香港中文大學
The Chinese University of Hong Kong

The dark side of the forces:
assessing non-conservative force models for atomistic machine learning

Bigi, Filippo, Marcel F. Langer, and Michele Ceriotti. arXiv:2412.11569 (2024).
Forthcoming in ICML 2025.

Jingyi Liu

26/03/2026

Outline

- **The Problem:** The rise of "direct" (non-conservative) force prediction.
- **Theory**
- **Failure Modes:** Geometry optimization and MD instabilities.
- **Analysis**
- **Solutions:** Hybrid training and evaluation strategies.

Research Background

Research Background

Physical constraints to obey

- **E(3)-invariance**: Interatomic potentials are invariant under the transformations of the Euclidean group in three dimensions E(3), including translations, rotations, and reflections.
- **Permutation invariance**: invariant with respect to permutations of atom indices.
- **Energy conservations**: interatomic forces are conservative, i.e., their mechanical work W over a closed loop is zero.

This is satisfied if and only if there exists a function (usually named the potential energy) of which the forces are the spatial derivatives

Nevertheless, differentiation incurs a computational overhead, typically 2−3×for inference and 3×for training; This overhead can be avoided by directly predicting forces during the forward pass.

Research Background

The Efficiency vs. Physics Trade-off

Conservative model

- Traditional MLIPs: Forces are exact gradients of energy ($F=-\nabla E$).
- Modern architectures (e.g., ORB -v2) often predict force vectors directly to avoid expensive backpropagation.

Non-conservative model

- Can models "learn" to be conservative from data alone?

Research Background

The dark side of the forces

- Direct force models break mathematical integrability, resulting in an asymmetric Jacobian (λ>0) where forces no longer derive from a consistent potential energy surface.
- This causes unphysical energy injection and systematic drift in MD simulations, effectively acting as a "phantom" heat source that thermostats cannot fully counteract.
- Furthermore, because these models lack the global context of an energy-gradient, they are more susceptible to geometric degeneracies and require larger cutoffs to maintain accuracy.

Research Background

Impact on geometry optimization: stalling & convergence issues

- Geometry optimization aims to find a local minimum. Forces are used during the optimization and stopping criteria.
- Without a true PES, optimizers like L-BFGS cannot find a stable global minimum.
- Non-conservative models show high residual forces and "thrashing" during relaxation.

Research Background

Impact on molecular dynamics: thermodynamic drift

- MD aims to at simulating the behavior of a microscopic system by solving its classical equations of motion numerically. Using a time step Δt, the simplest forms of molecular dynamics propagate a discretized version of Hamilton's equations.

- No underlying Hamiltonian can be defined for the dynamics generated by a non-conservative force field suggests that, in this case, artifacts might be more pronounced and harder to correct.

Proposed Mitigation Strategies

- Fine-tuning: Pre-train on direct forces for speed; fine-tune with energy gradients for physics.
- MTS (Multiple Time-Stepping): A hybrid approach using Non-constrain forces for fast steps and conservative corrections for slow steps.
- Using λ (the Frobenius norm of the antisymmetric component of the Jacobian to that of the Jacobian itself) as a criteria. $\lambda = 0$ → conservative forces; $\lambda = 1$, no conservative at all.

$$J_{i\alpha,j\beta} = \frac{\partial \mathbf{f}_{i\alpha}}{\partial \mathbf{r}_{j\beta}} = \frac{\partial \mathbf{f}_{j\beta}}{\partial \mathbf{r}_{i\alpha}} = J_{j\beta,i\alpha} \,.$$

$$\lambda = \frac{||\mathbf{J}_{\mathrm{anti}}||_{\mathrm{F}}}{||\mathbf{J}||_{\mathrm{F}}} \,,$$

Proposed Mitigation Strategies

- **PET / PET-NC:** The primary benchmark based on the Point Edge Transformer; **PET** is a standard **conservative model** ($F=-\nabla E$), while **PET-NC** is an unconstrained, **non-conservative** variant trained on the same bulk water dataset.

- **PET-M:** A hybrid architecture designed to predict **both conservative and non-conservative** force components simultaneously for direct comparison.

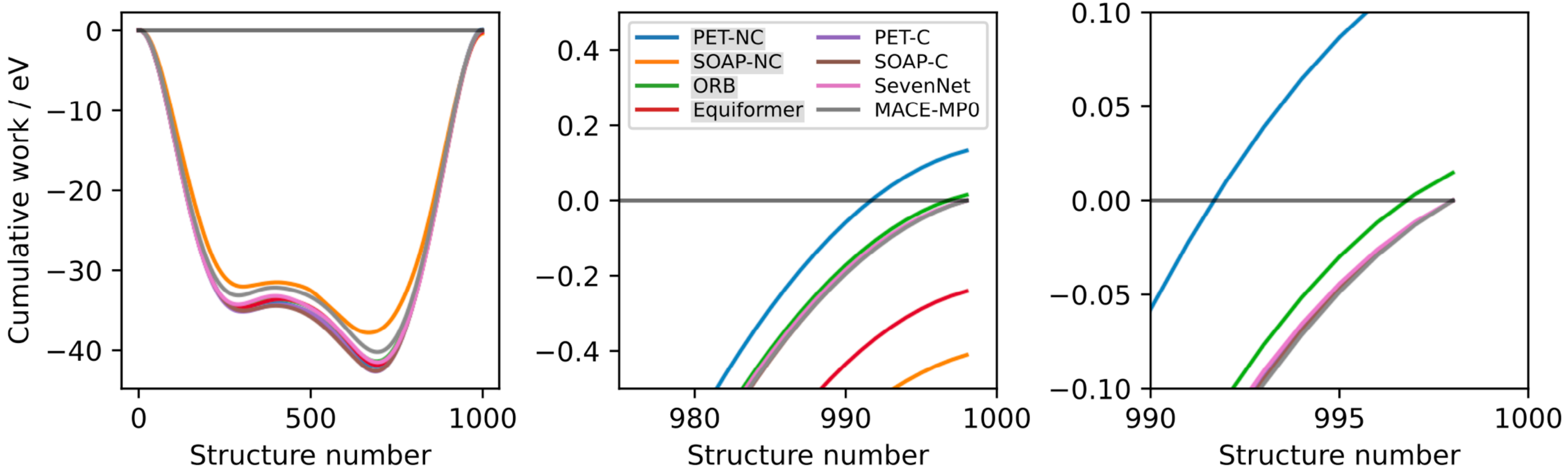
- **ORB-v2:** A state-of-the-art, **non-conservative** foundation model used to demonstrate that these physical inconsistencies persist in high-performing, real-world architectures.

- **SOAP-BPNN:** A "legacy" descriptor-based Behler-Parrinello Neural Network included to provide a **historical baseline for conservative force prediction.**

- **Foundation Models (MACE-MP-0, SevenNet, EquiformerV2):** A suite of pre-trained, "general purpose" models used to test the transferability of the observed non-conservative issues across diverse materials.

The data myth: can machines learn conservation?

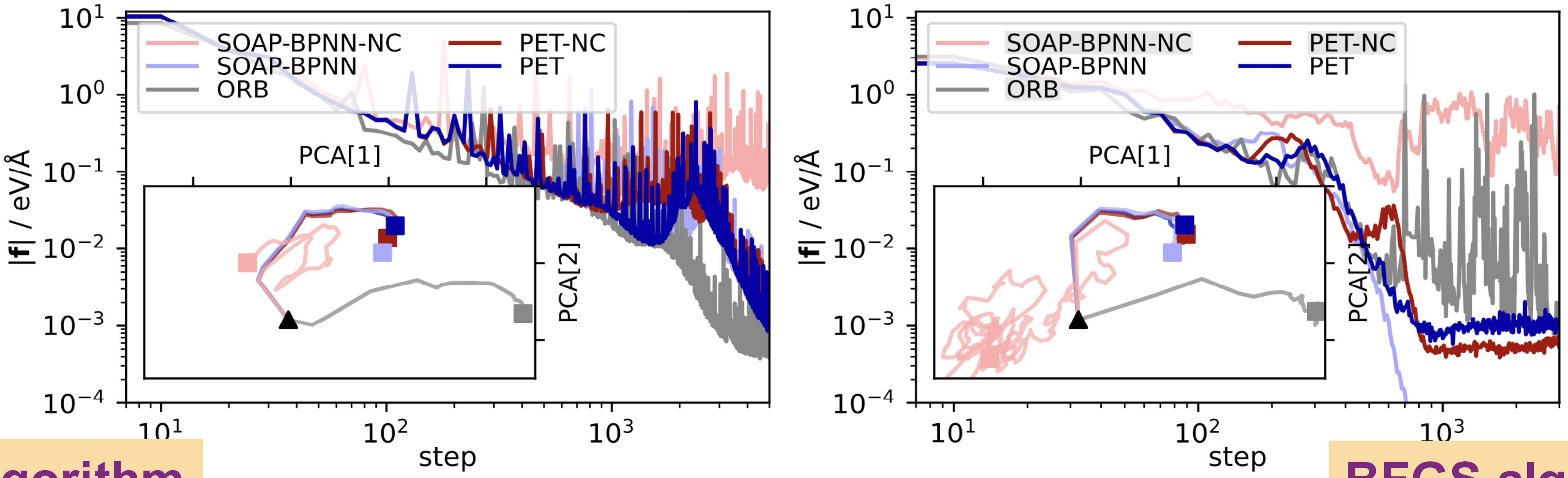
Figure 6. Cumulative work along a closed path for all models considered in the study. While the total work for the conservative models is zero up to machine precision, the non-conservative models exhibit a non-zero total work (ORB: 15 meV, Equiformer: -241 meV, PET-NC: 132 meV, SOAP-BPNN-NC: -410 meV). The first figure from the left shows the overall path, while the other two zoom in on the last part.



- Increasing training set size does not proportionally reduce the force curl.

Research Background

Proposed Mitigation Strategies: geometry optimization



FIRE algorithm

BFGS algorithm

Figure 7. Force modulus as a function of step along FIRE (left) and BFGS (right) optimization trajectories of a liquid water snapshot, for different MLIPs. The inset shows a latent-space projection of the trajectory in configuration space; the triangle indicates the starting configuration, and the squares the final one

- SOAP-BPNN, PET, PET-NC converge to a similar structure. SOAP-BPNN-NC never converges.
- ORB converges to a very different structure.

Research Background

Impact on molecular dynamics: thermodynamic drift

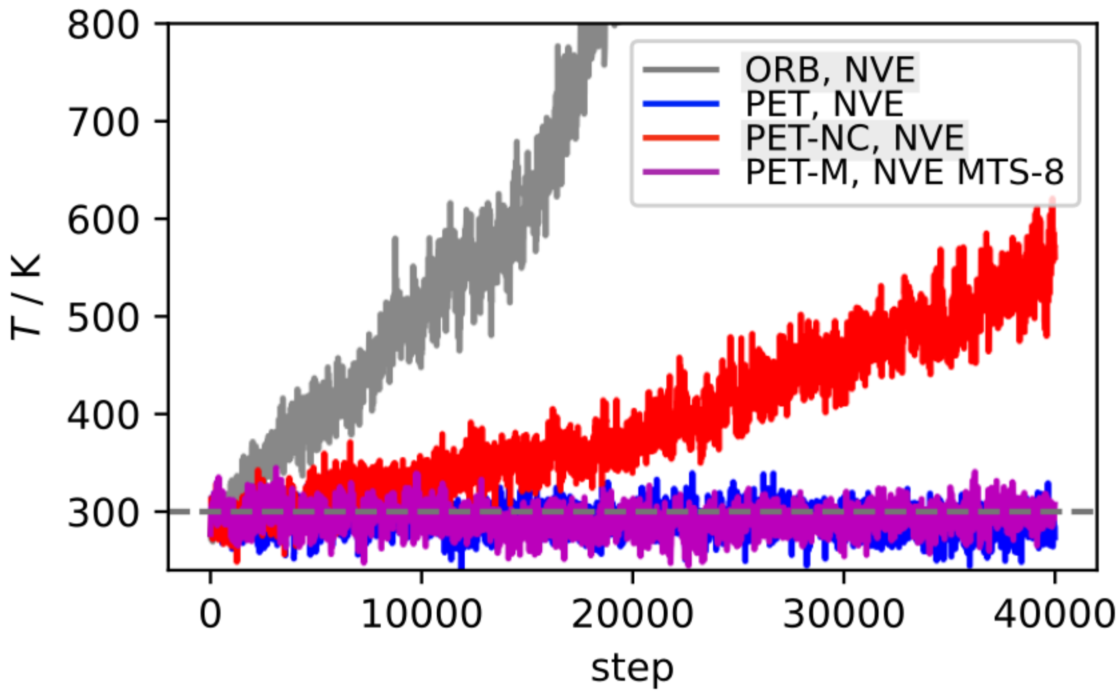


Figure 2. Time series for the kinetic temperature along a NVE MD trajectory, for ORB, the conservative and non-conservative PET models, and for a PET-M model using a multiple time-stepping (MTS) algorithm that evaluates conservative forces every 8 steps.

- NVE Ensembles: Rapid, unphysical energy increase (heating) or decrease (cooling).

- This unphysical drift corresponds to a rate of heating of about 7’000 billion degrees per second for the custom-trained PET-NC model, and another 10 times larger for the general-purpose ORB model.

Research Background

Proposed Mitigation Strategies

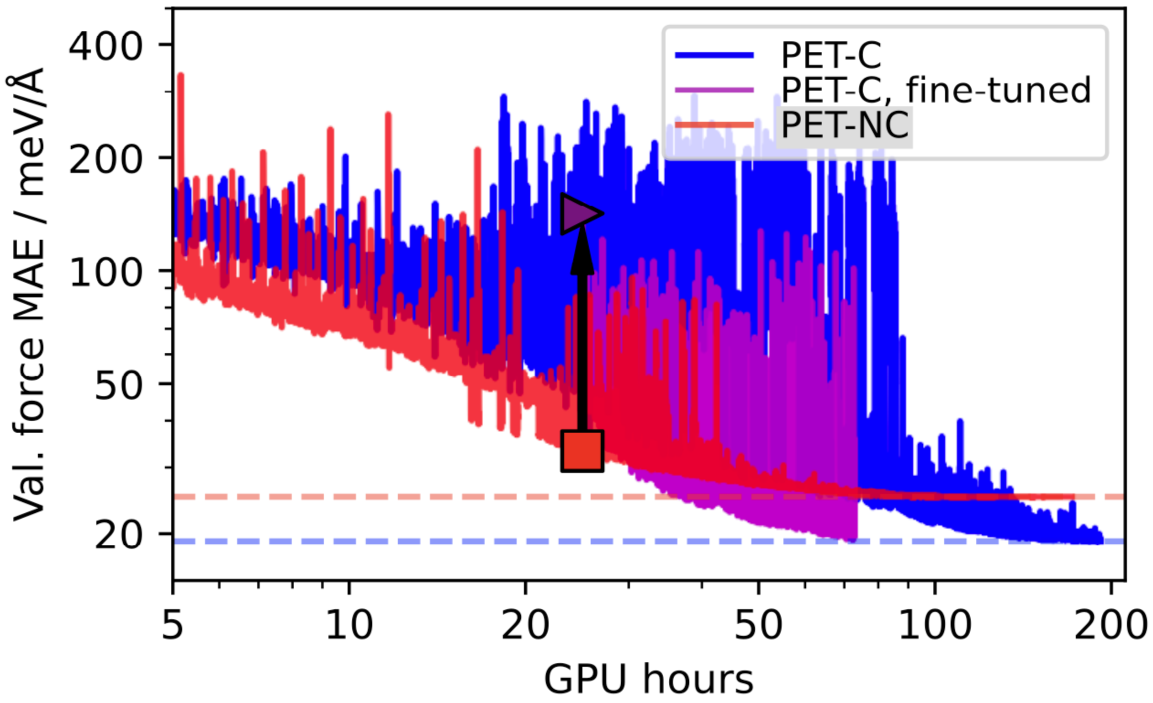


Figure 4. Training curves for PET-C and PET-NC models, and the conservative fine-tuning of a hybrid PET-M model initialized (epoch marked with arrow) from the potential energy head of the PET-NC model.

Conclusion

- Speed gains from direct force prediction come at the cost of thermodynamic validity.
- The "Dark Side" of forces is a fundamental limitation of unconstrained ML architectures.
- The best way to exploit the speed-up afforded by direct prediction of the forces is not to replace conservative models, but to augment them with a non-conservative head. That is, to supplement a conservative model with direct force predictions, using them to accelerate simulations that have well-defined, energy-conserving forces as the ground truth.

Questions? Comments?

Thank You