UNIVERSITY OF TURKU

Atomic Modelling of Oxidized Crystalline InAs (100) Surface with First Principles and Bayesian Optimization

Department of Mechanical and Materials Engineering
Faculty of Technology
Master's thesis

Author:
Shola Samuel Adeyemi

13.05.2024
Turku

Master's thesis

Subject: Materials Engineering: Modern Industrial Materials
Author(s): Shola Samuel Adeyemi
Title: Atomic Modelling of Oxidized Crystalline InAs(100) Surface with First Principles and Bayesian Optimization
Supervisor(s): Prof. Milica Todorović and Dr. Konstantinos Konstantinou
Number of pages: 62
Date: 13.05.2024

Abstract.

This study examines the surface reconstruction of the InAs(100)-ζ(4×2), the initial oxygen adsorption sites on the surface and their influence on the reconstructed optimized surface structure using a combined methodology of Density Functional Theory (DFT) calculations and Bayesian Optimization Structure Search (BOSS).

A detailed computational model of the InAs(100)-ζ(4×2) surface was systematically reconstructed and optimized across slab thicknesses ranging from 8 to 16 atomic layers, establishing 12 layers as the converged, stable thickness. Subsequently, an oxygen atom was placed near the reconstructed surface, and BOSS was employed to effectively map the potential energy landscape and predict the optimal adsorption configurations. A one-dimensional BOSS exploration optimized the oxygen atom's vertical position above the surface, and a two-dimensional BOSS exploration sampled possible lateral adsorption coordinates across the surface unit cell. Predicted atomic adsorption sites of the oxygen atoms identified by BOSS were fully relaxed with DFT to confirm stability and obtain precise adsorption energies.

The results reveal ten energetically favourable oxygen adsorption configurations on InAs(100)-ζ(4×2) reconstruction. In the most stable structures, oxygen binds between the surface indium-arsenic atoms, consistent with the general tendency of oxygen to preferentially oxidize group III-V elements. The BOSS methodology successfully located these adsorption sites with far fewer energy evaluations than a conventional potential energy surface search, demonstrating its efficiency for complex surface systems.

In future studies, these findings can be used as the foundation for further oxygen adsorption research on the surface to improve passivation strategies and support the development of more accurate models for InAs/oxide interfaces in electronic and optoelectronic devices.

Key words: InAs(100)-ζ(4×2), BOSS, DFT, reconstructed,

Table of contents

1 Introduction 6

1.1 Background 6

1.2 Research Problem 8

1.3 Scope of the Thesis 9

1.4 Significance of the Study 9

2 Theoretical Background 10

2.1 Atomistic Modelling 10

2.2 Density Functional Theory 10

2.3 Bayesian Optimizational Structural Search (BOSS) 14

2.3.1 Principles of BOSS 15

2.3.2 Acquisition Function 16

2.3.3 Gaussian Process, Kernel, and Hyperparameters 18

2.3.4 Integration with DFT Calculations 20

3 Methods 23

3.1 Bulk and Surface Model Preparation 23

3.2 Surface Slab Construction 25

3.3 Geometry Optimization 27

3.4 Oxygen Adsorption Test 28

3.5 BOSS implementation for oxygen adsorption sites 29

3.6 Relaxation of the oxidized model 31

4 Results and Discussion 33

4.1 Surface Reconstruction 33

4.2 Slab Optimization 33

4.2.1 Bond Length and Bond Angles 33

4.2.2 Oxygen Adsorption Test 37

4.3 Predicted Oxygen Adsorption Region 38

4.3.1 Global height 38

4.3.2 Lateral Direction 41

4.3.3 Local Minimum Search 44

4.4 Oxygen Adsorption Sites 46

5 Conclusion **52**

References **53**

Appendix: Atomic Coordinates of InAs(100)-ζ (4×2) surface **61**

Acronyms

1D – One Dimension

2D – Two Dimension

eLCB – exploratory lower confidence bound

As - Arsenic

BO – Bayesian Optimization

BOSS – Bayesian Optimization Structural Search

DFT – Density Functional Theory

GP – Gaussian Process

In – Indium

InAs – Indium Arsenide

O – Oxygen

PES – Potential Energy Surface

PH – Pseudo-Hydrogen

1 Introduction

1.1 Background

Semiconductor materials play a critical role in modern electronics and photoelectronic devices such as MOSFET, solar cells, and photodetectors [1] as the modern worlds rely on the efficiency of their performance. The unique electronic properties of semiconductors arise from their crystal structure, bandgap, and carrier mobility, which make them highly adaptable for a wide range of applications[2]. They are also commonly found in personal computers, smart devices, as well as automated systems, due to their rapid data transport speed, durability of the device, and the sustainable ease of manufacturing [3] [4]. Over several decades, the semiconductor industry has largely relied on silicon (Si) [5] due to its abundance, well-understood processing methods, and stable native oxide (SiO_2)[6], which enables high-quality interfaces for device fabrication. However, silicon-based devices are approaching fundamental physical limits, particularly in terms of speed, power consumption, and miniaturization [7].

To overcome these challenges, alternative semiconductor materials are being actively explored, and among a wide range of semiconductors, III-V have gained attention for more than three decades[8], which is particularly due to their high electron mobility, narrow bandgap, and strong optoelectronic properties [9]. They consist of elements from group III (such as Ga, In, Al) and group V (such as As, P, N) of the periodic table, and these materials exhibit direct bandgap, efficient light absorption and emission, as well as heterostructure compatibility, enabling precise tuning of the band structure [10] [11][12].

Indium Arsenide (InAs) is a III–V semiconductor in the zinc blende cubic structure [13] under ambient conditions. It has a cubic lattice of a constant of 6.058 Å, a direct narrow band gap of approximately 0.36 eV at room temperature[14] [15], and high electron mobility [16] , making it highly suitable for low-voltage electronics. InAs has been utilized in a broad range of technologies, including Infrared photodetectors and focal plane arrays, quantum dot and quantum wire structures, tunnelling field-effect transistors (TFETs), heterojunction bipolar transistors (HBTs), and nanoscale devices for quantum computing platforms [15] [16] [17] [20] [21] [22].

The surface of a semiconductor is a critical region due to its termination at the periodic crystal lattice. When a crystal is cleaved to expose a surface, atoms at the termination plane lose

bonding with neighbouring atoms, which leads to the formation of dangling bonds, and these unsaturated bonds contribute to a high surface energy and often lead to surface reconstructions, where the surface atoms rearrange to minimize the system's total energy[23]. Surface reconstructions are not merely geometric rearrangements, they have significant effects on the electronic properties of the material as they introduce surface states within the bandgap, act as electron or hole traps, and alter band alignment at interfaces [24] and these effects are particularly significant in nanoelectronics, where the surface-to-volume ratio is high, and surface states can dominate device behaviour.

The InAs(100) surface is one of the most widely used orientations in device fabrication due to its favourable epitaxial growth characteristics, as it exhibits complex reconstruction patterns in the clean state and is highly sensitive to environmental exposure [25]. Therefore, understanding the atomic-scale structure of InAs(100) surfaces, both in the clean and oxidized state, is crucial for advancing InAs-based device reliability and performance. Accurate modelling of surface reconstructions and oxidation mechanisms can provide key insights into optimizing surface passivation and dielectric integration, reducing interface defects, and improving carrier mobility in transistors. There have been several proposed energetically stable surface reconstructions for InAs(100) surface such as Beta 2(β2), Zeta(ζ), Zeta a(ζa), Zeta b(ζb), Alpha 2(α2), Alpha 3(α3) [26]. In study will focus on the ζ (4x2) reconstruction, which has been argued to be one of the most stable forms in InAs(100) surface [27].

The study of oxidized semiconductor surfaces is also a critical aspect of semiconductor research. While surface reconstructions occur on clean semiconductor surfaces, practical devices inevitably involve exposure to ambient environments or deliberate oxidation during the processing of the semiconductor device. Therefore, the formation of an oxide layer at the semiconductor interface, as shown in Figure 1, becomes significant, as it can significantly impact the device performance.

In silicon-based technologies, the formation of a stable and high-quality Si/SiO_2 interface is one of the many reasons for the success of the Complementary Metal Oxide Semiconductor (CMOS) technology. The native SiO_2 layer provides excellent electrical insulation, low interface trap density, and high thermal and chemical stability [26] [27].

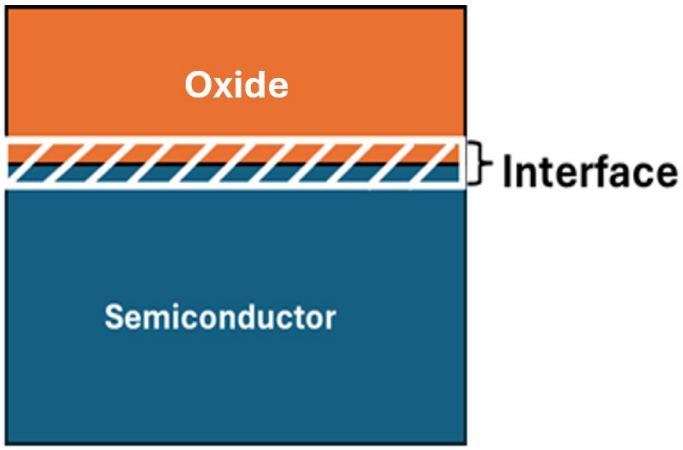

Figure 1: Schematic representation of a semiconductor and oxide interface

In contrast, III–V semiconductors lack a stable native oxide as they tend to form oxides of each constituent element present in the semiconductor [30]. Upon exposure to oxygen, unlike silicon, which forms a stable and protective oxide layer (SiO_2), the oxides on InAs can cause a chemical imbalance at the surface [31] . Despite the promising characteristics of InAs, the application in practical devices faces major challenges, especially concerning surface and interface stability. InAs readily oxidises upon air exposure, forming a mixture of In-rich and As-rich oxides at a temperature range of 290 – 330 °C [32], and this can affect the overall properties of the semiconductor [33]. By studying these processes at the atomic level through simulations, we can learn where oxygen atoms prefer to adsorb, how they change the surface structure, and how to potentially control or prevent the formation of unstable oxides.

1.2 Research Problem

The oxidation of the InAs(100) surface introduces structural disorder, which impacts charge transport and device reliability. Traditional surface reconstruction models have focused on clean, non-oxidised surfaces, but the detailed atomic structure of oxidised InAs(100) is yet to be explored. The research objectives of the thesis are given below:

a. Obtain the ζ(4×2) surface reconstruction of InAs(100)

b. Perform geometry optimization of the reconstructed surface model

c. Understand the initial oxidation of the InAs(100) surface

1.3 Scope of the Thesis

The thesis focuses on the investigation of the structural changes of the reconstructed InAs(100)-ζ(4×2) surface model under oxidation using computational modelling. The study highlights the reconstruction of the surface, optimization of the reconstructed surface, and oxygen adsorption on the optimized reconstructed surface. This is achieved through a combination of DFT and BOSS.

1.4 Significance of the Study

The results of this study will provide insights into the role of oxidation in III-V semiconductor surfaces, particularly for device applications in transistors, quantum electronics, and optoelectronic components. A deeper understanding of oxide-semiconductor interactions will contribute to the development of:

a. Improved surface passivation techniques to enhance device reliability.

b. Optimized oxide/semiconductor interfaces for high-performance transistors.

c. Techniques for minimizing defects at the surface and improving carrier mobility.

This research aligns with broader efforts in materials design, semiconductor engineering, and computational approaches combining DFT with AI-based optimization for materials science.

2 Theoretical Background

2.1 Atomistic Modelling

Atomistic modelling is a computational technique that simulates the behaviour of materials at the scale of individual atoms. They are becoming increasingly popular since the advance of computational power over the last 40 years, and are being deployed in various studies [34]. These methods are essential in materials science, engineering, and discovery for investigating structural, mechanical, as well as electronic properties of materials. These techniques can complement available knowledge from experiments and can also improve on existing experimental knowledge [35] [36] [37], since they can provide important information at the atomic level to tailor existing processes and propose new processes that can be implemented in experiments.

Atomistic simulations can be broadly classified into classical methods, such as molecular dynamics (MD) and Monte Carlo simulations, and quantum mechanical methods, such as density functional theory (DFT) [38]. The classical molecular dynamics method uses empirical interatomic potentials to model atomic interactions and is suitable for large-scale systems, often lacking the accuracy required to capture electronic effects, while the quantum mechanical method, which solves the electronic structure of materials from first principles, uses fewer atoms [39], making them more accurate but computationally expensive. In surface science, atomistic modelling enables the prediction of surface reconstructions, adsorption behaviour of atoms/molecules, formation of surface states and electronic defects, chemical reactions, and diffusion on surfaces [38] [39].

For this thesis, DFT is employed as the primary simulation tool due to its ability to accurately model both structural and electronic properties of clean and oxidized InAs(100) surfaces.

2.2 Density Functional Theory

Density functional theory (DFT) is a quantum mechanical modelling method for obtaining the physical properties of a system based on the electron density distribution in its ground state instead of its many-electron wave function [42]. To fully understand the electrical and structural characteristics of any material, the many-body system of interacting nuclei and electrons should be explained. To understand this, it should first be pointed out that at a fundamental level, the properties of ordinary matter originate from the quantum mechanics of

electrons and nuclei interacting under Coulomb forces [43]. Formally, one can describe a chemical or condensed matter system by its many-body wave function, which depends on the location and spins of every electron and, in principle, the nuclei. Solving the non-relativistic, time-independent Schrödinger equation [44] [45] written for such a system is key to understanding its thermodynamic and chemical properties.

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \quad (1)$$

where $\hat{H}$ is the Hamiltonian operator, $|\Psi\rangle$ is the quantum state vector of the system, and E is the energy eigenvalue corresponding to that state. However, the problem becomes traceable only under certain approximations, one of which is the Born–Oppenheimer approximation [46], which allows the total Hamiltonian to be separated into electronic and nuclear parts. The electrons move in the field of clamped nuclei, and the nuclear motion can be accounted for subsequently under the Born–Oppenheimer approximation [47] [48] [49]. The electronic Hamiltonian of an N-electron system in atomic units can be written as:

$$\widehat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i v(r_i) + \sum_{i<j} \frac{1}{|r_i - r_j|} = \hat{T} + \hat{V} + \widehat{W} \quad (2)$$

where ∇^2 is the Laplacian, $\hat{T}$ is the electronic kinetic energy operator, $\hat{V}$ is the external potential energy operator, and $\hat{W}$ is the electron–electron interaction energy operator. The electronic energy of the system is derived by solving the Born–Oppenheimer approximation, and the classical Coulombic nucleus–nucleus repulsion energy [50] is then added to obtain the total energy, which is provided in equation 3. The value of the ground-state energy expectation can be expressed as follows in standard quantum chemical nomenclature:

$$E[\Psi] = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \quad (3)$$

With the normalization $\langle\Psi|\Psi\rangle = 1$, by the variational principle, any trial wavefunction ψ gives an upper bound to the true ground-state energy E_0 and $E_0 \leq E[\psi]$. Minimizing $E[\psi]$ overall admissible wavefunction yields the exact ground state solution $|\Psi\rangle$ and E_0 [51]. This approximation has been employed because it makes computations easier and enables a clear definition of the key idea of potential energy in several studies [52]. However, the multi-dimensional nature of electron correlation makes a direct solution of this

many-body problem computationally prohibitive, necessitating further approximations beyond Born–Oppenheimer. One of the most fundamental approximations beyond BO is the Hartree–Fock (HF) method, wherein the electrons move in an average (mean) potential due to all other electrons. The HF wavefunction Ψ_{HF} for a fermionic system is typically approximated by a single Slater determinant, ensuring the required antisymmetry [53] [54] [55]:

$$\Psi_{HF}(x) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \cdots & \chi_N(x_1) \\ \vdots & \ddots & \vdots \\ \chi_1(x_N) & \cdots & \chi_N(x_N) \end{vmatrix} \quad (4)$$

where each row corresponds to an electron coordinate $x_j = (r_j, s_j)$ (space and spin), and each column corresponds to one electron spin orbital. The Slater determinant ensures that exchanging two electrons changes the sign of the wavefunction, reflecting the Pauli principle:

$$\Psi_{HF} = (\ldots, xi, \ldots, xj, \ldots) = - \Psi_{HF} (\ldots, xj, \ldots, xi, \ldots) \quad (5)$$

Using the variational principle once again:

$$E_{HF}[x_i] = \frac{\langle \Psi_{HF} | \hat{H} | \Psi_{HF} \rangle}{\langle \Psi_{HF} | \Psi_{HF} \rangle} \quad (6)$$

The selection of spin orbitals results in the HF total energy being reduced. By using Lagrange multipliers to impose these restrictions, a set of coupled one-electron equations for the spin orbitals is defined by the Hartree–Fock equations. Each electron experiences an average Coulomb and exchange potential from the other electrons, but electron correlation is only partially included through the exchange term. As a result, while HF captures exact exchange, it neglects dynamic correlation (the instantaneous correlated motions of electrons), requiring post-Hartree–Fock approaches to go beyond mean-field accuracy.

Hohenberg and Kohn proved two important theorems that form the basis of DFT. These theorems establish that the external potential $V_{\text{ext}}(r)$, and, consequently, every property of the system is precisely determined by the ground-state density n(r) [56] [57].

I: For a system of interacting particles, the ground-state density n(r) and the external potential $V_{\text{ext}}(r)$, are proportional to each other, up to a constant.

II The ground-state energy and density can be obtained by minimising a universal total energy functional E[n]

Total energy functional:

$$E[n] = \int n(r)V_{\text{ext}}(r)d^3r + F[n] \qquad (7)$$

Variational principle:

$$E_0 = \min_{n(r)} E[n] \qquad (8)$$

Here, F[n] is a universal function comprising the kinetic and electron–electron interaction energies. The Hohenberg-Kohn theorems do not explicitly state the form of this functional, but they do ensure its existence.

Kohn and Sham developed a practical implementation of DFT by adding a non-interacting electron auxiliary system that replicates the actual system's ground-state density [58]. This leads to the Kohn–Sham equations, which are given in equations 9 to12.

$$[-½\nabla^2 + V_{e\text{xt}}(r) + V_H(r) + V_{\text{xc}}(r)]\psi_i(r) = \varepsilon_i\psi_i(r) \qquad (9)$$

Electron density:

$$n(r) = \sum|\psi_i(r)|^2 \qquad (10)$$

Hartree potential:

$$V_H(r) = \int n(r') / |r - r'| \, d^3r'' \qquad (11)$$

Total Kohn–Sham energy:

$$E[n] = T_s[n] + E_H[n] + E_{xc}[n] + \int n(r)V_{ext}(r)d^3r \qquad (12)$$

Here, $T_s[n]$ is the non-interacting kinetic energy, $E_H[n]$ is the classical electrostatic energy, and $E_{xc}[n]$ is the exchange–correlation functional. The Kohn–Sham method allows the many-body problem to be reduced to a set of single-particle equations, greatly simplifying computations. The Kohn–Sham formalism offers several key advantages. It provides a single-particle framework, replacing the complicated many-body problem with a simpler system of non-interacting electrons. In theory, it is accurate, which means that the precise ground-state energies and electron densities would be obtained if the genuine exchange correlation V_{xc} potential were known. However, since the exact form of E_{xc}[n] is unknown; the accuracy of practical DFT depends on the quality of the chosen approximation.

The exchange correlation energy, denoted by $E_{xc}\,[n(r)]$, captures all many-body effects that go beyond the simpler Hartree (classical Coulomb) and non-interacting kinetic energy terms. Conceptually, E_{xc} is often decomposed into two parts:

- Exchange energy (E_x), which illustrates how electrons are fermionic and the Pauli exclusion principle works.
- Correlation energy (E_c), accounting for dynamic rearrangements of electrons and the resulting screening of Coulomb repulsion.

In the Hartree–Fock (HF) method, the exact exchange term can be expressed by the integral:

$$E_x^{HF} = \frac{1}{2}\sum_{i,j}\int\int \frac{\psi_i(\boldsymbol{r})\,\psi_j(\boldsymbol{r'})\,\psi_i(\boldsymbol{r'})\,\psi_j(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r'}|}\,d^3r\,d^3r' \qquad (13)$$

Here, $\psi_i(r)$ are spin orbitals, and the summation runs over all occupied orbitals i, j. This term results from the electronic wave function's antisymmetry, which causes electrons with parallel spins to effectively repel one another and apply the Pauli exclusion principle.

The correlation energy E_c is then defined as the remainder of E_{xc} after subtracting exchange:

$$E_c = E_{xc} - E_x \qquad (14)$$

It captures the additional energy lowering due to correlated electron motion, beyond any mean-field or single-determinant approximation. Because the exact form of E_{xc}[n(r)] is unknown; DFT relies on approximate functionals, and common examples include [59] :

1. Local Density Approximation (LDA), which is based on the homogeneous electron gas.
2. Generalized Gradient Approximation (GGA), which incorporates density gradients.
3. Meta-GGA and Hybrid Functionals, which capture higher-order effects or include a fraction of Hartree–Fock exchange, further enhancing accuracy. These approximations often yield results close to experimental data for molecules, bulk solids, and surfaces, making them very useful in computational materials science and quantum chemistry.

2.3 Bayesian Optimizational Structural Search (BOSS)

Traditional structure searches with DFT calculations rely on manually choosing configurations or exhaustive sampling, both of which are inefficient for large configurational spaces. Bayesian optimization offers a data-efficient alternative by modelling the energy landscape as a probabilistic surrogate function, guiding the search toward promising configurations with fewer evaluations. During this active machine learning process, different

atomic configurations are iteratively explored, aiming to efficiently identify the most energetically favorable structures. This approach enables the use of minimal computational resources in structural searches, maintaining both high accuracy and computational feasibility.

2.3.1 Principles of BOSS

Bayesian Optimization (BO) is used for global optimization of costly black-box functions. In materials science, this often means optimizing a function that can be obtained through a costly experiment or simulation, such as a DFT calculation. BO is designed to find the global minimum in as few evaluations as possible by building a probabilistic model of the unknown function and using it to decide where to sample next [60]. Unlike local gradient descent methods [61], BO treats the target function as a black box, and this makes BO well-suited for atomistic modelling problems where each function evaluation (a simulation of a specific atomic configuration) is time-consuming and where the underlying energy landscape may be complicated with many local minima.

The term “Bayesian” reflects that the method maintains a belief (probability distribution) about the unknown function and updates this belief after each evaluation, according to Bayes’ rule [62]. At the heart of BO is a surrogate model, a model that approximates the true function based on observed data. In modern BO implementations, this surrogate is often a Gaussian Process (GP) due to its flexibility and ability to provide uncertainty estimates [63] , and the surrogate’s predictions (mean and uncertainty) are used to determine where to sample next by optimizing an acquisition function. By balancing exploration and exploitation, BO prioritizes evaluations that are most likely to improve the current best result or reduce uncertainty in unexplored regions. In essence, the surrogate guides the search, it allows BO to peek at the shape of the objective function and make an informed guess for the next sample, instead of blindly sampling.

Figure 2 illustrates this iterative BO process, where the surrogate model is refined with each new sample, steadily converging towards the minimum. This iterative loop enables BO to learn about the function as it performs the runs, and early iterations cast a broad net to explore the landscape, while later iterations build on the most promising region as the surrogate’s estimate of the minimum becomes more confident. This makes it robust for complex tasks like atomistic structure search, where the true energy surface may be rugged and multimodal. However, BO does not guarantee finding the absolute global minimum with finite samples, as it often finds very good solutions with far fewer evaluations than random or grid search. This

effectiveness is essential for materials science applications: It has been demonstrated that, in comparison to brute-force scanning, BO reduces the number of costly simulations or parts of experiments by orders of magnitude. For example, Gutmann and Corander [64] demonstrated that BO could calibrate complex simulator-based models with several orders of magnitude fewer simulation runs by intelligently selecting informative points, therefore, highlighting the power of BO for expensive computational tasks.

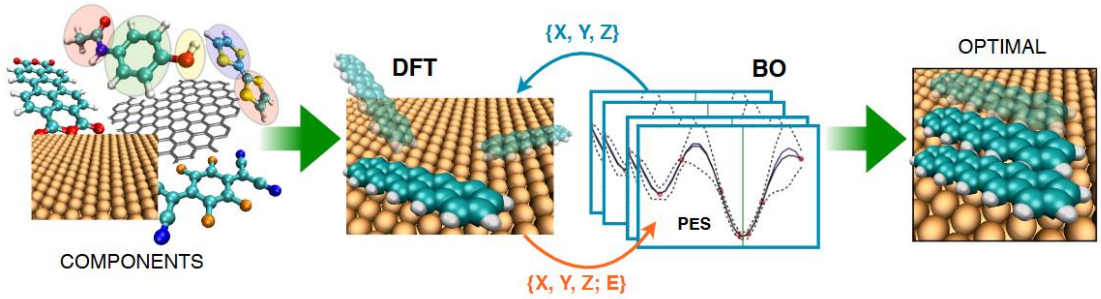

Figure 2: Schematic illustration of the Bayesian optimization loop, showing how the Gaussian Process surrogate model is updated after each sample, and how the acquisition function identifies the next query point [65]

2.3.2 Acquisition Function

An acquisition function guides the decision of where to sample next in Bayesian optimization. It is a mathematically defined utility function that the BO algorithm maximizes (for the next sample selection) using the current surrogate model [66]. The acquisition function is computed analytically or semi-analytically for any input x using the Gaussian Process predictions, and its role is critical to balance exploitation and exploration[67]. Exploitation means favoring inputs that the GP model currently predicts to have very good objective values. For instance, if the goal is to minimize energy, exploitation would sample near the lowest predicted energy found so far, refining the known minimum. This leverages the information gathered so far to converge on the most promising region. Exploration means favoring inputs about which the model is uncertain. These are regions not well sampled yet, where the function might conceivably yield even better results than the current best, and it helps avoid getting trapped in a local minimum by probing new areas of the design space

A good acquisition function achieves a trade-off: it tends to select points that either improve on the current best (exploitation) or greatly reduce uncertainty (exploration), or ideally both [68]. By adjusting this balance, the user can make the BO strategy more cautious in its exploration.

Several standard choices are used in BO literature, each with different ways of mixing exploration and exploitation, but only three will be discussed here. The first is Expected Improvement (EI), which is one of the most popular acquisition functions because it chooses points that maximize the expected improvement over the current best objective value. For a minimization problem with the current best (lowest) observed value indicated as f_{min}, the expected amount by which sampling at x would reduce the objective function, taking the expectation with respect to the GP's predictive distribution at x [69]. Intuitively, EI is high in regions where there is a significant probability of achieving a value better than the best-so-far. This can happen either because the GP mean, $\mu(x)$, is already low (suggesting x is likely good) or because of the uncertainty $\sigma(x)$ is large in meaning x could potentially yield a much better value than we have seen. EI naturally balances exploration and exploitation and is often a strong default choice. After each iteration, f_{min} is updated if a new best is found, and the EI landscape is recomputed for the next selection.

The second is the Lower Confidence Bound (LCB), this acquisition function picks the next point based on an adjustable confidence bound on the GP prediction, and for minimization problems, one defines the acquisition function, $a(x) = \mu(x) - \kappa, \sigma(x)$, where $\kappa > 0$ is a tunable parameter [70]. The algorithm then selects the point that minimizes this bound (equivalently, maximizes the negative of it). By choosing the lowest predicted lower bound on the objective function, LCB effectively targets points that could plausibly have very low true values, considering the GP's uncertainty $\sigma(x)$. The parameter κ controls the exploitation-exploration balance: a larger κ emphasizes the $\sigma(x)$ term, encouraging exploration (wider search), while a smaller κ relies more on the mean, leaning toward exploitation. LCB is simple and often used when one desires a more straightforward control over the exploration level.

The third is the Probability of Improvement (PI), which selects points based on the probability that they will improve over the current best. For minimization, $a(x) = P(f(x) < f_{min})$ given the GP distribution [71]. It is a more binary criterion than EI, which means that it doesn't account for the magnitude of improvement, only the chance of any improvement. PI can

sometimes be greedy by focusing on regions that are just barely likely to beat the best by any amount or too exploratory if adjusted, so it is less commonly used on its own. However, it introduces the concept of a user-defined threshold ξ to modulate aggressiveness (similar to κ in LCB). Other specialized acquisition functions exist (e.g. Thompson sampling, Knowledge Gradient, entropy search variants, etc.), but EI and LCB are most frequently used in practice and strike a good balance for many problems.

The choice of acquisition can be problem-dependent; for instance, EI is a good default that performs well in many settings, while LCB comes with guarantees on convergence in some cases. In materials science applications, one might experiment with both to see which finds optimal structures faster. The key point is that the acquisition function is computationally efficient to evaluate and can be globally optimized compared to an actual DFT calculation. Thus, we can afford an extensive search over the acquisition function to find the next best query point at every iteration. By dynamically steering the search through $a(x)$, BO efficiently explores the input space sampling where it matters most to uncover global optima with far fewer evaluations than conventional methods.

2.3.3 Gaussian Process, Kernel, and Hyperparameters

A Gaussian Process (GP) assumes the unknown objective function, $f(x)$, is a random function drawn from a Gaussian distribution specified in a function space. In practice, a GP is specified by a mean function and a covariance function (kernel). Given some observed data, the GP produces a posterior predictive distribution for $f(x)$ at any new point x, which is Gaussian function with a certain mean $\mu(x)$ and variance $\sigma^2(x)$ [72] [73]. This yields both a prediction and an uncertainty (standard deviation $\sigma(x)$) for each point. Because of this, GP are an ideal choice for BO. At each step, the GP's mean guides exploitation, and its variance guides exploration, and the implementation of GPs has become a standard tool for Bayesian modelling in machine learning and optimization.

The covariance or kernel function $k(x, x')$ is the core of a GP, encoding our assumptions about the function's behavior, such as smoothness and periodicity, by defining the covariance between $f(x) and f(x')$. In essence, $k(x, x')$ specifies how similar the outputs are expected to be at two input points x and x' based on their proximity in the input space [74]. If $k(x, x')$ is large, the GP assumes $f(x)$ and $f(x')$ should have similar values. A variety of kernel forms can be chosen, but a common default in materials optimization is the Radial Basis Function

(RBF) kernel (also called the squared-exponential or Gaussian kernel) due to its general-purpose smoothness assumption. The RBF kernel considers any small change in input to be significant only up to a characteristic length, beyond which points become essentially uncorrelated [75]. It is defined in equation 15 as:

$$K_{RBF}(x, x') = \exp\left(-\frac{|x-x'|^2}{2\ell^2}\right) \quad (15)$$

where ℓ is the length scale. The length scale ℓ controls the kernel's notion of distance: inputs closer than ℓ apart are considered highly correlated (the function does not change much over that range), whereas points far apart (relative to ℓ) have negligible correlation. In practical terms, ℓ determines the smoothness of the functions the GP can model, a small ℓ allows rapid variation and a large l enforces a smoother, slowly varying function. The parameter σ_f^2 is used to multiply the kernel function as it represents the variance of the function's output values; it sets the overall vertical scale of variation. If σ_f^2 is large, the GP expects the function values to vary widely; if small, the function is expected to stay near the mean. These two parameters ℓ and σ_f^2 are often called hyperparameters of the kernel, and they critically shape the surrogate's behavior.

Another commonly used kernel in materials problems is the periodic kernel, which is useful if the function is believed to be repeated with some period. For example, if one of the input variables is an angle or a composition fraction that has a natural periodicity, a periodic kernel can capture that behavior better than an RBF [75]. A standard periodic kernel can be defined in equation 16 as:

$$K_{per}(x, x') = \exp\left(-\frac{2sin^2\left(\frac{\pi|x-x'|}{p}\right)}{2\ell^2}\right) \quad (16)$$

where p is the period of repetition. This kernel ensures that $k(x, x')$ is high not only when x is close to x' (like RBF does), but also when x and x' are separated by an integer multiple of the period p (because $sin^2(\pi|x - x'|/p)$ will be zero if $|x - x'| = p, 2p, 3p$, etc.). The effect is that the GP's prior belief is that $f(x)$ is exactly repeating every p units in that input dimension. Such a kernel is appropriate, for instance, for modeling properties with respect to

crystallographic angle rotations or any scenario with inherent periodicity. In materials science, periodic kernels might be combined with RBF kernels if only some dimensions are periodic.

The choice of kernel reflects prior knowledge, that is, if we know the property of interest should be smooth and have no sharp discontinuities, an RBF is a safe choice. Also, if we suspect a periodic trend, we include a periodic component.

Hyperparameters and Learning: Kernel functions typically have a set of hyperparameters (such as ℓ, σ_f, p, etc.) that are not initially fixed. These hyperparameters can be learned from the data. In GP regression, a common approach is to maximize the marginal likelihood (also called the evidence) of the data with respect to these hyperparameters [76]. Intuitively, this finds the kernel parameters that make the observed data most probable under the GP model. Rasmussen and Williams [77] noted that hyperparameters such as the characteristic length-scale and variance can be inferred from data either by this Bayesian evidence maximization or by cross-validation. In practice, for BO, one often re-estimates the GP hyperparameters after each batch of new data is added, ensuring the surrogate remains a good fit. Software packages for BO automate this hyperparameter tuning. It is also possible to place initial values on hyperparameters or fix them based on previous knowledge of the computation run or literature values. For example, one might set a reasonable length-scale based on physical intuition of how far one must change a parameter to see a significant effect on the output.

2.3.4 Integration with DFT Calculations

Bayesian Optimization can be seamlessly integrated with DFT to accelerate atomistic structure search and materials discovery. In this context, the black-box function f(x) that BO optimizes is often the result of a DFT simulation, for example, f(x) might be the total energy of a particular atomic configuration, a formation enthalpy, or an electronic property of a material that we wish to optimized [57] [58]. The input x would represent a set of parameters defining a structure or material configuration. This could be a vector of continuous parameters (such as atomic positions, lattice constants, bond lengths/angles) and possibly discrete choices (such as atomic species or stoichiometry, though discrete choices often require special handling). The goal might be to find the lowest-energy structure (global minimum on the potential energy surface) or temperature, depending on the application. BO provides a strategy to conduct this search intelligently, using far fewer DFT calculations than conventional methods. The typical workflow to combine BO with DFT involves an outer loop

(the BO algorithm) that calls an inner loop (DFT simulation) for function evaluations. Practically, one can think of it as a closed-loop automated experiment, where each experiment is a DFT calculation guided by the BO logic.

First, one must define how to parameterize the atomic configuration or materials space for the GP model. For example, if searching for a crystal structure, one might parameterize by lattice parameters and a set of fractional coordinates of atoms in the unit cell. If searching for the adsorbed configuration of a molecule on a surface, x could encode the molecule's orientation and position on the surface. In some cases, a direct coordinate representation is high-dimensional. Techniques like using symmetry constraints, collective variables, or even fingerprint descriptors of structures may be used to reduce the dimensionality [60].

Secondly, one needs to generate an initial set of structures in the search space. These could be random structures or chosen via domain expertise (for example, known prototypes or low-energy structures for similar systems). Each structure is then evaluated using DFT to obtain the target quantity, such as the total energy. For structure optimization, DFT relaxations might be performed so that the energy corresponds to a locally minimized structure for each initial guess [78]. The results form the initial dataset for the BO algorithm. A GP model is initialized using the DFT results. For instance, the GP could model the formation energy as a function of structural parameters. If the structures have variable size or composition, one might need a fixed-length feature representation, such as using descriptors from materials informatics to feed them into a GP.

Now the Bayesian optimization loop begins. At each iteration, the acquisition function is computed based on the GP's current predictions, the BO algorithm finds the next candidate structure that maximizes the acquisition function. This selection might entail, for example, picking a structure that the GP predicts could have an even lower energy than those seen so far (exploitation) or a structure in a region of parameter space we have not sampled but which has high uncertainty (exploration). In practice, to propose a new structure, some algorithms directly optimize the continuous representation of the function. Others, especially if using a discrete set of candidate moves (which can be adding a molecule at a new site), may evaluate the acquisition function on a set of possible moves. Once the position is chosen, a DFT calculation is performed for that candidate position. At this point, one iteration is complete, and the process repeats: update acquisition function, select next structure, evaluate, and so on.

The loop continues until we have exhausted the iteration limit set for the DFT runs or until we are satisfied that we have found a good solution [79]. Because each iteration involves a full DFT calculation, which could take hours depending on the system size and computational resources, the number of iterations is usually limited to perhaps tens or a few hundred evaluations.

The best structure found across all iterations is taken as the result of the BO-guided search. Often, by the end of the run, the GP model is also a valuable by-product, as it provides an approximate model of the energy landscape, which can give insight into the problem, for example, revealing multiple low-energy basins or the relative importance of different structural parameters.

3 Methods

3.1 Bulk and Surface Model Preparation

The bulk InAs structure was prepared using the Atomic Simulation Environment (ASE), an open-source Python-based software designed for atomistic modelling. A bulk zinc-blende InAs unit cell, shown in Figure 3, was constructed employing the experimental lattice parameter of 6.06 Å [80]. Subsequently, this unit cell served as the basis to build a larger 4×2×1 surface supercell to ensure adequate representation of bulk interactions and periodic boundary conditions.

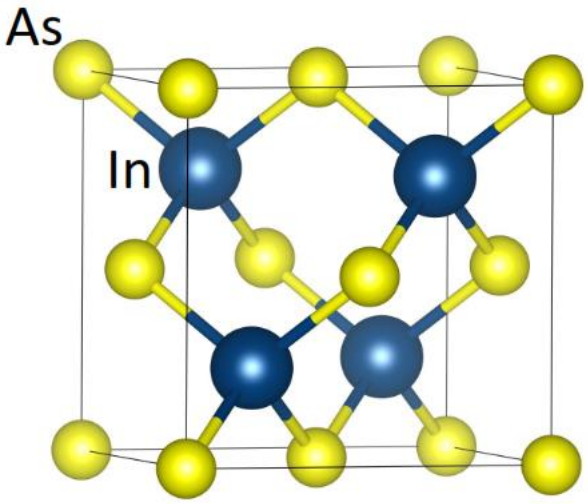

Figure 3: Atomic Crystal Structure of Bulk InAs

Upon establishing the bulk structure, the (100) crystal plane was cleaved utilizing ASE's built-in surface generation tools. However, the generated surface structure, shown in Figure 4, did not reflect known surface reconstructions from experimental or computational literature [26] [81] [82], specifically lacking the characteristic surface dimers and vacancy rows as shown in Figure 5. Therefore, manual reconstruction was performed based on previously reported atomic coordinates for ζ(4×2) [26]. Surface and subsurface atoms were repositioned into dimers aligned along the [01$\bar{1}$] crystallographic direction, and vacancy rows were introduced accordingly to represent the established ζ(4×2) reconstruction. This reconstructed surface geometry was essential for obtaining accurate adsorption properties in subsequent

calculations. In this work, the arsenic-terminated InAs (100) surface, under indium-rich conditions, was studied because it is representative of In-rich surface preparations and has been observed as a stable structure for InAs(100) [83]. The "ζ(4×2)" designation indicates a surface unit cell that is four times the bulk unit cell length in one in-plane direction and two times in the perpendicular in-plane direction, reflecting the long-range ordering of surface atoms in the reconstructed structure.

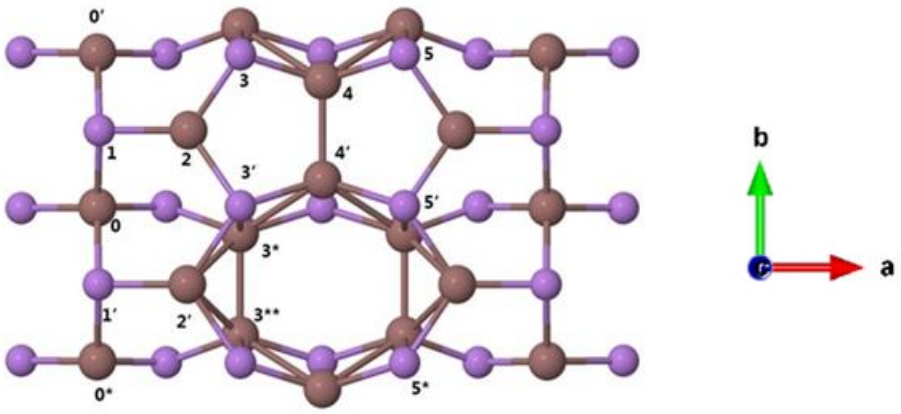

Figure 4: Top view of surface model of InAs(100) surface. The brown colour represents the In atoms, and the purple colour represents the As atoms

The As-terminated slab bottom was constructed by ensuring the topmost atomic layer consists of indium atoms arranged in the experimentally observed ζ(4×2) pattern. This involved replicating the conventional InAs unit cell to create a surface supercell of the desired size and then arranging the indium and arsenic atoms in the top four layers according to the proposed ζ(4×2) reconstruction geometry as shown in Figure 6. This preparation yielded a starting atomic model for InAs(100)-ζ(4×2) that was used in subsequent slab calculations and surface optimizations.

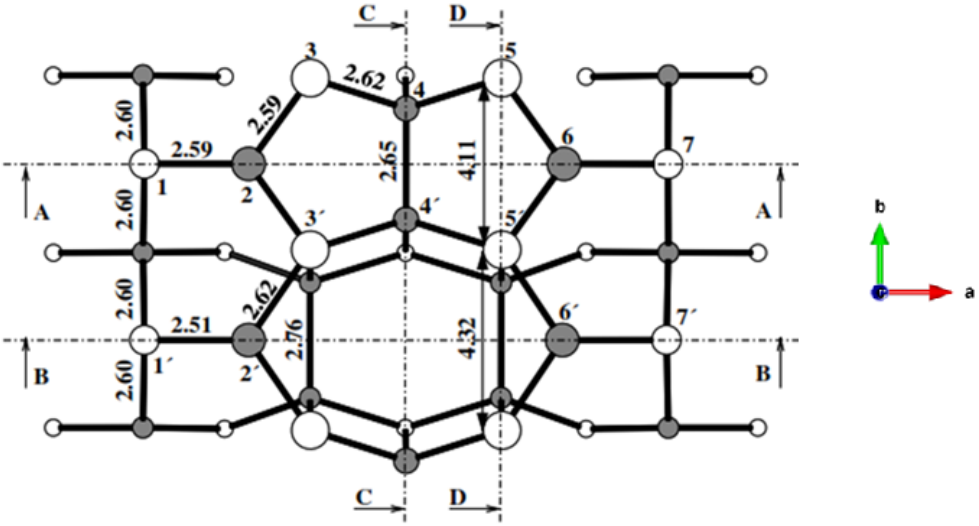

Figure 5: Top view of first four layers of InAs(100) ζ(4×2) surface [26]

3.2 Surface Slab Construction

The reconstructed surface cell was used to create slabs with varying thicknesses of 8, 10, 12, 14, and 16 atomic layers. The slabs contained layers of InAs stacked along the [100] direction as shown in Figure 6 (for 10 layers). This corresponds to stacks of indium layers and arsenic layers in alternating order, providing sufficient thickness so that the middle of the slab is bulk-like. A vacuum region of about 45 Å was added above the slabs to separate periodic images and prevent interaction between the slabs and their periodic duplicate in the out-of-plane direction. Also, this ensures negligible spurious interaction after any surface relaxation.

The lateral dimensions of the supercell were set to reproduce the ζ(4×2)surface periodicity, and the size of the surface unit cell is the lattice constant / $\sqrt{2}$ which corresponds to 4.307 Å. Consequently, this implies that the size of the 4 ×2 surface unit cell is 17.228 × 8.614 Å in the x and y directions, respectively.

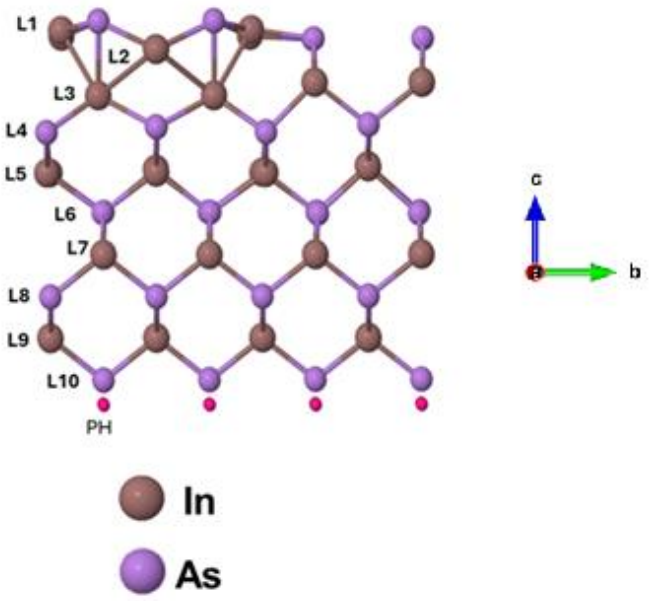

Figure 6: Side view of 10 layers of InAs(100) ζ(4×2)surface. L1 to L10 indicates 1st layer to 10th layer respectively.

The ζ(4×2) reconstruction may involve slight lateral relaxations; that is, one of the reasons we created different numbers of atomic layers to check for convergence with the atomic distances and bond angles. Any dangling bonds at the surface were left as is on the top side to allow for reconstruction during relaxation, while on the bottom side, the termination was handled to avoid introducing a large dipole moment by passivating the arsenic end with two pseudo-hydrogen atoms per As atom. The position of the pseudo-hydrogen atoms was chosen by fixing all other atoms and allowing only the pseudo-hydrogen atoms to move to their equilibrium position. The reason for the choice of pseudo-hydrogens is to stoichiometrically balance the terminated As atoms in the systems, and this method also agrees with the previous study [83] [84]. The passivation ensures that the bottom few atomic layers are kept in their bulk positions and the slab can accurately represent a single-terminated surface without spurious electrostatic interactions.

In each slab model, the bottom (last 4) layers were constrained, including the pseudo hydrogen atoms, and only the upper layers were allowed to relax. Constraining the movements of a portion of the slab in this way provides mechanical stability and mimics the presence of a bulk continuum beneath the surface while still permitting the surface region to reconstruct and relieve stress. This approach of using a thick slab with a frozen bottom region is standard for surfaces and ensures that the computed structure and energy are converged with respect to slab thickness. The slab constructions and atomic position modifications were

performed using established ASE tools in the Python library. The resulting InAs(100)-ζ(4×2) slab models served as the starting point for subsequent geometry optimizations and adsorption calculations.

3.3 Geometry Optimization

With the slab models prepared, computation of the relaxed atomic structure of the clean InAs(100)-ζ(4×2) surface through geometry optimization was performed. Geometry optimization was carried out using DFT-based total energy minimization. The atomic positions were iteratively adjusted to reduce forces, using a self-consistent cycle(SCF) until convergence criteria were met. During this process, the ζ(4×2) periodic boundary conditions in the surface plane were maintained for each slab, ensuring that the reconstruction features were consistently represented. This approach effectively models the surface as a continuous, repeating pattern, mimicking an infinitely extended and well-ordered structure. By preserving these periodic conditions, the simulation accurately captures the long-range ordering characteristic of the reconstructed surface. After optimization, the structural changes, bond length, and bond angles were analyzed to check for convergence of the slab surface. The relaxed, converged layer structure will provide a baseline for assessing stability and for comparison with the oxidized cases. The finally optimized structure is also compared with the other models [68][84] to check for agreement and discrepancies.

All calculations for geometry optimizations were performed with DFT. The generalized gradient approximation (GGA) in the Perdew–Burke–Ernzerhof (PBEsol) was employed for the exchange-correlation functional. A Pulay mixer was used to ensure stable convergence of the electronic self-consistency cycle. Table 1 gives a summary of the computation details.

The Geometry relaxation of the different slab models was carried out using the trust radius minimization (TRM) algorithm with a final atomic force threshold of 1×10^{-2} eV/Å. Brillouin zone integration was carried out using a Monkhorst-Pack k-grid of 2×4×1, chosen to reflect the anisotropic periodicity of the InAs(100)-ζ(4×2) surface [82]. This k-point grid provides sufficient sampling in the surface plane while minimizing unnecessary computations along the non-periodic surface. These settings were tested to ensure convergence of total energies and forces, and they were consistently applied across all stages of the study, including clean surface relaxation, oxygen adsorption tests, and Bayesian optimization search. The selected parameters ensured that the surface geometries and adsorption energies were computed with high reliability and accuracy.

Table 1: Computation Details for Geometry Optimization

Item	Details
Atomic Layers	8,10,12,14,16
Bottom Passivation	2 PH per As atom
Fixed Layers	4 bottom layer + PH
Vacuum	45 Å
Cell Size	17.228 × 8.614 × 100 Å
XC Functional	PBEsol
Tier	Light tier 1
k-grid	2×4×1
Geometry Optimization	trm algorithm, 10^{-2} eV/Å
DFT	FHI-aims code
HPC Computing	Mahti
Number of Nodes	1 (128CPUs)

3.4 Oxygen Adsorption Test

An initial test of oxygen adsorption was conducted to determine the optimal slab thickness for further adsorption studies. A single oxygen atom was placed at a fixed height of 2.0 Å above the surface, shown in Figure 7, over the highest site of the reconstructed slabs. Geometry optimization was performed, allowing the oxygen atom to freely move and bind with the surface atoms. This distance (2 Å) was chosen as a physically reasonable starting point, close enough to potentially form a bond during relaxation but not so close as to cause excessive repulsion at the outset.

All degrees of freedom for the O atom were allowed (it was free to move in all directions), and the top surface layers of InAs were again allowed to relax, while the bottom layers remained fixed as before. The convergence criteria were the same as for the clean surface relaxation. During the optimization, the oxygen atom was observed to move from its initial position as it bonded to the surface. By comparing the total energy of the relaxed O/InAs system to that of the clean surface (plus an isolated O atom reference), we can evaluate the adsorption energy with the relationship in equation 17.

$$E_{Ads} = E_{InAs+O} - (E_{InAs} + E_O) \qquad (17)$$

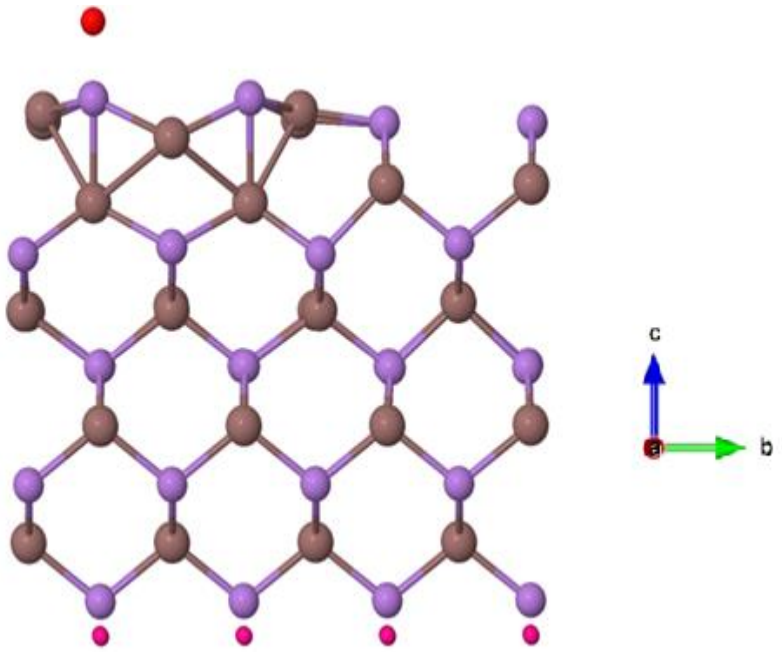

Figure 7: Placing the Oxygen atom exactly 2Å above the surface. The brown atom is In, purple atom As, pink atom PH, and red atom O.

Where E_{Ads} is the absorption energy, E_{InAs+O} is the total energy of the InAs slab with Oxygen adatom, E_{InAs} is the total energy of the InAs slab, and E_O is the total energy of the Oxygen adatom. The reason for this single-atom adsorption test was to ensure convergence of the absorption energy as the slab thickness increases, and the possibility of bonding with the In and As atoms at the surface set the stage for the more extensive structure search using BOSS, verifying that our computational setup can handle oxygen adsorption.

3.5 BOSS implementation for oxygen adsorption sites

To efficiently explore potential adsorption sites for oxygen atoms on the reconstructed InAs(100) surface, BOSS was applied to find the preferred binding site and configuration of an oxygen atom on the InAs(100)-ζ(4×2) surface. We defined the search space in terms of a few key structural parameters describing the O adatom position relative to the surface. Specifically, we considered the oxygen's vertical distance above the surface and its lateral position within the 4×2 unit cell as the parameters to optimize. The underlying idea is that the adsorption energy of the system will be treated as an objective function of the oxygen atom position. Then, BOSS can find the combination of oxygen atom coordinates that minimizes the objective function. BOSS is used here purely as a structural search tool; the total energy of the system is recalculated for each candidate point by DFT. The eLCB method acquisition

function, which was used in this study, balances exploration and exploitation, it selects points that minimize a combination of the predicted energy and uncertainty.

Two BOSS runs were carried out with increasing complexity: a one-dimensional search to optimize the oxygen height above a fixed site (1D BOSS), and a two-dimensional search to find the optimal lateral position on the surface (2D BOSS). In both cases, the surrogate model was updated iteratively with new DFT single-point calculations as data points. The surface slab was kept fixed in its clean, relaxed structure during the BOSS searches to define a static potential energy surface for the oxygen adatom; after the BOSS-identified minimum was found, a full relaxation of the O-surface system was performed to confirm the stability of that configuration.

The one-dimensional BOSS search was designed to determine the optimal vertical distance of a single oxygen atom above the InAs surface, assuming a particular lateral position. For this run, we fixed the lateral position of O directly above the highest surface atom based on the initial test done and allowed only the O height to vary. The BOSS algorithm was configured with the settings, as shown in Table 2, for the 1D search. The result of the 1D BOSS search was an efficiently sampled absorption energy curve from which we could extract the optimal height for an oxygen atom above an indium site. This optimal height serves as a guideline for the input in the 2D search.

Table 2: 1D BOSS Computation Details

Item	Details
Objective Function	Adsorption Energy
Bounds	1 – 5 Å
kernel	Radial basis function(rbf)
Acquisition function	Expected Lower Confidence Bound (eLCB)
Vacuum	45 Å
Initial sampling points	5
Subsequent iterations	25

With the vertical position fixed at the optimal height determined from the 1D BOSS search, 2D lateral search was then carried out to fully map the adsorption energy landscape across the reconstructed unit cell. Fractional coordinates spanning [0,1]×[0,1] across the entire periodic cell defined the search domain, with the details given in Table 3.

Table 3: 2D BOSS Computation Details

Item	Details
Objective Function	Adsorption Energy
Bounds	[0,1]×[0,1] - Fractional
kernel	standard periodic (stdp)
Acquisition function	Expected Lower Confidence Bound (eLCB)
Vacuum	45 Å
Initial sampling points	5
Subsequent iterations	300

Five initial sampling points were again generated with a Sobol sequence, followed by 300 iterative BO samples directed by the eLCB acquisition function. Each point corresponded to a constrained (in z-axis) DFT calculation in which only the lateral position of the oxygen atom was allowed to relax slightly, capturing realistic lateral relaxation effects. The search proceeded until the global minimum energy position and other significant local minima were robustly identified, achieving convergence within 0.01 eV.

The output of this search was effectively a surrogate energy map of the O adsorption energy across the surface. From this map, we could identify the coordinate (x,y) that gave the lowest predicted energy, the likely global minimum adsorption site. This information is useful for understanding where oxygen prefers to bind on the clean InAs surface. This comprehensive BOSS analysis thus provided reliable predictions of energetically favorable oxygen adsorption sites to which we can calculate the adsorption energy, significantly streamlining the otherwise prohibitively expensive exploration of the configurational space using conventional computational methods. After completing the BOSS searches, a thorough post-analysis was performed to validate and interpret the results. The main results from BOSS included the predicted optimal oxygen adsorption configurations: one determined from the 1D search, identifying the optimal adsorption height, and another from the 2D search, which captured both the optimal lateral position and implicitly the corresponding height.

3.6 Relaxation of the oxidized model

After determining the optimal oxygen adsorption site from BOSS, the optimized reconstructed InAs(100) surface slab was used as the initial configuration for the detailed

relaxation of the oxidized surface. A single oxygen atom was placed precisely at the energetically favored adsorption site determined by the previous BOSS analysis. This initial oxidized configuration was subsequently subjected to geometry relaxation through density functional theory (DFT) calculations. During this relaxation procedure, atomic positions within the slab, including the adsorbed oxygen atom, surface atoms, and subsurface layers, were permitted to move freely in order to accurately capture local atomic distortions induced by oxygen adsorption. The computational settings and numerical parameters employed in these oxidized-surface relaxations were identical to those utilized previously in the clean InAs surface relaxations.

Using the same convergence parameters ensured direct comparability between clean and oxidized surface calculations, enhancing the reliability and interpretability of the resulting data. Allowing all slab atoms to relax during the oxidized-surface simulations was critical for capturing the realistic physical behavior associated with oxygen adsorption. Oxygen atoms are known to interact strongly with the InAs surface atoms [31], potentially inducing considerable local lattice distortions, bond rearrangements, or structural strain in the immediate vicinity of the adsorption site. The fully relaxed single-oxygen atom configuration guaranteed that the computed adsorption geometry accurately reflected the local environment and atomic-scale interactions, and it serves as a critical baseline for future oxidation studies. Typically, oxide formation on semiconductor surfaces involves multiple sequential oxygen adsorption steps, each affecting the electronic and geometric structure differently. Beginning with a single-oxygen fully relaxed state as a reference provides a controlled starting point from which subsequent oxygen atoms can be added iteratively.

In future studies, each subsequent oxygen addition would similarly be guided by BOSS analysis, with the structure fully relaxed again after each oxygen incorporation. By using this iterative approach, it becomes possible to systematically isolate and quantify incremental changes to surface energetics, geometry, and electronic properties. This approach also significantly improves computational efficiency and accuracy by allowing each additional oxidation step to start from a physically meaningful and energetically stable configuration, thereby minimizing the use of computational resources.

4 Results and Discussion

4.1 Surface Reconstruction

Figure 8 shows the top and side views of the reconstructed surface slab of InAs(100)-ζ(4×2) before geometry optimization. The top-view comparison between the model and the reference literature [26] structure reveals a close visual match in terms of symmetry, dimer orientation, and periodicity. The indium dimer pairs are separated by the dimer row and spaced apart across the trough, and the arsenic atoms in the top layer occupy positions that complete the surface network, filling the troughs beside the indium rows in line with the reported (4×2) periodicity. This comparison confirms that the pre-optimization model used in this study accurately replicates the ζ(4×2) surface configuration. It serves as a reliable starting point for further structural optimization and oxygen adsorption studies, ensuring that the simulation begins from a realistic and experimentally supported surface geometry.

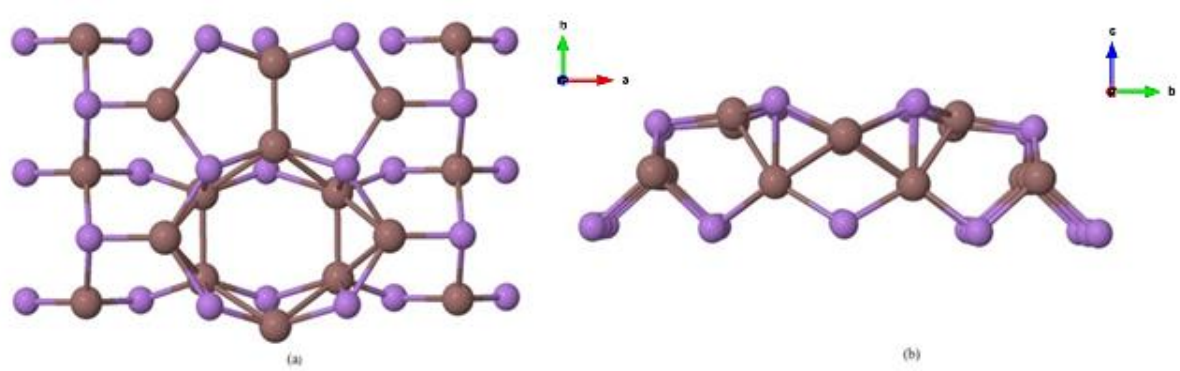

Figure 8: Top view (a) and Side view (b) of reconstructed InAs(100)- ζ (4×2) surface. The brown atom indicates indium and the purple colour indicates As.

4.2 Slab Optimization

4.2.1 Bond Length and Bond Angles

The structural features of the surface were relaxed, and optimized surfaces were compared to assess the convergence of surface bond lengths and bond angles with respect to slab thickness. Table 4 shows the comparison of the number of layers for each slab thickness with the simulation time in seconds, which indicates that there is an increase in computational time as the number of atoms increases, with 12 layers as the balance. Figure 9 provides a top view of the first four layers of the surface atomic arrangement for slabs of different thicknesses after

geometry optimization, offering a visual comparison of the surface geometry as the slab becomes thicker.

Table 4: Comparison of the number of layers and simulation time (s)

Layers	Total Atoms	Time (s)
8	78	130.201
10	94	206.4608
12	110	326.3269
14	126	463.4872
16	142	631.4403

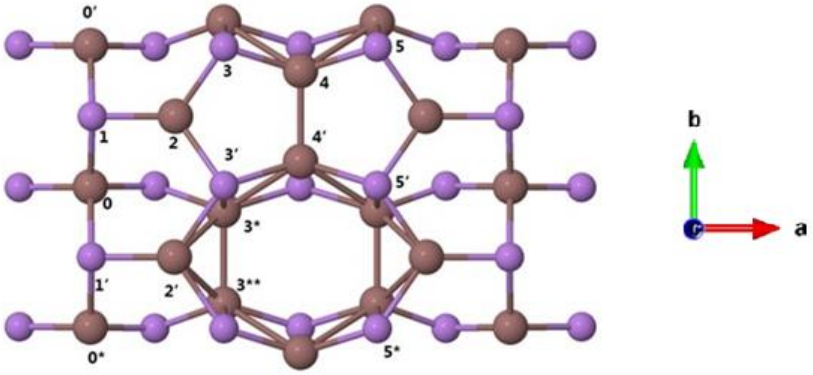

Figure 9: Top view of first four-layer view top view of the optimized slab

To quantify the structural convergence, the bond lengths and bond angles within the first four atomic layers from the surface are focused on. By examining the values of these bond lengths and key bond angles for each slab thickness, we can judge how the surface geometry evolves and at what point it stabilizes as the slab thickness increases. The surface bond lengths show a clear trend toward convergence as the slab thickness increases. Table 5 summarizes the key bond lengths for each slab, highlighting the surface In–As bond between an In atom and its nearest As neighbor (indicated as d0'1, d01', d1'2', d2'3', d12, d23, d34), As-As interatomic distance (indicated as d55', d5'5*) and the In–In dimer bond between two adjacent In atoms in the first layer and the fourth layer (indicated as 44’ and 3*3**). The bolded numbers on each slab show the point of convergence. The 8-layer slab exhibits the most pronounced deviation in some of the distance measurements in these bond lengths compared to thicker

slabs; for example, surface In–As (d01') bond is slightly longer than those in the 12-layer and thicker models, indicating an overly contracted surface due to the insufficient slab thickness.

Table 5: Bond length for the different slab thicknesses

Layers		8	10	12	14	16
Bond Length (Å)	d0'1	**2.66**	2.66	2.66	2.66	2.66
	d01'	2.68	**2.67**	2.67	2.67	2.67
	d1'2'	2.57	**2.56**	2.56	2.56	2.56
	d2'3'	**2.68**	2.68	2.68	2.68	2.68
	d12	**2.55**	2.55	2.55	2.55	2.55
	d23	2.66	**2.65**	2.65	2.65	2.65
	d34	2.67	**2.67**	2.67	2.66	2.67
	d44'	2.78	2.77	**2.78**	2.78	2.78
	d55'	**4.24**	4.24	4.24	4.24	4.24
	d5'5*	4.32	**4.38**	4.38	4.38	4.38
	d3*3**	2.76	2.86	**2.85**	2.85	2.85

By 10 layers, the surface In–As and In–In bond lengths are much closer to the converged values, differing only marginally from those of the thicker slabs. The 12-layer slab yields surface bond lengths essentially identical to those of the 14- and 16-layer slabs, demonstrating that the surface bonds have fully converged by 12 layers. Similarly, the bond angles in the near-surface region exhibit a convergent behavior with increasing slab thickness. Table 6 compares some selected bond angles involving surface and sub-surface atoms for the different layer models. One representative example is the As–In–As (indicated as 0-1-0') bond angle formed at a second-layer arsenic atom bonded to two indium atoms in the third layer below. In the 8-layer slab, this angle is noticeably different from that in the thicker slabs, reflecting the altered bonding environment and enhanced relaxation in the thinner slab. By 10 layers, the As–In–As angle moves closer to the value observed in the thicker slabs, and by 12 layers, it exactly matches the angle found in the 14 and 16 layer cases which also confirms that the geometrical configuration of the surface (in terms of bond orientations) has fully stabilized by the 12-layer slab thickness.

Table 6: Comparison of selected bond angles for the different slabs

Layers		8	10	12	14	16
Bond Angle (˚)	0-1-0'	108.9	108.6	**108.7**	108.7	108.7
	0-1-2	93.9	**92.8**	92.8	92.8	92.8
	1-2-3	**124.8**	124.8	124.8	124.8	124.8
	2-3-4	100.8	**101**	101	101	101
	3-4-4'	105.8	**105.9**	105.9	105.9	105.9

The interatomic distances and angles in the 12-layer relaxed surface also match well with literature values, as shown in Tables 7 and 8. In particular, the bond length between surface indium and arsenic atoms lies in the range 2.56–2.68 Å in our 12-layer slab model, falls within the reported range of 2.51–2.74 Å for this surface. The study that uses GGA for its calculation overestimated the bond length, the study that used LDA underestimated the bond length, and the results obtained from this study lie between the two in terms of bond length. Also, the local bonding geometry remains nearly close to the reported range of 105–126°, only slightly different from the angle obtained from my model, which is 105.9–124.8°. The good agreement in both bond lengths and angles validates our relaxed slab as an accurate representation of the InAs (100)- ζ (4×2) surface.

Table 7: Comparison of the Bond length of the 12 layer slab with the literature

		Layer 12 (PBEsol)	GGA [84]	LDA [26]
Bond Length (Å)	d0'1	2.66		2.60
	d01'	2.67		2.60
	d1'2'	2.56	2.60	2.51
	d2'3'	2.68	2.74	2.62
	d12	2.55	2.59	2.59
	d23	2.65	2.70	2.59
	d34	2.67		2.62
	d44'	2.78	2.82	2.65
	d55'	4.24	4.32	4.11
	d5'5*	4.38	4.45	4.32
	d3*3**	2.85	2.89	2.76

Table 8: Comparison of bond angle of the top layer surface slab with literature

		Layer 12 (PBEsol)	LDA [26]
Bond Angle (°)	0-1-0'	108.7	105-126
	0-1-2	92.8	
	1-2-3	124.8	
	2-3-4	101	
	3-4-4'	105.9	

Therefore, a 12-layer slab provides an optimal balance between accuracy and computational cost.

4.2.2 Oxygen Adsorption Test

As a further check on slab thickness adequacy, the calculation of oxygen adsorption energy on each slab was computed as given in Equation 17. Figure 10 shows the computed oxygen adsorption energy as a function of slab thickness.

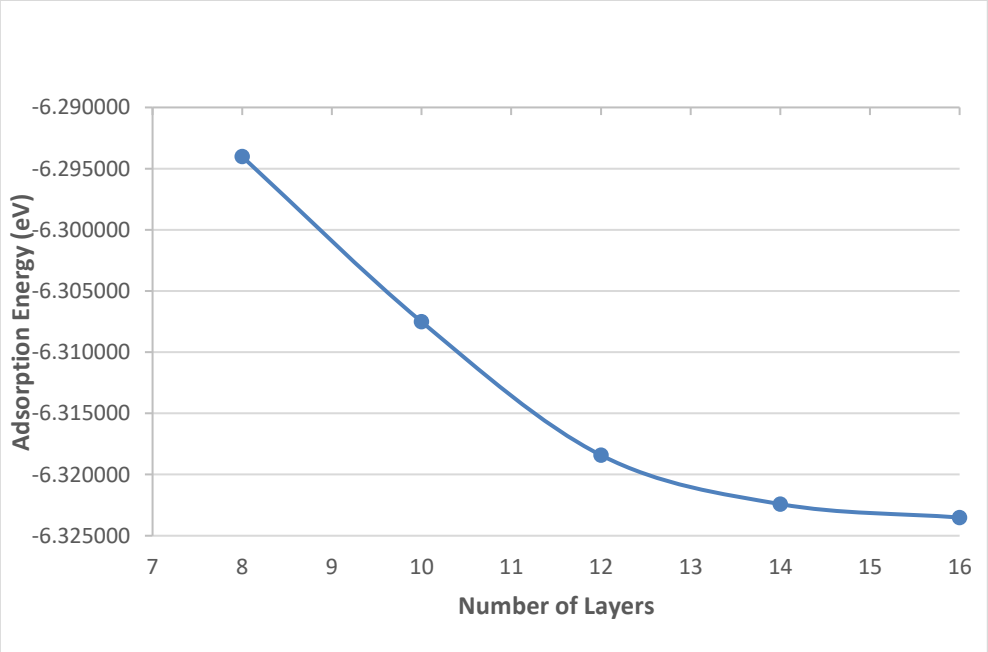

Figure 10: Oxygen adsorption energy vs. slab thickness. The adsorption energy of a single O atom (at 2 Å above a surface In site) is plotted for slab thicknesses of 8, 10, 12, 14, and 16 layers

The trend shows that the adsorption energy converges with increasing slab thickness; the 8-layer and 10-layer slabs yield slightly different adsorption energies, whereas the values level off by 12 layers. The test adsorption energy on the 8-layer slab is about - 6.29 eV, while on

the 12-layer slab, it was - 6.32 eV, and increasing to 14 and 16 layers also gives similar adsorption energy when evaluated to two decimal places. Thus, the 12-layer slab gave results within the convergence based on the absorption energy test. This confirms that the choice of a 12-layer slab is sufficient to obtain converged adsorption energetics. The 12-layer InAs(100)-ζ (4×2) model is thus adopted for all subsequent simulations in this work, ensuring that surface properties are well-converged without incurring unnecessary computational expense. A thinner slab would have led to noticeable errors in predicted adsorption energies, underscoring the importance of verifying slab thickness in surface studies.

4.3 Predicted Oxygen Adsorption Region

4.3.1 Global height

To obtain the global height to place the Oxygen atom, a 1D Bayesian optimization (BOSS) run was performed to refine the minimum vertical coordinate of an oxygen adatom above the reconstructed surface. Two laterally distinct positions were selected, indicated as site 1 and site 2, as shown in Figure 11. Site 1 is located directly above the topmost atom on the surface, and site 2, positioned above an empty space between two indium atoms in the unit cell. In each case, the oxygen atom was initially placed above the topmost layer, after which the BOSS routine was allowed to vary only the z-height while the lateral coordinates remained fixed. Figures 12 and 13 present surrogate-model snapshots after five and twenty BOSS evaluations at site 1 and site 2, respectively. The horizontal axis x_1, denotes the vertical O surface distance (Å), and the vertical axis y, represents the predicted adsorption energy. Five graphic elements are of interest: GP mean, $\mu(x)$, is the blue curve, the grey shading is the region of uncertainty, evaluated DFT points is the red circles (acq), current GP minimum is the red vertical line (x_{glmin}), next point proposed by the acquisition function is the green dashed line (x_{next}). After five evaluations, the GP still carried considerable uncertainty away from the sampled region, which is shown as the grey envelope broadens markedly beyond 3.0 Å. After twenty evaluations, the GP mean has converged to a single deep minimum indicated by the red line. The surrounding grey band is now extremely narrow across indicating that additional samples would change the predicted energy in this region by only a few meV. Most new evaluations (red circles) have clustered around the emerging minimum, reflecting the eLCB acquisition function's gradual shift from exploration to exploitation. An identical analysis for Site 2, shown in Figure 13, yielded a slightly higher minimum value

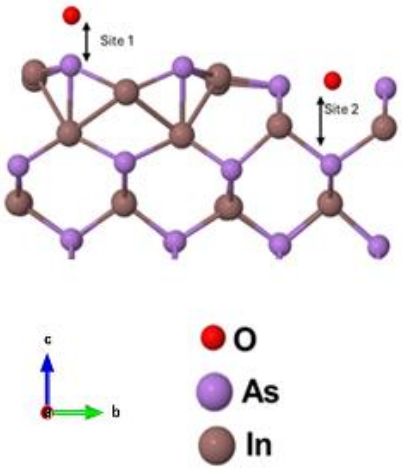

Figure 11: Top 6 layers of InAs(100)-ζ(4x2) showing the two sites the Oxygen adatom was placed for the global height search

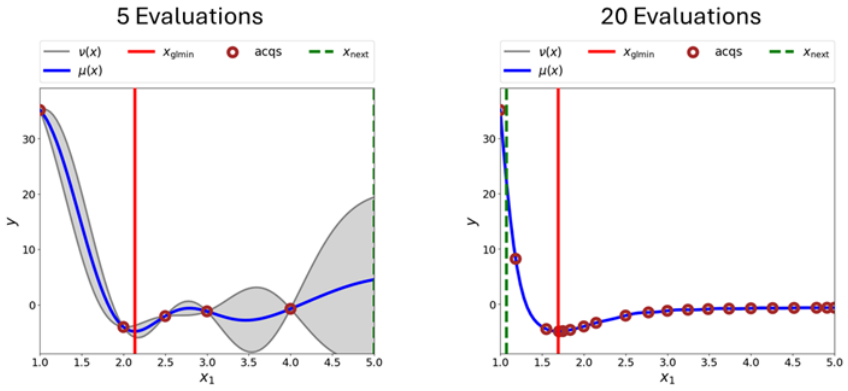

Figure 12: Adsorption energy of a single O atom at site 1 vs. height above a surface on InAs(100)-ζ(4×2). Red circles show the samled points chosen by the BOSS algorithm; the red line shows the predicted global minimum, the green dot line shows the next point to be sampled

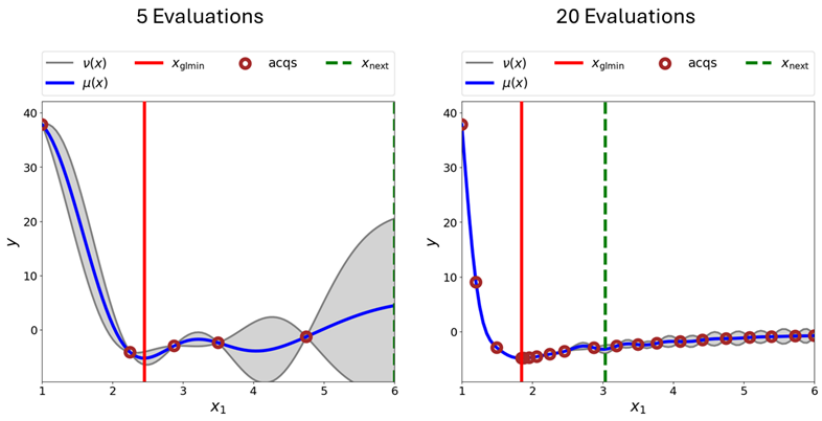

Figure 13: Adsorption energy of a single O atom at site 2 vs. height above a surface on InAs(100)-ζ(4×2). Red circles show the sampled points chosen by the BOSS algorithm; the red line shows the predicted global minimum, the green dot line shows the next point to be sampled

Figure 14 shows the acquisition model of the BOSS run for site 1 and site 2. The upper panels are the plot of adsorption energy predicted by the GP surrogate at the current global minimum against the number of iterations on the horizontal axis. The lower panels track the sampling history against the number of iterations in the horizontal axis, blue circles denote the heights selected for the DFT evaluation, while the solid blue line traces the running estimate of the global minimum height. In the exploration phase, both sites exhibit large vertical scatter in the evaluated heights, reflecting the eLCB acquisition function's deliberate sampling of high-uncertainty regions. As a consequence, the surrogate's energy predictions fluctuate markedly, and the global-minimum height drifts from its initial guess. In the exploitation phase, evaluations cluster in a narrow height window since the surrogate has identified a promising energy well.

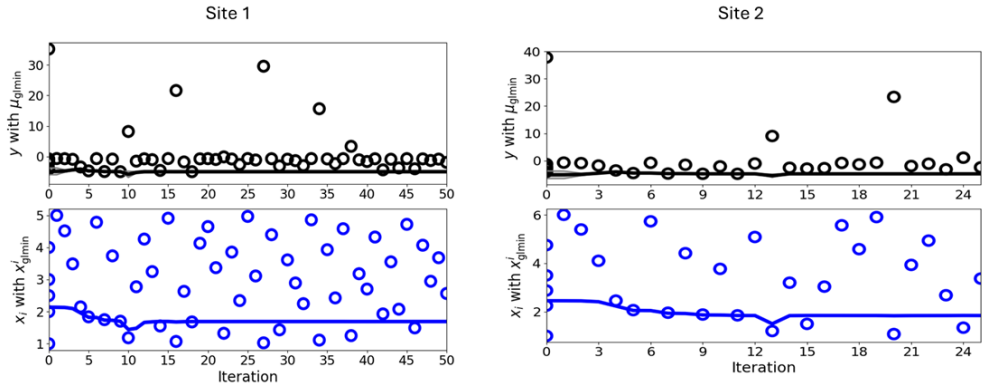

Figure 14: Data Acquisition plot of the 1D Model

For Site 1, sampling densities are between 1.6 Å and 1.8 Å, and for Site 2, between 1.8 Å and 2.0 Å. Further evaluations refine the minimum, and both the predicted energy and height begin to stabilise. The flat plateau regions beyond iteration 15 on both curves indicate that additional evaluations no longer shift the minimum height within the resolution of the GP model, signifying practical convergence of the 1D search. Figure 15 presents the evolution of $x_{gl\min}$ with the iteration number for both adsorption sites. A curve is not flat at the initial stage of the iteration because the eLCB occasionally proposes points that reveal a deeper minimum farther from the current estimate, causing sudden downward jumps. At site 1, the global minimum($x_{gl\min)}$ height decreases rapidly from 2.12 Å to 1.75 Å within the first eight iterations. A pronounced dip to 1.45 Å at iteration 10 corresponds to a single exploratory sample that proved slightly over-optimistic; by iteration 13, the estimate rebounds to 1.69 Å and thereafter fluctuates by less than 0.01 Å. At site 2, a similar descent is observed as global

minimum (x_{glmin}) falls from 2.45 Å to 1.85 Å by iteration 12, followed by a brief overshoot to 1.55 Å; the value ultimately stabilises at 1.84 Å from the surface atom at the valley position at iteration 15 onwards. The plateau regions beyond iteration 15 on both curves indicate that additional evaluations no longer shift the minimum height within the resolution of the GP model, signifying practical convergence of the 1D search.

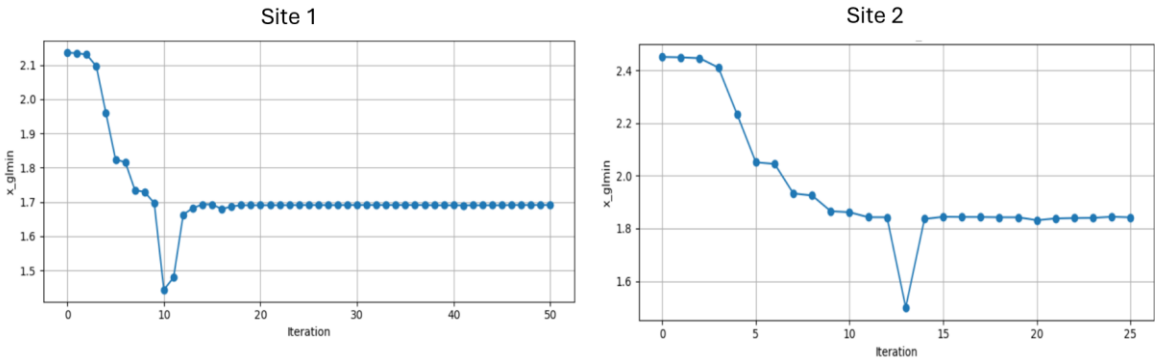

Figure 15: Convergence tracking of the Global minimum value vs adsorption energy

To maintain a single, transferable vertical coordinate for the subsequent two-dimensional lateral search, a compromise height of 1.50 Å above the highest surface atom was adopted. At this distance, it ensures balance between site 1 and site 2, which yields negative adsorption energies at both sites. The oxygen adatom is still securely within the attractive well for both sites, ensuring that lateral moves will probe adsorption configurations rather than risk desorption while avoiding repeated optimisation of the height coordinate.

4.3.2 Lateral Direction

Figure 16 compares the GP adsorption-energy maps obtained after 20, 100, and 150 BOSS evaluations from left to right with identical colour scales. Each heat-map depicts the adsorption energy $E(x,y)$ across the periodic 16.6 Å × 8.8 Å surface unit cell; cooler colours signify more exothermic adsorption energy. At only twenty DFT samples, the map remains comparatively coarse, a broad purple colour indicated as 1 on the graph signals the algorithm's early identification of a promising low-energy basin. Elsewhere, large teal and green regions indicate where uncertainty dominates; the GP has not yet sampled those areas, so predicted energies are relatively shallow. By one hundred evaluations, the surface shows markedly higher spatial resolution. Four symmetry-related rings of deep adsorption energy

(purple crescents indicated by 1-4) encircle the cell's centre, reflecting the mirror operations of the ζ(4 × 2) lattice. Although small areas of teal persist along the edges, the overall landscape is stabilizing, and the most exothermic region now lies in the ring, with the energy minimum having deepened to -5.0 eV. The growing similarity between symmetry replicas confirms that the periodic kernel is being informed by sufficient data. After 150 evaluations, little qualitative change is seen relative to the 100 point map. Outside the purple wells, the colour gradient is now smooth, indicating that the GP variance is not changing(or changes by few meV) across the entire unit cell. Subsequent sampling campaigns of 151 to 305 evaluations reproduce this topology almost exactly. Thus, the 150 evaluation map is deemed converged for all practical purposes.

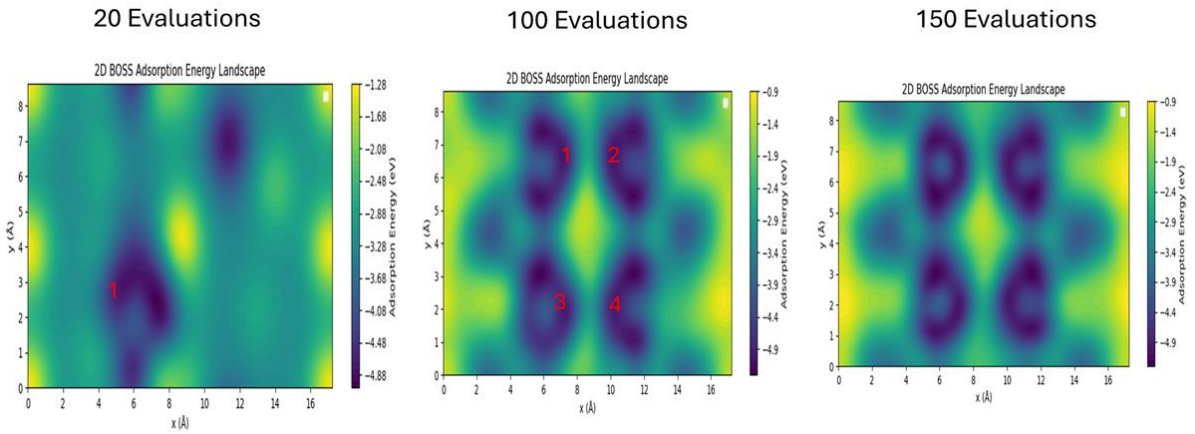

Figure 16: Evolution of Energy Landscape from the 2D BOSS run

Figure 17 shows the data-acquisition sequence, the upper panels summarize the difference between each newly acquired energy value (open circles) and the current best GP prediction ($\mu_{g\text{lmin}}$), which is shown as the solid black line. The early iterations probe a wide energy range, whereas by approximately 150 iterations, nearly all samples lie within 0.001 eV of the best-known minimum, indicating that the exploration has shifted decisively toward local refinement. In the lower panel of Figure 17, the sampled fractional coordinates are superimposed on the running estimate of the global minimum position, both coordinates oscillate among symmetry-related zones until iteration approximately reaches 150, after which the prediction stabilises. No further significant coordinate drift was recorded over the

final 150 evaluations, which is clearly shown in Figure 18, demonstrating practical convergence of the lateral search.

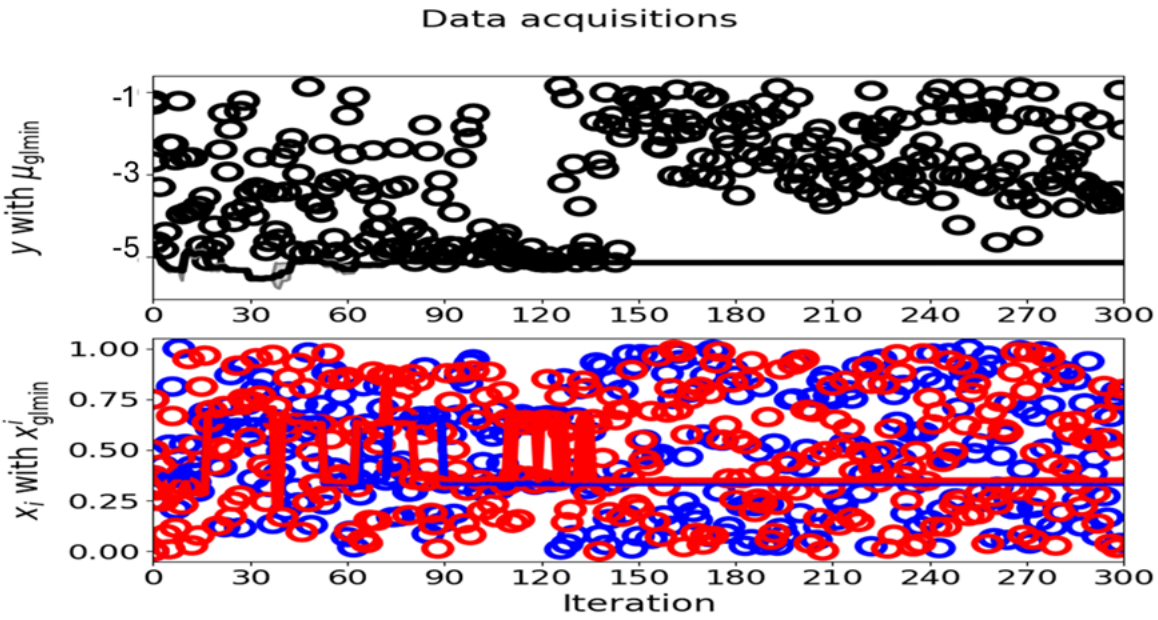

Figure 17: Data Acquisition sequence for the 2D BOSS run

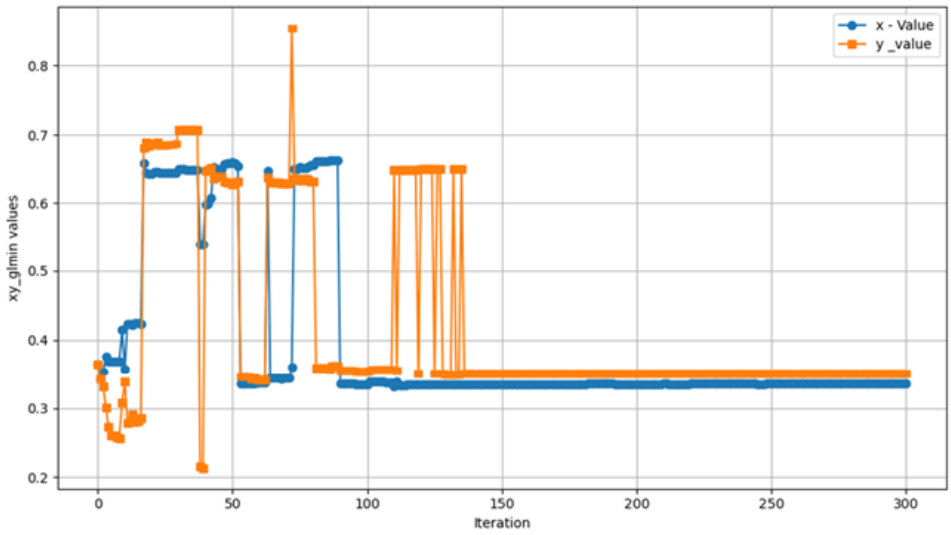

Figure 18: Global Minimum prediction of the 2D BOSS iteration

In Figure 19, the GP surface obtained after 305 direct BOSS evaluations was further refined by applying the four symmetry operations inherent to the ζ(4 × 2) unit cell. Each original evaluation was mirrored and translated to populate the equivalent positions in the periodic

cell, raising the effective information content to 1,220 symmetry augmented samples. A direct visual comparison highlights some improvements:

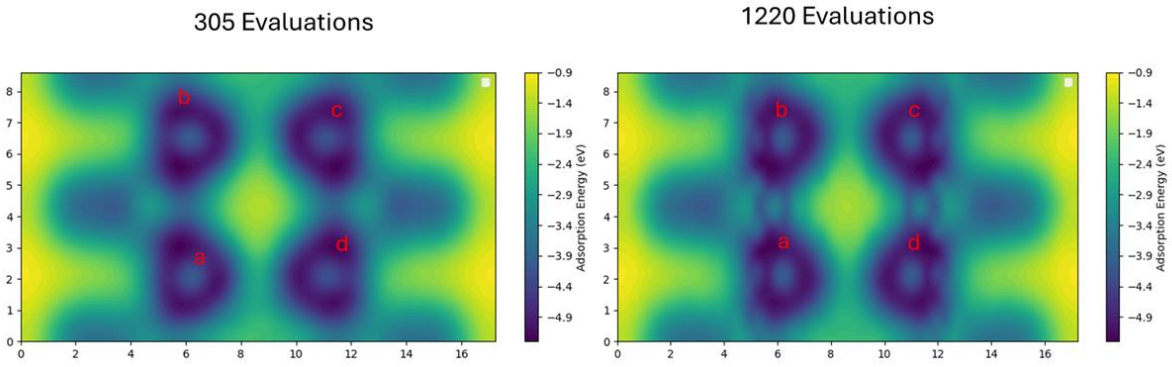

Figure 19: Energy Landscape for 1220 Evaluations and 305 Evaluations

The four indium-arsenic bridge basins (deep purple regions indicated as a-d) appear sharper and more radially symmetric after the symmetry embedding. Local contour inconsistencies present in the 305 sample map have been averaged out, indicating that residual surrogate noise has been suppressed. Secondly, the yellow to green plateau areas, which reflect weak adsorption, appear more uniform in the 1,220 evaluations map.

Despite the refined topology, the global minimum remains fixed with an adsorption energy of approximately -5.15 eV. The energy difference between the global minimum and the second deepest symmetry-related minimum is similar, demonstrating that the 305-evaluation surrogate had already captured the correct energetic ordering of the deepest sites, and therefore, the symmetry embedding enhances resolution rather than altering qualitative conclusions. These observations justify the workflow adopted in the present study: an initial direct-BOSS run of 305 evaluations is sufficient to locate all chemically relevant minima, while subsequent symmetry augmentation is computationally inexpensive once the DFT data have been generated, delivering a high-fidelity landscape suitable for guiding adsorption searches.

4.3.3 Local Minimum Search

To quantify how thoroughly BOSS run had explored the lateral energy landscape, all distinct adsorption energies predicted by the GP surrogate were catalogued at successive evaluations of 150, 200, 250, 305, and 1,220 evaluations. Figure 20 shows, for each evaluation, the energy of every unique minimum in ascending order; each coloured staircase therefore represents the depth distribution of the landscape resolved at that sampling level. In the graph,

the minima are shown as energy plateaus of varying lengths. The longer the plateau, the more the minimizers have ended up in that region of adsorption energy.

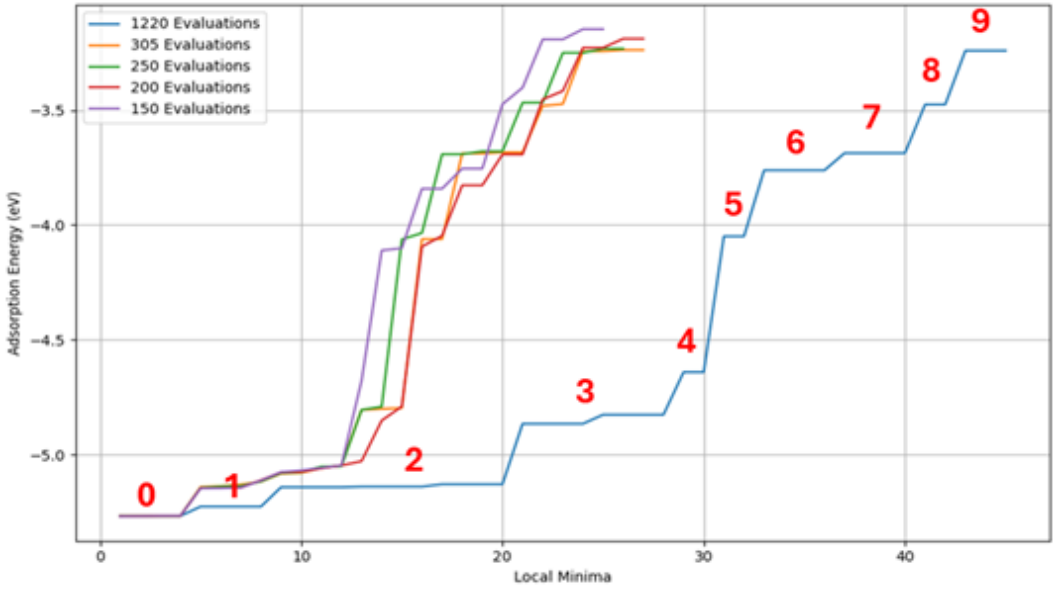

Figure 20: Adsorption energy of the lowest minima in 5 different evaluations of 150, 200,250,305, and 1220. The local minima are identified as points of identical energy, found by multiple minimizers, which are shown as energy plateaus. For the 1220 evaluations, it is indicated by 0-9 in the graph

All evaluations starting with only 150 identify the same deepest energy at -5.35 ± 0.02 eV. Subsequent sampling therefore adds more minima rather than rather than new energy points: the first and second energies are not replaced by deeper discoveries after the 150 evaluations. Increasing from 150 to 305 evaluations basically shifts the position of the staircase. When symmetry operations are applied to generate the 1,220 evaluations surrogate, a further 18 shallow basins are resolved, extending the catalogue to 46 unique points; however, all of these extra states reside in the –4.80 eV to –3.20 eV energy level (–5.3 eV to –3.20 eV for 1220 evaluations), which captures every adsorption basin that the underlying potential-energy surface can support at sub-monolayer coverage. The local minima search analysis confirms that the BO, symmetry-enhanced surrogate containing 1,220 effective points, provides the most complete and numerically stable map of the O/InAs(100) adsorption landscape, and is therefore adopted as the sole reference landscape for further DFT relaxation of the oxygen atom at the local minima positions.

A quartet of the cell was examined (x/2,y/2) since it is symmetry in both lateral directions. Table 9 gives the unique lateral minimum positions sorted by energy level from the local minima search for 1220 evaluation. Focusing the DFT step on these 10 most favourable predictions guarantees that every low-energy adsorption geometry revealed by the fully sampled landscape is examined, while avoiding the high cost of relaxing dozens of higher lying energy positions.

Table 9: The atomic position of oxygen atom indicated by 0-9 from the local minima search

	x(Å)	y(Å)	Eads (eV)
0	5.75159344	3.02230457	-5.26588791
1	5.35469725	2.86558225	-5.22518839
2	5.72819347	1.31195213	-5.14077521
3	5.2775562	1.49297355	-4.86456133
4	5.40130749	4.30698889	-4.64000267
5	3.22849796	4.30698672	-4.04905478
6	3.27543452	0.00000081	-3.76153218
7	2.7168839	0.00000081	-3.68677861
8	1.72279496	4.3069874	-3.47537548
9	1.72279496	0	-3.24124763

4.4 Oxygen Adsorption Sites

Figure 21 shows the 10 extracted positions on the graph, with 0 representing the global minimum. For each site, a single oxygen atom was positioned 1.50 Å (result from 4.3.1) above the surface of the optimized 12-layer slab, followed by geometry optimization to achieve a relaxed structure. Eight optimization points (0-7) converged to a chemically identical In-O-As bridge geometry, confirming the energetic dominance of this motif, while 2 optimization points (8-9) bonded to the As atom at the surface. The key quantitative results are collated in Table 10, while Figures 22-31 provides and top-view before and after optimisation of the final structures.

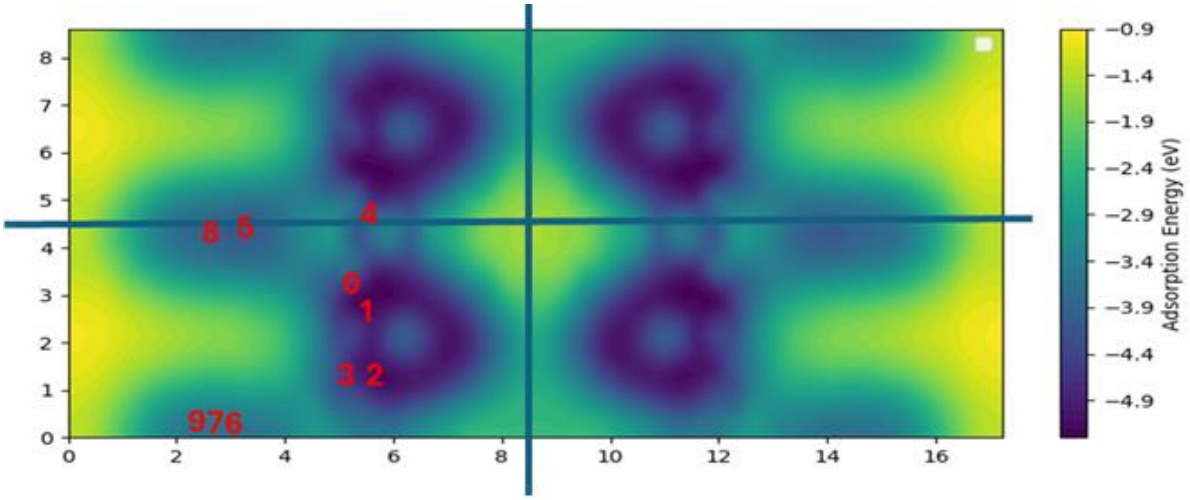

Figure 21: 1220 evaluations energy landscape showing the adsorption positions(indicated as 0-6) predicted by the local minima run

Table 10: Adsorption energies and local metrics

Site	E_{ads}(eV)	In-O	O-As	O-In-As	Final Structural Motif
0	-6.26713321	2.08	1.76	120.9	In-As Bridge
1	-6.26699441	2.08	1.76	120.8	In-As Bridge
2	-6.31438734	2.08	1.76	130.8	In-As Bridge
3	-6.31373631	2.08	1.76	130.4	In-As Bridge
4	-6.13852271	2.10	1.77	105.1	In-As Bridge
5	-6.08820121	2.02	1.81	130.3	In-As Bridge
6	-6.06482021	2.02	1.81	122.8	In-As Bridge
7	-6.06374041	2.02	1.82	122.7	In-As Bridge
8	-5.88264411	-	1.66	-	On Top As Atom
9	-5.75919491	-	1.68	-	On Top As Atom

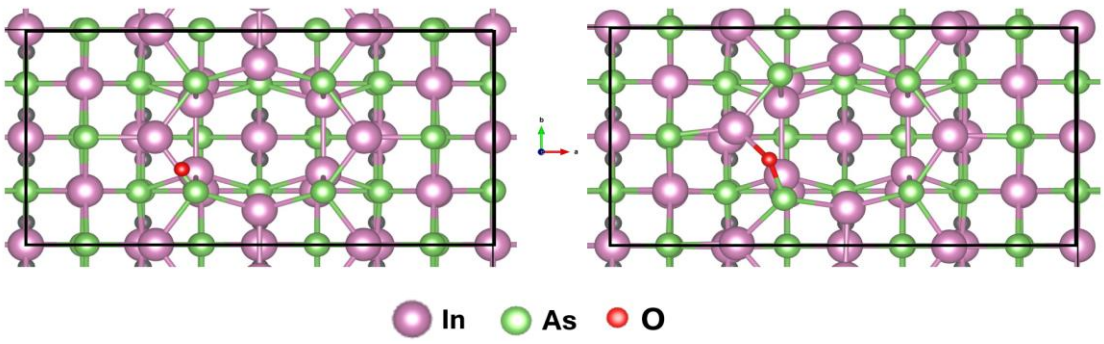

Figure 22: Before (left) and after (right) of the geometry optimization based on position 0 predicted.

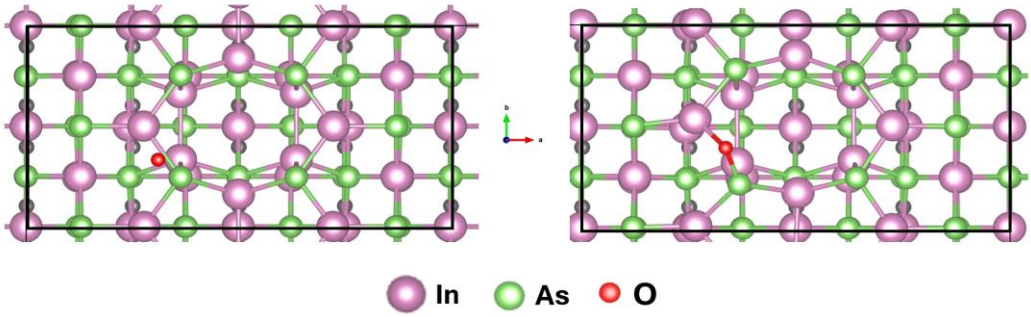

Figure 23: Before (left) and after (right) of the geometry optimization based on position 1 predicted.

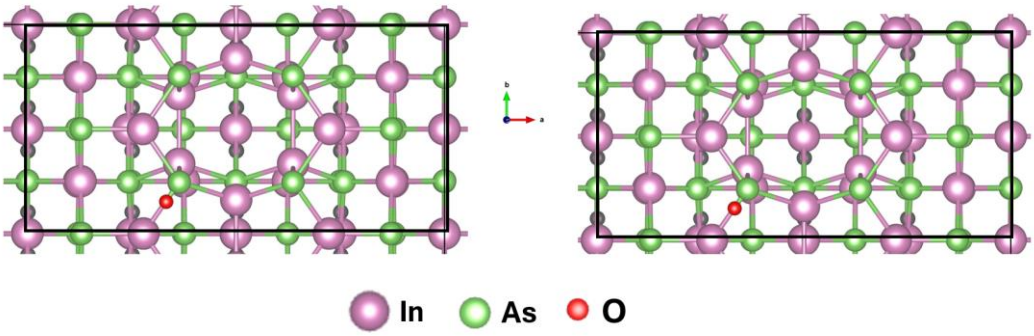

Figure 24: Before (left) and after (right) of the geometry optimization based on position 2 predicted.

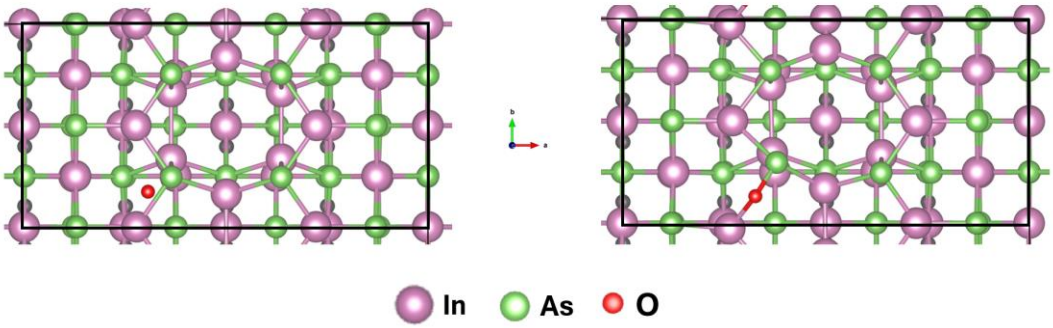

Figure 25: Before (left) and after (right) of the geometry optimization based on position 3 predicted.

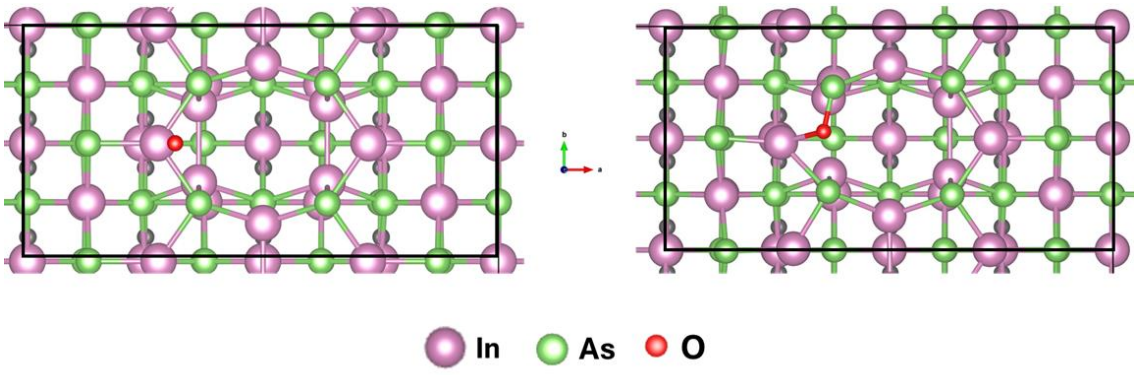

Figure 26: Before (left) and after (right) of the geometry optimization based on position 4 predicted.

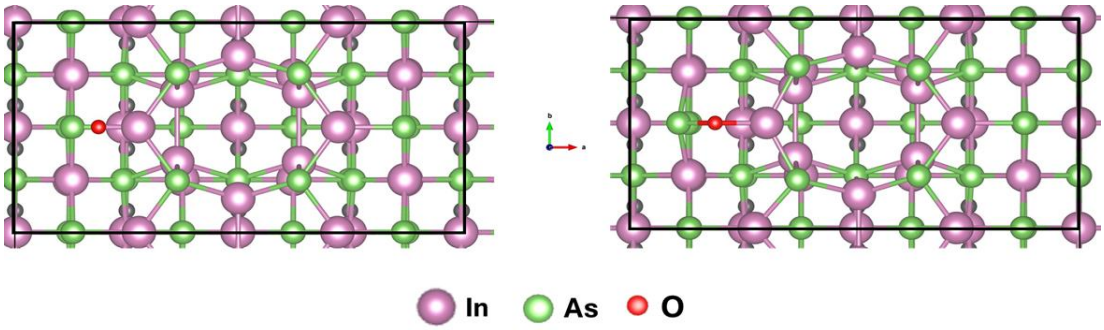

Figure 27: Before (left) and after (right) of the geometry optimization based on position 5 predicted.

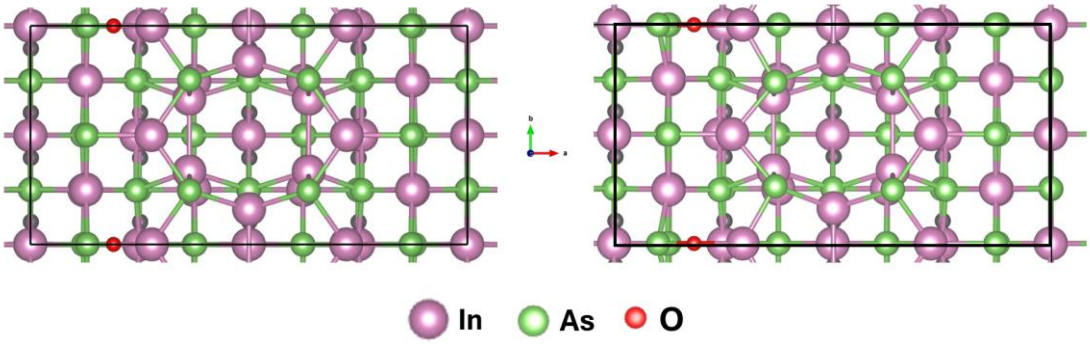

Figure 28: Before (left) and After(right) of the geometry optimization based on position 6 predicted.

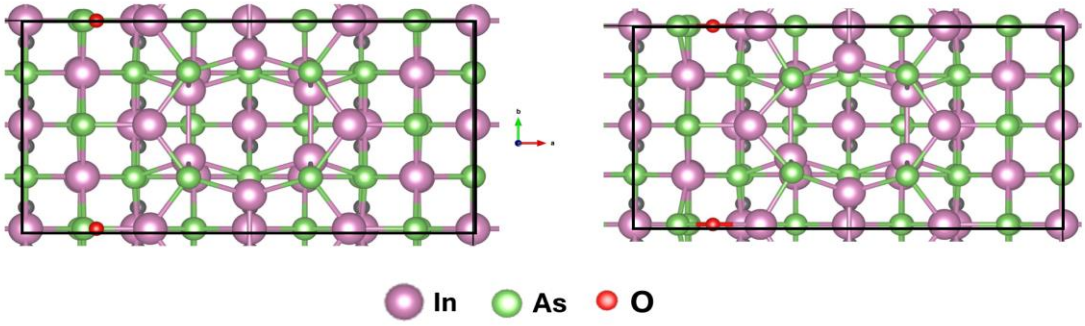

Figure 29: Before (left) and after (right) of the geometry optimization based on position 7 predicted.

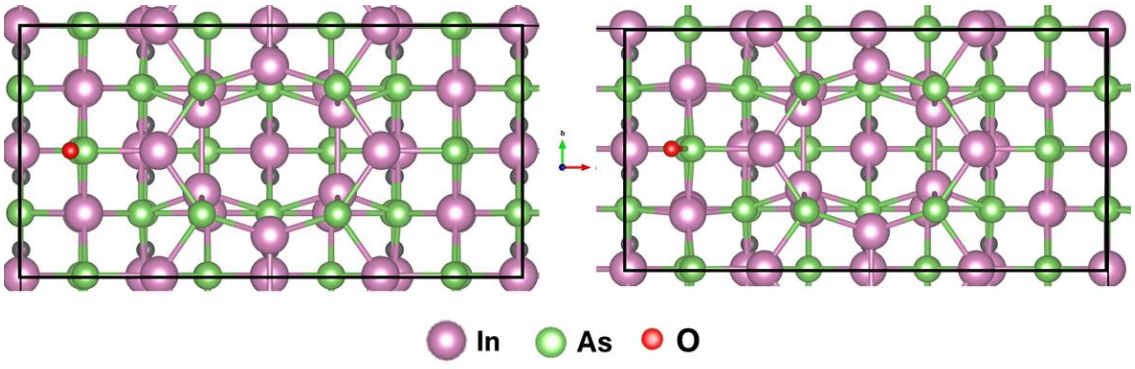

Figure 30: Before (left) and after (right) of the geometry optimization based on position 8 predicted.

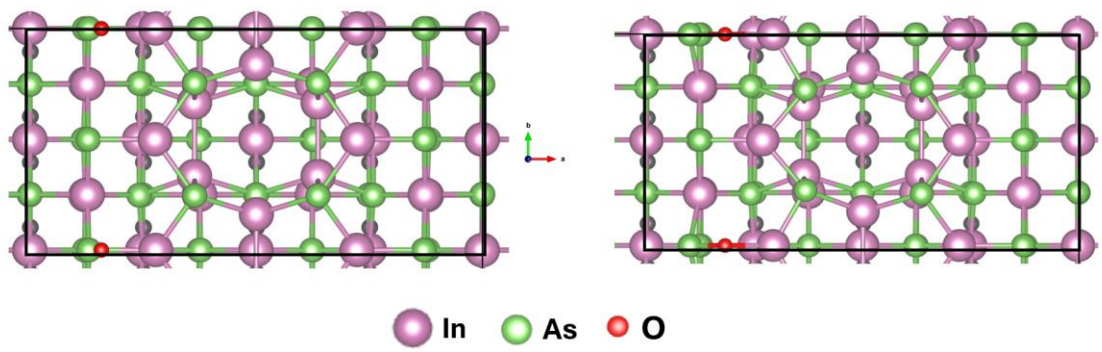

Figure 31: Before (left) and after (right) of the geometry optimization based on position 9 predicted.

The In–O bond lengths in the relaxed structures lie in the range of approximately 2.02–2.08 Å, while the O–As bond lengths lie around approximately 1.66–1.82 Å for all the configurations. Likewise, the O–In–As bond angle is very similar across the sites, falling between approximately 105° and 131°. These small variations are attributable to minor differences in the local surface environment, whether the bridge is closer to a row edge or at the middle of the surface. Importantly, however, eight O adatoms (positions 0-7) adopt the same qualitative bonding geometry that is a two-coordinate (bridging) oxygen with one bond to In and one to As. Two O adatoms bond weakly to the As atom(positions 8-9) at the surface. This indicates that despite the slight asymmetries of the ζ(4×2) reconstruction, oxygen adsorption drives each site to a chemically equivalent In–O–As configuration.

All ten relaxed structures exhibit only minor distortions of the surface relative to the clean ζ(4×2) lattice, indicating that oxygen adsorption does not significantly disrupt the overall reconstruction. In each case, the indium and arsenic atoms involved in the bridge adjust their positions slightly to accommodate the oxygen. The Indium atom tends to be pulled upwards by the bridging O, typically rising by approximately 0.25 Å from its original position, while the arsenic may shift subtly.

These minimal relaxations imply that no major bond rearrangements occur apart from the formation of the In–O–As bonds as well as the As-O bond(for positions 8-9). No other surface bonds are broken or formed beyond this insertion; the rest of the lattice remains intact. There is also no indication of the oxygen moving to a different site: each adatom stays in the vicinity of the initial bridge position, with only slight lateral drift during relaxation to find the local minimum. The prevalence of this configuration is further reflected in the nearly identical adsorption energies at the surfaces above.

The DFT results strongly suggest that the In–As bridge site is the dominant adsorption geometry for atomic oxygen on the In-rich InAs(100)-ζ (4×2) surface. This preferred geometry aligns well with both the BOSS global search predictions, as the algorithm had indeed ranked In–As bridge sites as the lowest-energy adsorption points, and the DFT relaxations confirm their stability.

5 Conclusion

This thesis has investigated the atomic-scale processes of oxygen adsorption and surface reconstruction on the InAs(100)-ζ (4×2) semiconductor surface using density functional theory (DFT) calculations combined with the Bayesian Optimization Structure Search (BOSS) technique. The study commenced with the construction of a literature-consistent ζ(4 × 2) model and a systematic slab-thickness analysis. Structural parameters, which are most critically the surface and subsurface (In–In dimer lengths, In-As atomic distance, and bond angles) were shown to converge at a twelve-layer depth. A preliminary single-oxygen test verified that adsorption energies likewise stabilise at this thickness, establishing a physically reliable yet computationally economical slab for all subsequent work.

An active-learning workflow, combining first-principles total-energy calculations with Bayesian Optimisation Structure Search (BOSS), was then deployed. A one-dimensional BOSS scan determined an adsorption height of 1.5 Å above the surface, whereas a two-dimensional search of 305 direct evaluations, symmetry-augmented to an effective 1,220 samples, mapped the entire lateral potential-energy landscape. The global adsorption site was clearly identified as a bridge position, where the oxygen atom inserts into the surface In–As bond.

Future work may extend the present framework in three directions. First, the stepwise addition of additional oxygen atoms, each guided by BOSS, will enable the exploration of cooperative effects and the gradual transition from sub-monolayer adsorption to the formation of a continuous oxide film. Second, finite-temperature ab initio molecular dynamics simulations, initialised from the relaxed bridge structure, could quantify kinetic barriers for diffusion and dimer-breaking events.

Finally, the approach can be transferable to other III-V surfaces and to alternative passivating compounds, offering a general strategy for rational interface design in next-generation high-electron mobility transistors and quantum devices. This thesis delivers both mechanistic insight and a validated computational framework for the atomistic engineering of oxide–semiconductor interfaces.

References

[1] A. M. Rahman, “A Review on Semiconductors Including Applications and Temperature Effects in Semiconductors,” Am. Sci. Res. J. Eng. Technol. Sci. ASRJETS, vol. 7, pp. 50–70, 2014.

[2] M. Jaros, “Electronic properties of semiconductor alloy systems,” Rep. Prog. Phys., vol. 48, no. 8, pp. 1091–1154, Aug. 1985, doi: 10.1088/0034-4885/48/8/001.

[3] Z. Qiu, X. Shen, and Z. Zhao, “Development Trends and Prospects of Semiconductor Devices and Technology,” Highlights Sci. Eng. Technol., vol. 81, pp. 374–380, Jan. 2024, doi: 10.54097/jc4btz06.

[4] R. S. Mahmood, M. Ali Shakir, Y. T. Mahmood, and D. H. Hussain, “Semiconductors between past and present,” *J. Adv. Sci. Eng. Technol.*, vol. 3, no. 1, pp. 54–56, Jan. 2022, doi: 10.32441/jaset03.01.05.

[5] K. Y. Kamal, “The Silicon Age: Trends in Semiconductor Devices Industry,” *J. Eng. Sci. Technol. Rev.*, vol. 15, no. 1, pp. 110–115, 2022, doi: 10.25103/jestr.151.14.

[6] U. Sharma, G. Kumar, S. Mishra, and R. Thomas, “Advancement of Gate Oxides from SiO_2 to High-k Dielectrics in Microprocessor and Memory,” *J. Phys. Conf. Ser.*, vol. 2267, no. 1, p. 012142, May 2022, doi: 10.1088/1742-6596/2267/1/012142.

[7] R. W. Keyes, “Fundamental limits of silicon technology,” *Proc. IEEE*, vol. 89, no. 3, pp. 227–239, Mar. 2001, doi: 10.1109/5.915372.

[8] M. Razeghi, “Optoelectronic devices based on III-V compound semiconductors which have made a major scientific and technological impact in the past 20 years,” *IEEE J. Sel. Top. Quantum Electron.*, vol. 6, no. 6, pp. 1344–1354, Nov. 2000, doi: 10.1109/2944.902188.

[9] T. F. Kuech, “III-V compound semiconductors: Growth and structures,” *Prog. Cryst. Growth Charact. Mater.*, vol. 62, no. 2, pp. 352–370, Jun. 2016, doi: 10.1016/j.pcrysgrow.2016.04.019.

[10] M. Asif Khan, A. Bhattarai, J. N. Kuznia, and D. T. Olson, “High electron mobility transistor based on a GaN-Al x Ga1− x N heterojunction,” *Appl. Phys. Lett.*, vol. 63, no. 9, pp. 1214–1215, Aug. 1993, doi: 10.1063/1.109775.

[11] C. Convertino, C. B. Zota, H. Schmid, A. M. Ionescu, and K. E. Moselund, “III–V heterostructure tunnel field-effect transistor,” *J. Phys. Condens. Matter*, vol. 30, no. 26, p. 264005, Jul. 2018, doi: 10.1088/1361-648X/aac5b4.

[12] Department of Electronics, Faculty of Electrical Engineering, University of Sidi Bel Abbes, Algeria and A. Mostefai, "Comparison Between Silicon (Si) and Gallium Arsenide (GaAs) Using MATLAB," *J. Nano- Electron. Phys.*, vol. 14, no. 4, pp. 04028-1-04028–4, 2022, doi: 10.21272/jnep.14(4).04028.

[13] O. Madelung, U. Rössler, and M. Schulz, Eds., *Group IV Elements, IV-IV and III-V Compounds. Part b - Electronic, Transport, Optical and Other Properties*, vol. b. in Landolt-Börnstein - Group III Condensed Matter, vol. b. Berlin/Heidelberg: Springer-Verlag, 2002. doi: 10.1007/b80447.

[14] Y. I. Diakite, Y. Malozovsky, C. O. Bamba, L. Franklin, and D. Bagayoko, "First Principle Calculation of Accurate Electronic and Related Properties of Zinc Blende Indium Arsenide (zb-InAs)," *Materials*, vol. 15, no. 10, p. 3690, May 2022, doi: 10.3390/ma15103690.

[15] S. Zwerdling, B. Lax, and L. M. Roth, "Oscillatory Magneto-Absorption in Semiconductors," *Phys. Rev.*, vol. 108, no. 6, pp. 1402–1408, Dec. 1957, doi: 10.1103/PhysRev.108.1402.

[16] A. G. Milnes and A. Y. Polyakov, "Indium arsenide: a semiconductor for high speed and electro-optical devices," *Mater. Sci. Eng. B*, vol. 18, no. 3, pp. 237–259, Apr. 1993, doi: 10.1016/0921-5107(93)90140-I.

[17] Z. Ye, J. C. Campbell, Z. Chen, E.-T. Kim, and A. Madhukar, "InAs quantum dot infrared photodetectors with In0.15Ga0.85As strain-relief cap layers," *J. Appl. Phys.*, vol. 92, no. 12, pp. 7462–7468, Dec. 2002, doi: 10.1063/1.1517750.

[18] E. Memisevic *et al.*, "Individual Defects in InAs/InGaAsSb/GaSb Nanowire Tunnel Field-Effect Transistors Operating below 60 mV/decade," *Nano Lett.*, vol. 17, no. 7, pp. 4373–4380, Jul. 2017, doi: 10.1021/acs.nanolett.7b01455.

[19] T. Xu *et al.*, "Recent progress on infrared photodetectors based on InAs and InAsSb nanowires," *Nanotechnology*, vol. 31, no. 29, p. 294004, May 2020, doi: 10.1088/1361-6528/ab8591.

[20] S. Datta, Shen Shi, K. P. Roenker, M. M. Cahay, and W. E. Stanchina, "Simulation and design of InAlAs/InGaAs pnp heterojunction bipolar transistors," *IEEE Trans. Electron Devices*, vol. 45, no. 8, pp. 1634–1643, Aug. 1998, doi: 10.1109/16.704357.

[21] A. I. Galimov *et al.*, "Source of Indistinguishable Single Photons Based on Epitaxial InAs/GaAs Quantum Dots for Integration in Quantum Computing Schemes," *JETP Lett.*, vol. 113, no. 4, pp. 252–258, Feb. 2021, doi: 10.1134/S0021364021040093.

[22] S. Bhattacharya, J. Ajayan, D.Nirmal, S. Tayal, S. Kollem, and L. M. I. L. Joseph, "Investigation on DC/RF Performance of LG = 19 nm Heterogeneous Integrated Ga0.15In0.85As/InAs/Ga0.15In0.85As Composite Channel InP HEMT on Silicon Substrate for Future Beyond 5G and Quantum Computing Applications," *Silicon*, vol. 14, no. 15, pp. 9581–9588, Oct. 2022, doi: 10.1007/s12633-022-01719-2.

[23] G. P. Srivastava, "Theory of semiconductor surface reconstruction," *Rep. Prog. Phys.*, vol. 60, no. 5, pp. 561–613, May 1997, doi: 10.1088/0034-4885/60/5/002.

[24] G. P. Srivastava, "Theory of semiconductor surface reconstruction," *Rep. Prog. Phys.*, vol. 60, no. 5, pp. 561–613, May 1997, doi: 10.1088/0034-4885/60/5/002.

[25] C. Ratsch, W. Barvosa-Carter, F. Grosse, J. H. G. Owen, and J. J. Zinck, "Surface reconstructions for InAs(001) studied with density-functional theory and STM," *Phys. Rev. B*, vol. 62, no. 12, pp. R7719–R7722, Sep. 2000, doi: 10.1103/PhysRevB.62.R7719.

[26] R. H. Miwa, R. Miotto, and A. C. Ferraz, "In-rich (4×2) and (2×4) reconstructions of the InAs(001) surface," *Surf. Sci.*, vol. 542, no. 1–2, pp. 101–111, Sep. 2003, doi: 10.1016/S0039-6028(03)00955-5.

[27] W. G. Schmidt, "III-V compound semiconductor (001) surfaces," *Appl. Phys. Mater. Sci. Process.*, vol. 75, no. 1, pp. 89–99, Jul. 2002, doi: 10.1007/s003390101058.

[28] Yuanning Chen, R. Myricks, M. Decker, J. Liu, and G. S. Higashi, "The origination and optimization of Si/SiO_2 interface roughness and its effect on CMOS performance," *IEEE Electron Device Lett.*, vol. 24, no. 5, pp. 295–297, May 2003, doi: 10.1109/LED.2003.812545.

[29] D. Bauza, "Extraction of $Si\text{-}SiO_2$ interface trap densities in MOS structures with ultrathin oxides," *IEEE Electron Device Lett.*, vol. 23, no. 11, pp. 658–660, Nov. 2002, doi: 10.1109/LED.2002.805008.

[30] M. J. Graham *et al.*, "Oxidation of III–V semiconductors," *Corros. Sci.*, vol. 49, no. 1, pp. 31–41, Jan. 2007, doi: 10.1016/j.corsci.2006.05.004.

[31] E. D. V. Gómez, M. A. Burgos Paci, S. Amaya-Roncancio, L. B. Avalle, and M. Cecilia Gimenez, "Adsorption and diffusion of O atoms on metallic (1 0 0) surfaces. Cluster and periodic slab approaches," *Comput. Theor. Chem.*, vol. 1208, p. 113556, Feb. 2022, doi: 10.1016/j.comptc.2021.113556.

[32] X. Qin, W.-E. Wang, M. S. Rodder, and R. M. Wallace, "*In situ* surface and interface study of crystalline (3×1)-O on InAs," *Appl. Phys. Lett.*, vol. 109, no. 4, p. 041601, Jul. 2016, doi: 10.1063/1.4959940.

[33] J. Wu, E. Lind, R. Timm, M. Hjort, A. Mikkelsen, and L.-E. Wernersson, "Al2O3/InAs metal-oxide-semiconductor capacitors on (100) and (111)B substrates," *Appl. Phys. Lett.*, vol. 100, no. 13, p. 132905, Mar. 2012, doi: 10.1063/1.3698094.

[34] R. LeSar, *Introduction to Computational Materials Science: Fundamentals to Applications*, 1st ed. Cambridge University Press, 2013. doi: 10.1017/CBO9781139033398.

[35] C. R. A. Catlow, R. G. Bell, and J. D. Gale, "Computer modelling as a technique in materials chemistry," *J. Mater. Chem.*, vol. 4, no. 6, p. 781, 1994, doi: 10.1039/jm9940400781.

[36] D. Alfè, G. D. Price, and M. J. Gillan, "Iron under Earth's core conditions: Liquid-state thermodynamics and high-pressure melting curve from *ab initio* calculations," *Phys. Rev. B*, vol. 65, no. 16, p. 165118, Apr. 2002, doi: 10.1103/PhysRevB.65.165118.

[37] K. Gall, M. F. Horstemeyer, M. Van Schilfgaarde, and M. I. Baskes, "Atomistic simulations on the tensile debonding of an aluminum–silicon interface," *J. Mech. Phys. Solids*, vol. 48, no. 10, pp. 2183–2212, Oct. 2000, doi: 10.1016/S0022-5096(99)00086-1.

[38] G. Lu and E. Kaxiras, "An Overview of Multiscale Simulations of Materials," 2004, *arXiv*. doi: 10.48550/ARXIV.COND-MAT/0401073.

[39] M. Orio, D. A. Pantazis, and F. Neese, "Density functional theory," *Photosynth. Res.*, vol. 102, no. 2–3, pp. 443–453, Dec. 2009, doi: 10.1007/s11120-009-9404-8.

[40] J. Chen, F. Min, L. Liu, and C. Liu, "Mechanism research on surface hydration of kaolinite, insights from DFT and MD simulations," *Appl. Surf. Sci.*, vol. 476, pp. 6–15, May 2019, doi: 10.1016/j.apsusc.2019.01.081.

[41] A. Liu *et al.*, "Investigation on the interfacial behavior of polyorganic inhibitors on a metal surface by DFT study and MD simulation," *Appl. Surf. Sci.*, vol. 541, p. 148570, Mar. 2021, doi: 10.1016/j.apsusc.2020.148570.

[42] W. Kohn, A. D. Becke, and R. G. Parr, "Density Functional Theory of Electronic Structure," *J. Phys. Chem.*, vol. 100, no. 31, pp. 12974–12980, Jan. 1996, doi: 10.1021/jp960669l.

[43] B. Keimer and J. E. Moore, "The physics of quantum materials," *Nat. Phys.*, vol. 13, no. 11, pp. 1045–1055, Nov. 2017, doi: 10.1038/nphys4302.

[44] A. Szabo and N. S. Ostlund, *Modern quantum chemistry: introduction to advanced electronic structure theory*. Mineola, N.Y: Dover Publications, 1996.

[45] M. V. Putz, Ed., *Advances in quantum chemical bonding structures*. Trivandrum: Transworld Research Network, 2008.

[46] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Ann. Phys.*, vol. 389, no. 20, pp. 457–484, Jan. 1927, doi: 10.1002/andp.19273892002.

[47] V. J. Härkönen, R. Van Leeuwen, and E. K. U. Gross, "Many-body Green's function theory of electrons and nuclei beyond the Born-Oppenheimer approximation," *Phys. Rev. B*, vol. 101, no. 23, p. 235153, Jun. 2020, doi: 10.1103/PhysRevB.101.235153.

[48] T. Jecko, "On the mathematical treatment of the Born-Oppenheimer approximation," *J. Math. Phys.*, vol. 55, no. 5, p. 053504, May 2014, doi: 10.1063/1.4870855.

[49] L. S. Cederbaum, "The exact molecular wavefunction as a product of an electronic and a nuclear wavefunction," *J. Chem. Phys.*, vol. 138, no. 22, p. 224110, Jun. 2013, doi: 10.1063/1.4807115.

[50] J. J. Sakurai and S. F. Tuan, *Modern quantum mechanics*, Rev. ed. Reading, Mass: Addison-Wesley Pub. Co, 1994.

[51] S. Oh, "Quantum computational method of finding the ground-state energy and expectation values," *Phys. Rev. A*, vol. 77, no. 1, p. 012326, Jan. 2008, doi: 10.1103/PhysRevA.77.012326.

[52] A. M. N. Niklasson and M. J. Cawkwell, "Fast method for quantum mechanical molecular dynamics," *Phys. Rev. B*, vol. 86, no. 17, p. 174308, Nov. 2012, doi: 10.1103/PhysRevB.86.174308.

[53] P. Echenique and J. L. Alonso, "A mathematical and computational review of Hartree–Fock SCF methods in quantum chemistry," *Mol. Phys.*, vol. 105, no. 23–24, pp. 3057–3098, Dec. 2007, doi: 10.1080/00268970701757875.

[54] J. G. Valatin, "Generalized Hartree-Fock Method," *Phys. Rev.*, vol. 122, no. 4, pp. 1012–1020, May 1961, doi: 10.1103/PhysRev.122.1012.

[55] W. A. Goddard, "Improved Quantum Theory of Many-Electron Systems. II. The Basic Method," *Phys. Rev.*, vol. 157, no. 1, pp. 81–93, May 1967, doi: 10.1103/PhysRev.157.81.

[56] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Phys. Rev.*, vol. 136, no. 3B, pp. B864–B871, Nov. 1964, doi: 10.1103/PhysRev.136.B864.

[57] U. V. Barth and L. Hedin, "A local exchange-correlation potential for the spin polarized case. i," *J. Phys. C Solid State Phys.*, vol. 5, no. 13, pp. 1629–1642, Jul. 1972, doi: 10.1088/0022-3719/5/13/012.

[58] W. Kohn and L. J. Sham, “Self-Consistent Equations Including Exchange and Correlation Effects,” *Phys. Rev.*, vol. 140, no. 4A, pp. A1133–A1138, Nov. 1965, doi: 10.1103/PhysRev.140.A1133.

[59] A. D. Becke, “Perspective: Fifty years of density-functional theory in chemical physics,” *J. Chem. Phys.*, vol. 140, no. 18, p. 18A301, May 2014, doi: 10.1063/1.4869598.

[60] M. Todorović, M. U. Gutmann, J. Corander, and P. Rinke, “Bayesian inference of atomistic structure in functional materials,” *Npj Comput. Mater.*, vol. 5, no. 1, p. 35, Mar. 2019, doi: 10.1038/s41524-019-0175-2.

[61] N. Ketkar, “Stochastic Gradient Descent,” in *Deep Learning with Python*, Berkeley, CA: Apress, 2017, pp. 113–132. doi: 10.1007/978-1-4842-2766-4_8.

[62] J. Joyce, “Bayes Theorem,” *Stanf. Encycl. Philos.*, 2008, [Online]. Available: https://philpapers.org/rec/JOYBT

[63] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*, 3. print. in Adaptive computation and machine learning. Cambridge, Mass.: MIT Press, 2008.

[64] C. Jukka and G. Michael U, “Bayesian Optimization for Likelihood-Free Inference of Simulator-Based Statistical Models,” *J. Mach. Learn. Res.*, vol. 17, pp. 1–47, 2016.

[65] M. Todorović, M. U. Gutmann, J. Corander, and P. Rinke, “Bayesian inference of atomistic structure in functional materials,” *Npj Comput. Mater.*, vol. 5, no. 1, p. 35, Mar. 2019, doi: 10.1038/s41524-019-0175-2.

[66] E. Brochu, V. M. Cora, and N. de Freitas, “A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning,” 2010, *arXiv*. doi: 10.48550/ARXIV.1012.2599.

[67] A. Candelieri, “A Gentle Introduction to Bayesian Optimization,” in *2021 Winter Simulation Conference (WSC)*, Phoenix, AZ, USA: IEEE, Dec. 2021, pp. 1–16. doi: 10.1109/WSC52266.2021.9715413.

[68] J. T. Wilson, R. Moriconi, F. Hutter, and M. P. Deisenroth, “The reparameterization trick for acquisition functions,” 2017, *arXiv*. doi: 10.48550/ARXIV.1712.00424.

[69] P. I. Frazier, “A Tutorial on Bayesian Optimization,” 2018, *arXiv*. doi: 10.48550/ARXIV.1807.02811.

[70] D. D. Cox and S. John, “A statistical method for global optimization,” in *[Proceedings] 1992 IEEE International Conference on Systems, Man, and*

Cybernetics, Chicago, IL, USA: IEEE, 1992, pp. 1241–1246. doi: 10.1109/ICSMC.1992.271617.

[71] H. J. Kushner, "A New Method of Locating the Maximum Point of an Arbitrary Multipeak Curve in the Presence of Noise," *J. Basic Eng.*, vol. 86, no. 1, pp. 97–106, Mar. 1964, doi: 10.1115/1.3653121.

[72] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012, doi: 10.1109/TIT.2011.2182033.

[73] M. Seeger, "GAUSSIAN PROCESSES FOR MACHINE LEARNING," *Int. J. Neural Syst.*, vol. 14, no. 02, pp. 69–106, Apr. 2004, doi: 10.1142/S0129065704001899.

[74] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian Optimization of Machine Learning Algorithms," 2012, *arXiv*. doi: 10.48550/ARXIV.1206.2944.

[75] M. Kuchelmeister, "Multi-fidelity Bayesian machine learning for global optimization," 2022, *Universität Stuttgart*. doi: 10.18419/OPUS-12396.

[76] R. Preuss and U. Von Toussaint, "Global Optimization Employing Gaussian Process-Based Bayesian Surrogates," *Entropy*, vol. 20, no. 3, p. 201, Mar. 2018, doi: 10.3390/e20030201.

[77] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. in Adaptive computation and machine learning. Cambridge, Mass: MIT Press, 2006.

[78] L. Kotthoff, H. Wahab, and P. Johnson, "Bayesian Optimization in Materials Science: A Survey," 2021, *arXiv*. doi: 10.48550/ARXIV.2108.00002.

[79] P. Lyngby, C. Larsen, and K. W. Jacobsen, "Bayesian optimization of atomic structures with prior probabilities from universal interatomic potentials," 2024, *arXiv*. doi: 10.48550/ARXIV.2408.15590.

[80] O. Madelung, U. Rössler, and M. Schulz, Eds., *Group IV Elements, IV-IV and III-V Compounds. Part b - Electronic, Transport, Optical and Other Properties*, vol. b. in Landolt-Börnstein - Group III Condensed Matter, vol. b. Berlin/Heidelberg: Springer-Verlag, 2002. doi: 10.1007/b80447.

[81] G. Goryl, D. Toton, M. Goryl, N. Tomaszewska, and J. J. Kolodziej, "Structure of the In-rich InAs (001) surface," *Surf. Sci.*, vol. 605, no. 23–24, pp. 2073–2081, Dec. 2011, doi: 10.1016/j.susc.2011.08.008.

[82] K. Kourchid, R. Alaya, A. Rebey, and M. Mbarki, "Study of Surface Stability and Electronic Structure of a Bi-terminated InAs (001) Surface Based on Ab Initio

Calculations," *J. Electron. Mater.*, vol. 50, no. 6, pp. 3527–3536, Jun. 2021, doi: 10.1007/s11664-021-08858-z.

[83] S. Yang *et al.*, "Electronic Structure of InAs and InSb Surfaces: Density Functional Theory and Angle-Resolved Photoemission Spectroscopy," *Adv. Quantum Technol.*, vol. 5, no. 3, p. 2100033, Mar. 2022, doi: 10.1002/qute.202100033.

[84] A. V. Bakulin, S. V. Eremeev, O. E. Tereshchenko, and S. E. Kulkova, "Chlorine adsorption on the InAs (001) surface," *Semiconductors*, vol. 45, no. 1, pp. 21–29, Jan. 2011, doi: 10.1134/S1063782611010040.

Appendix

Appendix: Atomic Coordinates of InAs(100)-ζ (4×2) surface

Atom	x(Å)	y(Å)	z(Å)
As	0	0	0
In	0	2.49	0.19
As	-2.05	3.97	0.76
In	-1.33	6.26	-0.27
As	-2.11	8.55	0.76
In	0	10.03	0.19
As	0	12.62	0
As	4.21	0.04	0.05
In	4.21	2.53	0.38
As	2.05	3.97	0.76
In	1.33	6.26	-0.28
As	2.05	8.55	0.76
In	4.21	10.03	0.38
As	4.21	12.54	0.05
In	-2.12	0	-1.51
As	-2.12	1.95	-3.23
As	-2.12	-1.95	-3.23
In	2.12	0	-1.51
As	2.12	-1.95	-3.23
As	2.12	1.95	-3.23
As	-2.12	6.26	-3.03
In	6.293	0	-1.51
As	6.293	-1.95	-3.23
As	6.293	1.95	-3.23
In	-2.12	12.62	-1.51
As	-2.12	10.67	-3.23
As	-2.12	14.57	-3.23
In	2.12	12.62	-1.51
As	2.12	10.67	-3.23
As	2.12	14.57	-3.23
In	6.293	12.62	-1.51
As	6.293	10.67	-3.23
As	6.293	14.574	-3.23

Atom	x(Å)	y(Å)	z(Å)
In	2.83	3.97	-1.75
As	6.37	3.97	0.76
In	5.59	3.97	-1.75
As	2.09	6.26	-3.05
As	6.3	6.26	-3.05
In	7.05	6.26	-0.28
In	2.83	8.55	-1.75
In	5.59	8.55	-1.75
As	6.37	8.55	0.76