Epitopea-QI: A Quantum–Inspired Convex–Nonconvex Framework for Learning the Class II Antigenic Landscape with Cytokine Policy Reinforcement

Francis Boabang, PhD†, Samuel Asante Gyamerah, PhD^
†Concordia Institute for Information and Systems Engineering (CIISE),
Concordia University, Montréal, QC, Canada
^Department of Mathematics, Toronto Metropolitan University, Toronto, Ontario, Canada
*Corresponding author: asante.gyamerah@torontomu.ca

***Abstract*—We introduce Epitopea-QI, a quantum–inspired computational immunology framework that extends the original Epitopea model by embedding a Hamiltonian energy regularizer into the Proximal Policy Optimization (PPO) cytokine controller. This addition penalizes high-energy immune policies and stabilizes learning through interference-based phase modulation. The model jointly learns peptide–TCR–MHC II recognition and cytokine control within a convex–nonconvex optimization landscape inspired by Waddington's developmental framework. A Graph Attention Network (ImmuneNet-GAT) encodes structural dependencies, while the PPO controller explores cytokine perturbation policies under a quantum-inspired energy constraint. This hybrid energy landscape enables smooth transitions between stable immune equilibria and adaptive exploration states, improving binding accuracy, cytokine diversity, and critic stability. The approach bridges reinforcement learning, quantum optimization, and systems immunology toward interpretable digital immune modeling.**

***Index Terms*—Quantum Machine Learning, Reinforcement Learning, Cytokine Policy, Convex Optimization, Immunoinformatics, Waddington Landscape**

I. Introduction

T cell recognition of peptide–MHC complexes is central to adaptive immunity. The *TCR–CD4–MHC II* axis initiates cytokine secretion and immune polarization [4]. Capturing both binding specificity and feedback adaptation remains a major challenge in computational immunology.

Previous works employ graph-based neural networks for MHC binding prediction [3], [1], [2]. However, these models are typically static and fail to incorporate dynamic cytokine feedback. Reinforcement learning (RL), especially Proximal Policy Optimization (PPO) [5], provides a biologically interpretable mechanism for adaptive cytokine regulation as a policy optimization task.

Epitopea-QI extends this approach by introducing a *Quantum-Inspired Hamiltonian Regularization* term into the PPO loss. This term acts as a soft energy constraint that aligns the immune policy with an underlying quantum Hamiltonian landscape, providing a physical analogy for immune adaptation across convex and nonconvex basins.

II. Methodology

A. ImmuneNet–QGE Representation

Each peptide–TCR–MHC II complex is represented as a dynamic graph $G = (V, E)$, where residues are nodes $v_i \in V$ and physical or energetic interactions define edges $(i, j) \in E$. Node embeddings are updated through stacked graph-attention transformer layers:

$$h_i^{(l+1)} = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}^{(l)} W^{(l)} h_j^{(l)}\right),$$

where $\alpha_{ij}^{(l)}$ are normalized attention coefficients and σ is a convex (ReLU, Softplus) or nonconvex (Swish, GELU) activation, depending on the optimization phase.

To capture higher-order entanglement between residues and allelic context, a quantum–variational coupling layer $\mathcal{Q}_\theta$ is applied on the hidden embeddings z from the ImmuneNet backbone:

$$\tilde{z} = \begin{cases} z + \mathcal{Q}_\theta(z), & \text{(residual coupling)} \\ (1-g)z + g\,\mathcal{Q}_\theta(z), & \text{(gated coupling)} \\ W_c[z, \mathcal{Q}_\theta(z)], & \text{(concatenative coupling).} \end{cases}$$

The variational block performs a cosine–sine feature map $\phi(x) = [\cos(Wx), \sin(Wx)]$ followed by orthogonally initialized linear transformations that mimic parametric quantum circuit layers. This Quantum Graph Encoder (QGE) enriches the biophysical latent space with interference-like superpositions analogous to variational quantum states.

B. Convex–Nonconvex Immune Landscape Dynamics

Immune differentiation is modeled as a two-phase hybrid optimization process, interpolating between convex and nonconvex response surfaces:

$$\phi(x, e) = \begin{cases} \text{Softplus}(x), & e < \tau_s, \\ x \cdot \sigma(x), & e \geq \tau_s. \end{cases}$$

Algorithm 1 Epitopea-QI: Quantum–Inspired PPO Distillation

Require: Teacher T, Student S, Data $\mathcal{D}$, Hamiltonian H, Epochs E

for $e = 1$ to E **do**
  Set activation mode (Convex, Two-Stage, or Nonconvex)
  for batch (x, y) in $\mathcal{D}$ **do**
    Compute teacher $(\mu_t, v_t) = T(x)$, student $(\mu_s, v_s) = S(x)$
    Compute $\mathcal{L}_{\text{PPO-Distill}}$ via KL and critic value losses
    Add Hamiltonian term: $\mathcal{L}_{QI} = \lambda_H \, \mathbb{E}[\mu_s^\top H \mu_s] + \lambda_\phi \sin^2(\phi(v_s - v_t))$
    Total loss: $\mathcal{L}_{Total} = \mathcal{L}_{\text{PPO-Distill}} + \mathcal{L}_{QI}$
    Update student $S \leftarrow S - \eta \nabla_S \mathcal{L}_{Total}$
  end for
end for

The epoch-dependent switching threshold τ_s defines a transition from early convex convergence (homeostatic regulation) to later nonconvex exploration (adaptive activation). This transition parallels biological shifts between naive, activated, and effector immune cell states. When coupled with the quantum–variational term, the effective immune landscape evolves under a composite potential:

$$\mathcal{L}_{\text{QI}} = \mathcal{L}_{\text{base}} + \lambda_E \langle \mu, H\mu \rangle + \lambda_P \sin^2(\omega(v - \hat{v})),$$

where the first term models conventional convex/nonconvex optimization, and the latter two correspond to the Hamiltonian energy and phase-interference penalties introduced by the quantum-inspired regulator. This hybrid structure encourages exploration of low-energy, high-stability basins analogous to epigenetic attractors in the Waddington immune landscape.

C. Quantum-Enhanced PPO Distillation

We define the total PPO–Distillation loss as:

$$\mathcal{L}_{\text{Total}} = \mathcal{L}_{\text{PPO-Distill}} + \mathcal{L}_{QI},$$

where $\mathcal{L}_{\text{PPO-Distill}}$ remains:

$$\mathcal{L}_{\text{PPO-Distill}} = \alpha \mathcal{L}_{\text{KL}} + (1 - \alpha)\mathcal{L}_V,$$

with convex, nonconvex, and two-stage formulations for $\mathcal{L}_V$. The Hamiltonian penalty adds a quantum-aware correction:

$$\mathcal{L}_V \leftarrow \mathcal{L}_V + \lambda_H \, \mathbb{E}[\psi^\top H \psi].$$

This quantum term reduces oscillatory divergence and improves convergence stability without eliminating biological plasticity.

III. Results

The Epitopea-QI system was evaluated using MHC-II recognition datasets. Experiments compared convex, nonconvex, and two-stage modes with and without the quantum regularizer.

These results indicate that quantum-inspired regularization provides a mechanistic analogy for cytokine interference, mirroring biological homeostasis through energy minimization.

IV. Discussion

The quantum Hamiltonian term acts as a soft constraint guiding the immune policy toward biologically meaningful attractors. It reflects the energetic balance of cytokine networks and supports the Waddington landscape analogy: convex basins correspond to stable tolerance states, while quantum interference facilitates tunneling between effector basins.

In future work, coupling quantum variational circuits with biological graph encoders could enable direct simulation of immune energy landscapes on quantum hardware.

V. Conclusion

Epitopea-QI merges convex–nonconvex optimization, PPO reinforcement learning, and quantum-inspired regularization into a unified framework for modeling immune adaptation. The Hamiltonian regularizer improves PPO stability, interprets immune dynamics in energy-space, and opens new directions for quantum biology-inspired machine learning.

References

[1] L. Huang and et al. Deepmhcii: a deep learning framework for peptide–mhc class ii binding prediction. *Bioinformatics*, 36(24):6183–6191, 2020.

[2] Vanessa Jurtz and et al. Netmhcpan-4.0: improved peptide–mhc class i interaction predictions integrating eluted ligand and peptide binding affinity data. *The Journal of Immunology*, 199(9):3360–3368, 2017.

[3] Alessio Montemurro, Philip C Stirling, and Enrico Radaelli. Tcr–peptide–mhc interactions: structural insights into antigen recognition and cross-reactivity. *Frontiers in Immunology*, 12:676575, 2021.

[4] Kenneth Murphy and Casey Weaver. *Janeway's Immunobiology*. Garland Science, 2016.

[5] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. In *arXiv preprint arXiv:1707.06347*, 2017.