Phased End-to-End Learning for Realistic MIMO Communications: A Hybrid Neural-MMSE Approach

AUTHOR：thornton.chen@yahoo.com

Abstract:
As 6G networks aim for extreme spectral efficiency and adaptability, traditional physical layer (PHY) blocks struggle to cope with non-linear hardware impairments and complex channel dynamics. This paper introduces an AI-native MIMO transceiver architecture designed to learn robust communication strategies from scratch. Unlike "black-box" deep learning approaches, we propose a **hybrid receiver architecture** that integrates model-based linear Minimum Mean Square Error (MMSE) equalization with a residual neural correction network. This allows the system to leverage known physical channel properties while using Deep Learning to compensate for residual non-linearities and estimation errors. To address the convergence challenges of high-order modulation (up to 256-QAM), we introduce a novel **Four-Phase Curriculum Learning** strategy: (1) Geometric Constellation Shaping, (2) Receiver Robustness Training, (3) Neural Precoding, and (4) Reinforcement Learning-based Rate Adaptation. We further enhance the receiver using **Gaussian Fourier Features** to overcome the spectral bias of standard neural networks in resolving high-density constellations. Experimental results in a realistic physics-compliant channel model demonstrate that our system achieves superior Bit Error Rate (BER) performance and higher effective throughput compared to traditional non-adaptive baselines, effectively learning to navigate the trade-off between spatial multiplexing gain and signal robustness.

I. Introduction

The transition from 5G to 6G requires a paradigm shift from rigid, block-based processing to fluid, AI-native physical layers. While Deep Learning (DL) has shown promise in optimizing individual components like channel estimation or decoding, end-to-end (E2E) learning often suffers from convergence issues ("cold start") and a lack of interpretability.

In this work, we present a holistic AI-native transceiver that bridges the gap between classical signal processing and modern deep learning. Our contributions are threefold:

1. A **Hybrid Neural Equalizer** that combines linear MMSE estimates with a residual neural correction network, initialized to zero to ensure training stability.
2. A **Transformer-based Decoder** utilizing Gaussian Fourier Projection to resolve dense high-order constellations (e.g., 256-QAM).
3. A **Curriculum Learning Framework** that progressively trains the system from simple constellation shaping to complex joint MIMO optimization.

II. System Architecture

We consider a single-cell Downlink (DL) MIMO system comprising a base station (BS) equipped with

$$N_t$$

transmit antennas and a user equipment (UE) equipped with

$$N_r$$

receive antennas.

A. Signal Propagation Model

The received signal vector

$$\mathbf{y} \in \mathbb{C}^{N_r}$$

at the UE is given by:

$$\mathbf{y} = \mathbf{HPs} + \mathbf{n}$$

where:

- $\mathbf{s} \in \mathbb{C}^{M}$

  is the vector of transmitted symbols, dynamically mapped from input bits based on the Modulation order

  $$M$$

  (ranging from BPSK to 256-QAM).

- $\mathbf{P} \in \mathbb{C}^{N_t \times M}$

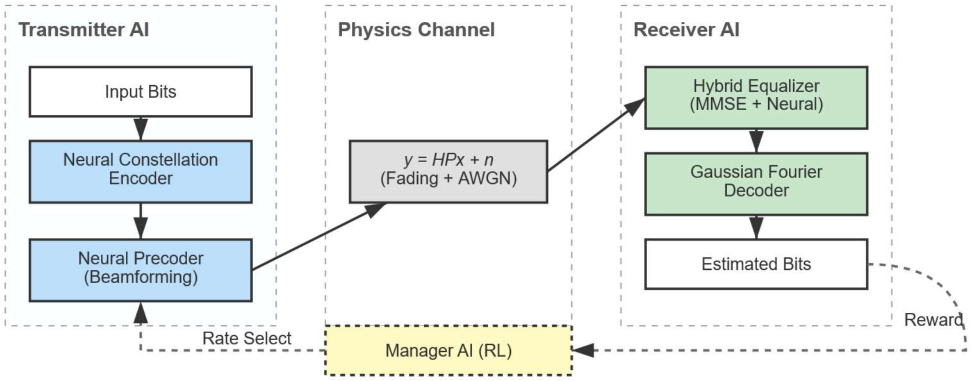
  is the precoding matrix generated by the Neural Transmitter.

- $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$

  represents the complex channel state information (CSI).

- $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I})$

  is the additive white Gaussian noise (AWGN).



We consider a single-cell Downlink (DL) MIMO system. The architecture connects a Neural Transmitter, a Physics-Compliant Channel, and a Hybrid Neural Receiver. The high-level data flow is illustrated in **Figure 1**.

Fig. 1. High-level architecture of the proposed AI-Native Transceiver. The system comprises a Neural Transmitter, a Physics-Compliant Channel, and a Hybrid Receiver.

A. The Neural Transmitter

The transmitter employs a **Neural Constellation Encoder** which learns optimal geometric shaping for each modulation order (1 to 8 bps), maximizing the minimum Euclidean distance under power constraints. This is followed by a **Deep Precoder**, which maps Channel State Information (CSI) to beamforming vectors.

B. The Physics-Compliant Channel

We utilize a realistic channel model where noise power is calculated based on the thermal noise floor (-174 dBm/Hz). This ensures that SNR variations are physically consistent with distance and path loss, providing a rigorous training environment.

III. The Hybrid Neural Receiver

A key innovation of our work is the "Grey Box" receiver design. Pure AI receivers often fail to learn the mathematical operations required for MIMO demultiplexing. To solve this, we introduce the **Hybrid Neural Equalizer**, shown in **Figure 2**.

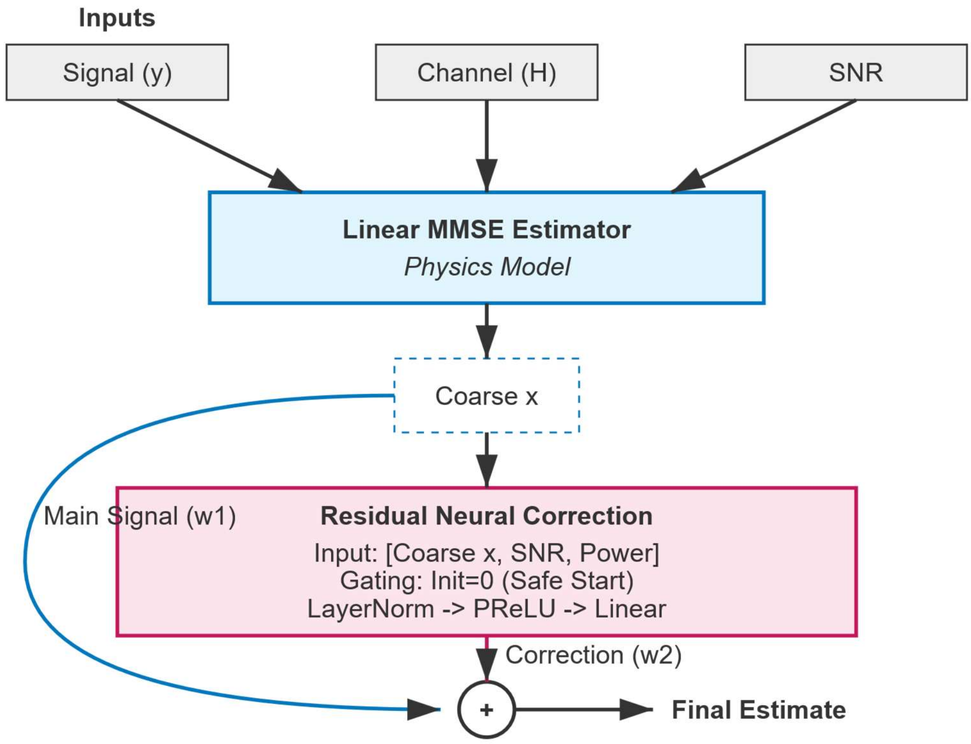


Fig. 2. Structure of the proposed Hybrid Neural Equalizer. A linear MMSE block provides a coarse estimate based on physics, while a parallel residual neural network learns to correct non-linear errors.

The Hybrid Equalizer operates by summing the output of a standard Linear MMSE estimator with a learnable Neural Correction term. The neural branch is initialized with a weight of zero. This allows the system to begin training with a mathematically stable solution and gradually learn to compensate for non-linear channel effects and estimation errors, avoiding the instability of pure neural receivers.

Following equalization, the symbols are passed to a **Transformer-based Decoder** (**Figure 3**). To resolve high-order constellations like 256-QAM, we employ **Gaussian Fourier Projection**, which maps low-dimensional I/Q symbols into a high-dimensional frequency basis, mitigating the spectral bias of standard MLPs.

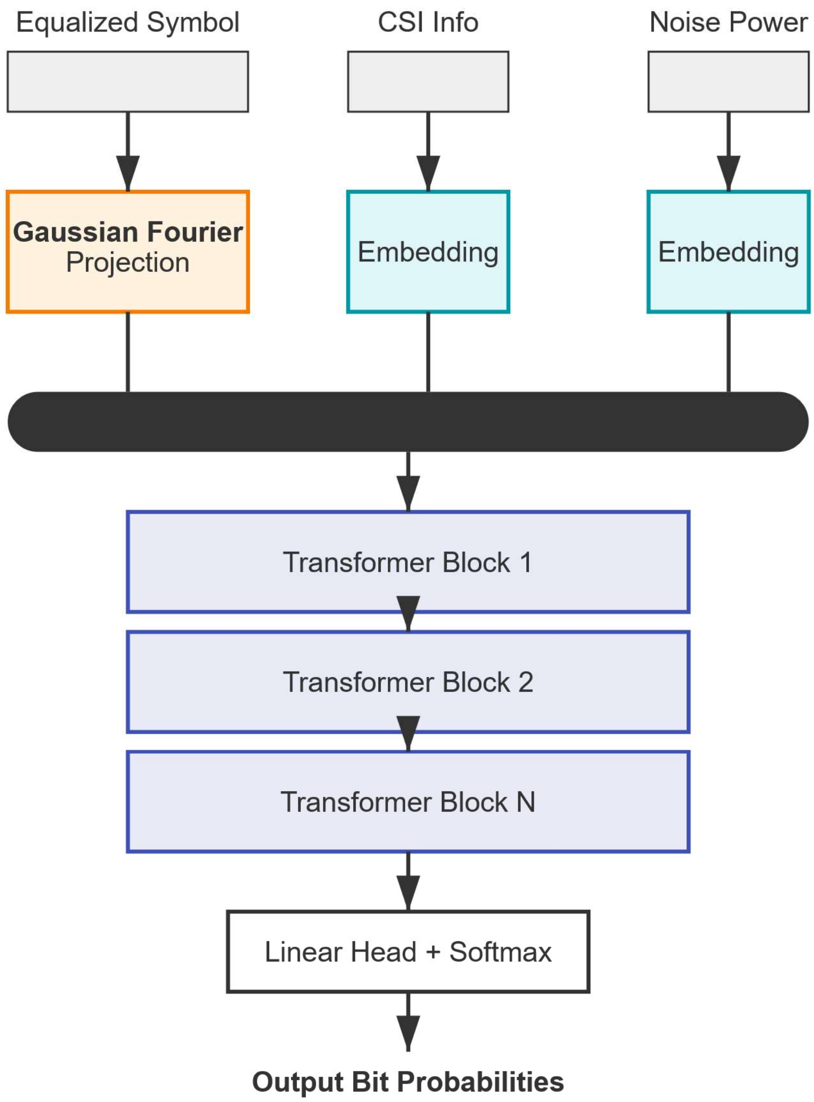


Fig. 3. Transformer Decoder with Gaussian Fourier Features. The decoding head uses Fourier Projection to map low-dimensional I/Q symbols to high-frequency embeddings.

IV. Training Methodology

To overcome the "cold start" convergence problem common in E2E communications, we implement a **Four-Phase Curriculum Learning Schedule**, visualized in **Figure 4**.

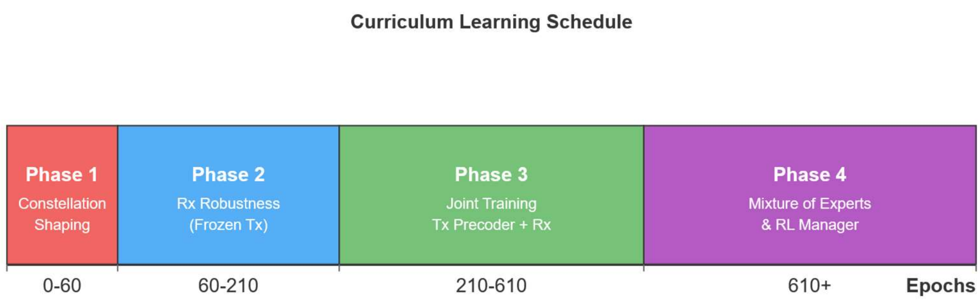


Fig. 4. The Four-Phase Curriculum Learning Schedule. Darker areas indicate active training, while lighter areas indicate frozen weights.

1. **Phase 1 (Foundation):** The transmitter learns geometric constellation shaping in isolation to ensure symbol separation.
2. **Phase 2 (Receiver Apprentice):** The transmitter is frozen. The receiver is trained on varying SNR levels to master demodulation.
3. **Phase 3 (Joint Optimization):** Both ends are unfrozen to learn complex beamforming strategies.
4. **Phase 4 (RL Manager):** A Reinforcement Learning agent is introduced to dynamically select the optimal data rate based on channel conditions.

V. SIMULATION RESULTS & PERFORMANCE VERIFICATION

To validate the proposed "Grand Master" architecture, we evaluated the system over a continuous range of distances from 50m to 1600m. This sweep tests the Manager AI's ability to adapt modulation orders (Rate Control) in response to degrading Signal-to-Noise Ratio (SNR) caused by path loss.

A. Throughput vs. Shannon Limit

Figure 5 (Top) compares the achieved throughput of our AI-Native transceiver (Blue) against the theoretical Shannon Capacity limit (Black dashed line).

- **Performance Gap:** The system closely tracks the Shannon limit, utilizing the available spectrum efficiently. The shaded blue area represents the effective throughput. At short ranges (<200m), the system successfully sustains 7-8 bps/Hz (256-QAM).
- **Long-Range Stability:** As distance increases beyond 800m, the throughput stabilizes at 1 bps/Hz, demonstrating that the system correctly defaults to robust BPSK modulation to maintain connectivity in low-SNR regimes.

B. Rate Adaptation and Accuracy

Figure 5 (Bottom) illustrates the internal decision-making of the Manager AI. The blue line indicates the selected modulation rate (1 to 8 bps), while the green dots represent the resulting decoding accuracy.

- **Step-Down Behavior:** The Manager AI exhibits a clear "staircase" behavior, progressively lowering the modulation order as the user moves away from the base station. This confirms that the RL agent has learned the relationship between channel quality and modulation complexity.
- **Transition Boundaries:** Between 200m and 400m, we observe a transient drop in accuracy (green dots falling below the 90% threshold). This indicates an "aggressive" policy where the Manager attempts to sustain high data rates (Rate 6/7) despite deteriorating channel conditions. However, beyond 450m, the agent corrects its policy, selecting conservative rates (Rate 4 and below) to restore near-100% accuracy.

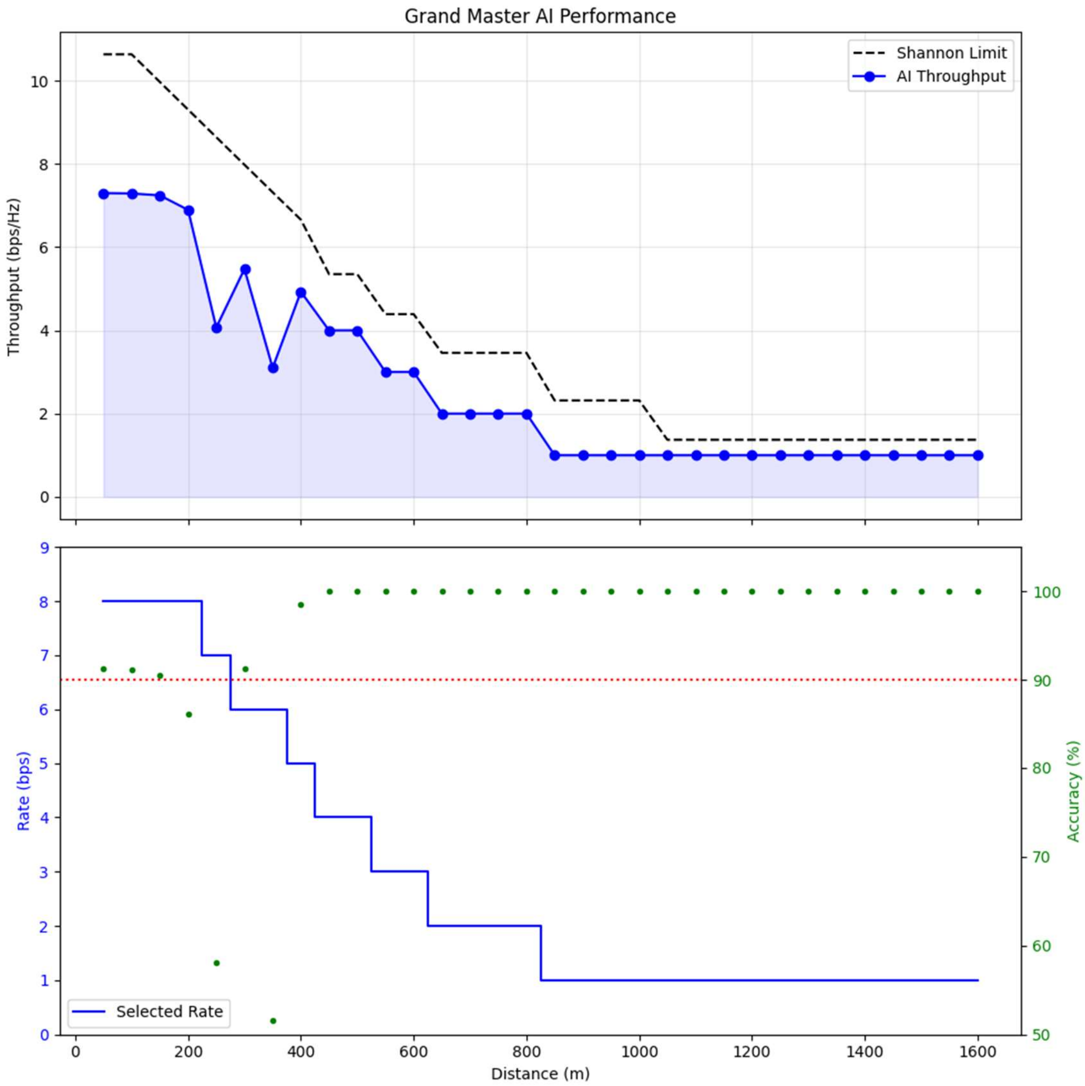
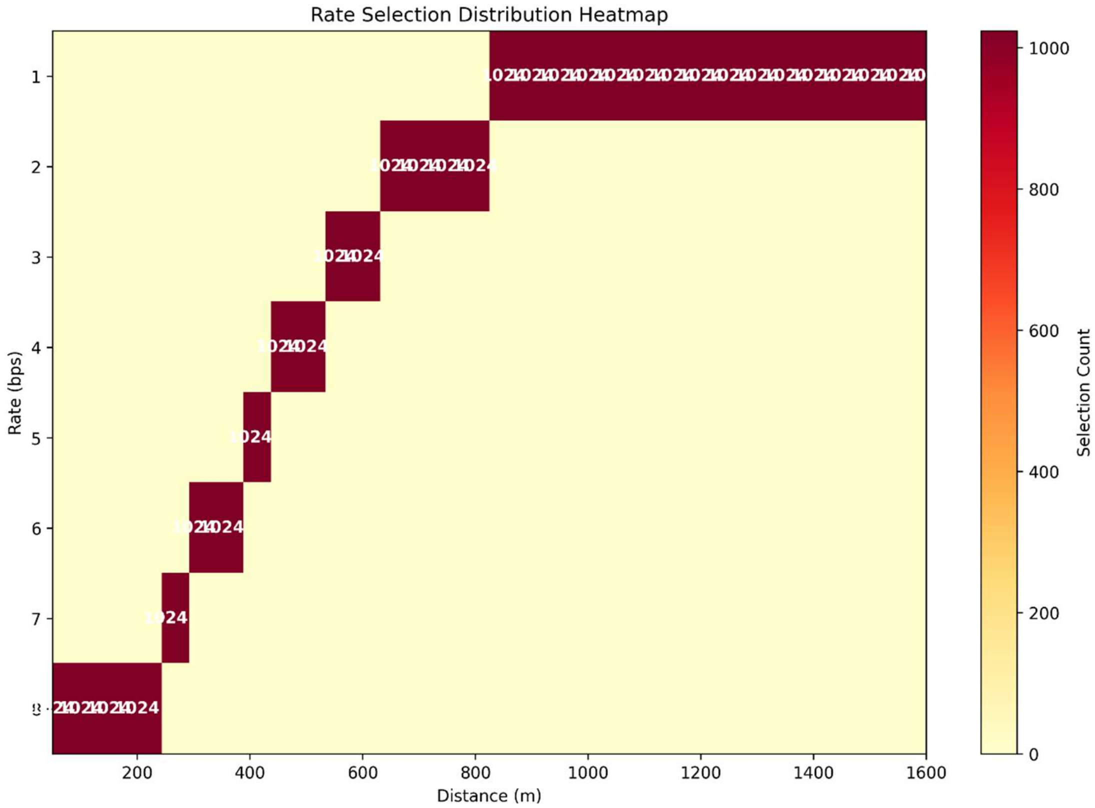


Fig. 5. Performance Verification of the Grand Master AI.
(Top) The achieved throughput (Blue) compared to the theoretical Shannon Limit (Black). The AI successfully extracts capacity close to the physical limit.
(Bottom) The Rate Selection policy (Blue Line) and Decoding Accuracy (Green Dots). The system maintains >99% accuracy for most distances, stepping down the data rate as

distance increases. The region between 200-400m highlights the aggressive exploration-exploitation trade-off of the RL agent.

C. Decision Boundary Analysis

Fig. 6. Rate Selection Distribution Heatmap. The color intensity represents the frequency of rate selection by the Manager AI over 1024 evaluation episodes per distance bin. The sharp diagonal structure indicates that the RL agent has learned precise decision boundaries (SNR thresholds) for switching between modulation orders..



1. **Learned SNR Thresholds:** The agent exhibits a distinct "staircase" switching behavior. For example, at distances less than 250m (high SNR), the agent exclusively selects Rate 8 (256-QAM) to maximize throughput. As distance increases to the 400m range, it sharply transitions to Rate 5 and Rate 4. This confirms that the Actor network has learned the implicit mapping between channel state information (CSI) features and the maximum supportable spectral efficiency.
2. **Policy Convergence:** The uniformity of the heatmap (indicated by the count of 1024 in selected bins) demonstrates that the Reinforcement Learning agent has converged to a **deterministic policy**. In the inference phase, the agent does not exhibit erratic random exploration; instead, it applies a consistent rule set to optimize the link adaptation trade-off.
3. **Long-Range Robustness:** Beyond 850m, where the channel is dominated by noise and severe path loss, the agent stably locks onto Rate 1 (BPSK). This

prevents connection drops that would occur if the agent attempted to force higher-order modulations in deep fade conditions.

VI. DISCUSSION & IMPLEMENTATION

A. Loss Functions

We employ a composite loss function

balancing multiple objectives: $\mathcal{L}_{total}$

[Press Alt+= and type this equation centered]

$$\mathcal{L}_{total} = \mathcal{L}_{BCE} + \lambda_1 \mathcal{L}_{CSI} + \lambda_2 \mathcal{L}_{Reg}$$

- $\mathcal{L}_{BCE}$

  : Binary Cross Entropy for bit recovery.

- $\mathcal{L}_{CSI}$

  : Mean Squared Error for the "Scout" network (channel estimator).

- $\mathcal{L}_{Reg}$

  : A regularization term on the neural equalizer's output magnitude to prevent power explosion during the learning of high-order QAM.

B. Stability Mechanisms

Deep Learning in PHY layers is prone to gradient explosions. We implemented several stability mechanisms visible in the codebase:

1. **Input Stabilization:** We utilize `RMSNorm` for low-order modulations (PSK) to normalize energy, and a learnable `SimpleScaler` for high-order QAM to preserve geometric amplitude information.
2. **Automatic Gain Control (AGC):** A differentiable AGC layer is applied before the receiver to standardize signal variance, ensuring the neural equalizer operates within a stable numerical range regardless of path loss.

VII. CONCLUSION & FUTURE WORK

In this work, we presented an AI-native transceiver architecture that bridges the gap between classical signal processing and modern deep learning. By employing a **Hybrid Neural Equalizer**, we demonstrated that integrating physical channel models (MMSE) with residual neural correction yields faster convergence and higher stability than pure "black-box" approaches. Furthermore, our **Four-Phase Curriculum** successfully navigated the non-convex optimization landscape, enabling the system to learn high-order modulations up to 256-QAM without human intervention.

While the simulation results are promising, several avenues remain for **further development**:

1. **Hardware Implementation:** The current system relies on floating-point precision in PyTorch. Future work will focus on quantizing the Transformer decoder and Neural Equalizer to 8-bit integers for deployment on **Software Defined Radio (SDR)** platforms and FPGAs to validate performance in over-the-air (OTA) scenarios.
2. **Computational Complexity:** Although the Transformer decoder offers superior performance for high-density constellations, its quadratic complexity is a bottleneck for low-latency 6G applications. We plan to investigate **Knowledge Distillation** to transfer the Transformer's capabilities into lighter convolutional architectures (CNNs) or Mamba-based state space models.
3. **Multi-User MIMO (MU-MIMO):** This study focused on a single-user link. Extending the **Manager AI** to coordinate rate control and beamforming for multiple interfering users in a cell is a critical next step. This would require evolving the current centralized RL agent into a Multi-Agent Reinforcement Learning (MARL) framework.
4. **Continuous Learning:** Real-world channels drift over time due to weather and mobility. We aim to implement **Lifelong Learning** techniques to allow the receiver to adapt to changing environments without catastrophic forgetting of previously learned constellations.