix}, \end{aligned} \tag{97}$$

and

$$\begin{aligned} \mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{b}\} &= \mathcal{E}_{\mathbf{n}}\{ \underbrace{\boldsymbol{A}^{\mathsf{T}}}_{=\boldsymbol{A}_0^{\mathsf{T}}+\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}};\ (13)} \quad \boldsymbol{\Phi} \quad \underbrace{\boldsymbol{b}}_{\boldsymbol{b}_0+\boldsymbol{\delta}_{\mathrm{b}};\ (15)} \} \\ &= \mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{b}_0 + \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{b}} + \boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{b}_0 + \boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{b}}\} \\ &= \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi} \underbrace{\boldsymbol{b}_0}_{=\boldsymbol{A}_0\boldsymbol{\theta}_0;\ (5)} + \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi} \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{\delta}_{\mathrm{b}}\}}_{=\boldsymbol{0};\ (91)} \\ &\quad + \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\}}_{=\boldsymbol{O};\ (90)}\boldsymbol{\Phi}\boldsymbol{b}_0 + \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{b}}\}}_{(97)} \\ &= \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}_0\boldsymbol{\theta}_0 - \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 \\ 0 \\ \mathrm{tr}(\boldsymbol{\Phi}\boldsymbol{D}(\boldsymbol{\delta})) \end{bmatrix}. \end{aligned} \tag{98}$$

From

$$\begin{aligned} \boldsymbol{A}^{\mathsf{T}}\boldsymbol{b} &= (\boldsymbol{A}_0^{\mathsf{T}} + \underbrace{\boldsymbol{\Delta}_{\mathrm{A}}}_{=\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n} \end{bmatrix};\ (16)}^{\mathsf{T}})(\boldsymbol{b}_0 + \underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{\approx -\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n};\ (18)}) \\ &\approx \boldsymbol{A}_0^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) + \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \end{bmatrix}, \end{aligned} \tag{99}$$

the product $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{T}}\boldsymbol{A}$ can be written as

$$\begin{aligned} \boldsymbol{A}^{\mathsf{T}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{T}}\boldsymbol{A} &= \boldsymbol{A}_0^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{A}_0 \\ &\quad + \boldsymbol{A}_0^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \end{bmatrix}\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{A}_0 \\ &\quad + \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \end{bmatrix}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big) \end{bmatrix}. \end{aligned} \tag{100}$$

From (98), the expectation of (99) can be found from

$$\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{b}\} = \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{A}_0\boldsymbol{\theta}_0 - \sigma_{\mathrm{n}}^2\,\mathrm{tr}(\boldsymbol{D}(\boldsymbol{\delta}))\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{101}$$

Further multiplication of (99) gives

$$\begin{aligned} \boldsymbol{A}^{\mathsf{T}}\boldsymbol{b}\boldsymbol{b}^{\mathsf{T}}\boldsymbol{A} &= \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}\boldsymbol{A}_0 - \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{A}_0 \\ &\quad - \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}\boldsymbol{A}_0 + \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{A}_0 \\ &\quad + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_0^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{n} \end{bmatrix} \\ &\quad + \begin{bmatrix} \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{0}^{\mathsf{T}} \\ \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{A}_0^{\mathsf{T}} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{n}^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{n} \end{bmatrix}. \end{aligned} \tag{102}$$

Basic calculus reveals that for $\mu_{\mathrm{n}} = 0$, we have

$$\int_{-\infty}^{\infty} n^3 \left(\frac{1}{\sigma_{\mathrm{n}}\sqrt{2\pi}}\right) \mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{n}}^2}(n-\mu_{\mathrm{n}})^2}\,\mathrm{d}n = 0, \tag{103}$$

and

$$\int_{-\infty}^{\infty} n^4 \left(\frac{1}{\sigma_{\mathrm{n}}\sqrt{2\pi}}\right) \mathrm{e}^{-\frac{1}{2\sigma_{\mathrm{n}}^2}(n-\mu_{\mathrm{n}})^2}\,\mathrm{d}n = 3\sigma_{\mathrm{n}}^4. \tag{104}$$

The expectation of each term in (102) becomes

$$\begin{aligned} &\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}_0^{\mathsf{T}}\big(\boldsymbol{b}_0 - \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\big)\big(\boldsymbol{b}_0^{\mathsf{T}} - \boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\big)\boldsymbol{A}_0\} \\ &= \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}\boldsymbol{A}_0 - \boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{b}_0 \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\}}_{=\boldsymbol{0}^{\mathsf{T}};\ (23a)} \boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{A}_0 \end{aligned}$$

$$
\begin{aligned}
&- A_0^{\mathsf{T}} D(\delta) \underbrace{\mathcal{E}_{\mathrm{n}}\{n\}}_{=0;\ (23a)} b_0^{\mathsf{T}} A_0 + A_0^{\mathsf{T}} D(\delta) \underbrace{\mathcal{E}_{\mathrm{n}}\{nn^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2 I_{K-1};\ (23b)} D(\delta) A_0 \\
&= A_0^{\mathsf{T}} b_0 b_0^{\mathsf{T}} A_0 + \sigma_{\mathrm{n}}^2 A_0^{\mathsf{T}} D(\delta^2) A_0, \qquad (105)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{A_0^{\mathsf{T}} (b_0 - D(\delta) n)(b_0^{\mathsf{T}} - n^{\mathsf{T}} D(\delta)) n\} \\
&= A_0^{\mathsf{T}} b_0 b_0^{\mathsf{T}} \underbrace{\mathcal{E}_{\mathrm{n}}\{n\}}_{=0;\ (23a)} - A_0^{\mathsf{T}} b_0 \mathcal{E}_{\mathrm{n}}\{ \underbrace{n^{\mathsf{T}} D(\delta) n}_{=\mathrm{tr}(D(\delta) nn^{\mathsf{T}})} \} \\
&\quad - A_0^{\mathsf{T}} D(\delta) \mathcal{E}_{\mathrm{n}}\{ \underbrace{n b_0^{\mathsf{T}} n}_{=nn^{\mathsf{T}} b_0} \} + A_0^{\mathsf{T}} D(\delta) \mathcal{E}_{\mathrm{n}}\{ nn^{\mathsf{T}} \underbrace{D(\delta) n}_{=D(n)\delta} \} \\
&= -A_0^{\mathsf{T}} b_0 \,\mathrm{tr}\,(D(\delta) \underbrace{\mathcal{E}_{\mathrm{n}}\{nn^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2 I_{K-1};\ (23b)}) - A_0^{\mathsf{T}} D(\delta) \underbrace{\mathcal{E}_{\mathrm{n}}\{nn^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2 I_{K-1};\ (23b)} b_0 \\
&\quad + A_0^{\mathsf{T}} D(\delta)\, \underbrace{\mathcal{E}_{\mathrm{n}}\{ \underbrace{nn^{\mathsf{T}} D(n)}_{=[nn_2^2 \ nn_3^2 \ \cdots \ nn_K^2]} \}}_{=O;\ (23a),\ (103)} \delta \\
&= -\sigma_{\mathrm{n}}^2 \,\mathrm{tr}\,(D(\delta)) A_0^{\mathsf{T}} b_0 - \sigma_{\mathrm{n}}^2 A_0^{\mathsf{T}} D(\delta) b_0, \qquad (106)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}} (b_0 - D(\delta) n)(b_0^{\mathsf{T}} - n^{\mathsf{T}} D(\delta)) A_0\} \\
&= \Big(\underbrace{\mathcal{E}_{\mathrm{n}}\{A_0^{\mathsf{T}} (b_0 - D(\delta) n)(b_0^{\mathsf{T}} - n^{\mathsf{T}} D(\delta)) n\}}_{=-\sigma_{\mathrm{n}}^2 \mathrm{tr}(D(\delta)) A_0^{\mathsf{T}} b_0 - \sigma_{\mathrm{n}}^2 A_0^{\mathsf{T}} D(\delta) b_0;\ (106)} \Big)^{\mathsf{T}} \\
&= -\sigma_{\mathrm{n}}^2 \,\mathrm{tr}\,(D(\delta)) b_0^{\mathsf{T}} A_0 - \sigma_{\mathrm{n}}^2 b_0^{\mathsf{T}} D(\delta) A_0, \qquad (107)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}} (b_0 - D(\delta) n)(b_0^{\mathsf{T}} - n^{\mathsf{T}} D(\delta)) n\} \\
&= \mathcal{E}_{\mathrm{n}}\{ \underbrace{n^{\mathsf{T}} b_0 b_0^{\mathsf{T}} n}_{=\mathrm{tr}(b_0 b_0^{\mathsf{T}} nn^{\mathsf{T}})} \} - \underbrace{\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}} b_0 n^{\mathsf{T}} D(\delta) n\}}_{=0;\ (86)} \\
&\quad - \underbrace{\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}} D(\delta) n b_0^{\mathsf{T}} n\}}_{=0;\ (87)} + \underbrace{\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}} D(\delta) nn^{\mathsf{T}} D(\delta) n\}}_{=\sigma_{\mathrm{n}}^4 \alpha_1;\ (88)} \\
&= \mathrm{tr}\,(b_0 b_0^{\mathsf{T}} \underbrace{\mathcal{E}_{\mathrm{n}}\{nn^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2 I_{K-1}}) + \sigma_{\mathrm{n}}^4 \alpha_1 \\
&= \sigma_{\mathrm{n}}^2 \|b_0\|_2^2 + \sigma_{\mathrm{n}}^4 \alpha_1. \qquad (108)
\end{aligned}
$$

According to (102), it follows from (105), (106), (107), and (108) that

$$
\begin{aligned}
\mathcal{E}_{\mathrm{n}}\{A^{\mathsf{T}} bb^{\mathsf{T}} A\} &= A_0^{\mathsf{T}} b_0 b_0^{\mathsf{T}} A_0 + \sigma_{\mathrm{n}}^2 A_0^{\mathsf{T}} D(\delta^2) A_0 \\
&\quad - \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 & 0 & \mathrm{tr}\,(D(\delta)) A_0^{\mathsf{T}} b_0 + A_0^{\mathsf{T}} D(\delta) b_0 \end{bmatrix} \\
&\quad - \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0^{\mathsf{T}} \\ 0^{\mathsf{T}} \\ \mathrm{tr}\,(D(\delta)) b_0^{\mathsf{T}} A_0 + b_0^{\mathsf{T}} D(\delta) A_0 \end{bmatrix} \\
&\quad + (\sigma_{\mathrm{n}}^2 \|b_0\|_2^2 + \sigma_{\mathrm{n}}^4 \alpha_1) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (109)
\end{aligned}
$$

Let us consider the expectations

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{(\delta + n)^{\mathsf{T}} \Phi \delta_{\mathrm{x}}\} \\
&= (\delta + \underbrace{\mathcal{E}_{\mathrm{n}}\{n\}}_{=0;\ (23a)})^{\mathsf{T}} \Phi \delta_{\mathrm{x}} \\
&= \delta^{\mathsf{T}} \Phi \delta_{\mathrm{x}}, \qquad (110)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{\delta_{\mathrm{x}}^{\mathsf{T}} \Phi (\delta + n)\} \\
&= \mathcal{E}_{\mathrm{n}}\{\delta_{\mathrm{x}}^{\mathsf{T}} \Phi (\delta + \underbrace{\mathcal{E}_{\mathrm{n}}\{n\}}_{=0;\ (23a)})\} \\
&= \mathcal{E}_{\mathrm{n}}\{\delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta\}, \qquad (111)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{(\delta + n)^{\mathsf{T}} \Phi (\delta + n)\} \\
&= \mathcal{E}_{\mathrm{n}}\{ \mathrm{tr}\,(\Phi (\delta + n)(\delta + n)^{\mathsf{T}}) \} \\
&= \mathrm{tr}\,(\Phi (\delta\delta^{\mathsf{T}} + \underbrace{\mathcal{E}_{\mathrm{n}}\{n\}}_{=0;\ (23a)} \delta^{\mathsf{T}} + \delta \underbrace{\mathcal{E}_{\mathrm{n}}\{n^{\mathsf{T}}\}}_{=0^{\mathsf{T}};\ (23a)} + \underbrace{\mathcal{E}_{\mathrm{n}}\{nn^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2 I_{K-1};\ (23b)})) \\
&= \mathrm{tr}\,(\Phi (\delta\delta^{\mathsf{T}} + \sigma_{\mathrm{n}}^2 I_{K-1})) \\
&= \delta^{\mathsf{T}} \Phi \delta + \sigma_{\mathrm{n}}^2 \,\mathrm{tr}(\Phi), \qquad (112)
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{E}_{\mathrm{n}}\{A^{\mathsf{T}} \Phi A\} \\
&= \begin{bmatrix}
\delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta_{\mathrm{x}} & \delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta_{\mathrm{y}} & \underbrace{\mathcal{E}_{\mathrm{n}}\{\delta_{\mathrm{x}}^{\mathsf{T}} \Phi (\delta + n)\}}_{=\delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta;\ (111)} \\
\delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta_{\mathrm{x}} & \delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta_{\mathrm{y}} & \underbrace{\mathcal{E}_{\mathrm{n}}\{\delta_{\mathrm{y}}^{\mathsf{T}} \Phi (\delta + n)\}}_{=\delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta;\ (111)} \\
\underbrace{\mathcal{E}_{\mathrm{n}}\{(\delta + n)^{\mathsf{T}} \Phi \delta_{\mathrm{x}}\}}_{=\delta^{\mathsf{T}} \Phi \delta_{\mathrm{x}};\ (110)} & \underbrace{\mathcal{E}_{\mathrm{n}}\{(\delta + n)^{\mathsf{T}} \Phi \delta_{\mathrm{y}}\}}_{=\delta^{\mathsf{T}} \Phi \delta_{\mathrm{y}};\ (110)} & \underbrace{\mathcal{E}_{\mathrm{n}}\{(\delta + n)^{\mathsf{T}} \Phi (\delta + n)\}}_{=\delta^{\mathsf{T}} \Phi \delta + \sigma_{\mathrm{n}}^2 \mathrm{tr}(\Phi);\ (112)}
\end{bmatrix} \\
&= \begin{bmatrix}
\delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta_{\mathrm{x}} & \delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta_{\mathrm{y}} & \delta_{\mathrm{x}}^{\mathsf{T}} \Phi \delta \\
\delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta_{\mathrm{x}} & \delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta_{\mathrm{y}} & \delta_{\mathrm{y}}^{\mathsf{T}} \Phi \delta \\
\delta^{\mathsf{T}} \Phi \delta_{\mathrm{x}} & \delta^{\mathsf{T}} \Phi \delta_{\mathrm{y}} & \delta^{\mathsf{T}} \Phi \delta + \sigma_{\mathrm{n}}^2 \,\mathrm{tr}(\Phi)
\end{bmatrix} \\
&= A_0^{\mathsf{T}} \Phi A_0 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{n}}^2 \,\mathrm{tr}(\Phi) \end{bmatrix}. \qquad (113)
\end{aligned}
$$

Substituting $\Phi = I_{K-1}$ into (113), we obtain

$$
\begin{aligned}
\mathcal{E}_{\mathrm{n}}\{A^{\mathsf{T}} A\} &= A_0^{\mathsf{T}} I_{K-1} A_0 + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_{\mathrm{n}}^2 \,\mathrm{tr}(I_{K-1}) \end{bmatrix} \\
&= A_0^{\mathsf{T}} A_0 + (K-1) \sigma_{\mathrm{n}}^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad (114)
\end{aligned}
$$

APPENDIX B
PROOF OF LEMMA 1

First, we consider

$$
\begin{aligned}
\hat{n}_{\ell_2} &= \arg\min_{n} \|n\|_2^2 \quad \text{s.t.} \quad A\theta - b = D(d) n \\
&= \arg_{n} \arg_{\lambda} \Big(\frac{\partial}{\partial n} \big(\|n\|_2^2 + \lambda^{\mathsf{T}} (A\theta - b - D(d) n) \big) = 0 \,\wedge \\
&\qquad \frac{\partial}{\partial \lambda} \big(\|n\|_2^2 + \lambda^{\mathsf{T}} (A\theta - b - D(d) n) \big) = 0 \Big) \\
&= \arg_{n} \arg_{\lambda} \,(2n - D(d)\lambda = 0 \ \wedge\ A\theta - b - D(d) n = 0) \\
&= \arg_{n} \arg_{\lambda} \Big(n = \frac{1}{2} D(d)\lambda \ \wedge\ A\theta - b = D(d) \underbrace{n}_{=\frac{1}{2} D(d)\lambda} \Big). \qquad (115)
\end{aligned}
$$

From (25), the solution of $\boldsymbol{n}$ in the minimal ℓ_2-norm sense becomes

$$\begin{aligned}\hat{\boldsymbol{n}}_{\ell_2} &= \arg\min_{\boldsymbol{n}}\Big(\boldsymbol{\lambda} = 2(\boldsymbol{D}(\boldsymbol{d})\boldsymbol{D}(\boldsymbol{d}))^{-1}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \wedge \\ &\qquad \boldsymbol{n} = \frac{1}{2}\boldsymbol{D}(\boldsymbol{d}) \underbrace{\boldsymbol{\lambda}}_{=2(\boldsymbol{D}(\boldsymbol{d})\boldsymbol{D}(\boldsymbol{d}))^{-1}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})}\Big) \\ &= \boldsymbol{D}(\boldsymbol{d})\big(\boldsymbol{D}(\boldsymbol{d})\boldsymbol{D}(\boldsymbol{d})\big)^{-1}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &= \boldsymbol{D}^{-1}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}). \end{aligned} \tag{116}$$

Substituting (116) into $\|\boldsymbol{n}\|_2^2$, we obtain

$$\begin{aligned}\|\hat{\boldsymbol{n}}_{\ell_2}\|_2^2 &= \Big(\boldsymbol{D}^{-1}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})\Big)^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &= (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{D}^{-1}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &= (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}). \end{aligned} \tag{117}$$

Substituting (116) and (117) into (29), we obtain

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\ell_2} &= \arg\min_{\boldsymbol{\theta}}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &\quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b} = \boldsymbol{D}(\boldsymbol{d})\boldsymbol{D}^{-1}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &= \arg\min_{\boldsymbol{\theta}}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &\quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b} = \boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b} \\ &= \arg\min_{\boldsymbol{\theta}}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}). \end{aligned} \tag{118}$$

By using few straightforward differentiations, we find that

$$\begin{aligned}\frac{\partial}{\partial\theta_1}\boldsymbol{D}(\boldsymbol{d}) &= \frac{\partial}{\partial \underbrace{\theta_1}_{=x-x_1;\ (7)}}\boldsymbol{D}(\boldsymbol{d}) \\ &= \boldsymbol{O}, \end{aligned} \tag{119}$$

$$\begin{aligned}\frac{\partial}{\partial\theta_2}\boldsymbol{D}(\boldsymbol{d}) &= \frac{\partial}{\partial \underbrace{\theta_1}_{=y-y_1;\ (7)}}\boldsymbol{D}(\boldsymbol{d}) \\ &= \boldsymbol{O}, \end{aligned} \tag{120}$$

$$\begin{aligned}\frac{\partial}{\partial\theta_3}\boldsymbol{D}(\boldsymbol{d}) &= \frac{\partial}{\partial \underbrace{\theta_1}_{=d_1;\ (7)}}\boldsymbol{D}(\boldsymbol{d}) \\ &= \boldsymbol{O}, \end{aligned} \tag{121}$$

$$\begin{aligned}\frac{\partial}{\partial\theta_n}\boldsymbol{D}^{-1}(\boldsymbol{d}) &= -\boldsymbol{D}^{-1}(\boldsymbol{d})\Big(\underbrace{\frac{\partial}{\partial\theta_n}\boldsymbol{D}(\boldsymbol{d})}_{=\boldsymbol{O}}\Big)\boldsymbol{D}^{-1}(\boldsymbol{d}) \\ &= \boldsymbol{O}; \quad n\in\{1,2,3\}, \end{aligned} \tag{122}$$

$$\begin{aligned}\frac{\partial}{\partial\theta_n}\boldsymbol{D}^{-2}(\boldsymbol{d}) &= \Big(\underbrace{\frac{\partial}{\partial\theta_n}\boldsymbol{D}^{-1}(\boldsymbol{d})}_{=\boldsymbol{O};\ (122)}\Big)\boldsymbol{D}^{-1}(\boldsymbol{d}) \\ &\quad + \boldsymbol{D}^{-1}(\boldsymbol{d})\Big(\underbrace{\frac{\partial}{\partial\theta_n}\boldsymbol{D}^{-1}(\boldsymbol{d})}_{=\boldsymbol{O};\ (122)}\Big) \\ &= \boldsymbol{O}; \quad n\in\{1,2,3\}. \end{aligned} \tag{123}$$

The gradient of the CTLS can be derived from

$$\begin{aligned}\boldsymbol{g}_\theta(\boldsymbol{\theta}) &= \frac{\partial}{\partial\boldsymbol{\theta}}f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \frac{\partial}{\partial\boldsymbol{\theta}}\Big((\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})\Big) \\ &= 2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ &\quad + \begin{bmatrix} (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\Big(\underbrace{\frac{\partial}{\partial\theta_1}\boldsymbol{D}^{-2}(\boldsymbol{d})}_{=\boldsymbol{O};\ (123)}\Big)(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\Big(\underbrace{\frac{\partial}{\partial\theta_2}\boldsymbol{D}^{-2}(\boldsymbol{d})}_{=\boldsymbol{O};\ (123)}\Big)(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \\ (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\Big(\underbrace{\frac{\partial}{\partial\theta_3}\boldsymbol{D}^{-2}(\boldsymbol{d})}_{=\boldsymbol{O};\ (123)}\Big)(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) \end{bmatrix} \\ &= 2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}). \end{aligned} \tag{124}$$

Solving (124) for $\boldsymbol{\theta}$ leads to (33). In other words, the problem in (30) can be solved as

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\ell_2} &= \arg\min_{\boldsymbol{\theta}} f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \arg_{\boldsymbol{\theta}}\big(\boldsymbol{g}_\theta(\boldsymbol{\theta}) = \boldsymbol{0}\big) \\ &= \arg_{\boldsymbol{\theta}}\Big(2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{\delta})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}) = \boldsymbol{0}\Big) \\ &= \arg_{\boldsymbol{\theta}}\Big(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\boldsymbol{\theta} - \boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{b} = \boldsymbol{0}\Big) \\ &= \big(\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\big)^{-1}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{b}. \end{aligned} \tag{125}$$

APPENDIX C
PROOF OF LEMMA 2

It is simple to show that

$$\begin{aligned}\boldsymbol{H}_{\theta\theta}(\boldsymbol{\theta}) &= \frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^{\mathsf{T}}}f_{\text{CTLS}}(\boldsymbol{\theta}) \\ &= \frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{g}_\theta^{\mathsf{T}}(\boldsymbol{\theta}) \\ &= \frac{\partial}{\partial\boldsymbol{\theta}}2(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A} \\ &= 2\Big(\underbrace{\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{\theta}^{\mathsf{T}}}_{=\boldsymbol{I}}\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A} - \underbrace{\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{b}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}}_{=\boldsymbol{O}}\Big) \\ &= 2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}. \end{aligned} \tag{126}$$

The expectation of the Hessian in (34) can be derived from

$$\begin{aligned}\bar{\boldsymbol{H}}_{\theta\theta}(\boldsymbol{\theta}) &= \mathcal{E}_{\mathbf{n}}\big\{2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\big\} \\ &= 2\underbrace{\mathcal{E}_{\mathbf{n}}\big\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\big\}}_{(113)}. \end{aligned} \tag{127}$$

APPENDIX D
PROOF OF LEMMA 3

Let us consider the outer product $(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}$ from

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}} &= (\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}-\boldsymbol{b}^{\mathsf{T}}) \\ &= \boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}} - \boldsymbol{b}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}} - \boldsymbol{A}\boldsymbol{\theta}\boldsymbol{b}^{\mathsf{T}} + \boldsymbol{b}\boldsymbol{b}^{\mathsf{T}}. \end{aligned} \tag{128}$$

Substituting (18), (13), and (15) into each term of (128), we obtain

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}} &= (\boldsymbol{A}_0+\boldsymbol{\Delta}_{\mathrm{A}})\boldsymbol{\theta}\boldsymbol{\theta}^{\mathsf{T}}(\boldsymbol{A}_0+\boldsymbol{\Delta}_{\mathrm{A}})^{\mathsf{T}}\\
&= (\boldsymbol{A}_0\boldsymbol{\theta}\boldsymbol{\theta}^{\mathsf{T}}+\boldsymbol{\Delta}_{\mathrm{A}}\boldsymbol{\theta}\boldsymbol{\theta}^{\mathsf{T}})(\boldsymbol{A}_0^{\mathsf{T}}+\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}})\\
&= \underbrace{\boldsymbol{A}_0\boldsymbol{\theta}}_{=\boldsymbol{b}_0}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}_0^{\mathsf{T}}}_{=\boldsymbol{b}_0^{\mathsf{T}}}+\underbrace{\boldsymbol{\Delta}_{\mathrm{A}}\boldsymbol{\theta}}_{=d_1\boldsymbol{n}}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}_0^{\mathsf{T}}}_{=\boldsymbol{b}_0^{\mathsf{T}}}\\
&\quad+\underbrace{\boldsymbol{A}_0\boldsymbol{\theta}}_{=\boldsymbol{b}_0}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}}_{=d_1\boldsymbol{n}^{\mathsf{T}}}+\underbrace{\boldsymbol{\Delta}_{\mathrm{A}}\boldsymbol{\theta}}_{=d_1\boldsymbol{n}}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}}_{=d_1\boldsymbol{n}^{\mathsf{T}}}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}+d_1^2\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}},
\end{aligned}\tag{129}$$

$$\begin{aligned}
\boldsymbol{b}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}} &= (\boldsymbol{b}_0+\boldsymbol{\delta}_{\mathrm{b}})\boldsymbol{\theta}^{\mathsf{T}}(\boldsymbol{A}_0+\boldsymbol{\Delta}_{\mathrm{A}})^{\mathsf{T}}\\
&= (\boldsymbol{b}_0\boldsymbol{\theta}^{\mathsf{T}}+\boldsymbol{\delta}_{\mathrm{b}}\boldsymbol{\theta}^{\mathsf{T}})(\boldsymbol{A}_0^{\mathsf{T}}+\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}})\\
&= \boldsymbol{b}_0\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}_0^{\mathsf{T}}}_{=\boldsymbol{b}_0^{\mathsf{T}}}+\underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{=-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}_0^{\mathsf{T}}}_{=\boldsymbol{b}_0^{\mathsf{T}}}\\
&\quad+\boldsymbol{b}_0\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}}_{=d_1\boldsymbol{n}^{\mathsf{T}}}+\underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{=-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}}\underbrace{\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}}_{=d_1\boldsymbol{n}^{\mathsf{T}}}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}-d_1\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}},
\end{aligned}\tag{130}$$

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{\theta}\boldsymbol{b}^{\mathsf{T}} &= (\underbrace{\boldsymbol{b}\boldsymbol{\theta}^{\mathsf{T}}\boldsymbol{A}^{\mathsf{T}}}_{(130)})^{\mathsf{T}}\\
&= \left(\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}-d_1\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\right)^{\mathsf{T}}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})+d_1\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}-d_1\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta}),
\end{aligned}\tag{131}$$

and

$$\begin{aligned}
\boldsymbol{b}\boldsymbol{b}^{\mathsf{T}} &= (\boldsymbol{b}_0+\boldsymbol{\delta}_{\mathrm{b}})(\boldsymbol{b}_0+\boldsymbol{\delta}_{\mathrm{b}})^{\mathsf{T}}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}+\underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{=-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}}\boldsymbol{b}_0^{\mathsf{T}}+\boldsymbol{b}_0\underbrace{\boldsymbol{\delta}_{\mathrm{b}}^{\mathsf{T}}}_{=-\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})}+\underbrace{\boldsymbol{\delta}_{\mathrm{b}}}_{=-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}}\underbrace{\boldsymbol{\delta}_{\mathrm{b}}^{\mathsf{T}}}_{=-\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})+\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta}).
\end{aligned}\tag{132}$$

Substituting (129), (130), (131), and (132) into (128), we obtain

$$\begin{aligned}
&(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}\\
&= \boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}+d_1^2\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\\
&\quad-(\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}+d_1\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}-d_1\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}})\\
&\quad-(\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})+d_1\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}-d_1\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta}))\\
&\quad+\boldsymbol{b}_0\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{b}_0^{\mathsf{T}}-\boldsymbol{b}_0\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})+\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\\
&= d_1^2\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}+\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})+d_1\boldsymbol{D}(\boldsymbol{\delta})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}+d_1\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{\delta})\\
&= (\underbrace{\boldsymbol{D}(\boldsymbol{\delta})+d_1\boldsymbol{I}}_{=\boldsymbol{D}(\boldsymbol{\delta})})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}(\underbrace{\boldsymbol{D}(\boldsymbol{\delta})+d_1\boldsymbol{I}}_{=\boldsymbol{D}(\boldsymbol{\delta})})\\
&= \boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d}).
\end{aligned}\tag{133}$$

The correlation of the CTLS gradient can be determined by

$$\begin{aligned}
\boldsymbol{R}_{\theta\theta} &= \mathcal{E}_{\mathbf{n}}\{\boldsymbol{g}_{\theta}(\boldsymbol{\theta})\boldsymbol{g}_{\theta}^{\mathsf{T}}(\boldsymbol{\theta})\}\\
&= \mathcal{E}_{\mathbf{n}}\{2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(2\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b}))^{\mathsf{T}}\}\\
&= 4\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\underbrace{(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})(\boldsymbol{A}\boldsymbol{\theta}-\boldsymbol{b})^{\mathsf{T}}}_{=\boldsymbol{D}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}(\boldsymbol{d});\ (133)}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\}\\
&= 4\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{A}\}.
\end{aligned}\tag{134}$$

Let $\boldsymbol{\Delta}\in\mathbb{R}^{3\times3}$ be a perturbation matrix, which is given by

$$\begin{aligned}
\boldsymbol{\Delta} &= \underbrace{\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathrm{A}}}_{(93)}+\underbrace{\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\Delta}_{\mathrm{A}}}_{(94)}+\underbrace{\boldsymbol{\Delta}_{\mathrm{A}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}_0}_{(95)}\\
&= \begin{bmatrix} 0 & 0 & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n}\\ 0 & 0 & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n}\\ \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n}+\boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{n}+\boldsymbol{n}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}\end{bmatrix}.
\end{aligned}\tag{135}$$

The noisy product $\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}$ can be written as

$$\begin{aligned}
\boldsymbol{A}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A} &= \underbrace{\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{A}_0}_{(92)}+\underbrace{\boldsymbol{\Delta}}_{(135)}\\
&= \begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{\Phi}(\boldsymbol{\delta}+\boldsymbol{n})\\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{\Phi}(\boldsymbol{\delta}+\boldsymbol{n})\\ (\boldsymbol{\delta}+\boldsymbol{n})^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{x}} & (\boldsymbol{\delta}+\boldsymbol{n})^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\delta}_{\mathrm{y}} & (\boldsymbol{\delta}+\boldsymbol{n})^{\mathsf{T}}\boldsymbol{\Phi}(\boldsymbol{\delta}+\boldsymbol{n})\end{bmatrix}.
\end{aligned}\tag{136}$$

Based on (136), the expectations of some scalars for $\boldsymbol{u}\in\mathbb{R}^{(K-1)\times1}$ and $\boldsymbol{v}\in\mathbb{R}^{(K-1)\times1}$ can be determined by

$$\begin{aligned}
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{v}\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}}_{=\sigma_{\mathrm{n}}^2\boldsymbol{I}_{K-1};\ (23b)}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{v}\\
&= \sigma_{\mathrm{n}}^2\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{v},
\end{aligned}\tag{137}$$

$$\begin{aligned}
&\mathcal{E}_{\mathbf{n}}\{\boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\mathcal{E}_{\mathbf{n}}\{\underbrace{\boldsymbol{n}}_{=\begin{bmatrix}n_{2,1}\\ n_{3,1}\\ \vdots\\ n_{K,1}\end{bmatrix};\ (17)}\underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K}\frac{1}{d_k}n_{k,1}^2;\ (17),\ (27)}\}\\
&= \boldsymbol{u}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\sum_{k=2}^{K}\frac{1}{d_k}\underbrace{\begin{bmatrix}\vdots\\ \mathcal{E}_{\mathbf{n}}\{n_{\tilde{k},1}n_{k,1}^2\}\\ \begin{cases}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k,1}^3\}}_{=0;\ (103)}, & \tilde{k}=k,\\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{\tilde{k},1}\}}_{=0;\ (23a)}\mathcal{E}_{\mathbf{n}}\{n_{k,1}^2\}, & \tilde{k}\neq k,\end{cases}\\ \vdots\end{bmatrix}}_{\tilde{k}\in\{2,3,\ldots,K\}}\\
&= 0,
\end{aligned}\tag{138}$$

$$\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{v}\}$$

$$= \underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\}}_{=\mathbf{0}^{\mathsf{T}};\ (138)}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{v}$$
$$= 0, \tag{139}$$

$$\mathcal{E}_{\mathbf{n}}\{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}\}$$
$$= \mathcal{E}_{\mathbf{n}}\{\underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K}\frac{1}{d_k}n_{k,1}^2;\ (17),\ (27)}\underbrace{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{D}^{-1}(\boldsymbol{d})\boldsymbol{n}}_{=\sum_{k=2}^{K}\frac{1}{d_k}n_{k,1}^2;\ (17),\ (27)}\}$$
$$= \sum_{k_1=2}^{K}\sum_{k_2=2}^{K}\frac{1}{d_{k_1}d_{k_2}}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_1,1}^2 n_{k_2,1}^2\}}$$
$$= \begin{cases} \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k,1}^4\}}_{=3\sigma_{\mathrm{n}}^4;\ (104)}, & k_1 = k_2 = k, \\ \underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_1,1}^2\}}_{=\sigma_{\mathrm{n}}^2;\ (23b)}\underbrace{\mathcal{E}_{\mathbf{n}}\{n_{k_2,1}^2\}}_{=\sigma_{\mathrm{n}}^2;\ (23b)}, & k_1 \neq k_2 \end{cases}$$
$$= \sum_{k_1=2}^{K}\left(\frac{1}{d_{k_1}^2}3\sigma_{\mathrm{n}}^4 + \sum_{k_2=2\neq k_1}^{K}\frac{1}{d_{k_1}d_{k_2}}\sigma_{\mathrm{n}}^4\right)$$
$$= \sigma_{\mathrm{n}}^4\underbrace{\left(2\sum_{k=2}^{K}\frac{1}{d_k^2} + \sum_{k_1=2}^{K}\sum_{k_2=2}^{K}\frac{1}{d_{k_1}d_{k_2}}\right)}_{\alpha_3}$$
$$= \sigma_{\mathrm{n}}^4\alpha_3. \tag{140}$$

From (136) and by using the results in (137), (138), (139), and (140), the matrix $\boldsymbol{R}_{\theta\theta}$ in (134) can be shown as

$$\boldsymbol{R}_{\theta\theta} = 4\sigma_{\mathrm{n}}^2\begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta} \\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta} \\ \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{\delta} + \alpha_2\sigma_{\mathrm{n}}^2 \end{bmatrix}$$
$$= 4\sigma_{\mathrm{n}}^2\left(\begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d}) \\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d}) \\ \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d}) \end{bmatrix}\underbrace{\begin{bmatrix}\boldsymbol{\delta}_{\mathrm{x}} & \boldsymbol{\delta}_{\mathrm{y}} & \boldsymbol{\delta}\end{bmatrix}}_{=\boldsymbol{A}_0} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3\sigma_{\mathrm{n}}^2 \end{bmatrix}\right)$$
$$= 4\sigma_{\mathrm{n}}^2\left(\underbrace{\begin{bmatrix} \boldsymbol{\delta}_{\mathrm{x}}^{\mathsf{T}} \\ \boldsymbol{\delta}_{\mathrm{y}}^{\mathsf{T}} \\ \boldsymbol{\delta}^{\mathsf{T}} \end{bmatrix}}_{=\boldsymbol{A}_0^{\mathsf{T}}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}\right)$$
$$= 4\sigma_{\mathrm{n}}^2\left(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}\right). \tag{141}$$

The error covariance matrix of the CTLS estimate can be calculated by

$$\boldsymbol{\Sigma}_{\theta\theta} = \mathcal{E}_{\mathbf{n}}\{(\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}}_{\mathrm{CTLS}} - \boldsymbol{\theta}_0)^{\mathsf{T}}\}$$
$$\simeq \mathcal{E}_{\mathbf{n}}\{\bar{\boldsymbol{H}}_{\theta\theta}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{g}_\theta(\boldsymbol{\theta}_0)\boldsymbol{g}_\theta^{\mathsf{T}}(\boldsymbol{\theta}_0)\bar{\boldsymbol{H}}_{\theta\theta}^{-1}(\boldsymbol{\theta}_0)\}$$
$$= \bar{\boldsymbol{H}}_{\theta\theta}^{-1}(\boldsymbol{\theta}_0)\underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{g}_\theta(\boldsymbol{\theta}_0)\boldsymbol{g}_\theta^{\mathsf{T}}(\boldsymbol{\theta}_0)\}}_{=\boldsymbol{R}_{\theta\theta};\ (134)}\bar{\boldsymbol{H}}_{\theta\theta}^{-1}(\boldsymbol{\theta}_0). \tag{142}$$

Substituting (35) and (141) into (142), we obtain

$$\boldsymbol{\Sigma}_{\theta\theta} = \Big(\underbrace{\bar{\boldsymbol{H}}_{\theta\theta}(\boldsymbol{\theta}_0)}_{=2\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\};\ (35)}\Big)^{-1}$$
$$\underbrace{\boldsymbol{R}_{\theta\theta}}$$
$$= 4\sigma_{\mathrm{n}}^2\left(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}\right);\ (141)$$
$$\Big(\underbrace{\bar{\boldsymbol{H}}_{\theta\theta}(\boldsymbol{\theta}_0)}_{=2\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\};\ (35)}\Big)^{-1}$$
$$= \sigma_{\mathrm{n}}^2\Big(\underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\}}_{(113)}\Big)^{-1}$$
$$\left(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0 + \sigma_{\mathrm{n}}^2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix}\right)$$
$$\Big(\underbrace{\mathcal{E}_{\mathbf{n}}\{\boldsymbol{A}^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}\}}_{(113)}\Big)^{-1}, \tag{143}$$

which yields (37).

APPENDIX E
PROOF OF COROLLARY 1

For a small σ_{n}^2, a simplified error covariance can be found from

$$\lim_{\sigma_{\mathrm{n}}^2\to 0}\boldsymbol{\Sigma}_{\theta\theta} \approx \sigma_{\mathrm{n}}^2\big(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\big)^{-1}\big(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\big)$$
$$\times\big(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\big)^{-1}$$
$$= \sigma_{\mathrm{n}}^2\big(\boldsymbol{A}_0^{\mathsf{T}}\boldsymbol{D}^{-2}(\boldsymbol{d})\boldsymbol{A}_0\big)^{-1}. \tag{144}$$

REFERENCES

[1] S. He, K. Shi, C. Liu, B. Guo, J. Chen, and Z. Shi, "Collaborative sensing in Internet of Things: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1435–1474, 3rd Quart., 2022.
[2] R. Jia, J. Wu, X. Wang, J. Lu, F. Lin, Z. Zheng, and M. Li, "Energy cost minimization in wireless rechargeable sensor networks," *IEEE/ACM Trans. Netw.*, vol. 31, no. 5, pp. 2345–2360, Oct. 2023.
[3] W. Dargie, J. Wen, L. A. Panes-Ruiz, L. Riemenschneider, B. Ibarlucea, and G. Cuniberti, "Monitoring toxic gases using nanotechnology and wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 11, pp. 12274–12283, Jun. 2023.
[4] M. Yang, N. Liu, Y. Feng, H. Gong, X. Wang, and M. Liu, "Dynamic mobile sink path planning for unsynchronized data collection in heterogeneous wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 17, pp. 20310–20320, Sep. 2023.
[5] S. He, D.-H. Shin, J. Zhang, and J. Chen, "Near-optimal allocation algorithms for location-dependent tasks in crowdsensing," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3392–3405, Apr. 2017.
[6] S. S. Mohar, S. Goyal, and R. Kaur, "Localization of sensor nodes in wireless sensor networks using bat optimization algorithm with enhanced exploration and exploitation characteristics," *J. Supercomput.*, vol. 78, no. 9, pp. 11975–12023, Feb. 2022.
[7] M. Ma, S. Xu, and J. Jiang, "A distributed gradient descent method for node localization on large-scale wireless sensor network," *IEEE J. Miniaturization Air Space Syst.*, vol. 4, no. 2, pp. 13774–13788, Jan. 2023.
[8] M.-M. Cheng, J. Zhang, D.-G. Wang, W. Tan, and J. Yang, "A localization algorithm based on improved water flow optimizer and max-similarity path for 3-D heterogeneous wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 12, pp. 13774–13788, Jun. 2023.

[9] M. Zhu, Z. Wei, C. Qiu, W. Jiang, H. Wu, and Z. Feng, "Joint data collection and sensor positioning in multi-UAV-assisted wireless sensor network," *IEEE Sensors J.*, vol. 23, no. 19, pp. 23664–23675, Oct. 2023.

[10] M. Shu, G. Chen, Z. Zhang, and L. Xu, "Indoor geomagnetic positioning using direction-aware multiscale recurrent neural networks," *IEEE Sensors J.*, vol. 23, no. 3, pp. 3321–3333, Feb. 2023.

[11] M. Fawad, M. Khan, K. Ullah, H. Alasmary, D. Shehzad, and B. Khan, "Enhancing localization efficiency and accuracy in wireless sensor networks," *Sensors*, vol. 23, no. 5, p. 2796, Mar. 2023.

[12] C. Huang, Z. Tian, W. He, K. Liu, and Z. Li, "Spotlight: A 3-D indoor localization system in wireless sensor networks based on orientation and RSSI measurements," *IEEE Sensors J.*, vol. 23, no. 21, pp. 26662–26676, Nov. 2023.

[13] Y. Gan, X. Cong, and Y. Sun, "Refinement of TOA localization with sensor position uncertainty in closed-form," *Sensors*, vol. 20, no. 2, p. 390, Jan. 2020.

[14] P. Wu, S. Su, Z. Zuo, X. Guo, B. Sun, and X. Wen, "Time difference of arrival (TDoA) localization combining weighted least squares and firefly algorithm," *Sensors*, vol. 19, no. 11, p. 2554, Jun. 2019.

[15] Q. Li, B. Chen, and M. Yang, "Time difference of arrival passive localization sensor selection method based on Tabu search," *Sensors*, vol. 20, no. 22, p. 6547, Nov. 2020.

[16] K. Dogancay and H. Hmam, "3D TDOA emitter localization using conic approximation," *Sensors*, vol. 23, no. 14, p. 6254, Jul. 2023.

[17] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.

[18] Y. Tian, J. Lv, S. Tian, J. Zhu, and W. Lu, "Robust least-square localization based on relative angular matrix in wireless sensor networks," *Sensors*, vol. 19, no. 11, pp. 1–26, Jun. 2019.

[19] K. Lee, S. Kim, and K. You, "Iterative regression based hybrid localization for wireless sensor networks," *Sensors*, vol. 21, no. 1, p. 257, Jan. 2021.

[20] J. Wang, J. Wen, J. Hua, X. Feng, Z. Xu, and Z. Ni, "Angle residual-based NLOS suppression and localization method in wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 13, pp. 14388–14396, Jul. 2023.

[21] G. Piccinni, G. Avitabile, and G. Coviello, "A novel distance measurement technique for indoor positioning systems based on Zadoff-Chu sequences," in *Proc. 15th IEEE Int. New Circuits Syst. Conf. (NEWCAS)*, Strasbourg, France, Jun. 2017, pp. 337–340.

[22] A. Florio, G. Avitabile, and G. Coviello, "Multiple source angle of arrival estimation through phase interferometry," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 674–678, Mar. 2022.

[23] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2568–2599, 3rd Quart., 2019.

[24] S. Zekavat, R. M. Buehrer, G. D. Durgin, L. Lovisolo, Z. Wang, S. T. Goh, and A. Ghasemi, "An overview on position location: Past, present, future," *Int. J. Wireless Inf. Netw.*, vol. 28, no. 1, pp. 45–76, Mar. 2021.

[25] J. Ma, J. Xian, H. Wu, Y. Yang, X. Mei, Y. Zhang, X. Chen, and C. Zhou, "Novel high-precision and high-robustness localization algorithm for underwater-environment-monitoring wireless sensor networks," *J. Mar. Sci. Eng.*, vol. 11, no. 9, pp. 1–28, Aug. 2023.

[26] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[27] *Comparison of Time-Difference-of-Arrival and Angle-of-Arrival Methods of Signal Geolocation*, document Report ITU-R SM.2211-2, International Telecommunication Union, Radiocommunication Sector of ITU, Jun. 2018. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-SM.2211-2-2018-PDF-E.pdf

[28] R. Ahmad, B. Rinner, R. Wazirali, S. K. M. Abujayyab, and R. Almajalid, "Two-level sensor self-calibration based on interpolation and autoregression for low-cost wireless sensor networks," *IEEE Sensors J.*, vol. 23, no. 20, pp. 25242–25253, Oct. 2023.

[29] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[30] W. Foy, "Position-location solutions by Taylor-series estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-12, no. 2, pp. 187–194, Mar. 1976.

[31] J. Smith and J. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, no. 12, pp. 1661–1669, Dec. 1987.

[32] A. W. H. Khong and M. Brookes, "The effect of calibration errors on source localization with microphone arrays," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2007, pp. 1–140.

[33] K. C. Ho, X. Lu, and L. Kovavisaruch, "Source localization using TDOA and FDOA measurements in the presence of receiver location errors: Analysis and solution," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 684–696, Feb. 2007.

[34] L. Yang and K. C. Ho, "An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4598–4615, Dec. 2009.

[35] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "A constrained least squares approach to mobile positioning: Algorithms and optimality," *EURASIP J. Adv. Signal Process.*, vol. 2006, no. 1, pp. 1–23, Dec. 2006.

[36] L. Lin, H. C. So, F. K. W. Chan, Y. T. Chan, and K. C. Ho, "A new constrained weighted least squares algorithm for TDOA-based localization," *Signal Process.*, vol. 93, no. 11, pp. 2872–2878, Nov. 2013.

[37] X. Qu and L. Xie, "An efficient convex constrained weighted least squares source localization algorithm based on TDOA measurements," *Signal Process.*, vol. 119, pp. 142–152, Feb. 2016.

[38] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada, "The constrained total least squares technique and its applications to harmonic superresolution," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1070–1087, May 1991.

[39] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numer. Anal.*, vol. 17, no. 6, pp. 883–893, Oct. 1980.

[40] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis* (Frontiers in Applied Mathematics), E. Greenspan, Ed. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1991.

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

[42] K. S. Arun, "A unitarily constrained total least squares problem in signal processing," *SIAM J. Matrix Anal. Appl.*, vol. 13, no. 3, pp. 729–745, Jul. 1992.

[43] P. Lemmerling, B. De Moor, and S. Van Huffel, "On the equivalence of constrained total least squares and structured total least squares," *IEEE Trans. Signal Process.*, vol. 44, no. 11, pp. 2908–2911, Nov. 1996.

[44] B. De Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Trans. Signal Process.*, vol. 42, no. 11, pp. 3104–3113, Nov. 1994.

[45] B. Schaffrin, "A note on constrained total least-squares estimation," *SIAM J. Matrix Anal. Appl.*, vol. 417, no. 1, pp. 245–258, Aug. 2006.

[46] D. Wang, L. Zhang, and Y. Wu, "Constrained total least squares algorithm for passive location based on bearing-only measurements," *Sci. China Ser. F, Inf. Sci.*, vol. 50, no. 4, pp. 576–586, Aug. 2007.

[47] K. Yang, J. An, X. Bu, and G. Sun, "Constrained total least-squares location algorithm using time-difference-of-arrival measurements," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1558–1562, Mar. 2010.

[48] L. Wang, A. Hu, G. Bai, and Y. Huang, "Enhanced constrained total least squares estimator in TDOA localization for WSNs," in *Proc. Int. Conf. Comput. Sci. Service Syst. (CSSS)*, Nanjing, China, Jun. 2011, pp. 389–392.

[49] D. Wang, R. Liu, J. Yin, Z. Wu, Y. Wang, and C. Wang, "A constrained-total-least-squares method for joint estimation of source and sensor locations: A general framework," *Math. Problems Eng.*, vol. 2018, pp. 1–23, Aug. 2018.

[50] R. Liu, D. Wang, J. Yin, and Y. Wu, "Constrained total least squares localization using angle of arrival and time difference of arrival measurements in the presence of synchronization clock bias and sensor position errors," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 7, Jul. 2019, Art. no. 155014771985859.

[51] Q. Li, B. Chen, and M. Yang, "Improved two-step constrained total least-squares TDOA localization algorithm based on the alternating direction method of multipliers," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13666–13673, Nov. 2020.

[52] Z. Zhou, Y. Rui, X. Cai, and J. Lu, "Constrained total least squares method using TDOA measurements for jointly estimating acoustic emission source and wave velocity," *Measurement*, vol. 182, Sep. 2021, Art. no. 109758.

[53] Q. Mao and Y. Li, "Robust localization based on constrained total least squares in wireless sensor networks," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–7, Mar. 2022.

[54] J. Yang, T. Yan, and W. Sun, "Polynomial fitting and interpolation method in TDOA estimation of sensors network," *IEEE Sensors J.*, vol. 23, no. 4, pp. 3837–3847, Feb. 2023.
[55] Z. Hua, Y. Lu, G. Pan, K. Gao, D. B. da Costa, and S. Chen, "Computer vision aided mmWave UAV communication systems," *IEEE Internet Things J.*, vol. 10, no. 14, pp. 12548–12561, Jul. 2023.
[56] L. Lv, S. Wu, H. Li, C. Jiang, and Y. Su, "DNN-aided distributed direct position determination for UAV cluster," *IEEE Trans. Veh. Technol.*, vol. 72, no. 12, pp. 16839–16844, Dec. 2023.
[57] X. Chen, D. Wang, R.-R. Liu, J.-X. Yin, and Y. Wu, "Structural total least squares algorithm for locating multiple disjoint sources based on AOA/TOA/FOA in the presence of system error," *Frontiers Inf. Technol. Electron. Eng.*, vol. 19, no. 7, pp. 917–936, Sep. 2018.
[58] C. Liao, X. Hu, S. Zhang, X. Li, Q. Yin, Z. Zhang, and L. Zhang, "Joint inversion of gravity, magnetotelluric and seismic data using the alternating direction method of multipliers," *Geophys. J. Int.*, vol. 229, no. 1, pp. 203–218, Oct. 2021.
[59] Z. Wang, S. Huo, X. Xiong, K. Wang, and B. Liu, "A maximally split and adaptive relaxed alternating direction method of multipliers for regularized extreme learning machines," *Mathematics*, vol. 11, no. 14, p. 3198, Jul. 2023.
[60] J. Zhang, K. Wang, Y. Wang, X. Cheng, and Z. Zhou, "A localization algorithm based on alternating direction multiplier method," in *Proc. 8th Int. Conf. Intell. Comp. Signal Process. (ICSP)*, Xi'an, China, Apr. 2023, pp. 1443–1447.
[61] H. Xiong, M. Peng, S. Gong, and Z. Du, "A novel hybrid RSS and TOA positioning algorithm for multi-objective cooperative wireless sensor networks," *IEEE Sensors J.*, vol. 18, no. 22, pp. 9343–9351, Nov. 2018.
[62] L. Ljung, *System Identification Theory for the User*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 1999.
[63] P. Bachmann, *Zahlentheorie: Zweiter Band—Die Analytische Zahlentheorie*. Leipzig, Germany: B. G. Teubner Verlag, 1894.
[64] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*. Leipzig, Germany: B. G. Teubner Verlag, 1909.
[65] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. Hoboken, NJ, USA: Wiley, 1987.
[66] Z. Li, D. C. Dimitrova, D. H. Raluy, and T. Braun, "TDOA for narrow-band signal with low sampling rate and imperfect synchronization," in *Proc. 7th IFIP Wireless Mobile Netw. Conf. (WMNC)*, Dimitrov, Portugal, May 2014, pp. 1–8.
[67] Z. Li, T. Braun, X. Zhao, Z. Zhao, F. Hu, and H. Liang, "A narrow-band indoor positioning system by fusing time and received signal strength via ensemble learning," *IEEE Access*, vol. 6, pp. 9936–9950, 2018.
[68] A. Mahmood, R. Exel, H. Trsek, and T. Sauter, "Clock synchronization over IEEE 802.11—A survey of methodologies and protocols," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 907–922, Apr. 2017.

BAMRUNG TAUSIESAKUL (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from Chulalongkorn University, Bangkok, Thailand, in 2002 and 2004, respectively, and the Dr.-Ing. degree in communication technology from Leibniz University Hannover, Hanover, Germany, in 2010.

He was a Postdoctoral Researcher with the School of Telecommunication Engineering, University of Vigo, Vigo, Spain, from 2011 to 2012, and a Research Associate with the School of Science and Technology, Department of Engineering, City, University of London, London, U.K., from 2013 to 2014. From 2015 to 2023, he was a Lecturer with the Department of Electrical Engineering, Faculty of Engineering, Srinakharinwirot University, Bangkok, Thailand. Since 2024, he has been a Lecturer with the Department of Electrical Engineering, Faculty of Engineering, Mahidol University, Nakhon Pathom, Thailand. His research interests include signal processing and wireless communications.

KRISSADA ASAVASKULKIET (Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering and the Ph.D. degree from the Department of Electrical Engineering, Chulalongkorn University, Bangkok, Thailand, in 2001, 2004, and 2011, respectively.

He was appointed as a Lecturer and an Assistant Professor with the Department of Electrical Engineering, Mahidol University, Nakhon Pathom, Thailand, in 2011 and 2014, respectively. His current research interests include image and video processing, signal processing, and digital signal processing (DSP) in telecommunication.