Measurement-converted Kalman filter tracking with Gaussian intensity attenuation signal in wireless sensor networks

Sha Wen¹,², Liqiang Xing¹, Xiaoqing Hu³ and Hui Zhang⁴

Abstract
In this article, the target tracking problem in a wireless sensor network with nonlinear Gaussian signal intensity attenuation model is considered. A Bayesian filter tracking algorithm is presented to estimate the locations of moving source that has unknown central signal intensity. This approach adopts a measurement conversion method to remove the measurement nonlinearity by the maximum likelihood estimator, and a linear estimate of the target position and its associated noise statistics obtained by the Newton–Raphson iterative optimization steps are applied into the standard Kalman filter. The Monte Carlo simulations have been conducted in comparison with the commonly used extended Kalman filter with an augmented state that consists of both the original target state and the augmentative central signal intensity. It is observed that the proposed measurement-converted Kalman filter can yield higher accurate estimate and nicer convergence performance over existing methods.

Keywords
Target tracking, wireless sensor networks, maximum likelihood estimation, extended Kalman filter, Gaussian attenuation

Introduction
Recent advances in micro-electro-mechanical systems, networking systems, and embedded microprocessor technologies have drawn tremendous interests and applications of wireless sensor networks (WSNs). A common issue in WSNs is the source localization, which is inherent to many monitoring applications, such as target tracking, infrastructure monitoring, habitat sensing, rescue, and emergency response.¹⁻³ The objective of source localization is to sequentially estimate the positions of a moving target by sensing the signal emitted from the target at a subset of distributed network sensors. These sensors collaborate by reporting sensing measurements to a signal processing center which subsequently estimates the source locations according to the received measurement data.

There exist many methods for source localization that are based on measurement models such as time delay of arrival (TDOA),⁴ direction of arrival (DOA),⁵ and received signal strength (RSS).⁶ Generally, the signal intensity measurements are very convenient and

¹China National Institute of Standardization (CNIS), Beijing, China
²Electronic Technology Information Research Institute, Ministry of Industry and Information Technology of China, Beijing, China
³Institute of Acoustics, Chinese Academy of Sciences, Beijing, China
⁴Department of Engineering Physics, Tsinghua University, Beijing, China
Sha Wen, China National Institute of Standardization (CNIS), Beijing 100191, China

Corresponding author:
Xiaoqing Hu, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, China.
Email: wensha0103@163.com; auxqhu@gmail.com

Creative Commons CC-BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (http://www.creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (http://www.uk.sagepub.com/aboutus/openaccess.htm).
economical to locate a moving target since no additional sensor functionalities and measurement features are required compared to DOA or TDOA modalities. It is known that the intensity of signal attenuates as a function of distance from the source. In the energy-based acoustic source localization applications, the acoustic energy decays exponentially with respect to the distance from source in a free space. In the medical ultrasound applications, the echo signal reflects attenuation and scattering properties of tissue that may correlate with various disease states, and the power spectra of received signals are approximately Gaussian. In the laser manufacturing applications, the nonlinear optical interactions with matter are available from laser beams because of their high power and spectral brightness, and the transverse profile of the optical intensity of the beam can be approximated by a Gaussian function. In the monitoring applications of forest fire, a time-varying elliptical forest fire spread model is adopted to fit to empirical data. This model considers a decreasing intensity Gaussian distribution from the head of fire to the fire tail. In this article, a type of Gaussian attenuation model is presented to facilitate more accurate characterization of signal intensity. Along this direction, we will consider the source localization problem of a moving target with a Gaussian attenuation model of signal intensity.

Previously, several least-squares methods for energy-based source localization using an isotropic signal energy attenuation model are proposed by Li and Hu and Deng and Liu. A more generalized statistical model of energy observation is derived by Meesookho et al., in which an arbitrary acoustic source is characterized by an unknown correlation function, and a weighted direct/one-step least-squares-based algorithm incorporating the dependence of unknown parameters is investigated to reduce the computational complexity. The energy-based sensor network source localization problem is solved by the maximum likelihood (ML) estimation coupled with incremental optimization method in the work by Rabbat and Nowak; however, this approach cannot converge globally and even diverge due to the problem being highly nonconvex. A new normalized incremental optimization algorithm is developed to avoid trapping into the local minima using a diminishing stepsize in the work by Shi and He. In the work by Blatt and Hero, the single-source localization problem is formulated as a convex feasibility problem instead of nonlinear least squares and giving a fast converging localization solution by finding a point at the intersection of some convex sets using projection onto convex sets method. A rough location estimation method is designed to localize an isotropic energy source using measurements from distributed sensors based on kernel averaging techniques. In the works by Niu and Varshney and Ozdemir et al., ML-based target localization approaches using the quantized data of signal intensity measurements are proposed to reduce the communication requirements in WSNs and employing a sequential quadratic programming–based grid search within the surveillance region and a reasonable interval of source signal power to find an approximate maximum location. In the work by Masazade et al., an iterative source localization based on the approximation of the posterior probability distribution function of source location using a Monte Carlo method is proposed for WSNs with quantized source energy measurements. However, due to the lack of knowledge of the target’s motion, these methods hardly obtain nice localization performance without using prediction and filtering algorithms. In the work by Sheng and Hu, a sequential source localization method using particle filter is put forward to estimate and track multiple target locations. The large number of particles to represent the non-Gaussian probability density function (pdf) of the target locations involve expensive calculations. A weighted extended Kalman filter (EKF) algorithm based on recursive weighted least-squares optimization for WSNs is presented by Wang and Wu, and the source location is calculated iteratively by taking a weighted average of the local estimates based on the tasking sensor nodes’ reliability. In the work by Mysorewala et al., an EKF-based localization method using an elliptical forest fire spread model is developed, but the linearization procedure of nonlinear measurement function in the EKF algorithm may incur larger errors in the true posterior mean and covariance of the target state.

To overcome the drawback of EKF, many measurement conversion methods have been proposed to transform the nonlinear measurement models into linear ones and estimate the covariances of the converted measurement noises and then use the Kalman filter, which clearly has better accuracy and consistency than the EKF. These conversion methods always require both the range and bearing measurements (and possibly other parameters) of the target, which cannot be applicable to the case where only the signal intensity measurements are available for target tracking. In this article, we propose a new measurement conversion method using ML estimation with Gaussian intensity attenuation model to compute the converted measurement noise. A linear estimate of the target position can be obtained by an iteration that amounts to relinearization of the measurement equation to reduce the effects of linearization errors, and an approximate covariance matrix of the converted measurement noise is achieved based on Gaussian approximation of the localization error.

We begin by introducing our signal models and formulating the target tracking problem in WSNs in section “Models and problem formulation,” where the signal
intensity attenuates Gaussianly with an unknown central signal intensity. We delineate the EKF benchmark that will come handy in the subsequent derivations, and an EKF with augmented state is designed to estimate the augmented state vector that consists of both the original target state and the augmentative central signal intensity parameter in section “Augmented-state EKF tracking algorithm.” In section “Measurement-converted Kalman filtering algorithm,” we lay out the measurement conversion method using ML estimation, and the new converted target position will be followed by a Kalman filter for recursive estimation of the target state. The performance evaluations of the proposed algorithms are conducted by simulation comparisons in section “Numerical simulations.” Finally, some conclusions are given in section “Conclusion.”

Models and problem formulation

Consider an ad hoc WSN that tracks a target in the two-dimensional plane. The state evolution of target is described by a multivariate discrete-time random process

\[ X_k = F_k X_{k-1} + G_k w_{k-1} \]

(1)

where \( X_k \in \mathbb{R}^N \) is the state vector of the target at the \( k \)th time step, consisting of the position vector \( x_k = [x(k), y(k)]^\top \) and the velocity vector \( \dot{x}_k = [\dot{x}(k), \dot{y}(k)]^\top \). The input noise \( w_k \) is a Gaussian random vector with zero mean and a covariance matrix \( Q_k \); \( F_k \) and \( G_k \) are the state transition matrix and input matrix, respectively.

Signal model

We assume that \( N \) identical sensor nodes with the same noise statistics are densely deployed over a sensing field. The position of the \( n \)th (\( 1 \leq n \leq N \)) sensor node, denoted by \( \rho_n = [x_n, y_n]^\top \), is assumed to be known. The signal emitted from a point source propagates omnidirectionally through the medium, and the intensity of the source will attenuate along the propagation path like an elliptical shape as shown in Figure 1, which can be represented by a Gaussian attenuation model

\[ h_n(x_k) = a \exp \left( -\frac{1}{2} (x_k - \rho_n)^\top \Sigma^{-1} (x_k - \rho_n) \right) \]

(2)

where \( a \) is the signal intensity at the center of point source and \( \Sigma \) is a variance that determines the signal propagation range and orientation.

Upon sensing a signal intensity greater than a trigger threshold value \( r \), the sensor acts as tasking node to perform tracking task, and the set of indices of tasking sensors at time \( k \) can be denoted by

\[ L_k = \{ n|h_n(x_k) \geq r, 1 \leq n \leq N \} \]

and then the signal intensity received at the \( n \)th sensor during time \( k \) can be expressed as

\[ z_n(k) = h_n(x_k) + v_n(k) \quad n \in L_k \]

(3)

where \( v_n(k) \) is the observation noise which is modeled as zero-mean additive white Gaussian noise random variable with variance \( \sigma^2_n \) and mutually independent with \( w_k \) and \( X_k \). As the source moves in the sensing field, \( L_k \) will change with respect to \( k \). Denote \( \ell_k \) to be the cardinal number of \( L_k \) (i.e. the number of elements in \( L_k \)); the sensor measurements may be represented in a matrix form as follows

\[ z_k = h_k(x_k) + v_k = \begin{bmatrix} h_1(x_k) \\ h_2(x_k) \\ \vdots \\ h_{\ell_k}(x_k) \\ e_1(x_k) \\ e_2(x_k) \\ \vdots \\ e_{\ell_k}(x_k) \end{bmatrix} + \begin{bmatrix} v_1(k) \\ v_2(k) \\ \vdots \\ v_{\ell_k}(k) \end{bmatrix} = a e_k(x_k) + v_k = a \begin{bmatrix} v_1(k) \\ v_2(k) \\ \vdots \\ v_{\ell_k}(k) \end{bmatrix} \]

(4)

where \( e_n(x_k) = \exp \{ -1/2(x_k - \rho_n)^\top \Sigma^{-1} (x_k - \rho_n) \} \), \( 1 \leq n \leq \ell_k \) is the Gaussian attenuation function of signal intensity with respect to the \( n \)th sensor, and \( e_n(x_k) = [e_1(x_k), \ldots, e_{\ell_k}(x_k)]^\top \). The subscripts of \( h \) and \( e \) in equation (4) refer to the indices within \( L_k \) rather than node indices among all \( N \) sensor nodes, and then, the covariance matrix of \( v_k \) is

\[ R_k = \text{diag} \left( \sigma^2_1, \ldots, \sigma^2_{\ell_k} \right) \]

(5)

The practical signal sources with such attenuation model can be found in ultrasounds,\(^{10}\) lasers,\(^{11}\) and fire sources.\(^{12}\)
**Problem formulation**

For simplicity, we only consider the problem of tracking a single target, but nevertheless our proposed tracking algorithm is also applicable to multi-target tracking as the targets can be classified based on the target signatures.\(^\text{23}\) When the target moves through the monitored area, these tasking nodes that have detected the target dynamically form a cluster. One of the nodes in the cluster will be selected as the cluster head, and each cluster member transmits its signal observation to the cluster head. The cluster head serves as the fusion center of the signal and information processing. It selects a set of tasking nodes from the cluster for tracking computation and then broadcasts the estimated target information to its neighbor nodes for the subsequent estimation.\(^\text{30}\) More elaborate clustering routing protocols\(^\text{31–33}\) can be considered but are beyond the scope of this article.

Let \(Z_{1:k}\) denote the measurements from initial time up to and including time \(k\), that is, \(Z_{1:k} = \{z_1, \ldots, z_k\}\). The problem is for the cluster head to estimate the target state \(X_k\), denoted by \(\hat{X}_{h(k)}\), given the measurements \(Z_{1:k}\) with the unknown central signal intensity parameter \(a\) by way of a recursive Bayesian filtering formulation.

**Augmented-state EKF tracking algorithm**

The basic filtering solution to the state estimation problem can be obtained by a two-stage recursive process of prediction and update. Since the unknown parameter \(a\) is not included in the ordinary state variables \(X_k\), the clairvoyant EKF cannot be directly applied to the additional parameter estimation case. The augmented model provides a simple method to adjust the state model without altering the filtering framework. In this section, we delineate an augmented-state extended Kalman filter (AEKF) algorithm with the Gaussian signal attenuation model, in which the unknown central signal intensity parameter will be to act on the augmented state, and in turn act on the original model states; then correspondingly the underlying EKF algorithm is not changed and retains most or all of its properties.\(^\text{34}\)

Let \(\theta_k = [x(k), \dot{x}(k), y(k), \dot{y}(k), a]^T\) be a new augmented state vector. As a constant, the parameter \(a\) can be assumed to vary slowly, and then, the corresponding augmentative evolution model of \(\theta_k\) is given by

\[
\theta_k = \tilde{F}_k \theta_{k-1} + \tilde{G}_k \tilde{w}_{k-1}
\]

where \(\tilde{w}_{k-1} \sim N\left(0, \tilde{Q}_k\right)\), \(\tilde{Q}_k = \text{diag}(Q_k, \sigma_a^2)\) in which the variance \(\sigma_a^2\) can be used as a tuning parameter, and

\[
\tilde{F}_k = \begin{bmatrix} 1 & \Delta_k & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta_k & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \tilde{G}_k = \begin{bmatrix} \Delta_k^1 & 0 & 0 \\ 0 & \Delta_k & 0 \\ 0 & 0 & \Delta_k \\ 0 & 0 & 0 \end{bmatrix}
\]

In the above equation, \(\Delta_k\) is the sampling time interval between two successive time steps \(k\) and \(k-1\). Using Bayes’ rule, one has

\[
p(\theta_k|Z_{1:k-1}) = \beta_k p(\theta_k|z_{1:k-1})p(Z_k|\theta_k)\tag{7}
\]

where the proportionality factor \(\beta_k^{-1} = p(Z_k|Z_{1:k-1})\) is independent on \(\theta_k\).

Assume that the previous estimate of target state \(\hat{X}_{k-1|k-1}\) and its error covariance matrix \(P_{k-1|k-1}\) are known at step \(k\). Correspondingly, the augmented state vector estimate at the time \(k-1\) can be denoted as \(\hat{\theta}_{k-1|k-1} = [\hat{X}_{k-1|k-1}, \dot{a}_{k-1}]^T\), and its associated error covariance \(\hat{P}_{k-1|k-1} = \text{diag}(P_{a_{k-1}}, \hat{Q}_{a_{k-1}})\), where \(\hat{a}_{k-1}\) and \(\hat{Q}_{a_{k-1}}\) are, respectively, the estimate and error covariance of unknown parameter \(a\). Now, given the distribution \(p(\theta_{k-1}|Z_{1:k-1}) = N(\hat{\theta}_{k-1|k-1}, \hat{P}_{k-1|k-1})\), the joint distribution of \(\theta_k\) and \(\theta_{k-1}\) conditioned on the previous measurements \(Z_{1:k-1}\) is

\[
p(\theta_{k-1}, \theta_k|Z_{1:k-1}) = p(\theta_k|\theta_{k-1})p(\theta_{k-1}|Z_{1:k-1}) = N\left(\begin{bmatrix} \hat{\theta}_{k-1|k-1} \\ \tilde{F}_k \hat{\theta}_{k-1|k-1} \end{bmatrix}, \tilde{G}_k \hat{P}_{k-1|k-1} \tilde{F}_k^T + \tilde{G}_k \hat{Q}_{a_{k-1}} \tilde{G}_k^T \right)\tag{8}
\]

and thus, the predictive distribution of \(\theta_k\), given the measurement history up to time step \(k-1\), can be calculated as

\[
p(\theta_k|Z_{1:k-1}) = N\left(\begin{bmatrix} \hat{\theta}_{k-1|k-1} \\ \tilde{F}_k \hat{\theta}_{k-1|k-1} \end{bmatrix}, \tilde{G}_k \hat{P}_{k-1|k-1} \tilde{F}_k^T + \tilde{G}_k \hat{Q}_{a_{k-1}} \tilde{G}_k^T \right)\tag{9}
\]

since it is the marginal distribution of \(\theta_k\) (see “Lemma A2” in Appendix 2). Consider a first-order Taylor series expansion approximation of the measurement function \(h_k\) around the predictive augmented state \(\hat{\theta}_{k|k-1}\) in equation (4)

\[
z_k \approx h_k(\hat{\theta}_{k|k-1}) + \bar{H}(\hat{\theta}_{k|k-1})[\theta_k - \hat{\theta}_{k|k-1}] + v_k\tag{10}
\]

where \(\bar{H}(\hat{\theta}_{k|k-1}) = (\partial h(x)/\partial x)|_{\theta_k - \hat{\theta}_{k|k-1}}\) is a \(\ell_k \times 5\) matrix, and its specific outcome can be found in Appendix 2. Note that \(\theta_k\) and \(v_k\) conditioned on the previous measurements \(Z_{1:k-1}\) are Gaussian and independent, and using equation (10), the joint distribution of \(\theta_k\) and \(z_k\) conditioned on \(Z_{1:k-1}\) can be approximated as
From equation (11), it is clear that the posterior distribution of \( \theta_k \) conditioned on \( Z_{1:k} \) can be found by the conditional distribution of jointly Gaussian random variables as (see “Lemma A2” in Appendix 2)

\[
p(\theta_k | Z_{1:k}) \simeq \mathcal{N}(\hat{\theta}_{k|k}, \hat{P}_{k|k})
\]

where

\[
\hat{\theta}_{k|k} = \hat{\theta}_{k|k-1} + \tilde{K}_k [z_k - h_k (\hat{\theta}_{k|k-1})]
\]

\[
\hat{P}_{k|k} = \hat{P}_{k|k-1} - \tilde{K}_k \tilde{H}(\hat{\theta}_{k|k-1}) \hat{P}_{k|k-1}
\]

\[
\tilde{K}_k = \hat{P}_{k|k-1} \tilde{H}^T_k (\hat{\theta}_{k|k-1}) [\tilde{H}(\hat{\theta}_{k|k-1}) \hat{P}_{k|k-1} \tilde{H}^T_k (\hat{\theta}_{k|k-1}) + R_k]^{-1}
\]

The above equations lead to the state update \( \hat{X}_{k|k} \) of the target state \( X_k \) with the current measurement \( z_k \) at time \( k \) from the augmented state estimate \( \hat{\theta}_{k|k} \) and its corresponding error covariance \( \hat{P}_{k|k} \) from \( \hat{P}_{k|k} \). The recursive process of computing \( \hat{X}_{k|k} \) is described by the following AEEKF in Algorithm 1, where \( A[\cdot] \) takes entries of the matrix \( A \).

**Measurement-converted Kalman filtering algorithm**

The linearization errors of measurement function approximated by the Taylor series expansion in the EKF algorithm may build up and result in the filtering divergence.\(^{35}\) In this section, we propose a measurement conversion method to transform the nonlinear measurement model into linear one by a ML estimation method, in which the iterated re-linearization of measurement equation is used to reduce the effects of linearization errors. With the converted linear measurement model, a standard Kalman filter is applied for recursive estimation of the target state.

The conditional pdf of the measurement \( z_k \), given \( x_k \) from equations (4) and (5), can be written as follows

\[
p(z_k | x_k) = \mathcal{N}(h_k(x_k), R_k)
\]

\[
= \frac{1}{(2\pi)^{d/2}(\det \{ R_k \})^{1/2}} \cdot \exp \left\{ -\frac{1}{2} (z_k - h_k(x_k))^T R_k^{-1} (z_k - h_k(x_k)) \right\}
\]

The negative log-likelihood function of \( p(z_k | x_k) \) is proportional to a quadratic form

\[
L(x_k) = \frac{1}{2} (z_k - h_k(x_k))^T R_k^{-1} (z_k - h_k(x_k))
\]

the above equation ignores the constant terms not dependent on \( x_k \), and then, the ML estimate of \( x_k \) can be obtained by minimizing \( L(x_k) \). Note that the measurement function \( h_k(x_k) \) has an unknown parameter \( a \), and the solution of the optimization problem must lie on a stationary point where

\[
\frac{\partial L(x_k)}{\partial x_k} = 0, \text{and} \frac{\partial L(x_k)}{\partial a} = 0
\]

The second condition will lead to the following relation of \( x_k \) and \( a \)

\[
a = [e_k^T(z_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T(z_k) R_k^{-1} z_k
\]

**Algorithm 1: Augmented-state EKF algorithm.**

**Require**: \( \hat{X}_{0|0}, \hat{P}_{0|0}, \hat{\theta}_0^0 \) and \( \hat{P}_0^0 \)

for \( k = 1, 2, \cdots \) do

1: Construct augmented state \( \hat{\theta}_{k-1|k-1} \)
2. Construct augmentative error covariance \( \hat{P}_{k-1|k-1} \)
3. Predict augmented state estimate \( \hat{\theta}_{k|k-1} \) using equation (9)
4. Predict augmentative error covariance \( \hat{P}_{k|k-1} \) using equation (9)

update step: \( \hat{\theta}_{k|k} \) and \( \hat{P}_{k|k} \)

5. Linearize measurement function \( h_k(x_k) \) using equation (10)
6. Update augmented state estimate \( \hat{\theta}_{k|k} \) by the measurement \( z_k \) using equation (13)
7. Update augmentative error covariance \( \hat{P}_{k|k} \) using equation (14)

**Ensure**: \( \hat{X}_{k|k} \) and \( \hat{P}_{k|k} \)

8. Obtain the target state estimate \( \hat{X}_{k|k} = \hat{\theta}_{k|k}[1:4] \)
9. Obtain the error covariance of target state \( \hat{P}_{k|k} = \hat{P}_{k|k}[1:4][1:4] \)

end for
Substituting equation (19) into equation (17), the unknown parameter \( a \) can be eliminated, and giving a modified negative log-likelihood function

\[
L'(x_k) = \frac{1}{2} \varepsilon_k^T (I - e_k(x_k)) [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T (x_k) R_k^{-1} T R_k^1 \cdot (I - e_k(x_k)) [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T (x_k) R_k^{-1} z_k
\]

\[
= \frac{1}{2} \varepsilon_k^T R_k^{-1} z_k - \frac{1}{2} \varepsilon_k^T e_k(x_k) [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T (x_k) R_k^{-1} z_k
\]

The minimization problem of equation (20) can be equivalently stated as a nonlinear unconstrained optimization problem, and the common solution to the optimization problem is found through the following Newton–Raphson iterative method

\[
x_{k+1} = x_k - \left[ \frac{\partial^2 L(x_k)}{\partial x_k^T} \right]^{-1} \frac{\partial L(x_k)}{\partial x_k}
\]

where \( t \) is the iteration step, and the initial value \( x_0 = x_{k+1} \). The ML estimation problem equivalently is a nonlinear least-squares problem \( L(x_k) = \frac{1}{2} \gamma^T \gamma = \gamma^T (I - e_k(x_k)) [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T (x_k) R_k^{-1} z_k \). The first-order Taylor series expansion of measurement function in the standard EKF is consistent with the Gauss–Newton method to solve the nonlinear least-squares problem. Unlike Newton–Raphson method, the Gauss–Newton algorithm approximates the Hessian matrix of \( L(x_k) \) by ignoring the second-order derivative terms of the residuals \( \gamma^T (I - e_k(x_k)) [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} e_k^T (x_k) R_k^{-1} z_k \). As a consequence, the convergence of the Gauss–Newton method depends on whether the omitted second-order derivative terms of the residuals are large parts of the Hessian. When

\[
\left| \frac{\partial \gamma^T}{\partial x_k} \right| \gg \left| \frac{\partial^2 \gamma}{\partial x_k^T} \right|
\]

cannot be satisfied, it is shown that the Gauss–Newton method may not be locally convergent all.\(^{38}\)

In equation (21), the \( i \)th component of the gradient of the logarithm function \( L(x_k) \) with respect to \( x_k[i] \) is easily given as

\[
\frac{\partial L(x_k)}{\partial x_k[i]} = [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-2} [e_k^T (x_k) R_k^{-1} e_k(x_k)]^{-1} [e_k^T (x_k) R_k^{-1} e_k(x_k)] e_k^T (x_k) R_k^{-1} z_k
\]

Then, the element at the \( i \)th row and \( j \)th column of the Hessian matrix of the function \( L'(x_k) \) can be computed as

\[
\frac{\partial^2 L(x_k)}{\partial x_k[i] \partial x_k[j]} = \left[ e_k^T (x_k) R_k^{-1} e_k(x_k) \right]^{-2} \left[ e_k^T (x_k) R_k^{-1} e_k(x_k) \right]^{-1} e_k^T (x_k) R_k^{-1} R_k^{-1} \left( \frac{\partial^2 e_k^T (x_k)}{\partial x_k[i]} \right) R_k^{-1} e_k(x_k)
\]

The detailed derivation of equation (23) is put in the Appendix 2.

Considering a second-order Taylor series expansion of \( L(x_k) \) around \( x'_{k+1} \) in equation (17) and combining with equation (18) yield

\[
L(x_k) \approx L'(x'_{k+1}) + \frac{1}{2} (x_k - x'_{k+1})^T \frac{\partial^2 L(x_k)}{\partial x_k^T} (x_k - x'_{k+1})
\]

Using the above approximation, it can follow from equations (16) and (17) that

\[
p (x_k | x_{k+1}) \approx \alpha \exp \left\{ -\frac{1}{2} \left( x_k - x'_{k+1} \right)^T \frac{\partial^2 L(x_k)}{\partial x_k^T} \left( x_k - x'_{k+1} \right) \right\}
\]

for some constant \( \alpha \). That is, the estimate error \( x_k - x'_{k+1} \) of the position state \( x_k \) can be approximated as a Gaussian distribution with zero mean and covariance matrix

\[
\sigma_k^2 (k) = \frac{\partial^2 L(x_k)}{\partial x_k^T} \left( x_k - x'_{k+1} \right)
\]
Algorithm 2: Measurement-converted Kalman filter.

Require: $X_{00}$ and $P_{00}$
for $k = 1, 2, \ldots$ do
1: Compute the converted measurement $\hat{x}_{k|k}$ using equation (21)
2: Compute the converted noise covariance $\sigma^2_z(k)$ using equation (26)
prediction step: $X_{k|k-1}$ and $P_{k|k-1}$
3: Predict state estimate $X_{k|k-1}$ using equation (28)
4: Predict error covariance $P_{k|k-1}$ using equation (28)
update step: $X_{k|k}$ and $P_{k|k}$
5: Update state estimate $X_{k|k}$ by the converted measurement $\hat{x}_{k|k}$ using equation (31)
6: Update error covariance $P_{k|k}$ using equation (32)
end for

which is described by the following linear representation in the target state $X_k$

$$\hat{x}_{k|k} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} X_k + \varepsilon_k = CX_k + \varepsilon_k$$ (27)

where $\varepsilon_k \sim N(0, \sigma^2_z(k))$. We consider the estimate of the position state $\hat{x}_{k|k}$ as a new converted measurement and then it can be found that the converted measurement function C is linear, correspondingly $\varepsilon_k$ is the converted measurement noise. The converted measurement model will be used in Kalman filtering for the recursive update of the target state $X_k$.

From equation (9), the predictive distribution of $X_k$, given the measurement history up to time step $k - 1$, is

$$p(X_k|Z_{1:k-1}) = \mathcal{N}(\hat{x}_{k|k-1}, P_{k|k-1})$$

By Bayes' rule, and using the fact $\hat{x}_{k|k}$ is independent with $Z_{1:k-1}$ given $X_k$, since $\hat{x}_{k|k}$ is generated from $\varepsilon_k$, then we have the following joint probability distribution of $(X_k, \hat{x}_{k|k})$, and

$$p(X_k, \hat{x}_{k|k}|Z_{1:k-1}) = p(\hat{x}_{k|k}|X_k)p(X_k|Z_{1:k-1})$$

$$\simeq \mathcal{N}(\hat{x}_{k|k}, C\hat{x}_{k|k-1})$$ (29)

thus, the posterior distribution of $X_k$ conditioned on the history measurements $Z_{1:k-1}$ and the current converted measurement $\hat{x}_{k|k}$ can be obtained in the same way as for the AEKF in equation (12)

$$p(X_k|Z_{1:k-1}, \hat{x}_{k|k}) \simeq \mathcal{N}(\hat{X}_{k|k}, P_{k|k})$$ (30)

where

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k[\hat{x}_{k|k} - C\hat{X}_{k|k-1}]$$ (31)

$$P_{k|k} = P_{k|k-1} - K_kCP_{k|k-1}$$ (32)

$$K_k = P_{k|k-1}C^T[C P_{k|k-1} C^T + \sigma^2_z(k)]^{-1}$$ (33)

The recursive computing process of the measurement-converted Kalman filter (MCKF) is summarized in the following Algorithm 2.

We remark that the AEKF is initialized by additionally guessing the initial estimate $\hat{a}_0$ and error covariance $\sigma^2_0$ of unknown parameter $a$, so that its convergence performance is extremely sensitive to the initialization accuracy. In contrast the MCKF not only has no additional assumption on the initial state of unknown parameter $a$, and the nice convergence can be partly assisted by the fact that the initial estimate of the iterative process of measurement conversion comes from the Kalman predictor which is typically good, as such the sampling time interval is not too long for convergence. Typically, only a few iterations are sufficient to get the extremum because the longer iteration steps cannot promise more accurate estimate.

Numerical simulations

To show the efficiency of the proposed MCKF, it is applied to a target tracking application in comparison with the AEKF, the traditional extended Kalman filter (TEKF) without measurement conversion, and the pure ML estimator without Kalman filtering. Error performance of the filters is evaluated with Monte Carlo simulations in this section.

The monitored area is 50 m × 50 m and covered by $N = 50$ sensor nodes randomly, the trigger threshold of each sensor $r = 0.1$, and the measurement noise covariance of each sensor $\sigma^2_n = 5$. A moving target travels along a circular arc, and the sampling interval $\Delta_t = 0.1$ s (10 Hz). The process noise $w_t$ corresponds to the variable acceleration of the target and can be approximated by a white Gaussian sequence with zero mean and covariance matrix $Q_t = \text{diag}(10, 10)$. The unknown parameter $a$ is assumed to vary slowly in the
AEKF, and its tuning covariance is set to $\sigma_a^2 = 0.00001$; in the TEKF, $\sigma_a$ will be approximately computed by the ML estimator before the state update step.

In the experiments, the source central signal intensity is set at $a = 25$, and $\Sigma = \begin{bmatrix} 40 & 10 \\ 5 & 30 \end{bmatrix}$. The initial state estimate and the corresponding covariance matrix for the MCKF and TEKF, respectively, are chosen to be

$$\hat{X}_{0|0} = \begin{bmatrix} 0 \text{m} \\ 5 \text{m/s} \\ 25 \text{m} \\ 5 \text{m/s} \end{bmatrix}^T, \quad P_{0|0} = 0.01 \times \text{diag}(1 \quad 1 \quad 1 \quad 1)$$

Additionally, the initial estimate of the variable $a$ in the AEKF is roughly assumed to the trigger threshold of the sensor $a_0 = r = 0.1$, and the associated variance is set to a large value $a_0^2 = 2$ so that the estimate converges quickly and the influence of the initial guess $a_0$ soon will be negligible.

In total, 100 Monte Carlo runs with the above noise condition are performed. The results are summarized in the following. Figure 2 shows the tracking trajectories of the four tracking algorithms over a typical simulation run. Figure 3 describes the corresponding mean value of the population of target estimated location biases $\| \hat{x}_k - \hat{x}_{2|k} \|$ for each tracking algorithm, where $\hat{x}_{2|k}$ comes from the estimated state $\hat{X}_{2|k}$ of the target in the MCKF and TEKF; in the AEKF, $\hat{x}_{2|k} \in \theta_{3|k}$, and in the ML and MCKF, $t = 2$. It is quite clear that the MCKF has the best tracking performance for most time steps, while the ML performs worse in comparison to the filters, and this is because the ML estimation method only utilizes the measurement information but the other filtering methods additionally utilize the target moving information. Figure 4 further depicts the average estimated location bias of the whole tracking trajectory in each realization: the solid line with marker represents the average estimated location bias for all the time steps.

Figure 5 gives the mean of quadratic sums of the position state estimation errors in each time step for the four methods, and the distributions of tracking errors of the position states along $x$- and $y$-directions together with the corresponding error covariance ellipses of position state are plotted in Figure 6. The error covariance of the position state is defined as
\[ P_{k|k,xy} = \begin{bmatrix} P_{x|k}[1,1] & P_{x|k}[1,3] \\ P_{x|k}[3,1] & P_{x|k}[3,3] \end{bmatrix} \]

where \( P_{k|k}[i,i] \) denotes the diagonal element of the error covariance matrix \( P_{k|k} \). The major and minor axes of each ellipse correspond to the square roots of eigenvalues of the error covariance matrix, and the angles between these axes and \( x \)-axis are determined by the orientation of corresponding eigenvectors. We can see that the mean-squared error (MSE) of the position state estimation in the ML is the worst among all the algorithms. The MSE of the MCKF is lower than the AEKF and TEKF, and this is because the MCKF can achieve smaller error covariance among them, as such the tracking errors of the position states in the MCKF for most realizations lie in the ellipse. In contrast, the MSE of the TEKF is larger than the AEKF and MCKF, and this is because the estimation error statistics of unknown central signal intensity parameter are not considered in the TEKF, which would reduce the filtering performance. This result indicates that the MCKF based on the measurement conversion method yields significant performance over the TEKF and AEKF based on the Taylor series linearization of measurement function.

Figure 7 shows the average estimate of \( a \) of 100 realizations: the black solid line without marker represents the real source signal intensity \( a \), and the average estimates of \( a \) in the AEKF are given for three different initial state cases of \( a \): (a) \( \hat{a}_0 = 0.1, \sigma^2_0 = 2 \), (b) \( \hat{a}_0 = 100, \sigma^2_0 = 2 \), and (c) \( \hat{a}_0 = 0.1, \sigma^2_0 = 0.02 \).
AEKF, TEKF, and the ML. The MCKF can yield higher accurate estimate and nicer convergence performance than the ML estimation approach. The discussion in this work is limited to the case of single target. Multiple target scenarios will be addressed in future works.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Foundation of President of the China National Institute of Standardization (No. 512016Y-4697) and the National Natural Science Foundation of China (No. 61501448).

Table 1. Average RMSEs under different measurement noise and monitoring range conditions with different initial state cases of α.

| Measurement noise | Average RMSE |
|-------------------|---------------|
|                   | ML            | MCKF          | TEKF          | AEKF $\alpha_0^2 = 0.02$ | AEKF $\alpha_0^2 = 0.2$ | AEKF $\alpha_0^2 = 2$ |
| $\sigma^2 = 0.1, r = 0.1$ | 0.0936        | 0.0737        | 0.0760        | 0.0799                  | 0.0754                  | 0.0749                  |
| $\sigma^2 = 0.1, r = 5$   | 0.1011        | 0.0786        | 0.0804        | 0.0853                  | 0.0799                  | 0.0789                  |
| $\sigma^2 = 1, r = 0.1$   | 0.2923        | 0.1866        | 0.2183        | 0.2557                  | 0.1972                  | 0.1889                  |
| $\sigma^2 = 1, r = 5$    | 0.3165        | 0.1986        | 0.2276        | 0.2718                  | 0.2098                  | 0.2010                  |
| $\sigma^2 = 5, r = 0.1$   | 0.6560        | 0.3476        | 0.4293        | 0.7432                  | 0.4101                  | 0.3572                  |
| $\sigma^2 = 5, r = 5$    | 0.7095        | 0.3694        | 0.4408        | 0.8014                  | 0.4357                  | 0.3791                  |

RMSE: root mean square of the target estimated position error; ML: maximum likelihood; MCKF: measurement-converted Kalman filter; TEKF: traditional extended Kalman filter; AEKF: augmented-state extended Kalman filter.

(RMSE) $\sqrt{\mathbb{E}[(x_t - \hat{x}_k^k)^2]}$ in all the time steps under different scenarios. As $r$ increases, the monitoring range of each sensor will diminish, and then, the number of sensors that detect the target decreases and the RMSEs of these tracking algorithms become higher. In the low measurement noise condition, the impact of the noise has negligible magnitude and then the observations are very accurate; thus, all the tracking algorithms attain the same RMSE. However, under the high measurement noise conditions, the measurement errors become bigger, the RMSE of the ML becomes the worst, the RMSE of the AEKF increases as the initial covariance of $a$ decreases, but the RMSE of the proposed MCKF is still lower than the AEKF and TEKF. These results indicate that the MCKF can yield higher accurate estimate and nicer convergence performance than the AEKF, TEKF, and the ML.

Conclusion

In this article, we consider a practical Gaussian attenuation model to facilitate more accurate characterization of source signal intensity. With the nonlinear Gaussian intensity attenuation model, two Bayesian filtering tracking algorithms (AEKF and MCKF) are developed. The Taylor series linearization error of nonlinear measurement function in the AEKF would lead to unstable tracking performance, and its convergence performance is extremely sensitive to the initialization accuracy. The measurement conversion method in the MCKF is used to remove the measurement nonlinearity using the ML estimation method, and a linear estimate of the target position is obtained by an iteration that amounts to re-linearization of the measurement equation to reduce the effects of linearization errors. The converted measurement and its associated noise statistics are then used in a standard Kalman filter for recursive update of the target state. Simulation results have shown that the MCKF can yield higher accurate estimate and nicer convergence performance than the AEKF and the TEKF as well as the commonly used ML estimation approach. The discussion in this work is limited to the case of single target. Multiple target scenarios will be addressed in future works.

References

1. Estrin D, Girod L and Srivastava GPM. Instrumenting the world with wireless sensor networks. In: Proceedings of the IEEE international conference on acoustics, speech, and signal processing, Salt Lake City, UT, 7–11 May 2001, pp.2033–2036. New York: IEEE.
2. Ho K. Bias reduction for an explicit solution of source localization using TDOA. ACM T Sensor Network 2006; 2(1): 1–38.
3. He T, Krishnamurthy S, Luo L, et al. VigilNet: an integrated sensor network system for energy-efficient surveillance. ACM T Sensor Network 2006; 2(1): 1–38.
4. Conner W, Krishnamurthy L and Want R. Making everyday life easier using dense sensor networks. In: Proceedings of the third international conference on ubiquitous computing, Atlanta, GA, 30 September–2 October 2001, pp.49–55. London: Springer.
5. Ho K. Bias reduction for an explicit solution of source localization using TDOA. IEEE T Signal Proces 2012; 60(5): 2101–2114.
6. Osborne RW, Bar-Shalom Y, George J, et al. Statistical efficiency of simultaneous target and sensors localization with position dependent noise. Proc SPIE 2012; 8392: 203–214.
Wen et al.

7. Hu YH and Sheng X. Dynamic sensor self-organization for distributed moving target tracking. *J Signal Process Sys* 2008; 51(2): 161–172.
8. Nguyen T, Septier F, Rajaona H, et al. New perspectives on multiple source localization in wireless sensor networks. *CoRR*, https://arxiv.org/abs/1504.05837.
9. Li D and Hu Y. Energy based collaborative source localization using acoustic micro-sensor array. *EURASIP J Appl Sig P* 2003; 2003(4): 321–337.
10. Wear K. A Gaussian framework for modeling effects of frequency-dependent attenuation, frequency-dependent scattering, and gating. *IEEE T Ultrason Ferr 2002; 49(11): 1572–1582.*
11. Miller D and Smith S. Variable attenuator for gaussian laser beams. *Appl Opt* 1978; 17(23): 3804–3808.
12. Mysorewala M, Popa D and Lewis F. Multi-scale adaptive sampling with mobile agents for mapping of forest fires. *J Intell Robot Syst 2009; 54: 535–565.*
13. Deng K and Liu Z. Weighted least-squares solutions of energy-based collaborative source localization using acoustic array. *Int J Comput Sci Netw Secur 2007; 7(1): 535–565.*
14. Meesookho C, Mitra U and Narayanan S. On energy-based acoustic source localization for sensor networks. *IEEE T Signal Proc 2008; 56(1): 365–377.*
15. Rabbat MG and Nowak RD. Decentralized source localization and tracking [wireless sensor networks]. In: *Proceedings of the IEEE international conference on acoustics, speech, and signal processing* (vol. 3), Montreal, QC, Canada, 17–21 May 2004, pp.921–924. *New York: IEEE.*
16. Shi Q and He C. A new incremental optimization algorithm for ML-based source localization in sensor networks. *IEEE Signal Proc Let 2008; 15: 45–48.*
17. Blatt D and Hero AO. Energy-based sensor network source localization via projection onto convex sets. *IEEE T Signal Proc 2006; 54(9): 3614–3619.*
18. Rabbat M, Nowak R and Bucklew J. Robust decentralized source localization via averaging. In: *Proceedings of the IEEE international conference on acoustics, speech, and signal processing*, Philadelphia, PA, 18–23 March 2005.
19. Niu R and Varshney PK. Target location estimation in sensor networks with quantized data. *IEEE T Signal Proc 2006; 54(12): 4519–4528.*
20. Ozdemir O, Niu R and Varshney PK. Channel aware target localization with quantized data in wireless sensor networks. *IEEE T Signal Proc 2009; 57(3): 1190–1202.*
21. Masazade E, Niu R, Varshney P, et al. Energy aware iterative source localization for wireless sensor networks. *IEEE T Signal Proc 2010; 58(9): 4824–4835.*
22. Liu J, Xiao W, Lewis F, et al. Energy-efficient distributed adaptive multisensor scheduling for target tracking in wireless sensor networks. *IEEE T Instrum Meas 2009; 58(6): 1886–1896.*
23. Sheng X and Hu Y. Sequential acoustic energy based source localization using particle filter in a distributed sensor network. In: *Proceedings of the IEEE international conference on acoustics, speech, and signal processing*, Montreal, QC, Canada, 17–21 May 2004.
24. Wang C and Wu D. Decentralized target positioning and tracking based on a weighted extended Kalman filter for wireless sensor networks. *Wirel Netw 2013; 19(8): 1915–1931.*
25. Mysorewala M, Popa D and Lewis F. Multi-scale adaptive sampling with mobile agents for mapping of forest fires. *J Intell Robot Syst 2009; 54(4): 535–565.*
26. Li X and Jilkov V. A survey of maneuvering target tracking—part III: measurement models. In: *Proceedings of the SPIE conference on signal and data processing of small targets* (vol. 4473), San Diego, CA, 29 July 2001, pp.423–446. *Bellingham: SPIE.*
27. Zhao Z, Li X and Jilkov V. Best linear unbiased filtering with nonlinear measurements for target tracking. *IEEE T Aero Elec Sys 2004; 40(4): 1324–1336.*
28. Li X and Jilkov V. A survey of maneuvering target tracking: dynamic models. In: *Proceedings of the SPIE conference on signal and data processing of small targets* (vol. 4473), Orlando, Fl, 24–27 April 2000, pp.212–235. *Bellingham: SPIE.*
29. Li X, Hu Y, Hu D, et al. Detection, classification, and tracking of targets. *IEEE Signal Proc Mag 2002; 19(2): 17–29.*
30. Hu X, Hu Y and Xu B. Energy balanced scheduling for target tracking in wireless sensor networks. *ACM T Sensor Network 2014; 11(1): 21: 1–21: 29.*
31. Liu X. A survey on clustering routing protocols in wireless sensor networks. *Sensors 2012; 12(8): 113–153.*
32. Abbasi A and Younis M. A survey on clustering algorithms for wireless sensor networks. *Comput Commun 2007; 30(14): 2826–2841.*
33. Akkaya K and Younis M. A survey on routing protocols for wireless sensor networks. *Ad Hoc Netw 2005; 3(3): 325–349.*
34. Maeder U. *Augmented models in estimation and control*. PhD Dissertation, ETH Zurich, Zurich, 2010.
35. Bar-Shalom Y, Li XR and Kirubarajan T. *Estimation with applications to tracking and navigation*. New York: John Wiley & Sons, 2001.
36. Hu X, Bao M, Zhang X, et al. Generalized iterated Kalman filter and its performance evaluation. *IEEE T Signal Proc 2015; 63(12): 3204–3217.*
37. Bell B and Cathey F. The iterated Kalman filter update as a Gauss-Newton method. *IEEE T Automat Contr 1993; 38(2): 294–297.*
38. Dennis JE and Schnabel RB (eds). *Numerical methods for unconstrained optimization and nonlinear equations*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1987.
39. Wang X, Fu M and Zhang H. Target tracking in wireless sensor networks based on the combination of KF and MLE using distance measurements. *IEEE T Mobile Comp 2012; 11(4): 567–576.*
40. Sarkka Y. *Bayesian estimation of time-varying processes: discrete-time systems*. Espoo: Aalto University, 2011.

**Appendix I**

**Notation**

\[ a \] source central signal intensity
\( \hat{a}_k \) estimate of the source central intensity at time \( k \)
\( C \) converted measurement function
\( e_k(x_k) \) Gaussian attenuation functions vector of \( \ell_k \) nodes
\( e_n(x_k) \) Gaussian attenuation function of node \( n \)
\( F_k, G_k \) state transition matrix and input matrix
\( F_{\ell_k}, G_{\ell_k} \) augmentative state transition matrix and input matrix
\( H(\theta_k) \) Jacobian matrix of \( h_k(x_k) \) with respect to \( \theta_k \)
\( h_k(x_k) \) intensity measurement functions vector of \( \ell_k \) nodes
\( h_n(x_k) \) intensity measurement function of node \( n \)
\( L(x_k) \) negative log-likelihood function of \( x_k \)
\( P_{\ell_k/k} \) error covariance matrix of \( X_{\ell_k} \) at time \( k \)
\( P_{\theta_{\ell_k}/k} \) error covariance matrix of \( \theta_{\ell_k} \) at time \( k \)
\( Q_k \) covariance of the input noise \( w_k \)
\( Q_{\ell_k} \) covariance of the augmentative input noise \( \tilde{w}_k \)
\( R_k \) covariance of the measurement noise \( r_k \) by \( \ell_k \) nodes
\( X_k \) target state at time \( k \)
\( X_{\ell_k/k} \) state estimate of \( X_k \) at time \( k \)
\( x_k \) target position state at time \( k \)
\( \dot{x}_k \) target velocity state at time \( k \)
\( \tilde{x}_{\ell_k/k} \) converted measurement at time \( k \)
\( Z_{1:k} \) intensity measurements up to time \( k \)
\( z_n(k) \) intensity measurements vector of \( \ell_k \) nodes
\( z_{\ell_n}(k) \) intensity measurement of node \( n \) at time \( k \)
\( \Delta_k \) system sampling time interval
\( \theta_k \) augmented target state vector at time \( k \)
\( \dot{\theta}_{\ell_k/k} \) state estimate of \( \theta_k \) at time \( k \)
\( \rho_n \) position vector of node \( n \)
\( \rho_{\ell_k} \) error covariance of \( \hat{a}_k \) at time \( k \)
\( \sigma_n^2 \) covariance of the measuring noise \( v_{n}(k) \)
\( \sigma_{\ell_k}^2 \) covariance of converted measurement noise \( \tilde{e}_{n}(k) \)

where \( \Xi_x = E\{x\} \) and \( \Xi_y = E\{y\} \), and then, the marginal and conditional densities of \( x \) and \( y \) are given as follows

\[
p(x) = N(\Xi_x, \text{Cov}\{x, x\}) \tag{35}
\]

\[
p(y) = N(\Xi_y, \text{Cov}\{y, y\}) \tag{36}
\]

\[
p(x|y) = N(\Xi_x + \text{Cov}\{x, y\}\text{Cov}^{-1}\{y, y\}(y - \Xi_y), \text{Cov}\{x, x\} - \text{Cov}\{x, y\}\text{Cov}^{-1}\{y, y\}\text{Cov}^T\{x, y\}) \tag{37}
\]

\[
p(y|x) = N(\Xi_y + \text{Cov}^T\{x, y\}\text{Cov}^{-1}\{x, x\}(y - \Xi_x), \text{Cov}\{y, y\} - \text{Cov}\{x, y\}\text{Cov}^{-1}\{x, x\}\text{Cov}^T\{x, y\}) \tag{38}
\]

**Derivation of \( \tilde{H}(\hat{\theta}_k) \)**

From equation (4), it follows that

\[
\tilde{H}(\hat{\theta}_k) = \frac{\partial h_k(x_k)}{\partial \theta_k} = \begin{bmatrix}
\frac{\partial e_1(x_k)}{\partial \theta_k} \\
\frac{\partial e_2(x_k)}{\partial \theta_k} \\
\vdots \\
\frac{\partial e_{\ell_k}(x_k)}{\partial \theta_k}
\end{bmatrix} \tag{39}
\]

where the \( n \)th row of \( \tilde{H}(\hat{\theta}_k) \) is computed as

\[
\frac{\partial h_k(x_k)}{\partial \theta_k} = \begin{bmatrix}
-h_n(x_k)[1, 0]^T\Sigma^{-1}(x_n - \rho_n) \\
0 \\
-h_n(x_k)[0, 1]^T\Sigma^{-1}(x_n - \rho_n) \\
0 \\
e_n(x_k)
\end{bmatrix} \tag{40}
\]

**Derivation of equation (23)**

Taking derivation with respect to \( x_k[i] \) at both sides of equation (19) yields

\[
0 = [e_n^T(x_k)R_k^{-1}e_n(x_k)]^{-1}\frac{\partial e_n^T(x_k)}{\partial x_k[i]}R_k^{-1}z_k - 2[e_n^T(x_k)R_k^{-1}e_n(x_k)]^{-1}\frac{\partial e_n^T(x_k)}{\partial x_k[i]}R_k^{-1}e_n(x_k)e_n^T(x_k)R_k^{-1}z_k
\]

then, we can obtain that

\[
\frac{\partial e_n^T(x_k)}{\partial x_k[i]}R_k^{-1}z_k = 2[e_n^T(x_k)R_k^{-1}e_n(x_k)]^{-1}
\]

\[
\frac{\partial e_n^T(x_k)}{\partial x_k[i]} - R_k^{-1}e_n(x_k)[e_n^T(x_k)R_k^{-1}z_k]
\]

From equation (22), it follows that
then substituting equation (41) into equation (42), we can obtain that

\[
\frac{\partial^2 L'(x_k)}{\partial x_k[j] \partial x_k[i]} = - \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-1} \frac{\partial e_k^T(x_k)}{\partial x_k[j]} R_k^{-1} e_k(x_k) + 2 \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-2} \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right] \frac{\partial e_k^T(x_k)}{\partial x_k[j]} R_k^{-1} e_k(x_k) - \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-1} \frac{\partial^2 e_k^T(x_k)}{\partial x_k[j] \partial x_k[i]} R_k^{-1} e_k(x_k) + \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-2} \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right] \frac{\partial e_k^T(x_k)}{\partial x_k[j]} R_k^{-1} e_k(x_k) \tag{42}
\]

then substituting equation (41) into equation (42), we can obtain that

\[
\frac{\partial^2 L'(x_k)}{\partial x_k[j] \partial x_k[i]} = \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-2} \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right] \frac{\partial^2 e_k^T(x_k)}{\partial x_k[j] \partial x_k[i]} R_k^{-1} e_k(x_k) + \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-2} \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right] \frac{\partial e_k^T(x_k)}{\partial x_k[j]} R_k^{-1} e_k(x_k) - \left[ e_k^T(x_k) R_k^{-1} e_k(x_k) \right]^{-1} \frac{\partial^2 e_k^T(x_k)}{\partial x_k[j] \partial x_k[i]} R_k^{-1} e_k(x_k) \tag{43}
\]