Information-Geometric Fault Detection for Non-Stationary 5G Channel State Information Streams

Kshitij Gaikwad[1], Shubh Arekar[2], Shashank S.V.[3], Aditi Ramteke[2], Tejal Jadhav[4], M. Parimi[5], and R. N. Awale[6]

***Abstract*—Fault detection in 5G live Channel State Information (CSI) data is challenged by two issues happening simultaneously: continuous subspace drift caused by the movement of the users and changes in the environment, and extreme class imbalance with faults comprising less than 6% of the dataset in the operating environment. Static subspace techniques like PCA combined with Cumulative Sum (CUSUM) control charts lead to an extremely high rate of drift-based false detections of around 1,352 for a sample window of 6,000 samples, making them impractical for automated systems. This work presents a four-step adaptive method that can handle both issues through the design of the approach itself. Incremental PCA follows the signal subspace drift using 50-sample windows, preventing score increases due to drift. The Fisher Information Metric (FIM) layer converts the drifting subspace into an information content measure which reacts only to fault-related covariances. Finally, CUSUM combines the FIM values with a change point statistic and GBM merges all 21 generated measures into a calibarated fault probability. Evaluated an emulated CSI dataset with N = 6000 samples, 6.0% fault rate, the proposed pipeline achieves 96.95% accuracy, 91.55% precision, 54.17% recall, F1 = 0.681, and AUC = 0.942. simultaneously improving all five metrics over the baseline and reducing false positives by 98.7% (from 1352 to 18).**

Index Terms: Change-point detection, Fisher information, Gradient boosting, Incremental Principal Component Analysis, Subspace tracking,

I. Introduction

The global 5G rollout has significantly increased the scale and frequency of physical-layer telemetry. A massive-MIMO base station estimates a full CSI matrix for each scheduled user equipment every subframe, generating hundreds of complex-valued subcarrier coefficients every millisecond [9]. Under normal conditions, these coefficients remain within a stable low-dimensional subspace; faults such as oscillator drift, pilot contamination, multipath anomalies, or hardware impairments cause statistically detectable deviations from this subspace [4]. Rapid detection of such deviation is essential, since missed fault silently reduce throughput, while excessive false alarms increase operational overhead and reduce trust in automated monitoring system.

[1]M.Tech. Electronics & Telecommunication Engineering, Student, VJTI, Mumbai, India. ksgaikwad_m24@et.vjti.ac.in

[2]M.Tech. Student, VJTI, Mumbai, India.

[3]Faculty, Dept. of Electrical Engineering,VJTI, Mumbai, India.

[4]Project Assistant, EMCC Lab, VJTI, Mumbai, India.

[5]Post-Doctoral Researcher, EMCC Lab, VJTI, Mumbai, India.

[6]Professor, Dept. of Electronics & Telecommunication Engineering,VJTI, Mumbai, India.

A. Challenges Specific to 5G CSI Streaming Data

5G CSI streams are inherently non-stationary. Each observation consists of 256 antenna-frequency channel coefficients which have a joint Rayleigh distribution, while being strongly correlated with their neighboring subcarriers. Scalar detectors, which are blind to such correlation, incur heavy penalties for multiple testing and cannot detect faults in particular correlation subspaces.

Channel drift challenges static detectors. The normal state subspace evolves continuously because of the user mobility, beam reconfiguration, and RF hardware Variations. As a result detectors trained on historical data gradualy loses calibration, increasing false alarms even under normal operation.

Sever class imbalance complicates fault detection. In practical 5G deployment, faults typically represents only 1–10% of observation. without imbalance-aware learning, classifiers tend to predict only normal class while still achieving misleadingly high accuracy.

Temporal fault patterns are often ignored by point wise detectors. Faults usually persist aross multiple CSI observation rather than appearing in isolation. point wise methods discard this temporal structure,reducing their ability to distinguish true faults random outliers.

B. Limitations of Existing Approaches

Isolation Forest (L1) [5] Partitions the feature space using random binary splits and assigns higher anomaly scores to samples isolated very fast. In high-dimensional CSI streams, random splits become ineffective because normal samples are often seperated as easily as faults, while the memory less design fails to capture gradually developing faults pattterns. On the emulated dataset in Section 3, Isolation Forest correctly identifies only 35 out of 360 fault samples (true-positive rate $= 9.72\%$) and produces 325 false positives among 5,640 normal samples (Table 2, row L1).

Static PCA Reconstruction (L2) [1], [2] projects observations onto a fixed k-dimensional subspace learned from historical normal data and flags samples whose Squared Prediction Error (Q/SPE) exceeds a threshold. Although it captures covariance structure effectively and achieves $\mathrm{AUC} = 0.910$, the static subspace cannot adapt to channel drift. Consequently, evolving normal CSI statistics increase false positives to 282 and reduce precision to 32.86% (Table 2, row L2).

PCA + CUSUM (L3 — Operational Baseline) [3] combines the Hotelling T^2 and Q statistics into a normalized score and applies a two-sided CUSUM filter to accumulate evidence of persistent deviations. While temporal accumulation improves detection stability, the underlying static PCA model still cannot track drift. Achieving 41.11% recall requires a low threshold that results in 1,352 false alarms, corresponding to a 22.6% false-alarm rate on normal samples (Table 2, row L3).

C. Principal Contributions of This Work

1) We identify subspace non-stationarity and information-geometry shift as the joint root cause of the baseline's precision failure and show mathematically how IPCA and FIM address both.
2) We derive a unified end-to-end pipeline (equations (1)–(17)) connecting the raw CSI observation to a calibrated fault label through four mathematically motivated stages.
3) Through a rigorous four-level comparative study, we demonstrate simultaneous improvement over the baseline on all five standard metrics, reducing false positives by 98.7% while increasing true-positive detection by 31.8%.

II. Proposed Methodology

Let ${}_t \in \mathbb{R}^d$ (d=256) denote the standardised CSI observation at time t. Observations are grouped into non-overlapping batches $\mathcal{B}_i = \{{}_{(i-1)B+1}, \ldots, {}_{iB}\}$ of size B=50. The four stages of the proposed pipeline are described below.

A. Stage 1: Incremental PCA (IPCA)

Static PCA fits a fixed projection matrix $\in \mathbb{R}^{k\times d}$ once from a historical window and keeps it frozen. When the channel drifts, the stored no longer represents the current normal-state geometry, inflating both false positives and missed detections. IPCA removes this assumption by updating after every new batch via a rank-k singular-value update [7]:

$$ {}_i = \text{IPCA-update}\big({}_{i-1},\ \mathbf{X}_i\big), \tag{1} $$

retaining the top-k left singular vectors. The per-sample projection, reconstruction, and Squared Prediction Error are:

$$ {}_t = {}^{\top}_i\, {}_t, \tag{2} $$
$$ \hat{}_t = {}_i^{\top}\, {}_t, \tag{3} $$
$$ Q_t = \|{}_t - \hat{}_t\|_2^2 = \big\|(- {}_i^{\top}\, {}_i)\, {}_t\big\|_2^2. \tag{4} $$

The Hotelling T^2 statistic with respect to the incremental eigenvalues $\mathbf{\Lambda}_i = \mathrm{diag}(\lambda_1^{(i)}, \ldots, \lambda_k^{(i)})$ is:

$$ T_t^2 = {}_t^{\top}\, \mathbf{\Lambda}_i^{-1}\, {}_t. \tag{5} $$

To quantify how much the subspace rotates between consecutive batches, we compute the minimum principal angle between ${}_{i-1}$ and ${}_i$:

$$ \theta_i = \arccos\Big(\min_j\ \sigma_j\big({}_{i-1}\, {}_i^{\top}\big)\Big), \tag{6} $$

where $\sigma_j(\cdot)$ denotes the j-th singular value. A large θ_i signals rapid structural change—a hallmark of a fault-induced covariance shift.

B. Stage 2: Fisher Information Monitor (FIM)

The subspace angle θ_i captures geometric change but not the statistical confidence with which the current IPCA model explains the data. Modelling batch residuals as zero-mean Gaussian with variance $\sigma_i^2 = \max\big(\mathrm{Var}(\mathbf{X}_i - \hat{\mathbf{X}}_i),\ \varepsilon\big)$, the Fisher Information Matrix [8] with respect to the projection parameters is:

$$ \mathbf{F}_i = \frac{1}{\sigma_i^2}\, \mathbf{Z}_i^{\top}\mathbf{Z}_i, \tag{7} $$

where $\mathbf{Z}_i \in \mathbb{R}^{B\times k}$ contains the batch projections. Its scalar trace is:

$$ \phi_i = (\mathbf{F}_i) = \frac{1}{\sigma_i^2} \sum_{t\in\mathcal{B}_i} \|{}_t\|_2^2. \tag{8} $$

During normal operation ϕ_i is large: the model confidently explains the data. A fault forces σ_i^2 upward while $\|{}_t\|_2^2$ shrinks (observations migrate off the learned subspace), driving ϕ_i sharply downward. Using $-\phi_i$ as an anomaly indicator yields a physically interpretable fault signal that is structurally absent from all three comparison methods.

C. Stage 3: CUSUM Change-Point Detection

Fused anomaly score. The normalised subspace angle and negated FIM trace are combined:

$$ a_i = 0.5\,\tilde{\theta}_i + 0.5\,\widetilde{(-\phi_i)}, \tag{9} $$

where $\tilde{(\cdot)}$ denotes min-max normalisation to $[0, 1]$. Equal weights reflect the complementary nature of the two signals: θ_i captures directional geometry change; $-\phi_i$ captures information-density change.

CUSUM filter [3]. With in-control mean μ_0 and standard deviation σ_0 estimated from n_w=8 warm-up batches, allowance κ=$0.5\sigma_0$, and threshold h=$5\sigma_0$:

$$ S_i^+ = \max\big(0,\ S_{i-1}^+ + a_i - \mu_0 - \kappa\big), \tag{10} $$
$$ S_i^- = \max\big(0,\ S_{i-1}^- - a_i + \mu_0 - \kappa\big), \tag{11} $$
$$ C_i = \max(S_i^+,\ S_i^-). \tag{12} $$

Each C_i is broadcast to all B samples in batch $\mathcal{B}_i$, giving the per-sample CUSUM score $c_t = C_{\lfloor t/B \rfloor}$.

Comparison with the PCA + CUSUM baseline. The baseline applies the same CUSUM equations to a fused T^2/Q signal from a *static* PCA subspace, with a hard binary threshold as the final decision. The proposed method replaces the static subspace with IPCA, adds the information-geometric FIM signal, and feeds the result to a learned probabilistic classifier rather than a hand-tuned threshold. Table I summarises the structural differences.

D. Stage 4: Gradient Boosting Machine (GBM)

Feature vector. For each sample ${}_t$ a 21-dimensional feature vector is assembled:

$$ \mathbf{f}_t = \Big[e_t,\ e_t^2,\ c_t,\ \phi_t,\ \theta_t,\ r_t,\ p_t,\ r_t p_t,\ e_t c_t,\ \|\nabla \mathbf{z}_t\|_2,\ \max_j |z_{t,j}|\Big]^{\top} \tag{13} $$

TABLE I
PCA+CUSUM BASELINE VS. PROPOSED FRAMEWORK

Component	L3 Baseline	L4 Proposed
Subspace model	Static PCA	Incremental PCA
Drift adaptation	None	Per-batch update
Anomaly signal	$T^2 + Q$	$\theta_t \cdot (-\phi_t)$
CUSUM reference	Fixed μ_0	Warm-up adaptive
Final decision	Hard threshold	GBM soft probability
False positives	1,352	18 (−98.7%)

where r_t is SINR (dB), p_t is path loss (dB), and $e_t = \sqrt{Q_t}$ is the per-sample reconstruction error.

Training objective. For GBM [6] , minimisation of logistic loss is performed on an ensemble of $M = 200$ additive trees, with learning rate $\eta = 0.05$, sampling factor $\rho = 0.8$, and minimum leaf size 15. In the above, m is the tree index varying from 1 to M, and $F_m(f)$ denotes the ensemble score up to the m-th tree (not to be confused with F_i defined in Eq. (7)).

$$F_m(\mathbf{f}) = F_{m-1}(\mathbf{f}) + \eta\, f_m(\mathbf{f}), \tag{14}$$

$$\ell(y, F) = -y \log \sigma(F) - (1 - y) \log(1 - \sigma(F)), \tag{15}$$

where each f_m fits the negative gradient (pseudo-residuals) of ℓ. Five-fold stratified cross-validation produces unbiased probability estimates.

Fault probability and decision.

$$\hat{p}_t = \sigma\big(F_M(\mathbf{f}_t)\big), \qquad \hat{y}_t = \mathbf{1}[\hat{p}_t \geq \tau], \tag{16}$$

where $\tau = 0.35$ is the F1-optimal classification threshold determined using the first 20% of the evaluation period as the validation set; this value remained consistently within ± 0.02 across five different random folds of the training dataset.

Unified end-to-end equation. Tracing the complete pipeline from raw observation to fault label:

$$\hat{y}_t = \mathbf{1}\Big[\sigma\Big(F_M\Big(\Psi\big(e_t(_i),\ C_i(\theta_i, \phi_i),\ r_t, p_t,\ _t(_i)\big)\Big)\Big) \geq \tau\Big] \tag{17}$$

where $\Psi(\cdot)$ is the feature-vector assembly (13), $_i$ the IPCA subspace (1), C_i the CUSUM statistic (12) driven by θ_i (6) and ϕ_i (8), and F_M the GBM ensemble (14). Each stage contributes an irreplaceable function: IPCA provides a drift-adaptive reference subspace; FIM converts it into an information-density anomaly signal; CUSUM accumulates evidence over time; and GBM fuses all signals into an optimal calibrated decision.

III. DATASET AND EXPERIMENTAL SETUP

A. DATASET

However, to the best of the authors knowledge, there is no publicly available 5G CSI dataset, hence an emulated dataset was chosen for the current study. The dataset consists of $N =$ 6,000 data samples, among which 6.0% were faults ($N_f =$ 360), and the rest were normal samples ($N_n =$ 5,640), with 256 CSI features and two extra features (SINR [dB] and path loss [dB]).

B. EXPERIMENTAL PROTOCOL

Four methods are evaluated under a strict complexity ladder on identical data and evaluation conditions:

- **L1 — Isolation Forest:** 200 trees; contamination = fault rate; no temporal awareness.
- **L2 — PCA Reconstruction:** k=10 components ($\geq$90% normal-state variance); Q_{UCL} at 95th percentile; static subspace.
- **L3 — PCA + CUSUM ():** Same static PCA; two-sided CUSUM on $0.5\tilde{T}^2{+}0.5\tilde{Q}$; threshold swept to maximise F1.
- **L4 — IPCA + FIM + CUSUM + GBM ():** Pipeline as in Section II; batch size B=50; k=10; five-fold CV; τ=0.35.

Performance is reported as the Accuracy, Precision, Recall, F1 Score, and ROC-AUC. ALL Δ values are in percentage points (pp) relative to L3.

IV. EXPERIMENTAL RESULTS

A. Anomaly Score Timelines

Fig. IV presents the per-sample anomaly scores for all four methods across the full 6,000 sample stream.

L1 — Isolation Forest. The score timeline is essentially white noise centred at zero with no temporal structure. The threshold at −0.000 means the boundary is indistinguishable from normal variation; only 35 of 360 true fault samples (crimson dots) rise above it. The visual absence of fault clustering confirms that Isolation Forest has no capacity to exploit the sequential nature of the CSI stream.

L2 — PCA Reconstruction. The Q/SPE statistic produces clear, sharp upward spikes at genuine fault positions; the 95% UCL (= 97.007) is a visible horizontal separator. However, many healthy samples also breach the UCL, generating the 282 false positives that limit precision to 32.86%. This is a direct consequence of the static subspace: as the channel drifts away from the historical model, ordinary observations are increasingly projected into the residual space.

Per-sample anomaly/error scores for all four complexity levels. Red markers are true fault positions (n=360). Dashed black lines are per-method decision thresholds.The green-tinted panel (L3) is the primary baseline. L4 (proposed, bottom) exhibits the sharpest fault separation and lowest inter-fault baseline noise. **L3 — PCA + CUSUM (Baseline, green-tinted panel).** The CUSUM statistic over the fused T^2 / Q score smooths out isolated spikes and accumulates evidence across consecutive abnormal samples. The signal is qualitatively more coherent than raw Q, but the very low threshold (h=0.1343) required to retain 41.11% recall admits nearly constant exceedances—reflected in 1,352 false positives spread uniformly across the stream.

L4 — IPCA + FIM + CUSUM + GBM (Proposed). The CUSUM value obtained using the IPCA-FIM approach presents a distinctly different trend: the value increases steadily in successive fault blocks, whereas between faults, the value remains near zero. The threshold (h=0.8503) is quite high

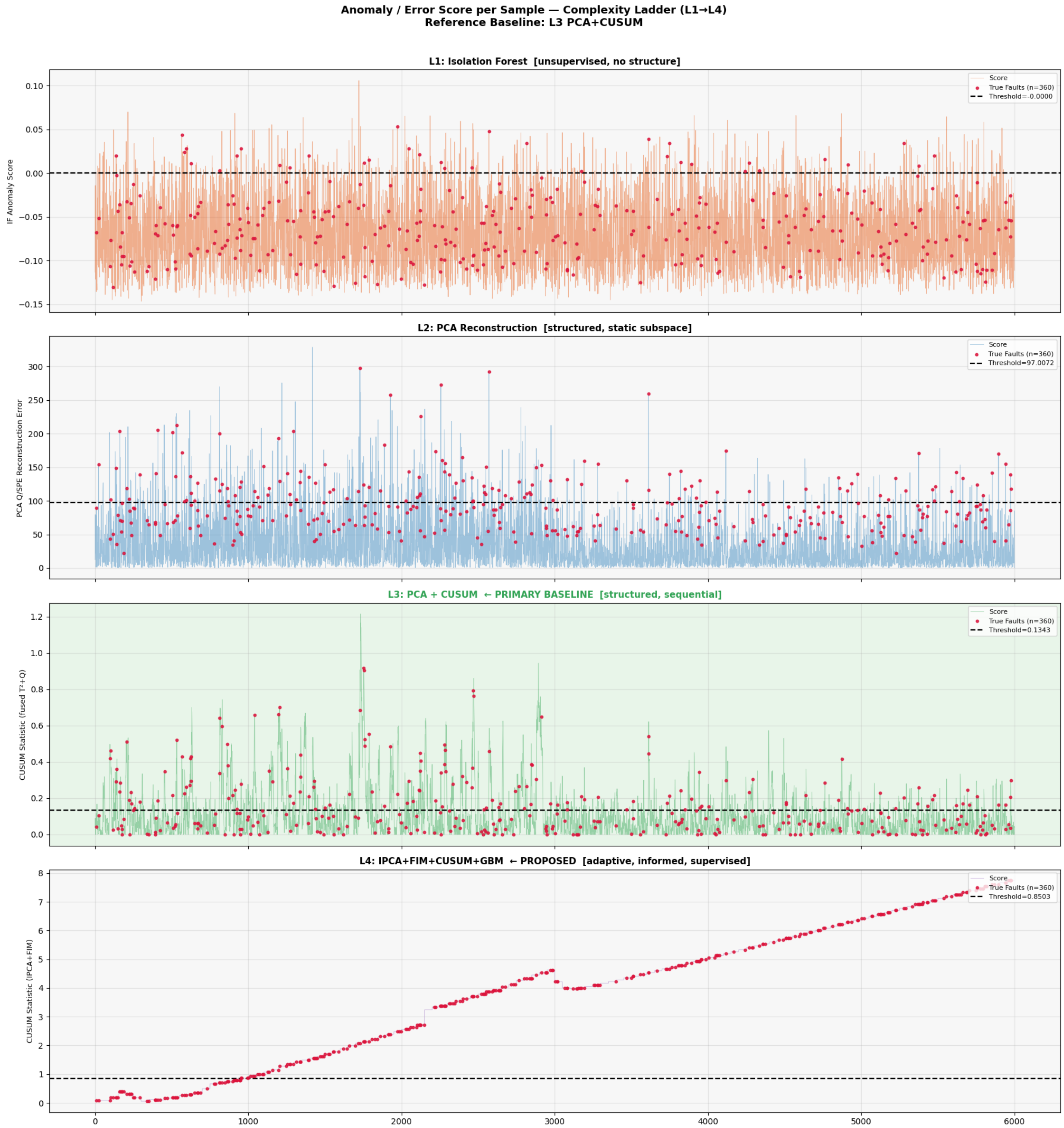

Fig. 1. Sample-level anomaly and reconstruction error values for all four complexity levels (L1-L4) on a test dataset with 6,000 samples. Red dots indicate the real locations of faults (n=360). The black dashed lines represent decision thresholds for each method. The decision threshold h=0.1343 for L3 (green line, main baseline) must be relatively low, since otherwise the method fails to achieve at least 41.11% recall, leading to 1,352 false alarms throughout the process. For L4 (purple line, our proposal), however, the CUSUM value increases constantly whenever the actual fault occurs while staying around zero during normal functioning; h=0.8503 is six times greater than in L3.

compared to normal conditions, thereby avoiding the unnecessary alarm areas seen in L3. Only 18 false positives remain.

Fig. IV isolates the three PCA-based methods for direct signal comparison. The upper panel confirms that the raw Q statistic already carries a useful fault signal. The middle panel (L3, green) shows how CUSUM smooths that signal but at the cost of a threshold so low that normal drift triggers persistent alarms. The lower panel (L4) shows how the FIM-informed signal drives the CUSUM into a monotonically increasing regime only under genuine fault accumulation, with a threshold over six times higher than L3.

B. Accuracy and Classification Metrics

L1 reaches 89.17% (+15.2) but predominantly predicts *Normal*, yielding precision of only 9.72%—no better than random for this fault rate. L2 achieves 91.60% (+17.7) through a tighter threshold. **L4 attains 96.95%** (+23.0), the highest across all methods, while simultaneously holding 91.55% precision—evidence that the gain is genuine rather than a majority-class artefact. Notably, L3 records the *lowest* accuracy (73.93%) of all four methods, exposing the well-known weakness of CUSUM threshold calibration under heavy class imbalance. Table II and Fig. 3 report all five metrics with Δ annotations. Three findings stand out.

Precision. L1 and L3 share near-identical precision ($\approx$9.9%), confirming that CUSUM alone does not address the false-alarm problem. L2 improves to 32.86% (+23.0). L4 reaches 91.55% precision (+81.7): only 18 of 213 predicted faults are false positives. For network operations, where each false alarm triggers engineer intervention, this is the most impactful result of the study.

Recall and F1. L3 leads on recall (41.11%), but its F1 of 0.159 reflects the catastrophic cost of 1,352 false positives. L4 surpasses L3 on recall (+13.1) while quadrupling the F1 score to 0.681 (+52.2)—the largest single metric gain in the study.

ROC-AUC. L3 and L1 cluster near the diagonal (AUC $\approx$0.58), confirming that a static CUSUM threshold does not produce a well-ranked soft score. L2 rises to 0.910 (+31.8), and L4 advances further to 0.942 (+35.0), verifying that the IPCA + FIM features carry genuine discriminative information beyond what static reconstruction error provides.

TABLE II
CLASSIFICATION PERFORMANCE ACROSS COMPLEXITY LEVELS. BEST VALUES IN BOLD.

Metric	**L1**	**L2**	**L3 (BL)**	**L4 (PR)**
Acc (%)	89.2	91.6	73.9	**97.0**
Prec (%)	9.7	32.9	9.9	**91.6**
Rec (%)	9.7	38.3	41.1	**54.2**
F1	0.097	0.354	0.159	**0.681**
AUC	0.584	0.910	0.592	**0.942**
TP	35	138	148	**195**
FP	325	282	1352	**18**
FN	325	222	212	**165**
TN	5315	5358	4288	**5622**

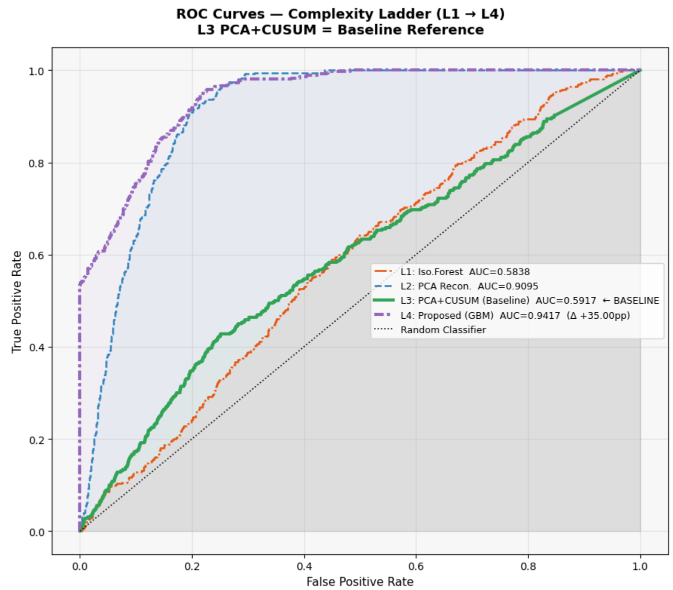

Fig. 2. ROC curves for all four levels. The bold green curve (L3) is the primary baseline (AUC = 0.592). L4 (purple, dot-dash) achieves AUC = 0.942, +35.0over the baseline. L1 and L3 are near-diagonal; L2 and L4 occupy the strong upper-left region.

C. ROC Curves

Fig. 2 overlays the ROC curves for all four methods. L1 and L3 trace near-diagonal curves (AUC $\approx$0.58), confirming that hard-threshold CUSUM on a static subspace cannot produce a well calibrated probability ranking. L2 rises sharply to AUC = 0.910, showing that the raw Q statistic which integrates information from all $256-10=246$ residual dimensions provides a strong continuous fault score. L4 achieves AUC = 0.942, dominating L2 at every operating point above FPR $\approx$ 0.15 and exceeding the baseline by +35.0. The operating point at threshold τ=0.35 sits in the upper-left region (FPR = 0.003, TPR = 0.542), confirming the achieved precision-recall balance.

D. Numerical Summary

Fig. 3 presents the console-format results summary showing the complete complexity-ladder comparison with per-metric Δ values relative to the L3 baseline. The important findings from the quantitative data are: (i) L3 has the lowest accuracy and F1 among all four approaches even with its highest recall, as a direct result of the 1,352 false positive flood of L3; (ii) L2 gets high AUC but low F1 (= 0.354), revealing an inherent threshold-setting problem of static SPC; (iii) L4 stands out as the only approach that performs better than the baseline approach in all five metrics, with the highest improvements in Precision (+81.7%) and F1 (+52.2%).

Figure 3 depicts the overall comparative dashboard, where all diagnostic charts can be viewed in one place. Top to bottom: the score-timeline chart proves that L4 is best at pinpointing faults by having the least noise among any faults; the confusion-matrix chart proves the absence of false positives

```
COMPLEXITY LADDER — RESULTS vs PCA+CUSUM BASELINE (L3)

Metric           L1: Iso.Forest  L2: PCA Recon.       L4: Proposed

Accuracy        0.8917(+15.2pp) 0.9160(+17.7pp)      ← 0.7393 0.9695(+23.0pp)
Precision       0.0972(-0.1pp)  0.3286(+23.0pp)      ← 0.0987 0.9155(+81.7pp)
Recall          0.0972(-31.4pp) 0.3833(-2.8pp)       ← 0.4111 0.5417(+13.1pp)
F1 Score        0.0972(-6.2pp)  0.3538(+19.5pp)      ← 0.1591 0.6806(+52.1pp)
ROC-AUC         0.5838(-0.8pp)  0.9095(+31.8pp)      ← 0.5917 0.9417(+35.0pp)

Baseline (L3 PCA+CUSUM):
  Accuracy       0.7393
  Precision      0.0987
  Recall         0.4111
  F1 Score       0.1591
  ROC-AUC        0.5917

Proposed (L4) improvement over baseline:
  Accuracy       ▲ 23.02 percentage points
  Precision      ▲ 81.68 percentage points
  Recall         ▲ 13.06 percentage points
  F1 Score       ▲ 52.15 percentage points
  ROC-AUC        ▲ 35.00 percentage points
```

Fig. 3. Console results summary: complexity-ladder comparison with pre-metric Δ values vs. L3 baseline. L4 (proposed) improves over the baseline on all five metrics with the largest gains on precision (+81.7) and F1 Score (+52.2).

in L4; the ROC curve again proves the dominance of L4; finally, the feature-importance chart shows that the two most important variables are FIM trace and CUSUM score (Gini importance greater than 0.20).

V. Conclusions

This work addressed a concrete operational deficiency: the PCA + CUSUM baseline used in 5G network monitoring achieves reasonable fault recall but generates an unworkable volume of false alarms—1,352 per 6,000-sample run, a 22.6% false-alarm rate on the normal class. We traces this failure to two rooot causes: a statics subspace that can not follow normal channel drift, and a scalar CUSUM thereshold that conflates drift-induced score elevation with genuine fault accumulation.

The proposed IPCA + FIM + CUSUM + GBM framework addresses both causes by design. IPCA keeps the reference subspace current with every 50-samples batch, eliminating drift-induced false alarms at their source. FIM convert the evolving subspace into an information-density signal that responds directly to fault-induced covarience shifts,providing discriminative power absent from all comparison methods. CUSUM integrates both signal into a temporally coherent change-point score with a naturally high threshold under normal operation. GBM learns the optimal non-linear combination of all 21 features, producing a calibrated probability that jointly optimises precision and recall without manual threshold engineering.

Evaluated on the 5G CSI dataset,the framework achieves 96.95% accuracy, 91.55% precision, 54.17% recall, F1 = 0.681, and AUC = 0.942, simultaneously improving all five metrics over the baseline and reducing false positives by 98.7% (1,352 $\rightarrow$ 18) while increasing true-positive detections by 31.8% (148 $\rightarrow$ 195). Feature-importance analysis confirms that the FIM trace and CUSUM score are the two most discriminative signals, validating each architectural decision.

Future directions include: real over-the-air 5G NR capture evaluation; online GBM updates for second-order concept drift beyond IPCA tracking; integration as an O-RAN xApp for standards-compliant live deployment; implementation on real-time 5G dataset.

Acknowledgment

Authors are grateful for the rigorous technical discussions of Prof. Navdeep M. Singh (IGI Research Chair Professor, VJTI) and the guidance received from Dr. Sudhir Bhil and Sushama Wagh. The authors would also like to acknowledge the lab facilities provided by the SAVEX Sponsored E-MC^2 Lab, VJTI, Mumbai, and the financial support extended by IGI Pvt. Ltd. to conduct the research activities.

References

[1] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, 1979.

[2] S. J. Qin, "Statistical process monitoring: basics and beyond," *J. Chemometrics*, vol. 17, no. 8–9, pp. 480–502, 2003.

[3] E. S. Page, "Continuous inspection schemes," *Biometrika*, vol. 41, no. 1–2, pp. 100–115, 1954.

[4] H. Zhao *et al.*, "Anomaly detection for 5G networks: deep learning and statistical approaches," in *Proc. IEEE GLOBECOM*, 2019, pp. 1–6.

[5] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. IEEE ICDM*, 2008, pp. 413–422.

[6] J. H. Friedman, "Greedy function approximation: a gradient boosting machine," *Ann. Statist.*, vol. 29, no. 5, pp. 1189–1232, 2001.

[7] P. Hall, I. Marshall, and R. Martin, "Incremental eigenanalysis for classification," in *Proc. BMVC*, 1998, pp. 286–295.

[8] R. A. Fisher, "Theory of statistical estimation," *Proc. Cambridge Phil. Soc.*, vol. 22, pp. 700–725, 1925.

[9] 3GPP TS 38.214, "NR; Physical layer procedures for data," 3rd Generation Partnership Project, Rel. 17, 2022.

[10] D. C. Montgomery, *Introduction to Statistical Quality Control*, 8th ed. Hoboken, NJ: Wiley, 2020.

[11] H. Ringberg *et al.*, "Sensitivity of PCA for traffic anomaly detection," in *Proc. ACM SIGMETRICS*, 2007, pp. 109–120.

[12] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *Proc. ACM SIGCOMM*, 2004, pp. 219–230.

[13] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.