

TIF345/FYM345 Project-2b: Bayesian Optimization: Searching for the global minima

Markus Utterström Salar Ghanbari

January 27, 2025

Introduction

Surface analysis plays a vital role in material science, aiding in the understanding of catalytic behavior, material stability, and interfacial phenomena. This project investigates the potential energy surface (PES)[1] of a gold atom on a gold surface, to efficiently locate its global minimum, the most stable configuration. We employ and compare computational techniques to minimize the cost of identifying this minimum. First, the PES is evaluated on a grid to establish a baseline and visually identify the global minimum. Next, we assess local search methods, analyzing their success rates from random starting points. We then implement Bayesian Optimization using Gaussian Processes (GPs) [2] to reduce the number of PES evaluations required. Finally, the GP model is used as a surrogate for the PES to explore transition pathways between minima and evaluate its accuracy. This structured approach highlights the efficiency of each method and demonstrates the utility of probabilistic models like GPs in computationally intensive material science problems.

Task 1, Analyzing the potential energy surface

The first task of this project involves direct calculation of the ad-atom energy. For this project, a function that calculates the ad-atom energy for a given position was provided and used on a dense grid. Only the primitive cell was checked, which was in the boundaries (0, 0) to (16.65653, 2.884996) Å. The direct calculation provides a visual location of where the global minimum is located. From this figure 1, we can see that there are multiple local minima, but there is a distinct global minimum located at $x = 3.2867$ Å, $y = 0.7237$ Å, with energy of 7.0761 eV.

The result is presented below in figure 1 where the white star marks the calculated global minimum. The coordinate z for the entire calculation was fixed to set the ad-atom on top of the surface. The used grid in the calculation was 300 times 300 grid points.

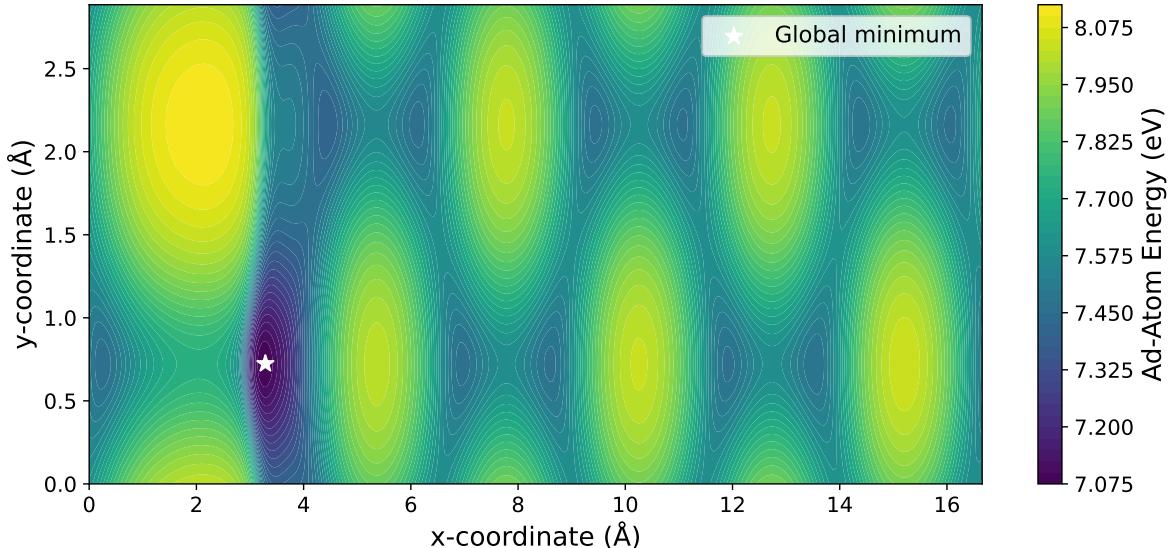


Figure 1: Contour plot over the PES for a set z-coordinate that fixed the ad-atom on top of the surface the calculation was made for a 300 times 300 grid.

This brute force method takes a long time to run and it is very dependent on the resolution to find a good local minimum. 90000 calculations were made to be able to acquire

this value for the global minimum.

Task 2, Local search

For this task, local searches were conducted for randomly selected points. By using gradient descent with `scipy`'s `minimize` function, 300 uniformly selected points on the grid were drawn and local searches were applied. There are 19 different local minima found from the 300 searches. 15.3% of them are around the global minimum calculated from the previous task.

The minima that were found with the local search are illustrated with red dots in figure 2. The global minimum from the previous task is marked with a white star. The global minimum is located at $x = 3.2591 \text{ \AA}$, $y = 0.7212 \text{ \AA}$, with an energy of 7.0751 eV. That is a difference of 0.001 from the calculated global minimum of the previous task.

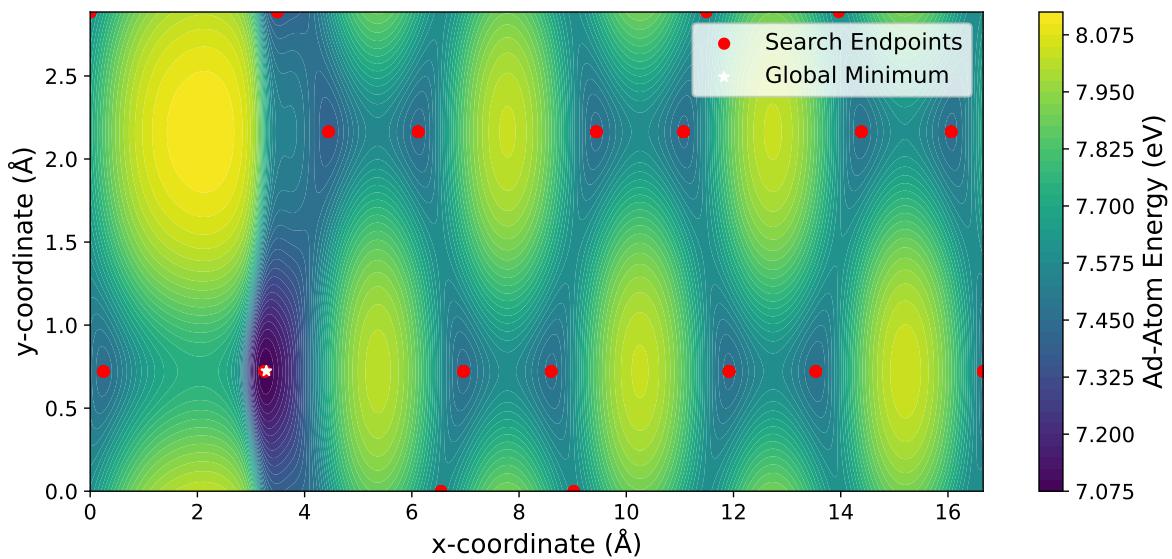


Figure 2: End points of local search showing clustering around local minima with a global minimum highlighted.

Using this method does find a fast way to evaluate the minima of the function but it is reliant on it not being too many red herrings of local minima. For this small surface, this method works well in finding the global minimum in quite a few iterations. Ideally, it will find the global minimum in less than 10 searches, but this is reliant on us already having an idea of how the PES looks and can visually determine how many local minima exist. Making this calculation required fewer iterations than the previous task and was quicker.

Task 3, Bayesian optimization

To find the minimum of the PES, a Bayesian process must also be tested. Here a Gaussian process was trained on two variables \mathbf{X} and \mathbf{E} , where \mathbf{X} is the matrix with positions and \mathbf{E} is the energies. At the start, five random samples were drawn to train the GP. The GP then calculates the mean and the variance from there, and a new sample $(x_{\text{new}}, y_{\text{new}})$ can be calculated from the acquisition function:

$$A(x, y) = -\mu(x, y) + \beta\sigma(x, y). \quad (1)$$

By selecting the (x, y) that maximizes the $A(x, y)$, the sample $(x_{\text{new}}, y_{\text{new}})$ can be added to the GP model and the energy at that point is calculated.

The choice of β heavily influences how well the GP performs, a lower value means less exploration and there is a high chance the GP gets stuck in a local minima and not the global one. A higher value means more exploration and thus also more iterations until it finds the global minimum. To find the best value, the GP was trained 25 times for the β -values 1 through 5. Then the average variance for each iteration was calculated the lowest energy for each iteration was extracted. Finally, the values were averaged over each beta and then plotted below in Figure 3.

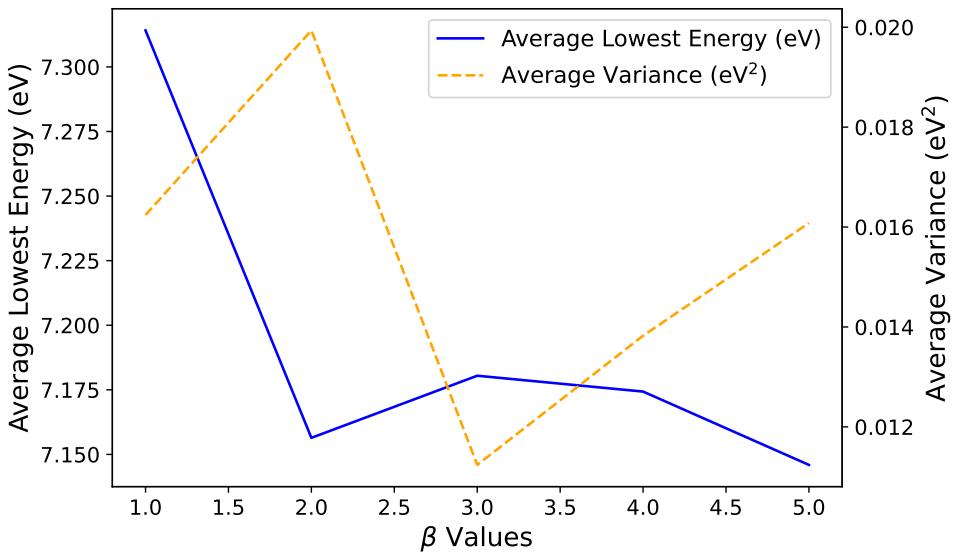


Figure 3: Relationship between β values, average lowest energy, and variance in Bayesian optimization, highlighting the exploration-exploitation trade-off.

Selecting a $\beta = 2$ implies a high rate of convergence to the global minimum but higher variance. For this project, $\beta = 5$ was selected as it could also explore how the function looks and give better estimations for the uncertainty for the whole PES and high chance of convergence to the global minimum. The average uncertainty is higher than other choices but the chance to converge to the global minimum is the highest for this choice, which is the central focus of this report.

The overall goal of this project is to find the global minimum of the energy in as few iterations as possible. Figure 1 already shows what the PES looks like and can thus help inform these decisions. This choice is made under the assumption that the previous tasks were not done and thus only reflect the success of the GP on its own. This influenced the choice of β as it is imperative to examine the full surface before concluding that a global minimum is reached. By selecting a lower value of β , the model converges more quickly but there is a high chance it converges somewhere else than the global minimum. To have an informative prior for the variance, an inverse gamma was chosen with $a = 2$ and $b = 1$. This also forces the variance to be positive which is needed per definition.

Here below is the mean and uncertainty of the final model at $\beta = 5$ that was ultimately selected. It had its global minimum at $x = 3.2610 \text{ \AA}$, $y = 0.7198 \text{ \AA}$, with an energy of 7.0751 eV. It is for all purposes the same as was acquired in the local search. Figure 4 shows the mean prediction and the uncertainty.

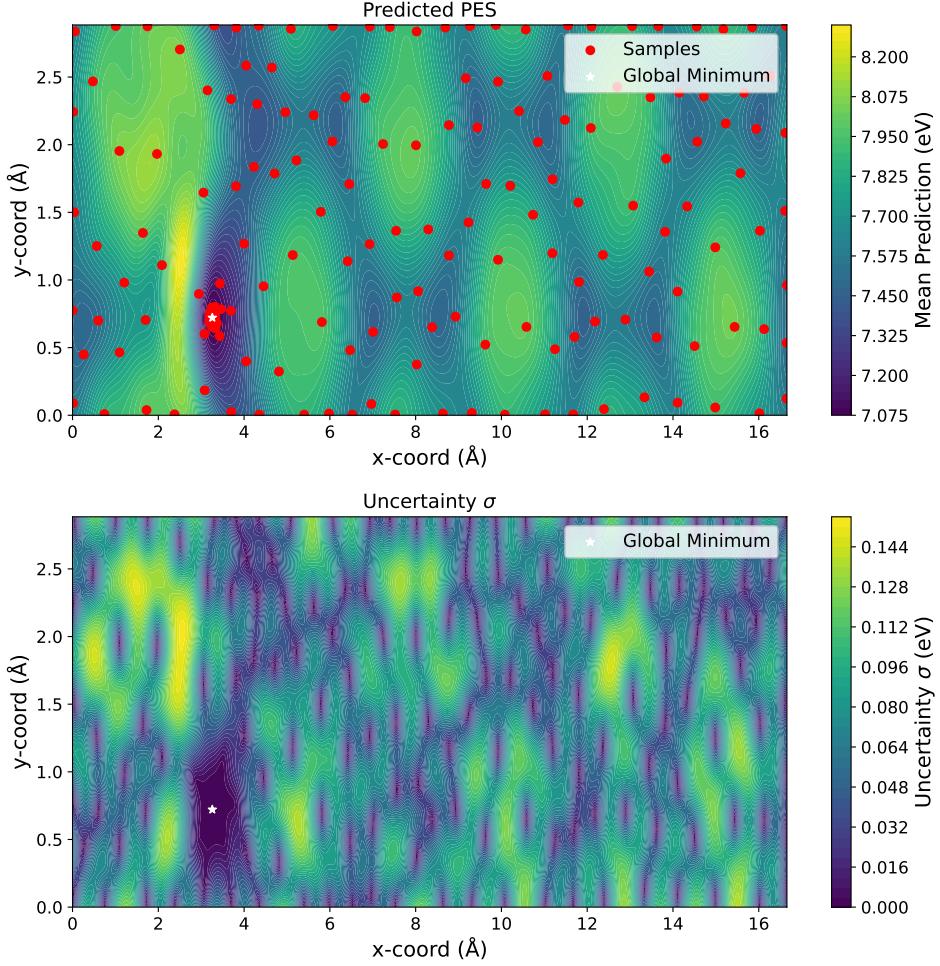


Figure 4: Predicted PES (top) and uncertainty σ (bottom) from the GP model, with sampled points and the global minimum highlighted.

The parts that show higher uncertainty are the parts that have not been explored as much as others. Notice that around the global minimum, the uncertainty is zero. The figure below, Figure 5 shows the acquisition function from the calculated samples. The figure presents it's highest values around the global minimum.

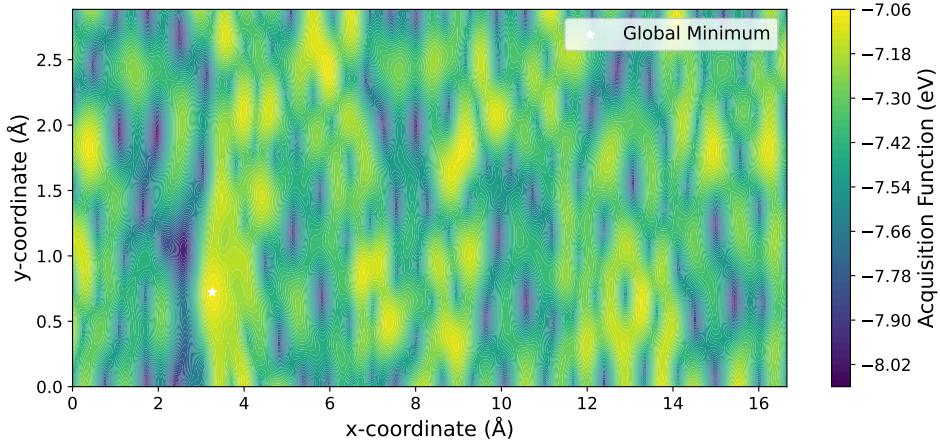


Figure 5: Acquisition function visualization highlighting the global minimum on the PES, guiding exploration during Bayesian optimization.

Task 4, Transition paths barriers

This task focuses on training a general-purpose Gaussian Process (GP) model to represent the full potential energy surface (PES) and analyze the energy and uncertainty along a transition path between the global minimum, $(x_{\text{start}}, y_{\text{start}})$, and a local minimum near $(x_{\text{end}}, y_{\text{end}}) = (11 \text{ \AA}, 2.1 \text{ \AA})$. The GP model was trained using initial samples from these two minima and additional random points to ensure broader coverage of the PES. To prioritize exploration and improve global accuracy, the lower confidence bound (LCB) acquisition function, was employed with a large β value to sample regions of high uncertainty. Approximately 100 samples were required to construct a globally accurate model. The transition path was parameterized as:

$$(x(\lambda), y(\lambda)) = (x_{\text{start}}, y_{\text{start}}) + \lambda(x_{\text{end}} - x_{\text{start}}, y_{\text{end}} - y_{\text{start}}), \quad (2)$$

where $\lambda \in [0, 1]$. The GP model was then used to compute the energy and its uncertainty along this path.

The root-mean-square error (RMSE) of the general-purpose GP model decreased significantly as the number of training samples increased. Starting from an initial RMSE of approximately 0.18 eV with minimal samples, it dropped below 0.05 eV after incorporating around 100 samples, demonstrating the model's ability to approximate the full PES with high accuracy as Figure 6 presents.

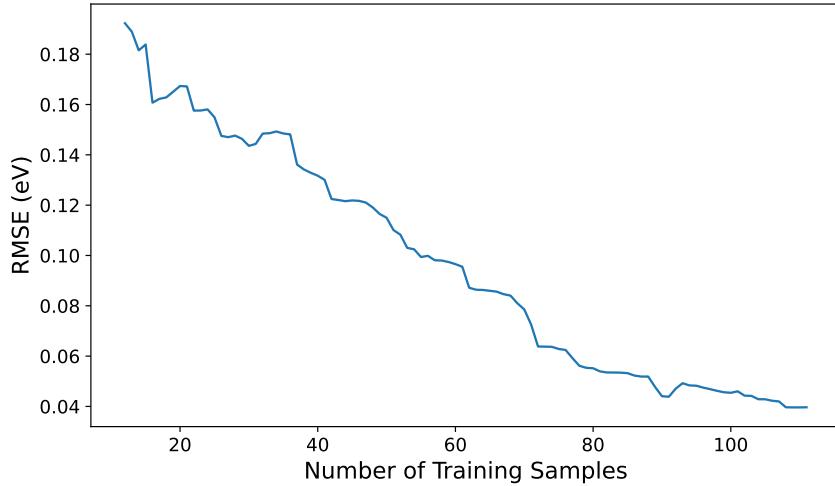


Figure 6: RMSE of the general-purpose GP model for the PES as a function of training samples, demonstrating improved accuracy with more samples.

Finally, the figure 7 illustrates The energy along the transition path, connecting the global minimum to the local minimum at $(11, 2.1) \text{ \AA}$, was computed using the general-purpose GP model, the Task 3 GP model, and direct EMT calculations. The general-purpose GP closely matched the EMT reference energies across the entire transition path, capturing both minima and energy barriers with small uncertainty bands . In contrast, the Task 3 GP model exhibited larger deviations from the EMT values, particularly away from the global minimum, highlighting its limitation in modeling the entire PES accurately.

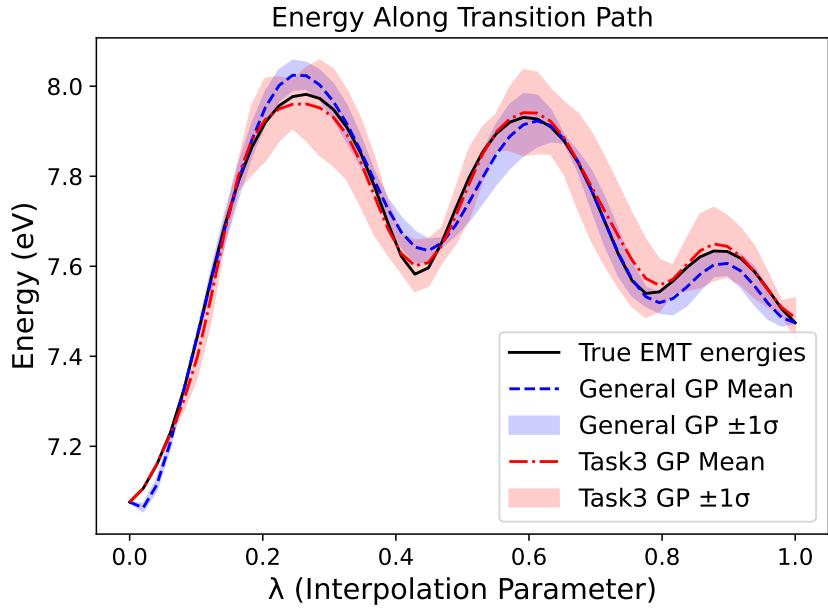


Figure 7: Energy along the transition path from the global to a local minimum, comparing EMT reference energies, the general GP model, and the Task 3 GP model, including their uncertainties ($\pm 1\sigma$).

Conclusions

This report discusses various methods for calculating the global minimum of a PES, first through direct calculation, then by doing multiple local searches and finally using a GP. The found result is that the lowest energy is 7.0751 eV. This result was found by both local searches and the GP. The method that could potentially find it in the least amount of iterations is the GP, but that relies on a lot of preparatory work as determining the best value for β .

References

- [1] Milica Todorović, Michael U. Gutmann, Jukka Corander, and Patrick Rinke. Bayesian optimization of materials informatics workflows. *npj Computational Materials*, 5:35, 2019.
- [2] Paul Erhart, Andreas Ekström, and Arkady Gonoskov. *Advanced Simulation and Machine Learning*. Lecture notes, 2024.