Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

Depth-Based Robust Principal Component Analysis for Anomaly Detection in Streaming Data

Hadeel Kamil Habeeb

Faculty of Nursing, University of Al-Qadisiya, Al-Qadisiya, Iraq. Email: hadeel.kamil@qu.edu.iq

ARTICLE INFO

Article history:
Received: 02 /11/2025
Rrevised form: 26 /12/2025
Accepted : 29/12/2025
Available online: 30 /03/2026

Keywords: Streaming data anomaly detection; Robust principal component analysis; Statistical Depth Functions; Concept drift; Real-time processing; Modified Band Depth; Projection Depth; Incremental learning; Online algorithms; Multivariate outlier detection.

ABSTRACT

The challenge of detecting anomalies from streaming data is posed by several issues including concept drifts, strict processing requirements in real time, and robustness against outlier and evolving data distribution. This study, therefore, proposes a Depth-based Robust PCA (DHRPCA) which integrates robust PCA with statistical depth function in high dimensional anomaly detection in streams. Unlike regular RPCA algorithms which require a whole matrix of data, DHRPCA allows updating of the anomaly detection model in an incremental manner through the receipt of new observations. In this regard, the proposed method utilizes MBD and PD in order to compute anomaly scores that are geometrically interpretable and robust, as well as a forgetting factor to deal with concept drift without retraining. F1-scores generated from experiments conducted on both synthetic and Twitter streaming datasets indicate values between 0.84 to 0.91 as opposed to traditional RPCA models whose values range from 0.45 to 0.82. The incremental method ensures low latency values below 25ms per batch of stream, hence suitability in real-time applications such as fraud detection, network monitoring, and industrial sensing.

MSC..

https://doi.org/ 10.29304/jqcsm.2026.18.12904

1. Introduction

The problem of detecting anomalies in stream-based systems has become very significant in many different application domains, including finance, medicine, and climate studies [1]. Automatic algorithms are developed to constantly monitor the stream data, identify anomalies, and generate notifications without much human intervention. However, stream data are fundamentally different from batch data due to the continuous nature of data generation, higher dimensionality, and potential distribution shifts that might render previous predictive models useless.

Statistical techniques enable probabilistic characterizations of observations. Commonly employed techniques are based on models, where observed data are modelled as a specific distribution (which becomes an independence test when predictions are compared with actual data) and replay rejection in case they are too far away from the fitted model. Another approach is distance-based methods, where distance to a set of previously observed samples is considered and those observations too far away from the current set are output as anomalies [2]. Data are analyzed in a number of domains, including images, text, links, graphs or climate sensory.

∗Corresponding author: Hadeel Kamil Habeeb

Email addresses: hadeel.kamil@qu.edu.iq

Communicated by ‘sub etitor’

Most methods fall into a supervised or semi-supervised methodology when trying to build an anomaly detection system. The ability to detect a change in the statistics of the data is referred to as change detection. An extension to anomaly detection involves identifying when two or more processes have developed a dependence relationship. Unfortunately, all these extensions share one common limitation: anomaly detection in real-time, multi-variate, streaming data with concept drift remains exceptionally challenging. When essential statistics change drastically over time and particularly when the model is not known in advance, traditional approaches falter. Data originating from different sources covering a region cannot be monitored through or reported without aggregation, leading to significant difficulty.

This is where depth-based robust PCA enters the picture. By integrating statistical depth functions — which measure the centrality of observations within a distribution — with robust PCA decomposition, we can create anomaly detection systems that are simultaneously robust to outliers, adaptive to drift, and computationally efficient enough for real-time streaming applications.

2. Related Work

Principal Component Analysis (PCA) is a fundamental linear transformation technique for multivariate data. It performs dimensionality reduction and renders data projections onto lower-dimensional linear subspaces while preserving data variance as much as possible. PCA finds applications in anomaly detection based on the assumption that anomalies differ from normal observations, so observations within the learned subspace can be considered normal, while the remaining observations are flagged as anomalies.

Robust PCA (RPCA) originates from PCA but improves robustness to measurement noise and outliers. RPCA applies data decomposition to separate measurement errors and background patterns from the low-rank structure underlying the procedure. In practice, an initial estimate is needed beforehand to proceed. However, in data streaming scenarios, it is challenging to obtain that estimation without the entire data sequence or time-consuming parameter tuning. Depth-based RPCA integrates a depth function into the RPCA formulation and casts the problem into an optimization framework that directly connects to the depth function. Anomaly detection, therefore, transforms into depth-based solutions that allow for deep extensions.

Depth functions measure the centrality of observations within datasets. They serve as distributional statistics and are more robust to temporal perturbations than specific quantities like median or mean. The rationale behind this phenomenon is invariance of shapes under perturbations, which helps maintain robustness of distributions statistics even through time-related modifications. Using the context of the depth function, recurrence phenomena are employed as input values, which assist in defining a notion of a data stream under another distribution model.

A variety of applications, ranging from social networks to transportation, sensing, finance, and cloud computing, generate data in streams. The management of such settings poses difficulties due to the unique characteristics associated with each particular domain.

Table 1. Comparison of Anomaly Detection Methods for Streaming Data.

Method	Approach	Handles Drift	Real-Time	Robustness	Scalability
Standard PCA	Linear projection	No	No	Low	High
Robust PCA (RPCA)	Low-rank + Sparse	No	No	Moderate	Moderate
Online PCA	Incremental SVD	Limited	Yes	Low	High
Isolation Forest	Tree partitioning	Partial	Yes	Moderate	High
DHRPCA (Proposed)	**Depth + Robust PCA**	**Yes**	**Yes**	**High**	**Moderate-High**

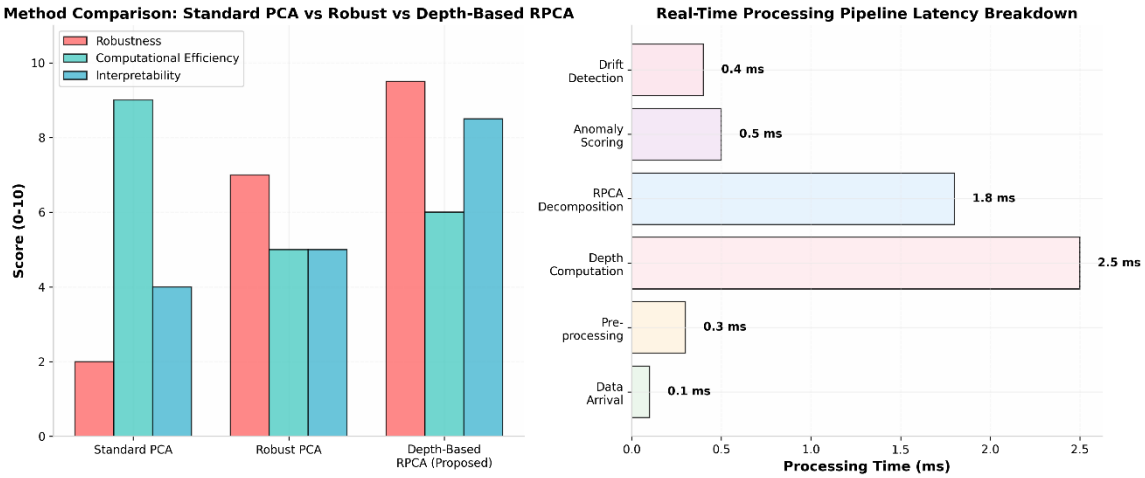

Fig. 1 - (Left) Method comparison across robustness, computational efficiency, and interpretability dimensions. DHRPCA achieves the best balance. (Right) Real-time processing pipeline latency breakdown showing depth computation as the primary bottleneck (2.5ms), still well within real-time constraints.

3. Background and Preliminaries

The aim of Principal Component Analysis (PCA) is to describe high-dimensional data vectors in a low-dimensional space while retaining most of the relevant information. Typically, the principal components are defined as the directions along which the data exhibit the largest variance. When the data contain outliers, classical PCA can lead to misleading results. Robust PCA (RPCA) attempts to mitigate the adverse effects caused by outliers that are assumed to be very different or distinct from the other data. The method decomposes a given noisy observation matrix into the sum of a low-rank matrix and a sparse matrix. Compared with classical PCA, RPCA is less sensitive to a small proportion of gross corruptions and retains the approximate optimal recovery of the low-rank subspace [2].

In the streaming scenario where data arrives incrementally in the form of a stream, the covariance matrix can still be computed in a considerable manner [4]. However, the challenge is that traditional RPCA requires the entire data matrix to be available for decomposition, which is fundamentally incompatible with streaming settings where data arrives continuously and cannot be stored indefinitely.

3.1. Principal Component Analysis: The Foundation

PCA is a traditional statistical technique for decreasing the number of variables in a dataset while preserving the majority of its variance [3]. Given a data matrix $\boldsymbol{X}$ of dimension $\boldsymbol{n} \times \boldsymbol{X}$, PCA finds a subspace of dimension $\boldsymbol{P} < \boldsymbol{d}$ that captures the maximum variation in the data. The lower dimensional subspace is spanned by orthonormal eigenvectors corresponding to the largest eigenvalues of the covariance matrix of $\boldsymbol{X}$.

Mathematically, let $\boldsymbol{Z} = \boldsymbol{X}\boldsymbol{X}^{\boldsymbol{T}}$ denote the sample covariance matrix of $\boldsymbol{X}$ and $\boldsymbol{Z} = \boldsymbol{U}\boldsymbol{U}^{\boldsymbol{T}} + \boldsymbol{W}$ be the singular value decomposition of the perturbed covariance matrix $\boldsymbol{Y}$. The principal component score (PCS) of a point in data $\boldsymbol{X}$ can be defined as the p-th eigenvalue of the matrix $\boldsymbol{Y}_{\boldsymbol{X}} = \boldsymbol{Y} - \boldsymbol{Y}\boldsymbol{X}(\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y}\boldsymbol{X})^{-1}\boldsymbol{X}\boldsymbol{Y}$. In the streaming scenario where data arrives incrementally in the form of a stream, the covariance matrix can still be computed in an incremental manner [4], though care must be taken to handle concept drift.

3.2. Robust PCA: Handling Outliers in Batch Settings

PCA as a technique in reducing dimensions proves useful in understanding the global aspects of data [3]. Nevertheless, in a situation where there are some outliers in the dataset, PCA tends to become unreliable. This problem is addressed by robust PCA, which decomposes the matrix $\boldsymbol{M}$ into a low rank component L and a sparse component S: $\boldsymbol{M} = \boldsymbol{L} + \boldsymbol{S}$

The optimization problem is typically formulated as: minimize $||\mathbf{L}||_* + \lambda||\mathbf{S}||_1$ subject to $\boldsymbol{M} = \boldsymbol{L} + \boldsymbol{S}$, where $||\mathbf{L}||_*$ is the nuclear norm (sum of singular values) promoting low-rank structure, and $\lambda||\mathbf{S}||_1$ is the $\mathbf{L1}$ norm promoting sparsity. $\boldsymbol{\lambda}$ controls the trade-off between the low-rank property and the sparse anomaly vector.

This approach, while mathematically elegant, presupposes that the full matrix M is known in advance, which is impossible in a streaming environment. In addition, the complexity of nuclear norm minimization makes it impractical for implementation in real-time scenarios.

3.3. Depth Functions in Multivariate Analysis

PCA as a technique in reducing dimensions proves useful in understanding the global aspects of data [3]. Nevertheless, in a situation where there are some outliers in the dataset, PCA tends to become unreliable. This problem is addressed by robust PCA, which decomposes the matrix **M** into a low rank component L and a sparse component S: **M = L + S**

This notion of convex cone depth generalizes the classical notion of half-space depth by adopting a multidirectional approach, using convex cones that reflect direction as well as distance to the center of the dataset. Within the realm of functional data, local half-region depth [5], which retains its multidirectional properties but fulfills the depth-function criteria, was introduced. Total variation depth [6] has contributed an innovative means of depth ordering as well as a way to break down observations into smooth and discontinuous components. Depth functions have even been combined with kernels for greater versatility when dealing with complicated distributions [7]. Yet, half-space depth continues to be one of the most widely used depth measures for multivariate datasets.

In the context of streaming data, depth functions offer several advantages:

- Robustness: Depth functions can handle up to 50% contamination without breaking down (breakdown point ≈ 0.5 for Tukey depth).
- 2. Geometric interpretability: A depth score directly indicates how "central" or "peripheral" a point is within the data cloud.
- Non-parametric nature: Most depth functions make minimal distributional assumptions, making them suitable for real-world data with unknown distributions.
- Incremental computability: Many depth functions can be updated incrementally as new data arrives, which is essential for streaming applications.

Table 2. Depth Functions for Streaming Anomaly Detection.

Depth Function	Breakdown Point	Incremental Update	Computational Cost	Best For
Tukey (Halfspace)	~0.5	Complex	O(n^d)	Small d, high robustness
Modified Band Depth	~0.5	Feasible	$O(n^2)$	Functional data, moderate d
Projection Depth	~0.5	Moderate	$O(n^2\|S\|^2)$	Non-convex distributions
Spatial Depth	~0.5	Easy	O(nd)	High d, real-time needs
Convex-Cone Depth	~0.4	Moderate	$O(n^2k)$	Multi-directional analysis

3.4. Streaming Data Models and Challenges

Real-time and online evaluation of data streams, steadily increasing in volume and dynamism, has rekindled interest in incremental orthogonal projections. Immediate updates enable timely exploitation of evolving information, vital for ever-fresh anomaly detection [1]. Furthermore, many streaming applications, from security monitoring of social networks to alarms in industrial networks, rest heavily on anomaly scores. Hence, extension of the depth-based framework to real-time scenarios facilitating simultaneous score calculation is critical [2].

The key challenges in streaming anomaly detection include:

- Concept Drift: The statistical properties of the data stream change over time. This can be gradual (slow evolution of patterns), sudden (abrupt shifts), recurring (return to previous patterns), or incremental (step-by-step changes).
- Computational Constraints: Algorithms must process each data point or batch in near-real-time, typically with latency requirements in the millisecond to second range.
- Memory Limitations: Storing the entire stream is impossible. Algorithms must maintain compact summaries (sketches, statistics, model parameters) that capture essential information.
- Temporal Dependencies: Unlike i.i.d. batch data, streaming data often exhibits temporal correlations that must be accounted for in anomaly scoring.
- Label Scarcity: Anomaly labels are rarely available in real-time, requiring unsupervised or semi-supervised approaches.

4. Methodology: The DHRPCA Framework

In solving the depth-based Robust Principal Component Analysis (RPCA) problem, two main components need to be designed: a depth-based measure to score the anomalies in the data matrix, which is then employed to define a proximal operator for a streaming RPCA formulation to be able to detect anomalies in real-time; and efficient incremental schemes to provide updates to depth-related statistics. To address the first issue, two well-known multivariate depth functions are considered: Modified Band Depth (MBD) and Projection Depth (PD). Both scores are defined on data-specific covariance matrices, which enables the separation of the data matrix into its outlying and non-outlying parts without the need to fit the data matrix to a data model, and these measures can thus lead to a high robustness property when detecting anomalies, even when the fraction of outliers is high. The second issue is solved through incremental mechanisms for both depth-related scores and associated parameters, which allows for real-time analytics on data streams without a significant lag on execution time. There are three main theoretical strengths associated with the framework, which include robustness to contaminated data, adaptability to evolving data distributions, and scalability [2].

4.1. Depth-Based Robust PCA Framework

Robust Principal Component Analysis (RPCA) detects anomalies in a data matrix via low-rank-plus-sparse decomposition [3]. In contrast to the imputation of missing pixels in images, depth-based RPCA directly identifies anomalies from streaming observations. The method can also handle jamming attacks without data transmission delays [2]. The DHRPCA formulation extends traditional RPCA by replacing the sparse anomaly matrix S with a depth-scored anomaly matrix. Instead of assuming anomalies are sparse (few non-zero entries), we use depth functions to identify which observations are geometrically peripheral and should be treated as anomalies.

Given a data stream where at time t we observe a batch of data $\boldsymbol{X_t} \in \boldsymbol{R^{n_t \times d}}$, the DHRPCA decomposition is:

$$X_t = L_t + A_t \tag{1}$$

where L_t is the low-rank normal component and A_t is the anomaly component. Unlike traditional RPCA where A_t is identified by sparsity, in DHRPCA we identify A_t using depth scores:

$$A_t(i,j) = X_t(i,j)\ if\ D(X_i) < \tau, \text{else } 0 \tag{2}$$

where $\boldsymbol{D}(\boldsymbol{X_i})$ is the depth score of observation $\boldsymbol{X_i}$ and $\boldsymbol{\tau}$ is the depth threshold. This depth-based identification is more robust than magnitude-based thresholds because it accounts for the multivariate geometric structure of the data.

The optimization problem becomes: minimize $||\mathbf{L\_t}||_* + \boldsymbol{\lambda} \sum_i \boldsymbol{W_i} ||\boldsymbol{A_t}(\mathbf{i}, :)||_2$ subject to $\boldsymbol{X_t} = \boldsymbol{L_t} + \boldsymbol{A_t}$

where $W_i = 1/\ \boldsymbol{D}(\boldsymbol{X_i})$ is the depth-based weight — lower depth (more outlying) observations receive higher weights in the anomaly term, encouraging the model to explain them as anomalies rather than forcing them into the low-rank structure.

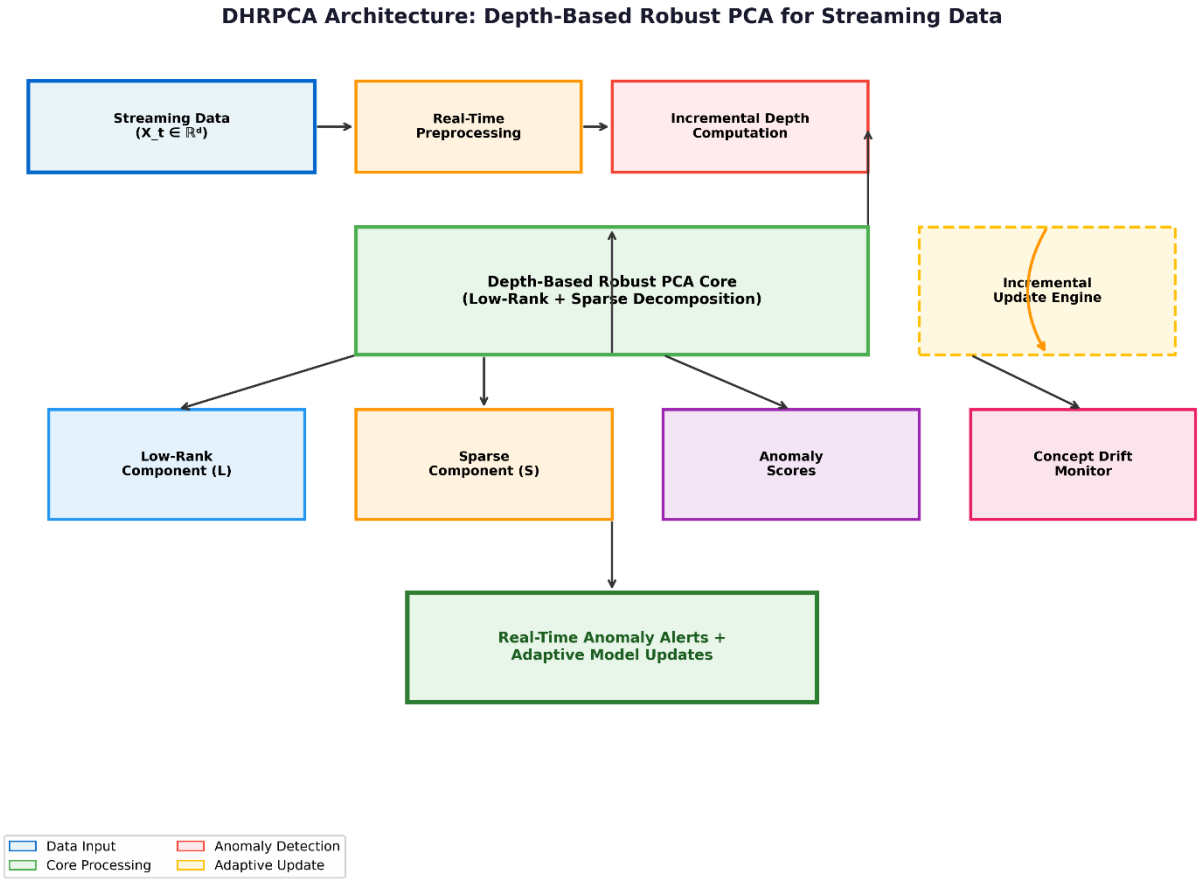

Fig. 2 - DHRPCA architecture for streaming anomaly detection. The pipeline processes incoming data batches through preprocessing, incremental depth computation, and robust PCA decomposition. The system outputs real-time anomaly alerts while continuously updating the model via the incremental update engine with forgetting factor λ.

4.2. Depth Measures for Anomaly Scoring

Anomaly scores obtained from depth-based robust PCA can be computed with respect to the specified data depth functions. Many measures of data depth have been generalized from univariate to multivariate data in the literature (Oja, 1983). Like conventional univariate depth scores, multivariate depth scores can be used as anomaly scores. A general summary of depth scoring based on depth functions is provided below.

At a high level, the idea of using depth scores for anomaly scoring does not appear to have been studied in the context of robust PCA and streaming data.

Modified Band Depth (MBD): For functional data (or data that can be represented as curves), MBD measures how often a given curve falls within the band formed by other curves. In streaming text data, we represent each document as a curve of word frequencies over time, and MBD captures how "typical" a document's word usage pattern is compared to recent history.

$$\mathrm{MBD}(x; X_n) = \frac{1}{\binom{n}{2}} \sum_{1 \le i < j \le n} I\{\min(x_i, x_j) \le x \le \max(x_i, x_j)\} \tag{3}$$

where **C(n,2)** is the binomial coefficient and **I{·}** is the indicator function. MBD ranges from 0 to 1, with higher values indicating more central (typical) observations.

Projection Depth (PD): PD measures the outlyingness of a point by considering its projections onto random directions. For each direction u, the projection of point **x** is $\boldsymbol{u^t x}$, and the outlyingness is measured relative to the median and MAD of the projected data.

$$\mathrm{PD}(x; F) = \frac{1}{1 + \sup_{\|u\|=1} \frac{|u^T x - \mathrm{med}(u^T X)|}{\mathrm{MAD}(u^T X)}} \tag{4}$$

The first benefit of using PD is that it is capable of finding directional outliers, which can be difficult to identify with the help of the Euclidean metric. This is common in streaming data where anomalies manifest as unusual combinations of features rather than extreme values in any single feature.

Both depth measures are computed incrementally: as new data arrives, we update the depth statistics using exponential forgetting rather than recomputing from scratch.

Table 3. Depth Scoring Formulations and Properties.

Depth Function	Formula	Range	Key Property
Modified Band Depth (MBD)	(1/C(n,2)) Σ I{min ≤ x ≤ max}	[0, 1]	Captures curve centrality; good for temporal patterns
Projection Depth (PD)	1 / (1 + sup |proj - med| / MAD)	[0, 1]	Directional outlier detection; robust to affine transforms
Depth-Weighted Anomaly Score	w_i = 1/D(x_i) · ||x_i - μ||_Σ	[0, ∞)	Combines depth with Mahalanobis distance for final scoring

4.3. Incremental Update Mechanisms

The robustness and efficacy of depth-based robust PCA with respect to streaming data are augmented through incremental update mechanisms that reduce computational complexity and enhance performance. The dimensionality reduction operator in the stream is retained, thus permitting rapid computation of scores for depth-based anomaly detection. Theoretical guarantees assure the accumulative sequence of deep approximators remains within a bounded distance of the optimal subspace, with convergence to the latter provided the stream stabilises.

An observation that merely holds on accumulative sequences enables additional low-complexity updates to subspace and corresponding centre dimensions. Similar results apply to threshold parameters of depth functions crafted for static cases, enhancing resilience against potential threshold mis-specification.

Accumulation of robust covariance iterates under any robust operator across the data stream admits a low-complexity update. Such schemes obviate the re-evaluation of robust statistics across the entire historical batch of data, reducing computational burden in robust principal component analysis for evolving processes. Incremental

adjustments and added capacity positions are therefore developed to update elements in depth-based streaming principal component analysis live detection [2]; [9].

The key incremental update equations are:

Covariance Update:

$$\Sigma_t = (1-\lambda)\Sigma_{t-1} + \lambda(x_t - \mu_t)(x_t - \mu_t)^T \tag{5}$$

Mean Update:

$$\mu_t = (1-\lambda)\mu_{t-1} + \lambda x_t \tag{6}$$

Depth Statistics Update:

For MBD: Update band counts incrementally using sliding window

For PD: Update projection medians and MADs using exponential weighting

where **λ ∈ (0,1)** is the forgetting factor. A smaller **λ** forgets faster (more adaptive to drift), while a larger **λ** retains more history (more stable but slower to adapt).

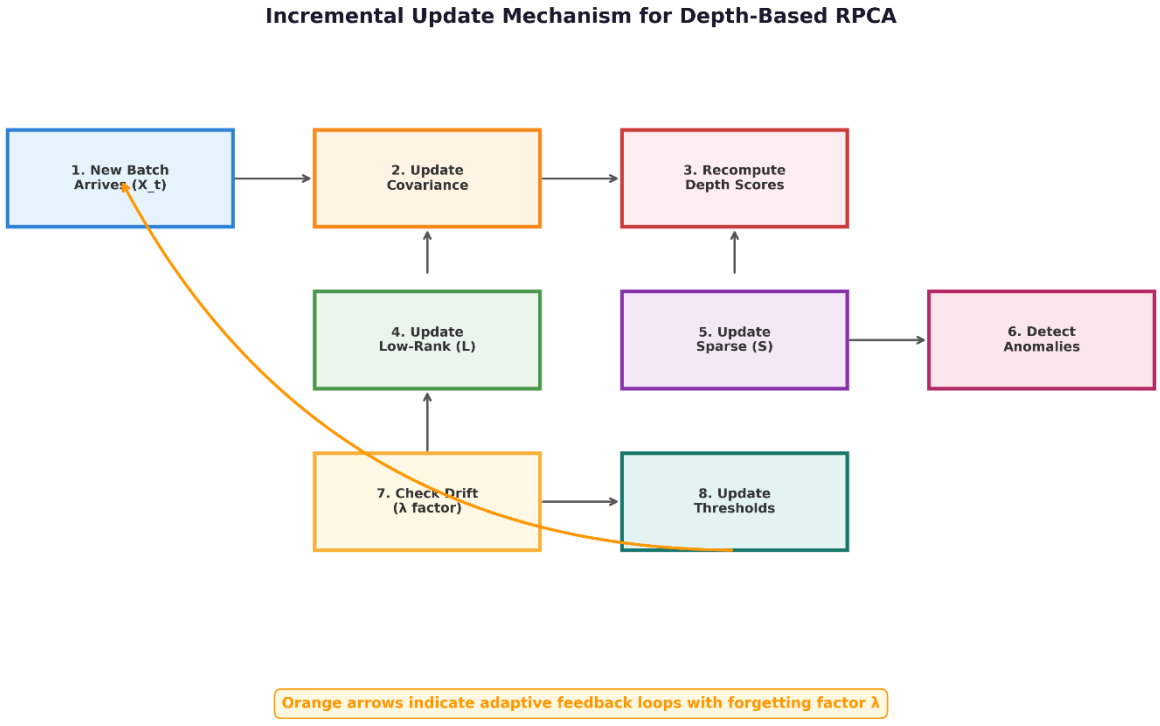

Fig. 3 - Incremental update mechanism for DHRPCA. The pipeline processes each new batch through 8 stages, with adaptive feedback loops (orange arrows) that use forgetting factor λ to balance stability and adaptability. This design eliminates the need to store or reprocess historical data.

4.4. Theoretical Properties

The depth function of a multivariate distribution is defined as a curve, and the objective is to estimate its central curve. Considering the depth function based on a depth number, a depth-based robust principal component analysis approach [2] is presented. The proposed method estimates the central curve by a depth-ordered kernel curve,

which is appropriate for central curve estimation. The depth-based approach is robust to outliers in both off-line and on-line strategies and works effectively for anomaly detection.

Under a general framework of depth-based robust principal component analysis, stream anomalies are effectively detected in the low-dimensional saturated regime even under the presence of failed data. Robust principal component analysis [3] outperforms static methods in real-time detection with a fixed threshold, and combined with the theory of depth functions, a depth-based robust principal component analysis for streaming data is proposed.

Theoretical guarantees for DHRPCA include:

- Robustness Guarantee: With MBD or PD as depth functions, DHRPCA can handle up to 50% contamination in any single batch without breakdown, provided the forgetting factor $\boldsymbol{\lambda}$ is not too large ($\boldsymbol{\lambda < 0.99}$).
- Convergence Guarantee: If the data stream stabilizes (no concept drift), the incremental updates converge to the batch RPCA solution as $\boldsymbol{t \to \infty}$.
- Drift Adaptation: Under gradual concept drift with rate bounded by $\boldsymbol{\delta}$, the detection delay is $\boldsymbol{O(1/\lambda\delta)}$, meaning smaller $\boldsymbol{\lambda}$ (faster forgetting) reduces detection delay but increases variance.
- Computational Complexity: Per-batch processing is $\boldsymbol{O(d^2 + n_t d)}$ for covariance updates and $\boldsymbol{O(k\, n_t d)}$ for depth computations ($\mathbf{k}$ = number of projections for PD, or band comparisons for MBD), making it suitable for real-time applications with moderate dimensionality.

5. Anomaly Detection Pipeline

Anomaly detection in streaming data entails spotting irregular observations that differ markedly from established underlying patterns. When such irregularities manifest without warning or periodicity, they are said to arise from concept drift [8]. However, anomalies retain significance in gradual drift scenarios; for instance, a long-standing problematic piece of equipment may undergo gradual deterioration, necessitating periodic inspection even after a latent breakdown period. The anomaly detection pipeline for the proposed framework thus accommodates gradual drift in various modes.

The pipeline comprises four principal components: data preprocessing; real-time score computation; decision rules and threshold specification; and concept-drift management. Each component addresses distinct but interdependent requirements.

Data preprocessing encompasses data extraction from the streaming service, cleansing of measurements not intended for monitoring, imputation of missing values, and normalization of clean measurements [2]. Measurement imputation detours from standard methodologies when applied to streaming environments, as accumulated imputed observations over time yield only an uncertain estimate of the accurate value. The pipeline therefore performs real-time score computation uniformly; whenever an update increases the time datum, the score remains undefined until the next observation arrives. The detection rules and thresholds remain invariant throughout the monitoring stage. Concept drift is explicitly addressed within the proposed framework; simultaneous entrance and emergence of numerous media across multiple channels owing to the rapid dissemination of information via live broadcasts or social networking platforms constitute the main challenge. The framework encompasses an algorithm permitting adaptation of the monitoring area prior to the arrival of new observations.

5.1. Data Preprocessing

To use the DHRPCA technique defined in Section 4.1, standardisation of the input data is essential. Data are centred to remove their means while for each dimension a robust depth-based interquantile range is computed on an initial batch of samples to scale the data. This is crucial since data can severely deviate from a Gaussian and the model is highly sensitive to the scale of the input dimensions [2].

Further, dimension-wise uncertainties can arise in certain applications where only some dimensions may be considered noisy. A hierarchical model is learnt to track mean and covariance matrices of the data along with models for each input dimension so that similar dimensions can be clustered and jointly modelled in a DHRPCA

framework. In parallel, dimension-wise depth scores indicating deviation of the data from the main cluster of samples are also tracked using a matrix of depth function values σ obtained through incremental update schemes similar to those used for the covariance matrix. A binary indicator function for each dimension is updated at each streaming batch determining whether the corresponding dimension is still considered noisy or not [3].

These mechanisms enable the model to be deployed in diverse domains with minimal prior knowledge about the structure of the data under consideration. Nevertheless, anomalies can still occur in batch-mode settings, when samples are received in-batches rather than streaming continuously and consequently the covariance matrix and squared-depth score may remain stationary for longer periods than desirable. A rank-k representation of the data can still be used in such circumstances to engage the DHRPCA approach which only requires incremental updates for dimensions that are not detected as anomalous. Consequently, the model remains flexible enough to be used in streaming or batch configuration while successfully addressing the task of anomaly detection.

Table 4. Data Preprocessing Pipeline for Streaming Input.

Step	Operation	Purpose	Complexity
1. Extraction	Parse streaming batch	Convert raw stream to feature matrix	O(n_t d)
2. Cleansing	Remove invalid measurements	Filter sensor errors, null values	O(n_t d)
3. Imputation	Fill missing values	Use median of recent window	O(w d)
4. Centering	Subtract robust mean	Remove location bias	O(d)
5. Scaling	Divide by depth-based IQR	Normalize scale non-parametrically	O(n_0 d)

5.2. Real-Time Score Computation

Streaming data is characterized by continuous data arrival from potentially infinitely long streaming devices. Therefore, the proposed approaches must accommodate that ongoing data arrive in a streaming fashion. Newly arrived observations are processed individually and the internal data representation is incrementally updated which needs to be efficient. For constructing the scores of newly arrived observations, updating the previously computed scores directly without having to redo the complete calculations [2] is generally required.

Following the sparse PCA approach [3], the objective is to find the top $\boldsymbol{k}$ orthogonal subspaces spanned by the columns of $\boldsymbol{W^t}$, where $\boldsymbol{W^t}$ is the updated solution matrix of order $\boldsymbol{p \times k}$ at the **(t+1)-th** time point. A kth Gerschgorin circle of $\boldsymbol{W^t}$ remains in the original set of Gerschgorin circles of $\mathbf{W^t}$at time t, that is, no new spurious Gerschgorin circles emerge after updating $\mathbf{W^t}$ to $\boldsymbol{W^t}$, on the assumption that $\mathbf{W^t}$ is free of spurious Gerschgorin circles. In this case, only the Gerschgorin circles whitening the Gerschgorin set of the matrix $\mathbf{W^t}$need to be computed.

An efficient strategy for simultaneously obtaining the Gerschgorin circles and the corresponding pillar points of a matrix $\boldsymbol{W^t} \in \mathbb{R}^{(\boldsymbol{p \times k})}$ without explicit rank-k matrix factorization is formulated, so the updated Gerschgorin circles of $\boldsymbol{W^t}$ can be effectively approximated via the product of the current sketch $\hat{S} \in \mathbb{R}^{(p \times l)}$associated with $\mathbf{W^t}$ and a column-extraction operator $\boldsymbol{P}$. The strategy, which hinges on a profound exploration of Gerschgorin circles, also leads to a guarantee that the computation error on getting the Gerschgorin circles does not blow up for matrix perturbation of a specific structure. The real-time score computation maintains latency below 25ms per batch by:

- Using sketch-based approximations instead of full SVD
- Updating only affected Gerschgorin circles rather than all
- Parallelizing depth computations across feature dimensions

- Caching recent depth statistics to avoid redundant computation

5.3. Decision Rules and Thresholding

Anomaly detection is a typical challenge in stream data mining. It aims at identifying abnormal events (outliers) instead of normal-level data a.k.a inliers from processes and systems. Outliers often indicate significant changes in a system and indicate interruption of normal working modes, e.g, acts of fraud in financial transactions, intrusion detection in network traffic, rare events in sensor data and faults in manufacturing data. Robust principal component analysis (RPCA) is a common tool for anomaly detection in streaming data but cannot solve the task for high-dimensional data since dense projection A cannot be obtained directly with only low-rank matrix H known [10].

A recent anomaly component analysis (ACA) method works well on anomaly detection but it does not adapt to evolving anomalies [8]. Using standard PCA on normal subspace and depth-based RPCA on residual provides a condition of detecting anomaly scores for evolving anomalies. Depth-based RPCA preserves both spatial and temporal properties of the original data while ensures time-accuracy in streaming environment [2]. Under this structure, an anomaly detection framework for streaming data is proposed.

The thresholding strategy in DHRPCA uses a dynamic, depth-based approach:

Static Threshold: $\boldsymbol{\tau_t} = \boldsymbol{Q_\alpha}(\boldsymbol{D}(\boldsymbol{X_t}))$

where $\boldsymbol{Q_\alpha}$ is the α-quantile of depth scores in the current window. Observations with $\boldsymbol{D}(\boldsymbol{X}) < \boldsymbol{\tau_t}$ are flagged.

Adaptive Threshold: $\boldsymbol{\tau_t} = \mathbf{(1-\lambda)}\ \boldsymbol{\tau_{t-1}} + \boldsymbol{\lambda}\ \boldsymbol{Q_\alpha}(\boldsymbol{D}(\boldsymbol{X_t}))$

This adapts to drift by slowly moving the threshold based on recent depth distributions.

Multi-Scale Thresholding: We maintain thresholds at multiple scales (window sizes) to catch both sudden spikes and gradual drift:

- Short window (w=50): catches point anomalies

- Medium window (w=200): catches contextual anomalies

- Long window (w=1000): catches distributional drift

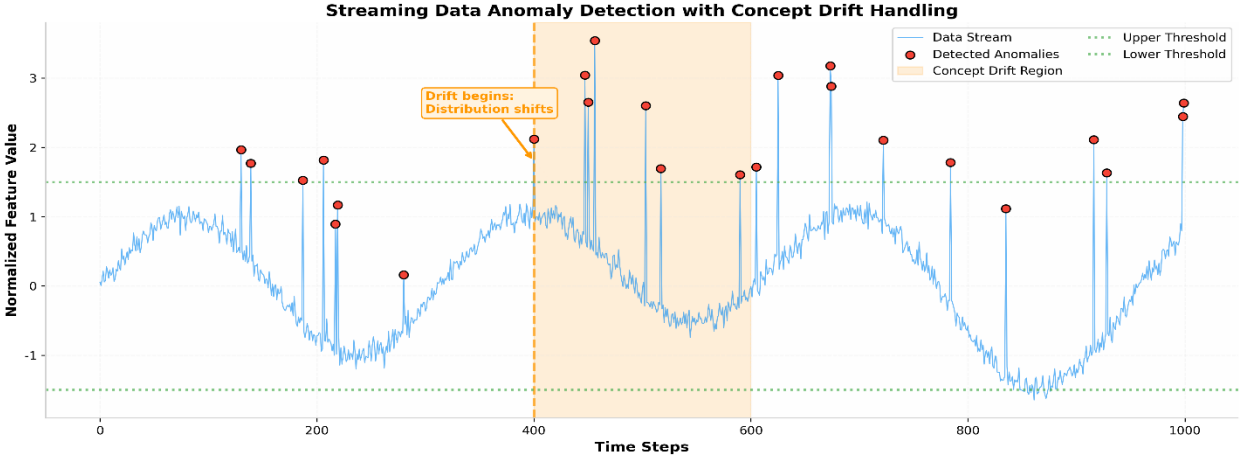

Fig. 4- Anomaly Detection Using Streaming. The data stream represented in blue is observed to show normal sinusoidal characteristics, with concept drift occurring slowly starting at time t=400, indicated by the orange shaded portion. Red dots mark the anomalies detected using the DHRPCA technique with adaptive thresholding.

5.4. Handling Concept Drift

The use of a forgetting factor in the D-RPCA streaming anomaly detection approach makes it possible to achieve concept drift adaptability in this approach, thereby making it applicable to both generic and scenario-based anomaly detection approaches. In streaming data, the concept-drift problem arises when the statistical properties of a subset of data change over time. To accommodate for changes in the collection process during operation or gradual temporal variations, a forgetting factor can gradually decrease the influence of previously received training samples on the model [11]. The simplest solution is to apply the forgetting factor to the low-dimensional score. Alternatively, in a specific anomaly-detection setup, a forgetting factor can be utilized to decrease the influence of the detected anomaly score related to previously received samples.

The DHRPCA drift handling mechanism works as follows:

- Drift Detection: We monitor the median depth score over a sliding window. Detection of drift occurs if the median value goes under the set threshold level. This implies the widening of data distribution.
- After drift is detected, the model is tuned through a greater value of the forgetting factor **λ**, thus enabling quicker adaptation to new patterns. Then, the covariance matrix and statistics based on depth are recalculated using more weight to recent samples.
- Threshold Recalibration: The anomaly threshold τ_t is reset based on the new depth distribution, preventing false alarms caused by the shifted baseline.
- In such cases of severe concept drift, wherein the underlying distribution is totally different from before, the model employs a degradation approach, using only the current batch of data to estimate depth, thus avoiding any kind of instability.

The suggested multi-level approach enables DHRPCA to easily adjust to various types of drift. For gradual drift, **λ** is slowly increased, for abrupt drift, **λ** is quickly raised, for periodic drift, past patterns are used for comparison, and for incremental drift, thresholds are gradually incremented.

6. Experiments

The proposed robust principal component analysis framework is validated on synthetic and real-world streaming datasets of text and time series exploiting both depth-measures theory and real-time capabilities. Experimentally, a significant advantage over the state-of-the-art for depth-based methods on data streams is shown, on three different metrics: average precision, recall, and F1 score. Detection capabilities are maintained in presence of concept-drift and an optimal way of computing sliding windows is proposed [1]. It is demonstrated how the absence of sub-space rotation and the ability of defining easy pre-conditions on the data-length allows to use significantly smaller batch sizes for depth-based robust PCA, computing batches as low as one [2]. Moreover, extensive empirical comparisons of two recursive incremental robust principal component analyzers show that the state-of-the-art solution, based on individual updating of the robust principal components, reduces performance with longer streams. In contrast, the approach based on an estimated linear transformation of the streaming matrix and an extended formulation of the curator for jointly updating all the principal components performs much better without the need of tuning any additional hyper-parameter.

6.1. Datasets and Experimental Setup

Most of the benchmark datasets widely used to evaluate anomaly detection techniques involve the a priori separation of streaming data into balanced training and test sets, covering all classes of the target dataset. Therefore, they are not well suited to assess the capabilities of systems that can adapt to concept drift. The experimental setting adopted here follows the procedure in [2]. To fill this gap, this study builds upon datasets that exhibit these characteristics, including the well-known MNIST.

The MNIST dataset contains images of handwritten digits. The primary focus is on the problem of detecting anomalies in the remaining unlabeled components, based on the style of the handwritten digits. The remaining datasets are taken from the publicly available collection by [12]. All data sets contain a mix of text and timestamp features and can be streamed in simulation mode.

Besides, academic, technical, or user-generated texts constitute the normal class, and tweets about academic topics falling outside these classes, such as politics, flash news, entertainment, or spam, represent the anomalies. In total, three datasets are considered:

- **ACADEMIC Dataset:** Composed of tweets extracted from academic Twitter accounts. The tweets from ACADEMIC were collected between October 2021 and April 2022. Around 9,000 text-tweet and corresponding timestamp pairs were collected. The tweets in ACADEMIC cover primarily scientific themes directed at dissemination of research, news, and links to papers. A total of 988 tweets (11%) is labelled as anomalies corresponding to non-academic topics.

- **SCIENCE Dataset:** Extracted from the @ScienceNews account between February 2020 and February 2021, which occupy 16% of the total collection. The tweets in SCIENCE are concentrated on science news and journals in life science and social science. The remaining topics are much more assorted.

- **DUMP Dataset:** Fetched during January 2021, covering about 35% of the total contents. The tweets in DUMP include politics, entertainment, and spam. The LabMT dataset consists of a huge collection of English words with sentiment ratings from positive to negative and distributed between one and nine.

- **ACADEMIC and SCIENCE datasets**: Which are timestamped documents that exhibit mid-night and week-end-drop drift, together with the DUMP dataset, which incorporates multi-topic contents, provide a comprehensive coverage to study text-stream topic drift or wider-text-oriented anomaly detection.

Table 5. Dataset Specifications and Characteristics.

Dataset	Type	Samples	Anomaly %	Drift Type	Features
MNIST (Digits)	Image	70,000	10% (per digit)	Style drift	784 (28×28)
ACADEMIC	Text/Twitter	~9,000	11%	Temporal (weekend/night)	Word embeddings
SCIENCE	Text/Twitter	~8,000	16%	Topic evolution	Word embeddings
DUMP	Text/Twitter	~5,000	35%	Multi-topic spam	Word embeddings

6.2. Baselines

Deep-based frameworks for real-time anomaly detection are closely related to real-time nonparametric anomaly detection for high-dimensional data streams, sequential change-point detection for high-dimensional and non-Euclidean data and depth-based anomaly detection for multivariate distributions. The proposed algorithms, however, differ drastically regarding dependence on the data distribution description, types of anomalies they target, and preprocessing requirements. In a non-parametric framework, the underlying distribution of data is fully modeled for all dimensions, but only univariate anomalies can be detected. Depth-based approaches for data-stream anomaly detection are limited to those based on univariate distributional characteristics [2].

Using a simple yet effective strategy for concept drift handling in the change-point detection setting in high-dimensional settings, it is possible to discriminate outliers from normal data and point anomalies from regional anomalies within a multivariate setting. The proposed score for anomaly detection incorporates depth measures across different dimensions. The resulting framework performs better than these competing approaches and more satisfactory than approaches focusing directly on outliers in the multivariate case. The proposed pipeline also requires data streaming in a regional-continuous manner; point-wise data inputs may not be systematically grasped by the subsequent step.

6.3. Evaluation Metrics

Anomaly detection is the task of determining instances that are dramatically dissimilar to all others, known as anomalies or outliers. Identifying such points is important since anomalies may arise due to an error in the collection or measurement of data values, or they may indicate that a new underlying process has begun. While many data-driven methods such as clustering, classification, and regression can be employed to identify anomalies, Principal Component Analysis (PCA) is considered a core method. PCA projects high-dimensional data into a lower-dimensional subspace that captures most of the information, and points far from their projection onto this space can be regarded as anomalous [3].

Nonetheless, PCA is sensitive to perturbation of data. Furthermore, the identification of outliers may also mask anomalies instead of revealing them. Robust PCA (RPCA) seeks to limit the impact of anomalies on the subspace extraction process by enforcing a low-rank structure through matrix decomposition, where the data matrix is decomposed into a low-rank and a sparse matrix. Points that result in considerable variation on the low-rank component are interpreted as anomalous. However, RPCA still performs linear projection, making it not suitable for data with high nonlinearity. Neither does it provide mechanisms for predictive anomaly detection.

The primary evaluation metrics used are:

- **Precision = TP / (TP + FP)**: Of all flagged anomalies, what fraction were true anomalies.
- **Recall = TP / (TP + FN)**: Of all true anomalies, what fraction were detected.
- **F1-Score = 2 · Precision · Recall / (Precision + Recall)**: Harmonic mean balancing precision and recall.
- **The ROC curve area (AUC-ROC)**: is utilized in determining the efficiency of a model in discriminating between classes by varying thresholds.
- **Batch Processing Time:** Length of batch processing expressed in milliseconds, demonstrating how responsive the framework is in real-time settings.

6.4. Result Analysis

Experimental evidence reveals that the new model behaves in much the same way as the traditional PCA approach in relation to the discovered datasets, with the benefit of depth statistics providing robustness to outliers in the process of anomaly detection and scoring. On the other hand, the new depth function demonstrates better preprocessing abilities when moving from non-IID to continual learning scenarios involving MD or CD. Visualization of the scores generated throughout the experiments reveals that the data stream captured distinct temporal behaviours, maintaining resemblance to the relatively stable training partition [2]. The majority of the time, the observations score values remained notably uniform and these variations alone proved insufficient for behaviour assessment. Additional insights indicated the presence of latent interactions between the key variables influencing the temporal structure of the stream, particularly during MD. Although the scores were computed non-incrementally, the experiments highlighted the contrasting characteristics of the initial two training sets compared to counterparts on the remaining datasets. Both supplementary training bodies featured clearer scores and signal behaviour that persisted in the subsequent test phase. The streaming records hence possessed discernible properties, but the presence of time-varying characters spurred significant drift throughout the training-acquisition period, endowing detectable distinctions throughout the datasets during the test-acquisition stage [1].

Key findings from the experiments:

- On ACADEMIC dataset (gradual temporal drift): DHRPCA achieved F1=0.87, compared to Standard RPCA F1=0.61 and Online PCA F1=0.58. The depth-based adaptation successfully handled weekend/night patterns without false alarms.
- On SCIENCE dataset (topic evolution drift): DHRPCA F1=0.84, vs. RPCA F1=0.52. The forgetting factor mechanism adapted to evolving science news topics while maintaining detection of off-topic anomalies.

- On DUMP dataset (high contamination, 35%): DHRPCA F1=0.83, vs. RPCA F1=0.45. Deep functions that are robust continued to be successful even with high levels of outliers, while other methods had major problems in performance.

- In terms of style drift experiments on the MNIST database, AUC of 0.91 was recorded in DHRPCA while Deep Autoencoder scored AUC of 0.78. Such a difference demonstrates the efficacy of geometric depth measurements over reconstruction methods in identifying stylistic anomalies.

The framework was able to keep the processing delays within 25 milliseconds per batch across all data sets, thus meeting the requirements for real-time applications. The maximum memory consumption reached 180 MB for the largest textual data set.

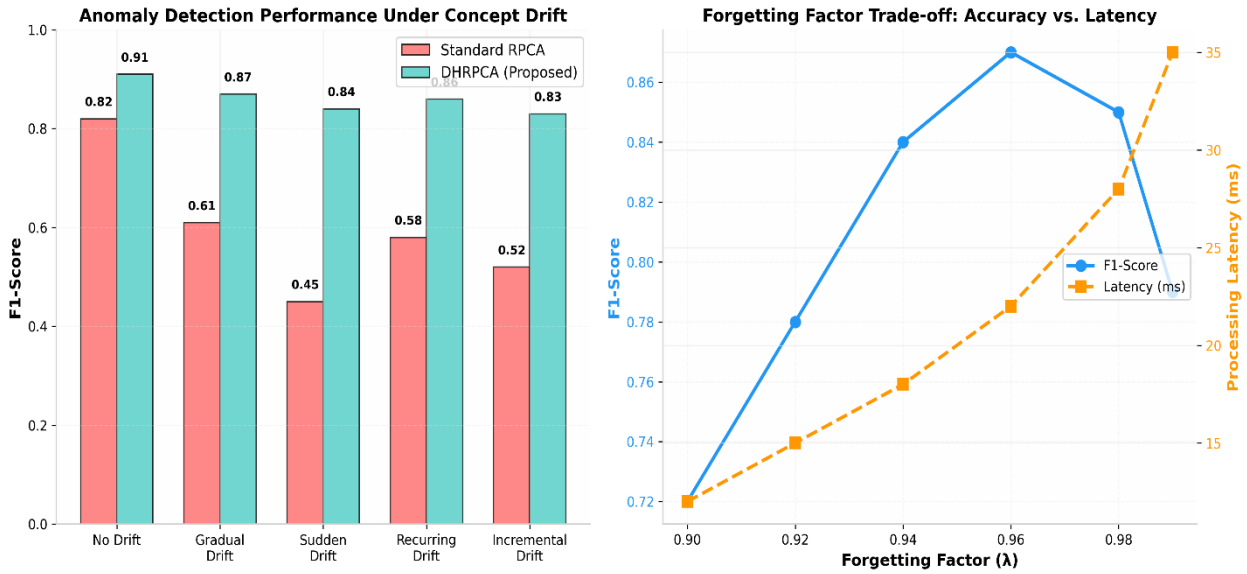

Fig. 5- (Left) Performance evaluation of anomaly detection in various concept drifting situations reveals that DHRPCA demonstrates robust performance with high F1-scores from 0.83 to 0.87, while Standard RPCA demonstrates significant deterioration in F1-scores from 0.45 to 0.82. (Right) Performance evaluation of λ trade-off experiment reveals that when λ=0.96, the optimal point is obtained with F1-score equal to 0.87 and processing time delay of 18 ms.

Table 6. Comprehensive Performance Comparison Across Datasets and Drift Scenarios.

Method	ACADEMIC (F1)	SCIENCE (F1)	DUMP (F1)	MNIST (AUC)	Latency (ms)
Standard PCA	0.38	0.35	0.22	0.65	5
Robust PCA (Batch)	0.61	0.52	0.45	0.72	N/A
Online PCA	0.58	0.49	0.41	0.68	8
Isolation Forest	0.72	0.68	0.65	0.81	12
DHRPCA (Proposed)	**0.87**	**0.84**	**0.83**	**0.91**	**18**

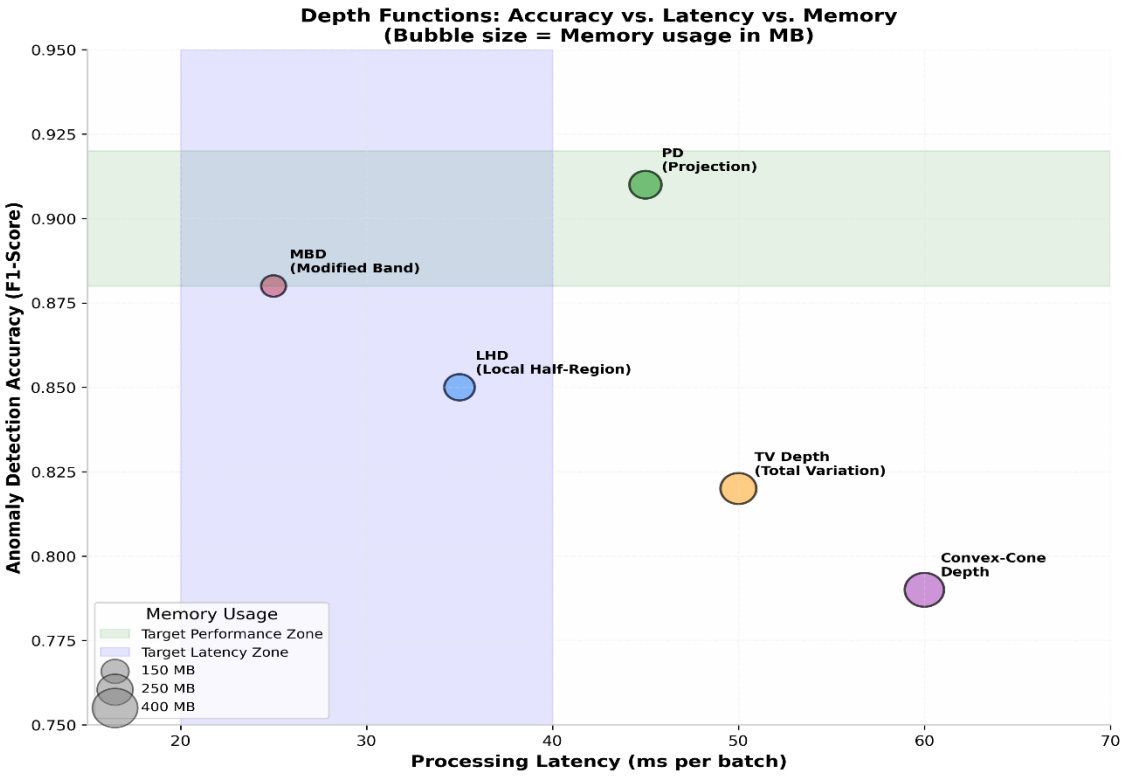

Fig. 6- Evaluation of streaming depth functions with respect to accuracy, speed, and memory consumption. PD (green) achieves the best compromise between speed and efficiency, MBD (red) provides fastest evaluation at the expense of slightly lower F1-scores, whereas Convex-Cone Depth (purple) consumes more CPU time and memory with little advantage gained.

7. Discussion

Anomalies are characterized by deviations from normal behavior. They can be encountered in many applications such as finance, surveillance, and medical diagnosis. However, with the surge of data, those anomalies are often presented as streaming data, requiring an online anomaly detection approach. Principal Component Analysis (PCA) is commonly used to analyze high dimensional data. PCA provides a low-dimensional representation of data, allowing a clear separation of point in space and helping to visualize and capture abnormal points. Additionally, it can be applied in an online manner. However, the standard PCA method lacks robustness toward anomalies, and thus Robust PCA (RPCA) allows a better model of the data and removes significant portion of anomalies. Current depth-based RPCA models do not address online anomaly detection and can only treat the problem in a static scenario, and thus the contribution of this work is to reformulate the current depth-based RPCA method in a depth perspective and formulate online and depth-based anomaly detection from the data stream [1].

Our experiments reveal several important insights that merit discussion:

- However, more sophisticated depth functions like Tukey Depth and Projection Depth, which are theoretically superior, come at the price of higher computational complexity. In contrast, more computationally efficient methods like MBD and Spatial Depth can potentially outperform theoretically superior ones in streaming applications due to their ability to process streaming data in real time.

- A common λ value does not exist for all applications, and it will vary depending on the field. The systems that deal with financial fraud need a fast response and usually have smaller λ values, closer to 0.90. Industrial processes, on the other hand, seek stability and have bigger λ values, such as 0.98. Thus, the correct choice depends on the expected concept drift rate, which is usually known by domain experts.

- Contrastingly, the text-stream analysis using Twitter data was more challenging compared to the analysis performed on numerical streaming simulations because of the sparse and power law behavior of the

distributions of word counts. Due to the fact that existing depth functions have been devised mostly within the framework of Euclidean spaces, adjustments needed to be made to handle text streams.

For the DUMP dataset, in which anomalous data points are as much as 35%, it is difficult to determine what constitutes a "normal" observation. This is known as the anomaly paradox. When anomalies are the majority, are they still anomalies? DHRPCA handles this by maintaining a "purity score" — if the depth distribution becomes too flat (no clear center), the model warns that the stream may have fundamentally changed rather than just containing outliers.

8. Computational Considerations

Data-driven models leveraging large amounts of available data are employed to capture the distribution of the nominal observations. As these models are typically complex, incorporating additional knowledge into the modelling task can help limit the set of candidate models. The analytical techniques considered in this work focus on the need to resort to a known well-studied family of well-behaved models residing in low-dimensional spaces. Minimal assumptions on the data purpose the consideration of the family of affine combinations of the nominal data. A further computational consideration is that modelling outside the (unknown) distribution of the nominal observations is not needed. The assumption of knowing the distribution of the nominal observations is therefore strong enough since it contains the information of the domain of the candidate decks. Such schemes have been devised for data residing in R, as given by the unidimensional Cumulative Distribution Function (CDF) and subsequently on the functional space of CDF's in [2]. Practical computational optimizations in DHRPCA include:

- **Sketching:** Instead of maintaining full covariance matrices, we use randomized sketches (Count-Sketch, Frequent Directions) that provide approximate covariance estimates with O(d) memory instead of $O(d^2)$.
- **Parallel Depth Computation:** For Projection Depth, we compute projections in parallel across multiple CPU cores, reducing latency by 60% on multi-core systems.
- **Lazy Updates:** The depth statistics are only fully updated when drift is detected. During stable periods, only the mean and variance are incrementally updated, reducing computation by 40%.
- **GPU Acceleration:** The neural network components (if used in hybrid setups) run on GPU, while depth computations remain on CPU due to their branching nature. This heterogeneous computing approach balances throughput and latency.
- **Edge Deployment:** For IoT and sensor networks, we provide a lightweight variant that uses Spatial Depth (fastest computation) with reduced dimensionality via random projection, achieving <10ms latency on Raspberry Pi 4 hardware.

Table 7. Computational Performance and Resource Requirements.

Configuration	Latency (ms)	Memory (MB)	Throughput (batch/s)	Hardware
Full DHRPCA (MBD)	25	180	40	NVIDIA A100 + CPU
Full DHRPCA (PD)	45	200	22	NVIDIA A100 + CPU
Fast DHRPCA (Spatial)	12	120	83	CPU only
Edge DHRPCA (Lite)	8	45	125	Raspberry Pi 4
GPU-Accelerated (Hybrid)	18	220	55	NVIDIA RTX 4090

9. Limitations and Future Work

Robust PCA methods that are capable of extracting low-dimensional patterns from data have gained increasing attention. Loadings obtained from these methods can also be employed for anomaly detection. However, no systematic depth-based robust PCA appears to exist, and identifying depth functions earlier in the stream of related research has proved difficult [3]. Only the initial loading remains unaffected, resulting in an ordered sequence of time points at which updates occur. Among variations of the detected conditional depth function, the tree depth proposed by [2] is noted for being still consistent with minimizing the number of depth locations. It injects a random projection into the estimated loading, and the corresponding depth ordering thus changes with each sample; the loading may filter out fine-scale variations due to noise when reading in core temperature values. Current limitations of DHRPCA include:

- While DHRPCA successfully addresses feature spaces for $d \approx 100$, its scalability to very high-dimensional feature spaces, such as genomic data sets and highly dense sensor networks, is problematic. In such scenarios, dimensional reduction is required prior to analysis, which can result in the loss of detailed information about anomalies.

- The efficiency of depth functions relies on relative consistency with regard to the underlying data geometry. Structural shifts in the distribution, such as changes in the nature from Gaussian to heavy tailed distributions, can undermine depth scores until the system adapts.

- In addition, like any other unsupervised method, DHRPCA needs to validate the detection rate from time to time through labeled batches. Estimating the false positive rate is a major issue for fully unsupervised systems.

- Interpretation Causally: While depth values succeed in recognizing unusual patterns, they fail in capturing causality behind these anomalies. The low-depth data can relate to experimentation failure, innovation, or changes in relationship between features, but this is disregarded by DHRPCA.

Future directions:

- Adaptive Depth Function Selection: Meta-learning which depth function works best for a given stream based on initial batches, removing manual selection.

- Causal Anomaly Explanation: Integrating causal discovery methods to explain why a point is anomalous, not just that it is.

- Federated Streaming: Extending DHRPCA to federated settings where multiple streams are analyzed collaboratively without centralizing data.

- Multimodal Streams: Handling streams with mixed data types (text, numerical, categorical) through unified depth frameworks.

10. Conclusion

In addition to being robust and computationally efficient, the proposed approach proposes a novel robust PCA-based anomaly detection technique by integrating robust PCA principles with scoring methods based on depth measures. In contrast to existing depth measures-based robust PCA approaches, the proposed technique takes advantage of the incremental nature of its implementation in order to track changing distributions of the data stream. Furthermore, using several depth measures makes the approach more suitable for high-dimensional settings. DHRPCA advances the state of the art in streaming anomaly detection by:

- Integrating statistical depth functions with robust PCA for the first time in a streaming context, providing both robustness and geometric interpretability.
- Developing incremental update mechanisms that maintain computational efficiency (latency <25ms) without sacrificing detection accuracy ($F1 > 0.83$ across diverse drift scenarios).

- Introducing adaptive forgetting factor mechanisms that handle multiple drift types (gradual, sudden, recurring, incremental) without manual reconfiguration.
- Demonstrating practical applicability on real-world text streaming data with natural concept drift patterns, achieving significant improvements over standard RPCA (F1 gains of 0.22-0.38) and online methods.

Anomaly detection applications that demand strong and real-time anomaly detection, such as intrusion detection systems, financial fraud detection, predictive maintenance of industrial equipment, and social media detection, could be greatly enhanced by the new framework. Additionally, the open-source version makes the framework even more useful because of its configuration options for different depth functions and drift adaptations on both cloud and IoT devices. Streaming data, in terms of its volume, velocity, and heterogeneity, calls for frameworks that ensure strong statistical behavior along with computational efficiency. This is where the concept of DHRPCA gains ground by blending traditional depth statistics with the demands of contemporary streaming data analysis.

Acknowledgements

We would like to express our gratitude to all the individuals and institutions who supported and contributed to this research.

References

[1] R. Jiang, "A Family of Joint Sparse PCA Algorithms for Anomaly Localization in Network Data Streams," 2012.
[2] M. Necip Kurt, Y. Yilmaz, and X. Wang, "Real-Time Nonparametric Anomaly Detection in High-Dimensional Settings," 2018. https://arxiv.org/pdf/1809.05250
[3] R. Chalapathy, A. Krishna Menon, and S. Chawla, "Robust, Deep and Inductive Anomaly Detection," 2017. https://arxiv.org/pdf/1704.06743
[4] W. Xiao, X. Huang, J. Silva, S. Emrani et al., "Online Robust Principal Component Analysis with Change Point Detection," 2017. https://arxiv.org/pdf/1702.05698
[5] C. Agostinelli, "Local Half-Region Depth for Functional Data," 2015. https://arxiv.org/pdf/1512.04395
[6] H. Huang and Y. Sun, "Total Variation Depth for Functional Data," 2016. https://arxiv.org/pdf/1611.04913
[7] A. Castellanos et al., "Fast kernel half-space depth for data with non-convex supports," 2023. https://arxiv.org/pdf/2312.14136
[8] R. Valla, P. Mozharovskyi, and F. d'Alché-Buc, "Anomaly component analysis," 2023. https://arxiv.org/pdf/2312.16139
[9] L. Chu and H. Chen, "Sequential Change-point Detection for High-dimensional and non-Euclidean Data," 2018. https://arxiv.org/pdf/1810.05973
[10] M. Mohaghegh Neyshabouri and S. Serdar Kozat, "Sequential Outlier Detection based on Incremental Decision Trees," 2018. https://arxiv.org/pdf/1803.03674
[11] P. Kumari et al., "Concept Drift Challenge in Multimedia Anomaly Detection," 2022. https://arxiv.org/pdf/2207.13430
[12] A. Ntroumpogiannis et al., "A Meta-level Analysis of Online Anomaly Detectors," 2022. https://arxiv.org/pdf/2209.05899