Empirical Analysis of Copy-Number Effects in Helstrom Quantum Classification for Titanic Survival

Mujirin, Jae Uk Roh, Wook Park, and Hyundong Shin
Department of Electronics and Information Convergence Engineering, Kyung Hee University, Republic of Korea
Email: hshin@khu.ac.kr

***Abstract*—We investigate the practical gains of multi-copy quantum discrimination on a real-world dataset. Using the Titanic survival data, we embed every two-feature pair into a qubit space via three schemes: Amplitude, phase, and angle encoding, and form class density operator by averaging. For each encoder, we apply the Helstrom measurement to n-copy tensor products ($1 \leq n \leq 8$), recording test accuracy and the commutator between two density classes. Contrary to asymptotic theory, our experiments reveal no strong copy-dependent accuracy gains ($r = 0.042$) and only a weak positive link between non-commutativity of the class density operator and accuracy ($r = 0.174$). Surprisingly, the commutator norm itself slightly decreases with n ($r = -0.114$). These results show that, in low-dimensional, finite-sample settings, simply increasing the number of copies does not guarantee improved classification performance.**

I. Intorduction

Quantum hypothesis testing provides a rigorous framework for binary classification by leveraging the Helstrom measurement [1], which is provably Bayes-optimal for distinguishing between two quantum states. Theoretically, when the two density operators do not commute, preparing n copies of each state in the asymptotic limit yields an exponential decay of the minimum error probability—known as the quantum Chernoff bound [2]—implying that additional copies should monotonically improve classification accuracy. However, real-world datasets often violate these ideal conditions [3], and empirical guidance on how copy number, operator non-commutativity, and embedding choice interact in practice remains scarce.

In this paper, we systematically explore these interactions on the classical Titanic survival dataset by evaluating every two-feature combination of passenger attributes using the Helstrom Quantum Classifier (HQC) [4] and, for comparison, the classical algorithms Decision Tree, Random Forest (RF), Logistic Regression (LR), Support Vector Machine (SVM), and k-Nearest Neighbors (k-NN) [5]. For the quantum part, we embed each two-dimensional feature vector into a quantum state via three distinct schemes—amplitude, phase, and angle encoding [3]—and construct class prototypes by simple averaging. For each encoder, we then apply the Helstrom measurement on n-copy tensor products (with n ranging from 1 to 8) and record both the test accuracy and the the commutator. Finally, we compute the correlations to quantify (a) how non-commutativity grows with n, (b) how accuracy grows with n, (c) how accuracy relates to non-commutativity, and (d) how HQC compares to classical algorithm.

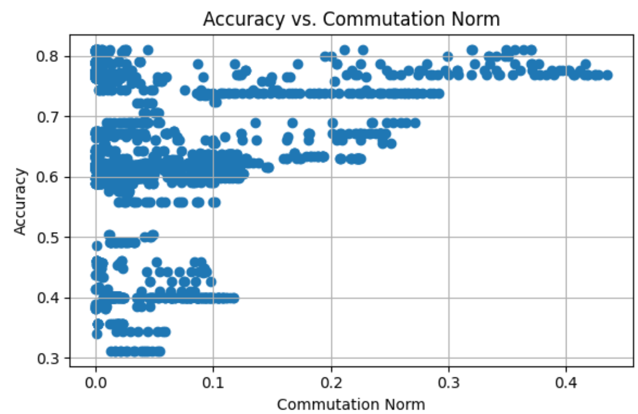

Figure 1: Plot of accuracy vs commutation norm of class density operators of HQC.

Our key empirical findings are that, across hundreds of two-feature tests and all three embeddings, (a) the commutator norm shows a slight, negative correlation with copy number ($r = -0.114$, see Fig. 1), (b) accuracy shows virtually no correlation with copy number ($r = 0.042$), (c) accuracy has only a weak positive correlation with non-commutativity ($r = 0.174$), and (d) the HQC achieves accuracy comparable to that of classical algorithms (see Fig. 2). These results challenge the straightforward application of the quantum Chernoff intuition.

II. Quantum Helstrom Binary Classifier

A. Quantum Data Embeddings

We map each 2D feature $x = (x_0, x_1)$ to a quantum state by first normalizing [3]

$$u = \frac{x}{\|x\|}, \qquad \|x\| = \sqrt{x_0^2 + x_1^2}, \tag{1}$$

and substituting $\frac{1}{2}I$ if $\|x\|$ falls below a threshold.

1) Amplitude Encoding: Use the real components of u as amplitudes [3]:

$$|\psi_{\text{amp}}\rangle = u_0 |0\rangle + u_1 |1\rangle, \quad \rho_{\text{amp}} = |\psi_{\text{amp}}\rangle\langle\psi_{\text{amp}}|. \tag{2}$$

2) Phase Encoding: Keep $|u_0|$ and $|u_1|$ but add a phase $\phi = \pi\, u_1$ [3]:

$$|\psi_{\text{ph}}\rangle = u_0 |0\rangle + e^{i\phi} u_1 |1\rangle, \quad \rho_{\text{ph}} = |\psi_{\text{ph}}\rangle\langle\psi_{\text{ph}}|. \tag{3}$$

3) Angle Encoding: Interpret u as spherical angles [4]:

$$\theta = 2\arccos(u_0), \quad \varphi = \arg(u_0 + i\, u_1), \tag{4}$$

$$|\psi_{\text{ang}}\rangle = \cos\tfrac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\tfrac{\theta}{2}\,|1\rangle, \quad \rho_{\text{ang}} = |\psi_{\text{ang}}\rangle\langle\psi_{\text{ang}}|. \tag{5}$$

B. Class Density Operators by Simple Averaging

Given a training set $\{(x_i, y_i)\}$ with binary labels $y_i \in \{0,1\}$, one computes for each class the empirical density matrix [4]

$$\rho_0 = \frac{1}{N_0}\sum_{i:y_i=0}\rho(x_i)\,, \qquad \rho_1 = \frac{1}{N_1}\sum_{i:y_i=1}\rho(x_i), \tag{6}$$

where N_k is the number of samples in class k. This "prototype" construction requires no iterative parameter fitting—just a weighted average of the embedded states.

C. Minimum-Error Measurement (Helstrom POVM)

The goal is to decide between the two class states ρ_0 and ρ_1 with lowest possible error, assuming equal prior probabilities. One defines the Helstrom operator Δ [1], and diagonalizes it in its eigenbasis. The optimal two-element POVM for binary discrimination is

$$\Delta = \tfrac{1}{2}(\rho_0 - \rho_1),\ E_0 = \sum_{\lambda_i > 0}|v_i\rangle\langle v_i|\ ,\ E_1 = I - E_0, \tag{7}$$

where $\{\lambda_i, |v_i\rangle\}$ are the eigenpairs of Δ. Measuring an unknown state $\rho(x)$ with this POVM yields outcome 0 (class 0) with probability $\mathrm{Tr}[\rho(x)\,E_0]$, and 1 otherwise. This prescription is known to minimize the average error probability $P_{\text{err}} = \frac{1}{2}\Big(1 - \|\Delta\|_1\Big)$, where $\|\Delta\|_1 = \sum_i |\lambda_i|$ is the trace norm of Δ.

D. Multi-Copy Extension

To exploit an exponential convergence of error, one can prepare n identical copies of each embedded state and form [2]

$$\rho_0^{\otimes n}, \quad \rho_1^{\otimes n}, \quad \Delta^{(n)} = \tfrac{1}{2}\big(\rho_0^{\otimes n} - \rho_1^{\otimes n}\big). \tag{8}$$

Repeating the POVM construction on the 2^n-dimensional space yields an error probability scaling as $P_{\text{err}}^{(n)} \approx \exp(-\xi\, n)$, where $\xi > 0$ is the quantum Chernoff exponent. The classification proceeds by computing $\mathrm{Tr}[\rho(x)^{\otimes n}\, E_0^{(n)}] \geq 0.5$.

E. Non-Commutativity and Encoding Quality

Because ρ_0 and ρ_1 need not commute, the Helstrom POVM can exploit genuinely quantum interference effects. In practice, one can quantify the degree of non-commutativity by the Frobenius norm [2] of their commutator

$$\big\|[\rho_0, \rho_1]\big\|_F = \big\|\rho_0\,\rho_1 - \rho_1\,\rho_0\big\|_F. \tag{9}$$

A small commutator norm indicates nearly classical (commuting) prototypes—often leading to limited quantum advantage—whereas larger norms signal richer quantum discrimination potential.

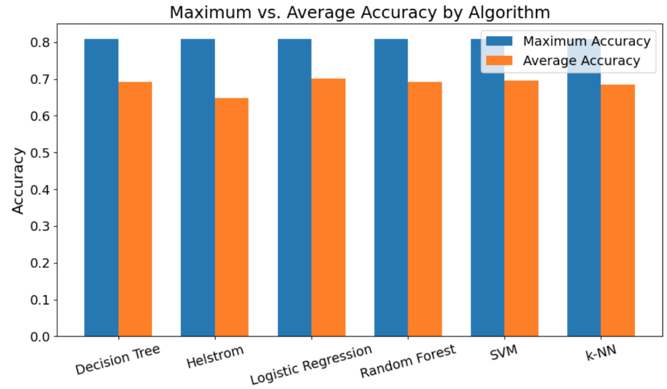

Figure 2: Maximum (blue) and average (orange) accuracies for each algorithm.

III. Conclusion

Our extensive empirical study reveals that the theoretical promise of multi-copy Helstrom discrimination—namely, monotonic accuracy gains with increasing n—does not automatically materialize on real data. In particular, amplitude encoding yields no clear benefit beyond the single-copy case, while phase and angle embeddings produce modest improvements up to a small optimal n before plateauing. The near-zero correlation between accuracy and copy number ($r = 0.042$) and only a weak link to non-commutativity ($r = 0.174$) indicate that additional copies can amplify sampling noise without unlocking proportional quantum advantage unless the embedding induces sufficiently strong non-commuting class states. As future work, we plan to further investigate regions where the class density operators are nearly commuting, and to implement Helstrom multi-copy classification on an FPGA to accelerate the diagonalization process, as in [6].

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) under RS-2025-00556064, by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2025-2021-0-02046) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), and by a grant from Kyung Hee University in 2023 (KHU-20233663).

References

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory*. Academic Press, 1976.

[2] K. M. R. Audenaert *et al.*, "Discriminating states: The quantum chernoff bound," *Phys. Rev. Lett.*, vol. 98, no. 16, p. 160501, Apr. 2007.

[3] V. Havlíček *et al.*, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, pp. 209–212, Mar. 2019.

[4] M. Schuld and F. Petruccione, *Supervised Learning with Quantum Computers*. Springer, 2019.

[5] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. Springer, 2009.

[6] T. Bonny and A. Haq, "Emulation of high-performance correlation-based quantum clustering algorithm for two-dimensional data on fpga," *Quantum Information Processing*, vol. 19, no. 179, 2020. [Online]. Available: https://link.springer.com/article/10.1007/s11128-020-02683-9