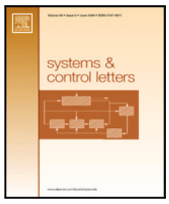
Systems & Control Letters 204 (2025) 106198

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle



Stable kernel size adaptation-based maximum correntropy Kalman filter

Vahid Azimi [a,*], Simona Onori [b]

[a] Gatik, Mountain View, CA, USA
[b] Department of Energy Resources Engineering, Stanford University, Stanford, CA, USA

ARTICLE INFO

Keywords:
Adaptive maximum correntropy Kalman filter
Kernel size adaptation
Uniform ultimate boundedness
Deterministic Lyapunov stability analysis

ABSTRACT

The state-of-the-art maximum correntropy Kalman filter (MCKF) that incorporates Gaussian kernels into its formulation is a powerful estimation technique to robustify the traditional Kalman filter (KF) against non-Gaussian noises. However, improper selection of *kernel sizes* may either degrade the estimation performance or slow down the convergence rate. This paper extends and encompasses the baseline MCKF through the development of a projection-based adaptive MCKF (PAMCKF) for continuous-time linear systems with a class of non-Gaussian noises: measurement outliers that are mainly due to sensor failures, cyber-attacks, etc We suggest an adaptation mechanism that exploits the innovation signals to update "on-line" the *kernel sizes*. The proposed adaptive method is then unified with a modified version of MCKF to achieve robust estimation with low computational effort against measurement outliers without the need for knowing their bounds *a priori*. The main contribution of this paper is that state and *kernel size* estimation errors are uniformly ultimately bounded which is formally proven using a deterministic Lyapunov stability argument. Simulations and comparisons to KF and MCKF are carried out to validate the theoretical results and demonstrate the benefits of the proposed filter.

1. Introduction

The Kalman filter (KF) [1] is still the most effective tool for estimating the states of linear systems with noise processes. Despite the broad literature covering the KF applications [2–4], a current challenge is to address estimation in the presence of non-Gaussian measurement noises. In real-time applications, measurements could be impacted by large deviations due to sensor faults (e.g., bias, drift, and abrupt failure), wrong replacement of measure, cyber-attacks, etc. These annoying signals are called *outliers* in different fields including, but not limited to, power system [5], heart surgery [6], system identification [7], fault detection [8], cloud management [9], state estimation [10], and vehicle system [11,12]. The outliers lead to measurement noises to be non-Gaussian; this, in turn, not only degrades the KF estimates but also delegitimizes convergence proofs [13]. Over the years, an enormous amount of research has been conducted to enhance the robustness of Kalman-type filters regardless of noise models. In general, these approaches can be divided into three groups: (i) Filters with non-Gaussian noise distributions such as heavy-tailed and t-distributions [14,15], (ii) Multiple-model filters that approximate a non-Gaussian distribution with a finite sum of Gaussian distributions [16], and (iii) Monte Carlo filters, including particle filters [17] and unscented Kalman filter [18], that approximate the posterior distribution of states with a set of random samples with associated weights. Although these approaches improve the estimation performance against non-Gaussian noises, they add computational complexity that hinders their real-time implementations.

Maximum correntropy Kalman filter (MCKF) is a recently developed estimation technique that can robustify the KF against non-Gaussian measurement noises while keeping computational requirements low [13,19–23]. In contrast to the ordinary KF that only uses the second-order information from the measurements (that is due to the use of quadratic objective function), MCKF utilizes the information from higher-order signal statistics by incorporating Gaussian kernels into formulation to improve the state estimation. However, under MCKF, the improper choice of *kernel sizes* either degrades the estimation performance or slows down the rate of convergence. In particular, in [13], a fixed-point algorithm was utilized to update the posterior estimations based on which a sufficient condition was found to guarantee the convergence of the fixed-point iterations in MCKF. It was proved that the convergence of MCKF is heavily dependent on the choice of the *kernel size* such that the convergence is ensured only if the *kernel size* is larger than a certain value; if the *kernel size* is too small, the algorithm presented in [13] will diverge or converge very slowly. To mitigate this issue, few attempts have been recently made suggesting *heuristic* algorithms for updating the *kernel sizes* [24–29]. Despite the usefulness, those algorithms lack a formal guarantee of stability.

* Corresponding author.
E-mail address: vahid.azimi@gatik.ai (V. Azimi).

https://doi.org/10.1016/j.sysconle.2025.106198
Received 3 April 2024; Received in revised form 20 May 2025; Accepted 29 June 2025
Available online 16 July 2025

To the best of our knowledge, to date, no study has focused on introducing an MCKF that can select the optimal *kernel sizes* "on-line" under which the stability of the system is formally ensured. To address this gap, the current paper extends and encompasses the baseline MCKF to continuous-time linear systems whose measurements are disturbed by unknown large outliers (systems with state equations impacted by the outliers are beyond the scope of this paper). Our main contributions are thus twofold: (i) The unification of MCKF with a *kernel size* adaptation mechanism to cope with unknown measurement outliers, and (ii) The proof of uniform ultimate boundedness of state and kernel size estimation errors through a rigorous deterministic Lyapunov stability analysis.

In this paper, we begin by revisiting the KF and then discuss the impact of measurement outliers on its estimation performance. As a remedy for heuristic tuning of the *kernel sizes*, a projection-based adaptation mechanism is suggested that updates them by the use of the *innovations*. The proposed adaptation law is then unified with a modified version of MCKF to derive the projection-based adaptive MCKF (PAMCKF) whose solutions are uniformly ultimately bounded. This is formally proven by employing a deterministic Lyapunov argument. Simulations and comparisons to KF and MCKF support the main results.

The organization of the paper is as follows. Section 2 provides the notations and definitions used in this paper. Section 3 formulates the proposed filter. Section 4 provides the stability analysis. Section 5 presents simulation results. Section 6 concludes the paper and provides suggestions for future works.

Acronyms	
CARE	Continuous algebraic Riccati equation
DRE	Differential Riccati equation
KF	Kalman filter
MCKF	Maximum correntropy KF
PAMCKF	Projection-based adaptive MCKF
RMS	Root-mean-square
MMSE	Minimum mean square error
MC	Monte Carlo

2. Preliminaries

This section provides an overview of basic preliminaries, definitions, and lemmas that will be later used when developing the proposed filter.

Given a real n-dimensional vector x, the $\mathcal{L}_2$-norm of x is defined by $\|x\| = \sqrt{x^T x}$. The root-mean-square (RMS) of x is given by $\text{RMS}_x = \frac{1}{\sqrt{n}}\|x\|$. $E(\cdot)$ is the expectation operation. $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ denote the minimal and maximal eigenvalues of a matrix, respectively, whose eigenvalues are real.

Definition 1. The solutions of $\dot{x}(t) = f(x(t), t)$ with the continuous function $f : \mathfrak{R}^n \times [0, \infty) \longrightarrow \mathfrak{R}^n$ are uniformly ultimately bounded with ultimate bound b if $\exists\ b, c > 0$ and for every $a \in (0,\ c)$, $\exists\ T = T(a, b) > 0$ such that $\|x(t_0)\| \le a \Rightarrow \|x(t)\| \le b,\ \forall\ t_0 \ge 0$ and $t \ge t_0 + T$ [30].

Lemma 1. *Let $g(x) : \mathfrak{R}^n \longrightarrow \mathfrak{R}$ be a convex function. Then, the subset $\Omega_\lambda = \{x \in \mathfrak{R}^n | g(x) \le \lambda\}$ is convex for any constant $\lambda > 0$.*

Lemma 2. *Let $g(x) : \mathfrak{R}^n \longrightarrow \mathfrak{R}$ be a continuously differentiable convex function. Also, let x^* be an interior point of Ω_λ such that $g(x^*) < \lambda$ and x be a boundary point of Ω_λ such that $g(x) = \lambda$. Then, one holds $(x^* - x)^T \nabla g(x) \le 0$, where $\nabla g(x) = \left(\frac{\partial g(x)}{\partial x_1} \dots \frac{\partial g(x)}{\partial x_n}\right)^T$ is the gradient vector of $g(x)$ evaluated at x. This implies that the gradient evaluated at the boundary always points away from the subset Ω_λ.*

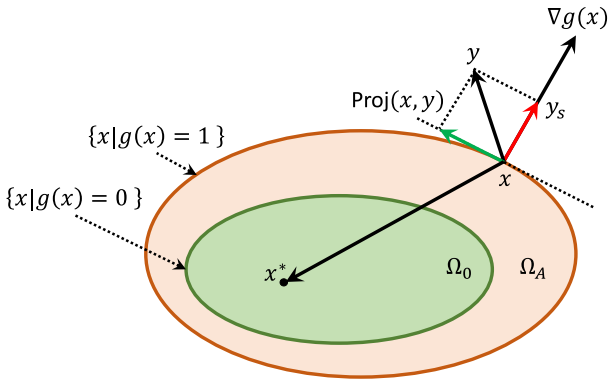


Fig. 1. Projection operator in $\mathfrak{R}^2$ defined by Eq. (1).

Definition 2. The projection operator for two vectors $x, y \in \mathfrak{R}^n$ is defined as [31]

$$\text{Proj}(x, y) = \begin{cases} y - \underbrace{\dfrac{\nabla g(x)(\nabla g(x))^T}{\|\nabla g(x)\|^2} y}_{y_s}\ g(x), \\ \text{if } g(x) > 0 \ \wedge\ y^T \nabla g(x) > 0 \\ y, \quad \text{otherwise}, \end{cases} \tag{1}$$

where $g(x) = \frac{x^T x - x_{max}^2}{\nu_x x_{max}^2} : \mathfrak{R}^n \longrightarrow \mathfrak{R}$ is a continuously differentiable convex function; x_{max} is the norm bound enforced on x; and $\nu_x > 0$ is the projection tolerance.

To provide a geometrical interpretation of (1), let us define the convex sets $\Omega_0 = \{x \in \mathfrak{R}^n | g(x) \le 0\}$ and $\Omega_1 = \{x \in \mathfrak{R}^n | g(x) \le 1\}$ from which one holds $\Omega_0 \subset \Omega_1$. From (1), y is not altered when $x \in \Omega_0$. However, in the set $\Omega_A = \{x \in \mathfrak{R}^n | 0 < g(x) \le 1\}$, if $y^T \nabla g(x) > 0$, the projection operator subtracts a scaled component of y, named y_s, that is normal to the boundary $\{x|g(x) = \kappa\}$, from y; there is a smooth transformation from $\text{Proj}(x, y) = y$ when $\kappa = 0$ to $\text{Proj}(x, y)$ tangent to the boundary $\{x|g(x) = 1\}$ when $\kappa = 1$, where y_s, normal to the boundary $\{x|g(x) = 1\}$, is completely subtracted from y. This concept is illustrated in Fig. 1.

The projection operator (1) has the following properties: (i) if $g(x) \le 1$, one holds

$$\|x\| \le x_{max}\sqrt{1 + \nu_x} \tag{2}$$

and (ii) given $x^* \in \Omega_0$, one ensures

$$-\tilde{x}^T \big(y - \text{Proj}(x, y)\big) \le 0, \tag{3}$$

where $\tilde{x} = x - x^*$.

3. Proposed filter

This section is comprised of two distinct subsections. Section 3.1 revisits the traditional KF and discusses the impact of measurement outliers on its performance. Section 3.2 derives the novel PAMCKF through the unification of a *kernel size* adaptation mechanism with a modified version of MCKF presented in [19].

3.1. Kalman Filter (KF)

Consider the following continuous-time linear system

$$\begin{aligned} &\dot{x}(t) = Ax(t) + Bu(t) + w(t) \\ &y(t) = \bar{y}(t) + Du(t) \quad \text{with} \quad \bar{y}(t) = Cx(t) + v(t) \\ &w(t) \sim (0, Q) \quad \text{and} \quad v(t) \sim (0, R), \end{aligned} \tag{4}$$

where $x(t) \in \mathfrak{R}^n$ is the vector of the state variables; $u(t) \in \mathfrak{R}^m$ is the vector of the inputs; $y(t) \in \mathfrak{R}^l$ is the vector of the measurements; $w(t) \in \mathfrak{R}^n$ and $v(t) \in \mathfrak{R}^l$ are the continuous-time zero-mean process and measurement noises with covariance matrices $Q \in \mathfrak{R}^{n\times n} \succ 0$ and $R \in \mathfrak{R}^{l\times l} \succ 0$, respectively; and $A \in \mathfrak{R}^{n\times n}$, $B \in \mathfrak{R}^{n\times m}$, $C \in \mathfrak{R}^{l\times n}$, and $D \in \mathfrak{R}^{l\times m}$ are the state transition, input, measurement, and feedthrough matrices, respectively.

For the system (4), the continuous-time KF equations are given as [32][Chapter 8, Section 8.2]

$$\hat{x}(0) = E[x(0)] \tag{5}$$

$$P(0) = E\left[e(0)(e(0))^T\right] \tag{6}$$

$$K(t) = P(t)C^T R^{-1} \tag{7}$$

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(t)r(t) \tag{8}$$

$$\dot{P}(t) = AP(t) + P(t)A^T + Q - K(t)CP(t), \tag{9}$$

where $\hat{x}(t) \in \mathfrak{R}^n$ is the estimate of $x(t)$; $x(0) \in \mathfrak{R}^n$ and $\hat{x}(0) \in \mathfrak{R}^n$ are the initial values of the state and its estimate, respectively; the positive definite matrix $P(t) \in \mathfrak{R}^{n\times n}$ is the covariance of the estimation error $e(t) = \hat{x}(t) - x(t)$; $P(0) \in \mathfrak{R}^{n\times n}$ is the covariance of the initial estimation error $e(0) = \hat{x}(0) - x(0)$; $K(t) \in \mathfrak{R}^{n\times l}$ is the Kalman gain; and $r(t)$ is the difference between the measurement $y(t)$ and the prediction of the measurement $\hat{y}(t) = \bar{\hat{y}}(t) + Du(t)$ with $\bar{\hat{y}}(t) = C\hat{x}(t)$ that is also called the *innovations* defined as $r(t) = \bar{y}(t) - C\hat{x}(t)$.

Assumption 1. We assume that (A, C) is detectable and $(A, Q^{\frac{1}{2}})$ is controllable in the closed left half plane [32][Chapter 8, Theorem 29].

Lemma 3. *Due to Assumption 1, $P(t)$ for some $P(0) > 0$ in (9) will reach a steady-state value, resulting in the continuous algebraic Riccati equation (CARE): $AP + PA^T + Q - K(t)CP = 0$ that has at least one positive definite solution P [32][Chapter 8, Theorem 29].*

Since KF is derived based upon the minimum mean square error (MMSE) criterion (it utilizes only second-order information of signal), it is an unbiased filter which minimizes the trace of P (the KF is optimal) only when the noises w and v are both uncorrelated and Gaussian [32]. The existence of measurement outliers, as a specific class of non-Gaussian noises, could potentially degrade the performance of the traditional KF that loses the optimality in this case. To mitigate this issue, the next section is devoted to develop an adaptive KF, that uses the information from higher-order signal statistics, to robustify the baseline KF when measurements are impacted by large outliers.

3.2. Projection-based adaptive maximum correntropy Kalman Filter (PAMCKF)

This subsection describes the PAMCKF. With the goal of developing an adaptive KF with uniformly ultimately bounded solutions, we present the first steps towards unifying a projection-based adaptation mechanism with the baseline MCKF [19] for systems with measurement outliers without the need for knowing their bounds *a priori*.

Consider the system (4), where the measurement $y(t)$ is disturbed by an unknown but bounded outlier $d(t) \in \mathfrak{R}^l$

$$\begin{aligned} &\dot{x}(t) = Ax(t) + Bu(t) + w(t)\\ &y_d(t) = \bar{y}_d(t) + Du(t) \quad \text{with} \quad \bar{y}_d(t) = \bar{y}(t) + d(t)\\ &w(t) \sim (0, Q) \quad \text{and} \quad v(t) \sim (0, R), \end{aligned} \tag{10}$$

where $y_d(t)$ is the *disturbed measurement* that is available instantaneously for the later use in the PAMCKF algorithm.

Assumption 2. The unknown measurement outlier $d(t)$ is uniformly bounded for all $t \geq 0$ such that $\|d(t)\| \leq \bar{d}$ for some unknown upper bound $\bar{d} > 0$. Note that the above assumption assumes only the boundedness of $d(t)$ so that the value of $\bar{d}$ is not required for use in the proposed filter [33].

Assumption 3. The process and measurement noises w and v are assumed to be uncorrelated (the covariance matrices Q and R are diagonal), Gaussian, and zero-mean [32]. However, the outlier $d(t)$ is enforced to the measurement noise v to model a class of non-Gaussian noises – measurement outliers – defined as

$$\begin{aligned} &w(t) \sim \mathcal{N}(0, Q)\\ &v(t) \sim \mathcal{N}(0, R) + d(t) \end{aligned} \tag{11}$$

for which the notation $\mathcal{N}$ is used to indicate that w and v are Gaussian.

For the system (10), the continuous-time MCKF equations are suggested as

$$\hat{x}(0) = E[x(0)] \tag{12}$$

$$P(0) = E\left[e(0)(e(0))^T\right] \tag{13}$$

$$K(t) = P(t)C^T S(t)R^{-1} \tag{14}$$

$$r_d(t) = \bar{y}_d(t) - C\hat{x}(t) \tag{15}$$

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(t)r_d(t) \tag{16}$$

$$\dot{P}(t) = AP(t) + P(t)A^T + Q - K(t)CP(t), \tag{17}$$

where $r_d(t)$ is the *disturbed innovations* and $S \in \mathfrak{R}^{l\times l}$ is the correntropy matrix

$$S(t) = \begin{bmatrix} \mathcal{G}_\sigma\left(\frac{|r_{d_1}|}{\sqrt{R_{11}}}\right) & & 0\\ & \ddots & \\ 0 & & \mathcal{G}_\sigma\left(\frac{|r_{d_l}|}{\sqrt{R_{ll}}}\right) \end{bmatrix} \tag{18}$$

for which the $r_{d_i}(t)$ and the Gaussian kernel $\mathcal{G}_\sigma\left(|r_{d_i}|/\sqrt{R_{ii}}\right)$ are defined as

$$r_{d_i}(t) = \bar{y}_{d_i}(t) - C_i\hat{x}(t) \tag{19}$$

$$\mathcal{G}_\sigma\left(\frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right) = \exp\left(-\frac{r_{d_i}^2(t)}{2\mu_i^2 R_{ii}}\right), \tag{20}$$

where $\bar{y}_{d_i}(t)$, C_i, and R_{ii} for $i = 1, \ldots, l$ are the ith element of $\bar{y}_d(t)$, the ith row of the measurement matrix C, and the ith diagonal element of the matrix R, respectively; and $\mu_i > 0$ for $i = 1, \ldots, l$ is the *kernel size*.

Remark 1. Noting that μ_i is a positive scalar, one has $0 < \mathcal{G}_\sigma(.) \leq 1$ for all r_{d_i} with $i = 1, \ldots, l$, leading to S to be always a positive definite matrix. Therefore, the property stated in Lemma 3 also applies to the differential Riccati equation (DRE) (17) (that is, the CARE has $P > 0$); it is because (17) resembles (9) except that the proposed DRE (17) uses the Kalman gain (14) that replaces the positive definite matrix SR^{-1} with R^{-1}.

The performance of the baseline MCKF is degraded by improper selection of the *kernel size* μ_i such that (i) When μ_i is large, $S \to I$ and MCKF reduces to the traditional KF; and (ii) Small μ_i has the effect of slowing down the convergence rate and even leads to the filter that does not update. This is due to the fact that when μ_i is too small, then $S, K \to 0$ which, in turn, implies that the filter ignores the information provided by the *disturbed innovations* r_d; the estimates cannot be improved or estimation rate becomes slow.

Thus, to maintain the estimation performance in the presence of $d(t)$, a proper selection of μ_i is required. However, the optimal value of the *kernel sizes* is not known *a priori* in real-world applications subject to unknown measurement outliers. To tackle this issue and relieve the engineer of the need to manually tune the *kernel sizes* against

unknown outliers $d(t)$ for different practical systems, the following projection-based adaptation mechanism is suggested

$$\dot{\sigma}_i(t) = \sigma_{0_i}\left(\text{Proj}\left(\sigma_i(t), \frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right) - \eta_i\sigma_i(t)\right) \tag{21}$$

with initial condition $\sigma_i(t=0) = \sigma_i(0)$, where $\sigma_i = 1/\mu_i$, $\sigma_{0_i}, \eta_i > 0$ for $i = 1, \dots, l$, and Proj(.,.) is realized by Eq. (1) for the scalar arguments $\sigma_i, |r_{d_i}|/\sqrt{R_{ii}} \in \mathfrak{R}$. The initial value of μ_i is picked by the user as an initial guess of the *kernel size*. The above adaptation law comprises two terms. The first term is a projection-based adaptation law the sole usage of which ensures $\sigma_i(t) \in \Omega_1 \ \forall t \geq 0$ if $\sigma_i(0) \in \Omega_1$ and $\sigma_{0_i} = 1$. The second term is a *sigma-modification* term [34] that adds damping to the first term. Thus, the adaptation mechanism becomes a linear time-varying ordinary differential equation with bounded input whose solution σ_i is always bounded by $\max\left(\frac{x_{max_i}\sqrt{1+v_{x_i}}}{\eta_i}, \sigma_i(0)\right)$ with the tightest asymptotic bound $\frac{x_{max_i}\sqrt{1+v_{x_i}}}{\eta_i}$. Thus, the mechanism provides a safeguard by bounding σ_i to prevent instability.

Remark 2. The adaptation parameters σ_{0_i}, x_{max_i}, v_{x_i}, and η_i are tuning parameters to be set for different applications. The exponential decay rate is $\sigma_{0_i}\eta_i$, and choosing a larger value results in faster decay of σ_i. Moreover, by choosing a larger η_i and a smaller x_{max_i}, the tightest asymptotic bound becomes smaller.

The proposed algorithm defined by Eqs. (12)–(21) extends and encompasses the baseline MCKF [19] in such a way that when large measurement outliers are encountered, the diagonal elements of the matrix S exponentially decrease to avoid the divergence of the filter while the decaying rate is adjusted by the proposed adaptation mechanism (21). Thus, PAMCKF estimates the states while rejecting the measurement outlier $d(t)$. The next section guarantees the uniform ultimate boundedness of all system solutions $(e(t), \tilde{\sigma}(t))$ under the proposed PAMCKF in the presence of unknown but bounded outliers $d(t)$.

Remark 3. The overall asymptotic computational complexity of KF is $O(n^3 + n^2 m)$. In the proposed PAMCKF, the inclusion of the correntropy matrix S and the adaptation mechanism (21) both incur a computational cost of $O(m)$, which is negligible compared to the dominant covariance update term. Thus, despite the extensions, the proposed filter retain the same overall asymptotic computational complexity of $O(n^3 + n^2 m)$, both theoretically and in practice.

4. Stability analysis

To show the uniform ultimate boundedness of all system solutions $(e, \tilde{\sigma})$, let us first use the system (10), the dynamics (16), and the *disturbed innovations* (15) to construct the estimation error dynamics

$$\begin{aligned}\dot{e} =& A\hat{x} + Bu + Kr_d - \big(Ax + Bu + w\big)\\ =& \big(A - KC\big)e + K\big(d + v\big) - w,\end{aligned} \tag{22}$$

where throughout this section, we drop the argument t for ease of exposition.

To conduct a deterministic stability analysis, throughout this section, the following assumption is made which is common when performing stability analyses of Kalman filter families in a deterministic setting [35].

Assumption 4. We assume that $w = v = 0$, which implies that the matrices Q^{-1} and R^{-1} are the confidence in the system model and measurements, respectively, but not the covariances of the process and measurement noises w and v. This leads to the stability analysis to be performed in a deterministic setting.

From the error dynamics (22) and the adaptation mechanism (21), the following scalar positive definite Lyapunov function is suggested

$$V(e, \tilde{\sigma}) = e^T P^{-1} e + 0.5\tilde{\sigma}^T H^{-1}\tilde{\sigma}, \tag{23}$$

where $\tilde{\sigma} = \sigma - \sigma^*$ with $\sigma = [\sigma_1, \dots, \sigma_l]^T \in \mathfrak{R}^l$ and σ^* as the true value of σ; and $H = \text{diag}(\sigma_{0_1}, \dots, \sigma_{0_l}) > 0 \in \mathfrak{R}^{l\times l}$. Note that only the existence of σ^* is assumed; true knowledge of σ^* is not required.

Theorem 1. *Consider the Lyapunov function (23) and the PAMCKF defined by Eqs. (12)–(21). Under Assumptions 1–4 and Lemma 3, the proposed filter ensures the uniform ultimate boundedness of $(e, \tilde{\sigma})$ for all initial conditions $(e(0), \tilde{\sigma}(0)) \in \mathfrak{R}^{n+l}$ and any bounded outlier $d(t) \in \mathfrak{R}^l$ without knowing its bound a priori.*

Proof. Taking the time derivative of the Lyapunov function (23) yields

$$\dot{V} = 2e^T P^{-1}\dot{e} + e^T \frac{d}{dt}(P^{-1})e + \tilde{\sigma}^T H^{-1}\dot{\sigma}. \tag{24}$$

Using the matrix inequality

$$\frac{d}{dt}(PP^{-1}) = \dot{P}P^{-1} + P\frac{d}{dt}(P^{-1}) = 0 \tag{25}$$

we obtain

$$\dot{V} = 2e^T P^{-1}\dot{e} - e^T(P^{-1}\dot{P}P^{-1})e + \tilde{\sigma}^T H^{-1}\dot{\sigma} \tag{26}$$

for which from (17), (21), and (22), it yields

$$\begin{aligned}\dot{V} =& 2e^T P^{-1}\Big((A - KC)e + Kd\Big)\\ &- e^T P^{-1}\Big(AP + PA^T + Q - KCP\Big)P^{-1}e\\ &+ \sum_{i=1}^{l}\tilde{\sigma}_i\left(\text{Proj}\left(\sigma_i, \frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right) - \eta_i\sigma_i\right).\end{aligned} \tag{27}$$

By expanding all the terms, incorporating the Kalman gain K from (14), and canceling the similar terms, one can write

$$\begin{aligned}\dot{V} =& -e^T Me + 2e^T C^T SR^{-1}d\\ &+ \sum_{i=1}^{l}\tilde{\sigma}_i\left(\text{Proj}\left(\sigma_i, \frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right) - \eta_i\sigma_i\right),\end{aligned} \tag{28}$$

where $M = C^T SR^{-1}C + P^{-1}QP^{-1}$ is a positive definite matrix. Noting that $0 < \mathcal{G}_\sigma(.) \leq 1$ for all r_{d_i} with $i = 1, \dots, l$, the operator norm[1] of S has the property that

$$\|S\|_{OP} \leq 1 \tag{29}$$

using which one can obtain

$$\begin{aligned}\dot{V} \leq& -e^T Me + 2\|e\|\|C^T R^{-1}\|\|d\|\\ &+ \sum_{i=1}^{l}\tilde{\sigma}_i\left(\text{Proj}\left(\sigma_i, \frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right) - \eta_i\sigma_i\right).\end{aligned} \tag{30}$$

Now, using the projection operator (1) in case of its first condition, one has

$$\begin{aligned}\dot{V} \leq& -e^T Me + 2\|e\|\|C^T R^{-1}\|\|d\|\\ &+ \sum_{i=1}^{l}\tilde{\sigma}_i\Big(\frac{|r_{d_i}|}{\sqrt{R_{ii}}} - \frac{|r_{d_i}|}{\sqrt{R_{ii}}}g(\sigma_i) - \eta_i\sigma_i\Big).\end{aligned} \tag{31}$$

Using the property (3), on can say

$$-\tilde{\sigma}_i\frac{|r_{d_i}|}{\sqrt{R_{ii}}}g(\sigma_i) = -\tilde{\sigma}_i\left(\frac{|r_{d_i}|}{\sqrt{R_{ii}}} - \text{Proj}\left(\sigma_i, \frac{|r_{d_i}|}{\sqrt{R_{ii}}}\right)\right) \leq 0$$

[1] For a matrix A, the operator norm $\|A\|_{OP}$ is the square root of the largest eigenvalue of $A^T A$, where A^T is A's transpose.

using which Eq. (31) becomes

$$\dot{V} \leq -e^T M e + 2\|e\|\|C^T R^{-1}\|\|d\| + \sum_{i=1}^{l} \tilde{\sigma}_i \Big(\frac{|r_{d_i}|}{\sqrt{R_{ii}}} - \eta_i \sigma_i \Big). \tag{32}$$

Obviously, under the second condition of (1), again, (30) becomes (32).

By using the relations $r_{d_i} = d_i - C_i e$ and $\sigma_i = \tilde{\sigma}_i + \sigma_i^*$, one obtains

$$\dot{V} \leq -\lambda_{\min}(M)\|e\|^2 - \lambda_{\min}(\eta)\|\tilde{\sigma}\|^2 + 2\|e\|\|C^T R^{-1}\|\bar{d} + \|\eta\|\|\tilde{\sigma}\|\|\sigma^*\| + \sum_{i=1}^{l} \tilde{\sigma}_i \Big(\frac{|d_i - C_i e|}{\sqrt{R_{ii}}} \Big), \tag{33}$$

where $\eta = \text{diag}(\eta_1, \dots, \eta_l) \in \Re^{l\times l}$.

The last term can be bounded above using norms and the triangle inequality to have

$$\dot{V} \leq -\lambda_{\min}(M)\|e\|^2 - \lambda_{\min}(\eta)\|\tilde{\sigma}\|^2 + 2\|e\|\|C^T R^{-1}\|\bar{d} + \|\eta\|\|\tilde{\sigma}\|\|\sigma^*\| + \sum_{i=1}^{l} \frac{\tilde{\sigma}_i}{\sqrt{R_{ii}}} \Big(\|d_i\| + \|C_i\|\|e\| \Big). \tag{34}$$

Let us define $\bar{\bar{d}} = [\|d_1(t)\|, \dots, \|d_l(t)\|]^T \in \Re^l$ and $\bar{\bar{C}} = [\|C_1\|, \dots, \|C_l\|]^T \in \Re^l$ to write

$$\dot{V} \leq -\lambda_{\min}(M)\|e\|^2 - \lambda_{\min}(\eta)\|\tilde{\sigma}\|^2 + 2\|e\|\|C^T R^{-1}\|\bar{d} + \|\eta\|\|\tilde{\sigma}\|\|\sigma^*\| + \tilde{\sigma}^T R^{-\frac{1}{2}} \bar{\bar{d}} + \tilde{\sigma}^T R^{-\frac{1}{2}} \bar{\bar{C}}\|e\| \tag{35}$$

that reduces to

$$\dot{V} \leq -\lambda_{\min}(M)\|e\|^2 - \lambda_{\min}(\eta)\|\tilde{\sigma}\|^2 + 2\|e\|\|C^T R^{-1}\|\bar{d} + \|\eta\|\|\tilde{\sigma}\|\|\sigma^*\| + \|\tilde{\sigma}\|\|R^{-\frac{1}{2}}\|\|\bar{\bar{d}}\| + \|\tilde{\sigma}\|\|R^{-\frac{1}{2}}\|\|\bar{\bar{C}}\|\|e\|. \tag{36}$$

For sufficiently large $\lambda_{\min}(M)$ and $\lambda_{\min}(\eta)$, $\dot{V} \leq 0$ everywhere outside of a compact set. To show that the set is actually compact, note that (36) is quadratic in $\|e\|$ and $\|\tilde{\sigma}\|$. We can then derive conservative estimates for the positively invariant set (within which the errors are bounded) by solving the quadratic inequalities for each variable at a time. Beginning with $\|e\|$, the bound that enforces $\dot{V} \leq 0$ is derived in (37). We then consider $\|\tilde{\sigma}\|$ to define the bound (38) to enforce $\dot{V} \leq 0$. By doing so, the compact set outside of which $\dot{V} \leq 0$ is derived by

$$\|e\| \geq \frac{\Phi_1 + \sqrt{\Phi_1^2 + 4\lambda_{\min}(M)\Phi_2}}{2\lambda_{\min}(M)} \triangleq B_e \tag{37}$$

$$\|\tilde{\sigma}\| \geq \frac{\Phi_3 + \sqrt{\Phi_3^2 + 4\lambda_{\min}(\eta)\Phi_4}}{2\lambda_{\min}(\eta)} \triangleq B_{\tilde{\sigma}}, \tag{38}$$

where

$$\begin{aligned} \Phi_1 &= 2\|C\|\|R^{-1}\|\bar{d} + \|\tilde{\sigma}\|\|R^{-\frac{1}{2}}\|\|\bar{\bar{C}}\| \\ \Phi_2 &= \|\tilde{\sigma}\| \left(-\lambda_{\min}(\eta)\|\tilde{\sigma}\| + \|\eta\|\|\sigma^*\| + \|R^{-\frac{1}{2}}\|\|\bar{\bar{d}}\| \right) \\ \Phi_3 &= \|\eta\|\|\sigma^*\| + \|R^{-\frac{1}{2}}\| \left(\|\bar{\bar{d}}\| + \|\bar{\bar{C}}\|\|e\| \right) \\ \Phi_4 &= \|e\| \left(-\lambda_{\min}(M)\|e\| + 2\|C\|\|R^{-1}\|\bar{d} \right). \end{aligned} \tag{39}$$

The curves represented by (37) and (38) are guaranteed to intersect. To this end, noting that (23) is a common Lyapunov candidate [36] across all time intervals, the error trajectories $(e, \tilde{\sigma})$ ultimately enter the compact set formed by the intersection of the curves (37) and (38)

$$\Gamma = \left\{ (e, \tilde{\sigma}) \in \Re^n \times \Re^l : \|e\| \leq B_e \wedge \|\tilde{\sigma}\| \leq B_{\tilde{\sigma}} \right\} \tag{40}$$

in finite time T and remain there for $\forall t \geq T$; Γ is positively invariant. This, in turn, according to Definition 1, proves the uniform ultimate boundedness of all system solutions under the proposed PAMCKF.

Remark 4. In view of the compact set (40), B_e and $B_{\tilde{\sigma}}$ are analytical curves that are guaranteed to intersect. Although the outlier d is unknown (B_e and $B_{\tilde{\sigma}}$ cannot be directly computed), for any bounded d with unknown bound $\bar{d}$, we can ensure that the system solutions are uniformly ultimately bounded.

Remark 5. The ultimate bound on the estimation error e, i.e., B_e given in (37), can be reduced by increasing $\lambda_{\min}(M)$ that is dependent on the choice of parameters Q, R, $P(0)$, η, and σ_0. In particular, in the absence of the outlier $d(t)$ (i.e., $\|\bar{d}\| = 0$) when $\|\tilde{\sigma}\| = 0$ (that results in $\Phi_1 = \Phi_2 = 0$), one has $\|e\| = 0$, implying the convergence of the estimation error to zero. Note that $\dot{V} = 0$ when both trajectories reach the boundary of Γ.

Remark 6. As an alternative to the proposed filter, one may think of estimating d and x (as $\hat{d}$ and $\hat{x}$) to compute the innovations $r = \hat{y} - y = -Ce + (d - \hat{d})$ and, in turn, directly update the estimate of x using (8). That is, a method that estimates d and x but does not update the Kalman gain and the error covariance from (7) and (9). Despite effectiveness in simulation, such a method is not applicable to real-world applications in which the ground truth states are not accessible using which the performance of estimator could be evaluated; instead, error covariance P is the only source of information to be used to assess the estimation performance. On the contrary, under the proposed filter, any disturbance applying to the measurements impacts on the innovations, that is captured by the correntropy gain and in turn tunes the Kalman gain based on which the error covariance is updated in terms of the disturbance. This implies that the proposed filter is applicable to real-world applications.

5. Simulation results

In this section, we verify the benefits of our proposed filter through simulations and comparisons to the traditional KF and the baseline MCKF [19] on (i) an equivalent circuit battery model used in hybrid electric vehicles and (ii) a longitudinal dynamical model for autonomous vehicles.

The reason we compare PAMCKF only with KF and MCKF—and not with other Monte Carlo filters such as UKF and Particle Filter—is twofold: (i) PAMCKF is conceptually derived from KF and MCKF, sharing similar algorithmic foundations and structure, and (ii) their computational complexities are comparable (see Remark 3), allowing for a fair, performance-focused evaluation without introducing other factors related to differing computational loads. In contrast, both UKF and PF face significant challenges in terms of scalability, computational cost, and memory usage, which limit their practicality in real-time and resource-constrained environments—especially in linear systems where their added complexity offers no clear advantage.

The following parameter tuning and simulation design are applied to both examples.

(1) Note that R could be known from the physics of sensors and Q can compromise between estimation accuracy and time lag. $P(0)$ and Q can be seen as tuning factors to be picked in such a way that the best performance is achieved.

(2) For the sake of having a fair comparison, the design parameters of all filters—proposed PAMCKF defined in Eqs. (12)–(21), the baseline MCKF presented in [19], and the traditional KF—are chosen to be the same. Note that under the baseline MCKF, the *kernel size* is constant and set to $\frac{1}{\sigma(0)}$.

(3) To assess the robustness of different filters in the presence of outliers with different magnitudes and different number of occurrences during the simulation, 20 Monte Carlo (MC) simulations are carried out for each filter under the following two cases, where the outlier $d(t)$ is considered to be occasional impulses (the encounter moment of impulses is randomly selected).

- *Case 1*: for each MC simulation, $N_d = 100$ impulses are enforced (with duration dt [s]) whose magnitudes $\bar{d}$ are drawn from a uniform distribution over the interval $[5, 20]$ (different $\bar{d}$s but the same N_ds).
- *Case 2*: for each MC simulation, the number of impulses N_d is drawn from a uniform distribution over the interval $[20, 200]$ (the duration of all impulses is dt [s]) while all taking the same magnitude $\bar{d} = 10$ (different N_ds but the same $\bar{d}$s).

(4) We use MATLAB to implement the proposed PAMCKF for which the following built-in functions are used.

- *randn* generates random numbers to follow a Gaussian distribution in order to model process and measurement noises mentioned in Assumption 3: $w(t) = \sqrt{Q}randn(n,t) \in \mathfrak{R}^n$ and $v(t) = \sqrt{R}randn(l,t) \in \mathfrak{R}^l$.
- *randi* draws a uniformly distributed pseudorandom scalar integer introducing (a) the random encounter moment of impulses during the simulation, (b) the random magnitude of each impulse, and (c) the random number of impulses.
- We use *the Fourth Order Runge Kutta's method (RK4)* [37] to solve the system dynamics, and *the Forward Euler's method (explicit)* to solve the dynamics (17) and (21).

5.1. Battery model of hybrid electric vehicles

This model can predict the dynamics of a lithium-ion battery and is comprised of an open circuit voltage (OCV) source $V_{B_{oc}}$ connected in series with a high-frequency internal resistance R_{B_0}, and two resistor–capacitor pairs $\left(R_{B_1}, C_{B_1}, R_{B_2}, C_{B_2}\right)$ that replicate the charge transfer and diffusion dynamics of the battery. The state-space representation of the battery is described as [38]

$$\begin{aligned}
\dot{V}_{B_1}(t) &= -\frac{V_{B_1}(t)}{R_{B_1}C_{B_1}} + \frac{I_B(t)}{C_{B_1}} + w_1(t)\\
\dot{V}_{B_2}(t) &= -\frac{V_{B_2}(t)}{R_{B_2}C_{B_2}} + \frac{I_B(t)}{C_{B_2}} + w_2(t)\\
y_{B_m}(t) &= \underbrace{V_{B_{oc}} - V_{B_1}(t) - V_{B_2}(t) - I_B(t)R_{B_0}}_{V_{cell}(t)}\\
&\quad + v(t) + d(t),
\end{aligned} \tag{41}$$

where $x(t) = [V_{B_1}(t), V_{B_2}(t)]^T$ is the vector of the state variables including the voltages $(V_{B_1}(t), V_{B_2}(t))$ [V] across two resistor–capacitor pairs; $I_B(t)$ [A] is the input current to the battery; $w(t) = [w_1(t), w_2(t)]^T$ is the vector of process noises; $y_{B_m}(t)$ [V] is the *disturbed measurement* that is impacted by the uncorrelated Gaussian noise $v(t)$ plus the large outlier $d(t)$ modeled as an impulsive function; and $V_{cell}(t)$ [V] is the terminal battery voltage excluding the noise and outlier. The model parameters are considered to be constant given as $R_{B_1} = 5 \times 10^{-4}$ [Ohm], $R_{B_2} = 9\times10^{-4}$ [Ohm], $C_{B_1} = 25000$ [F], $C_{B_2} = 32000$ [F], $R_{B_0} = 16\times10^{-4}$ [Ohm], and $Q_B = 15$ [Ah]. The sampling time is $dt = 1$ [s]. In hybrid electric vehicles, the battery is constrained to operate over a predefined and narrow range of state-of-charge due to which the OCV (i.e., $V_{B_{oc}}$) is generally taken as constant and equal to its average value over that state-of-charge range of operation. In particular, the average OCV of graphite anode/NMC cathode lithium ion batteries over the state-of-charge range of $[0.5, 0.8]$ is equal to 3 [V] [39]. The input current profile is shown in Fig. 2.

In view of (41) and by defining $y_d(t) = y_{B_m}(t) - V_{B_{oc}}$, one can obtain

$$A = \begin{bmatrix} \frac{-1}{R_{B_1}C_{B_1}} & 0 \\ 0 & \frac{-1}{R_{B_2}C_{B_2}} \end{bmatrix}, \quad B = \begin{bmatrix} \frac{1}{C_{B_1}} \\ \frac{1}{C_{B_2}} \end{bmatrix},$$
$$C = \begin{bmatrix} -1 & -1 \end{bmatrix}, \qquad D = -R_{B_0}. \tag{42}$$

The filter parameters are set to $Q = P(0) = 1 \times 10^{-6} I_2$, $R = 0.01$, $\sigma_0 = 0.001$, $\sigma(0) = 0.01$, $\eta = 20$, $\sigma_{max} = 1$, and $\nu_\sigma = 0.01$.

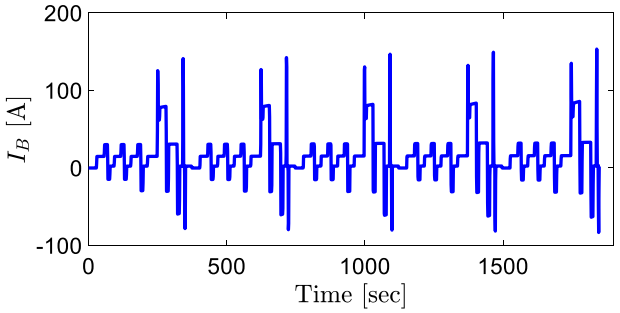


Fig. 2. Time evolution of input current $I_B(t)$ in Example 1.

5.2. Longitudinal model of autonomous vehicles

This model simulates the longitudinal dynamics of an autonomous vehicle using a simplified two-state model that incorporates actuator dynamics. We describe the vehicle's longitudinal behavior using the following state-space representation

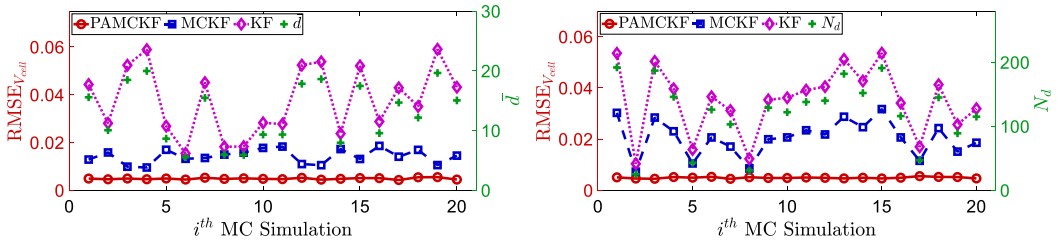
$$\begin{aligned}
\dot{v}_l(t) &= a_l(t) + w_1(t)\\
\dot{a}_l(t) &= -\frac{1}{\tau}a_l(t) + \frac{1}{\tau}a_c(t) + w_2(t)\\
y_{d_1}(t) &= v_l(t) + v_1(t) + d_1(t)\\
y_{d_2}(t) &= a_l(t) + v_2(t) + d_2(t),
\end{aligned}$$

where v_l [m/s] and a_l [m/s^2] are the *actual* vehicle's longitudinal velocity and acceleration, respectively; a_c [m/s^2] is the commanded acceleration. A first-order system with time constant $\tau = 0.5$ [s] is used to model the delayed response of the acceleration actuator. The sampling time is $dt = 0.02$ [s]. The vector of process noises is $w(t) = [w_1(t), w_2(t)]^T$; $y_d(t) = [y_{d_1}(t), y_{d_2}(t)]^T$ is the vector of *disturbed measurements* impacted by Gaussian noise $v(t) = [v_1(t), v_2(t)]^T$ and the impulse vector $d(t) = [d_1(t), d_2(t)]^T$ ($d(t)$ is modeled as an impulsive function); and $v_l(t)$ and $a_l(t)$ are the outputs excluding the noise and impulses. The input acceleration a_c is to simulate a driving scenario where the vehicle starts from stationary and accelerates at 1.5 [m/s^2] from 10 to 30 [s], then cruises at a constant speed from 30 to 60 [s], and finally decelerates to a stop at -1 [m/s^2] from 60 to 90 [s].

The filter parameters are set to $Q = P(0) = 1 \times 10^{-6} I_2$, $R = 0.01 I_2$, $\sigma_{0_1} = \sigma_{0_2} = 0.1$, $\sigma_1(0) = \sigma_2(0) = 0.025$, $\eta_1 = \eta_2 = 50$, $\sigma_{max_1} = \sigma_{max_2} = 1$, and $\nu_{\sigma_1} = \nu_{\sigma_2} = 0.01$.

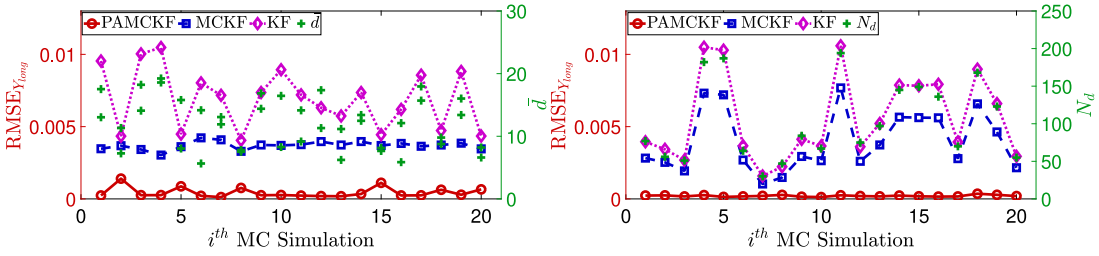
Figs. 3 and 4 both illustrate the estimation performance for different filters under *Case 1* and *Case 2* when the measurements are corrupted by the large outliers $d(t)$. It can be seen that under both cases in both examples, the proposed PAMCKF provides accurate and consistent estimation while significantly decreases RMSE values over the other filters. It is observed that although MCKF outperforms KF, under both of which, RMSE values are increased when either the magnitude or the number of the outliers encountered during the simulation increases. On the contrary, PAMCKF shows a relatively insensitive behavior when either *Case 1* or *Case 2* is enforced. These improvements are mainly due to the adaptation mechanism (21) using which the *kernel size* is evolved towards robustifying the filter against the large outliers $d(t)$. An example of such evolution is demonstrated in Figs. 5 and 6 where 20 randomly distributed outliers are considered. It is seen that σ immediately increases to compensate the outlier encountered and then decays with the exponential decay rate $\sigma_0\eta$ (see Remark 2).

Taken altogether, these results provide convincing evidence of applicability of the proposed filter as an alternative for the baseline MCKF and KF to be applied to the linear systems impacted by the large measurement outliers. Results provided also support the claim of Theorem 1 in which uniform ultimate boundedness of all system errors $(e(0), \tilde{\sigma}(0))$ is guaranteed.



(a) Case 1 (b) Case 2

Fig. 3. Example 1. Comparison results based on 20 MC simulations. RMSE$_{V_{cell}}$ returns RMS of the difference between $V_{cell}(V_{B_1}, V_{B_2})$ and $V_{cell}(\hat{V}_{B_1}, \hat{V}_{B_2})$ form which the state estimation performance is measured.



(a) Case 1 (b) Case 2

Fig. 4. Example 2. Comparison results based on 20 MC simulations. RMSE$_{Y_{long}}$ returns the sum of the RMS of $v_l - \hat{v}_l$ and the RMS of $a_l - \hat{a}_l$ form which the state estimation performance is measured.

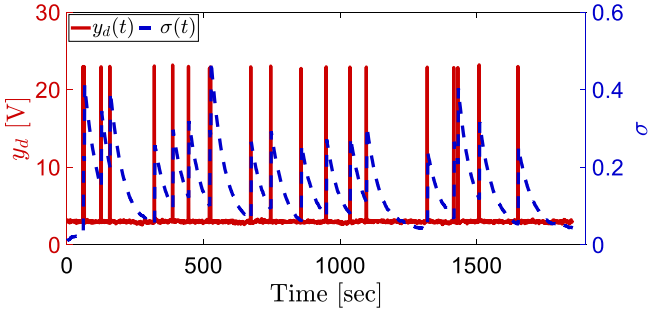


Fig. 5. Example 1. *Disturbed measurement* $y_d(t)$ affected by $N_d = 20$ randomly distributed outliers with the same magnitude $\bar{d} = 20$ [V], along with the parameter $\sigma(t)$ adjusted by the proposed PAMCKF.

6. Conclusions and future work

This work introduced a PAMCKF to be applied to linear systems subject to unknown measurement outliers. Under our filter, the *kernel size* is updated on-line, the computational cost is low, and the uniform ultimate boundedness of all system errors is guaranteed in a deterministic setting. The effectiveness of PAMCKF over the baseline KF and MCKF has been demonstrated through simulations. Results showed that measurement outliers degrade the estimates under the baseline filters, whereas the estimation performance with the proposed technique is consistent and relatively insensitive to large outliers. Despite its effectiveness, an important direction for future research is the development of a discrete-time version of the proposed filter, which is crucial for deployment in real-time embedded systems and other practical applications. However, transitioning to a discrete-time implementation introduces several challenges—most notably, the difficulty of conducting a rigorous stability analysis in a stochastic setting.

CRediT authorship contribution statement

Vahid Azimi: Writing – review & editing, Validation, Resources, Formal analysis, Visualization, Writing – original draft, Supervision, Methodology, Conceptualization, Software, Investigation. **Simona Onori:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

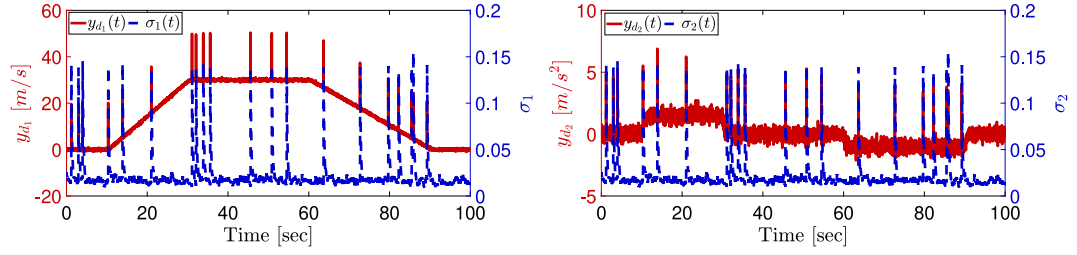


Fig. 6. Example 2. *Disturbed measurements* y_{d_1} and y_{d_2} each affected by $N_d = 20$ randomly distributed outliers, with outlier magnitudes of $\bar{d}_1 = 20$ [m/s] for y_{d_1} and $\bar{d}_2 = 5$ [m/s^2] for y_{d_2}, along with the parameter $\sigma_1(t)$ and $\sigma_2(t)$ adjusted by the proposed PAMCKF.

Data availability

Data will be made available on request.

References

[1] R.E. Kalman, A new approach to linear filtering and prediction problems, ASME. J. Basic Eng. 82 (1) (1960) 35–45.
[2] R.R. Bitmead, M. Hovd, M.A. Abooshahab, A Kalman-filtering derivation of simultaneous input and state estimation, Automatica 108 (2019) 108478.
[3] H. Zhang, R. Ayoub, S. Sundaram, Sensor selection for Kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms, Automatica 78 (2017) 202–210.
[4] B. Ding, Z. Ren, H. Fang, Simultaneous input and state estimation with multi-step delay for linear stochastic systems based on infinity filtering and smoothing, Systems Control Lett. 185 (2024) 105752.
[5] J. Zhao, G. Zhang, A robust prony method against synchrophasor measurement noise and outliers, IEEE Trans. Power Syst. 32 (3) (2017) 2484–2486.
[6] T. Ortmaier, M. Groger, D. Boehm, V. Falk, G. Hirzinger, Motion estimation in beating heart surgery, IEEE Trans. Biomed. Eng. 52 (10) (2005) 1729–1740.
[7] R. Pearson, Outliers in process modeling and identification, IEEE Trans. Control Syst. Technol. 10 (1) (2002) 55–63.
[8] H. Ferdowsi, S. Jagannathan, M. Zawodniok, An online outlier identification and removal scheme for improving fault detection performance, IEEE Trans. Neural Networks Learn. Syst. 25 (5) (2014) 908–919.
[9] S. Meng, L. Liu, Enhanced monitoring-as-a-service for effective cloud management, IEEE Trans. Comput. 62 (9) (2013) 1705–1720.
[10] B. Cui, W. Chen, D. Weng, J. Wang, X. Wei, Y. Zhu, Variational resampling-free cubature Kalman filter for GNSS/INS with measurement outlier detection, Signal Process. 237 (2025) 110036.
[11] B. Wang, J. Wang, B. Zhang, W. Chen, Z. Zhang, Leader–follower consensus of multivehicle wirelessly networked uncertain systems subject to nonlinear dynamics and actuator fault, IEEE Trans. Autom. Sci. Eng. 15 (2) (2018) 492–505.
[12] B. Wang, B. Zhang, R. Su, Optimal tracking cooperative control for cyber-physical systems: Dynamic fault-tolerant control and resilient management, IEEE Trans. Ind. Inform. 17 (1) (2021) 158–167.
[13] B. Chen, X. Liu, H. Zhao, J.C. Principe, Maximum correntropy Kalman filter, Automatica 76 (2017) 70–77.
[14] A. Harvey, A. Luati, Filtering with heavy tails, J. Amer. Statist. Assoc. 109 (507) (2014) 1112–1122.
[15] Y. Huang, Y. Zhang, N. Li, Z. Wu, J.A. Chambers, A novel robust student's t-based Kalman filter, IEEE Trans. Aerosp. Electron. Syst. 53 (3) (2017) 1545–1554.
[16] D. Alspach, H. Sorenson, Nonlinear bayesian estimation using gaussian sum approximations, IEEE Trans. Autom. Control 17 (4) (1972) 439–448.
[17] X. Zhang, J. You, L. Kang, A novel augmented particle filter for extended target tracking with star-convex random hypersurface model, Measurement 253 (2025) 117602.
[18] S.J. Julier, J.K. Uhlmann, Unscented filtering and nonlinear estimation, Proc. IEEE 92 (3) (2004) 401–422.
[19] R. Izanloo, S.A. Fakoorian, H. Sadoghi, D. Simon, Kalman filtering based on the maximum correntropy criterion in the presence of non-Gaussian noise, in: 50^{th} Annual Conference on Information Science and Systems, Princeton, New Jersey, 2016, pp. 530–535.
[20] M. Kulikova, Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise, Systems Control Lett. 108 (2017) 8–15.
[21] L. Dang, B. Chen, S. Wang, Y. Gu, J.C. Príncipe, Kernel Kalman filtering with conditional embedding and maximum correntropy criterion, IEEE Trans. Circuits Syst. I. Regul. Pap. 66 (11) (2019) 4265–4277.
[22] M. Zhang, W.X. Zheng, X. Song, H. Yuan, Two efficient Kalman filter algorithms for measurement packet dropping systems under maximum correntropy criterion, Systems Control Lett. 175 (2023) 105515.
[23] M. Kulikova, Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-Gaussian noise, Systems Control Lett. 108 (2017) 8–15.
[24] Q. Ge, X. Bai, P. Zeng, Gaussian-Cauchy mixture kernel function based maximum correntropy criterion Kalman filter for linear non-Gaussian systems, IEEE Trans. Signal Process. 73 (2025) 158–172.
[25] Z. Liu, Z. Zhao, Y. Qiu, B. Jing, C. Yang, H. Wu, Enhanced state of charge estimation for li-ion batteries through adaptive maximum correntropy Kalman filter with open circuit voltage correction, Energy 283 (2023) 128738.
[26] S. Fakoorian, R. Izanloo, A. Shamshirgaran, D. Simon, Maximum correntropy criterion Kalman filter with adaptive kernel size, in: 2019 IEEE National Aerospace and Electronics Conference, NAECON, 2019, pp. 581–584.
[27] S. Li, B. Xu, L. Wang, A.A. Razzaqi, Improved maximum correntropy Cubature Kalman filter for cooperative localization, IEEE Sensors J. 20 (22) (2020) 13585–13595.
[28] G.T. Cinar, J.C. Príncipe, Hidden state estimation using the correntropy filter with fixed point update and adaptive kernel size, in: The 2012 International Joint Conference on Neural Networks, IJCNN, 2012, pp. 1–6.
[29] B. Hou, Z. He, X. Zhou, H. Zhou, D. Li, J. Wang, Maximum correntropy criterion Kalman filter for α-Jerk tracking model with non-Gaussian noise, Entropy 19 (12) (2017) 13585–13595.
[30] M. Corless, G. Leitmann, Continuous state feedback guaranteeing uniform ultimate boundedness for uncertain dynamic systems, IEEE Trans. Autom. Control 26 (5) (1981) 1139–1144.
[31] J.-B. Pomet, L. Praly, Adaptive nonlinear regulation: estimation from the Lyapunov equation, IEEE Trans. Autom. Control 37 (6) (1992) 729–740.
[32] D. Simon, Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches, John Wiley & Sons, 2006.
[33] V. Azimi, S. Hutchinson, Robust adaptive control Lyapunov-Barrier function for non-collocated control and safety of underactuated robotic systems, Internat. J. Robust Nonlinear Control 32 (13) (2022) 7363–7390.
[34] P. Ioannou, P. Kokotovic, Adaptive Systems with Reduced Models, Springer-Verlag, New York, 1983.
[35] S. Bonnabel, J. Slotine, A contraction theory-based analysis of the stability of the deterministic extended Kalman filter, IEEE Trans. Autom. Control 60 (2) (2015) 565–569.
[36] D. Liberzon, Switching in Systems and Control, Birkhäuser, Boston, MA, 2003.
[37] B. Carnahan, et al., Applied Numerical Methods, John Wiley & Sons, 1969.
[38] A. Chu, A. Allam, A. Cordoba Arenas, G. Rizzoni, S. Onori, Stochastic capacity loss and remaining useful life models for lithium-ion batteries in plug-in hybrid electric vehicles, J. Power Sources 478 (2020) 228991.
[39] S. Onori, L. Serrao, G. Rizzoni, Hybrid Electric Vehicles Energy Management Strategies, Springer Briefs in Electrical and Computer Engineering, Springer 1st ed, 2016.