**Proposal: Enhancing the Boltz Diffusion Framework with Moment-Tensor Features for Physically Accurate and GPU-Efficient Binding Predictions**

Team QURE

**1. Motivation**

Current Boltz-2 type diffusion models achieve remarkable throughput in molecular binding prediction and ligand placement, yet they remain largely statistical in their geometric reasoning. They capture pairwise correlations through SE(3)-invariant embeddings but often miss higher-order angular and many-body dependencies essential for precise energy reconstruction.

To reach parity with physically motivated interatomic potentials such as the Moment Tensor Potential (MTP) family, we propose integrating tensorial descriptors that encode multi-body geometry directly into the Boltz-2 framework. This addition will (*1) improve physical fidelity, (2) provide interpretable geometric–energetic relationships, and (3) preserve Boltz’s signature GPU scalability.*

This hybridization aligns Boltz-2 with the modern trend toward physics-aware AI: models that retain the efficiency of deep diffusion networks while embedding analytic representations of molecular geometry, as in MTP and NequIP. The resulting model—Boltz-MTP—can serve as both a GPU-efficient learning platform and a fast energy surrogate competitive with neural potentials used in materials simulation.

**2. Scientific Rationale**

The MTP formalism expresses the total energy as a linear combination of local tensor contractions:

Each site’s energy contribution is linear in coefficients but nonlinear in geometry via the analytic tensor basis. This yields linear scaling with system size (O(Nk)) while retaining full multi-body sensitivity through the chosen tensor ranks.

By concatenating these moment-tensor invariants to Boltz atom embeddings, we inject explicit local structure information—effectively a low-cost, differentiable measure of curvature in the energy landscape. The diffusion model then learns on a physically meaningful latent manifold, improving extrapolation to unseen geometries and ligand classes.

In physical terms, the tensorial features provide a surrogate to the local Hessian of the potential energy surface. Their inclusion should increase the sensitivity of Boltz energy changes (ΔE) to geometric perturbations, enabling the model to reproduce the correct stiffness of molecular interactions without deeper neural layers or expensive self-attention.

**3. Methodology and Implementation Plan**

Data generation via controlled “wiggles.”

Generate small displacements (rigid + torsional) around reference docked poses. For each, record Boltz-predicted and AutoDock Vina (score-only) energies, defining the target ΔE.

These small perturbations create a local energy-geometry map for regression testing.

Tensorial feature construction.

Compute per-atom MTP descriptors 𝑀0,1, 𝑀0,2, traces, and norms within a 6 Å contact window. Pool to a 12–24 D invariant vector per complex. Concatenate Δb = b(wiggle) – b(ref) to existing Boltz embeddings.

Hybrid energy head.

Train a light ridge or MLP energy head on Δb while freezing diffusion layers. Evaluate on held-out complexes for 𝑅2, MAE, Spearman ρ, and sign accuracy.

Efficiency benchmarking.

Measure GPU memory and wall time per batch with and without tensorials. Expected overhead < 5%, since tensorial preprocessing is CPU-side and scales linearly.

Comparative metrics.

Success = ≥ 0.10 increase in R² or ≥ 10% drop in MAE vs. baseline, negligible GPU cost, and improved correlation of ΔE with Vina energies.

**4. Expected Outcomes and Competitive Edge**

* **Physics-grounded learning:** Boltz learns energy curvature rather than only positional likelihoods, enabling smoother and more transferable potentials.
* **Linear-scaling GPU profile:** Unlike equivariant attention or higher-order message passing, tensorial preprocessing keeps inference cost linear in system size.
* **Enhanced interpretability:** Tensor invariants correspond to physically intuitive quantities—neighbor counts, dipoles, quadrupoles, and angular anisotropy—enabling explainable predictions.
* **Benchmark superiority:** The hybrid model can bridge the gap between neural docking scorers and ML interatomic potentials, potentially outperforming baseline Boltz and NequIP-style networks on small-geometry sensitivity tasks.

**5. Why these matter?**

This project positions Boltz-2 as a **next-generation GPU-ready diffusion model** that competes on both scientific rigor and computational speed. Immediate next steps:

1. Implement feature concatenation and ridge testing pipeline (proof-of-concept).
2. Validate against Vina or DFT-like energies for 20–30 ligand–receptor pairs.
3. Release a benchmark report highlighting accuracy-vs-efficiency gains.
4. Integrate into the Boltz-2 GitHub as an *optional “tensorial energy head.”*

If successful, Boltz-2 would represent the first diffusion-based docking model to achieve **explicit tensorial geometric sensitivity with linear GPU scaling**, combining the strengths of deep learning and analytic potentials—a compelling advantage in any competitive evaluation.

**References:**

A.V. Shapeev, “Moment Tensor Potentials: a class of systematically

improvable interatomic potentials,” <https://arxiv.org/pdf/1512.06054>

A typical parity plot for energy prediction using MTP for a physical problem (R2> 0.99):

