

Insights on Adaptive Robust Filtering for Navigation under Harsh Time-Varying Environments

Paul Chauchat
Aix-Marseille Univ, CNRS, LIS
Marseille, France

Andrea Bellés
German Aerospace Center
Neustrelitz, Germany

Daniel Medina
German Aerospace Center
Neustrelitz, Germany

Jordi Vilà-Valls
ISAE-SUPAERO
Toulouse, France

Abstract—Standard filtering techniques operate under the assumption that the system is perfectly known: system matrices/functions, noise statistics and inputs. But such strong assumption does not typically hold in real-world applications. Indeed, when the assumed model does not perfectly align with the true system dynamics (i.e., model mismatch) the optimality properties of the Kalman filter and its nonlinear extensions are compromised, and the filter performance can be significantly degraded, reason why robust solutions must be accounted for. This contribution explores how a recently introduced adaptive robust regression framework can be adapted to the recursive filtering case, being then able to deal with time-varying outliers in the observation model. Methodological and practical insights are given regarding the design and implementation of the method. An illustrative navigation example is provided to highlight the filters' advantages and limits, and support the discussion.

I. INTRODUCTION

It is well known that state estimation is a fundamental task in several engineering applications [1]–[3], and in particular for navigation systems [4]. The goal is to infer the dynamic states of a system from a set of available noisy observations, thus resorting to a stochastic state-space representation of the system dynamics. Such problem formulation can be linear/nonlinear, Gaussian/non-Gaussian, and possibly taking into account system inputs, but in any case, standard filtering techniques operate under the assumption that the system is perfectly known: system matrices/functions, noise statistics and inputs; an assumption that does not typically hold in real-world applications, and therefore one must deal with a model mismatch (i.e., the assumed model does not perfectly align with the true system dynamics). This applies to both standard linear Kalman filter (KF) and extended KF (EKF) [3], but also to sigma-point Gaussian filters [5] and particle filters [6].

A plethora of robust filtering techniques have been developed in the last decade in order to mitigate the inherent model mismatch impact on the final filter performance: i) constrained filtering approaches to cope with process/measurement functions' parametric mismatch [7], [8], ii) filters directly designed for a specific non-Gaussian noise distribution [9]–[12], iii) outlier detection and rejection methods [13], [14], or iv) robust statistics-based filters [15]–[18], being the class of robust filters of interest in this contribution.

This research was partially supported by DLR's FuturePorts project and the DGA/AID project 2022.65.0082.

The idea behind estimation methods based on robust statistics is to use robust loss functions (RLFs) able to correctly penalise observation outliers not fitting the nominal noise statistics. This applies to batch linear/nonlinear least squares (LS) solutions, but also to KF-like techniques. While in the former the ℓ_2 -norm minimisation of the residuals is replaced by an iterative reweighted LS (with appropriate weights), in the latter the weights are used to modify the measurement noise covariance or the innovation sequence [15].

There exist several RLFs in the literature (see [19]), the most common being the monotone Huber and the redescending bisquare Tukey [15]. The main limitation of these functions is that their behaviour, or their sensitivity to detect outliers, is driven by a tuning hyperparameter (i.e., typically set to achieve a given efficiency under nominal Gaussian conditions). There has been prior work on self-tuning robust estimators [20], but as pointed out in [19], this hyperparameter is most often a priori selected without knowing the actual distribution of the residuals (and in some cases disregarding a non-zero mode), reason why adaptive loss functions have been proposed to avoid such tuning [19], [21], [22]. The limitations of a fixed hyperparameter were also discussed in [23] for a robust navigation problem, where the authors propose a supervised learning approach to infer this parameter depending on the navigation context. The adaptive loss of [21] was also used in a filtering framework in some recent works, but without giving much insights or implementation details [24], [25].

The main contribution of this article is to further explore the use of a recently introduced adaptive RLF [19] (designed for batch LS problems) in a filtering context, i.e., how to embed the adaptive weighting in a KF-like method. In particular, the article focuses on the impact of the output size, and of the way the residuals are normalised. For the latter, a specific scheme is proposed for the recursive estimation framework. Such a solution would overcome the limitations of standard robust M-type KFs, and adapt to time-varying scenarios without need of a supervised learning scheme as in [23].

II. SYSTEM MODEL AND PROBLEM FORMULATION

In general, the state-space model (SSM) representation of the system dynamics is given by

$$\mathbf{x}_t = \mathbf{f}_{t-1}(\mathbf{x}_{t-1}) + \mathbf{v}_{t-1}, \quad (1)$$

$$\mathbf{y}_t = \mathbf{h}_t(\mathbf{x}_t) + \mathbf{n}_t, \quad (2)$$

where $\mathbf{x}_t \in \mathbb{R}^{n_x}$ are the hidden state of the system to be inferred from the available observations $\mathbf{y}_t \in \mathbb{R}^{n_y}$; $\mathbf{f}_{t-1}(\cdot)$ and $\mathbf{h}_t(\cdot)$ are known functions of the state; both noise sequences $\mathbf{v}_{t-1}$ and $\mathbf{n}_t$ are assumed to be independent. While the process noise is typically assumed to be zero-mean Gaussian distributed, $\mathbf{v}_{t-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{t-1})$, in real-life time-varying harsh environments the measurement noise distribution may deviate from Gaussianity. A way to account for measurement outliers is to consider a parametric non-Gaussian distribution (e.g., Student-t, Laplace or skew-t [10]), but the resulting robust filter is then constrained to the imposed noise model and not flexible to time-varying environments. In addition, given the random nature of outliers, the underlying non-Gaussian distribution is usually unknown. Then, one may consider the typical contamination model arising in robust statistics [15]: a proportion $1 - \beta$ of observations under nominal Gaussian noise, and another proportion $0 \leq \beta \leq 1$ of observations contaminated by an unknown arbitrary distribution $\mathcal{H}$,

$$n_{i,t} \sim (1 - \beta) \mathcal{N}(0, \sigma_{i,t}^2) + \beta \mathcal{H}. \quad (3)$$

Under these conditions, the problem is to recursively approximate the posterior $p(\mathbf{x}_t | \mathbf{y}_{1:t})$, therefore to design a robust filter able to perform close to optimal under nominal Gaussian conditions, that does not break down under outliers (i.e., optimality vs. robustness trade-off), and providing an adaptive solution to time-varying environments.

III. STANDARD AND ADAPTIVE ROBUST LOSS FUNCTIONS

The linear regression problem, $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, with $\mathbf{x} \in \mathbb{R}^{n_x}$ a vector of unknown parameters to be estimated, and $\mathbf{n}$ noise of known covariance, is typically solved applying a weighted LS estimator (minimization of the ℓ_2 -norm of the residuals, $r_i(\mathbf{x}) = y_i - \mathbf{H}_i^\top \mathbf{x}$),

$$\hat{\mathbf{x}}_{\text{LS}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \Rightarrow \arg \min_{\mathbf{x}} \sum_{i=1}^{n_x} \left(\frac{r_i(\mathbf{x})}{\sigma_i} \right)^2, \quad (4)$$

with σ_i^2 the variance of the noise components n_i (i.i.d. along the set of observations), which is optimal if the Gaussian assumption holds for $\mathbf{n}$. In the case of measurement outliers, the regression problem can be generalised as,

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \sum_{i=1}^{n_x} \rho \left(\frac{r_i(\mathbf{x})}{\sigma_i} \right), \quad (5)$$

with $\rho(x)$ a RLF that allows to penalise outliers that deviate from the nominal (assumed) noise distribution. The regression problem is then classically solved via an iterated reweighted LS (IRLS) approach. At each step, the following LS is solved

$$\hat{\mathbf{x}}^{(i)} = \arg \min_{\mathbf{x}} \sum_{i=1}^{n_x} w \left(\frac{r_i(\mathbf{x})}{\sigma_i} \right) \left(\frac{r_i(\mathbf{x})}{\sigma_i} \right)^2, \quad w(r) = \frac{\rho'(r)}{r}. \quad (6)$$

Several RLFs exist in the literature, the so-called Huber RLF presented hereafter being the most widespread.

A. Huber robust loss function

The family of Huber RLFs, and the associated weight functions, are defined for an arbitrary scalar variable r as

$$\rho_H(r) = \begin{cases} r^2 & \text{if } |r| \leq c \\ 2c|r| - c^2 & \text{if } |r| > c \end{cases}, \quad (7)$$

$$w_H(r) = \min \left\{ 1, \frac{c}{|r|} \right\}. \quad (8)$$

Notice that the Huber RLF is a combination of the ρ_{ℓ_2} and ρ_{ℓ_1} , behaving as a LS for small errors and as a least-absolute value (ℓ_1) for larger ones. The behaviour of the RLF is controlled by the parameter c , which is typically chosen to achieve a given efficiency to the normal distribution. The most common value is $c = 1.345$, for a 95% normal efficiency [15].

B. Adaptive robust loss function

A number of authors proposed various RLFs to replace the Huber family [15]. Lately, the focus was put on trying to adapt the parameters of these RLFs given current residuals. The notion of self-tuning estimators had already been explored [20]. For the Huber loss function this was for instance done through machine learning [23]. Recently, more general classes of RLFs are emerging. For instance, a class of parameterised loss functions was proposed for machine learning purposes in [21]. Its main parameter $\alpha \in]-\infty, 2]$ allows to recover ℓ_2 , Cauchy, Geman-McClure and Welsch losses as special cases, and proposes an infinity of other losses. Its main interest is that it is possible to adapt α to given data. This RLF was adapted to the robotics community, focusing on point cloud registration, and proposing improvements [19], [22]. We recall hereafter the main concepts of these works: the class of RLF and associated weights was originally given by [21], [22]

$$\tilde{\rho}(r, \alpha) = \frac{|\alpha - 2|}{\alpha} \left(\left(\frac{r^2}{|\alpha - 2|} + 1 \right)^{\alpha/2} - 1 \right), \quad (9)$$

$$\tilde{w}(r, \alpha) = \left(\frac{r^2}{|\alpha - 2|} + 1 \right)^{\alpha/2 - 1}, \quad (10)$$

with special cases for $\alpha = 2$ (standard square loss) and $\alpha = 0$ (Cauchy loss). Note that these functions are even in r , only depending on $|r|$. It was proposed in [19] to add an additional parameter to the RLF family, setting a threshold below which residuals remain unweighted. This leads to

$$\bar{\rho}(r, \alpha, \varepsilon) = \tilde{\rho}(\max(|r| - \varepsilon, 0), \alpha), \quad (11)$$

$$\bar{w}(r, \alpha, \varepsilon) = \tilde{w}(\max(|r| - \varepsilon, 0), \alpha). \quad (12)$$

This choice was motivated by the fact that [19] considered residuals coming from vector norms, hence having non-zero mean. As will be explained in Section IV-B2, this choice can also be interesting in the case of residuals linked to 1D measurements (i.e., as in our illustrative example).

C. Comparison of the RLFs' parameters

As explained in [21], [22], the parameter α of $\tilde{\rho}$ decides how hard the downweighting of the residuals is. This clearly appears in the top plot of Figure 1.

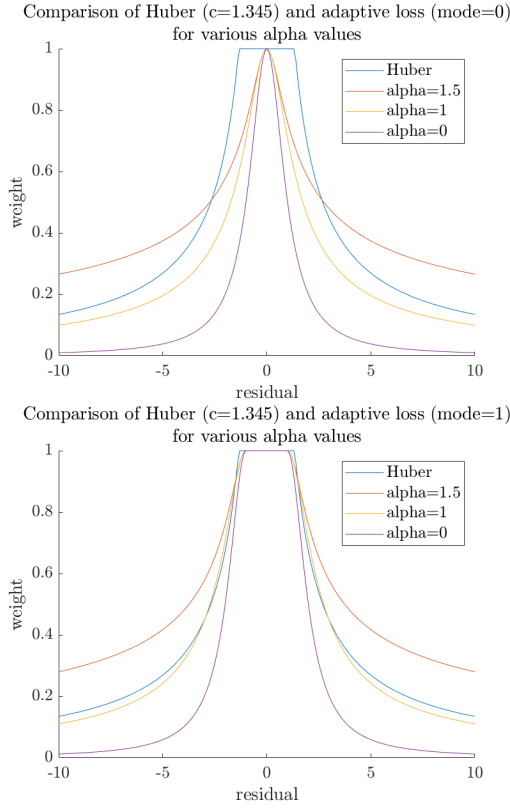


Fig. 1. Comparison of the Huber loss ρ_H ($c = 1.345$) and adaptive loss $\bar{\rho}$ functions. Top: $\varepsilon = 0$. Bottom: $\varepsilon = 1$.

However, we also see that $\tilde{w}(r, \alpha) < 1$ if $r \neq 0$, whatever the value of α . This is why the additional parameter ε was introduced in [19]: It defines an “inlier zone” in which the residuals are left unweighted, i.e. $\tilde{w}(r, \alpha, \varepsilon) = 1$ if $|r| < \varepsilon$. Their motivation was the fact that they considered residuals defined as vector norms, i.e. $r_i(\mathbf{x}) = \|\mathbf{r}_i(\mathbf{x})\|$ for some function $\mathbf{r}_i$. Indeed, in this case the residual distribution has a non-zero mode, therefore most inliers ended up strongly downweighted. It turns out that this second parameter can actually make sense even for 1D residuals. Indeed, as clearly appears in the bottom plot of Figure 1, it mimics parameter c of the Huber RLF (7), which has exactly the same role. Thus, for $\bar{\rho}$, ε and α define, respectively, which residuals are fully considered, and how quickly the others get downweighted.

IV. ADAPTIVE ROBUST KF-BASED FILTERS

A. Robust KF-based filters

In this section we recall how robust M-estimation based KFs are implemented [16]. Only the update step is impacted. Let $\hat{\mathbf{x}}_{t|t-1}$, $\mathbf{P}_{t|t-1}$ be the state prediction and prediction error covariance, respectively. It is now well-known that the standard KF update step computation of $\hat{\mathbf{x}}_{t|t}$ is equivalent to solving the following LS problem,

$$\hat{\mathbf{x}}_{t|t} = \arg \min_{\mathbf{x}_t} \left(\left\| \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1} \right\|_{\mathbf{P}_{t|t-1}}^2 + \left\| \mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t) \right\|_{\mathbf{R}_t}^2 \right). \quad (13)$$

Robust KF-based filters consider a modified optimisation, where some or all of the squared terms are replaced with a RLF ρ , as in Sec. III. Herein, outliers are only present in outputs, and then the considered optimisation problem becomes,

$$\hat{\mathbf{x}}_{t|t} = \arg \min_{\mathbf{x}_t} \left(\left\| \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1} \right\|_{\mathbf{P}_{t|t-1}}^2 + \rho \left(\mathbf{R}_t^{-1/2} (\mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t)) \right) \right), \quad (14)$$

with $\mathbf{R}_t^{-1/2}$ a Cholesky decomposition of $\mathbf{R}_t$. Solving (14) is commonly done through an IRLS scheme. At the i -th step, $\hat{\mathbf{x}}_{t|t}^{(i+1)}$ is computed through a modified version of (13):

$$\hat{\mathbf{x}}_{t|t}^{(i+1)} = \arg \min_{\mathbf{x}_t} \left(\left\| \mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1} \right\|_{\mathbf{P}_{t|t-1}}^2 + \left\| \mathbf{y}_t - \mathbf{h}_t(\mathbf{x}_t) \right\|_{\tilde{\mathbf{R}}_t^{(i)}}^2 \right), \quad (15)$$

with

$$\tilde{\mathbf{R}}_t^{(i)} = \mathbf{R}_t^{1/2} (\mathbf{W}_t^{(i)})^{-1} \mathbf{R}_t^{\top/2}, \quad (16)$$

$$\mathbf{W}_t^{(i)} = \text{diag} \left[w \left(\mathbf{R}_t^{-1/2} (\mathbf{y}_t - \mathbf{h}_t(\hat{\mathbf{x}}_{t|t}^{(i)})) \right) \right]. \quad (17)$$

where $w(\cdot)$ is the weight function obtained from the RLF ρ . To apply this methodology with the adaptive RLF $\bar{\rho}$, the main question then becomes: how do we choose ε and α ?

B. Choice of the RLF parameters

1) *State-of-the-art*: The RLF $\tilde{\rho}$, without mode ε , has been applied to both regression [21], [22] and filtering problems. Since α represents the distribution of the residuals, it needs to be estimated from them. Usually, the normalised, unitless, residuals are considered in regression, as for (17). Filtering methods using $\tilde{\rho}$ do not give further details.

Let $\mathbf{r}_t$ be the normalised residuals, then the parameter α is obtained by maximum likelihood estimation:

$$\alpha^* = \arg \min_{\alpha \in [-A, 2]} \left(n_y \log(Z(\alpha)) + \sum_{k=1}^{n_y} \tilde{\rho}(r_t^k, \alpha) \right), \quad (18)$$

where $-A$ is a chosen lower bound, and $Z(\alpha) = \int \exp(-\tilde{\rho}(r, \alpha)) dr$ is the normalisation constant. This approach is similar to the general one proposed in [20]. Note that, for $\alpha < 0$, the integral needs to be truncated since it diverges [22]. The additional mode parameter ε has only been used in a regression framework [19]. This was introduced to model a specific class of (inlier) residuals: those coming from the norm of vectorial errors. Indeed, these residuals follow a Maxwell-Boltzmann distribution, hence have a non-zero mode. In [19] authors propose a two-stage estimation process: first, ε is estimated independently, and then α is obtained using again a maximum likelihood approach.

2) *On the filtering residuals normalisation and mode*: It turns out that the normalisation procedure used to estimate α is not as straightforward as the one used for the weighting in (17). Indeed, consider a linear measurement model $\mathbf{h}_t(\mathbf{x}_t) = \mathbf{H}\mathbf{x}_t$. Then, the i -th residual is given by

$$r_t^i = y_i - \mathbf{h}_i^\top \hat{\mathbf{x}}_{t|t-1} = \mathbf{h}_i^\top (\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}) + n_t^i. \quad (19)$$

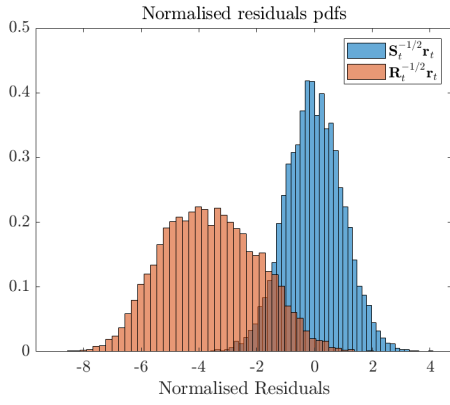


Fig. 2. Histograms of normalised residuals, for 10000 random measurements with Gaussian noise and without outliers, considering measurement covariance $\mathbf{R}_t$ or innovation covariance $\mathbf{S}_t$.

This highlights two facts:

- First, the residual covariance is not $\mathbf{R}_t$, but the innovation covariance $\mathbf{S}_t = \mathbf{H}\mathbf{P}_{t|t-1}\mathbf{H}^\top + \mathbf{R}_t$. Therefore, estimating the distribution shape α using $\mathbf{R}_t^{-1/2}\mathbf{r}_t$ may lead to overestimating the presence of outliers. This is illustrated in Figure 2, which compares for inlier residuals $\mathbf{r}_t$ both normalisations $\mathbf{R}_t^{-1/2}\mathbf{r}_t$ and $\mathbf{S}_t^{-1/2}\mathbf{r}_t$. It appears that, using the former, a too small α may be chosen. This may strongly downweight inliers, thus degrading the estimation. At the extreme, the filter slowly stops updating, as it sees all measurements as outliers.
- The second observation coming from (19) is that, from a filtering point of view, $\mathbf{r}_t$ is biased, due to the prediction error, $\mathbf{x}_t - \hat{\mathbf{x}}_{t|t-1}$. This is expected, since it is the purpose of the filtering update to correct this bias. However, this also highlights the interest of considering the mode parameter ε , and thus the adaptive RLF $\bar{\rho}$, in all cases, even with a set of 1D measurements.

3) *Proposed adaptive robust update step:* Taking into account the findings stated above, we advocate to 1) use residuals normalised with innovation covariance $\mathbf{S}_t$ to estimate α , but 2) keep residuals normalised by $\mathbf{R}_t$ for a consistent update procedure (15)-(17). Regarding the mode value ε , since we consider 1D measurements, there is no advantage using the Maxwell-Boltzmann approach of [19], as it would always yield $\varepsilon = 0$. Trying to estimate this parameter jointly with α using a maximum likelihood approach did not produce convincing results, while requiring heavier computations. Therefore, based on our experiments and considering the relation with the Huber loss shown in Figure 1, we advocate setting it to $\varepsilon = 1$.

We propose the following update procedure to build an adaptive robust KF based on $\bar{\rho}$:

- 1) set $\varepsilon = 1$ and estimate α^* from (18) using $\mathbf{r}_t = \mathbf{S}_t^{-1/2}(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1})$,
- 2) solve (15)-(17), using $\bar{w}$ with $\varepsilon = 1$, α^* and $\mathbf{R}_t^{-1/2}(\mathbf{y}_t - \mathbf{H}_t\hat{\mathbf{x}}_t^{(i)})$,

which we call Adaptive M-type KF (AMKF).

V. SIMULATION RESULTS

We illustrate the impact of using the adaptive RLF $\bar{\rho}$ in a KF framework using a simulated linear system, in order to eliminate other disturbance sources such as nonlinearities. We consider a constant velocity model with linearised pseudorange-type measurements,

$$\mathbf{x}_t = \begin{pmatrix} \mathbf{p}_t \\ \mathbf{v}_t \end{pmatrix}, \quad \mathbf{x}_t = \begin{bmatrix} \mathbf{I} & \Delta t \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \mathbf{x}_{t-1} + \begin{pmatrix} 0 \\ \mathbf{v}_{t-1} \end{pmatrix}, \quad (20)$$

$$\mathbf{y}_t = [\mathbf{H}_t^p \quad \mathbf{0}] \mathbf{x}_t + \mathbf{n}_t + \mathbf{o}_t, \quad (21)$$

with $\mathbf{H}_t^p$ is a set of random unit vectors; a Gaussian process noise $\mathbf{v}_t \sim \mathcal{N}(0, \text{diag}(0.1, 0.1, 0.01))$; a Gaussian nominal measurement noise $\mathbf{n}_t \sim \mathcal{N}(0, \sigma_y \mathbf{I})$, with $\sigma_y = 0.3$; and $\mathbf{o}_t$ is an additional noise accounting for possible outliers. The latter is present in three given time periods (i.e., the grey zones in Fig. 3) and affects a random proportion of measurements ranging from 0 to 90%:

- In the first outlier period, $\mathbf{o}_t \sim \mathcal{N}(0, 100\sigma_y \mathbf{I})$ is a white noise with large variance.
- In the second outlier period, $\mathbf{o}_t \sim \mathcal{N}(0, 5\sigma_y \mathbf{I})$ represents medium intensity outliers.
- In the third outlier period, $\mathbf{o}_t \sim \mathcal{ST}(0, \sigma_y^2, 20\sigma_y, 3)$ is a skewed t-distribution [9], [10], representing typical disturbances in navigation systems such as multipath.

The performance of the proposed AMKF is compared with the standard KF and the self-tuning MKF (based on [20]). For the latter, it turns out that the self-tuning MKF and the standard MKF with fixed $c = 1.345$ produces similar results. The position mean square error (MSE), obtained from 50 Monte Carlo runs, is shown in Fig. 3. Indeed, since the main objective of the AMKF is to estimate α , i.e., to fit a distribution, this requires to have enough data. To highlight what “enough” means in this example, results in Fig. 3 are given for $n_y = 10$ (top plot) and $n_y = 50$ (bottom plot). First, one can see that in the low dimension case $n_y = 10$, the AMKF fails as it is not able to correctly estimate α . On the other hand, in the large dimension case $n_y = 50$, the AMKF outperforms the MKF. In particular, notice that the AMKF maintains good performance even in the skewed noise period, where the MKF is not able to completely mitigate the impact of such outliers. Considering the third outlier region, the mean MSE for both AMKF and MKF and different n_y is shown in Fig. 4. For $n_y = 20$ both AMKF and MKF provide similar performance, but afterwards (increasing n_y) the AMKF takes advantage of the additional measurements and provides the best performance.

VI. CONCLUSIONS

The development of adaptive RLFs is an active research line for robust estimation, that is, to design estimation methods able to deal with observation outliers. In this contribution, we further explored the use of a new class of RLFs, recently proposed for regression problems, within a recursive filtering framework. Such application is not straightforward, in particular the normalisation stage. We proposed a way to implement an adaptive robust KF, and investigated the impact of the

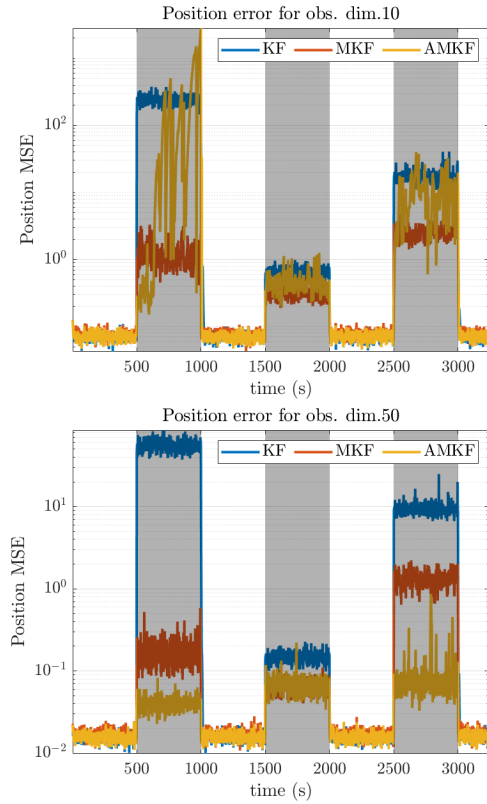


Fig. 3. Position estimation MSE of KF, MKF and AMKF. The zones in which outliers are present are depicted in grey. Top: For $n_y = 10$. Bottom: For $n_y = 50$.

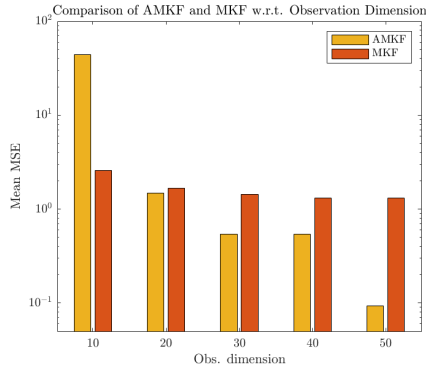


Fig. 4. Mean of the AMKF and MKF position MSE on the third outlier period for different values of $n_y = 10, 20, 30, 40$ and 50 .

measurement dimension in the filter performance. The impact of different types of outliers, and the capacity to mitigate them, were also illustrated with a numerical example. In conclusion, the new method is a promising approach for robust filtering.

REFERENCES

- [1] J. L. Crassidis and J. L. Junkins, *Optimal Estimation of Dynamic Systems*, CRC Press, Taylor & Francis Group, 2nd edition, 2012.
- [2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, Springer Science+Business Media, 2006.
- [3] D. Simon, *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*, John Wiley & Sons, 2013.

- [4] J. Duník, S. K. Biswas, A. G. Dempster, T. Pany, and P. Closas, "State Estimation Methods in Navigation: Overview and Application," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 35, no. 12, pp. 16–31, Dec. 2020.
- [5] I. Arasaratnam and S. Haykin, "Cubature Kalman Filters," *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, June 2009.
- [6] P. Djurić et al., "Particle Filtering," *IEEE Signal Processing Magazine*, vol. 140, no. 2, pp. 19–38, Sept. 2003.
- [7] J. Vilà-Valls, E. Chaumette, F. Vincent, and P. Closas, "Robust Linearly Constrained Kalman Filter for General Mismatched Linear State-Space Models," *IEEE Trans. Autom. Control*, vol. 67, no. 12, pp. 6794–6801, Dec. 2022.
- [8] P. Chauchat, D. Medina, J. Vilà-Valls, and E. Chaumette, "Robust Linearly Constrained Filtering for GNSS Position and Attitude Estimation under Antenna Baseline Mismatch," in *Proc. of FUSION*, 2021.
- [9] H. Nurminen, T. Ardesi, R. Piché, and F. Gustafsson, "Robust Inference for State-Space Models with Skewed Measurement Noise," *IEEE Signal Process. Lett.*, vol. 22, no. 11, pp. 1898–1902, May 2015.
- [10] J. Vilà-Valls and P. Closas, "NLOS Mitigation in Indoor Localization by Marginalized Monte Carlo Gaussian Smoothing," *EURASIP Journal on Adv. in Sig. Process.*, vol. 62, Aug. 2017.
- [11] Y. Huang, Y. Zhang, P. Shi, Z. Wu, J. Qian, and J.A. Chambers, "Robust Kalman Filters Based on Gaussian Scale Mixture Distribution with Application to Target Tracking," *IEEE Trans. Systems, Man. And Cybernetics*, vol. 49, no. 10, pp. 2082–2096, 2019.
- [12] Y. Huang, Y. Zhang, Y. Zhao, and J.A. Chambers, "A Novel Robust Gaussian-Student's t Mixture Distribution Based Kalman Filter," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3606–3620, 2019.
- [13] H. Wang, H. Li, J. Fang, and H. Wang, "Robust Gaussian Kalman Filter With Outlier Detection," *IEEE Signal Process. Lett.*, vol. 25, no. 8, pp. 1236–1240, 2018.
- [14] H. Li, D. Medina, J. Vilà-Valls, and P. Closas, "Robust Variational-Based Kalman Filter for Outlier Rejection with Correlated Measurements," *IEEE Trans. Signal Process.*, vol. 69, pp. 357–369, 2020.
- [15] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, Eds., *Robust Statistics for Signal Processing*, Cambridge University Press, London, UK, 2018.
- [16] D. Medina, H. Li, J. Vilà-Valls, and P. Closas, "Robust Filtering Techniques for RTK Positioning in Harsh Propagation Environments," *Sensors*, vol. 21, no. 4, pp. 1250, 2021.
- [17] A. Bellés, D. Medina, P. Chauchat, and J. Vilà-Valls, "Reliable GNSS Joint Position and Attitude Estimation in Harsh Environments via Robust Statistics," in *Proc. of the IEEE Aerospace Conference*, 2022.
- [18] A. Bellés, D. Medina, P. Chauchat, S. Labsir, and J. Vilà-Valls, "Robust M-Type Error-State Kalman Filters for Attitude Estimation," in *Proc. of EUSIPCO*, 2023.
- [19] T. Hitchcox and J. R. Forbes, "Mind the Gap: Norm-Aware Adaptive Robust Loss for Multivariate Least-Squares Problems," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 7116–7123, 2022.
- [20] G. Agamennoni, P. Furgale, and R. Siegwart, "Self-Tuning M-Estimators," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 4628–4635.
- [21] J. T. Barron, "A General and Adaptive Robust Loss Function," in *Proc. of the IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019.
- [22] N. Chebrolu, T. Labe, O. Vysotska, J. Behley, and C. Stachniss, "Adaptive Robust Kernels for Non-Linear Least Squares Problems," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 2240–2247, 2021.
- [23] Y. Ding, P. Chauchat, G. Pages, and P. Asseman, "Learning-Enhanced Adaptive Robust GNSS Navigation in Challenging Environments," *IEEE Robot. Autom. Lett.*, vol. 7, no. 4, pp. 9905–9912, 2022.
- [24] S. Chen, Q. Zhang, D. Lin, and S. Wang, "Generalized Loss Based Geometric Unscented Kalman Filter for Robust Power System Forecasting-Aided State Estimation," *IEEE Signal Process. Lett.*, vol. 29, pp. 2353–2357, 2022.
- [25] W. Yan, S. Chen, D. Lin, and S. Wang, "Variational Bayesian-Based Generalized Loss Cubature Kalman Filter," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, 2024.