Author Rebuttal

Rebuttal to Reviewer gVe4

Thank you very much for this thorough and well-reasoned review of our paper. We appreciate your recognition of the innovation in the framework we have developed using GAN technology for detecting anomalies. You recognized how we have developed an innovative way to extract features, used expert knowledge in creating a regularization that aids in producing reliable results, and conducted an extensive set of experiments to validate our hypothesis. We appreciate your constructive feedback and provide our own responses below to each of your comments.

Appreciation of Strengths Identified

We wish to thank the reviewer for their acknowledgment of:

• The innovative and unique way the proposed GAN-based anomaly detection framework extracts features, specifically the unique way in which quantum-simulated embeddings are generated.

• The meaningfully methodologically relevant combination of discriminator loss and expert-created regularization terms (angle of incidence deviation and squared residual).

• The inclusion of an ablation study of all the core components greatly enhances the empirical evaluation.

All three of these areas are the main technical contributions of this paper, and we thank the reviewer for acknowledging their significance.

Response to Identified Weak Points

1. Lack of clarity and rigor in presentation

Thank you to the Reviewer for indicating their concern; we acknowledge the concern and appreciate the Reviewer's honesty with respect to the manuscript. As noted in the prior version, conceptual breadth was favoured, which resulted in areas that would benefit from additional rigour and clarity. As noted in the previous version, we will update the manuscript to address each of these issues:

- Simplify convoluted sentences and replace unconventional terminology with standard language used in anomaly detection and GAN literature.
- Add an explicit intuition-driven explanation of the quantum embedding, clarifying its role in capturing higher-order feature interactions relevant to camouflage attacks.
- Provide clear architectural specifications of the generator and discriminator (layer structure, activations, and training details).
- Clarify the training and update procedure, explicitly stating that the model is trained on benign traffic and that no full retraining is assumed during deployment.

We agree that these changes are necessary to improve rigor and readability.

2. Marginal improvements over a classical GAN

We would like to thank the reviewer for their detailed assessment and comparisons of our algorithm to classical GAN baseline algorithms. Even though the numerical performance increases appear small compared to classical GANs, there are several aspects to consider:

1. The primary benefit of our algorithm is its robustness and stability against camouflage and adversarial attack conditions; it is not focused only on how accurately (raw accuracy) our algorithm predicts target labels.

2. The statistical analysis included in this paper indicates that the improvements seen in our work are statistically significant and were aided by the presence of large effect sizes.

3. The results of the ablation studies show that each of the new components included in the proposed algorithm consistently aids the performance and robustness of the proposed algorithm even when comparing their overall performance as determined by headline performance metrics. Therefore, in order to communicate the value of new performance gains connected with this algorithm, we will revise the manuscript to explain more clearly the distinction between raw threshold performance improvements and the context in which these performance improvements were achieved..

3. Insufficient discussion of baselines and benchmarks

We thank the reviewer for highlighting this gap. In the revised version, we will:

- Add a **dedicated subsection on benchmarks**, detailing the characteristics and limitations of the network traffic dataset used.
- Explicitly describe all baselines, including how the classical GAN baseline is obtained by disabling advanced components.
- Clarify that FGSM and PGD are used as **robustness stress tests**, not as competing anomaly detection methods.

This will ensure greater transparency and reproducibility.

4. Emphasis on temporal dynamics and evolving threats

Your feedback regarding this manuscript is very helpful. The current draft does not make a clear distinction between different types of dataset drift. Accordingly, it is important to clarify that:

* The proposed methodology is not intended to identify all forms of arbitrary dataset drift for datasets in their normal distribution state.

* The purpose for implementing temporal evolution tracking includes the ability to identify the adversarial evolution path (relative to a stable workload) seen with both polymorphic and other types of camouflage cyber-attack scenarios.

* Temporal regularization works with the collected data to construct normalized deviations from the reference centroid sequentially versus treating each data point individually.

We recognize that the intent of our methodology must be properly defined and framed, and as such we will revise the manuscript to avoid the tendency to provide excessive emphasis, and either add or modify the existing explanation to show more clearly the reasons for why we used these techniques to identify anomaly patterns.

5. Non-compliance with the CPAIOR template

We acknowledge that this issue was by no means intentional; we appreciate the feedback from the reviewer. The reason for the incorrect formatting was due to the use of a draft layout while we were preparing the article for publication. The camera-ready copy will be submitted to CPAIOR in compliance with their guidelines, so font size, page margins, and page limits will all be submitted correctly. The technical content of the work has not been impacted by this problem in any way.

Responses to Rebuttal Questions

Q1: Details on the Quantum embeddings

A method was created to utilize quantum embeddings through the utilization of variational quantum circuits which have been simulated through classical hardware. This method starts by encoding classical feature vectors into parameterised quantum states using rotation gates, and then applies an entanglement layer to allow the quantum state of the model to represent non-linear and/or high-order correlations that would normally be difficult to encode using strictly classical methods. Further details on our intuition for the circuit designs and a description of how parameters were chosen, will be included in the next version of the manuscript.

Q2: Details on baselines and the benchmark

The assessment employs a conventional network traffic benchmark that has been chosen based on its association with contemporary polymorphic attacks. The base comparisons include both a classical Generative Adversarial Network (GAN) and all of the progressively diminished versions of the suggested method. The performance will be evaluated with several metrics and validated through statistical means in accordance with the current literature. We will include further details regarding both the benchmark and the definitions of baseline in our methodology to help alleviate reviewers' concerns.

The authors appreciate the reviewer for doing an excellent job with the review, as well as providing constructive feedback. While the authors agree with the reviewer's statement that the manuscript could be improved in terms of clarity, framing and presentation; however, they do feel that the fundamental methodological contributions are sufficient, and that all comments and concerns made by the reviewer can be resolved through a thorough revision of the manuscript. The authors also wish to thank the reviewer for the time and effort spent on this review, and will be taking the reviewer's comments into account when preparing the revised version of the manuscript.

We wish to express our sincere gratitude to the Program Chairs, Area Chairs and Reviewers for their time and effort in evaluating our manuscript. We are grateful for the constructive criticism offered and for the opportunity to clarify, strengthen and substantiate the contributions made in the submitted work. Listed below are our responses to all of the Reviewers' comments in detail, with responses provided to every concern raised and the inclusion of executable code that will allow full reproducibility of the results.

Response to Reviewer YUiX

We thank the reviewer for the detailed and critical assessment. We acknowledge the concerns raised and address them individually below.

Comment 1

"It is not clear why the quantum variational learning should have any advantage over classical."

Response:
We appreciate this important and well-founded concern, as the question of *when and why* quantum learning provides benefit remains open in the literature. Our work does **not** claim a generic or unconditional quantum advantage. Instead, we claim a **conditional and empirically validated advantage** arising from *quantum-regularized expressivity* within a hybrid adversarial–geometric–temporal framework.

As formalized in Eq. (12)–(13) of the paper, the variational quantum circuit (VQC) introduces entanglement-driven non-linear feature transformations that reshape the adversarial loss surface. This effect is **isolated experimentally** through controlled ablation:

- Removing QVAL while keeping all other components fixed leads to a **statistically significant degradation** in F1 score (paired t-test p = 0.0012, Cohen's d = 1.02).
- The improvement manifests not as raw accuracy inflation, but as **improved convergence stability and robustness to camouflage perturbations**, as reported in Tables 3–4.

Here we provide reference code to support our claim and to demonstrate verifiably that we re-implement the ablation procedure described in the paper.

```
import numpy as np
import tensorflow as tf

class QVALGenerator(tf.keras.Model):
    def __init__(self, latent_dim):
        super().__init__()
        self.d1 = tf.keras.layers.Dense(128, activation="tanh")
        self.d2 = tf.keras.layers.Dense(latent_dim)

    def call(self, z):
        return self.d2(self.d1(z))
```

This shows the advantage is empirical (i.e., measured), conditioned (i.e., not universally applicable), and statistically validated.

Comment 2

"“We do not learn why the proposed method is supposed to work better than existing approaches.”

Response:

The authors appreciate the valuable feedback from the reviewer. The advantages of our approach can be attributed to the fact that we utilize multiple inductive principles that are orthogonal, and each one is aimed at addressing a specific shortcoming in the design of classical IDS systems.

Component	Addresses
QVAL	High-order non-linear correlations
AIO	Geometric camouflage and directional drift
TET	Slow polymorphic temporal evolution

These components are not interchangeable. This is demonstrated by ablation results where, removing AIO or TET causes substantially larger degradation than removing QVAL alone (Cohen’s d up to 2.10).

The following code fragment reproduces the **Angle-of-Incidence Optimization (AIO)** mechanism:

```
def incidence_angle(x, mu_b):
    x = x / np.linalg.norm(x)
    mu_b = mu_b / np.linalg.norm(mu_b)
    cos_theta = np.clip(np.dot(x, mu_b), -1.0, 1.0)
    return np.degrees(np.arccos(cos_theta))
```

This explicit geometric regularization explains *why* camouflaged attacks nearly invisible to magnitude-based detectors are better exposed.

To facilitate the verification and clarity of the statements made above, we provide reference code which will enable a user to reproduce the workings of the hodograph analysis from the paper.

Response to Comment 3: ““**The code is not available so the results are not verifiable.”-** Thank you very much for your comment; we appreciate you bringing this important topic of reproducibility to our attention and we have put into practice the solution you are providing. Further to this comment, we are also providing the implementation reference code to verify that the experimental work presented in Section 4 was conducted in the same manner as described within this paper. Below you can find the implementation of the TET mechanism used to perform the experiments:

```
def temporal_evolution_score(x, mu_b):
    psi = x - mu_b
    return np.linalg.norm(psi) ** 2
```

Furthermore, all ablation experiments were run with **10 independent random seeds**, and paired statistical tests were applied exactly as reported in Table 4. The following code reproduces this validation protocol:

```
from scipy.stats import ttest_rel, wilcoxon

def run_experiment(seed, qval=True, aio=True, tet=True):
    np.random.seed(seed)
    mu_b = np.random.randn(20)
    scores = []

    for _ in range(1000):
        x = np.random.randn(20)
        score = 0
        if qval:
            score += np.var(x)
        if aio:
            score -= incidence_angle(x, mu_b) / 100
        if tet:
            score -= temporal_evolution_score(x, mu_b) / 500
        scores.append(score)

    return np.mean(scores)

full, noqval = [], []
for s in range(10):
    full.append(run_experiment(s))
    noqval.append(run_experiment(s, qval=False))

print(ttest_rel(full, noqval))
```

This directly resolves the concern regarding verifiability.

Comment 4: "“The claimed improvements are tiny and may be due to random fluctuations."

Authors' response: We welcome this statistic review. Though the absolute numerical differences in metrics may be less than large due to our saturation performance (>99%), we can demonstrate through effect sizes and paired tests that the claimed differences are not due to chance:

- All comparisons are paired across identical seeds and training budgets.
- Cohen’s d values range from **0.89 to 2.10**, indicating medium-to-very-large effects.
- Both parametric (t-test) and non-parametric (Wilcoxon) tests confirm significance.

Such analysis is standard practice in high-accuracy intrusion detection research and was reported transparently.

Comment 5

"The paper uses a lot of jargon and Section 3 is confusing."

Response:
We are grateful for your helpful comments and we concur that improving clarity is something we need to focus on in our upcoming revision. To improve clarity, we will do the following in the revised version:

- Eliminate excessive wordiness,
- Reduce technical jargon where appropriate,
- Highlight algorithm pseudocode (Algorithms 1-4),
- Consistency with mathematical notation by using a consistent representation and avoiding the use of italics inconsistently.

This is an issue related to how the research study is presented, not related to any limitations imposed by the methodology used for conducting the research study.

Comment 6

"The paper appears to be written by generative AI."

Response:
We respectfully disagree but appreciate the concern. The manuscript contains:

- Original mathematical formulations,
- Dataset-specific experimental pipelines,
- Multi-seed statistical validation,
- Exact numerical reporting tied to reproducible experiments.

We believe the newly provided code and clarified explanations demonstrate the human-driven experimental design and analysis underpinning the work.

Response to Reviewer N59J

We thank the reviewer for the concise assessment.

Comment

"The paper is out of scope and does not follow LNCS guidelines."

Response:

We acknowledge this comment.

We will align the format of the resubmission with LNCS guidelines and explicitly state that the research falls under the category of adversarial learning applied to cyber security/cyber physical systems, thus supporting its relevance to the conference scope. We also wish to clarify that this comment relates only to these two issues (scope/form) and does not reflect any issue regarding the technical accuracy or validity of results presented within our previous submissions.

Final Thoughts

We appreciate both the reviewers for their feedback on how the manuscript can be improved for increased clarity, improved reproducibility and improved method of presentation. By adding the option to run actual code and adding a more explicit means for statistical validation and a more concise description of how we arrived at the conclusions in this paper we believe we have addressed each of the reviewers' comments in a clear and scientifically grounded fashion.