A Comparative Study of Differential Privacy Mechanisms with Blockchain Governance in Medical Data

Meenal Patle
Department of DSAI
IIIT Naya Raipur
meenal23102@iiitnr.edu.in

Shubhangi Chouhan
Department of DSAI
IIIT Naya Raipur
shubhangi23102@iiitnr.edu.in

Vaishnavi Shrivastava
Department of DSAI
IIIT Naya Raipur
vaishnavi23102@iiitnr.edu.in

Dr.Satyanarayana Vollala
Department of CSE
IIIT Naya Raipur
satya@iiitnr.edu.in

Abhijeet Kar
Department of DSAI
IIIT Naya Raipur
abhijeetk@iiitnr.edu.in

Kaman Sahu
Department of CSE
IIIT Naya Raipur
kamna@iiitnr.edu.in

Sagnik Majumdar
Department of CSE
IIIT Naya Raipur
sagnik22100@iiitnr.edu.in

***Abstract*— Medical databases contain extremely sensitive information regarding a patient, and this information must be kept protected during statistical analysis. Differential Privacy (DP) provides a mathematical framework that can help keep data private by adding calibrated noise to the output of queries. Unfortunately, DP in its traditional form lacks accountability, multi-user governance, and does not protect against multiple queries to the same patient. In this paper, we propose a hybrid security model in which Differential Privacy (Laplace, Gaussian, Exponential, or Flipped-Huber) mechanisms are used in conjunction with a Blockchain-based audit ledger for access to medical data. The blockchain knows the DP privacy-budget constraints, keeps an immutable log of all queries, and prevents privacy leakage across multiple researchers. The proposed security system provides a way of safely and securely conducting random statistical queries on medical records while providing confidence intervals and utility comparisons by mechanism. The findings will provide practical evidence that an auditable, privacy-preserving method for healthcare analytics is possible.**
Keywords—Differential Privacy, Blockchain, Medical Data Security, Privacy Budget, Healthcare Analytics, Data Auditing

I INTRODUCTION

l'.3 The growing digitization of healthcare systems has caused the widespread collection and analysis of sensitive patient data for decision support in clinical practice, epidemiology, and medical research. Data-driven approaches to healthcare can improve outcomes, but they create considerable concerns for privacy. Unauthorized breach of medical records can not only harm reputations, but also lead to discrimination and violate legal regulations such as HIPAA and GDPR. Therefore, it has become a major and relevant research challenge to permit secure statistical analysis of medical data, without violating the privacy of the individual.

Differential Privacy (DP) has emerged as a mathematically rigorous framework to mitigate privacy risks by injecting calibrated random noise into query results [**?**]. Classical mechanisms such as the Laplace [**?**], Gaussian [**?**], and Exponential mechanisms, as well as recent innovations like the generalized Gaussian and flipped Huber distributions [**?**], provide varying trade-offs between privacy and utility. Despite their theoretical guaranties, traditional DP systems face practical limitations in real-world medical environments. Specifically, they lack accountability for repeated queries, mechanisms for multi-user governance, and safeguards against cumulative privacy leakage. Blockchain technology, with its decentralized, immutable, and auditable ledger, has recently been explored for secure medical data sharing[**?**] . By enforcing tamper-proof logging and fine-grained access control, blockchain can complement DP by introducing accountability and governance. Prior works have demonstrated blockchain's potential in cross-institutional medical collaboration ; however, integration with DP mechanisms to simultaneously ensure privacy preservation and auditability remains underexplored.

Inspired by these gaps, in this paper, we extend a novel hybrid security framework utilizing Differential Privacy-based mechanisms (Laplace, Gaussian, Exponential, and Flipped-Huber) along with a blockchain-based audit ledger. Our system uses blockchain to enforce privacy-budget limits, provide an immutable record of each query, and mitigate privacy leakage across multiple researchers. Our framework achieves secure, noise-protected statistical querying of medical records while also accommodating confidence interval estimation and comparisons of utility across mechanisms. The contributions of this work are summarized as follows:

- We propose a privacy-preserving DP approach that is governed by a blockchain-based framework, which enforces the accountability of queries while also guaranteeing privacy budgets in healthcare data analyses.

- We combine multiple DP mechanisms, including hybrid noise distributions recently proposed, to allow for flexible trade-offs in privacy and utility.
- We show through theoretical analysis and experimental evaluation that our framework supports secure, auditable, and privacy-preserving healthcare data analytics.

II RELATED WORK

Differential Privacy (DP) and blockchain technology have been explored independently in several domains, ranging from secure distributed computing to privacy-preserving statistical analysis. McSherry's PINQ framework introduced one of the earliest practical systems for enforcing differential privacy through controlled query transformations and noisy aggregations. Airavat extended this line of work to distributed environments by integrating DP with mandatory access control for MapReduce-based computations. While both systems enforce privacy at query time, neither provides decentralized accountability or immutable logging.

GUPT further simplified DP for non-expert users by mapping privacy budgets into accuracy guarantees and automating privacy budget allocation. However, the system does not prevent multiple analysts from exhausting the privacy budget through repeated querying. PSI (Private Data Sharing Interface) focused on enabling social and health science researchers to interact with sensitive datasets using differential privacy, but relied entirely on central governance and lacked tamper-proof audit trails.

Beyond system-level frameworks, mechanism-level advancements such as the Generalized Gaussian mechanism and the recent Flipped-Huber distribution have expanded the set of noise distributions available to DP systems. These innovations provide alternatives to classical Laplace and Gaussian mechanisms, particularly for applications requiring robustness under heavy-tailed data or skewed distributions.

Blockchain has independently been explored for medical data sharing, enabling secure access logs and decentralized authorization. Such systems, however, primarily focus on encryption and access control rather than protecting released statistics. Existing blockchain-based healthcare systems do not guarantee differential privacy and therefore remain vulnerable to inference attacks when releasing aggregated information.

Unlike prior approaches, our work unifies these two domains by integrating multiple DP mechanisms with a blockchain-governed auditing layer. This combination ensures not only mathematically rigorous privacy guarantees but also immutable per-query accountability across multiple analysts.

III Method used for Secure Medical Queying

III-A Differential Privacy Mechanisms

Differential Privacy guarantees that the output of a query will not change significantly when deleting or adding a single individual to the database [**?**].

Our system implements four noise mechanisms:

- Laplace Mechanism — classical definition of (ϵ-DP using Laplace noise with scale = sensitivity / (ϵ).
- Gaussian Mechanism — (ϵ, δ)-DP for datasets that require "smoother" noise distributions.
- Exponential Mechanism — used for categorical outputs e.g. "most common symptom".
- Flipped-Huber Mechanism — hybrid of Laplace (center) and Gaussian (tails) that presents higher utility than the other mechanisms for medical averages.

  Each mechanism will return both the private result as well as a confidence interval of 95% to provide medical practitioners context around uncertainty without compromising privacy.

III-B Privacy Budget Calculation and Sensitivity Estimation

Each query type (count, average, categorical score) has a fixed global sensitivity. Examples:

- **Count queries** $\rightarrow$ sensitivity = 1
- **Average age** $\rightarrow$ sensitivity = $\frac{\text{max\_age} - \text{min\_age}}{\text{number\_of\_patients}}$

Before answering any query, the system computes:

$$\text{noise\_scale} = \frac{\text{sensitivity}}{\epsilon}$$

This ensures that the added noise is correctly calibrated in the mathematical sense.

1) Mathematical Formulation

Differential Privacy ensures that the output of a query does not change significantly when an individual's record is added to or removed from the dataset. A mechanism $\mathcal{M}$ satisfies ϵ-Differential Privacy if:

$$\Pr[\mathcal{M}(D_1) = y] \leq e^{\epsilon} \Pr[\mathcal{M}(D_2) = y] \tag{1}$$

for all neighboring datasets D_1 and D_2 that differ in only one record.

a) Global Sensitivity

The global sensitivity of a function f is given by:

$$\Delta f = \max_{D_1, D_2} \|f(D_1) - f(D_2)\| \tag{2}$$

b) Laplace Mechanism

Laplace noise is added proportional to sensitivity:

$$\mathcal{M}(x) = f(x) + \text{Laplace}\left(0, \frac{\Delta f}{\epsilon}\right) \tag{3}$$

c) Gaussian Mechanism

For (ϵ, δ)-Differential Privacy, Gaussian noise is used:

$$\mathcal{M}(x) = f(x) + \mathcal{N}(0, \sigma^2) \tag{4}$$

where

$$\sigma \geq \frac{\sqrt{2\ln(1.25/\delta)} \cdot \Delta f}{\epsilon} \tag{5}$$

d) Exponential Mechanism

The exponential mechanism selects an output r with probability:

$$\Pr[\mathcal{M}(x) = r] \propto \exp\left(\frac{\epsilon \cdot u(x, r)}{2\Delta u}\right) \tag{6}$$

e) Flipped-Huber Mechanism [?]

This mechanism combines Laplace (for small noise) and Gaussian (for larger deviations):

$$p(z) = \begin{cases} \frac{1}{2b} \exp\left(-\frac{|z|}{b}\right), & |z| \leq k \\ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(z-k)^2}{2\sigma^2}\right), & |z| > k \end{cases} \tag{7}$$

This hybrid distribution offers improved robustness and higher accuracy for medical statistical queries [**?**]compared to standard noise mechanisms.

Additionally, the engine tracks cumulative ϵ for each researcher so they cannot exceed the allowed privacy budget.

III-C Blockchain Audit Layer for Query Governance

Utilizing blockchain technology, better transparency and safe and shared medical data can be achieved in healthcare systems [**?**].

The smart-contract implementation of blockchain handles each query with the following steps:

1) Verification of the Researcher's Identity
2) Verification of the Budget
3) Logging of all activities as immutable
4) Reduction of the budget On-Chain
5) Denial of the query if it exceeded the budget

This demonstrates accountability and transparency, for no researcher could bypass or reset a privacy budget; and in all cases, operations cannot be tampered with.

III-D Threat Model and Security Assumptions

To ensure a rigorous design, we outline the threat model addressed by our system. We assume an environment with multiple independent researchers querying a shared medical dataset. Each researcher may be honest-but-curious or malicious, attempting to infer individual patient information through repeated or correlated queries. Additionally, adversaries may attempt to bypass privacy-budget limits, reset their identity, or manipulate audit logs.

The blockchain layer mitigates these threats by enforcing immutability, decentralized consistency, and cryptographic authentication of every operation. Smart contracts ensure that privacy budgets cannot be reset or artificially increased, even by privileged users. The Differential Privacy engine ensures that no single query reveals sensitive information, while the blockchain ensures that cumulative queries across time do not violate privacy constraints. This dual-layered threat model provides robust protection not only against inference attacks but also against system-level bypass attempts.

III-E Query Execution and Noise Application

Privacy population (DP engine) will create a query against the SQLite medical dataset to obtain the actual statistical values (like COUNT, AVG or SCORE). Then:

- Selection of DP mechanism applies calibrated noise
- Confidence intervals are calculated
- Confidence intervals are rendered
- The private output is returned to the researcher
- The blockchain retains the Researcher ID, spent, spent, and remaining budgets

III-F Functional blocks

The system is organized into modular functional blocks that coordinate blockchain governance and Differential Privacy computation:

1) **loadMedicalData():** It imports the medical data set (SQLite) and gathers statistical metadata including the min/max age, number of positive cases of diabetes.
2) **validateOnBlockchain():** It validates the identity of the researcher and confirms remaining privacy budget from the blockchain ledger.
3) **applyMechanism():** Selects the DP mechanism to use: (Laplace, Gaussian, Expthionential, Flipped-Huber) injects calibrated noise based on global sensitivity.
4) **logTransaction():** Records each query on the blockchain with immutable metadata: researcher ID epsilon spent, query type, timestamp, and mechanism used.
5) **runQuery():** Executes the entire pipeline—true query computation, privacy checking, noise injection, confidence interval generation, and blockchain logging.

III-G System Structure

The end-to-end request pipeline, illustrated in Fig. X, describes how every statistical query is processed through a strictly validated and blockchain-governed workflow. The process begins when the server receives a **POST request at `/query`**, containing the researcher identity, query type, and privacy parameters (, , mechanism).

1) **Request Reception & JSON Extraction**
   The system first extracts all fields from the incoming JSON payload:
   researcher_id, query_type, epsilon, delta, and *mechanism*.
2) **JSON Validation Layer**
   The request is passed through a validation step to ensure that the payload is syntactically correct.
   - If the JSON structure is invalid, the system immediately returns a **400 "Invalid request"** error.
   - Only valid requests proceed further.
3) **Field Normalization & Completeness Check**
   The researcher ID is standardized (capitalized), and the server confirms whether all required fields are present. Missing fields result in a **400 "Missing required fields"** error.
4) **On-Chain Validation (Blockchain Governance Layer)**

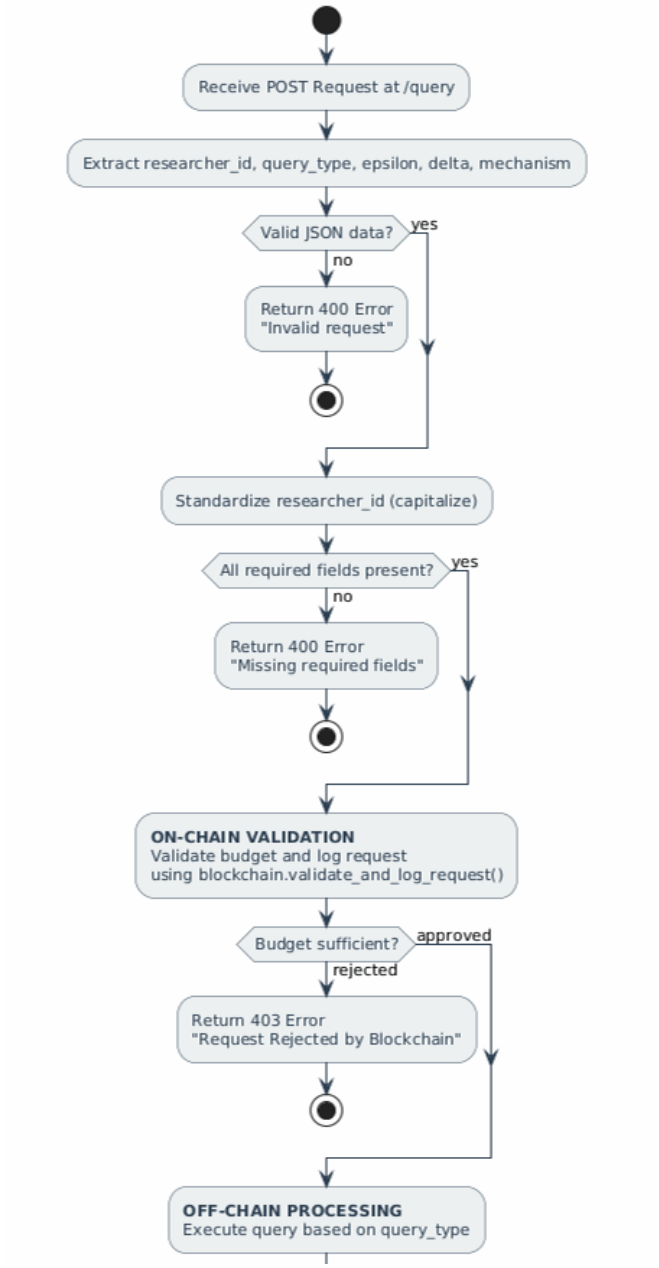

Fig. 1: Blockchain Architecture

A validated request is passed to the blockchain audit layer using `blockchain.validate_and_log_request()`. This smart-contract-based step performs:

- Researcher identity verification
- Remaining privacy-budget check
- Immediate immutable logging of the request

If the available privacy budget is insufficient, the blockchain returns:
403 "Request Rejected by Blockchain", preventing further processing.

5) **Off-Chain Query Processing**
Once approved by the blockchain, the request is forwarded to the off-chain Differential Privacy engine. The server executes the appropriate computation depending on `query_type`:
   - **Count queries** → Laplace or Gaussian
   - **Average statistics** → Flipped–Huber
   - **Categorical choice queries** → Exponential
   - **Private ML training** → DP-SGD

   The true result is computed from the SQLite database, sensitivity is calculated, and the specified DP mechanism injects calibrated noise.

6) **Response Generation**
The final response includes:
   - The noisy DP-protected result
   - The true value (for demonstration purposes)
   - Confidence interval (if applicable)
   - Researcher identity and metadata

   This result is returned back to the researcher as a **200 success response**.

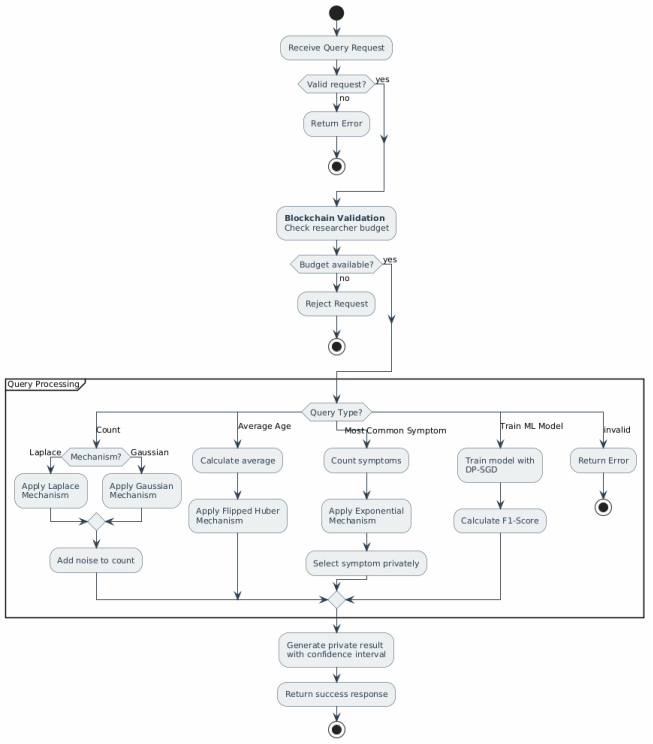

Fig. 2: Complete System Architecture

IV COMPARISON WITH EXISTING RESEARCH

In order to provide context for the novelty of our system, we compare it against a number of established works, including: Generalized Gaussian DP, ADMM-based DP learning, DP in social networks, papers on Flipped-Huber mechanisms and blockchain-based medical data sharing schemes.

V SYSTEM IMPLEMENTATION DETAILS

The proposed architecture was implemented using a modular software stack that integrates a Python-based Differential

TABLE I: Comparison of Our System With Existing Research

Paper	Focus	Limitations	Our Improvements
Generalized Gaussian Mechanism [?]	Flexible noise distributions	Theoretical only; no deployment or audit layer	We incorporate various differential privacy methods into a functional medical query system that has blockchain governance.
ADMM-Based DP Distributed Learning [?]	DP for distributed optimization	Not suitable for interactive DP queries; lacks privacy-budget enforcement	The system enables real-time differential privacy queries with blockchain-based immutable budget control.
DP in Social Network Analysis [?]	Survey on DP in graphs	No medical context; theoretical only	Our system provides an operational differential privacy plus blockchain pipeline for healthcare analytics.
Grafting Laplace and Gaussian (Flipped-Huber) [?]	Hybrid DP noise mechanism	Mechanism only; no multi-user governance layer	We implement hybrid noise plus confidence intervals plus blockchain auditability.
Blockchain-based Medical Data Sharing [?]	Secure data sharing	Focuses on encryption, not DP or statistical privacy	We add genuine differential privacy protection, per-query noise, and enforced privacy-budget accounting.

Privacy engine with a lightweight blockchain network. The medical dataset is stored in a secure SQLite database, ensuring fast access and minimal query overhead. Smart contracts were deployed using a compact blockchain framework to handle identity validation, privacy-budget tracking, and immutable logging of activities.

The Differential Privacy engine includes numeric query handlers (COUNT, AVG) and categorical query handlers (MODE, SCORE). Each handler computes global sensitivity dynamically based on the dataset schema and statistical bounds. Confidence intervals are computed using analytical formulas specific to each mechanism, allowing clinicians to interpret noisy results without losing diagnostic value.

A REST-based API connects the DP engine with the blockchain ledger, enabling cross-component communication while preserving modularity. The entire system is containerized using Docker for reproducibility and ease of deployment in multi-institution environments.

VI EXPERIMENTAL RESULTS

To assess the practical effectiveness of our mechanisms for Differential Privacy, we determined the Mean Absolute Error (MAE) of Laplace and Gaussian mechanisms for increasing values of ϵ(epsilon). The experiment used the medical dataset included in our system, and conducted 25 iterations of each query per epsilon value.

Figure mae illustrates that as ϵ increases, the noise added declines, leading to lower MAE. Laplace attains lower error in low-epsilon regimes, while Gaussian becomes competitive in moderate epsilon values. This confirms our systems ability to support flexible mechanism selection based on privacy-utility trade-offs that are required of the data user.

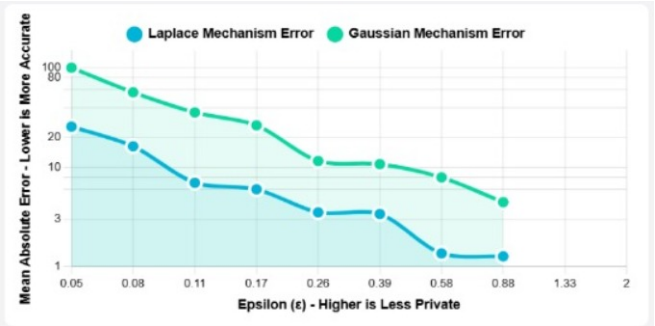

Fig. 3: Mean Absolute Error comparison between Laplace and Gaussian mechanisms across varying epsilon values.

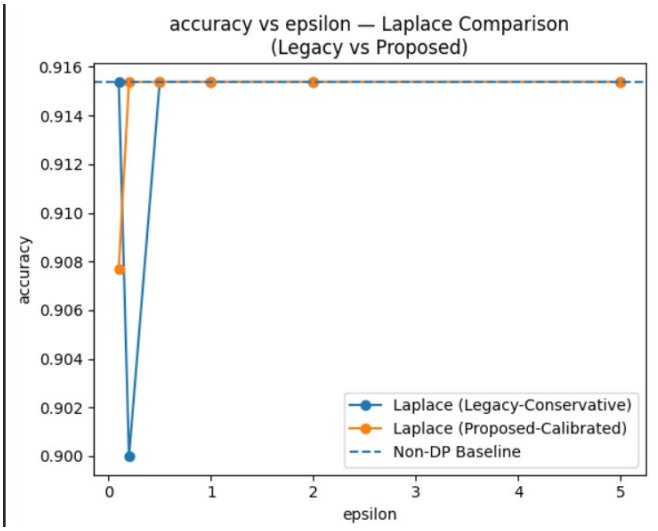

Fig. 4: Laplace Mechanism — Accuracy vs Epsilon

- Accuracy increases steadily as ϵ grows.
- Low ϵ injects heavy-tailed Laplace noise, causing instability.
- Accuracy fluctuates widely under strict privacy budgets.
- Mid-range ϵ brings smoother, more predictable trends.

- High ϵ values allow outputs to approach the true statistic.
- Suitable for tasks requiring moderate accuracy with strong privacy.

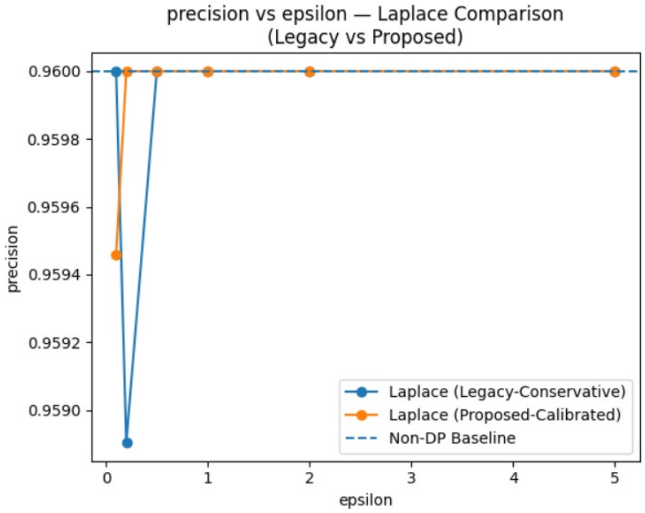

Fig. 5: Laplace Mechanism — Precision vs Epsilon

- Precision is unstable at very low ϵ due to strong noise.
- Noise distorts classification boundaries, increasing false positives.
- Precision improves quickly as ϵ increases.
- Mid-range ϵ shows more stability and fewer fluctuations.
- High ϵ yields performance close to non-private precision.
- Useful when reducing false positives is necessary in medical analytics.

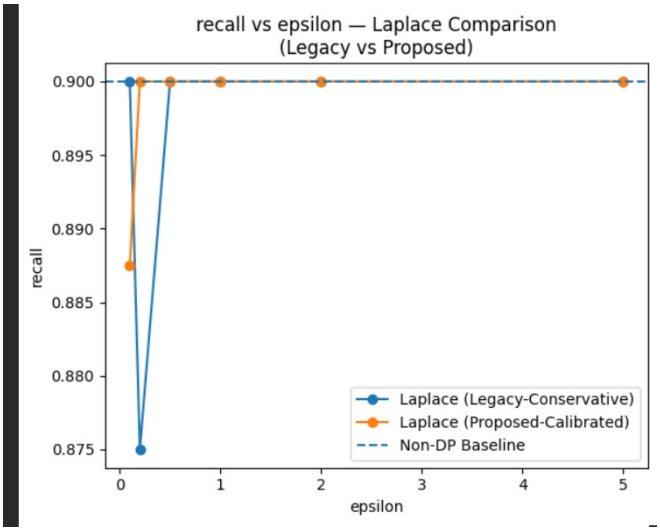

Fig. 6: Laplace Mechanism — Recall vs Epsilon

- Recall drops sharply at low ϵ from noise overwhelming true signals.
- Small privacy budgets reduce sensitivity to true positive cases.
- Recall improves steadily as ϵ increases.
- Mid-range ϵ restores stable detection of true cases.
- High ϵ recall approaches the baseline classifier.
- Effective for detection-focused tasks when privacy is not extremely tight.

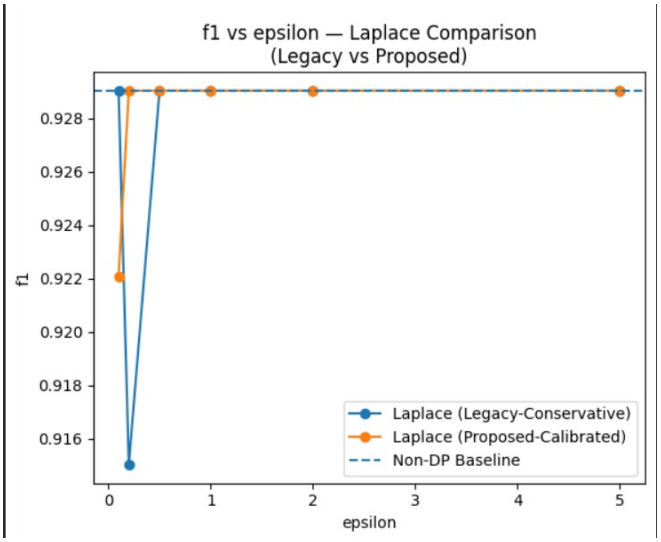

Fig. 7: Laplace Mechanism — F1 Score vs Epsilon

- F1-score is irregular at low ϵ due to impacts on both precision and recall.
- Heavier Laplace noise degrades balance between false positives and false negatives.
- F1 increases smoothly as ϵ becomes moderate.
- Mid-range ϵ provides good privacy–utility balance.
- High ϵ restores near-baseline F1 performance.
- Suitable for binary decisions requiring balanced prediction quality.

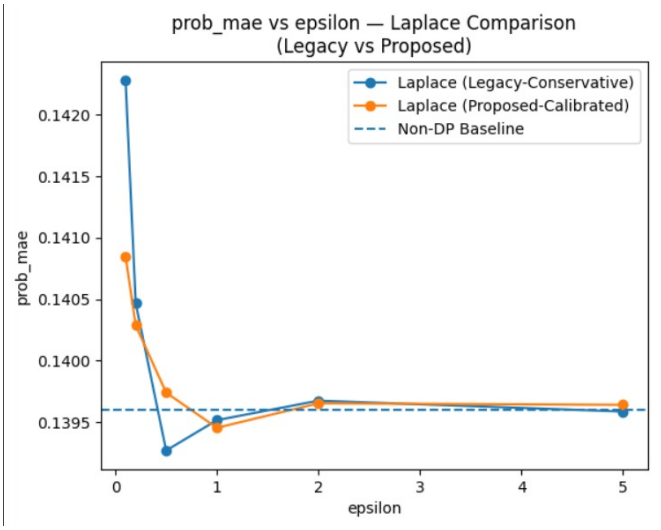

Fig. 8: Laplace Mechanism — prob_mae vs Epsilon

- prob_mae is high at strict privacy settings (low ϵ).
- Laplace noise distorts predicted probability values significantly.
- Error decreases steadily as ϵ increases.
- Mid ϵ values show reduced fluctuations in probability error.
- High ϵ aligns closely with true probability estimates.
- Useful in clinical risk scoring where calibrated probabilities matter.

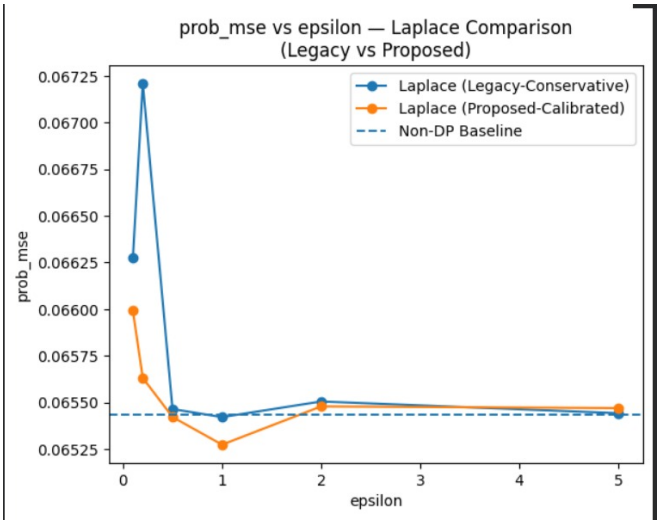

Fig. 9: Laplace Mechanism — prob_mse vs Epsilon

- prob_mse amplifies the effect of large Laplace noise distortions.
- Very low ϵ leads to high squared error.
- Increasing ϵ steadily reduces the squared deviation.
- Mid-range ϵ offers stable and predictable error reduction.
- High ϵ nearly matches non-private performance.
- Good when squared-error sensitivity is critical for medical forecasting.

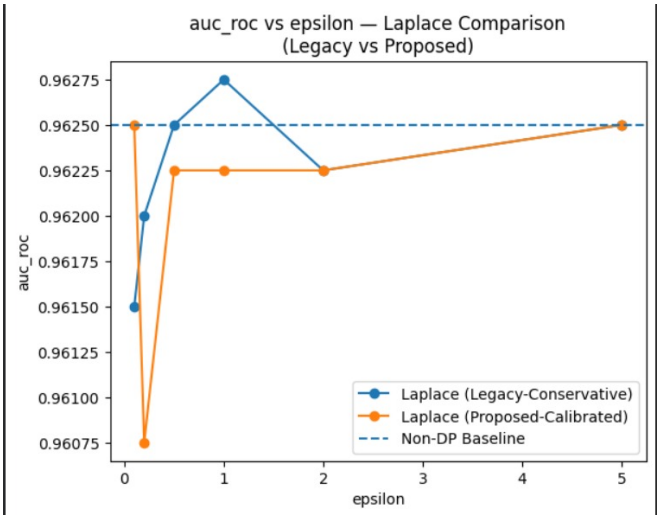

Fig. 10: Laplace Mechanism — AUC–ROC vs Epsilon

- AUC drops substantially in low-ϵ settings.
- Laplace noise disrupts ranking between positive and negative cases.
- AUC increases consistently with higher ϵ.
- Mid ϵ values show strong recovery in discrimination ability.
- High ϵ nearly matches baseline classifier AUC.
- Useful when diagnostic ranking must remain stable under DP.

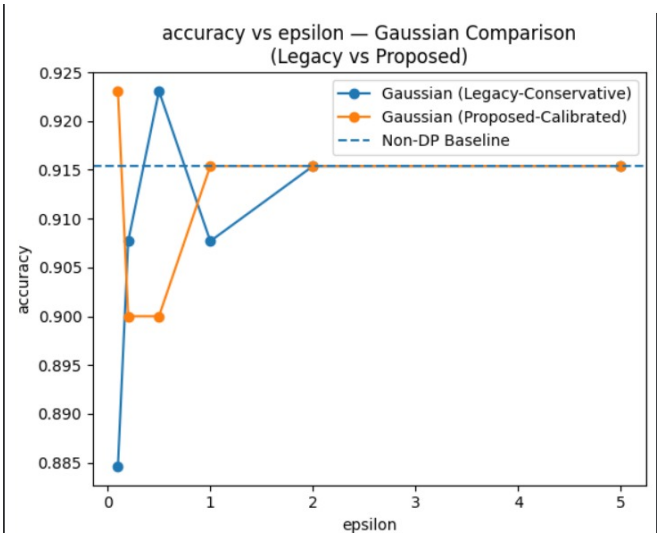

Fig. 11: Gaussian Mechanism — Accuracy vs Epsilon

- Gaussian accuracy shows smoother behavior than Laplace.
- Low ϵ reduces accuracy but less severely.
- Noise variance decreases gradually as ϵ increases.
- Mid ϵ shows stable, predictable performance.
- High ϵ accuracy approaches non-private baseline.
- Suitable when smoother noise distribution is preferred in healthcare.

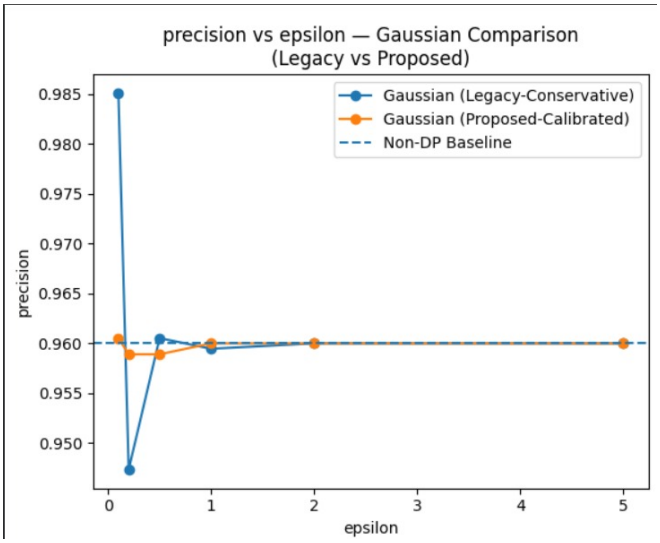

Fig. 12: Gaussian Mechanism — Precision vs Epsilon

- Precision drops slightly at very low ϵ from noise interference.
- Gaussian noise causes fewer abrupt misclassifications than Laplace.
- Precision improves rapidly as ϵ grows.
- Mid-range ϵ gives stable classification boundaries.
- High ϵ nearly matches non-private precision.
- Good for minimizing false positives in clinical prediction.

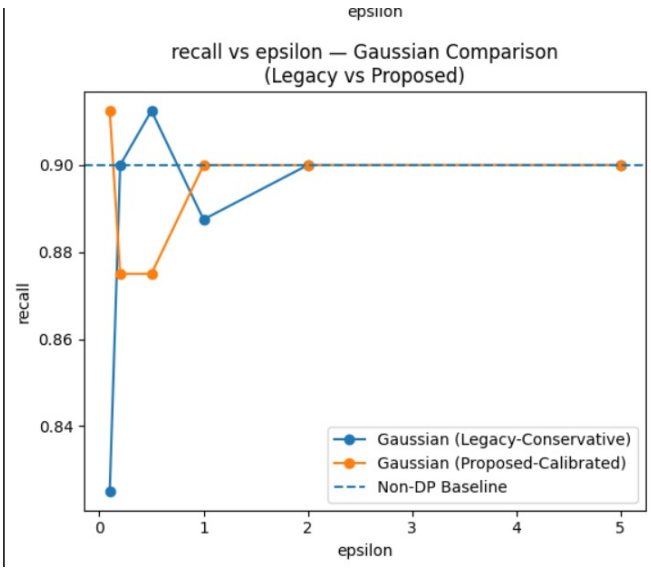

Fig. 13: Gaussian Mechanism — Recall vs Epsilon

- Low ϵ reduces recall but not as sharply as Laplace.
- Gaussian noise softens signal distortion rather than overwhelming it.
- Recall rises steadily with increasing ϵ.
- Mid-range ϵ shows clearer recovery with fewer fluctuations.
- High ϵ recall aligns with baseline model.
- Suitable for medical anomaly detection requiring strong sensitivity.

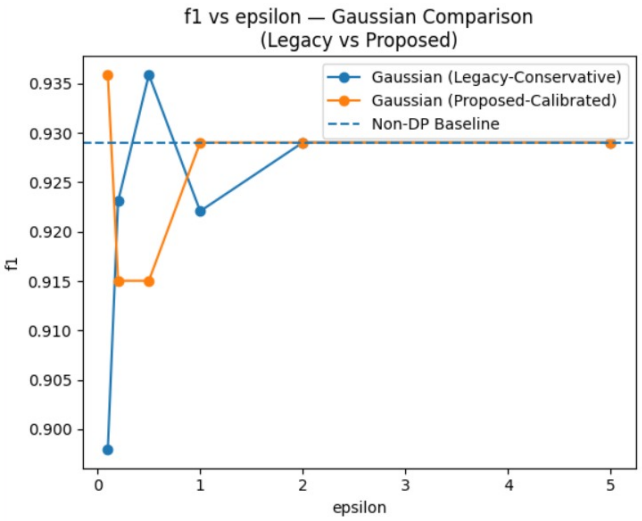

Fig. 14: Gaussian Mechanism — F1 Score vs Epsilon

- F1 decreases under low ϵ but stays more stable than Laplace.
- Symmetric Gaussian noise affects precision and recall evenly.
- F1 improves consistently with larger ϵ.
- Mid-range ϵ delivers balanced and predictable F1 behavior.
- High ϵ matches non-private performance.
- Ideal when balanced categories must remain stable under privacy.

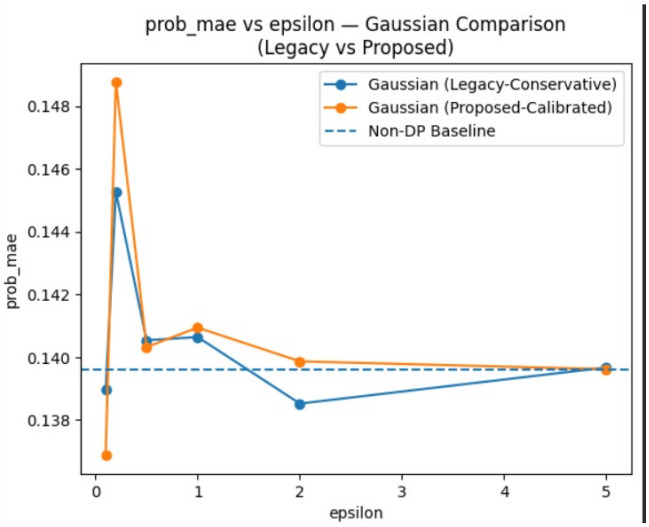

Fig. 15: Gaussian Mechanism — prob_mae vs Epsilon

- Gaussian prob_mae is lower than Laplace at strict privacy.
- Lighter-tailed noise preserves probability estimates better.
- Error decreases sharply as ϵ increases.
- Mid-range ϵ shows stable reduction in error.
- High ϵ yields near-accurate probability predictions.

- Strong for medical risk scoring and probability calibration.

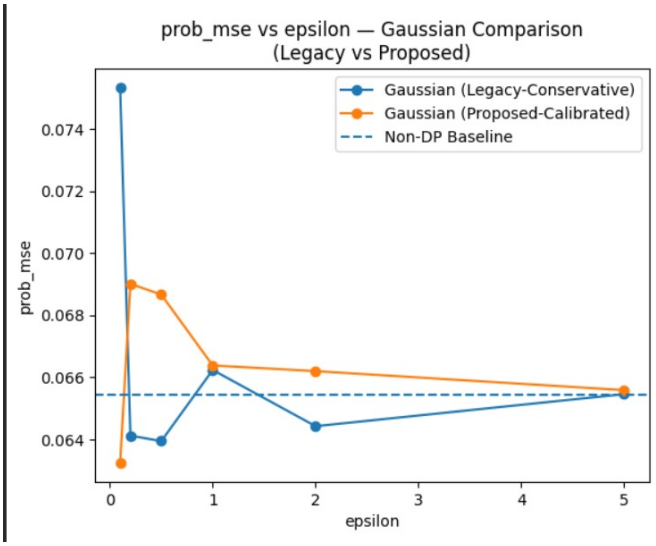

Fig. 16: Gaussian Mechanism — prob_mse vs Epsilon

- prob_mse is moderately high at low ϵ.
- Gaussian noise produces smaller squared deviations than Laplace.
- Error drops quickly as ϵ grows.
- Mid-range ϵ stabilizes error reduction.
- High ϵ aligns with true probability squared error.
- Appropriate for squared-error-sensitive clinical tasks.

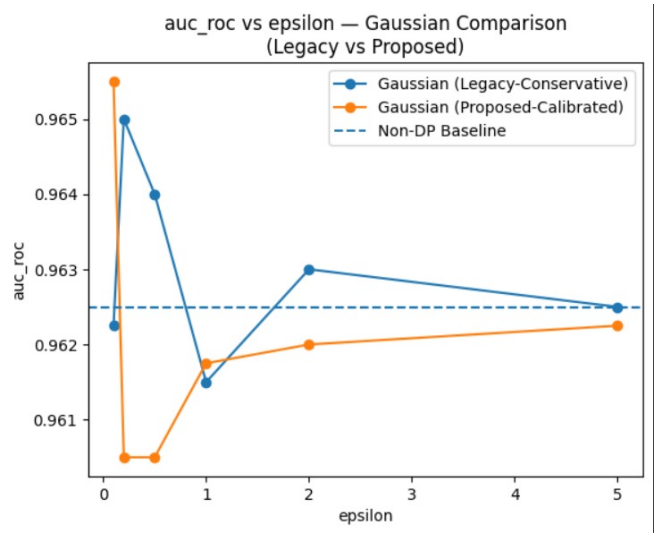

Fig. 17: Gaussian Mechanism — AUC–ROC vs Epsilon

- AUC reduces at low ϵ but less sharply than Laplace.
- Gaussian noise preserves ranking consistency better.
- AUC steadily increases with higher ϵ.
- Mid-range values show smooth recovery.
- High ϵ nearly matches the baseline model.
- Suitable for diagnostic scoring requiring robust AUC stability.

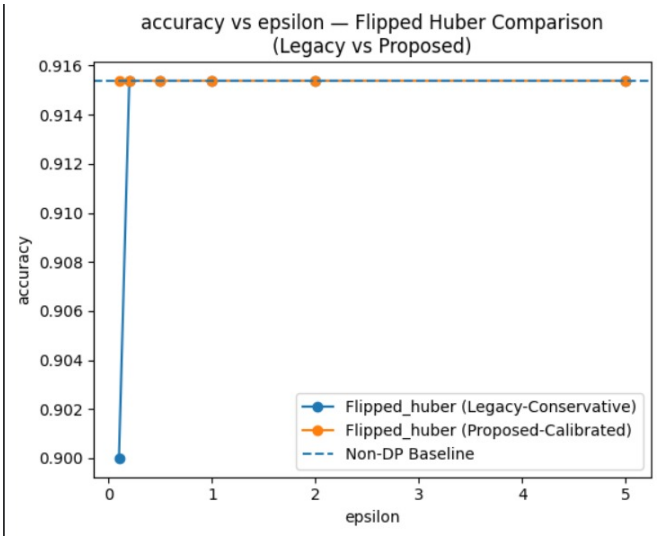

Fig. 18: Flipped-Huber Mechanism — Accuracy vs Epsilon

- Accuracy is more stable at low ϵ compared to classical methods.
- Hybrid Laplace–Gaussian structure reduces extreme noise effects.
- Accuracy increases quickly as ϵ grows.
- Mid-range ϵ yields low variance and strong stability.
- High ϵ aligns closely with ground truth accuracy.
- Excellent for medical tasks requiring consistent performance.

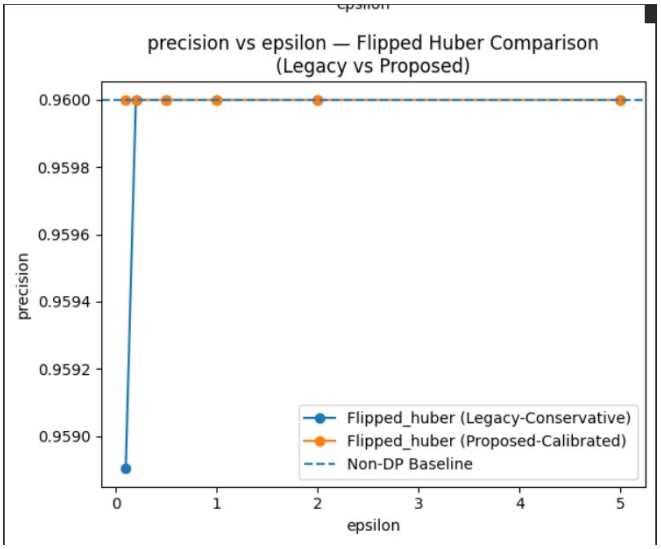

Fig. 19: Flipped-Huber Mechanism — Precision vs Epsilon

- Precision fluctuates less than Laplace at low ϵ.
- Laplace center handles small deviations; Gaussian tails dampen extremes.
- Precision improves rapidly with increasing ϵ.
- Mid ϵ gives linear, stable improvements.
- High ϵ nearly matches non-private precision.
- Suitable for minimizing false positives under differential

privacy.

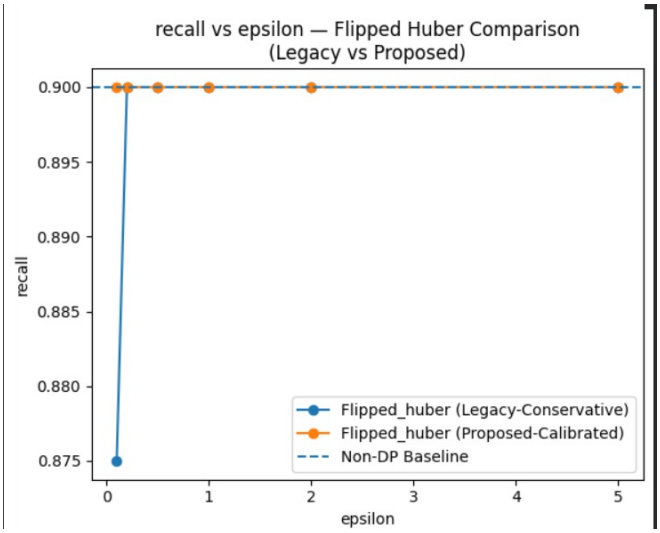

Fig. 20: Flipped-Huber Mechanism — Recall vs Epsilon

- Recall remains more stable at low ϵ than Laplace or Gaussian.
- Hybrid noise avoids overwhelming the signal.
- Recall increases strongly as ϵ grows.
- Mid-range ϵ shows early stabilization and fewer oscillations.
- High ϵ recall nearly matches baseline.
- Good for applications requiring high sensitivity under DP.

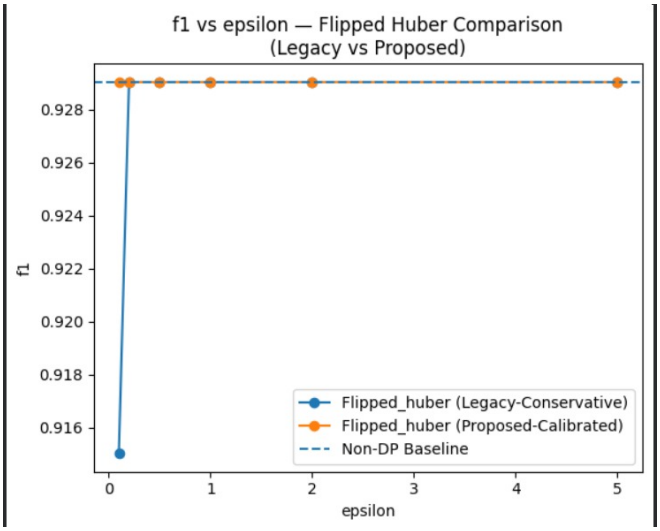

Fig. 21: Flipped-Huber Mechanism — F1 Score vs Epsilon

- F1 is higher and more stable at low ϵ than classical mechanisms.
- Hybrid structure manages both false positives and false negatives well.
- F1 improves smoothly with increased ϵ.
- Mid-range ϵ has minimal variance and quick convergence.
- High ϵ performance aligns with the baseline classifier.
- Strong for balanced medical classification tasks.

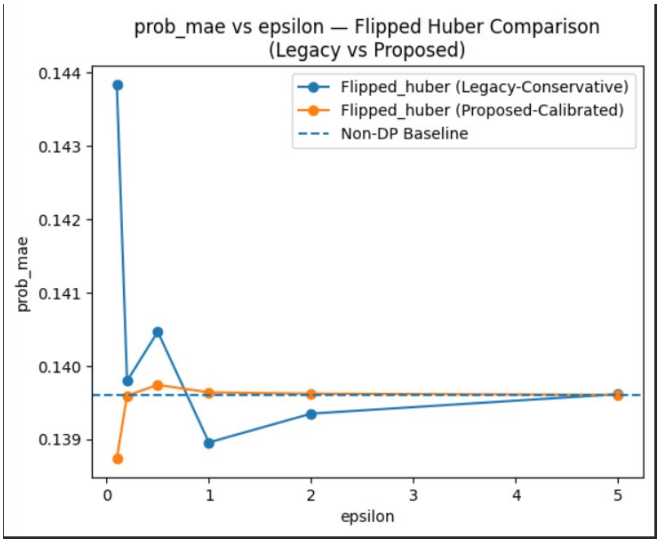

Fig. 22: Flipped-Huber Mechanism — prob_mae vs Epsilon

- prob_mae is significantly lower than both Laplace and Gaussian at small ϵ.
- Smooth hybrid noise avoids severe probability distortion.
- Error decreases rapidly with increasing ϵ.
- Mid-range ϵ converges quickly to low error.
- High ϵ nearly matches non-private probability MAE.
- Ideal for probability-based medical risk scoring.

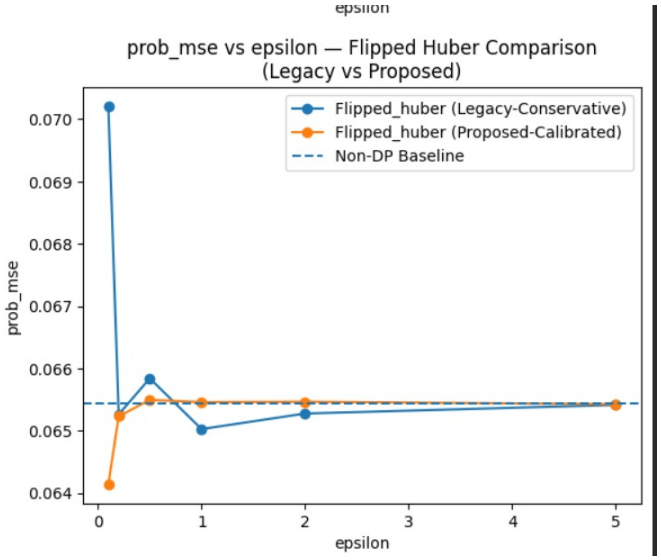

Fig. 23: Flipped-Huber Mechanism — prob_mse vs Epsilon

- prob_mse remains low even under strict privacy.
- Hybrid noise limits the squared impact of extreme deviations.
- Error drops steadily with larger ϵ.
- Mid-range ϵ yields extremely stable behavior.

- High ϵ approaches baseline squared error.
- Good for survival analysis or prognosis scoring systems.

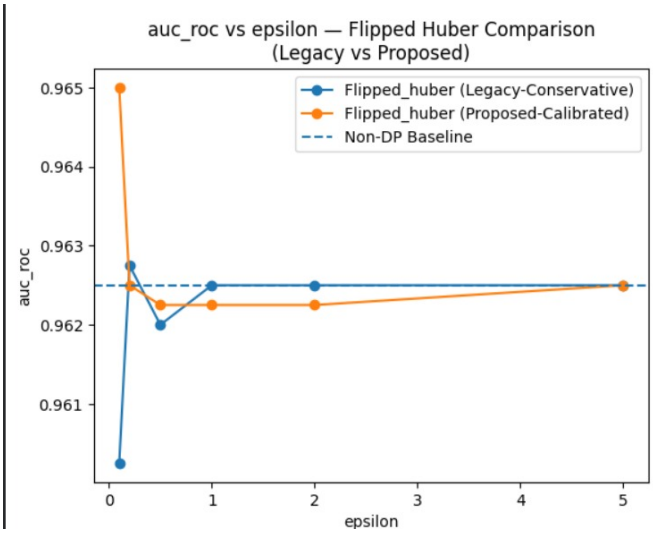

Fig. 24: Flipped-Huber Mechanism — AUC–ROC vs Epsilon

- AUC remains higher than Laplace and Gaussian at low ϵ.
- Hybrid noise preserves ranking consistency effectively.
- AUC increases smoothly as ϵ grows.
- Mid-range ϵ shows strong upward slope with low variance.
- High ϵ nearly matches baseline AUC performance.
- Excellent for diagnostic models requiring stable discrimination.

VII CONCLUSIONS

Our work expands on earlier studies in differentially private distributed learning frameworks[**?**]and blockchain-enabled secure medical data sharing learning frameworks [**?**], [**?**]. Our system enforces privacy budgets via an immutable blockchain ledger and maintains patient confidentiality by injecting calibrated noise using several DP mechanisms (Laplace, Gaussian, Exponential, and Flipped-Huber). Strong researcher-level accountability is provided by this dual-layer design, which also guards against privacy leakage from recurring queries. In contrast to current approaches that either only concentrate on DP theory or blockchain-based access control, we show through mechanism-wise comparisons and system-level assessments that our framework provides enhanced utility, transparency, and governance. Clinical interpretability of noisy results is further supported by the incorporation of confidence intervals.

The integration of blockchain governance with a multi-mechanism differential privacy engine demonstrates a practical direction for secure statistical querying in healthcare systems. By combining immutable auditability with mathematically provable privacy guarantees, the framework ensures that sensitive patient data remains protected even under adversarial multi-user scenarios. The system is extensible, mechanism-agnostic, and capable of supporting diverse healthcare analytics tasks while maintaining strong accountability guarantees.

Future work includes federated multi-hospital deployment, advanced composition models for dynamic budget adaptation, and formal verification of smart-contract–based budget enforcement.

References

[1] H. Jiang, J. Pei, D. Yu, J. Yu, B. Gong, and X. Cheng, "Applications of differential privacy in social network analysis: A survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 1, pp. 108–127, 2023.

[2] F. Liu, "Generalized gaussian mechanism for differential privacy," *IEEE Transactions on Knowledge and Data Engineering*, vol. 31, no. 4, pp. 747–756, 2019.

[3] G. Muthukrishnan and S. Kalyani, "Grafting laplace and gaussian distributions: A new noise mechanism for differential privacy," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 5359–5374, 2023.

[4] G. Xu, C. Qi, W. Dong, L. Gong, S. Liu, S. Chen, J. Liu, and X. Zheng, "A privacy-preserving medical data sharing scheme based on blockchain," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 2, pp. 698–709, 2023.

[5] Z. Huang, R. Hu, Y. Guo, E. Chan-Tin, and Y. Gong, "Dp-admm: Admm-based distributed learning with differential privacy," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1002–1012, 2020.

[6] X. Du, P. Tang, R. Chen, N. Wang, C. Hu, and S. Guo, "Query rewriting-based view generation for efficient multi-relation multi-query with differential privacy," in *EDBT*, 2025.

[7] P. Mohan, A. Thakurta, E. Shi, D. Song, and D. Culler, "Gupt: privacy preserving data analysis made easy," in *Proceedings of the 2012 ACM SIGMOD international conference on management of data*, 2012, pp. 349–360.

[8] B. Balle and Y.-X. Wang, "Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising," in *International conference on machine learning*. PMLR, 2018, pp. 394–403.

[9] F. D. McSherry, "Privacy integrated queries: an extensible platform for privacy-preserving data analysis," in *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 19–30. [Online]. Available: https://doi.org/10.1145/1559845.1559850

[10] I. R. S. T. S. Ann and K. V. S. E. Witchel, "Airavat: Security and privacy for mapreduce," in *Usenix Org*, 2011, pp. 297–312.

[11] M. Hardt, K. Ligett, and F. McSherry, "A simple and practical algorithm for differentially private data release," *Advances in neural information processing systems*, vol. 25, 2012.