

Training Language Models and Embeddings for Hybrid Classical/Quantum Computing

Damir Cavar, Peter Sullivan, Shane Sparks, Sam Gray Miller, Ronit Jha, Ayomide Jeje, Yanin Charoenpornasawat

NLP-Lab, Indiana University, Bloomington, IN 47405, USA
Luddy School, Indiana University, Bloomington, IN 47405, USA

Introduction

- Real-valued transformers rely on high dimensionality for expressiveness, motivating more compact complex-valued architectures
- Complex numbers enable magnitude–phase decomposition, encoding multiple features within a single dimension
- Phase-based representations support dimension-independent encoding, reducing reliance on large hidden spaces
- Compact complex representations lower computational cost while preserving semantic capacity
- Native compatibility with unitary operations enables classical–quantum hybridization, supporting quantum-aligned NLP models

Objective

Enable Hybrid Classical-Quantum AI Model Inference using classical or quantum hardware without any data or model conversion.

Methods

- Convert BERT to Complex-Valued BERT
- Build a reusable complex tensor library (PyTorch/JAX) with:
 - Complex linear, layer norm, dropout
 - Complex attention (Q, K, V, softmax surrogate)
 - Complex activations (modReLU, zReLU, CReLU, phase-gated variants)
- Training:
 - English Wikipedia, Bookcorpus

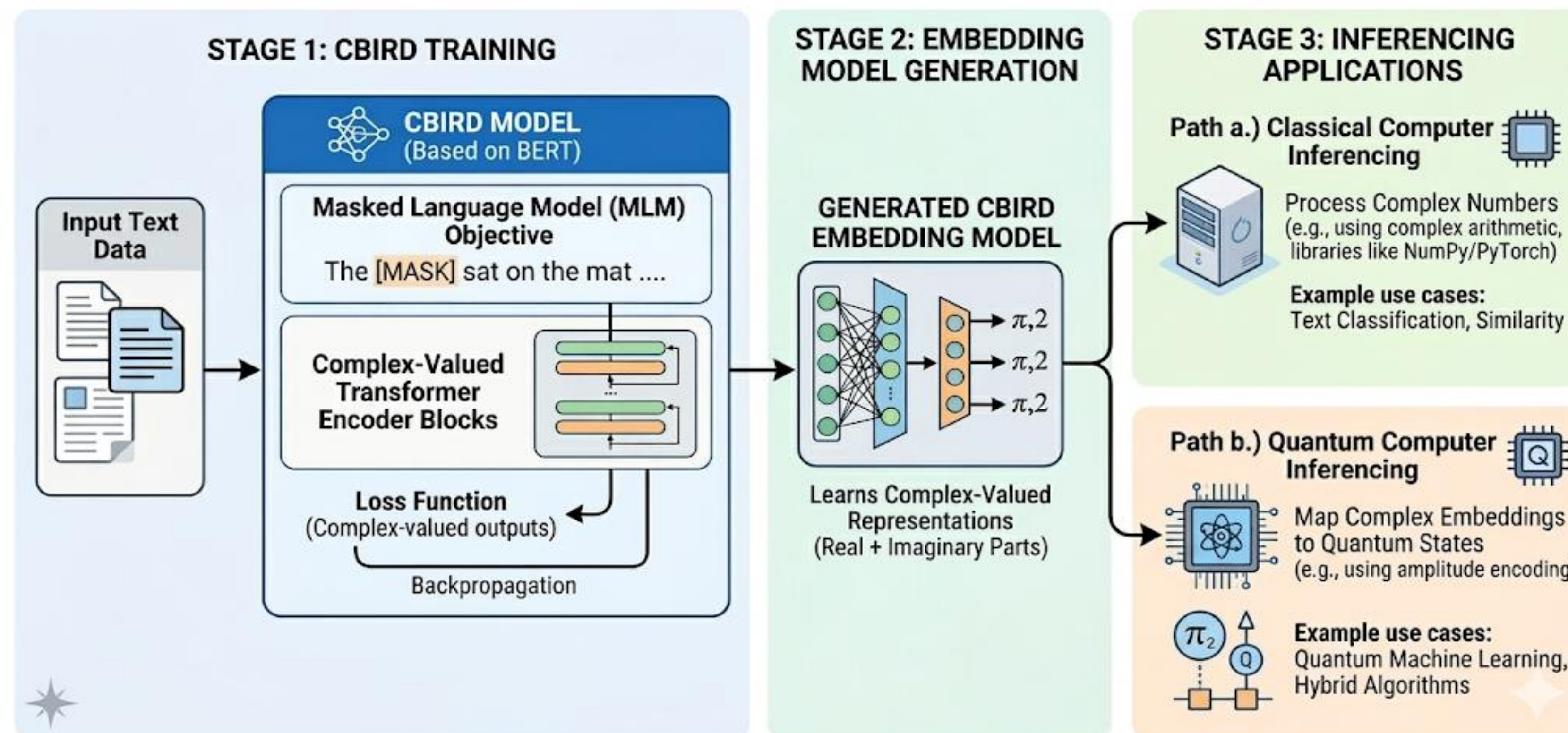
Results

- Core building blocks for complex-valued tensor models
 - Backpropagation (Wirtinger/complex derivatives, split real–imag, Cauchy–Riemann–aware variants)
 - Activation functions (modReLU, CReLU, phase-gated)
 - Architectures (Complex-valued CNNs, RNNs attention blocks)
- Doubling size of classical representations, not for quantum due to complex-valued nature of QC

Discussion and Future Work

- Many application domains where complex models are competitive or superior
 - Signal and communication
 - Speech and audio
 - Physics-aware tasks and imaging
- World Models

CBIRD: Complex Bidirectional Inducer for Representation Dynamics - Diagram



References

- Devlin et al. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Association for Computational Linguistics.
- Cavar et al. (2025) Old Wine in New Bottles: Using Classical Word Embeddings in Gate-Based Quantum NLP Systems. Springer Nature.
- Arjovsky, M., Shah, A., & Bengio, Y. (2015). Unitary Evolution Recurrent Neural Networks.

Acknowledgements

- Daniel Quigley
- Soren DeHaan
- Center for Quantum Technologies (CQT) initial grant funding