# Uncertainty in AI driven physical simulation

Dimitrios TZIVRAILIS

Laboratoire de Physique Théorique
et Modèles Statistiques

March 25, 2025

# Introduction

- Machine learning models are used to approximate interatomic potentials [1]

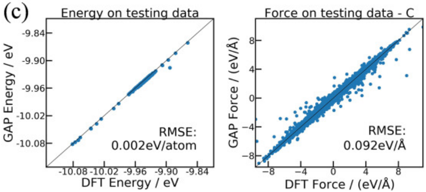



Figure 1: Energy and force predictions from NN approximation [3]

# Introduction

- Machine learning models are used to approximate interatomic potentials [1]

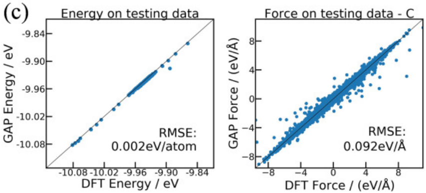
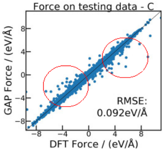



Figure 1: Energy and force predictions from NN approximation [3]

- These models introduce epistemic uncertainty [5] to their predictions due to their training protocols.



# Scientific Questions

- What are the consequences of the noise to the numerical simulations?

# Scientific Questions

- What are the consequences of the noise to the numerical simulations?

- What is the nature of the noise?

# Scientific Questions

- What are the consequences of the noise to the numerical simulations?

- What is the nature of the noise?

- How can the effect of the noise be mitigated? And how is uncertainty Quantification important for this task?

# What are the consequences of the noise to the numerical simulations?

## The $\phi^4$ model

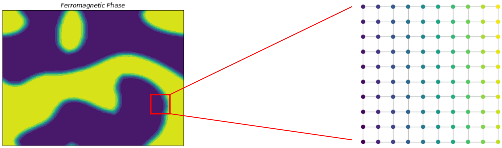



Figure 2: A $\vec{\phi}$ configuration in the ferromagnetic phase (left) and a zoomed-in view of the lattice structure (right)

$$S[\vec{\phi}] = \sum_{x \in L} \left[ - \sum_{\kappa=1}^{2} \beta \phi_x \phi_{x+e_\kappa} + \phi_x^2 + g \left( \phi_x^2 - 1 \right)^2 \right]$$

- $-\sum_{\kappa=1}^{2} \beta \phi_x \phi_{x+e_\kappa}$ interaction between neighbors.

# What are the consequences of the noise to the numerical simulations?

## The $\phi^4$ model

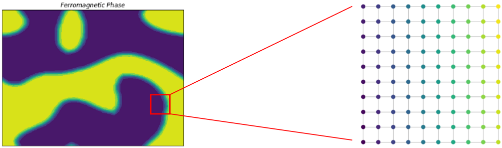



Figure 2: A $\vec{\phi}$ configuration in the ferromagnetic phase (left) and a zoomed-in view of the lattice structure (right)

$$S[\vec{\phi}] = \sum_{x \in L} \left[ -\sum_{\kappa=1}^{2} \beta \phi_x \phi_{x+e_\kappa} + \phi_x^2 + g\left(\phi_x^2 - 1\right)^2 \right]$$

- $-\sum_{\kappa=1}^{2} \beta \phi_x \phi_{x+e_\kappa}$ interaction between neighbors.
- $\phi_x^2 + g\left(\phi_x^2 - 1\right)^2$ interaction term of the field.

# What are the consequences of the noise to the numerical simulations?

## Metropolis Hastings's (MH) Algorithm [4]

The MH algorithm is an MCMC method used to sample configurations for computing physical observables in high-dimensional spaces.

- Starts from a configuration $\vec{\phi}$ and we suggest a $\vec{\phi}+\delta\vec{\phi}$.

# What are the consequences of the noise to the numerical simulations?

## Metropolis Hastings's (MH) Algorithm [4]

The MH algorithm is an MCMC method used to sample configurations for computing physical observables in high-dimensional spaces.

- Starts from a configuration $\vec{\phi}$ and we suggest a $\vec{\phi}+\delta\vec{\phi}$.
- We calculate the difference of action $dS$.

$$dS = S[\vec{\phi}+\delta\vec{\phi}] - S[\vec{\phi}] = \delta\vec{\phi}\cdot\nabla S[\vec{\phi}] + \text{extra terms} \quad (1)$$

# What are the consequences of the noise to the numerical simulations?

## Metropolis Hastings's (MH) Algorithm [4]

The MH algorithm is an MCMC method used to sample configurations for computing physical observables in high-dimensional spaces.

- Starts from a configuration $\vec{\phi}$ and we suggest a $\vec{\phi} + \delta\vec{\phi}$.
- We calculate the difference of action $dS$.

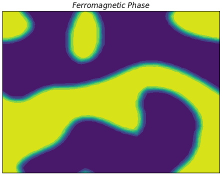
$$dS = S[\vec{\phi} + \delta\vec{\phi}] - S[\vec{\phi}] = \delta\vec{\phi} \cdot \nabla S[\vec{\phi}] + \text{extra terms} \quad (1)$$

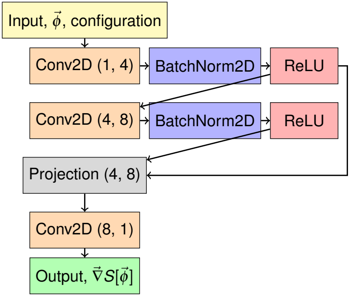
- We accept the move if:

$$u \geq min(1, exp(-dS)), \text{where u} \in [0,1] \text{ is an uniform random number} \quad (2)$$

# What are the consequences of the noise to the numerical simulations?

## Residual Convolutional Neural Network (RCNN) approximation of the $\nabla S[\vec{\phi}]$




- **Training protocol**: 200 independent samples, 180 training, 20 test, with batch size 24, for 50 epochs, with **Mean Squared Error (MSE)** loss function.

## Predictions of the $\nabla S[\vec{\phi}]$ with the RCNN

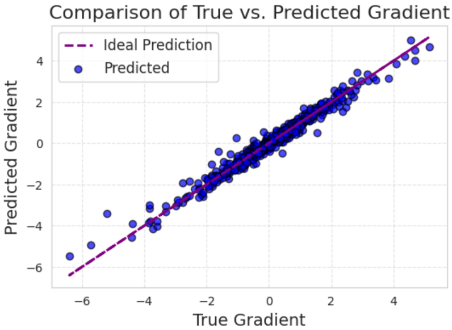



Figure 3: Predictions of the Gradient of the lattice field with good accuracy.

# How does noise affect simulations?

## Markov chain Trajectories

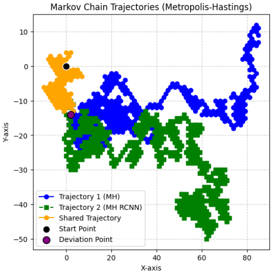



Figure 4: 2D Hypothetical Markov chain (MC) Trajectories

- Shared trajectory
  $min(1, exp(-\delta\vec{\phi} \cdot \nabla S[\vec{\phi}])) \simeq min(1, exp(-\delta\vec{\phi} \cdot \nabla S_{RCNN}[\vec{\phi}]))$

## Markov chain Trajectories

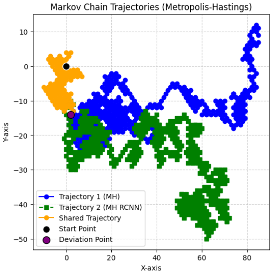



Figure 4: 2D Hypothetical Markov chain (MC) Trajectories

- Shared trajectory
  $min(1, exp(-\delta\vec{\phi} \cdot \nabla S[\vec{\phi}])) \simeq min(1, exp(-\delta\vec{\phi} \cdot \nabla S_{RCNN}[\vec{\phi}]))$
- $min(1, exp(-\delta\vec{\phi} \cdot \nabla S_{RCNN}[\vec{\phi}]))$, MH with RCNN

# How does noise affect simulations?

## Markov chain Trajectories

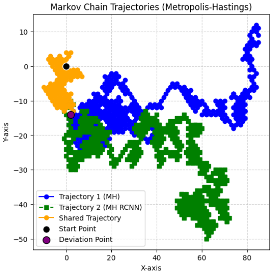



Figure 4: 2D Hypothetical Markov chain (MC) Trajectories

- Shared trajectory
  $min(1, exp(-\delta\vec{\phi} \cdot \nabla S[\vec{\phi}])) \simeq min(1, exp(-\delta\vec{\phi} \cdot \nabla S_{RCNN}[\vec{\phi}]))$
- $min(1, exp(-\delta\vec{\phi} \cdot \nabla S_{RCNN}[\vec{\phi}]))$, MH with RCNN
- $min(1, exp(-\delta\vec{\phi} \cdot \nabla S[\vec{\phi}]))$, MH

# How does noise affect simulations?

## MH sampling configurations for observables with RCNN

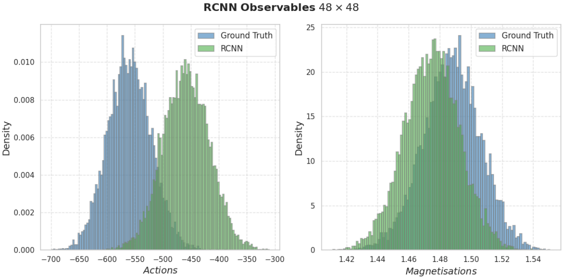



Figure 5: Discrepancy of the computations of the observables, using the RCNN (green) and the Ground Truth (blue) algorithm.

# What is the nature of the noise?

We trained the RCNN model using the MSE loss function:

$$\mathscr{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\nabla S[\vec{\phi}]_i - \nabla \hat{S}[\vec{\phi}]_i\right\|_2^2 \quad (3)$$

MSE assumes that the residuals follow a Gaussian distribution because it corresponds to the maximum likelihood estimate under this assumption.

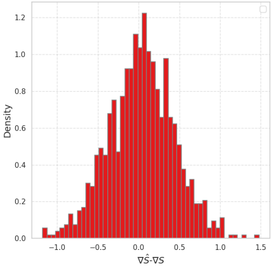



Figure 6: 1000 predictions for specific $\vec{\phi}$ of one component of the $\nabla S[\vec{\phi}]$

# How can the effect of the noise be mitigated?

**Penalty Technique [2]**

The RCNN approximation can be written as a normal distribution around the true value.

$$\nabla \hat{S}_i - \nabla S_i = \varepsilon_i \sigma, \quad \varepsilon_i \sim \mathcal{N}(0,1) \quad , \quad \forall i \in L \tag{4}$$

The introduced noise can be corrected with a penalty factor assuming that it follows a Gaussian distribution with variance $\sigma^2$. Thus, the new acceptance criteria is:

$$u \geq min(1, e^{-\nabla \hat{S}_i \cdot \delta \vec{\phi} - \frac{\sigma^2 (\delta\phi)^2}{2} - \cdots}), \quad \text{u is uniform random number } u \in [0,1] \tag{5}$$

## Ensemble technique plus penalty

We train N = 10 machines to predict the $\nabla S[\vec{\phi}]$ and we estimate the mean and the variance of the predictions with better precision trying to eliminate biases of the predictions.

$$\nabla \hat{S}_{mean} = E[\nabla \hat{S}_N], \quad \sigma^2_{penalty} = \boldsymbol{V}[\nabla \hat{S}_N], \quad min(1, e^{-\nabla \hat{S}_{mean} \cdot \vec{\delta\phi} - \frac{\sigma^2_{penalty}(\delta\phi)^2}{2} - \dots})$$

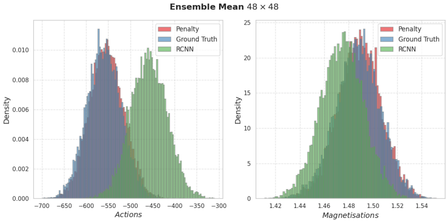
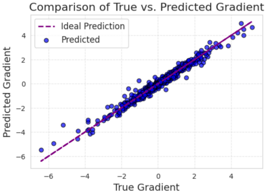
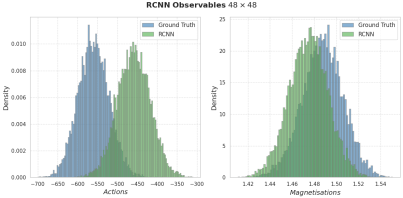



Figure 7: Penalty plus ensemble method can mitigate the effect of the noise, enabling correct sampling.

# Conclusions

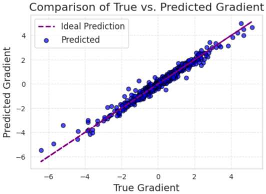
- High-accuracy machine learning approximations influence the sampling process in the MH algorithm.

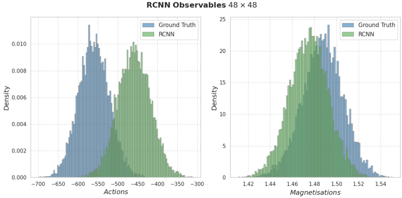


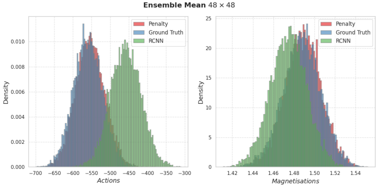
# Conclusions

- High-accuracy machine learning approximations influence the sampling process in the MH algorithm.




- We can mitigate the noise (Gaussian one) using a penalty technique and an ensemble method.



# Future works

- Extend our work by tackling the Metropolis adjusted Langevin equation.

# Future works

- Extend our work by tackling the Metropolis adjusted Langevin equation.

- Write an article to publish our approach and results.

# Future works

- Extend our work by tackling the Metropolis adjusted Langevin equation.
- Write an article to publish our approach and results.
- Use more advanced techniques for uncertainty quantification (e.g. Bayesian neural networks.)

# References


[1] Dylan M. Anstine and Olexandr Isayev. "Machine Learning Interatomic Potentials and Long-Range Physics". In: *The Journal of Physical Chemistry A* 127.11 (2023). PMID: 36802360, pp. 2417–2431. DOI: 10.1021/acs.jpca.2c06778. eprint: https://doi.org/10.1021/acs.jpca.2c06778. URL: https://doi.org/10.1021/acs.jpca.2c06778.

[2] D. M. Ceperley and M. Dewing. "The penalty method for random walks with uncertain energies". In: *The Journal of Chemical Physics* 110.20 (May 1999), 9812â9820. ISSN: 1089-7690. DOI: 10.1063/1.478034. URL: http://dx.doi.org/10.1063/1.478034.

[3] Ruiqiang Guo et al. "Small-data-based machine learning interatomic potentials for graphene grain boundaries enabled by structural unit model". In: *Carbon Trends* 11 (2023), p. 100260. ISSN: 2667-0569. DOI: https://doi.org/10.1016/j.cartre.2023.100260. URL: https://www.sciencedirect.com/science/article/pii/S2667056923000159.

[4] W. K. Hastings. "Monte Carlo sampling methods using Markov chains and their applications". In: *Biometrika* 57.1 (Apr. 1970), pp. 97–109. ISSN: 0006-3444. DOI: 10.1093/biomet/57.1.97. eprint: https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf. URL: https://doi.org/10.1093/biomet/57.1.97.

[5] Andrew A. Peterson, Rune Christensen, and Alireza Khorshidi. "Addressing uncertainty in atomistic machine learning". en. In: *Physical Chemistry Chemical Physics* 19.18 (2017), pp. 10978–10985. ISSN: 1463-9076, 1463-9084. DOI: 10.1039/C7CP00375G. URL: https://xlink.rsc.org/?DOI=C7CP00375G (visited on 10/30/2024).

Thank you for your attention.