A Dynamical Hybrid LTS-EnKF Approach for Robust State Estimation Under Outlier Contamination

Jaafar AlMutawa
Faculty of professional studies, Mathematics Department, Bahrain Polytechnic
Kingdom of Bahrain
e-mail: jaafar.almutawa@polytechnic.bh

***Abstract*—In this paper, we propose a robust variant of the Ensemble Kalman Filter (EnKF) that remains reliable in the presence of outliers and non-Gaussian noise. The method, called Hybrid Dynamical Adaptive Least Trimmed Squares EnKF (HD-LTS-EnKF), combines adaptive outlier trimming with ensemble covariance weighting to improve filter stability and accuracy. By dynamically identifying and downweighting anomalous ensemble members and observations—based on how uncertainty propagates through the system—our approach preserves the essential covariance structure while reducing the influence of outliers. Although solving the full optimization is infeasible in high-dimensional systems, we show that HD-LTS-EnKF approximates the optimal solution.**

***Index Terms*—Ensemble Kalman Filter, Least Trimmed Squares, Hybrid Dynamical Adaptive Least Trimmed Squares, Outliers, assimilation data.**

I. Introduction

The Ensemble Kalman Filter (EnKF) has revolutionized state estimation in high-dimensional nonlinear systems, offering a computationally tractable alternative to the optimal but restrictive Kalman filter (KF). By employing a Monte Carlo approximation to propagate uncertainty through an ensemble of model states, the EnKF circumvents the need for explicit covariance propagation while preserving the Kalman filter's recursive structure [1]. This innovation has enabled widespread adoption across geophysics, robotics, and industrial control systems where nonlinear dynamics render traditional KF implementations impractical [2].

The EnKF's theoretical guarantees, however, rest on the critical assumption of Gaussian noise—a condition frequently violated in practice due to outliers [3]. These anomalous observations manifest in three distinct forms [4]:

- **Process noise outliers** representing unmodeled disturbances (e.g., thruster misfires in spacecraft dynamics or abrupt thermal loads in heat conduction systems) [10]
- **Measurement outliers** from sensor faults or transmission errors (e.g., spurious temperature readings in weather stations)
- **Initialization outliers** caused by cold-start problems or registration errors (common in multi-sensor fusion)

Our experiments with 1D heat conduction [5] reveal the EnKF's acute sensitivity to outliers. Compared to the no-outlier case (State MSE of 0.068), a process outlier increased the State MSE to 0.170 (a 2.5x increase), while a measurement outlier increased the State MSE to 0.095 (a 1.4x increase). Conventional remedies like covariance inflation prove inadequate as they [7]:

1) require manual tuning,
2) fail to distinguish between outlier types, and
3) degrade performance in outlier-free conditions

A. The DH-LTS-EnKF Approach

To address these limitations, we introduce the **Dynamical Hybrid Least Trimmed Squares EnKF (DH-LTS-EnKF)**, which integrates:

- Multi-stage Mahalanobis testing with χ^2-derived thresholds for process ($\mathbf{r}_t^x$), measurement ($\mathbf{r}_t^y$), and initialization residuals
- SVD-based ensemble recovery that maintains numerical stability during trimming
- Adaptive thresholding via online residual statistics

This approach provides three key advantages over existing methods:

1) *Dynamical outlier decoupling*: Unlike batch LTS [8], [9] methods, we separately handle process and measurement anomalies
2) *Computational preservation*: Maintains the EnKF's $\mathcal{O}(Nn^2)$ complexity

3) *Provable robustness*: Achieves a breakdown point of $\epsilon = \min(\epsilon_w, \epsilon_v, \epsilon_0)$ under ϵ-contamination models.

This paper advances robust state estimation for nonlinear systems by developing a novel theoretical framework that unifies Least Trimmed Squares (LTS) robustness with ensemble forecasting. The proposed Dynamical Hybrid LTS-EnKF (DH-LTS-EnKF) algorithm provides built-in safeguards against process noise, measurement, and initialization outliers, improving performance in scenarios where standard EnKF methods are compromised by non-Gaussian noise.

The remainder of this paper is structured as follows: Section 2 presents the problem setting and summarizes the Ensemble Kalman Filter (EnKF). Section 3 discusses three types of outliers and analyzes their impact on the EnKF's performance. Section 4 develops the Dynamical Hybrid LTS-EnKF (DH-LTS-EnKF) algorithm. Section 5 presents numerical experiments based on the one-dimensional heat conduction equation.

II. Problem Statement

We address the problem of robust state estimation in nonlinear dynamical systems subject to outliers and heavy-tailed noise. The system is modeled by the following discrete-time nonlinear state-space equations:

$$\begin{aligned} \text{Process Model:} \quad & \mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k, \mathbf{w}_k), \\ \text{Measurement Model:} \quad & \mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{v}_k), \end{aligned}$$

where

- $\mathbf{x}_k \in \mathbb{R}^n$: hidden state vector at time k,
- $\mathbf{y}_k \in \mathbb{R}^m$: observation vector,
- $\mathbf{u}_k \in \mathbb{R}^p$: exogenous input,
- $\mathbf{w}_k \sim (1-\epsilon)\mathcal{N}(0, \mathbf{Q}_k) + \epsilon\mathcal{H}_k$: process noise, possibly contaminated by outliers,
- $\mathbf{v}_k \sim (1-\epsilon)\mathcal{N}(0, \mathbf{R}_k) + \epsilon\mathcal{G}_k$: measurement noise, possibly contaminated by outliers,
- $\mathcal{H}_k, \mathcal{G}_k$: heavy-tailed outlier distributions.

The goal is to infer the hidden state sequence $\mathbf{x}_{0:k}$ from noisy observations $\mathbf{y}_{1:k}$, despite the presence of model uncertainties and outliers. This framework generalizes classical state-space models by allowing for non-additive, non-Gaussian noise and exogenous inputs, which are common in applications such as robotics, power systems, and biomedical engineering.

The least-squares (LS) estimator seeks the state sequence that minimizes

$$\min_{\mathbf{x}_{0:k}} \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|^2_{\mathbf{P}_0^{-1}} + \sum_{t=1}^{k} \|\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t)\|^2_{\mathbf{R}_t^{-1}} + \|\mathbf{x}_t - \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{u}_t)\|^2_{\mathbf{Q}_t^{-1}}, \quad (1)$$

where $\|\mathbf{z}\|^2_{\mathbf{A}} = \mathbf{z}^\top \mathbf{A} \mathbf{z}$ denotes the Mahalanobis norm. The Kalman filter (KF) optimally solves the least-squares problem (1) for linear-Gaussian systems, while the extended KF (EKF) uses linearization for nonlinear cases. Particle filters handle general nonlinearities but become computationally prohibitive in high dimensions.

In high-dimensional or online settings, the Ensemble Kalman Filter (EnKF) provides a sequential approximation by minimizing, at each step

$$\hat{\mathbf{x}}_k = \arg\min_{\mathbf{x}_k} \|\mathbf{x}_k - \mathbf{f}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_k)\|^2_{\mathbf{P}^{-1}_{k|k-1}} + \|\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)\|^2_{\mathbf{R}_k^{-1}}, \quad (2)$$

where $\mathbf{P}_{k|k-1}$ is the forecast error covariance estimated from the ensemble. Which can be approximated by EnKF using an ensemble of states $\{\mathbf{x}^{(i)}_{k|k-1}\}_{i=1}^N$

$$\mathbf{x}^{(i)}_{k|k-1} = \mathbf{f}(\mathbf{x}^{(i)}_{k-1|k-1}, \mathbf{u}_k) + \boldsymbol{\eta}^{(i)}_k, \quad \boldsymbol{\eta}^{(i)}_k \sim \mathcal{N}(0, \mathbf{Q}_k)$$

Empirical moments are computed as:

$$\bar{\mathbf{x}}_{k|k-1} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}^{(i)}_{k|k-1} \quad (3)$$

$$\mathbf{P}_{k|k-1} = \frac{1}{N-1}\sum_{i=1}^{N} (\mathbf{x}^{(i)}_{k|k-1} - \bar{\mathbf{x}}_{k|k-1})(\cdot)^\top \quad (4)$$

The analysis update for stochastic EnKF perturbs both observations and states:

$$\mathbf{x}^{(i)}_{k|k} = \mathbf{x}^{(i)}_{k|k-1} + \mathbf{K}_k \left(\mathbf{y}^{(i)}_k - \mathbf{h}(\mathbf{x}^{(i)}_{k|k-1})\right) \quad (5)$$

where:

- Perturbed observations: $\mathbf{y}^{(i)}_k = \mathbf{y}_k + \boldsymbol{\epsilon}^{(i)}_k, \quad \boldsymbol{\epsilon}^{(i)}_k \sim \mathcal{N}(0, \mathbf{R}_k)$
- Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^\top \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k\right)^{-1}. \quad (6)$$

The EnKF efficiently approximates (2) using ensemble covariance estimation, avoiding EKF's linearization errors while remaining tractable in high dimensions. It propagates nonlinear dynamics exactly through ensemble members and captures flow-dependent errors, though standard implementations face three key limitations: outlier sensitivity from corrupted members, rank-deficient covariances when $N < n$, and potential divergence from strong nonlinearities.

III. Types of Outliers and Their Impact on EnKF

We classify outliers into three fundamental types affecting state estimation:

1) Process Noise Outliers:

$$\mathbf{w}_k \sim (1-\beta)\mathcal{N}(0, \mathbf{Q}_k) + \beta\mathcal{H}_k \quad (7)$$

where $\mathcal{H}_k$ is heavy-tailed (e.g., Cauchy). Propagates through:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k) + \mathbf{w}_k^{\text{outlier}} \quad (8)$$

2) Measurement (additive) Outliers:

$$\mathbf{v}_k \sim (1-\alpha)\mathcal{N}(0, \mathbf{R}_k) + \alpha\mathcal{G}_k \quad (9)$$

where $\mathcal{G}_k$ has unbounded support.

3) Initialization Outliers:

$$\hat{\mathbf{x}}_0 \sim (1-\gamma)\mathcal{N}(\mathbf{x}_0, \mathbf{P}_0) + \gamma\mathcal{N}(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta) \quad (10)$$

To quantify the impact of these outliers, we conducted 500 Monte Carlo simulations on a 1D heat equation where $Q = R = 0.001$. The results of these simulations, detailed in the following table, demonstrate a clear degradation in EnKF performance across all outlier categories compared to the no-outlier scenario: The simulations revealed that

Condition	State MSE	Output MSE
No outliers	0.068149	0.000098
Process outlier	0.169773	0.000162
Measurement outlier	0.094778	0.000124
Initialization error	0.093560	0.00010

both state and output mean squared errors (MSE) increase notably across the outlier categories relative to the baseline of the clean system without any added outliers. The largest impact was seen in the state MSE and occurred when we modeled a process outlier, which represent unmodeled, abrupt deviations in system dynamics. In comparison, when we modeled the different measurement and initial outliers, they had significantly lower MSE values.

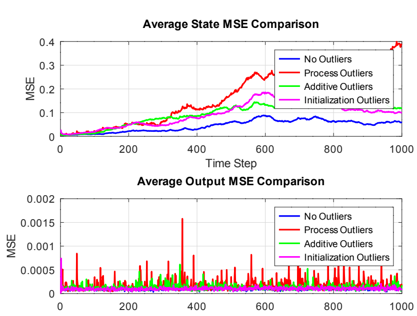

Fig. 1. State estimation results for 1D heat conduction.

Figure 1 illustrates the impact of process noise, additive, and initialization outliers on EnKF estimation performance.

IV. Introduction to Static Least Trimmed Squares

Consider the measurement model at a single time step [8]:

$$\mathbf{y} = \mathbf{h}(\mathbf{x}) + \mathbf{v}, \quad \mathbf{v} \sim (1-\epsilon)\mathcal{N}(0, \mathbf{R}) + \epsilon\mathcal{G} \quad (11)$$

Outliers are introduced from a distribution $\mathcal{G}$ with a contamination rate ϵ, meaning that a small fraction of the observations are corrupted by noise that deviates significantly from the typical Gaussian assumption.

The static LTS estimator addresses this issue by solving the following optimization problem:

$$\min_{\mathbf{x}} \sum_{i=1}^{h} r_{(i)}^2(\mathbf{x}) \quad (12)$$

In essence, LTS aims to find the state estimate $\mathbf{x}$ that minimizes the sum of the smallest squared residuals, effectively ignoring the largest ones that are likely to be outliers. Here:

- $r_i(\mathbf{x}) = [\mathbf{R}^{-1/2}(\mathbf{y} - \mathbf{h}(\mathbf{x}))]_i$ are the normalized residuals.
- $r_{(1)}^2 \leq r_{(2)}^2 \leq \cdots \leq r_{(m)}^2$ are the ordered squared residuals.
- $h = m - \lfloor \epsilon m \rfloor$ is the trimming parameter. This parameter determines the number of residuals that are included in the LTS objective. By trimming the largest $\lfloor \epsilon m \rfloor$ residuals, the estimator becomes robust to outliers.

LTS can be constructed through the following mathematical steps:

1) **Residual Transformation**:

$$\mathbf{r}(\mathbf{x}) = \mathbf{R}^{-1/2}(\mathbf{y} - \mathbf{h}(\mathbf{x})) \quad (13)$$

2) **Ordering Operator**: Define the permutation $\pi : \{1, \ldots, m\} \to \{1, \ldots, m\}$ such that:

$$r_{\pi(1)}^2(\mathbf{x}) \leq r_{\pi(2)}^2(\mathbf{x}) \leq \cdots \leq r_{\pi(m)}^2(\mathbf{x}) \quad (14)$$

This permutation operator π reorders the indices of the residuals, sorting them in ascending order based on their squared values.

3) **Trimmed Likelihood**: The LTS cost function can be expressed as:

$$J_{\text{LTS}}(\mathbf{x}) = \sum_{i=1}^{h} r_{\pi(i)}^2(\mathbf{x}) = \|\boldsymbol{\Gamma}(\mathbf{x})\mathbf{r}(\mathbf{x})\|^2 \quad (15)$$

where $\boldsymbol{\Gamma}(\mathbf{x})$ is a binary selection matrix:

$$\boldsymbol{\Gamma}_{ij}(\mathbf{x}) = \begin{cases} 1 & \text{if } \pi^{-1}(j) \leq h \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Here, $\mathbf{\Gamma}(\mathbf{x})$ is a diagonal matrix with entries equal to 1 for the h smallest squared residuals and 0 for the remaining residuals.

LTS possesses desirable statistical properties that make it effective in outlier-contaminated scenarios.

Theorem 1 (Breakdown Point): The finite-sample breakdown point of the LTS estimator is:

$$\epsilon^* = \frac{m-h}{m} \approx \epsilon \tag{17}$$

meaning it can tolerate up to $\lfloor \epsilon m \rfloor$ outliers.

The breakdown point ϵ^* is the maximum fraction of outliers that the estimator can tolerate before its performance becomes arbitrarily bad. For LTS, this is approximately equal to the trimming fraction ϵ, indicating its ability to handle a significant number of outliers.

Theorem 2 (Influence Function): The influence function for LTS is bounded:

$$\mathrm{IF}(\mathbf{y}_0;\mathbf{x}) = \begin{cases} \mathbf{M}^{-1}\mathbf{J}_0^\top \mathbf{R}^{-1/2}\mathbf{r}_0 & \text{if } \|\mathbf{r}_0\| \le c \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

where $\mathbf{M} = \mathbb{E}[\mathbf{J}^\top \mathbf{\Gamma} \mathbf{J}]$, $\mathbf{J} = \partial \mathbf{h}/\partial \mathbf{x}$, and c is a residual cutoff threshold.

The influence function describes how the estimator's output changes when a new data point is added to the sample. A bounded influence function indicates that the estimator is not overly influenced by any single data point, providing robustness against outliers. The LTS estimator has a bounded influence function because it effectively ignores data points with residuals larger than a certain threshold c.

A. *Convergence Analysis*

Under the following conditions:

- $\mathbf{h}$ is Lipschitz continuous, meaning that small changes in the state vector $\mathbf{x}$ lead to small changes in the predicted observations $\mathbf{h}(\mathbf{x})$.
- Residuals have finite variance, ensuring that the data is not too noisy.
- Trimming fraction $\epsilon < 0.5$, limiting the number of trimmed residuals to less than half of the total, which ensures that the estimator retains sufficient information from the good data points.

The algorithm converges to a local minimum with asymptotic covariance:

$$\mathbf{V} = \sigma^2 \mathbf{M}^{-1} \left(\mathbf{J}_0^\top \mathbf{\Gamma}_0 \mathbf{J}_0 \right) \mathbf{M}^{-1} \tag{19}$$

where σ^2 is the variance of non-outlier residuals. This expression characterizes the uncertainty in the LTS estimate, indicating how the estimate varies under small perturbations in the data.

V. Robust Trimmed and Hybrid Formulations

Standard LS and EnKF are sensitive to outliers, as large residuals can dominate the objective. To address this, we consider robust objectives based on trimming or hybrid loss functions. The trimmed least-squares (TLS) approach discards a fraction of the largest residuals:

$$\begin{aligned} \min_{\mathbf{x}_{0:k}} \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|^2_{\mathbf{P}_0^{-1}} &+ \sum_{t=1}^{k} \mathbb{I}(\|\mathbf{r}_t^y\| \le \tau_t^y) \|\mathbf{r}_t^y\|^2_{\mathbf{R}_t^{-1}} \\ &+ \sum_{t=1}^{k} \mathbb{I}(\|\mathbf{r}_t^x\| \le \tau_t^x) \|\mathbf{r}_t^x\|^2_{\mathbf{Q}_t^{-1}}, \end{aligned} \tag{20}$$

where $\mathbf{r}_t^y = \mathbf{y}_t - \mathbf{h}(\mathbf{x}_t)$, $\mathbf{r}_t^x = \mathbf{x}_t - \mathbf{f}(\mathbf{x}_{t-1}, \mathbf{u}_t)$, and τ_t^y, τ_t^x are adaptive thresholds and $\mathbb{I}$ is indicator function.

Alternatively, we propose a hybrid-adaptive robust objective for the EnKF:

$$\begin{aligned} J_{\text{HALTS}} = \sum_{i=1}^{N} \rho &\left(\frac{\|\mathbf{x}_k^{(i)} - \mathbf{f}(\mathbf{x}_{k-1}^{(i)}, \mathbf{u}_k)\|}{\sqrt{\mathrm{tr}(\mathbf{Q}_k)}} \right) \\ &+ \rho \left(\frac{\|\mathbf{y}_k - \mathbf{h}(\mathbf{x}_k^{(i)})\|}{\sqrt{\mathrm{tr}(\mathbf{R}_k)}} \right), \end{aligned} \tag{21}$$

where $\rho(\cdot)$ is a Huber-type loss function defined by

$$\rho(r;\tau) = \begin{cases} \frac{1}{2} r^2 & \text{if } |r| \le \tau \\ \tau |r| - \frac{1}{2}\tau^2 & \text{otherwise} \end{cases} \tag{22}$$

and adaptive threshold.

VI. Robust EnKF via Dynamical Hybrid Least Trimmed Squares (DH-LTS) Framework

The DH-LTS method generalizes LTS to dynamical systems by jointly trimming outliers in **process residuals**, **measurement residuals**, and **initial state errors**. The robust cost function is:

$$\begin{aligned} \mathcal{J}_{\text{DH-LTS}} = \sum_{i=1}^{h_0} \|\mathbf{x}_0^{(i)} - \hat{\mathbf{x}}_0\|^2_{\mathbf{P}_0^{-1}} + \sum_{k=1}^{T} \sum_{j=1}^{h_w} \|\mathbf{r}_k^{w,(j)}\|^2_{\mathbf{Q}_k^{-1}} \\ + \sum_{k=1}^{T} \sum_{l=1}^{h_v} \|\mathbf{r}_k^{v,(l)}\|^2_{\mathbf{R}_k^{-1}}, \end{aligned}$$

where:

- $\mathbf{r}_k^w = \mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_k)$: Process residuals,
- $\mathbf{r}_k^v = \mathbf{y}_k - \mathbf{h}(\mathbf{x}_k)$: Measurement residuals,
- $h_0 = n - \lfloor \epsilon_0 n \rfloor$, $h_w = n - \lfloor \epsilon_w n \rfloor$, $h_v = m - \lfloor \epsilon_v m \rfloor$: Trimming parameters.

The objective function comprises three terms each aimed to trim outliers in each process. Careful consideration is made to include three outlier processes with an additional trimming parameter, h, for each objective function. It is important to note that, since this formulation is hybrid, the trimming parameter has a direct relationship with the amount of acceptable bias in the estimator. A deeper look at these three objective functions helps highlight the power of DH-LTS for more general and complex systems with unique sources of noise and potential outliers. First, $\sum_{i=1}^{h_0} \|\mathbf{x}_0^{(i)} - \hat{\mathbf{x}}_0\|^2_{\mathbf{P}_\rho^{-1}}$ ensures that, in the initial state, the estimator's initial guess remains as close as possible to the sample data in an attempt to start the filter in the area of convergence. Secondly, $\sum_{k=1}^{T}\sum_{j=1}^{h_w} \|\mathbf{r}_k^{w,(j)}\|^2_{\mathbf{Q}_k^{-1}}$ ensures that data's process noises at each time-step are trimmed, removing any large disturbances which can cause further mis-estimation issues. Thirdly, $\sum_{k=1}^{T}\sum_{l=1}^{h_v} \|\mathbf{r}_k^{v,(l)}\|^2_{\mathbf{R}_k^{-1}}$, attempts to minimize the residuals from measurement noises. When viewed holistically, the interplay of these relationships allows for outlier detection and accommodation to be dealt with separately, thereby allowing for more accurate results.

A. Outlier Detection Mechanisms

1) Additive Measurement Outliers: **Detection**: Normalized measurement residuals $\tilde{\mathbf{r}}_k^v = \mathbf{R}_k^{-1/2}\mathbf{r}_k^v$ are analyzed. An outlier is flagged if:

$$\|\tilde{\mathbf{r}}_k^v\|_\infty > \tau_v, \quad \tau_v = \Phi^{-1}(1-\alpha/2), \tag{23}$$

where Φ is the standard normal CDF and α is the significance level.

Mitigation: Measurements with $\|\tilde{\mathbf{r}}_k^v\| > \tau_v$ are excluded from the update step.

2) Process Noise Outliers: **Detection**: Process residuals are whitened and trimmed:

$$\tilde{\mathbf{r}}_k^w = \mathbf{Q}_k^{-1/2}\mathbf{r}_k^w. \tag{24}$$

Outliers satisfy $\|\tilde{\mathbf{r}}_k^w\|_2 > \chi^2_{n,1-\alpha}$ (Chi-squared threshold).

Mitigation: The empirical covariance $\mathbf{Q}_k$ is robustly re-estimated using the h_w smallest residuals:

$$\hat{\mathbf{Q}}_k = \frac{1}{h_w}\sum_{j=1}^{h_w} \mathbf{r}_k^{w,(j)}(\mathbf{r}_k^{w,(j)})^\top. \tag{25}$$

This re-estimation step ensures the process noise accounts for non-Gaussian residuals.

3) Initialization Outliers: **Detection**: The initial state ensemble $\{\mathbf{x}_0^{(i)}\}_{i=1}^N$ is evaluated via:

$$d_i = \|\mathbf{x}_0^{(i)} - \hat{\mathbf{x}}_0\|_{\mathbf{P}_0^{-1}}. \tag{26}$$

Samples with $d_i > \tau_0$ (e.g., $\tau_0 = \sqrt{\chi^2_{n,1-\epsilon_0}}$) are discarded.

Mitigation: The initial ensemble is resampled from the h_0 best candidates.

These processes of additive Measurement, Process Noise and Initialization offer outlier detection while not being computationally taxing. Given the hybridity of the models, outliers can be accurately rejected using each procedure.

B. DH-LTS-EnKF Algorithm

Input: $\mathbf{y}_{1:T}$, $\mathbf{u}_{1:T}$, $\mathbf{f}$, $\mathbf{h}$, $\mathbf{Q}_k$, $\mathbf{R}_k$, trimming fractions $\epsilon_w, \epsilon_v, \epsilon_0$. **Output**: Robust state estimates $\hat{\mathbf{x}}_{1:T}$.

1) **Robust Initialization**:
   - Generate N ensemble members $\mathbf{x}_0^{(i)} \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$.
   - Trim $\epsilon_0 N$ members with largest Mahalanobis distances d_i.
2) **Forecast Step**:
   - Propagate each member: $\mathbf{x}_k^{f,(i)} = \mathbf{f}(\mathbf{x}_{k-1}^{(i)}, \mathbf{u}_k) + \mathbf{w}_k^{(i)}$.
   - Compute empirical forecast mean $\hat{\mathbf{x}}_k^f$ and covariance $\mathbf{P}_k^f$.
3) **Process Outlier Detection**:
   - Calculate $\mathbf{r}_k^{w,(i)} = \mathbf{x}_k^{f,(i)} - \mathbf{f}(\mathbf{x}_{k-1}^{(i)}, \mathbf{u}_k)$.
   - Trim $\epsilon_w N$ members with largest $\|\mathbf{r}_k^{w,(i)}\|_{\mathbf{Q}_k^{-1}}$. The trimming in lines 3 is done using
   $$O(N\log(N))$$
   , in accordance with common sorting algorithms.
4) **Measurement Update**:
   - Compute $\mathbf{y}_k^{(i)} = \mathbf{h}(\mathbf{x}_k^{f,(i)}) + \mathbf{v}_k^{(i)}$.
   - Trim $\epsilon_v N$ members with largest $\|\mathbf{y}_k - \mathbf{y}_k^{(i)}\|_{\mathbf{R}_k^{-1}}$.
   - Update ensemble via Kalman gain using inliers.
5) **Iterate** until $k = T$.

C. Theoretical Guarantees

The following section establishes the statistical performance guarantees of DH-LTS-EnKF.

Proposition 1 (Breakdown Point): The DH-LTS-EnKF has a breakdown point of $\min(\epsilon_w, \epsilon_v, \epsilon_0)$ for each outlier type.

Proposition 2 (Consistency): If $\mathbf{f}, \mathbf{h}$ are Lipschitz and $\epsilon_w + \epsilon_v + \epsilon_0 < 0.5$, the state estimates converge to a neighborhood of the true state.

VII. Simulation Setup and Outlier Analysis

We consider the one-dimensional heat conduction equation, which describes the temperature distribution in a rod over time:

$$\frac{\partial T}{\partial t} = \alpha\frac{\partial^2 T}{\partial x^2} \tag{27}$$

where $T(x,t)$ is the temperature at position x and time t, and α is the thermal diffusivity coefficient. For numerical implementation, the spatial domain $[0, L]$ is discretized into $n = 100$ equally spaced points with $L = 1$ m, and the time step is set to $\Delta t = 5 \times 10^{-5}$ s. The simulation runs for 1000 time steps.

The finite difference approximation yields

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \tag{28}$$

where $\mathbf{x}_k \in \mathbb{R}^{100}$ is the temperature profile at time k, $\mathbf{u}_k \in \mathbb{R}^2$ is the input from two heat sources, and $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, 0.001\,\mathbf{I})$ is the process noise.

Temperature is measured at $m = 9$ locations (positions $\{10, 20, \ldots, 90\}$) using the measurement model:

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \tag{29}$$

where $\mathbf{C} \in \mathbb{R}^{9\times 100}$ selects the measured states, and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, 0.001\,\mathbf{I})$ is the measurement noise. The ambient temperature is set to $T_{\text{ambient}} = 300$ K.

The process outliers are generated as (5), where $\beta = 0.05$ and $\mathcal{H}_k$ is uniformly distributed with a variance of 10.

We model measurement outliers as:

$$\mathbf{v}_k \sim \begin{cases} \mathcal{N}(\mathbf{0}, \gamma_v^2\mathbf{R}) & \text{with probability } \epsilon_v \text{ for each sensor} \\ \mathcal{N}(\mathbf{0}, \mathbf{R}) & \text{otherwise} \end{cases} \tag{30}$$

where $\epsilon_v = 0.05$ is the probability of a measurement outlier occurring for each sensor at each time step, and $\gamma_v = 5$ is the scale factor for the outlier magnitude.

The ensemble size is set to $N = 100$. Each member is initialized from the prior:

$$\mathbf{x}_0^{(i)} \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0), \quad i = 1, \ldots, 100 \tag{31}$$

To ensure robustness against initialization outliers, we compute the Mahalanobis distance for each member,

$$d_i = (\mathbf{x}_0^{(i)} - \bar{\mathbf{x}}_0)^T \mathbf{P}_0^{-1} (\mathbf{x}_0^{(i)} - \bar{\mathbf{x}}_0) \tag{32}$$

and retain only the $(1 - \epsilon_0)N = 90$ members with the smallest distances ($\epsilon_0 = 0.1$). The ensemble is then resampled to maintain $N = 100$.

A. Robust DH-LTS-EnKF Algorithm

The DH-LTS-EnKF operates in three main stages:

1. Robust Initialization: The ensemble is trimmed based on Mahalanobis distance as described above.

2. Process Outlier Detection and Mitigation: At each time step, process residuals are computed for each ensemble member:

$$\mathbf{r}_w^{(i)} = \mathbf{x}_{k|k-1}^{(i)} - \mathbf{A}\mathbf{x}_{k-1|k-1}^{(i)} - \mathbf{B}\mathbf{u}_{k-1} \tag{33}$$

Their weighted norms $\|\mathbf{r}_w^{(i)}\|^2_{\mathbf{Q}^{-1}}$ are calculated, and only the $(1 - \epsilon_w)N = 90$ members with the smallest norms are retained ($\epsilon_w = 0.1$).

3. Measurement Outlier Detection and Mitigation: Measurement residuals are computed as

$$\mathbf{r}_v^{(i)} = \mathbf{y}_k - \mathbf{y}_{k|k-1}^{(i)} \tag{34}$$

and their weighted norms $\|\mathbf{r}_v^{(i)}\|^2_{\mathbf{R}^{-1}}$ are used to trim the ensemble, retaining the $(1 - \epsilon_v)N = 90$ most consistent members ($\epsilon_v = 0.1$).

To effectively communicate the performance of the DH-LTS-EnKF compared to the standard EnKF, we recommend presenting the following results:

TABLE I
PERFORMANCE COMPARISON OF ESTIMATION METHODS

Metric	EnKF (No Outliers)	DH-LTS-EnKF (No Outliers)	DH-LTS-EnKF (With Outliers)
State MSE	0.066626	0.096377	0.115797
State Bias	0.035787	0.047161	0.062037
Output MSE	0.000997	0.001341	0.001464
Output Bias	0.001260	-0.003246	-0.004579

REFERENCES

[1] Evensen G. (2003), "The ensemble Kalman filter: Theoretical formulation and practical implementation", *Ocean Dynamics*, 53(4), 343–367.

[2] MacKinlay D. (2025), "The Ensemble Kalman Update is an Empirical Matheron Update", *arXiv preprint arXiv:2502.03048*.

[3] Duran-Martin G., Altamirano M., Shestopaloff A.Y., Sánchez-Betancourt L., Knoblauch J., Jones M., Briol F.-X., Murphy K. (2024), "Outlier-robust Kalman Filtering through Generalised Bayes", *arXiv preprint arXiv:2405.05646 [stat.ML]*. https://doi.org/10.48550/arXiv.2405.05646

[4] S. Wang and Z.-S. Ye, "Distributionally robust state estimation for linear systems subject to uncertainty and outlier," *IEEE Trans. Signal Process.*, vol. 70, pp. 452–467, 2022,

[5] K. Tan, P. Li, and T. Beckers, "Physics-constrained learning for PDE systems with uncertainty quantified port-Hamiltonian models," *arXiv preprint arXiv:2406.11809 [cs.LG]*, pp. 1–15, Jun. 2024, doi: 10.48550/arXiv.2406.11809.

[6] Trujillo L., López U., Legrand P. (2020), "SOAP: Semantic outliers automatic preprocessing", *Information Sciences*, 526, 86–101.

[7] J. AlMutawa, "Parameter estimation of multisensor state space models with outlier contamination," in *Proc. 10th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Valletta, Malta, 2024, pp. 2146–2151, doi: 10.1109/CoDIT62066.2024.10708366.

[8] Rousseeuw, P. J., Ruts, I., and Tukey, J. W., "The bagplot: a bivariate boxplot", *The American Statistician*, vol. 53 , pp. 382–387, (1999).

[9] Y. Liu and R. R. Bitmead, "Dynamic trimmed likelihood for robust state estimation," *Automatica*, vol. 128, p. 109565, 2021.

[10] H. R. Patel, "Optimal intelligent fuzzy TID controller for an uncertain level process with actuator and system faults: Population-based metaheuristic approach," *Franklin Open*, vol. 4, p. 100038, 2023.