Student's t Information Filter with Adaptive Degree of Freedom for Multi-Sensor Fusion

Joelle Al Hage, Philippe Xu, Philippe Bonnifait

Student's t Information Filter with Adaptive Degree of Freedom for Multi-Sensor Fusion

Joelle Al Hage, Philippe Xu and Philippe Bonnifait

***Abstract*— Safety-critical applications such as autonomous driving require a high-integrity localization system that bounds the errors of the estimation process. In this paper, the classical Kalman filter used for multi-sensor data fusion, which is unable to consistently bound estimation errors with a low probability risk, is replaced by a Student's *t* filter. The degree of freedom of the *t* distribution offers a way of shaping the heavy tail of the distribution that makes the estimation process more robust in the presence of non-detectable bias and results in a more consistent confidence interval computation. We make use of the heavy-tailed property of the *t* distribution by introducing a novel real-time adaptive computation of the degree of freedom. The filtering process is formalized through an informational form, since this makes it easier to include a fault detection and exclusion step where a bank of filters is generated. The performance of the proposed approach is evaluated through a localization problem using data acquired from an experimental vehicle equipped with multiple sensors: a GNSS receiver, wheel-speed sensors, a yaw rate gyro and a smart camera that can detect several lane markings, together with high-definition maps.**

I. Introduction

The social acceptability of autonomous vehicles depends on their ability to operate seamlessly in today's road networks. Accurate and reliable localization is therefore vital. Localizing autonomous vehicles generally relies on sensors such as GNSS (Global Navigation Satellite System) receivers, perception sensors (camera, LiDAR) and dead-reckoning sensors. It also makes use of High Definition (HD) road maps [1], [2] used to perform a map matching by associating the estimated position (obtained via GNSS or dead-reckoning) with the map [3], [4]. Additional information from exteroceptive sensors such as cameras can be used to detect lane markings and to improve positioning.

One of the most commonly used filters for combining data from different sensors is the Kalman Filter (KF). This filter is optimal in terms of minimum mean square error for a linear model with Gaussian noise. However, for autonomous vehicles, the assumption of a Gaussian distribution is often not justified, especially in urban environments where the GNSS signals are affected by multi-path and non-line of sight. Large errors with small probabilities can appear in these environments, generating heavy-tailed distributions.

In this work, we are looking for a state estimator that provides consistency in terms of confidence intervals. Consistency is here a more important criterion than optimality. The Student's t Filter (StF) can be seen as a modification of KF for managing process and measurement noise that has a heavy-tailed distribution. This filter is known to be more robust to large errors and outliers [5], [6]. In this paper we develop an informational form of the StF to take advantage of its update step. This form is better suited to multi-sensor data fusion and to decentralized architectures. The development of the Student's t Information Filter (StIF) is done in a similar way to the development of the Information Filter (IF) [7]. Although the Student's t filter is quite robust to large errors, they still affect the localization procedure. For this reason we add a Fault Detection and Exclusion stage (FDE) to exclude erroneous measurements that have a direct effect on the estimation procedure. The Student's t distribution is characterized by a degree of freedom (dof) that shapes the tail of the distribution. It directly affects the performance of the filter and the computed confidence interval. In order to avoid setting a fixed degree of freedom, which is difficult to do when the environment changes, the residual used in the FDE stage is used to adapt the dof of the StIF in real-time. The literature has little to say about the determination of the appropriate dof of a Student's t filter, apart from the work done in [8], where the authors propose using a multiple model adaptive estimation.

We apply our proposed robust StIF with adaptive degree of freedom to the localization of an autonomous vehicle. Dead-reckoning measurements are merged with GNSS measurements (GPS and GLONASS) obtained from a low-cost mono-frequency receiver and with measurements from a smart camera. The camera returns measurements of detected lane markings that are used to enhance the localization procedure after solving the data association with the HD map.

A second step involves computing a confidence interval that bounds the position errors with a given risk. Compared to a classical Gaussian assumption, the Student's t distribution is well adapted to this kind of computation. The heavy tail of the t distribution takes into account errors that have a small probability. Since we are dealing with autonomous vehicles, it is interesting to express the confidence interval in the along-track (AT) and cross-track (CT) directions. This shows the contribution made by the camera in reducing the confidence interval in the CT direction.

This paper is organized as follows. After an introduction on Student's t distributions, section 2 presents the main contributions of the research, namely the Student's t information filter (StIF) and the robust StIF with adaptive degree of freedom. In this section the FDE stage is also described

The authors are with Heudiasyc laboratory, UMR 7253, Université de Technologie de Compiègne, Sorbonne Universités, 60200 Compiègne, France. The authors are also members of the SIVALab, joint laboratory between Renault/Heudiasyc/CNRS/UTC

in detail. Section 3 presents the localization of a self-driving car using the proposed robust filter. The computation of confidence intervals applied to the localization is also presented in this section. Section 4 reports an experimental study and the performance of the method on a recorded dataset in Rambouillet, France. Our conclusion and prospects for future work are given in section 5.

II. Multi-sensor data fusion using the informational form of Student's t filter

A. Student's t distribution

A random variable X is said to follow a multivariate t distribution with mean vector μ, scale matrix P and degree of freedom (dof) ν, if it can be written in the form [9]:

$$X = \mu + \omega^{-\frac{1}{2}} y \tag{1}$$

where ω is distributed according to a Gamma distribution $\Gamma\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ and y has a zero-mean Gaussian distribution with covariance P. It is denoted as $X \sim St(\mu, P, \nu)$.

The degree of freedom of the Student's t distribution has a direct effect on the tails of the distribution and if $\nu \to \infty$ the t distribution converges to a Gaussian distribution. The covariance of the t distribution is defined only if $\nu > 2$, and in this case it can be written as $\Sigma = \frac{\nu}{\nu-2} P$. The Student's t distribution can also be viewed as a mixture of normal density [10].

The Student's t distribution shares some properties with the Gaussian distribution [9]. For example, for any a and any non-singular matrix B, the random variable $z = BX + a$ follows a t distribution in the form: $St(B\mu + a, BPB^T, \nu)$. This property is used in the derivation of Student's t Kalman Filter (StF).

B. System modeling

Consider the following evolution and observation models used in a multi-sensor data fusion problem:

$$X_{k+1} = f(X_k, u_k) + v_k \tag{2}$$

$$Z_k = h(X_k) + w_k \tag{3}$$

X is the state vector with scale matrix P

u_k is the input vector (for instance speed or acceleration)

Z is the observation vector (typically exteroceptive sensors measurements)

v_k and w_k are the process and measurements noises.

Suppose that the initial state and the noises follow Student's t distributions [11], [12], [6]:

$$p(X_0) = St(X_0; \hat{X}_0, P_0, \nu_0) \tag{4}$$

$$p(v_k) = St(v_k; 0, Q_k, \gamma_k) \tag{5}$$

$$p(w_k) = St(w_k; 0, R_k, \delta_k) \tag{6}$$

where $\hat{X}_0$, P_0 and ν_0 correspond to the initial guess. A scale matrix Q_u is associated with the input vector u_k to manage the input noise.

C. Student's t Filter

Similarly to the Kalman Filter, the StF includes two steps: time update and measurement update.

At time k, suppose that $p(X_k|Z_{1:k})$ follows a t distribution $St(X_k, \hat{X}_{k/k}, P_{k/k}, \nu_k)$ and suppose that $p(v_k)$ follows equation 5. In order to derive the StF, these two distributions have to share a common degree of freedom denoted ν'_k [11], [12]:

$$p(X_k, v_k|Z_{1/k}) = St\left(\begin{bmatrix} X_k \\ v_k \end{bmatrix}; \begin{bmatrix} \hat{X}_{k/k} \\ 0 \end{bmatrix}, \begin{bmatrix} P'_{k/k} & 0 \\ 0 & Q'_k \end{bmatrix}, \nu'_k\right) \tag{7}$$

X_k and v_k are assumed to be uncorrelated.

In order to preserve the heaviest tails between X_k and v_k, ν'_k can be chosen to be the minimum value between (ν_k, γ_k) as proposed in [11]. Then, $P'_{k/k}$ and Q'_k are the adjusted values of $P_{k/k}$ and Q_k given the common degree of freedom ν'_k. Different methods can be used to adjust the matrices, such as moment matching or minimizing the Kullback-Leibler divergence [11].

Suppose that the uncertainty propagation is well modeled by a first-order Taylor approximation. Then, since linear transformations preserve the degree of freedom of a Student's t distribution (see section II-A), the predicted density can be approximated as

$$p(X_{k+1}|Z_{1:k}) = St(X_{k+1;}\hat{X}_{k+1/k}, P_{k+1/k}, \nu'_k) \tag{8}$$

The predicted estimators $\hat{X}_{k+1/k}$ and $P_{k+1/k}$ are obtained as in the case of an Extended KF, using Jacobian matrices (F_k and B_k):

$$\hat{X}_{k+1/k} = f(\hat{X}_{k/k}, u_k) \tag{9}$$

$$P_{k+1/k} = F_k P'_{k/k} F_k^T + B_k Q'_u B_k^T + Q' \tag{10}$$

Notice that $P_{k+1/k}$ represents the scale matrix and not the covariance matrix. The latter can be easily retrieved as a function of the scale matrix: $\Sigma_{k+1/k} = \frac{\nu'_k}{\nu'_k - 2} P_{k+1/k}$.

Similarly to the time update, the measurement update requires an assumption regarding the degree of freedom. Hence, the joint distribution between the predicted state and the measurement noise can be written as follows:

$$p(X_k, w_k|Z_{1:k-1}) = St\left(\begin{bmatrix} X_k \\ w_k \end{bmatrix}; \begin{bmatrix} \hat{X}_{k/k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} P'_{k/k-1} & 0 \\ 0 & R'_k \end{bmatrix}, {\nu''}_k\right) \tag{11}$$

where X_k and w_k are assumed to be uncorrelated and $P'_{k/k-1}$ and R'_k are the adjusted matrices corresponding to the joint degree of freedom ${\nu''}_k = \min(\nu'_{k-1}, \delta_k)$.

The joint distribution of the state and measurement can be written as

$$p(X_k, Z_k|Z_{1:k-1}) = St\left(\begin{bmatrix} X_k \\ Z_k \end{bmatrix}; \begin{bmatrix} \hat{X}_{k/k-1} \\ \hat{Z}_k \end{bmatrix}, \begin{bmatrix} P'_{k/k-1} & P'_{k/k-1} H_k^T \\ H_k {P'}_{k/k-1} & S_k \end{bmatrix}, {\nu''}_k\right) \tag{12}$$

where $S_k = H_k P'_{k/k-1} H_k^T + R'_k$ and H_k is the Jacobian matrix corresponding to the measurement model.

The updated degree of freedom, state vector and scale matrix are [11]

$$\nu_k = \nu''_k + d_Z \tag{13}$$

$$\hat{X}_{k/k} = \hat{X}_{k/k-1} + P'_{k/k-1} H_k^T S_k^{-1} (Z_k - h(\hat{X}_{k/k-1})) \tag{14}$$

$$P_{k/k} = \frac{\nu''_k + \Delta_k^2}{\nu''_k + d_Z} (P'_{k/k-1} - P'_{k/k-1} H_k^T S_k^{-1} H_k P'_{k/k-1}) \tag{15}$$

$$= \frac{\nu''_k + \Delta_k^2}{\nu''_k + d_Z} (I - W_k H_k) P'_{k/k-1} \tag{16}$$

where d_Z is the dimension of the observation vector, $\Delta_k^2 = (Z_k - h(\hat{X}_{k/k-1}))^T S_k^{-1} (Z_k - h(\hat{X}_{k/k-1}))$ and $W_k = P'_{k/k-1} H_k^T S_k^{-1}$. The $P_{k/k}$ matrix is the same as that obtained in the case of KF, but scaled by a term that depends on the quality of the observations.

From equation 13 and after the first iteration, the dof increases to ν_k. In the absence of a limitation of this dof such as that presented in the prediction step ($\nu'_k = \min(\nu_k, \gamma_k)$), it will tend to infinity and the StF will converge to a KF.

D. Informational form of the Student's t Filter

A derivation of the informational form of the StF can be given similarly to the informational form of the KF, known as the Information Filter (IF) [7]. The IF is well adapted to decentralized multi-sensor data fusion and to multiple Fault Detection and Exclusion (FDE) thanks to its update step that is modeled as a simple summation of the information contributions of the different observations [13]. Likewise, the IF does not need to inverse a matrix that has the dimension of the observation vector, meaning that it has a shorter execution time than the KF in the case of multi-sensor fusion. In this section we derive the Student's t Information Filter (StIF) and examine its advantages in relation to an StF.

The StIF features an information matrix and an information vector obtained from the scale matrix (and not the covariance matrix) and the state vector:

$$Y_k = P_k^{-1} \tag{17}$$

$$y_k = Y_k X_k \tag{18}$$

The time update remains unchanged relatively to the IF. For the measurement update, equation 15 can be expressed as

$$P_{k/k} = \frac{\nu''_k + \Delta_k^2}{\nu''_k + d_Z} (P_{k/k-1}^{-1} + H_k^T R_k^{-1} H_k)^{-1} \tag{19}$$

This equation is obtained in a similar way to the computation of the information matrix in the case of the IF.

By setting $c_k = \frac{\nu''_k + \Delta_k^2}{\nu"_k + d_Z}$, the information matrix can be written as

$$Y_{k/k} = c_k^{-1} (Y_{k/k-1} + H_k^T R_k^{-1} H_k) \tag{20}$$

The derivation of the information vector is expressed in the same way as for IF and using the following equations [7]:

$$\begin{aligned}(I - W_k H_k) &= (P_{k/k-1} - W_k S_k W_k^T) P_{k/k-1}^{-1} \\ &= c_k^{-1} P_{k/k} P_{k/k-1}^{-1}\end{aligned} \tag{21}$$

$$\begin{aligned}W_k &= P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1} \\ W_k (H_k P_{k/k-1} H_k^T + R_k) &= P_{k/k-1} H_k^T \\ W_k R_k &= (I - W_k H_k) P_{k/k-1} H_k^T \\ W_k &= (I - W_k H_k) P_{k/k-1} H_k^T R_k^{-1} \qquad (22) \\ &= c_k^{-1} P_{k/k} H_k^T R_k^{-1} \qquad (23)\end{aligned}$$

Therefore, using equation 21 and 23, in the linear case, $X_{k/k}$ can be written as

$$\hat{X}_{k/k} = (I - W_k H_k) \hat{X}_{k/k-1} + W_k Z_k \tag{24}$$

$$\hat{X}_{k/k} = c_k^{-1} P_{k/k} [Y_{k/k-1} \hat{X}_{k/k-1} + H_k^T R_k^{-1} Z_k] \tag{25}$$

and the information vector is

$$\hat{y}_{k/k} = c_k^{-1} \hat{y}_{k/k-1} + c_k^{-1} H_k^T R_k^{-1} Z_k \tag{26}$$

In the case of multi-sensor data fusion and in a similar way to the IF, the equations of the StIF are obtained as

$$Y_{k/k} = c_k^{-1} \left(Y_{k/k-1} + \sum_{i=1}^{N} I_{i,k} \right) \tag{27}$$

$$\hat{y}_{k/k} = c_k^{-1} \left(\hat{y}_{k/k-1} + \sum_{i=1}^{N} i_{i,k} \right) \tag{28}$$

where

N is the number of observation vectors, assumed to be uncorrelated with each other,

$I_{i,k} = H_{i,k}^T R_{i,k}^{-1} H_{i,k}$ and $i_{i,k} = H_{i,k}^T R_{i,k}^{-1} Z_{i,k}$ are the information contributions of the observation $Z_{i,k}$.

Note that equations 27 and 28 distribute the computation of $I_{i,k}$ and $i_{i,k}$ thanks to the summation part. However, the quantity c^{-1} is present in both terms and in its present form it cannot be computed in a distributed manner since it needs to determine the value of Δ_k^2, and this involves inversing a matrix S that has the dimension of the observation vector. Note that S contains correlation terms. A new expression of Δ_k^2 can be determined using the matrix inversion lemma:

$$\begin{aligned}S^{-1} &= (H_k P_{k/k-1} H_k^T + R_k)^{-1} \qquad (29) \\ &= R_k^{-1} - R_k^{-1} H_k (Y_{k/k-1} + H_k^T R_k^{-1} H_k)^{-1} H_k^T R_k^{-1} \\ &= R_k^{-1} - R_k^{-1} H_k (Y_{k/k-1} + \sum_{i=1}^{N} H_{i,k}^T R_{i,k}^{-1} H_{i,k})^{-1} H_k^T R_k^{-1}\end{aligned}$$

The term $(Y_{k/k-1} + \sum_{i=1}^{N} H_{i,k}^T R_{i,k}^{-1} H_{i,k})$ has already been computed and the new expression of S^{-1} avoids the inversion of a non-diagonal matrix with the dimension of the observation vector. The matrices R_k and H_k present in this expression are obtained by concatenation.

Algorithm 1 StIF with fixed degree of freedom.

Input: $\hat{X}_{k-1/k-1}, Y_{k-1/k-1}, \nu_{k-1},$

We assume that $\gamma = \delta$ (no approximation of the dof is needed when going from time update to measurement update)

$\nu'_{k-1} = \min(\nu_{k-1}, \gamma) = \gamma$

Moment matching to adjust the matrices Q and $P_{k-1/k-1}$:

We suppose that the matrices Q and R are the same as in the Gaussian case

$$P'_{k-1/k-1} = \frac{\nu'_{k-1} - 2}{\nu'_{k-1}} \frac{\nu_{k-1}}{\nu_{k-1} - 2} P_{k-1/k-1} \tag{31}$$

$$Q' = \frac{\nu'_{k-1} - 2}{\nu'_{k-1}} Q \tag{32}$$

Time update

$X_{k/k-1} = f(X_{k-1/k-1}, u_{k-1}) + v_k$

$P_{k/k-1} = F_{k-1} P'_{k-1/k-1} F_{k-1}^T + B_{k-1} Q'_{u,k-1} B_{k-1}^T + Q'_{k-1}$

F_k and B_k are the Jacobian matrices

Measurement update

$$R' = \frac{\nu'_{k-1} - 2}{\nu'_{k-1}} R$$

$$c_k = \frac{\nu'_{k-1} + \Delta_k^2}{\nu'_{k-1} + d_Z}$$

$$Y_{k/k} = c_k^{-1} (Y_{k/k-1} + \sum_{i=1}^{N} I_{i,k})$$

$$\hat{y}_{k/k} = c_k^{-1} \left(y_{k/k-1} + \sum_{i=1}^{N} i_{i,k} \right)$$

$$\nu_k = \nu'_{k-1} + d_Z$$

Covariance matrix: $\Sigma_{k/k} = \frac{\nu_k}{\nu_k - 2} Y_{k/k}^{-1}$

$X_{k/k} = Y_{k/k}^{-1} y_{k/k}$

Output: $\hat{X}_{k/k}, Y_{k/k}, \nu_k$

For non-linear systems, $i_{i,k}$ in equation 28 becomes

$$i_{i,k} = H_{i,k}^T R_{i,k}^{-1} \left[\left(Z_{i,k} - \hat{Z}_{i,k} \right) + H_{i,k} \hat{X}_{k/k-1} \right] \tag{30}$$

The algorithm for the StIF with a fixed degree of freedom ($\gamma = \delta$) is given in algorithm 1. Another algorithm will be given in section II-E after developing a fault detection and exclusion stage.

E. Robust StIF with adaptive degree of freedom

Although Student's t Filter is designed to be more robust to faults than KF insofar as larger errors are taken into account via the heavier tail of the density function, these faults still influence the quality of the filter estimation. Therefore we propose adding a fault detection and exclusion stage in order to exclude these errors from the fusion procedure. Note that in the StF, when a fault is presented in the system, Δ^2 increases, leading to a significant increase in the value of the covariance matrix.

The first step of an FDE is the generation of the residual. In this work, a residual based on the Mahalanobis distance between the prediction (obtained from dead-reckoning) and the update estimation (obtained after merging the measurements Z) is constructed as

$$r_k = (\hat{X}_{k/k} - \hat{X}_{k/k-1})^T Y_{k/k} (\hat{X}_{k/k} - \hat{X}_{k/k-1}) \tag{33}$$

where $\hat{X}_{k/k}$ and $Y_{k/k}$ are obtained from a filter that uses all available measurements at instant k.

Similarly to the use of the Chi-squared distribution in the Gaussian case for threshold computation, an F-distribution is employed in the Student's t case [5]. The variable $\frac{r_k}{d_X}$ follows an F-distribution (F_{d_X,ν_k}) with parameter d_X (dimension of the state vector) and $\nu_k = \nu'_k + d_Z$, and therefore

$$r_k \sim d_X F_{d_X,\nu_k} \tag{34}$$

The threshold value for the statistical test over r_k corresponding to a given false alarm probability P_{Fa} is given as

$$Th_{P_{Fa}} = d_X \mathcal{F}^{-1}_{1-P_{Fa}}(F_{d_X,\nu_k}) = d_X \mathcal{F}^{-1}_{P_{1-Fa}}(F_{d_X,\nu'_k+d_Z}) \tag{35}$$

where $\mathcal{F}^{-1}$ is the inverse cumulative distribution.

If the residual r_k indicates the presence of faults by exceeding the threshold value, residuals used for the exclusion stage are generated by creating a bank of StIF where each one uses only one observation. These residuals are computed as

$$r_{i,k} = (\hat{X}_{i,k/k} - \hat{X}_{k/k-1})^T Y_{i,k/k} (\hat{X}_{i,k/k} - \hat{X}_{k/k-1}) \tag{36}$$

where $\hat{X}_{i,k/k}$ and $Y_{i,k/k}$ are obtained from a filter that uses only the observation i in the update step.

A measurement i is excluded if

$$r_i > Th_i = d_X \mathcal{F}^{-1}_{1-P_{Fa}} \left(F_{d_X,\nu'_k+d_{Z,i}} \right) \tag{37}$$

Then, the measurement i is excluded from the fusion procedure by subtracting its information contribution from the main filter and by updating the c term. This exclusion step is more complicated to do when the KF (or StF) is used.

Regarding the degree of freedom of the Student's t filter, instead of fixing it as proposed in section II-C, it can be adapted in real-time by taking into account the value of the residual. This residual can provide an indication about the quality of the measurements on this particular sample. Intuitively, when the residual is large, one may wish to be more cautious in the estimation process and when computing the confidence interval, and for this purpose a heavy-tailed distribution may be preferred. Inversely, when the residual is small, a Gaussian distribution may be more reasonable. Starting from large values of the residual, the degree of freedom increases slowly up to a given limit when the residual decreases (e.g dof=20), and if r_k is greater than a given value the dof should be limited to 2.1. We propose an expression of the degree of freedom according to the residual as follows:

$$\begin{cases} \text{dof}_k(r_k) = a.\exp(b.r_k) & \text{if } r_k < d \\ \text{dof}_k(r_k) = 2.1 & \text{if } r_k > d \end{cases} \tag{38}$$

Algorithm 2 Multi-sensor data fusion with FDE

Input: $\hat{X}_{k-1/k-1}, Y_{k-1/k-1}, \nu_{k-1}, \gamma_k = \delta_k$
if $k = 0$ **then**
  $\nu'_{k-1} = \nu_{k-1} = \gamma_k = \delta_k$
else
  $\gamma_k = \delta_k = \nu'_{k-1}$, ($\nu'_{k-1}$ obtained in the measurement update)
Moment matching: Compute $P'_{k-1/k-1}$ and Q' as in algorithm 1
Time update: same as algorithm 1
Measurement update: Compute R', c_k,$Y_{k/k}$ and $y_{k/k}$ as in algorithm 1.

$$\nu_k = \nu'_{k-1} + d_z \tag{39}$$

Compute the corresponding residual r_k
if $r_k > Th$ **then**
  compute the set of residuals $r_{i,k}$ using a bank of StIF
  exclude the measurements that have $r_{i,k} > Th_i$,
  compute the new value of c_k, $Y_{k/k}$ and $\hat{y}_{k/k}$
$\hat{X}_{k/k} = Y^{-1}_{k/k}\hat{y}_{k/k}$
if measurements available **then**
  compute the new value of ν'_k that will be used in the next iteration, as in equation 38.
else
  $\nu'_k = \nu_k = \nu'_{k-1}$
Output: $\hat{X}_{k/k}, Y_{k/k}, \nu_k, \nu'_k$

We note that in the case study given in section III, a , b and d were tuned with the following values: a =20.1137, $b = -0.0565$ and $d = 40$. With this tuning, the degree of freedom tends to 20 if r_k tends to 0 and to 2.1 if the residual is greater than d.

Algorithm 1 is then modified to algorithm 2 in order to take into account the FDE stage and the computation of the adaptive degree of freedom.

III. Case Study

A. Localization for self-driving vehicles

Let consider the localization problem of a self-driving vehicle equipped with wheel-speed sensors, a gyro, a low-cost GNSS receiver, an intelligent camera for lane-marking detection and an HD map. The localization is defined relatively to a Cartesian ENU frame (East, North, Up) denoted R_O. The vehicle pose is defined at the body frame located at the middle of the rear wheel axis (frame R_B) (see figure 1).

At instant k, the state vector is considered to be the position and the heading of the vehicle:

$$X = [\ x \quad y \quad \theta\ ]^T \tag{40}$$

The GNSS observation is

$$Z_1 = \begin{bmatrix} x_{\text{GNSS}} \\ y_{\text{GNSS}} \\ \theta_{GNSS} \end{bmatrix} = \begin{bmatrix} t_x \cos\theta - t_y \sin\theta + x \\ t_x \sin\theta + t_y \cos\ \theta + y \\ \theta \end{bmatrix}, \tag{41}$$

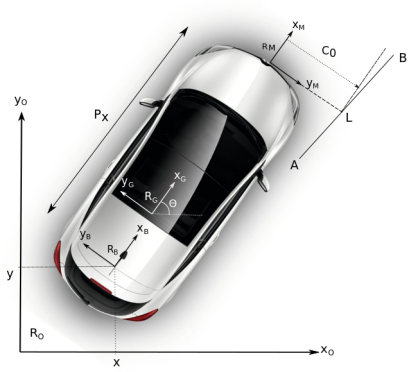


Fig. 1: Frames representation.

where $[t_x\ t_y]$ is the translation of the antenna (located at frame R_G) with respect to the body frame.

A smart camera detects the lane markings and returns the coefficients of a Taylor's expansion of a clothoid in the frame R_M[14]:

$$y = C_3x^3 + C_2x^2 + C_1x + C_0 \tag{42}$$

The camera is able to detect up to two markings at a time on each side (right and left) of the car. In this paper, we consider only the use of the lateral distance to the lane markings (the C_0 parameter):

$$Z_{2,j} = C_0 \tag{43}$$

where $j \in \{1, 2, 3, 4\}$ is an index representing the detected marking: right$_1$, right$_2$, left$_1$ or left$_2$ if they exist.

A camera observation model was developed in [15], [14] and is given by

$$C_0 = \frac{(P_x \sin\theta + y - y_A)x_{AB} - (P_x \cos\theta + x - x_A)y_{AB}}{x_{AB} \cdot \cos\theta + y_{AB} \cdot \sin\theta}$$

where the lane marking is detected on the segment [AB] at coordinates $[x_L, y_L]$ expressed in frame R_O as

$${}^OL = {}^O\,T_B.{}^BT_M.{}^ML \tag{44}$$

${}^OT_B$ and ${}^BT_M$ are the transformation matrices between (R_B and R_O) and (R_M and R_B) respectively. It results that

$$\begin{bmatrix} x_L \\ y_L \end{bmatrix} = \begin{bmatrix} P_x \cos\theta + C_0 \sin\theta + x \\ P_x \sin\theta - C_0 \cos\theta + y \end{bmatrix} \tag{45}$$

where ${}^ML = [0, C_0]$ and P_x is the distance between R_B and R_M (figure 1).

For state estimation, the robust StIF developed in algorithm 2 is used. Therefore, the residual r_k defined in equation 33 is computed in order to detect the presence of faults. Note that the prediction is obtained from the wheel speed sensors and the gyro while the update is done using the GNSS fix and the Mobileye measurements. Mobileye measurements are not always available, since the camera cannot operate in highly curved roads. For the fault exclusion, a set of residual $r_{i,k}$ is computed (equation 36) for each measurement i, which can be the GNSS measurement or one of the four Mobileye measurements.

Let us choose the example of the residual corresponding to the GNSS measurement ($i = gnss$, $d_{Z|gnss} = 3$) as illustration. $\hat{X}_{i,k/k}$ and $Y_{i,k/k}$ are obtained from the following equations:

$$S^{-1} = R_i^{-1} - R_i^{-1} H_i (Y_{k/k-1} + H_{i,k}^T R_{i,k}^{-1} H_{i,k})^{-1} H_i^T R_i^{-1}$$

$$c_{i,k} = \frac{\nu'_k + \Delta_k^2}{\nu'_k + 3}$$

$$Y_{i,k/k} = c_{i,k}^{-1} (Y_{k/k-1} + H_{i,k}^T R_{i,k}^{-1} H_{i,k})$$

$$\hat{y}_{i,k/k} = c_{i,k}^{-1} \left(\hat{y}_{k/k-1} + i_{i,k} \right)$$

$$\hat{X}_{i,k/k} = Y_{i,k/k}^{-1} \hat{y}_{i,k/k}$$

If $r_{gnss} > Th_{gnss} = d_X \mathcal{F}^{-1}_{1-P_{Fa}}(F_{d_X,\nu'_k+3})$, the GNSS measurements are excluded from the fusion procedure by subtracting their information contributions from the main filter and by updating the c_k term.

B. External integrity and computation of protection level

External integrity is an emergent concern for road localization. External localization integrity was initially developed in aeronautical applications in order to provide a measure of confidence for GNSS navigation solutions [16]. It is associated with a target integrity risk that represents the maximum probability that the error in position exceeds a limit without warning the user. This limit is known as the Protection Level (PL). In this paper we are seeking to bound the errors by computing a consistent PL that makes use of Student's t distributions. For this computation, the matrix $P_{k/k}$ obtained in the output of algorithm 2 is used. The heavy tail property of Student's t distribution is essential for computing a consistent PL without underestimating its value, which may happen in the Gaussian case [17]. For this purpose, the matrix $P_{k/k}$ is adjusted according to the minimum degree of freedom between ν'_k and ν_k (denoted M in equations 48 and 49). If this minimum value is greater than a limit value fixed according to the application and to the environment, the degree of freedom for the PL computation is taken to be this limit value. For example, in urban environments the degree of freedom for the PL computation should not go above $M = 5$. After updating the degree of freedom, the matrix $P_{k/k}$ is adjusted using the moment matching. The adjusted matrix is denoted P''_k and is only used for the PL computation. The reason for choosing the degree of freedom in this way is to preserve a heavy tail, which is necessary in order to obtain a consistent protection level (for more details the reader can refer to [17]).

For road vehicles, we would like to be able to bound the errors in the along track (AT) and cross track (CT) directions. Therefore, the shape matrix in the ENU frame is projected onto the (AT, CT) frame:

$$P''_{k/k,proj} = R_{proj} {P''}_{k/k(1,2),(1,2)} R_{proj}^T \quad (46)$$

with

$$R_{proj} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \quad (47)$$

where θ is the orientation of the vehicle.

The protection levels are obtained as

$$PL_{AT} = K(\alpha, N).\sqrt{M}.[\text{eigenvalue}({P''}_{k/k,proj})]_1 \quad (48)$$

$$PL_{CT} = K(\alpha, N).\sqrt{M}.[\text{eigenvalue}({P''}_{k/k,proj})]_2 \quad (49)$$

where the indexes 1 and 2 denote the first and second elements of the vector (eigenvalue$({P''}_{k/k,proj})$) and $K(\alpha, M)$ is obtained from Student's t distribution with degree of freedom M according to a given confidence level α [18].

IV. Experimental results

Fig. 2: Test trajectory of 3.5 km.

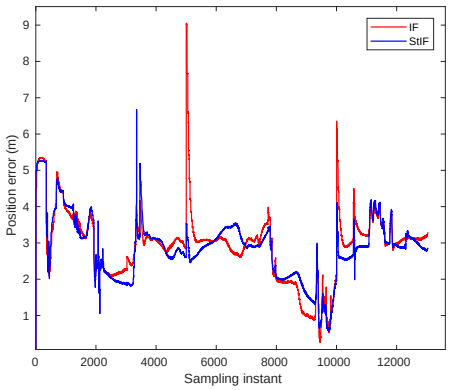
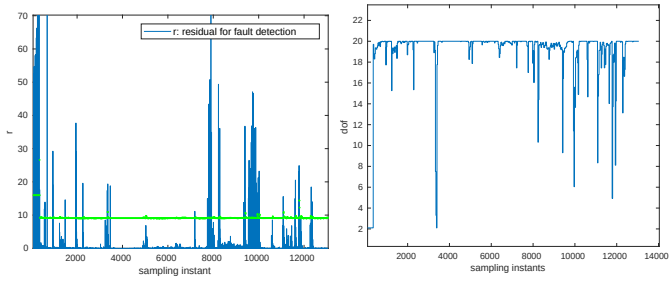


Fig. 3: Error with additive biases on GNSS fix: StIF with adaptive dof (in blue) and IF (in red). Biases were injected in the intervals 5000 to 5010 and 10000 to 10010.

In order to evaluate the performance of the proposed approach, results were processed using sensor data recorded by an experimental vehicle equipped with wheel-speed sensors and the yaw rate gyro of the trajectory stability system of the car, a Ublox 8T which is a single frequency GPS/GLONASS constellation receiver, a Mobileye camera for lane-marking detection and a NovAtel SPAN-CPT with network RTK corrections used as a ground truth. Experiments were carried out in Rambouillet, France (figure 2). The results shown were obtained in a challenging environment where the GNSS measurements were affected by significant bias throughout the trajectory. The purpose of choosing this kind of trajectory was to study the performance of our proposed approach in this kind of challenging environment and to compare the StIF to the IF. The processing of the data was handled



(a) Residual for FDE

(b) The dof in real-time

Fig. 4: (a) Residual for the fault detection (in blue), threshold (in green), (b) the associated degree of freedom.

asynchronously. The wheel-speed sensor and the gyro data were available at a frequency of 50 Hz, the GNSS at 2 Hz and the frequency of the data from Mobileye camera was down-sampled to 3.3 Hz. The sampling period of the filter is the same as for the dead-reckoning data (0.02s).

Figure 3 shows the position errors obtained using an IF and an StIF without an FDE stage. To provide a better understanding of the results, simulated biases on GNSS measurements were injected between the sampling instants 5000 to 5010 and 10000 to 10010. It can be remarked that after the error source is removed and without an FDE stage, the Student's t filter (StIF) convergence is faster than a Gaussian filter (IF). Although the StIF behaves better than the IF in the presence of errors, large errors still affect its behavior. It should be noted that the StIF used for this figure is the one presented in algorithm 2 without an FDE stage. The degree of freedom is adapted in real-time according to the residual value. The values chosen for a , b and d in equation 38 were a =20.1137, $b = 0.0565$, and $d = 40$.

Figure 4 shows the residual used for the fault detection with the associated threshold value. This threshold varies with the degree of freedom computed in real-time (figure 4b). When r indicates the presence of a fault, the set of residuals used for the fault exclusion are generated (if available) and are shown in figure 5. Figure 5a shows the residual corresponding to the GNSS measurements and figures 5b and 5c show the residuals corresponding to the camera measurements on the left and right-hand sides respectively.

The errors in the CT and AT directions are shown in figure 6 with the corresponding PL value. The target integrity risk was fixed to 10^{-3}. The contribution made by the Mobileye camera in reducing the error and the PL values in the CT direction (highlighted zone) can be seen. When there is no camera measurement, the non-observable biases from GNSS lead to inconsistent PL despite the use of Student's t distribution. Moreover, this sensor does not have any effect on the AT direction, since it is not used to detect markings in this direction. Therefore, the PL in the AT direction is not consistent. The limit value of the dof for PL computation was fixed to $M = 5$ (see section III-B). Note that if this limit were fixed to $M = 3$, the PLs value would be consistent in the AT and CT directions.

Table I shows the errors and the percentile values before and after the FDE stage for both the StIF and the IF. For this experiment where a bias was present throughout the trajectory, even though the StIF and the FDE stage do not significantly improve the results, the position estimation using StIF was more accurate and more consistent than when using IF. For this particular case, the IF after the FDE stage results in a larger error than before FDE. The reason is that the algorithm excludes many camera measurements; using these measurements would give better results. For this trajectory, StIF with dof=3 was slightly more accurate than StIF with adaptive dof. However, on other trajectories, the StIF with dof=3 may give the poorest results, according to the error model. This shows that the choice of a fixed dof may not be straightforward in real applications. In this experiment, according to the percentile values, it can be remarked that the largest errors that occur are smaller when using StIF with FDE than when using an IF (a reduction from 4.15m to 3.8m for the 97.5 percentile).

From a consistency point of view, table I also shows the measured integrity risk obtained from the experiment in the AT and CT directions for a target integrity risk of 10^{-3}. Given the number of available samples in the experiment, we cannot perform tests for a target integrity risk below 10^{-3}. However, the Student's t distribution and the Student's t filter are more advantageous and more consistent for small integrity risk in comparison to the Gaussian cases, especially when using the moment matching to adjust distributions. Regarding the measured IR, the StIF is more consistent than the IF, and the StIF with adaptive dof leads to better results since the dof is modified according to the residual, and heavier tails are used when the environment is too noisy.

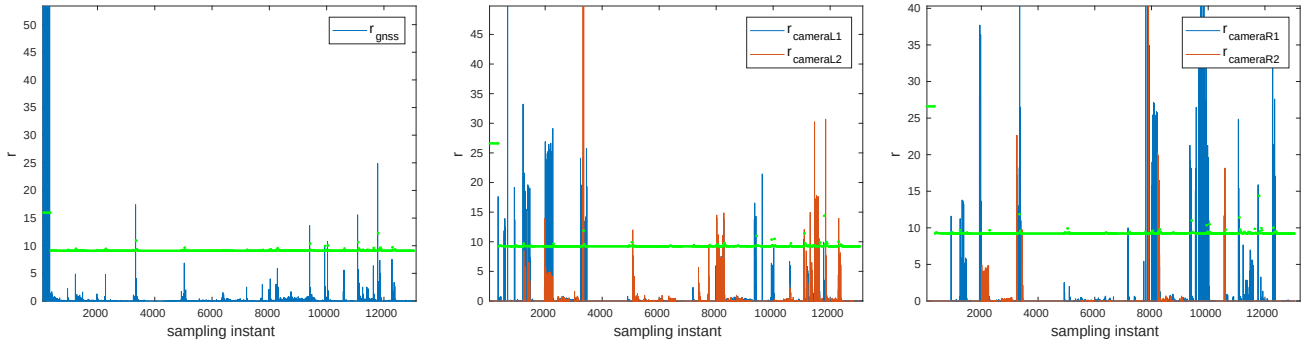
V. Conclusion and Future Work

In this paper a new robust Student's t information filter was proposed. The StIF, which is the informational form of the StF, was developed. Using this formulation, a fault detection and exclusion stage was added in order to exclude erroneous measurements that have a direct influence on the state estimation. The computed residual gives an indication regarding the environmental conditions and the quality of the measurements in real-time. This residual is used to adapt the degree of freedom and to avoid setting it *a priori*. Experimental results using a vehicle equipped with different sensors showed that the StIF with an adaptive degree of freedom improves performance, particularly in terms of integrity.

In future works, additional studies on the variation of the degree of freedom in real-time will be done and the approach will be tested with trajectories in different environments. Likewise, the results using the moment matching will be compared with other approaches, such as minimizing the Kullback-Leibler divergence.

References

[1] D. Betaille and R. Toledo-Moreo, "Creating enhanced maps for lane-level vehicle navigation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 4, pp. 786–798, 2010.



(a) r_{GNSS} (b) $r_{\text{camera Left},1}$ (blue), $r_{\text{camera Left},2}$ (red) (c) $r_{\text{camera Right},1}$ (blue), $r_{\text{camera Right},2}$ (red)

Fig. 5: Residuals for fault exclusion (without unit), in green the thresholds.

error (m)	without FDE				with FDE			
	StIF(adapt. dof)	StIF(dof=3)	StIF(dof=8)	IF	StIF(adapt. dof)	StIF(dof=3)	StIF(dof=8)	IF
Root mean square	2.82 ± 0.77	2.78 ± 0.72	2.84 ± 0.69	2.91 ± 1.2	2.75 ± 0.64	2.62 ± 0.72	2.69 ± 0.61	2.98 ± 0.64
50 %-th percentile	2.92	2.84	2.88	2.92	2.93	2.78	2.86	3.02
75 %-th percentile	3.23	3.19	3.24	3.57	3.1	3.15	3.17	3.54
97.5%-th percentile	5.2	4.97	4.94	5.26	3.8	4.05	3.90	4.15
measured CT-IR	0.003	0.01	0.01	0.08	0.010	0.014	0.011	0.27
measured AT-IR	0.21	0.28	0.31	0.43	0.20	0.27	0.30	0.40

TABLE I: Statistical values of the horizontal positioning errors.

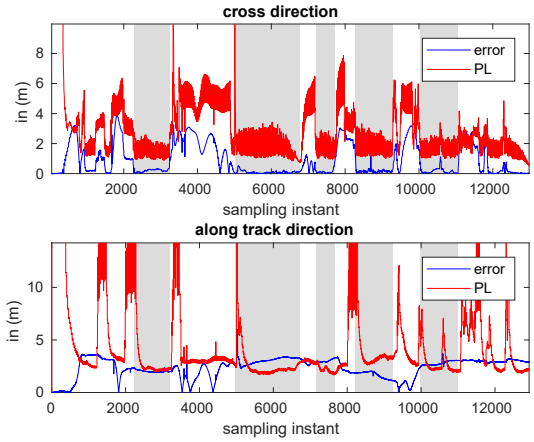


Fig. 6: The error (in blue) and the PL (in red) in the AT and CT directions (case of StIF with adaptive dof). Highlighted zones correspond to straight line where the Mobileye detect well the markings.

[2] M. A. Quddus, W. Y. Ochieng, and R. B. Noland, "Current map-matching algorithms for transport applications: State-of-the art and future research directions," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 5, pp. 312–328, 2007.

[3] C. E. White, D. Bernstein, and A. L. Kornhauser, "Some map matching algorithms for personal navigation assistants," *Transportation Research Part C: Emerging Technologies*, vol. 8, no. 1-6, pp. 91–108, 2000.

[4] A. Kasmi, D. Denis, R. Aufrere, and R. Chapuis, "Map matching and lanes number estimation with Openstreetmap," in *International Conference on Intelligent Transportation Systems*, 2018, pp. 2659–2664.

[5] M. Roth, "Kalman filters for nonlinear systems and heavy-tailed noise," phdthesis, Linköping University Electronic Press, 2013.

[6] Y. Huang, Y. Zhang, N. Li, and J. Chambers, "Robust student's t based nonlinear filter and smoother," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2586–2596, Oct. 2016.

[7] H. Durrant-Whyte and T. C. Henderson, "Multisensor data fusion," *Springer handbook of robotics*, pp. 585–610, 2008.

[8] Q. Li, Y. Ben, J. Tan, S. M. Naqvi, and J. Chambers, "Robust selection of the degrees of freedom in the student's distribution through multiple model adaptive estimation," *Signal Processing*, vol. 153, pp. 255–265, 2018.

[9] S. Kotz and S. Nadarajah, *Multivariate T-Distributions and Their Applications*. Cambridge university, Feb. 2004.

[10] S. Nadarajah and S. Kotz, "Mathematical properties of the multivariate t distribution," *Acta Applicandae Mathematica*, vol. 89, no. 1, pp. 53–84, 2005.

[11] M. Roth, T. Ardeshiri, E. Ozkan, and F. Gustafsson, "Robust bayesian filtering and smoothing using student's t distribution," *arXiv:1703.02428 [cs, stat]*. [Online]. Available: http://arxiv.org/abs/1703.02428

[12] M. Roth, E. Ozkan, and F. Gustafsson, "A student's t filter for heavy tailed process and measurement noise," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 5770–5774.

[13] J. Al Hage, M. E. El Najjar, and D. Pomorski, "Multi-sensor fusion approach with fault detection and exclusion based on the Kullback-Leibler divergence: Application on collaborative multi-robot system," *Information Fusion*, vol. 37, pp. 61–76, 2017.

[14] Z. Tao, Ph. Bonnifait, V. Frémont, J. Ibañez Guzman, and S. Bonnet, "Road-centered map-aided localization for driverless cars using single-frequency GNSS receivers," *Journal of Field Robotics*, vol. 34, no. 5, pp. 1010–1033, 2017.

[15] Z. Tao, Ph. Bonnifait, V. Frémont, and J. Ibañez Guzman, "Mapping and localization using GPS, lane markings and proprioceptive sensors," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2013, pp. 406–412.

[16] N. Zhu, J. Marais, D. Betaille, and M. Berbineau, "GNSS position integrity in urban environments: A review of literature," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 9, pp. 2762–2778, Jan. 2018.

[17] J. Al Hage, Ph. Xu, and Ph. Bonnifait, "Bounding localization errors with Student distribution for road vehicles," in *International Technical Symposium on Navigation and Timing*, 2018.

[18] P. F. Navarro Madrid, "Device and method for computing an error bound of a Kalman filter based gnss position solution," patent, Apr., 2016.