Laplace $\ell_1$ robust Student’s T-filter for attitude estimation of satellites

Xu Zhang$^1$, Xu Zhang$^2$, He Wang$^2$, Pengyu Guo$^3$ and Bing Xiao$^4$

Abstract
A Laplace $\ell_1$ robust student’s T-filter is presented for satellites to estimate their attitude state despite severe measurement noise and modeling error. Although the student’s T-filter (STF) has the capability of handling measurement noise, it cannot address the unknown modeling error. It is further sensitive to the degree of freedom (DOF). Hence, the measurement covariance is updated by using the maximum correntropy criterion to accommodate the covariance of unknown modeling error in robust filtering design. Moreover, the Laplace distribution is derived to reduce the influence of the DOF parameter by forming a surrogate function of an optimization problem. Then, the majorization minimization approach is formulated to solve such an optimization problem and present the proposed filtering algorithm in the STF framework. Numerical simulation of applying the proposed attitude estimation scheme is performed by comparing it with the third/fifth-order cubature Kalman filters.

Keywords
Student’s T-filter, maximum correntropy criterion, Laplace distribution, attitude estimation, measurement noise, modeling error

Date received: 2 September 2020; accepted: 16 September 2020

Introduction
Nonlinear filters play a vital role in signal processing,$^1$ target tracking,$^2$–$^4$ and navigation.$^5$–$^7$ In the Bayesian estimation framework, filters are designed by solving the posterior probability density function (PDF). However, there does not exist an optimal solution to the posterior PDF for nonlinear systems, since the closed PDF does not exist.$^8$ The posterior PDF is thus assumed to satisfy the Gaussian distribution. Accordingly, many approximation methods are applied to design the suboptimal nonlinear filters such as the classical Kalman filter (KF) and its extensions.$^9$–$^{14}$ For those filters, both the system states and the measurement predictions are assumed to have Gaussian characteristics. Considering space missions such as the orbit tracking or the attitude determination, the noise and the unknown perturbations may be induced. The measurement outliers may also exist.$^{15}$ The noise of the system and its measurement may have a heavier tail than the distribution with the Gaussian characteristics.$^{16}$ Moreover, unknown perturbations bring satellites with modeling error. This error will further increase the nonlinearity and the uncertainty of the noise. This complicates the state estimation further and deteriorates the estimation performance of KF.

The past decade has witnessed significant development of the filter theory in dealing with the heavy-tailed noise. Many investigations are available and they can be categorized into two types. The first is the regression-based robust filters.$^{17}$ The Huber theory$^{18,19}$ is applied to design a robust KF. The maximum correntropy criterion$^{20,21}$ is applied to maximize the prediction error and the residual correntropy with the non-Gaussian noise handled. However, their performance is sensitive to the estimation gains. This drawback drives the researchers to design alternative solutions to the heavy-tailed noise problem.

Some achievements focused on the PDF distributions describing the heavy-tailed noise. For example, the particle filter (PF) adopts a stochastic approximation method to address the non-Gaussian noise. For this filter, random sampling points are adopted to approximate the PDF via the sequence of the Monte Carlo sampling
technique.\textsuperscript{22,23} However, the PF has a computational burden because the considered particles’ number will increase exponentially along with the state’s dimension. The mixture Gaussian KF was derived by assuming that the sum of the Gaussians can approximate any density with arbitrary accuracy.\textsuperscript{24,25} However, it is not easy to determine the number of filters and the corresponding computation burden is large.

An efficient approach to solve the above issue is the application of the Student’s T distribution to approximate the posterior PDF. This scheme is named as the Student’s T-filter (STF), in which the predicted and the posterior PDF are supposed to satisfy the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.\textsuperscript{27–29} To deal with the unknown modeling error, the maximum correntropy criterion is considered here to update the STF’s measurement covariance. This can compensate for the covariance of the unknown modeling error to suppress it for the robust estimation design. Besides, the estimation performance is also dependent on the degree-of-freedom (DOF) of the Student’s T distribution. The selection of DOF for STF estimation design. Besides, the estimation performance is also dependent on the Student’s T distribution.\textsuperscript{26} Therefore, the STF has significant advantages when dealing with the heavy-tailed noise. Many investigations focusing on this method have been presented in recent years.

### Mathematical model of a class of nonlinear systems

The following discrete-time nonlinear system is considered in this study:

\[
x_{k+1} = f(x_k) + w_k
\]

where \( k \in \mathbb{N}_0 \) is the discrete-time index, \( x_k \in \mathbb{R}^n \) is the system state at the discrete-time \( t_k \), \( y_k \in \mathbb{R}^m \) is the system’s measurement at \( t_k \), \( w_k \in \mathbb{R}^p \) is the process noise at \( t_k \), and \( v_k \in \mathbb{R}^q \) is measurement noise at \( t_k \), \( f(x_k) \) and \( h(x_k) \) are known, and they are the process and the measurement functions, respectively.

Assuming that those noises \( w_k \) and \( v_k \) satisfy the heavy-tailed distribution, then the PDFs of those noises are supposed to be the following Student-T distribution:

\[
p(w_k) = \text{St}(w_k; 0, Q_k, v_1)\quad (2)
\]

\[
p(v_k) = \text{St}(v_k; 0, R_k, v_2)\quad (3)
\]

where \( \text{St}(\cdot; \mu, \Sigma, v) \) is the Student’s T PDF with mean \( \mu \), scale matrix \( \Sigma \), and DOF \( v \). Moreover, the specific expression of the Student’s T PDF with any variable \( z \) could be given by

\[
p(z) = \text{St}(z; \mu, \Sigma, v) = \frac{1}{\sqrt{|\Sigma|} \Gamma(v/2)} \left[ 1 + \frac{1}{v} (z - \mu)^T \Sigma^{-1} (z - \mu) \right]^{-\frac{v+
u}{2}} \quad (4)
\]

For \( v > 2 \), the mean and the variance of \( z \) can be calculated by \( E(z) = \mu \) and \( \text{var}(z) = E((z - \mu)(z - \mu)^T) = \frac{v}{v-2} \Sigma \), respectively.

It is known from the above definition that \( Q_k \) and \( v_1 \) are the scale matrix and the process noise’s DOF, respectively. \( R_k \) and \( v_2 \) are the scale matrix and the measurement noise’s DOF, respectively. Then, the initial value of \( x_0 \) can be assumed as the distribution of the Student’s T characteristics with mean \( x_{00} \), scale matrix \( P_{00} \), and DOF \( v_3 \) as

\[
p(x_0) = \text{St}(x_0; x_{00}, P_{00}, v_3) \quad (5)
\]

where \( x_0 \), \( w_k \), and \( v_k \) are mutually uncorrelated.
Student's T filtering theory

When applying the Bayesian theory framework to develop the Student's T-filter, the posterior PDF \( p(x_k|Y_k) \) is calculated via the Bayesian relation

\[
p(x_k|Y_k) = \frac{p(x_k|Y_{k-1})p(y_k|x_k)}{\int_{y_k} p(x_k|Y_{k-1})p(y_k|x_k)dy_k}
\]

\[
= \frac{p(x_k, y_k|Y_{k-1})}{\int_{x_k} p(x_k|Y_{k-1})dx_k}
\]

which represents the measurement update for the estimation.

Based on (6), it is ready to get the predicted PDF \( p(x_k|Y_{k-1}) \) by the Chapman-Kolmogorov equation for the time update of the estimation, which is specified as

\[
p(x_k|Y_{k-1}) = \int_{x_k} p(x_k|x_{k-1})(x_{k-1}|Y_{k-1})dx_{k-1}
\]

(7)

The PDFs \( p(x_k|x_{k-1}) \) and \( p(y_k|x_k) \) are got from the process function and the measurement function in (1). The PDFs of the correlated noises are updated as

\[
p(x_k|x_{k-1}) = \text{St}(x_k; f(x_{k-1}), Q_k, v_k)
\]

(8)

\[
p(y_k|x_k) = \text{St}(y_k; h(x_k), R_k, v_k)
\]

(9)

To obtain a solution to the posterior PDF (7), the jointly predicted PDF \( p(x_k, y_k|Y_{k-1}) \) based on the state \( x_k \) and the measurement \( y_k \) is assumed as the Student’s t-distributed 26.

\[
p(x_k, y_k|Y_{k-1}) = \text{St}
\left[
\begin{bmatrix}
  x_k \\
  y_k
\end{bmatrix};
\begin{bmatrix}
  \hat{x}_{k|k-1} \\
  \hat{y}_{k|k-1}
\end{bmatrix},
\begin{bmatrix}
  P_{k|k-1} & P_{k|k-1}^Y \\
  P_{k|k-1}^Y & P_{k|k-1}^Y
\end{bmatrix},
\begin{bmatrix}
  v_k
\end{bmatrix}
\right]
\]

(10)

where \( \hat{x}_{k|k-1} \) is the mean vector, and \( P_{k|k-1} \) is the scale matrix of the predicted PDF \( p(x_k|Y_{k-1}) \). \( \hat{y}_{k|k-1} \) and \( P_{k|k-1}^Y \) are the mean and the scale matrix of the PDF \( p(y_k|Y_{k-1}) \), respectively. \( P_{k|k-1}^Y \) is the cross-scale matrices of the state system, and \( P_{k|k-1}^Y \) is the cross-scale matrices of the measurement.

To this end, the posterior PDF \( p(x_k|Y_k) \) in (7) can be calculated by the Student’s T distribution. In accordance, the generalized algorithm of STF can be summarized as follows.

1. Time update:

\[
\hat{x}_{k|k-1} = E(x_k|Y_{k-1}) = \\
\int_{x_k} f(x_k|x_{k-1})\text{St}(x_k|x_{k-1}; \hat{x}_{k|k-1}, P_{k|k-1}, v_k)dx_{k-1}
\]

(11)

\[
P_{k|k-1}\times = \frac{v_k}{v_k+2}E((x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T|Y_{k-1})
\]

\[
= \frac{v_k}{v_k+2} \int_{x_k} f(x_k|x_{k-1})(f(x_k|x_{k-1})^T\text{St}(x_k|x_{k-1}; \hat{x}_{k|k-1}, P_{k|k-1}, v_k)dx_{k-1}
\]

(12)

(2) Measurement update:

\[
\hat{y}_{k|k-1} = E(y_k|Y_{k-1}) = \\
\int_{y_k} h(x_k)\text{St}(x_k; \hat{x}_{k|k-1}, P_{k|k-1}, v_k)dx_k
\]

(13)

\[
P_{k|k-1}^Y = \frac{v_k}{v_k+2}E((y_k - \hat{y}_{k|k-1})(y_k - \hat{y}_{k|k-1})^T|Y_{k-1})
\]

\[
= \frac{v_k}{v_k+2} \int_{y_k} h(x_k)(h(x_k))^T\text{St}(x_k; \hat{x}_{k|k-1}, P_{k|k-1}, v_k)dx_k
\]

\[
= \frac{v_k}{v_k+2}\hat{y}_{k|k-1}\hat{y}_{k|k-1}^T + \frac{v_k}{v_k+2}v_kQ_k
\]

(14)

\[
\hat{y}_{k|k-1} = \hat{y}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1})
\]

(15)

\[
K_k = \frac{P_{k|k-1}^Y}{P_{k|k-1}^Y + K_k}
\]

(16)

The generalized STF could be obtained by calculating (11)–(15) with the Student’s T weighted integral, and then by updating the estimated state and the scale matrix (16)–(19). However, the DOF parameters of the STF are difficult to select and they will have an important effect on the estimation accuracy. This issue will be addressed in the next section.

Development of the Laplace \( \ell_1 \) robust Student’s T-filter

Maximum correntropy criterion design

In this section, the maximum correntropy criterion is adopted to update the measurement covariance of the unknown modeling error for the robust estimation scheme design. Therefore, the maximum correntropy criterion is introduced first. For any random variables \( G_1 \) and \( G_2 \), their correntropy is denoted as 31

\[
V(G_1, G_2) = E(\kappa(G_1, G_2)) = \int \kappa(g_1, g_2)df_{G_1, G_2}(g_1, g_2)
\]

(20)

where \( \kappa(\cdot) \) denotes the kernel function and it is positive-definite. The function \( f_{G_1, G_2}(g_1, g_2) \) denotes the density of the joint probability.

In practical engineering, the joint probability cannot be obtained directly. It is usually represented by a series
of data samples \( \{ (g_1, g_2), i = 1, 2, \ldots, N \} \) with finite numbers. Then, one can write the correntropy as

\[
V(G_1, G_2) = \frac{1}{N} \sum_{i=1}^{N} \kappa(g_i, g_2)
\]

(21)

According to Wang et al., the following Gaussian kernel function can be employed for (19).

\[
\kappa_G(\epsilon_i) = \kappa_G(g_i - g_2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{\epsilon_i^2}{2\sigma^2}\right)
\]

(22)

where \( \sigma \) is the correntropy’s kernel width. It is found that the kernel \( \kappa_G(\epsilon_i) \) will have its maximum value when \( \epsilon_i = 0 \), which can be used as the cost function of the maximum correntropy criterion.

\[
J_{MCC} = \min \sum_{i=1}^{N} (\kappa_G(\epsilon_i) - \kappa_G(\epsilon_i))
\]

(23)

Considering that the investigation focuses on updating the measurement covariance to suppress the unknown modeling error, the new cost function is proposed by combining (23) and the STF framework with the definition of \( \epsilon_i = (\frac{u}{v_{0j}} - R_j) \) as \( (y_k - H_jx_k) \) as

\[
\hat{x}_{k|k} = \arg \min \left( \frac{1}{2} ||x_k - \hat{x}_{k|k-1}||^2 (\frac{u}{v_{0j}} - P_{i0j})^{-1} + \frac{1}{2} \sum \kappa_G(0) - \kappa_G(\epsilon_i)) \right)
\]

(24)

with \( \frac{1}{2} ||x_k - \hat{x}_{k|k-1}||^2 (\frac{u}{v_{0j}} - P_{i0j})^{-1} = (x_k - \hat{x}_{k|k-1})^T (\frac{u}{v_{0j}} - 2) \).

\( P_{k|k-1}^{-1} (x_k - \hat{x}_{k|k|1}) \). \( \hat{x}_{k|k-1} \) is the system state estimation after the measurement update. \( \hat{x}_{k|k-1} \) is system state estimation before measurement update. \( P_{k|k-1} \) is the likelihood of the scalar matrix of the PDF \( p(x_k|Y_{k-1}) \). \( H_j \) is the Jacobian matrix of the nonlinear function \( h(\cdot) \) for the following theoretical derivation.

Differentiating (24) with respect to \( x_k \) yields

\[
(\frac{u}{v_{0j}} - P_{k|k-1})^{-1}(x_k - \hat{x}_{k|k-1}) - \sum_{i=1}^{m} \frac{\partial \kappa_G(\epsilon_i)}{\partial \epsilon_i} \frac{\partial \epsilon_i}{\partial x_k} = 0
\]

(25)

Inserting (22) into (23), one has

\[
(\frac{u}{v_{0j}} - P_{k|k-1})^{-1}(x_k - \hat{x}_{k|k-1}) - \sum_{i=1}^{m} \psi_i(x_k, \epsilon_i) \frac{\partial \epsilon_i}{\partial x_k} = 0
\]

(26)

where \( \psi_i = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{\epsilon_i^2}{2\sigma^2}\right) \), \( i = 1, 2, \ldots, m \).

Define the matrix \( \Psi_k \) as

\[
\Psi_k = \text{diag}((\psi_{k,1} \psi_{k,2} \ldots \psi_{k,m})^T),
\]

(25) can be rewritten as

\[
(\frac{u}{v_{0j}} - P_{k|k-1})^{-1}(x_k - \hat{x}_{k|k-1}) - \Psi_k \frac{\partial \epsilon_i}{\partial x_k} = 0
\]

(27)

The maximum posterior estimation of \( x_k \) is then obtained by solving the following cost function, which is derived and transformed from (27).

\[
\hat{x}_{k|k} = \arg \min \left( \frac{1}{2} ||x_k - \hat{x}_{k|k-1}||^2 (\frac{u}{v_{0j}} - P_{i0j})^{-1} + \frac{1}{2} ||y_k - H_jx_k||^2 (\frac{u}{v_{0j}} - R_j)^{-1} \right)
\]

(28)

where \( R_k = R_k^T \psi_k (R_k^T)^T \). The modified cost function can be treated as the integration of the STF and the maximum correntropy criterion with the updated covariance. This can update the covariance according to the influence of unknown modeling error for robustness design.

**The Laplace \( \ell_1 \) distribution design**

Based on the above result, the Laplace \( \ell_1 \) distribution is introduced to formulate the proposed \( L1 \)-RSTF. Firstly, the normalized residual vector is denoted as \( \hat{x}_k = (\frac{u}{v_{0j}} - R_j)^{-1}(y_k - H_jx_k) = [\hat{x}_{k,1} \hat{x}_{k,2} \cdots \hat{x}_{k,m}]^T \). Then, the new cost function (28) can be simplified by the Laplace \( \ell_1 \) distribution as

\[
\hat{x}_{k|k} = \arg \min \left( \frac{1}{2} ||x_k - \hat{x}_{k|k-1}||^2 (\frac{u}{v_{0j}} - P_{i0j})^{-1} + \sqrt{2}||\hat{x}_k||_1 \right)
\]

(29)

Because the differentiation of the \( \ell_1 \) norm function in (29) does not exist at \( \hat{x}_k = 0 \), it fails to apply the traditional gradient techniques to solve (29) directly.

To solve this above difficulty, an iterative surrogate function is to be presented for re-describing the cost function (29). The majorization minimization (MM) method will be designed further to solve this issue by calculating the minimum value of the surrogate function. Consider a positive function \( g(\cdot) \) and a positive scalar \( a \), it follows that

\[
\sqrt{g(x)} = \frac{a}{2} + \frac{g(x)}{2a}
\]

(30)

Moreover, \( \sqrt{g(x)} = \frac{a}{2} + \frac{g(x)}{2a} \) holds only when \( g(x) = a^2 \). Then, one has

\[
||\hat{x}_k||_1 = \sum_{i=1}^{m} ||\hat{x}_{k,i}|| \leq \sum_{i=1}^{m} \left( \frac{x_i}{2} + \frac{x_i^2}{2x_i} \right),
\]

(31)

\[\forall x_i > 0, i = 1, \ldots, m\]

Substituting (31) into (29), it leaves the surrogate function as
\[
\begin{aligned}
\hat{x}_{k|k} &= \arg \min \mathbf{J}(x_k, \Xi) \\
\mathbf{J}(x_k, \Xi) &= \frac{1}{2}||x_k - \hat{x}_{k|k-1}||_{P_{k|k-1}}^{-1} + \frac{1}{2}||y_k - H_k x_k||_{\frac{N}{v_y^2 + \nu_k}}^{-1} + \frac{N}{2} \sum_{i=1}^{m} \xi_i, \quad (32)
\end{aligned}
\]

where
\[
\hat{R}_k = \frac{\nu_k}{N} \hat{R}_k \Xi (\hat{R}_k)^T
\]
and
\[
\Xi = \text{diag}(\xi_1, \xi_2, \ldots, \xi_m)^T.
\]

Now, the original optimization problem has been changed into solving the surrogate function (32) via the MM technique. It is described as
\[
\hat{x}^{(e+1)}_{k|k} = \arg \min_{x_k} \mathbf{J}(x_k, \hat{\xi}^{(e+1)}_{k|k})
\]
\[
\hat{\xi}^{(e+1)}_{k|k} = \arg \min_{\xi_k} \mathbf{J}(\hat{x}^{(e+1)}_{k|k}, \hat{\xi}^{(e+1)}_{k|k}) \quad (33)
\]
Aiming of applying the equivalent condition (31), the solution of (34) could be approximated by
\[
\hat{\xi}^{(e+1)}_{k|k} = \left( \left( \frac{\nu_k}{\nu_y^2} \hat{R}_k \right)^{-1} (y_k - H_k \hat{x}^{(e)}_{k|k}) \right) + \epsilon_i
\]
where \(\epsilon_i\) is a very small non-zero constant (e.g. \(10^{-6}\) to \(10^{-8}\)) to avoid \(\xi^{(i)} = 0\) with \(c\) representing the iterations’ number. Generally, a starting weighting matrix \(\Xi^{(0)}\) could be selected first to obtain \(\hat{x}^{(1)}_{k|k}\) through (33). Then, the estimation \(\hat{x}^{(1)}_{k|k}\) could be used to update and obtain a new weighting matrix \(\Xi^{(1)}\). In this work, the identity matrix \(I\) is treated as the initial weighting matrix for simplification.

Note that the surrogate function (32) has some similarities as the cost function of STF for a given \(\Xi^{(c)}\). Hence, the function (32) is equivalent to the following form
\[
\hat{x}^{(c)}_{k|k} = \arg \min_{x_k} \mathbf{J}(x_k, \hat{\xi}^{(c)}_{k|k})
\]
\[
\frac{1}{2}||x_k - \hat{x}_{k|k-1}||_{\frac{\nu_k}{\nu_y^2} \hat{R}_k}^{-1} + \frac{1}{2}||y_k - H_k x_k||_{\frac{N}{v_y^2 + \nu_k}}^{-1}
\]

It obtains the upper bound of this cost function in the STF framework. Once getting the estimation state \(\hat{x}^{(c)}_{k|k}\), one can update and iterate a new weighting matrix \(\Xi^{(c+1)}\) for obtaining \(\hat{x}^{(c+1)}_{k|k}\).

### The Laplace \(\ell_1\) robust Student’s T-filter

Based on section “The Laplace \(\ell_1\) distribution design,” we are ready to present the L1-RSTF in the following algorithm. The proposed L1-RSTF has a similar structure to the STF in the measurement updating. The major difference is in the state updating. Therefore, the new process of state updating for proposed L1-RSTF is described as follows.

1. **State updating for L1-RSTF:**

The initial value of the iteration and the scalar matrix for the measurement covariance is given by
\[
\Xi^{(0)} = I, c = 0 \quad (37)
\]
\[
P_{k|k-1} = \frac{v_y}{v_y^2 - \nu_k} P_{k|k-1} + \frac{v_y(v_y - 2)}{v_y^2 - 2 \nu_k} \hat{R}_k
\]
\[
\hat{R}_k = \frac{\nu_k}{N} \hat{R}_k \Xi (\hat{R}_k)^T
\]
\[
K^{(c)}_{k|k-1} = \frac{\nu_k}{N} \hat{R}_k \Xi (\hat{R}_k)^T
\]
\[
\hat{x}^{(c)}_{k|k} = \hat{x}^{(c)}_{k|k-1} + K^{(c)}_{k|k-1}(y_k - \hat{y}^{(c)}_{k|k-1})
\]
\[
\hat{\xi}^{(c+1)}_{k|k} = \text{diag}(\hat{\xi}^{(c+1)}_{1} , \hat{\xi}^{(c+1)}_{2}, \ldots, \hat{\xi}^{(c+1)}_{m})^T
\]
\[
c = c + 1
\]

Until the desired convergence performance index is satisfied.

Up to now, the final estimation for the states and their scalar matrix of the covariance can be determined as
\[
\hat{x}_{k+1|k} = \hat{x}^{(c)}_{k+1|k+1} \quad (50)
\]
\[
P_{k+1|k} = \frac{v_y}{v_y^2 - \nu_k} (P_{k+1|k} - K^{(c)}_{k+1} P_{k+1|k} K^{(c)}_{k+1})
\]

It can be known that the proposed L1-RSTF is similar to the standard STF via the Laplace \(\ell_1\) theory. However, it is difficult to choose the convergence criterion. Hence, one can determine the number of iterations according to the precision requirements.

### Numerical simulation

In this section, the application of the L1-RSTF developed in section “Development of the Laplace \(\ell_1\) robust Student’s T-filter” to achieve attitude estimation for small satellites will be done.
Mathematical model of a rigid satellite

The unit quaternion $\mathbf{q} = [q_1 \; q_2 \; q_3 \; q_4]^T = [q_1 \; q_2 \; q_3 \; q_4]^T$, $\mathbf{q}_{24} = [q_2 \; q_3 \; q_4 \; q_1]^T$, is adopted to represent the attitude with respect to the inertial frame. Then, the kinematics of a rigid satellite is given by

$$\dot{\mathbf{q}} = \frac{1}{2} \Phi(\mathbf{q}) \omega$$  \hspace{1cm} (52)

where $\omega = [\omega_1 \; \omega_2 \; \omega_3]^T$ is the angular velocity in the body frame with respect to the inertial frame. $\Phi(\mathbf{q})$ is given by

$$\Phi(\mathbf{q}) = \begin{bmatrix} -q_1^3 & q_2 q_3 & q_2 q_4 & q_3 q_4 \\ q_2 q_3 & -q_1 q_4 & q_1 q_2 & -q_3 q_4 \\ q_2 q_4 & q_1 q_3 & -q_1 q_2 & q_3 q_4 \\ q_3 q_4 & -q_2 q_3 & q_2 q_4 & -q_1 q_3 \end{bmatrix}$$ \hspace{1cm} (53)

For small satellites, their attitude dynamic model is established as

$$\dot{\omega} = J^{-1}(\tau - \omega \times J \omega + d)$$ \hspace{1cm} (54)

where $\tau \in \mathbb{R}^3$ is the total torque and $J \in \mathbb{R}^{3 \times 3}$ is the inertia of the satellite. Based on (50), the attitude dynamics without any gyro measurement could be written as:

$$\dot{x} = \begin{bmatrix} 0.5 \Phi(\mathbf{q}) \omega^T J^{-1} (\tau - \omega \times J \omega) + J^{-1} d \end{bmatrix} + w$$ \hspace{1cm} (55)

where $w$ denotes the heavy-tailed noise and $x = [\mathbf{q}^T \; \omega^T]^T$.

Moreover, considering cost limitation, a three-axis magnetometer and a sun sensor are fixed to measure the states of the attitude control system. For these sensors, the measurement model could be given by

$$y = h(X) = \begin{bmatrix} C_{11}(\mathbf{q}) S_1 \; + \; \Delta \nu_1 \\ C_{12}(\mathbf{q}) B_2 \; + \; \Delta \nu_2 \end{bmatrix}$$ \hspace{1cm} (56)

where $C_{11} \in \mathbb{R}^{3 \times 3}$ represents the transition matrix from the inertial frame to the body-fixed frame, and it describes the attitude information. $S_1 \in \mathbb{R}^3$ and $B_2 \in \mathbb{R}^3$ are the sun vector and the magnetic field, respectively. $\Delta \nu_1 \in \mathbb{R}^3$ and $\Delta \nu_2 \in \mathbb{R}^3$ are the heavy-tailed measurement noise of the sensors, respectively.

To this end, the satellite’s attitude control system described by (51) and (52) can be written into the form of (1) after discretization. Consequently, the proposed L1-RSTF can be applied to estimate the attitude state with $\dot{x} = [\mathbf{q}^T \; \omega^T]^T$ denoting the estimation of the state $x$.

Simulation results

To validate the performance of the proposed L1-RSTF, simulations are conducted by comparing it with the CKF. In the simulation, the initial states of the satellites are $\mathbf{q}(0) = [1 \; 0 \; 0 \; 0]^T$ and $\omega(0) = [0 \; 0 \; 0]^T$ deg/s. The states’ initial estimation is set to be $\hat{\mathbf{q}}(0) = \sqrt{\frac{3}{4}}$, $\hat{\omega}(0) = \sqrt{\frac{3}{4}} [\sqrt{3} \cos(0.5 \theta) \sin(0.5 \theta) \sin(0.5 \theta) \sin(0.5 \theta) \sin(0.5 \theta) \sin(0.5 \theta)]^T$, $\theta = \frac{\pi}{2}$, and $\omega(0) = [0 \; 0 \; 0]^T$ deg/s. The inertia of the satellite has a value of $J = \text{diag}(49.96 \; 55.40 \; 63.0)^T$ kg m$^2$. The processing and the measurement noises are chosen as

$$w = \begin{bmatrix} \chi(0, Q) \; w.p. \; p_1 \\ \chi(0, 100Q) \; w.p. \; 1 - p_1 \end{bmatrix}$$

$$v = \begin{bmatrix} \chi(0, R) \; w.p. \; p_2 \\ \chi(0, 100R) \; w.p. \; 1 - p_2 \end{bmatrix}$$

where $Q$ represent the covariance of the process noise with $p_1 = 0.9$. $R$ is the measurement noise’s covariance with $p_2 = 0.9$. The measurement accuracy of the magnetometer and the sun sensor is 50 nT and 0.2 degrees, respectively. Their covariances are represented by $R_B$ and $R_S$, respectively. Based on Cao et al. \textsuperscript{33} the 13th-order Gaussian sphere harmonic function of IGRF is considered to simulate the real $y_B$. The 7th-order of IGRF is applied to simulate $B_2$ in the satellite system. The measurement error between the 13th-order and 7th-order of the IGRF is about 208 nT. This error is also considered. In the simulation, the following two scenarios are considered.

Remark: Let $\hat{x} = [\mathbf{q}^T \; \omega^T]$, $\hat{\mathbf{q}}$ denote the state estimation error, where $\mathbf{q} = q - \hat{q}$ and $\omega_{\text{error}} = [W_{x_{\text{error}}} \; W_{\omega_{\text{error}}} \; W_{\omega_{\text{error}}}]^T = \omega - \hat{\omega}$. In satellite engineering, the unit quaternion $\mathbf{q}$ can be mathematically transformed into the Euler attitude angles to explicitly describe the satellite’s attitude, that is,

$$\Theta = [\phi \; \theta \; \psi]^T = f(x)$$ \hspace{1cm} (58)

where $f(\cdot)$ is the transformation function. $\phi$, $\theta$, and $\psi$ are the roll, the pitch, and the yaw attitude Euler angles with respect to $\mathbf{q}$, respectively. Hence, the attitude estimation error $\hat{\mathbf{q}}$ can be transformed into the attitude Euler angles estimation error denoted as $\Theta_{\text{error}} = [\theta_{\text{error}} \; \theta_{\text{error}} \; \psi_{\text{error}}]^T$.

1) The first scenario: This scenario is to validate the developed filter’s effectiveness in handling the heavy-tailed noise. Figures 1 and 2 show the resulted estimation errors. Because the CKF is widely used in practice, the 3rd-CKF and the 5th-CKF are also employed to compare with the proposed L1-RSTF. It is found in Figures 1 and 2 that the convergence rate of L1-RSTF is much faster and more stable than the CKFs without the obvious fluctuations. Although the advantage of L1-RSTF is not very obvious in the angular velocity estimation, this method still has a certain precision superiority. Moreover, there are a few estimated mutations for the L1-RSTF, which is mainly due to the mutation of heavy-tailed distribution with a small probability. However, due to the theoretical advantages of proposed L1-RSTF, the estimation error divergence can be rapidly suppressed and return to the normal estimated trajectory.
Table 1 lists the RMSE results of the estimation error between the 3rd-CKF, the 5th-CKF, and the L1-RSTF. It is listed that the developed L1-RSTF has higher estimation precision than the 3rd-CKF and the 5th-CKF, which is especially true in attitude estimation. In the aspect of angular velocity, the theoretical
The advantage of L1-RSTF is not obvious, but it still has the highest accuracy. Regarding the computational burden, the 3rd-CKF and the 5th-CKF need 12 and 73 sampling points to calculate and propagate the state estimation, respectively. However, the proposed L1-RSTF only needs 12 sampling points and the number of iterations is five. Therefore, the L1-RSTF also has obvious advantages in computational complexity.

(2) The second scenario: The attitude dynamics of small satellites are characterized by modeling error. Hence, the modeling error in the attitude dynamics is considered. It is represented by

\[ d = \Delta N_e + \Delta N_c, \]

where \( \Delta N_c \) and \( \Delta N_e \) are assumed as

\[ N_c = \begin{bmatrix} 0.0001 & 0.0002 & 0.0004 \end{bmatrix}^{\top} \text{Nm} \]

\[ \Delta N_c = \begin{bmatrix} 0.0003 \cos(10\omega_c t) & 0.0002 \cos(10\omega_c t) & -0.0001 \cos(10\omega_c t) \end{bmatrix}^{\top} \text{Nm} \]

Therefore, the modeling errors and the heavy-tailed noise (57) are both taken into consideration in this case and introduced into the simulation system, which can be used to test the robust ability of presented L1-RSTF to the modeling error and the noise.

Figures 3 and 4 show that the estimation accuracy of the second scenario is inferior to the accuracy in the first scenario. This coincides with the preceding theoretical result. Especially for the CKFs, the estimated convergence curve deviates seriously from the zero-value due to the existence of model error in Figure 3. However, the L1-RSTF still fluctuates slightly around zero and maintains a high precision compared with the CKFs. The angular velocity estimate accuracy of L1-RSTF also has better convergence properties than the CKFs, as we can see in Figure 4.

Moreover, the comparison results listed in Table 2 show that the estimation accuracy obtained from the 3rd-CKF and the 5th-CKF is severely decreased due to the modeling error. The reason is that the nonlinearity and the uncertainty of the satellite system are increased by modeling error and noise. Only the proposed L1-RSTF maintains acceptable high precision. This validates the superiority of L1-RSTF.

The above simulation results verify that the presented scheme is robust to the heavy-tailed noise and the modeling error with low computation required. For the CKFs, only the angular velocity estimation can get acceptable performance. It is concluded that the proposed scheme is less dependent on the model of the uncertainty and the noise. This lets the designed filter be superior to the traditional Kalman filter because the latter relies heavily on the known noise level. Therefore, it can be anticipated that this method has important application value and will be more widely used in the future.

**Conclusion**

This work presented a Laplace \( \ell_1 \) robust Studen’s T-filter to solve the attitude estimation problem of small satellites subject to heavy-tailed noise and modeling errors. The maximum correntropy criterion and the Laplace distribution were used to model the measurement process. The majorization minimization technique was applied to solve the maximum posterior estimation...
via an iterative algorithm. Numerical simulation with the application of the proposed L1-RSTF was conducted by comparing it with the 3rd-CKF and the 5th-CKF. The effectiveness of the proposed approach in both the accuracy of system estimation and the system robustness with less calculation burden was validated. As one of future work, using the estimated states provided by the proposed filtering approach, the attitude controller design should be carried out to ensure high-accuracy attitude control. This could be done by invoking the robust methods reported in.35–38

**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) received no financial support for the research, authorship, and/or publication of this article.

**ORCID iD**

Bing Xiao https://orcid.org/0000-0003-2389-0707

**References**

1. Haykin S and Arasaratnam I. *Adaptive signal processing*. New York: John Wiley & Sons, 2010.
2. Linares R, Jah MK, Crassidis JL, et al. Space object shape characterization and tracking using light curve and angles data. *J Guid Control Dyn* 2014; 37(1): 13–25.
3. Graham MC, How JP and Gustafson DE. Robust state estimation with sparse outliers. *J Guid Control Dyn* 2015; 38(7): 1229–1240.
4. Wang YD, Zheng W, Sun SM, et al. Robust information filter based on maximum correntropy criterion. *J Guid Control Dyn* 2016; 39(5): 1126–1131.
5. Cao L, Chen YQ, Zhang ZD, et al. Predictive smooth variable structure filter for attitude synchronization estimation during satellite formation flying. *IEEE Trans Aerosp Electron Syst* 2017; 53(3): 1375–1383.
6. Cao L, Ran D, Chen XQ, et al. Huber second-order variable structure predictive filter for satellites attitude estimation. *Int Control Autom Syst* 2019; 17(7): 1781–1792.
7. Cao L and Li HN. Unscented predictive variable structure filter for satellite attitude estimation with model errors when using low precision sensors. *Acta Astronaut* 2016; 127: 505–513.
8. Anderson BDO and Moore JB. *Optimal filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

---

**Figure 4.** The angular velocity estimation error in the second scenario.

**Table 2.** The RSME comparison in the second scenario.

| RSME     | φ(degrees) | θ(degrees) | ψ(degrees) | ωx(rad/s) | ωy(rad/s) | ωz(rad/s) |
|----------|------------|------------|------------|-----------|-----------|-----------|
| 3rd-CKF  | 0.0527     | 0.1081     | 0.0748     | 0.00120   | 0.000924  | 0.001100  |
| 5th-CKF  | 0.0475     | 0.0973     | 0.0674     | 0.00100   | 0.000832  | 0.000991  |
| L1-RSTF  | 0.0074     | 0.0101     | 0.0103     | 0.000613  | 0.000467  | 0.000608  |
9. Cao L, Dong Q and Chen XO. Laplace L1 Huber based cubature Kalman filter for attitude estimation of small satellite. Acta Astronaut 2018; 148: 48–56.

10. Norgaard M, Poulsen NK and Ravn O. New developments in state estimation for nonlinear systems. Automatica 2000; 36(11): 1627–1638.

11. Arasaratnam I and Haykin S. Cubature Kalman filter. IEEE Trans Automat Contr 2009; 54(6): 1254–1269.

12. Cao L, Dong Q, Han L, et al. Strong tracking sigma point predictive variable structure filter for attitude synchronization estimation. J Navig 2018; 71(3): 607–624.

13. Jia B, Xin M and Cheng Y. Sparse Gauss-Hermite quadrature filter with application to spacecraft attitude estimation. J Guid Control Dyn 2011; 34(2): 367–379.

14. Jia B, Xin M and Cheng Y. Anisotropic sparse Gauss-Hermite quadrature filter. J Guid Control Dyn 2012; 35(3): 1014–1023.

15. Huang Y, Zhang Y, Li N, et al. A robust Gaussian approximate fixed-interval smoother for nonlinear systems with heavy-tailed process and measurement noises. IEEE Signal Process Lett 2016; 23(4): 468–472.

16. Wang HW, Li HB, Zhang W, et al. Laplace l1 robust Kalman filter based on majorization minimization. In: The 20th International conference on information fusion, Xi’an, China, 10–13 July 2017, pp.10–13. New York: IEEE.

17. Huber PJ. Robust statistics. New York: Wiley, 2004.

18. Gandhi MA and Mili L. Robust Kalman filter based on a generalized maximum likelihood-type estimator. IEEE Trans Signal Process 2010; 58(5): 2509–2520.

19. Wu H, Chen S, Yang B, et al. Robust derivative-free cubature Kalman filter for bearings-only tracking. J Guid Control Dyn 2016; 39(8): 1866–1871.

20. Chen C, Huang J, He L, et al. Fast iteratively reweighted least squares algorithm for analysis-based sparsity reconstruction. Medical Image Analysis 2018; 49: 141–152.

21. Principe JC. Information theoretic learning-Renyi’s entropy and kernel perspectives. New York: Springer, 2010, pp.385–455.

22. Wang YD, Zheng W, Sun SM, et al. Robust information filter based on maximum correntropy criterion. J Guid Control Dyn 2016; 39(5): 1126–1131.

23. Cao L, Chen XQ and Sheng T. An algorithm for high precision attitude determination when using low precision sensors. Sci China Inf Sci 2012; 55(3): 626–637.

24. Sidi MJ. Spacecraft dynamics and control: a practical engineering approach. New York: Cambridge University Press, 1997.

25. Xiao B and Yin S. Large-angle velocity-free attitude tracking control of satellites: an intelligent observer-free framework. IEEE Trans Cybern. Epub ahead of print 2019. DOI: 10.1109/TCYB.2019.2945844.

26. Chen CC, Qian CJ, Sun ZY, et al. Global output feedback stabilization of a class of nonlinear systems with unknown measurement sensitivity. IEEE Trans Automat Contr 2018; 63(7): 2212–2217.

27. Xiao B, Cao L, Xu SY, et al. Robust tracking control of robot manipulators with actuator faults and joint velocity measurement uncertainty. IEEE ASME Trans Mechatron 2020; 25(3): 1354–1365.

28. Cao L, Xiao B and Golestani M. Robust fixed-time attitude stabilization control of flexible spacecraft with actuator uncertainty. Nonlinear Dyn 2020; 100(3): 2505–2519.