LatticeZero: Physics-Derived Scoring for Protein-Ligand Docking

First-Principles Dispersion with Calibrated Sterics and Pocket-Aware Classification

RFT Cosmology Research Group
contact@latticezero.com
https://latticezero.com

December 2025
Version 1.0 (Draft)

Abstract

Empirical scoring functions for protein-ligand docking suffer from over-parameterization, poor transferability across target classes, and sensitivity to minor steric clashes in rigid-receptor poses. We present **LatticeZero**, a geometric bypass scoring engine that replaces empirical terms with physics-derived components: (1) dispersion energies from Resonant Field Theory lattice calculations yielding first-principles C_6 coefficients, (2) steric repulsion with a docking-calibrated scale factor s_R derived from potential matching rather than AUC optimization, and (3) optional HBQ mode adding directional hydrogen bonding and Coulomb electrostatics. Benchmarked on a 6-target DUD-E subset spanning kinases, proteases, and GPCRs, LatticeZero achieves **0.70 AUC on AA2AR** using physics alone—competitive with empirical methods. We introduce the **ΔClash pocket descriptor**, which predicts optimal steric calibration from pocket geometry and classifies binding sites into four regimes (tight, neutral, discriminating, broken). Rust-accelerated kernels deliver 117× speedup for van der Waals scoring. LatticeZero provides an interpretable physics baseline for downstream machine learning and a principled framework for pocket-aware scoring.

Keywords: molecular docking, scoring function, dispersion, van der Waals, pocket geometry, protein-ligand binding

Contents

1 Introduction

1.1 The Docking Scoring Problem

Molecular docking is a cornerstone of computational drug discovery, enabling virtual screening of large compound libraries against protein targets [Kitchen et al., 2004]. The process involves two distinct challenges: *pose prediction* (finding plausible binding geometries) and *scoring* (ranking poses and compounds by predicted affinity). While sampling algorithms have matured considerably, scoring functions remain a fundamental bottleneck [Li et al., 2019].

The disconnect between docking scores and experimental binding affinities is well-documented. Correlations between predicted and measured ΔG values rarely exceed $r = 0.5$ even for well-behaved congeneric series [Wang et al., 2015]. More critically for virtual screening, enrichment of active compounds over decoys varies dramatically across target classes, with methods that excel on kinases often failing on GPCRs or proteases [Mysinger et al., 2012].

1.2 Limitations of Empirical Scoring Functions

Contemporary scoring functions such as AutoDock Vina [Trott & Olson, 2010], Glide [Friesner et al., 2004], and GOLD [Verdonk et al., 2003] rely on empirical parameterization against experimental binding data. While pragmatically successful, this approach introduces several fundamental limitations:

Over-parameterization. Empirical functions employ 10–50+ adjustable parameters fit to historical binding affinity datasets. These parameters encode statistical correlations rather than physical principles, leading to unstable extrapolation beyond the training domain. The PDBbind refined set [Wang et al., 2004], while valuable, over-represents certain target classes (kinases, proteases) and under-represents others (GPCRs, ion channels).

Decoy bias. Benchmark datasets like DUD-E [Mysinger et al., 2012] construct decoys with matched physicochemical properties but distinct topologies. However, empirical functions trained on such benchmarks may learn to exploit subtle distributional artifacts rather than genuine binding physics. Reported AUC values of 0.8–0.9 often fail to replicate in prospective screens.

Steric clash sensitivity. Rigid-receptor docking protocols from engines like Smina [Koes et al., 2013] produce poses with minor atom-atom overlaps ($\sim$0.2–0.5 Å). Empirical functions with steep repulsive walls severely penalize such clashes, often assigning infinite or near-infinite energies that obscure favorable dispersion and electrostatic contacts.

Target class inhomogeneity. Different binding pockets exhibit different physics. Kinase ATP-binding sites require precise hydrogen bond geometry to the hinge region. GPCR transmembrane pockets are dominated by hydrophobic packing. Protease catalytic sites demand electrostatic complementarity to the oxyanion hole. A single set of empirical weights cannot optimize for all regimes without pocket-aware calibration.

1.3 Our Contribution

We present **LatticeZero**, a scoring engine that addresses these limitations through first-principles physics rather than empirical fitting. Our contributions are:

1. **Physics-derived dispersion.** Van der Waals attraction computed from Resonant Field Theory (RFT) lattice calculations, yielding atom-type-specific C_6 coefficients without empirical parameterization (Section 2.1).

2. **Calibrated sterics.** A steric scale factor s_R derived from potential matching between LatticeZero's CCSD(T)-calibrated repulsion and Smina's soft-sphere model, providing principled clash tolerance without AUC-based tuning (Section 2.2).

3. **HBQ scoring mode.** Optional hydrogen bond geometry and Coulomb electrostatics for polar pockets, with Rust-accelerated kernels achieving 2300× speedup (Section 2.3).
4. **ΔClash pocket descriptor.** A novel geometric feature that predicts optimal steric calibration and classifies pockets into four regimes, enabling pocket-aware scoring without per-target fitting (Section 2.4).
5. **Competitive physics-only performance.** On a 6-target DUD-E benchmark, LatticeZero achieves 0.70 AUC on AA2AR using pure physics—demonstrating that principled scoring can compete with empirical methods on favorable targets.

LatticeZero is designed as a *rescoring* engine: it accepts poses from any docking program (Smina, Vina, Glide) and re-evaluates them with physics-derived energetics. This modular architecture separates pose generation from scoring, enabling systematic comparison of scoring approaches on identical pose ensembles.

2 Methods

2.1 RFT Dispersion Model

The attractive component of the van der Waals interaction arises from correlated electron fluctuations between non-bonded atoms. In LatticeZero, we compute dispersion energies using C_6 coefficients derived from Resonant Field Theory (RFT) lattice calculations rather than empirical fitting.

2.1.1 Theoretical Foundation

The London dispersion energy between atoms i and j at separation r_{ij} takes the familiar form:

$$E_{\text{disp}} = -\sum_{i\in\text{ligand}}\sum_{j\in\text{protein}}\frac{C_6^{ij}}{r_{ij}^6} \tag{1}$$

The C_6^{ij} coefficient depends on the dynamic polarizabilities $\alpha_i(\omega)$ and $\alpha_j(\omega)$ of the interacting atoms:

$$C_6^{ij} = \frac{3}{\pi}\int_0^\infty \alpha_i(i\omega)\alpha_j(i\omega)\,d\omega \tag{2}$$

In RFT, atomic polarizabilities emerge from the lattice structure of electron correlation rather than being fit to experimental data. This provides a principled, transferable basis for dispersion calculations across atom types.

2.1.2 Damping Function

At short range, Equation 1 diverges as $r \to 0$. We apply a Fermi-type damping function to regularize this behavior:

$$f_{\text{damp}}(r_{ij}) = \frac{1}{1+e^{-\beta(r_{ij}-r_0^{ij})}} \tag{3}$$

where $r_0^{ij} = s_R\cdot(\sigma_i+\sigma_j)$ defines the onset of damping in terms of atomic radii σ_i, σ_j and a global scale factor s_R. The steepness parameter $\beta = 12$ Å^{-1} ensures rapid but smooth transition.

The damped dispersion energy becomes:

$$E_{\text{disp}} = -\sum_{i,j}\frac{C_6^{ij}}{r_{ij}^6}\cdot f_{\text{damp}}(r_{ij}) \tag{4}$$

2.1.3 Atom Type Classification

We assign atoms to hybridization-aware types following standard SYBYL conventions: C.3 (sp^3 carbon), C.2 (sp^2 carbon), C.ar (aromatic carbon), N.3, N.2, N.ar, O.3, O.2, S.3, P.3, and halogens (F, Cl, Br, I). Each type receives a pre-computed C_6 coefficient from RFT lattice calculations. Hydrogen atoms are treated implicitly through the polarizability of their heavy-atom parents.

2.2 Calibrated Steric Repulsion

Steric repulsion prevents atoms from overlapping. We use a standard r^{-12} repulsive potential:

$$E_{\text{rep}} = \sum_{i,j} k_{\text{rep}} \left(\frac{s_R \cdot \sigma_{ij}}{r_{ij}} \right)^{12} \tag{5}$$

where $\sigma_{ij} = (\sigma_i + \sigma_j)/2$ is the combined atomic radius, k_{rep} sets the energy scale, and s_R is the critical steric scale factor.

2.2.1 The Steric Calibration Problem

Atomic radii σ_i in LatticeZero are calibrated against CCSD(T) calculations on noble gas dimers (He_2, Ne_2, Ar_2, Kr_2), achieving 0.25 kcal/mol mean absolute error on equilibrium well depths. However, this *ab initio* calibration assumes perfect geometries.

Docking poses from Smina and similar programs contain minor steric clashes because these engines use *soft-sphere* potentials during optimization. Smina's internal scoring allows partial atom overlap with finite penalty, enabling poses that would be catastrophically penalized by LatticeZero's strict CCSD(T)-calibrated radii.

2.2.2 Soft-Sphere Compatibility Coefficient

Rather than tuning s_R to maximize benchmark AUC (which would constitute empirical fitting), we derive s_R from *potential matching*: minimizing the discrepancy between LatticeZero's repulsive potential $V_{\text{LZ}}(r; s_R)$ and Smina's $V_{\text{Smina}}(r)$ over typical contact distances.

Let $V_{\text{LZ}}(r; s_R)$ denote LatticeZero's repulsive potential with scale factor s_R, and let $V_{\text{Smina}}(r)$ denote Smina's soft-sphere potential. We seek s_R^* that minimizes:

$$s_R^* = \arg\min_{s_R} \int_{r_{\min}}^{r_{\max}} \left[V_{\text{LZ}}(r; s_R) - V_{\text{Smina}}(r) \right]^2 dr \tag{6}$$

where $r_{\min} = 0.8 \cdot r_{\text{contact}}$ and $r_{\max} = 1.2 \cdot r_{\text{contact}}$ bracket the typical contact region.

This analysis yields $s_R^{\text{Smina}} \approx 0.50$, which we adopt as the default for rescoring Smina-generated poses. Importantly, this is a *pose-engine-specific* constant derived from physical potential matching, not a per-target or per-benchmark fit.

2.2.3 Empirical Validation

As an independent check, we swept s_R from 0.45 to 0.60 across all benchmark targets and observed that $s_R = 0.50$ indeed maximizes mean AUC while maintaining reasonable clash rates (Table **??**). The concordance between the potential-matching derivation and empirical optimum provides confidence that the calibration is physically grounded.

2.3 HBQ Scoring Mode

For polar binding pockets where hydrogen bonding and electrostatics dominate, we provide an extended scoring mode (HBQ) that adds two terms to the base van der Waals energy:

$$E_{\text{HBQ}} = E_{\text{disp}} + E_{\text{rep}} + E_{\text{Coulomb}} + E_{\text{H-bond}} \tag{7}$$

2.3.1 Coulomb Electrostatics

Electrostatic interactions are computed using Gasteiger partial charges [Gasteiger & Marsili, 1980] with a distance-dependent dielectric:

$$E_{\text{Coulomb}} = \sum_{i,j} \frac{332 \cdot q_i q_j}{\varepsilon(r_{ij}) \cdot r_{ij}} \tag{8}$$

where q_i, q_j are partial charges in electron units, r_{ij} is in Å, and the factor 332 converts to kcal/mol. The distance-dependent dielectric $\varepsilon(r) = 4r$ accounts for solvent screening in the binding site, increasing from $\varepsilon \approx 8$ at contact ($r \approx 2$ Å) to $\varepsilon \approx 40$ at 10 Å.

2.3.2 Directional Hydrogen Bonding

Hydrogen bonds are scored geometrically based on distance and angle criteria:

$$E_{\text{H-bond}} = \sum_{\text{H-bonds}} k_{\text{HB}} \cdot f_d(d_{\text{HA}}) \cdot f_\theta(\theta_{\text{DHA}}) \tag{9}$$

where d_{HA} is the hydrogen-acceptor distance, θ_{DHA} is the donor-hydrogen-acceptor angle, and $k_{\text{HB}} = -2.5$ kcal/mol is the maximum H-bond strength.

The distance factor f_d is:

$$f_d(d) = \begin{cases} 1 & \text{if } d < 2.5 \text{ Å} \\ \frac{3.0-d}{0.5} & \text{if } 2.5 \leq d < 3.0 \text{ Å} \\ 0 & \text{if } d \geq 3.0 \text{ Å} \end{cases} \tag{10}$$

The angular factor f_θ penalizes non-linear geometries:

$$f_\theta(\theta) = \begin{cases} 1 & \text{if } \theta > 150^\circ \\ \frac{\theta-120^\circ}{30^\circ} & \text{if } 120^\circ \leq \theta \leq 150^\circ \\ 0 & \text{if } \theta < 120^\circ \end{cases} \tag{11}$$

2.3.3 Rust Acceleration

All pairwise energy calculations are implemented in Rust via PyO3 bindings, enabling significant speedups over pure Python:

Kernel	**Python (ms)**	**Rust (ms)**	**Speedup**
vdW pairwise	219	1.9	117×
Coulomb	8.5	0.20	43×
H-bond geometry	16	0.007	2,376×

Table 1: Rust acceleration benchmarks for scoring kernels on a representative complex (∼5,000 protein atoms, ∼50 ligand atoms).

The dramatic H-bond speedup arises from vectorized distance-angle calculations that avoid Python's interpreter overhead.

2.4 Pocket Regime Classification

A key insight from our benchmark analysis is that different binding pockets respond differently to steric calibration. We formalize this observation through the **ΔClash descriptor**.

2.4.1 ΔClash Definition

For a given target and steric setting s_R, we define the clash rate as the fraction of poses with repulsive energy exceeding a threshold (1000 kcal/mol). The ΔClash descriptor captures the *differential* clash behavior between actives and decoys:

$$\Delta\text{Clash} = \text{Clash\%}_{\text{actives}} - \text{Clash\%}_{\text{decoys}} \tag{12}$$

2.4.2 Four Pocket Regimes

Based on ΔClash, we classify pockets into four regimes:

1. **Tight** ($\Delta\text{Clash} < -10\%$): Active ligands clash more than decoys. This indicates a constrained pocket where true binders must adopt strained conformations. *Softer sterics help.*
2. **Neutral** ($-10\% < \Delta\text{Clash} < +5\%$): Similar clash rates for actives and decoys. The pocket provides little geometric discrimination. *Default s_R works.*
3. **Discriminating** ($+5\% < \Delta\text{Clash} < +20\%$): Decoys clash more than actives. The pocket geometry itself provides selectivity—active ligands fit cleanly while decoys collide. *Physics-only scoring excels.*
4. **Broken** ($\Delta\text{Clash} > +20\%$ or total clash rate $> 60\%$): Extreme clash differential or universal clashing. Steric scoring is unreliable. *HBQ or ML heads required.*

2.4.3 Predictive Utility

The pocket regime predicts, *a priori*, how well physics-only scoring will perform:

- **Discriminating pockets** like AA2AR achieve high AUC (0.70) with pure vdW scoring because the pocket shape naturally rejects decoys.
- **Tight pockets** like P38A require softer sterics ($s_R = 0.45$) to allow actives to "breathe."
- **Broken pockets** like Thrombin and BACE1 cannot be rescued by steric tuning alone and require electrostatic terms (HBQ) or learned corrections (ML heads).

This classification enables pocket-aware scoring without per-target parameterization.

3 Experiments

[Section 4: Dataset description, docking protocol, evaluation metrics. To be completed.]

4 Results

[Section 5: Steric calibration sweep, per-target AUCs, pocket regime validation, HBQ results. To be completed pending HBQ data.]

5 Discussion

[Section 6: Physics vs empirical trade-offs, limitations, future work. To be completed.]

6 Conclusion

[Section 7: Summary. To be completed.]

References

Friesner, R.A., et al. (2004). Glide: A new approach for rapid, accurate docking and scoring. *Journal of Medicinal Chemistry*, 47(7):1739–1749.

Gasteiger, J. & Marsili, M. (1980). Iterative partial equalization of orbital electronegativity. *Tetrahedron*, 36(22):3219–3228.

Kitchen, D.B., et al. (2004). Docking and scoring in virtual screening for drug discovery. *Nature Reviews Drug Discovery*, 3(11):935–949.

Koes, D.R., Baumgartner, M.P., & Camacho, C.J. (2013). Lessons learned in empirical scoring with smina. *Journal of Chemical Information and Modeling*, 53(8):1893–1904.

Li, J., et al. (2019). An overview of scoring functions used for protein-ligand interactions. *Interdisciplinary Sciences*, 11:320–328.

Mysinger, M.M., et al. (2012). Directory of useful decoys, enhanced (DUD-E). *Journal of Medicinal Chemistry*, 55(14):6582–6594.

Trott, O. & Olson, A.J. (2010). AutoDock Vina: Improving the speed and accuracy of docking. *Journal of Computational Chemistry*, 31(2):455–461.

Verdonk, M.L., et al. (2003). Improved protein-ligand docking using GOLD. *Proteins*, 52(4):609–623.

Wang, R., et al. (2004). The PDBbind database: Collection of binding affinities. *Journal of Medicinal Chemistry*, 47(12):2977–2980.

Wang, L., et al. (2015). Accurate and reliable prediction of relative ligand binding potency. *Journal of the American Chemical Society*, 137(7):2695–2703.