

New Material Discovery Using Physics-Inspired Reinforcement Learning Techniques

A thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the
requirements for the BS-MS Dual Degree Programme

by

Garvit Bansal

20201013



Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,
Pashan, Pune 411008, India.

April, 2025

Supervisor: Chandrajit Bajaj

All Rights Reserved

Certificate

This is to certify that this dissertation entitled New Material Discovery Using Physics-Inspired Reinforcement Learning Techniques towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Garvit Bansal at University of Texas, Austin under the supervision of Professor Chandrajit Bajaj, Oden Institute, during the academic year 2024-2025.

Chandrajit Bajaj

(Project Guide)

Committee:



Chandrajit Bajaj

(Project Guide)



Prasenjit Ghosh

(TAC Expert)

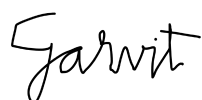
Date: 21 March 2025

“To optimize the whole, we must sub-optimize the parts”

William Edwards Deming

Declaration

I hereby declare that the matter embodied in the report entitled “New Material Discovery Using Physics-Inspired Reinforcement Learning Techniques” are the results of the work carried out by me at the Oden Institute, University of Texas (UT) Austin, under the supervision of Professor Chandrajit Bajaj, and the same has not been submitted elsewhere for any other degree. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgement of collaborative research and discussions.

A handwritten signature in black ink, appearing to read "Garvit". The script is cursive and fluid, with the first letter 'G' being particularly large and stylized.

Name of Student: Garvit Bansal

Roll Number: 20201013

Abstract

The accelerated discovery of novel materials with targeted properties is crucial for advancing technological fields such as photovoltaics, catalysis, and energy storage. Traditional methods for material discovery often rely on experimental intuition or computational brute-force screening, both of which become increasingly inefficient as the complexity and dimensionality of the search space increase. To overcome these limitations, this thesis introduces a novel computational framework based on physics-inspired reinforcement learning (RL), specifically leveraging Stochastic Hamiltonian Dynamics (SHD), to efficiently explore high-dimensional search spaces for materials optimization.

The properties of materials can be calculated via computational techniques like Density Functional Theory (DFT) calculations, which, although accurate, are computationally demanding. To address this challenge, we incorporate local surrogate modeling to approximate expensive DFT computations rapidly and cost-effectively. These inherently differentiable surrogate models significantly accelerate property evaluations, enabling efficient gradient-based optimization within both structural and chemical spaces. Additionally, to seamlessly navigate categorical variables associated with chemical compositions, we employ the Gumbel-Softmax reparameterization technique, transforming discrete choices into differentiable, continuous variables.

Optimal control theory, particularly the Stochastic Pontryagin Maximum Principle (PMP), underpins our approach, providing a systematic framework for stable convergence and global optimization. By adapting SHD within this control-theoretic framework, we combine deterministic momentum-driven exploration with controlled stochastic perturbations, effectively avoiding local minima and promoting efficient global optimization.

Ultimately, this thesis advances computational methodologies in materials science by establishing a robust, scalable, and versatile discovery framework. By integrating physical insights, reinforcement learning strategies, surrogate modeling, and optimal control, we provide a transformative approach that addresses the critical computational barriers and accelerates the discovery of new materials with optimal properties.

Acknowledgements

I would like to extend my deepest gratitude to my supervisor, Professor Chandrajit Bajaj from the University of Texas at Austin, for his invaluable guidance, support, and mentorship throughout the course of this research. His expertise, patience, and insightful suggestions have significantly contributed to shaping my thesis and my understanding of computational material science. Also, I sincerely thank Professor Prasenjit Ghosh from IISER Pune for evaluating my thesis and providing constructive feedback, which has greatly enhanced the clarity and depth of this work.

I am deeply grateful to the University of Texas at Austin for granting me the incredible opportunity to carry out my research in a stimulating and resourceful environment, as well as providing generous funding to support my work with the support of my supervisor. Additionally, I express my heartfelt thanks to the Indian Institute of Science Education and Research (IISER), Pune, my home university, for offering me a foundational academic environment and continuous support throughout my degree.

I warmly acknowledge my coworkers and peers — Yi Wang, Fabian Conrad, Luke McLennan, Ashwin Vinod, Ryan Farell, Ryan Roby, Shubham Bharadwaj, and Aditya Sai — for their invaluable discussions, collaborative spirit, and camaraderie, which have made my research experience both productive and enjoyable.

I extend my sincere appreciation to the Department of Science and Technology's Inspire Scholarship for Higher Education (SHE), whose financial support has been instrumental throughout the tenure of my degree. Moreover, special thanks are due to the Physics Department at IISER Pune for accommodating my thesis under their department, facilitating access to resources, and providing an encouraging environment for conducting research.

Finally, I wish to thank Professor Ivano Castelli for generously sharing his insights about his database of organometal halide perovskites and for his patient guidance in clarifying my queries regarding the GPAW library.

Each contribution mentioned has been pivotal in the successful completion of this thesis, and I am immensely grateful to all who have supported me in this academic journey.

Contents

1	Introduction	8
2	Methods	11
2.1	Electron Structure Calculations	11
2.1.1	Density Functional Theory	11
2.1.1.1	Inputs for DFT Calculations	13
2.1.1.2	Choice of Exchange Correlation Potential	15
2.1.1.3	The Self Consistent Flow Pipeline	16
2.1.1.4	Pulay Mixing	18
2.1.1.5	Bandgap Estimation Using DFT	19
2.1.1.6	Projector Augmented Wave Method	21
2.1.1.7	Lagrangian Method for Molecular Dynamics in DFT	22
2.1.2	Density Functional Perturbation Theory	22
2.2	Properties and Loss Functions for Photovoltaic Materials	24
2.2.1	Choice of Target Bandgap	26
2.3	The Search Space for Material Discovery	26
2.4	Optimization Methods	27
2.5	Optimal Control Theory	27
2.5.1	Pontryagin Maximum Principle	27
2.5.1.1	Deterministic PMP	27
2.5.1.2	Stochastic PMP	29
2.5.1.3	Using Stochastic PMP for Black-Box Minimization	30
2.5.2	Hamilton-Jacobi-Bellmann Method	32
2.5.3	Optimal Control in Fokker-Planck Framework	33
2.6	Surrogate Models	35
2.7	The Gumbel-Softmax Reparametrization Trick	35
2.8	Our Algorithm	36
3	Results	39
3.1	Case 1: Cubic Lattices	39
3.2	Case 2: Cuboidal Lattices	40

4	Discussion	42
4.1	Challenges in High-Dimensional Discovery via Traditional Approaches .	42
4.2	Advantages of Our Physics-Inspired Reinforcement Learning Approach	43
5	Conclusion and Future Outlook	45
	References	46
A	Derivation of Deterministic PMP Through Lagrange Multipliers	54
B	Proof of Uniqueness of Optimal Control in Stochastic Black Box Optimization	55
C	Defining the PAW Transformation	55

List of Figures

1	Steps that go into discovering a new material	8
2	Jacob's Ladder for Exchange-Correlation Potentials	16
3	The Self-Consistent Flow Pipeline	17
4	The Density Function Perturbation Theory Pipeline	24
5	Evolution of $\mathcal{F}(\mathbf{q})$ for search over cubic lattices	40
6	Evolution of Kohn-Sham Energy for search over cuboidal lattices	41
7	Evolution of $\mathcal{F}(\mathbf{q})$ for search over cuboidal lattices	41

1 Introduction

The accelerated discovery of new materials possessing targeted and optimal properties is a pivotal endeavor driving advancement in numerous scientific and technological domains, including photovoltaics [1], semiconductor technology [2, 3], catalysis [4], and energy storage [5]. Traditional methods of material discovery, historically reliant on experimental intuition, trial-and-error exploration, or computational brute-force screening, face significant limitations when confronted with increasingly complex and high-dimensional search spaces. These methodologies, while occasionally successful, are typically resource-intensive, time-consuming, and computationally prohibitive, particularly as the dimensionality and complexity of the search increase due to structural and compositional variations. Consequently, there exists a substantial demand for innovative and computationally efficient approaches that not only accelerate the materials discovery process but also reliably guide the exploration toward globally optimal solutions.

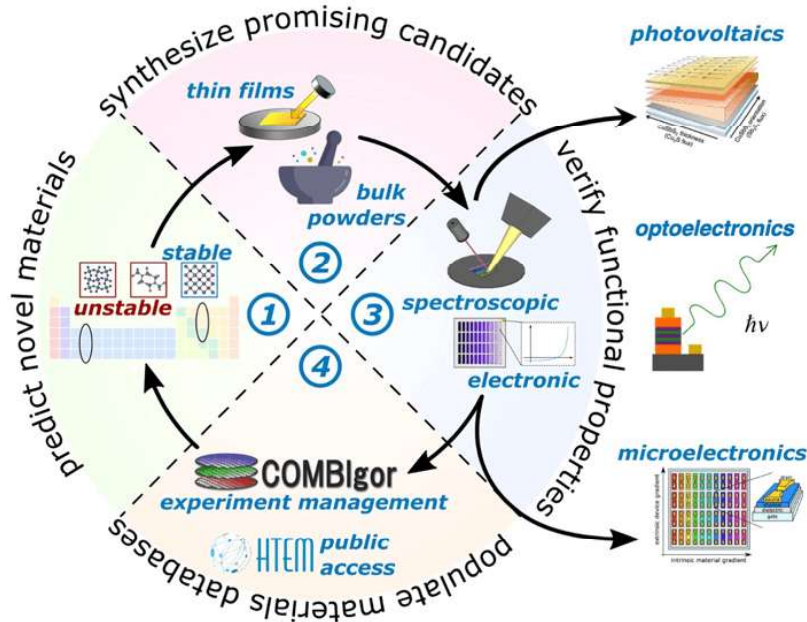


Figure 1: Steps that go into discovering a new material. We focus on step 1, which is predicting new materials [6]

Recent advancements in computational approaches, notably in the application of machine learning and artificial intelligence (AI) techniques, have demonstrated potential to address these challenges [7–11]. Reinforcement learning (RL), in particular, has

emerged as a powerful strategy for complex decision-making problems characterized by sequential, strategic actions in uncertain and high-dimensional environments [12, 13]. The potential of RL in materials science is substantial, however, conventional RL strategies often suffer from inefficiencies in exploration and convergence, especially when applied directly to materials discovery. To overcome these barriers, a new paradigm of physics-inspired RL [14, 15] has been proposed, leveraging principles from physical and dynamical systems to provide structured, efficient, and stable exploration in materials optimization.

This thesis advances the field by developing a novel framework that integrates physics-inspired reinforcement learning with sophisticated optimal control strategies, specifically leveraging Stochastic Hamiltonian Dynamics (SHD). Optimal control theory, particularly through approaches such as the Pontryagin Maximum Principle (PMP), provides systematic methodologies for achieving optimal actions under defined dynamical constraints. By adapting and extending these control-theoretic principles to stochastic environments, our approach systematically navigates the search space, effectively combining deterministic momentum-driven exploration with controlled stochastic perturbations to avoid local minima and promote global optimization.

Moreover, given the computational cost associated with Density Functional Theory (DFT) calculations (which is the standard method for estimating properties of materials), we integrate local surrogate modeling techniques [16] to approximate these expensive computations efficiently. Surrogate modeling has become a critical tool in computational science, enabling rapid and cost-effective estimations of complex properties while significantly reducing computational burdens. By employing inherently differentiable surrogate models and adopting advanced computational techniques such as the Gumbel-Softmax reparameterization [17], our approach efficiently addresses the discrete categorical variables associated with chemical compositions, thus facilitating seamless gradient-based optimization of categorical variables in high-dimensional search spaces.

To demonstrate the downsides of traditional approaches and compare them with our novel approach, we have considered the optimization of photovoltaic materials, focusing

particularly on fine-tuning the fundamental bandgap, which is important to understand the electro-optic properties of materials [18].

In conclusion, this thesis not only contributes to advancing computational methods in materials discovery but also provides a robust, scalable, and versatile platform suitable for diverse materials optimization objectives. Our methodology combines physical insights, reinforcement learning, optimal control, and advanced surrogate modeling to offer a transformative solution to the challenges inherent in modern materials science exploration.

2 Methods

2.1 Electron Structure Calculations

2.1.1 Density Functional Theory

Density Functional Theory (DFT) is a computational quantum mechanical method widely used in chemistry, physics, and materials science to study the electronic structure of atoms, molecules, and solids. Unlike traditional wavefunction-based methods like Hartree-Fock, which describe the quantum state of a system in terms of a complex mathematical object (the wavefunction), DFT provides the description of the electron structure based on its ground-state electron density.

Density functional theory is based on the Hohenberg-Kohn theorems [19, 20] which state that the potential acting on a system is uniquely determined by the ground-state electron density of the system and that it can be calculated by minimizing the Kohn-Sham energy functional $E_{KS}[n(\mathbf{r})]$. To make this problem more computable, Kohn and Sham came up with a strategy of mapping a system with interacting electrons to a system with non-interacting electrons with the same ground state density as interacting electrons [21]. Then, they decomposed the ground-state electron density into occupied Kohn-Sham energy eigenstates as:

$$n(\mathbf{r}) = \sum_{j \in \text{occupied}} |\phi_j(\mathbf{r})|^2 \quad (1)$$

Where $\phi_j(\mathbf{r})$ are the Kohn-Sham states. By applying Bloch theorem to periodic lattices [22], we get $\phi_j(\mathbf{r})$ of the form:

$$\phi_j(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{m,\mathbf{k}}(\mathbf{r}) \quad (2)$$

Where m is the band index which takes values in positive integers and $\mathbf{k}$ is a vector in the reciprocal space and without loss of generality, can be assumed to lie in the first Brillouin zone of the lattice. Here, the index j is actually a pair of indices ($j = (m, \mathbf{k})$). In the real space, periodic boundary conditions are implemented while performing calculations for lattices.

In our Kohn-Sham calculations, we need to sample enough $\mathbf{k}$ points from the entire Brillouin zone such that we get accurate estimates for electron density. The usual choice for this sampling is the Monkhorst-Pack grid [23], which is a uniform grid of $\mathbf{k}$ -points in the first Brillouin zone. People also typically use symmetries in the lattice to reduce the number of $\mathbf{k}$ -points sampled by just sampling in the irreducible Brillouin zone (which is the smallest component of the Brillouin zone which can generate the entire Brillouin zone via symmetry actions) and then, symmetries are applied to get eigenstates for symmetry equivalent $\mathbf{k}$ -points.

The Kohn-Sham energy functional is given by:

$$\begin{aligned}
E_{KS}[n(\mathbf{r})] = & \underbrace{T_{KS}[n(\mathbf{r})]}_{\text{Kohn-Sham Kinetic Energy}} + \underbrace{\int n(\mathbf{r})V_{ext}(\mathbf{r})d^3\mathbf{r}}_{\text{External Potential Energy (Electron-Ion Interaction)}} \\
& + \underbrace{E_H[n(\mathbf{r})]}_{\text{Hartree Energy (Electron-Electron Interaction)}} \\
& + \underbrace{E_{xc}[n(\mathbf{r})]}_{\text{Exchange-Correlation Energy}} + \underbrace{E_{ion-ion}}_{\text{Ion-Ion Interaction}}
\end{aligned} \tag{3}$$

These terms can further be expanded as:

$$T_{KS}[n(\mathbf{r})] = \sum_{j \in \text{occupied}} \int d^3\mathbf{r} \phi_j^*(\mathbf{r}) \left(-\frac{1}{2} \nabla_{\mathbf{r}}^2 \right) \phi_j(\mathbf{r}) \tag{4}$$

$$E_H[n(\mathbf{r})] = \frac{1}{2} \int d^3\mathbf{r} \int d^3\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{5}$$

Variational minimization of the Kohn-Sham energy functional gives us the Kohn-Sham equations of the form:

$$\hat{H}_{KS}[n(\mathbf{r})]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{6}$$

With the Kohn-Sham Hamiltonian $H_{KS}[n(\mathbf{r})]$ given as:

$$\hat{H}_{KS}[n(\mathbf{r})] = \hat{T}_{KS} + \hat{V}_{ext}(\mathbf{r}) + \hat{V}_H[n(\mathbf{r})] + \hat{V}_{xc}[n(\mathbf{r})] \tag{7}$$

With the different terms given as:

$$\hat{T}_{KS} = -\frac{1}{2} \nabla_{\mathbf{r}}^2 \tag{8}$$

$$\hat{V}_H[n(\mathbf{r})] = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \quad (9)$$

$$\hat{V}_{xc}[n(\mathbf{r})] = \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \quad (10)$$

From the solutions of the Kohn-Sham equations, the Kohn-Sham energy can be updated as:

$$E_{KS}[n(\mathbf{r})] = \sum_{j \in \text{occupied}} \epsilon_j - E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] - \int \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} n(\mathbf{r}) d^3\mathbf{r} \quad (11)$$

Extensions of DFT include Time-Dependent DFT, which is based on the Runge-Gross theorem [24] and extends DFT to time dependent and excited state systems, and DFT+U [25], which adds a Hubbard-U correction to the Hamiltonian.

2.1.1.1 Inputs for DFT Calculations In the density functional theory flow (PAW formalism with plane wave basis), the inputs required are:

- **Structure of Material (Including boundary conditions):** The structure of the material is required as input for the DFT calculations. The ase.Atoms object [26] is a data structure used to represent material structures in python
- **PAW dataset and Pseudopotentials:** These are necessary inputs for the DFT calculation in the PAW framework and are required to define the PAW transformation for the calculation (See sections 2.1.1.6 and C for more details).
- **Exchange-Correlation:** Depending on the quantity of interest, the choice of approximation of the exchange-correlation potential is really important (See section 2.1.1.2 for more details and section 2.1.1.5 for good choices of exchange-correlation for bandgap estimation).
- **K-Point Sample:** By Bloch theorem we can write the eigenstates of the Kohn-Sham Hamiltonian as:

$$\phi_{m,\mathbf{k}}(r) = u_{m,\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \quad (12)$$

Thus, the ground state electron density becomes:

$$n(\mathbf{r}) = \sum_{m=1}^{VB} \int_{\text{Brillouin Zone}} |\phi_{m,\mathbf{k}}(\mathbf{r})|^2 d^3\mathbf{k} = \sum_{m=1}^{VB} \int_{\text{Brillouin Zone}} |u_{m,\mathbf{k}}(\mathbf{r})|^2 d^3\mathbf{k} \quad (13)$$

Thus, to estimate the integral over k-points, we need to sample the k-points in the Brillouin zone to convert the integral to a finite sum.

- **Kinetic Energy cutoff:** By the Bloch theorem, the Kohn-Sham states can be expanded as equation 12, where $u_{m,\mathbf{k}}(\mathbf{r})$ is a function having the periodicity of the lattice. Thus, the Fourier basis (which is a good basis for periodic functions) is a natural choice for the basis for the Kohn-Sham states. In the Fourier basis, every Kohn-Sham state can be expanded into the plane wave fourier basis as:

$$\phi_{m,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} C_{m,\mathbf{k},\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \quad (14)$$

Where $\mathbf{G}$ are vectors in the reciprocal space given by:

$$\mathbf{G} = \left(2\pi n_a \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, 2\pi n_b \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, 2\pi n_c \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})} \right) \quad (15)$$

Where $n_a, n_b, n_c \in \mathbb{Z}$. Thus, for a fixed K-point grid, the Fourier basis expansion forms a countably infinite basis set. To make the basis finite for computability, we bound the basis by a parameter called kinetic energy cutoff (ϵ_{cut}) and restrict the basis to elements satisfying:

$$\frac{||\mathbf{k} + \mathbf{G}||^2}{2} \leq \epsilon_{cut} \quad (16)$$

Article [27] also defines an automated scheme for finding the optimal values of the kinetic energy cutoff and k-point sampling.

- **Density mixing parameters:** These are really important in parameters for the SCF step and determine how the electron densities get updated between the SCF iterations, thus influencing the rate of convergence of the SCF step. Pulay mixing scheme is one density mixing scheme which has 3 input parameters -

the linear mixing parameter β , the number of old densities mixed n_{max} , and the weight parameter w for defining the Pulay loss metric (see section 2.1.1.4 for more details).

- **Occupation Numbers (Density Smearing):** In metallic systems, there are instabilities and discontinuities in the calculations close to the Fermi level [28]. To tackle this problem, people usually smear the electron density close to the Fermi-level by changing the occupation numbers of the Kohn-Sham states such that states above the Fermi level also have a finite probability to be occupied (common choices are Fermi-Dirac smearing and Marzari-Vanderbilt smearing). As our applications deal with semiconducting materials, we don't smear our electron densities and use 0 temperature Fermi-Dirac occupation for the Kohn-Sham states.

2.1.1.2 Choice of Exchange Correlation Potential

The exchange-correlation potential incorporates the exchange and correlation effects of electrons. The exact form of the exchange correlation potential is unknown, which makes the exact exchange-correlation very difficult to compute (although articles like [29, 30] discuss computations of the exact exchange-correlation functional, their approaches face scalability issues, thus making these methods infeasible). Because of the difficulty of the computation of the exact exchange-correlation, the exchange-correlation functional has to be approximated. Depending on the application and required accuracy, the exchange-correlation potential has to be carefully chosen. The hierarchy between exchange-correlation potentials is given by Jacob's ladder [31] (Figure 2).

The different rungs of Jacob's ladder are explained below:

- **Local Density Approximation (LDA):** These are the simplest exchange-correlation potentials, that only require the information of the electron density at each point. They are cheapest to compute and are local in nature.
- **Generalized Gradient Approximation (GGA):** These are a bit more complicated than the LDA potentials as they require the information about the gradients of the electron density. These exchange-correlation potentials are semi-local

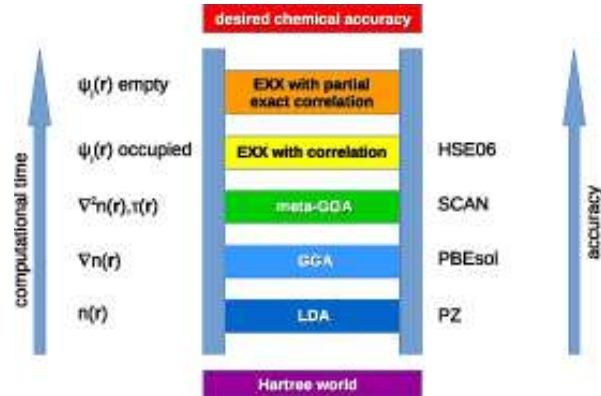


Figure 2: Jacob’s Ladder for Exchange-Correlation Potentials [31]

in nature due to the requirement of gradients.

- **Meta-Generalized Gradient Approximation (meta-GGA):** These are more expensive than the GGA potentials as they require the information of second derivatives in their computation. They are also semi-local in nature as they require information of first and second-order derivatives.
- **Hybrid Functionals:** These are much more expensive than the meta-GGA potentials as they require the information of all the occupied Kohn-Sham states to be calculated. These are non-local potentials with non-local exact Hartree-Fock exchange and local correlation.
- **Doubly Hybrid Functionals:** These are the most expensive exchange-correlation potentials. These require the knowledge of unoccupied Kohn-Sham states as well. These potentials are highly non-local with non-local Hartree-Fock exchange and non-local correlation.

2.1.1.3 The Self Consistent Flow Pipeline

As Kohn-Sham equations are non-linear eigenvalue equations, they are difficult to solve directly. Thus, they are usually solved self-consistently. Figure 3 shows the flow pipeline for the SCF calculations in DFT calculations.

The SCF calculations give us the ground state electron density and the occupied Kohn-Sham eigenstates of the system. Article [33] talks about the nature of the convergence of the SCF iterations. The update step of the electron density in the SCF calculations

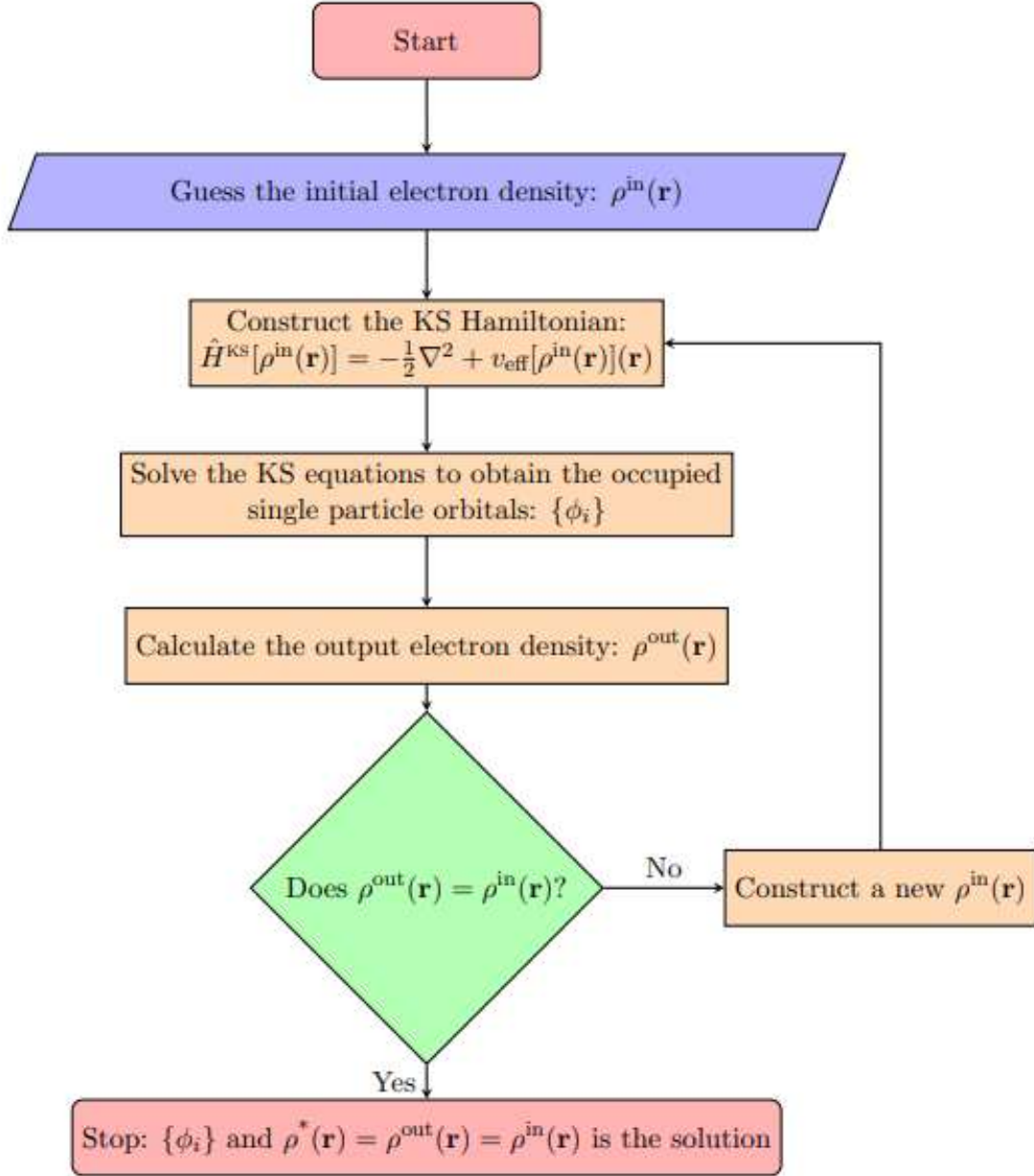


Figure 3: The Self-Consistent Flow Pipeline [32]

is really important as the update scheme of the electron density guarantees rate of convergence of the SCF calculations. Reference [32] discusses and compares various density update schemes for the SCF step. In our implementations, we use the Pulay Mixing scheme. For diagonalizing the Kohn-Sham Hamiltonian, we can use iterative diagonalization methods like Krylov-subspace based methods [34].

2.1.1.4 Pulay Mixing

Pulay mixing [35] (also called the direct inversion in the iterative subspace method) is one of the possible strategies to update the electron densities during the SCF step. In this density update scheme, we store a history of residues $\mathcal{R}[n_{in}(\mathbf{r})]$ to extrapolate the density that should minimize the modulus of the residual. The residual function is defined as:

$$\mathcal{R}[n(\mathbf{r})] = n_{new}[n(\mathbf{r})] - n(\mathbf{r}) \quad (17)$$

Where $n_{new}[n(\mathbf{r})]$ is the new value of the electron density after diagonalizing the Kohn-Sham Hamiltonian $H_{KS}[n(\mathbf{r})]$. In the Pulay mixing scheme, it is assumed that the new electron density should be of the form:

$$n^{(i+1)}(\mathbf{r}) = \sum_{j=\max(0, i-n_{max})}^i c_j n^{(j)}(\mathbf{r}) \quad (18)$$

Where $n^{(j)}$ is the electron density input of the j^{th} iteration. Moreover, the Kohn-Sham density update functional n_{new} is approximated to a linear function, allowing us to write:

$$n_{new}[n^{(i+1)}(\mathbf{r})] = \sum_{j=\max(0, i-n_{max})}^i c_j n_{new}[n^{(j)}(\mathbf{r})] \quad (19)$$

Thus, we the residue function also follows:

$$\mathcal{R}[n^{(i+1)}(\mathbf{r})] = \sum_{j=\max(0, i-n_{max})}^i c_j \mathcal{R}[n^{(j)}(\mathbf{r})] \quad (20)$$

Due to the preservation of normalization of electron densities, we have a constraint on c_j given by:

$$\sum_{j=\max(0, i-n_{max})}^i c_j = 1 \quad (21)$$

The loss functional to find optimal values of basis coefficients is given by:

$$\mathcal{L}[n^{(i+1)}(\mathbf{r})] = \int |\mathcal{R}[n^{(i+1)}(\mathbf{r})]|^2 d^3\mathbf{r} \quad (22)$$

In practice, if the residual loss in the initial conditions is large, this search may give suboptimal results. Thus, we regularize the density update as:

$$n^{(i+1)}(\mathbf{r}) = \sum_{j=\max(0, i-n_{\max})}^i c_j (n^{(j)}(\mathbf{r}) + \beta \mathcal{R}[n^{(j)}(\mathbf{r})]) \quad (23)$$

Where β is a linear mixing parameter (that lies between 0 and 1).

The implementation described above gives equal weights to all previous density iterations. But, in practice, we wish to give higher weights to newer densities as compared to older densities. This is done in the GPAW library [36] by introducing a metric F_w to change the loss function as:

$$\mathcal{L}[n^{(i+1)}(\mathbf{r})] = \int \langle \mathcal{R}[n^{(i+1)}(\mathbf{r})], F_w \mathcal{R}[n^{(i+1)}(\mathbf{r})] \rangle d^3\mathbf{r} \quad (24)$$

Where F_w is parametrized by the weights w . There are also extensions of Pulay Mixing like the periodic Pulay Mixing [37] which give even faster convergence.

2.1.1.5 Bandgap Estimation Using DFT

The bandgap [38] refers to the minimum energy required to excite an electron in a quantum system. Metals have a 0 bandgap while semiconductors and insulators have finite non-zero bandgaps. In the context of Kohn-Sham theory, there are a few kinds of bandgaps that are relevant. These are:

- **Kohn-Sham Gap:** This is the simplest kind of bandgap that can be calculated directly using Kohn-Sham equations. The Kohn-Sham gap can only give us qualitative idea of material properties and the value of the Kohn-Sham gap doesn't have much physical significance. The Kohn-Sham gap Δ_{KS} is given as:

$$\Delta_{KS} = \epsilon_{CBM} - \epsilon_{VBM} \quad (25)$$

Where ϵ_{CBM} is the conduction band minimum energy and ϵ_{VBM} is the valence

band minimum energy, defined as:

$$\epsilon_{CBM} = \inf_{m, \mathbf{k} \in \text{unoccupied}} \epsilon_{m, \mathbf{k}} \quad (26)$$

$$\epsilon_{VBM} = \sup_{m, \mathbf{k} \in \text{occupied}} \epsilon_{m, \mathbf{k}} \quad (27)$$

When $\arg \max_{\mathbf{k} \in \text{occupied}} \epsilon_{m, \mathbf{k}} = \arg \min_{\mathbf{k} \in \text{unoccupied}} \epsilon_{m, \mathbf{k}}$, the material is called a direct bandgap material, otherwise, it is called an indirect bandgap material.

- **Fundamental Bandgap:** The fundamental bandgap [39] tells us the difference in energy between adding a conduction electron and removing a valence electron from the system. Thus, the fundamental bandgap Δ_{fun} is defined as:

$$\Delta_{fun} = I - A \quad (28)$$

Where I is the ionization energy and A is the electron affinity. In the context of Kohn-Sham calculations, it can be calculated as:

$$\Delta_{fun} = \Delta_{KS} + \Delta_{xc} \quad (29)$$

Here, Δ_{xc} is the derivative discontinuity correction which arises due to the discontinuity of the derivative of the Kohn-Sham energy with respect to addition and removal of an electron.

The fundamental bandgap gives us details of the electrical properties of the material and is also used to classify materials into metals, semiconductors, and insulators.

- **Optical bandgap:** The optical bandgap is the minimum energy of a photon required to excite an electron in a material. From the DFT calculations, the optical bandgap Δ_{opt} can be calculated as:

$$\Delta_{opt} = \epsilon_{CBM}^{SOC} - \epsilon_{VBM}^{SOC} + \Delta_{xc} - \Delta_{e-h} = \Delta_{fun} - \Delta_{SOC} - \Delta_{e-h} \quad (30)$$

Here, ϵ_{CBM}^{SOC} and ϵ_{VBM}^{SOC} are spin-orbit corrected eigenvalues, Δ_{SOC} is the spin-orbit

correction to the bandgap ($\epsilon_{CBM}^{SOC} - \epsilon_{VBM}^{SOC} = \Delta_{KS} - \Delta_{SOC}$), and Δ_{e-h} is the electron-hole interaction (exciton binding energy). The spin-orbit correction is computed in the NSCF step of the DFT calculations and can be calculated in 2 ways - we can compute the corrections to each eigenvalue separately (in which we use the first form), or we calculate the correction to the net bandgap (in which we use the second form) (See article [36] for more details). The exciton binding energy has to be calculated using the Bethe-Salpeter equations [40] or Time-Dependent DFT [24].

While calculating the bandgap using DFT, we need to carefully sample the k-points from the Brillouin zone so that our sampling captures the conduction band minimum and valence band maximum correctly. Moreover, to get good estimates of bandgap, selection of a good exchange-correlation potential is really important. Articles [41, 42] compare various exchange-correlation functionals for bandgap calculations. Out of these, the GLLB-sc [43] (GGA), HLE17 [44] (meta-GGA), and HSE06 [45] (Hybrid) are good choices of exchange correlation potentials for bandgap calculations. Other ways of calculating the bandgap of the material include the GW approximation [46] (which is the most accurate method for bandgap calculations in DFT framework), and the $K \cdot P$ method [47] (which is a fully wavefunction based method).

2.1.1.6 Projector Augmented Wave Method

The Projector Augmented Wave (PAW) method [48] is a method of simplifying the Kohn-Sham equations. At the core of PAW lies an invertible linear transformation ($\hat{\mathcal{T}}$) that smoothes out all the quantum states near the atomic core region. The smoothening is given by:

$$|\phi_i\rangle = \hat{\mathcal{T}} |\tilde{\phi}_i\rangle \quad (31)$$

Here, $|\tilde{\phi}_i\rangle$ are the smoothened Kohn-Sham states and $|\phi_i\rangle$ are the actual Kohn-Sham states. The transformation modifies the Kohn-Sham equations to:

$$\hat{\mathcal{T}}^\dagger H_{KS} \hat{\mathcal{T}} |\tilde{\phi}_i\rangle = \epsilon_i \hat{\mathcal{T}}^\dagger \hat{\mathcal{T}} |\tilde{\phi}_i\rangle \quad (32)$$

For the exact definition of these transformation operators, see appendix section C.

In the PAW formalism, the frozen-core approximation is very common, which is an approximation that removes the need for projector states for the core electronic states. We can also do relaxed core PAW [49] to get a better approximation of the Kohn-Sham solutions. Also, physical operators in PAW lose some of their uniqueness. There is a freedom to add $B - \sum_{i,j} |p_i\rangle \langle \tilde{\phi}_i| B | \tilde{\phi}_j\rangle \langle p_j|$ to any quantum operator without changing any results (B is an arbitrary operator that is localized in the augmentation regions). Using this freedom, we can also simplify potentials to find smoother pseudopotentials that we can use for our calculations.

2.1.1.7 Lagrangian Method for Molecular Dynamics in DFT A way to perform the PAW Kohn-Sham equations as a Lagrangian optimization to construct the Lagrangian given by:

$$\mathcal{L} = \sum_i m_i \langle \dot{\tilde{\psi}}_i | \dot{\tilde{\psi}}_i \rangle + \sum_a \frac{1}{2} M_a ||\dot{\mathbf{R}}||_a^2 - E(|\tilde{\psi}\rangle, \mathbf{R}) + \sum_{i,j} (\langle \tilde{\psi}_i | \tilde{O} | \tilde{\psi}_j \rangle - \delta_{ij}) \Lambda_{i,j} \quad (33)$$

Here, E is the total Kohn-Sham energy, $\tilde{O}$ is the overlap operator, and a is the nuclear index. This method maps the Kohn-Sham equations as a dynamical problem, and the dynamics with respect to $\tilde{\psi}_i$ will give the Kohn-Sham equations while the dynamics with respect to R will make it so that the forces on the atoms become 0 and the lattice becomes unstrained. The last term is a Lagrange multiplier constraint that ensures the orthogonality of the Kohn-Sham states. We can also add a damping term to this Lagrangian to make sure that the dynamics converge. The Euler-Lagrange equations of this system can be solved using the Verlet algorithm [50].

2.1.2 Density Functional Perturbation Theory

Density Functional Perturbation Theory [51, 52] (DFPT) is a method for derivative calculation in the DFT framework. Consider that Kohn-Sham states can be expanded as:

$$\phi_{n,k} = \sum_{\alpha} C_{\alpha,n,k} \chi_{\alpha} \quad (34)$$

Where χ_i are the Fourier basis elements. Then, we can write the Kohn-Sham equations in matrix form as:

$$\sum_{\beta} H_{\alpha\beta} C_{\beta\gamma} = \epsilon_{\gamma} \sum_{\beta} S_{\alpha\beta} C_{\beta\gamma} \quad (35)$$

Here, γ is the pair (n, k) , $H_{ij} = \langle \chi_i | H_{KS} | \chi_j \rangle$, and $S_{ij} = \langle \chi_i | \chi_j \rangle$.

From the definition of the basis expansion, we can say that the gradient of the Kohn-Sham states is given as:

$$\nabla_q \phi_{\gamma} = \sum_{\alpha} (C_{\alpha,\gamma} (\nabla_q \chi_{\alpha}) + (\nabla_q C_{\alpha,\gamma}) \chi_{\alpha}) \quad (36)$$

Also, the gradient of the electron density is given as:

$$\nabla_q n(\mathbf{r}) = 2Re \left(\sum_{\gamma} (\nabla_q \phi_{\gamma}^*(\mathbf{r})) \phi_{\gamma}(\mathbf{r}) \right) \quad (37)$$

The derivatives of the Kohn-Sham states follow the Sternheimer equation, given by:

$$(H_{KS} - \epsilon_{\gamma}) (\nabla_q |\phi_{\gamma}\rangle) = -(\nabla_q H_{KS} - \nabla_q \epsilon_{\gamma}) |\phi_{\gamma}\rangle \quad (38)$$

The derivatives of the Kohn-Sham energies are given as:

$$\nabla_q \epsilon_{\gamma} = \langle \phi_{\gamma} | (\nabla_q H_{KS}) | \phi_{\gamma} \rangle \quad (39)$$

Putting the basis expansion into the Sternheimer equations, we get:

$$\sum_{\beta} ((H_{\alpha\beta} - \epsilon_{\gamma} S_{\alpha\beta}) (\nabla_q C_{\beta\gamma}) - \epsilon_{\gamma} (\nabla_q S_{\alpha\beta}) C_{\beta\gamma}) = - \sum_{\beta} (\nabla_q H_{\alpha\beta} - (\nabla_q \epsilon_{\gamma}) S_{\alpha\beta}) C_{\beta\gamma} \quad (40)$$

One way to solve these equations is to use the Coupled Perturbed Self-Consistent Field (CPSCF) method, which converts the derivative operator to matrix form (see eqns 16 and 17 in [52]). It is better to simply use self-consistent Sternheimer equations instead of using CPSCF expansion as the CPSCF solution requires knowledge of excited states for accuracy.

The gradient of the Kohn-Sham Hamiltonian is given by:

$$\nabla_q H_{KS}(\mathbf{r}) = \nabla_q V_{ext}[\mathbf{q}](\mathbf{r}) + \frac{\delta V_{Hxc}}{\delta n} \nabla_q n(\mathbf{r}) \quad (41)$$

Reference [53] talks about self-consistency in DFPT calculations using Localized Basis sets. The amount of information required for derivative calculation via perturbative techniques is given by the $2n + 1$ theorem, which states that the information of derivatives of the eigenstates upto the n^{th} order for a linear Hamiltonian gives us access to the $2n + 1^{th}$ derivatives of the eigenvalues.

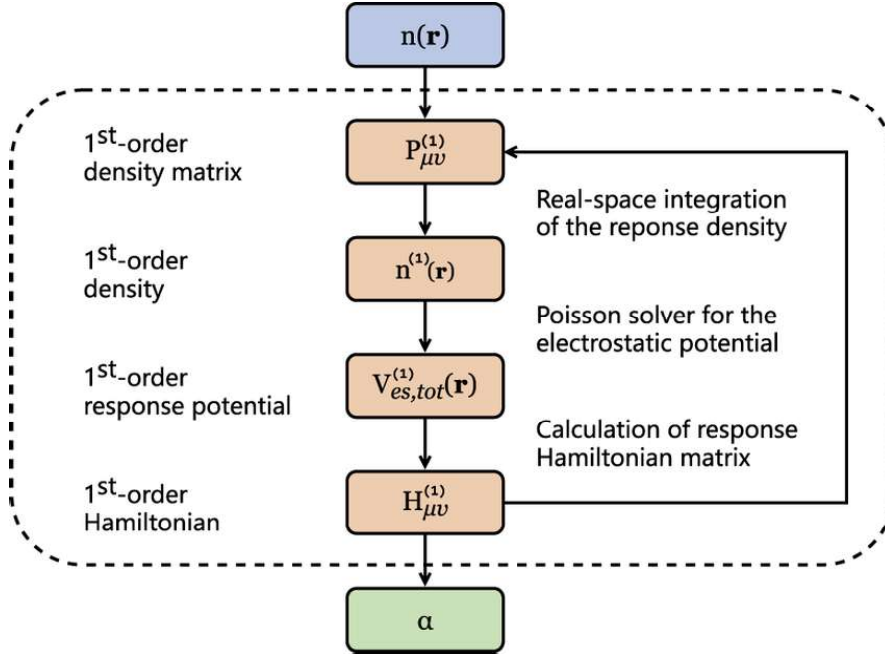


Figure 4: The Density Function Perturbation Theory Pipeline [54]

2.2 Properties and Loss Functions for Photovoltaic Materials

Article [55] gives various task-specific loss functions for photovoltaic materials, which they optimize via a genetic algorithm. In their approach, they consider a 1-D nanocrystal (20 atoms in length) with fixed atomic positions and vary the chemical types of the atoms in the first period of the periodic table to find the optimal chemical profile that optimizes one of four quantities relevant for HOMO (highest occupied molecule orbital) - LUMO (lowest unoccupied molecule orbital) transitions in the nanocrystal. These quantities are:

- **DFT bandgap:** The DFT bandgap is given by the difference in the Kohn-Sham eigenvalues corresponding to the HOMO and LUMO eigenstates. Thus, the DFT band gap is given as:

$$\Delta_{KS} = \epsilon_L - \epsilon_H \quad (42)$$

Where ϵ_L is the LUMO energy (also called the Conduction Band Minimum energy (ϵ_{CBM}) for periodic lattices), and ϵ_H is the HOMO energy (also called the valence band maximum energy (ϵ_{VBM}) for periodic lattices). For this problem, they minimize the loss functional given as:

$$\mathcal{L} = |\Delta_{KS} - \Delta^*|^2 \quad (43)$$

Here Δ^* is the desired value of the band gap.

- **Charge Transfer:** The charge transfer determines how far an electron can travel after being excited. For photovoltaic materials, a large charge transfer value is desirable. Thus, they maximize the functional value given by:

$$-\mathcal{L} = \langle \phi_L | \mathbf{r} \cdot \mathbf{e} | \phi_L \rangle - \langle \phi_H | \mathbf{r} \cdot \mathbf{e} | \phi_H \rangle \quad (44)$$

Here, ϕ_H and ϕ_L are the HOMO and LUMO eigenstates, respectively, $\mathbf{r}$ is the position operator and $\mathbf{e}$ is the unit vector in the direction of charge transfer. The $-$ sign on the LHS represents that this is a maximization and not a minimization.

- **Lifetime of Excited State:** For photovoltaic applications, long lifetimes are required to be able to harvest excited electrons from the system. For this application, they minimize the loss function given as:

$$\mathcal{L} = ||[H_{KS}[n_{exc}(\mathbf{r})], \rho_{exc}]||_{HS}^2 \quad (45)$$

Here, $||\cdot||_{HS}$ represents the Hilbert–Schmidt norm ($||A||_{HS} = \sqrt{\text{tr}(A^\dagger A)}$), $[\cdot, \cdot]$ represents the commutator, $n_{exc}(r)$ is the excited state electron density ($n_{exc}(\mathbf{r}) = n(\mathbf{r}) - \phi_H^*(\mathbf{r})\phi_H(\mathbf{r}) + \phi_L^*(\mathbf{r})\phi_L(\mathbf{r})$), and ρ_{exc} is the excited state density matrix ($\rho_{exc} = (\sum_{j \in \text{occupied}} |\phi_j\rangle \langle \phi_j|) - |\phi_H\rangle \langle \phi_H| + |\phi_L\rangle \langle \phi_L| = \rho - |\phi_H\rangle \langle \phi_H| + |\phi_L\rangle \langle \phi_L|$).

- **Electron-hole Overlap:** A low value of electron-hole overlap is required for good photovoltaic materials. The loss functional associated with this quantity is:

$$\mathcal{L} = \int |\phi_H(\mathbf{r})|^2 |\phi_L(\mathbf{r})|^2 d^3\mathbf{r} \quad (46)$$

2.2.1 Choice of Target Bandgap

Bandgap is a very important property for photovoltaic cells. The value of the bandgap of a material gives us the maximum theoretic efficiency of a solar cell. The functional form of this dependency is given by a detailed balance limit for the solar cell. This limit depends on the type of solar cell being used. The Shockley-Queisser limit [56, 57] gives a detailed balance limit for simple p-n junction solar cells, thus giving a relationship between bandgap of the cell and its efficiency, while the Vos limit [58, 59] gives a detailed balance limit for tandem solar cells, thus, giving a relationship between the efficiency of the solar cell with the bandgaps of the various layers.

2.3 The Search Space for Material Discovery

The space that we search for new materials is mainly divided into 3 classes. These are:

- **Structural Space:** The structural space determines the physical structure of the material. This space includes details like the boundary conditions (periodic or regular), dimensionality of the lattice [60], the size of the unit cell, and the atomic positions in the material. In the structural space, we can also allow the search for strained lattices [61, 62] which can be obtained through epitaxial growth on a carefully chosen substrate [63, 64], or by changing external environmental conditions like pressure [60, 65].
- **Chemical Space:** The chemical space includes change in the chemical composition of the material, which comes as a result of atomic or ionic substitutions [60].
- **Aggregation Space:** The aggregation space is the space of non-stoichiometric/topological changes in the material like the presence of grain boundaries [66],

defects [67], vacancies [68], or impurities [69] in the material.

2.4 Optimization Methods

Optimization is a fundamental problem in most computational problems as it requires the selection of the best possible solution out of a possible set of solutions. Usually, an optimization problem is defined by a loss (or reward) functional which has to be minimized (or maximized) over the search space. Simple optimization techniques like standard gradient descent [70] have a tendency to get stuck in local minima thus, we need to use more advanced methods for global minimization. Global minimization is a NP-hard problem [71] and thus, we need to find intelligent strategies that give approximate convergence to global optimizers most of the time. There are mainly 2 approaches to global minimization: adding momentum or adding stochasticity to the search (both of which allow us to escape local minima), or a combination of both. ADAM [72] and MT-ADAM [73] are some of the optimization techniques that use stochasticity in their minimizations to find optimal solutions. On the other hand, stochastic PMP-based optimizers (section 2.5.1.3) and langevin dynamics minimization [74, 75] are some of the methods that use a combination of momentum and stochasticity in their minimization strategies. Other global minimization strategies include simulated annealing (start with high amount of stochasticity (parametrized by temperature) and reduce stochasticity with time) [75–77], genetic algorithms (based on the principle on natural selection) [78], conformational flooding (traversing the search space over multiple runs and reducing depth of visited minima after each run) [79], and minima hopping (sample various local minima by traversing search space with momentum and based on if the minimum has been sampled before, update the search strategy) [80].

2.5 Optimal Control Theory

2.5.1 Pontryagin Maximum Principle

2.5.1.1 Deterministic PMP

Consider we wish to find the optimal control that extremizes the action functional

$(J(\mathbf{u}))$ given by:

$$J(\mathbf{u}) = \int_0^T L(\mathbf{q}, \mathbf{u}(t), t) dt + h(T, \mathbf{q}(T)) \quad (47)$$

Here, $\mathbf{q} \in \mathbb{R}^m$ is the state vector, $\mathbf{u} : [0, T] \rightarrow \mathcal{U} \subseteq \mathbb{R}^n$ is the control vector, $t \in [0, T]$ is the time, $L(\mathbf{q}, \mathbf{u}(t), t)$ is the lagrangian cost, and $h(T, q(T))$ is the terminal cost. Consider that the system is constrained by dynamics given by:

$$\dot{\mathbf{q}} = \mathbf{f}(\mathbf{q}, \mathbf{u}(t), t) \quad (48)$$

With the initial conditions given by:

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (49)$$

Then, by the Pontryagin Maximum Principle (PMP) [81], we can define an Optimal Control Hamiltonian functional ($\mathcal{H}$) (which we call the Optimaltonian), given by:

$$\mathcal{H}(\mathbf{q}, \mathbf{p}, \mathbf{u}, \lambda) = \langle \mathbf{p}, \mathbf{f}(\mathbf{q}, \mathbf{u}, t) \rangle + \lambda L(\mathbf{q}, \mathbf{u}, t) \quad (50)$$

Here, $\mathbf{p}$ is a vector called the costate (also sometimes called the momentum vector) [82] (which is also related to the sensitivity [83]) in the dual space of the state space, and $\lambda \in \mathbb{R}$ is an arbitrary real number that just rescales the Optimaltonian and the momentum (λ is usually chosen to be 1 or -1) (in further discussion, we consider $\lambda = 1$ unless stated otherwise). The dynamics of the system can be written as:

$$\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} \nabla_q \mathcal{H} \\ \nabla_p \mathcal{H} \end{bmatrix} \quad (51)$$

The terminal condition for the momentum is given by:

$$\mathbf{p}(T) = \nabla_q h(T, \mathbf{q}(T)) \quad (52)$$

The optimal control is thus given by:

$$\nabla_u \mathcal{H} = 0 \quad (53)$$

Assuming that $Hess_u(\mathcal{H})$ is non-singular, then by the implicit function theorem [84], we can write the optimal control $\mathbf{u}^*(t)$ uniquely as:

$$\mathbf{u}^*(t) = \mathbf{u}_0(\mathbf{q}, \mathbf{p}, t) \quad (54)$$

We can substitute this into the Optimaltonian to get the reduced Optimaltonian as:

$$\tilde{\mathcal{H}}(\mathbf{q}, \mathbf{p}, t) = \mathcal{H}(\mathbf{q}, \mathbf{p}, \mathbf{u}_0(\mathbf{q}, \mathbf{p}, t), t) \quad (55)$$

The symplectic Hamiltonian dynamics then get reduced to:

$$\begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} \nabla_q \tilde{\mathcal{H}} \\ \nabla_p \tilde{\mathcal{H}} \end{bmatrix} \quad (56)$$

2.5.1.2 Stochastic PMP

Consider we wish to find the optimal control that extremizes the action functional ($J(\mathbf{u})$) given by:

$$J(\mathbf{u}) = \mathbb{E} \left(\int_0^T L(\mathbf{q}, \mathbf{u}(t), t) dt + h(T, \mathbf{q}(T)) \right) \quad (57)$$

Here, $\mathbf{q} \in \mathbb{R}^m$ is the state vector random variable, $\mathbf{u} : [0, T] \rightarrow \mathcal{U} \subseteq \mathbb{R}^n$ is the control vector stochastic process, $t \in [0, T]$ is the time, $L(\mathbf{q}, \mathbf{u}(t), t)$ is the lagrangian cost, and $h(T, q(T))$ is the terminal cost. Consider that the system is constrained by dynamics given by:

$$d\mathbf{q} = \mathbf{f}(\mathbf{q}, \mathbf{u}(t), t)dt + \vec{\sigma}(\mathbf{q}, \mathbf{u}(t), t)d\mathbf{W}_t \quad (58)$$

With the initial conditions given by:

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (59)$$

Then, by the stochastic pontryagin maximim principle [85], we can define the Optimal Control Hamiltonian (Optimaltonian) ($\mathcal{H}$) as :

$$\mathcal{H}(\mathbf{q}, \mathbf{u}, \mathbf{p}, \vec{K}, t) = L + \langle \mathbf{p}, \mathbf{f} \rangle + \sum_{i,j} K_{ij}(t) \sigma_{ij}(\mathbf{q}, \mathbf{u}(t), t) \quad (60)$$

Where $(\mathbf{p}, \vec{K})$ is a unique stochastic process (which is like the momentum). Given that the control domain $\mathcal{U}$ is convex and $\mathbf{f}$ and $\vec{\sigma}$ are C^1 with respect to u , then, for the optimal control u^* , almost everywhere and almost surely, we have:

$$\langle \nabla_u \mathcal{H}(\mathbf{q}^*, \mathbf{u}^*, \mathbf{p}^*, \vec{K}^*, t), (\mathbf{u} - \mathbf{u}^*) \rangle \geq 0 \quad \forall \mathbf{u} \in \mathcal{U} \quad (61)$$

One of the solutions of this equation is given by $\nabla_u \mathcal{H} = 0$. The dynamics of the system are then given by:

$$\begin{bmatrix} d\mathbf{q} \\ d\mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} \nabla_q \mathcal{H} \\ \nabla_p \mathcal{H} \end{bmatrix} dt + \begin{bmatrix} \vec{\sigma}(\mathbf{q}, \mathbf{u}, t) \\ \vec{K}(t) \end{bmatrix} d\mathbf{W}_t \quad (62)$$

The terminal condition for the momentum is given by:

$$\mathbf{p}(T) = \nabla_q h(T, \mathbf{q}(T)) \quad (63)$$

2.5.1.3 Using Stochastic PMP for Black-Box Minimization

Consider we wish to minimize a function $\mathcal{F}(\mathbf{q})$. We then convert this into a PMP optimization problem taking the stochastic action as:

$$J(\mathbf{u}) = \mathbb{E} \left(\int_0^T e^{\gamma t} \left(\frac{1}{2} \langle \mathbf{u}, \vec{M} \mathbf{u} \rangle - \mathcal{F}(\mathbf{q}) \right) dt \right) \quad (64)$$

The system is assumed to be subject to dynamics given by:

$$d\mathbf{q} = \mathbf{u} dt + \vec{\epsilon} e^{-\alpha t} d\mathbf{W}_t \quad (65)$$

With initial condition:

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (66)$$

Here, $\vec{M}$ is the mass/inertia tensor (which takes care of amount of movement each step, and is symmetric positive definite), γ is a damping constant (which ensures that the drift part of the dynamics converges), $\vec{\epsilon}$ is the initial stochasticity in the dynamics, and α is the discount factor for the stochasticity that ensures that the system starts

converging to deterministic dynamics after some time. This formulation of action has been motivated by the action of a particle moving through a potential $\mathcal{F}(\mathbf{q})$ and experiencing a viscous drag γ . Similar formulations of stochastic actions are obtained in the continuous formulation of reinforcement learning [15]. The Optimaltonian for this system is given by:

$$\mathcal{H}(\mathbf{q}, \mathbf{u}, \mathbf{p}, \vec{K}, t) = e^{\gamma t} \left(\frac{1}{2} \langle \mathbf{u}, \vec{M} \mathbf{u} \rangle - \mathcal{F}(\mathbf{q}) \right) + \langle \mathbf{p}, \mathbf{u} \rangle + e^{-\alpha t} \langle \vec{\epsilon}, \vec{K} \rangle \quad (67)$$

Where the tensor inner product $\langle \vec{A}, \vec{B} \rangle = \text{tr}(\vec{A}^T \vec{B})$. Thus, $\nabla_u \mathcal{H}$ is given by:

$$\nabla_u \mathcal{H}(\mathbf{q}, \mathbf{u}, \mathbf{p}, \vec{K}, t) = e^{\gamma t} \vec{M} \mathbf{u} + \mathbf{p} \quad (68)$$

Thus, the Optimal Control $\mathbf{u}^*$ is given by:

$$\langle e^{\gamma t} \vec{M} \mathbf{u}^* + \mathbf{p}, (\mathbf{u} - \mathbf{u}^*) \rangle \geq 0 \quad \forall \mathbf{u} \quad (69)$$

Thus, one of the possible solutions for this inequality is given by:

$$\mathbf{u}^* = -e^{-\gamma t} (\vec{M})^{-1} \mathbf{p} \quad (70)$$

This solution of the inequality is also unique (see appendix section B for proof). As this solution is equivalent to putting $\nabla_u \mathcal{H} = 0$, we can calculate the reduced Optimaltonian as:

$$\tilde{\mathcal{H}} = - \left(e^{-\gamma t} \frac{1}{2} \langle \mathbf{p}, (\vec{M})^{-1} \mathbf{p} \rangle + e^{\gamma t} \mathcal{F}(\mathbf{q}) \right) + e^{-\alpha t} \langle \vec{\epsilon}, \vec{K} \rangle \quad (71)$$

The stochastic-symplectic Hamilton's equations for this system would be given by:

$$\begin{bmatrix} d\mathbf{q} \\ d\mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} -e^{\gamma t} \nabla_q \mathcal{F}(\mathbf{q}) \\ -e^{-\gamma t} (\vec{M})^{-1} \mathbf{p} \end{bmatrix} dt + \begin{bmatrix} \vec{\epsilon} e^{-\alpha t} \\ \vec{K}(t) \end{bmatrix} d\mathbf{W}_t \quad (72)$$

We are also able to write dynamics of the optimal control as:

$$d\mathbf{u} = -\gamma \mathbf{u} dt - e^{-\gamma t} (\vec{M})^{-1} d\mathbf{p} = -\gamma \mathbf{u} dt - (\vec{M})^{-1} \nabla_q \mathcal{F} dt - e^{-\gamma t} (\vec{M})^{-1} \vec{K}(t) d\mathbf{W}_t \quad (73)$$

Motivating that the dynamics should converge to deterministic dynamics, we parametrize $\vec{K}(t) = \vec{\kappa}e^{(\gamma-\beta)t}$. Thus, the dynamics of the optimal control reduce to:

$$d\mathbf{u} = -\gamma\mathbf{u}dt - (\vec{M})^{-1}\nabla_q\mathcal{F}dt - (\vec{M})^{-1}\vec{\kappa}e^{-\beta t}d\mathbf{W}_t \quad (74)$$

And the symplectic-stochastic dynamics reduce to:

$$\begin{bmatrix} d\mathbf{q} \\ d\mathbf{p} \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} -e^{\gamma t}\nabla_q\mathcal{F}(\mathbf{q}) \\ -e^{-\gamma t}(\vec{M})^{-1}\mathbf{p} \end{bmatrix} dt + \begin{bmatrix} \vec{\epsilon}e^{-\alpha t} \\ \vec{\kappa}e^{(\gamma-\beta)t} \end{bmatrix} d\mathbf{W}_t \quad (75)$$

2.5.2 Hamilton-Jacobi-Bellmann Method

The Hamilton-Jacobi-Bellmann (HJB) method [86] is an alternative method for solving stochastic optimal control problems. Consider we wish to find the optimal control that extremizes the action functional ($J(\mathbf{u})$) given by:

$$J(\mathbf{u}) = \mathbb{E} \left(\int_0^T L(\mathbf{q}, \mathbf{u}(t), t) dt + h(T, \mathbf{q}(T)) \right) \quad (76)$$

Here, $\mathbf{q} \in \mathbb{R}^m$ is the state vector random variable, $\mathbf{u} : [0, T] \rightarrow \mathcal{U} \subseteq \mathbb{R}^n$ is the control vector stochastic process, $t \in [0, T]$ is the time, $L(\mathbf{q}, \mathbf{u}(t), t)$ is the lagrangian cost, and $h(T, q(T))$ is the terminal cost. Consider that the system is constrained by dynamics given by:

$$d\mathbf{q} = \mathbf{f}(\mathbf{q}, \mathbf{u}(t), t)dt + \vec{\sigma}(\mathbf{q}, \mathbf{u}(t), t)d\mathbf{W}_t \quad (77)$$

With the initial conditions given by:

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (78)$$

The optimal control is the control that minimizes the stochastic action, that is:

$$\mathbf{u}^* = \operatorname{argmin}_{\mathbf{u} \in \mathcal{U}} J(\mathbf{u}) \quad (79)$$

Then, we define the cost to go functional as:

$$C(s, \mathbf{x}) = \min_{\mathbf{u} \in \mathcal{U}} \mathbb{E} \left(\int_s^T L(\mathbf{q}, \mathbf{u}(t), t) dt + h(T, \mathbf{q}(T)) \right) | \mathbf{q}(s) = \mathbf{x} \quad (80)$$

By the dynamic programming principle [87], this is the same as:

$$C(s, \mathbf{x}) = \mathbb{E} \left(\int_s^T L(\mathbf{q}, \mathbf{u}^*(t), t) dt + h(T, \mathbf{q}(T)) \right) | \mathbf{q}(s) = \mathbf{x} \quad (81)$$

Moreover, by definition, $C(0, \mathbf{q}_0) = J(\mathbf{u}^*)$. Then, we can define an Optimal Control Hamiltonian (Optimaltonian) as:

$$\mathcal{H}(\mathbf{q}, \mathbf{u}, t) = \langle \nabla_{\mathbf{q}} C(t, \mathbf{q}), \mathbf{f}(\mathbf{q}, \mathbf{u}, t) \rangle + L + \frac{1}{2} \langle \vec{\Sigma}, \text{Hess}_{\mathbf{q}}(C) \rangle \quad (82)$$

Here, $\vec{\Sigma}$ is given by:

$$\vec{\Sigma} = (\vec{\sigma})(\vec{\sigma})^T \quad (83)$$

The optimal control is given by:

$$\nabla_{\mathbf{u}} \mathcal{H} = 0 \quad (84)$$

Moreover, the cost to go function follows the HJB dynamics, given by:

$$\frac{\partial C}{\partial t} + \mathcal{H} = 0 \quad (85)$$

2.5.3 Optimal Control in Fokker-Planck Framework

Unlike the previous methods of stochastic optimal control, which deal with dynamics of random variables, we can also deal with optimal control by understanding flows of probability densities. The Fokker-Planck equation gives the evolution of the probability density undergoing a Gaussian stochastic process [88]. The formulation of optimal control within the Fokker-Planck formalism is provided in article [86]. Consider that a random variable $\mathbf{q}$ undergoes dynamics given by:

$$d\mathbf{q} = \mathbf{f}(\mathbf{q}, \mathbf{u}(t), t) dt + \vec{\sigma}(\mathbf{q}, \mathbf{u}(t), t) d\mathbf{W}_t \quad (86)$$

With the initial conditions given by:

$$\mathbf{q}(0) = \mathbf{q}_0 \quad (87)$$

Then, the Fokker-Planck equation is given by:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{f}) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial q_i \partial q_j} (\rho \Sigma_{ij}) \quad (88)$$

Here, $\rho(\mathbf{q}, t)$ is the probability density of random variable $\mathbf{q}$ at time t , and $\vec{\Sigma}$ is defined as:

$$\vec{\Sigma} = (\vec{\sigma})(\vec{\sigma})^T \quad (89)$$

The initial condition for the probability density is given by:

$$\rho(\mathbf{q}, 0) = \rho_0(\mathbf{q}) = \delta^m(\mathbf{q} - \mathbf{q}_0) \quad (90)$$

In this formalism, we are also able to take a more generalized initial condition for the probability density which would be equivalent to initializing with an ensemble of random variables sampled from the initial distribution.

Here, the functional to be optimized is given by:

$$J(\mathbf{u}) = \int_{\mathbf{q} \in \mathbb{R}^m} \left(\int_0^T \rho(\mathbf{q}, t) L(\mathbf{q}, \mathbf{u}, t) dt + \rho(\mathbf{q}, T) h(\mathbf{q}, T) \right) d^m \mathbf{q} \quad (91)$$

To find the optimal control for this problem, we define a function $\eta(\mathbf{q}, t)$ which acts as a lagrange multiplier for the dynamics constraint. Then, we can define an optimal control Hamiltonian (Optimaltonian) as:

$$\mathcal{H} = \mathbf{f} \cdot \nabla_{\mathbf{q}} \eta + L + \frac{1}{2} \langle \vec{\Sigma}, \text{Hess}_{\mathbf{q}}(\eta) \rangle \quad (92)$$

The optimal control is given by the condition:

$$\nabla_{\mathbf{u}} \mathcal{H} = 0 \quad (93)$$

Then, the optimal dynamics are given by:

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{f}) + \frac{1}{2} \sum_{i,j} \frac{\partial^2}{\partial q_i \partial q_j} (\rho \Sigma_{ij}) \quad (94)$$

$$\frac{\partial \eta}{\partial t} + \mathcal{H} = 0 \quad (95)$$

With boundary conditions:

$$\rho(\mathbf{q}, 0) = \rho_0(\mathbf{q}) \quad (96)$$

$$\eta(\mathbf{q}, T) = h(\mathbf{q}, T) \quad (97)$$

Comparing this with the Hamilton-Jacobi-Bellman method (section 2.5.2), we find that the lagrange multiplier η is the same as the cost to go function.

2.6 Surrogate Models

Surrogate models are computationally efficient approximations of expensive functions [16]. The Kohn-Sham calculations are pretty expensive calculations, and hence, it is really expensive to use the actual Kohn-Sham calculations for our search, especially when searching for new materials in a high dimensional search space. Surrogate models provide a way to approximate these expensive computations while keeping errors in a certain range. These surrogate approximations are much cheaper and faster than the actual calculators, making our search for new materials much more faster. Moreover, surrogate models are usually taken to be inherently differentiable, making it easier to perform differential optimization with these surrogate models. An example of surrogate modeling for bandgap is given by article [89], which uses a variational autoencoder to project the Kohn-Sham states into a lower dimensional space called the latent space and then uses that latent space information to learn a supervised model for GW bandgaps of materials.

2.7 The Gumbel-Softmax Reparametrization Trick

The chemical space, which represents the change in the atomic/ionic types in a material is represented by categorical variables (type of atom/ion present at a site). Thus,

it is not possible to perform differentiable optimization in the chemical space. Thus, to tackle this problem, we can use the gumbel-softmax reparametrization trick [17], which converts categorical distributions into continuous distributions. In this trick, we first encode the categorical variables using a one-hot encoding [90], along with providing a probability distribution for these one-hot vectors. After this, we reparametrize the one-hot vectors by softening the probability distribution using the gumbel-softmax distribution (the amount of softening is given by a 'temperature' parameter) to get a differentiable sample of the categorical variables.

2.8 Our Algorithm

We design a discovery algorithm that continuously learns to discover new materials with enhanced properties while accelerating its search each time it learns. Our algorithm is based on the Hamiltonian-based Stochastic Pontryagin Maximum Principle (Section 2.5.1.3), which is symplectic in nature. During the discovery dynamics, we locally learn the surrogate model $\tilde{\mathcal{F}}_\theta(\mathbf{q})$ for our target property $\mathcal{F}(\mathbf{q})$ (θ is the set of parameters that we learn to learn the surrogate and $\mathbf{q}$ is the representation of the material structure). The learning in our motion strategy is done using 2 neural networks - P_{NET} and H_{NET} . The P_{NET} learns the parameters for the motion strategy in our discovery (the initial momentum $\mathbf{p}_0$, the stochastic parts in the $\mathbf{q}$ and $\mathbf{p}$ ($\vec{\epsilon}$ and $\vec{\kappa}$), and the discount factors on these stochastic parts (α and β)), while the H_{NET} learns the local environment in the material search space (the surrogate $\tilde{\mathcal{F}}_\theta(\mathbf{q})$, the inertia tensor $\vec{M}$, and the viscous drag coefficient γ) (See section 2.5.1.3 for more details of these parameters). The loss function for learning these networks is given by:

$$\begin{aligned} \mathcal{L} = & c_1 \|\mathbf{p}(T)\|^2 + c_2 \|\vec{\kappa}\|^2 + c_3 \|\vec{\epsilon}\|^2 + c_4 \|\nabla_q \mathcal{F}(\mathbf{q}(T))\|^2 \\ & + \sum_i \left(c_5^{(i)} |\mathcal{F}(\mathbf{q}(t_i)) - \tilde{\mathcal{F}}_\theta(\mathbf{q}(t_i))|^2 + c_6^{(i)} \mathcal{F}(\mathbf{q}(t_i)) + c_7^{(i)} E_{KS}(\mathbf{q}(t_i)) + c_8^{(i)} \|\mathbf{p}(t_i) - \mathbf{P}_{NET}^{(p)}(\mathbf{q}(t_i))\|^2 \right) \end{aligned} \quad (98)$$

Where t_i are some checkpoint times during the dynamics which act as the data for training H_{NET} and P_{NET} . The first term ensures that the learned dynamics stop at the end of the dynamics (which is a result of us choosing to take the terminal cost to be 0 in PMP dynamics). The second and third terms penalize dynamics with high amount of stochasticity (thus ensuring that we traverse the search manifold with the lowest

amount of stochasticity to minimize the loss). The 4th term penalizes us if the stopping point has a finite gradient, thus ensuring that we stop at a critical point. The 5th term is the training loss between the actual loss and the surrogate, thus minimizing it would mean that we are learning a better surrogate. The 6th term represents our property loss at intermediate steps, thus minimizing it would lead to better properties. The 7th term is the Kohn-Sham energy, which would help us penalize less stable materials in our search. Finally, the 8th term is a gap between the momentum obtained from the dynamics and the momentum that we get from the learned network. Minimizing this loss would help us learn the momentum required to start from intermediate positions in the dynamics.

To initialize the surrogate, we can use data that we generate using pre-run Kohn-Sham calculations or we can use data from online databases like Computational Materials Repository [91] to train an initial surrogate. Before calculating each property of the material, we perform fixed-cell relaxations [92] (which relaxes the atomic positions in the unit cell without changing the unit cell, thus allowing us to discover strained lattices). We then run a few batches of stochastic PMP dynamics to collect the data from runs. After these few runs, we update the H_{NET} and P_{NET} until desired results are achieved. For solving these stochastic PMP dynamics, we use symplectic solvers like the ones mentioned in article [93]. More details for this are given in algorithm 1.

Algorithm 1 Our Discovery Algorithm

- 1: **Input:** $\mathbf{q}_0$ (Initial structures to start discovery)
 - 2: Initialize P_{NET} and H_{NET}
 - 3: Set the input parameters for GPAW calculation. (section 2.1.1.1)
 - 4: Initialize the dynamics using H_{NET} and P_{NET}
 - 5: Take a minibatch of the structures
 - 6: Run Stochastic PMP dynamics on these structures (Algorithm 2)
 - 7: Collect data from some checkpoints on these dynamic trajectories
 - 8: Perform GPAW calculations on this data to get the actual values of the property loss $\mathcal{F}(\mathbf{q})$ (section 2.2) and the Kohn-Sham energy E_{KS} for these data points.
 - 9: Update the P_{NET} and H_{NET} by minimizing loss given by equation 98 using MTADAM (Section 2.4).
 - 10: **if** $\mathcal{F}(q(T)) > tol$ or $nit \leq n_{max}$ **then** Go to 5.
-

Algorithm 2 Discovery Dynamics during Search

- 1: **Input:** $\mathbf{q}_0$ (Minibatch of initial structures to start discovery)
 - 2: Perform fixed cell relaxations to obtain stable atomic configurations in the (possibly strained) lattice.
 - 3: Set $\mathbf{u} = -e^{-\gamma t}(\vec{M})^{-1}p$
 - 4: Calculate the property loss functional by computing the surrogate $\tilde{\mathcal{F}}_\theta(\mathbf{q})$ along with its gradient.
 - 5: Update the structure $\mathbf{q}$ and momentum $\mathbf{p}$ by applying a stochastic symplectic solver on the dynamics described by equation 75.
 - 6: **if** $t \leq T$ **then** Go to 2
-

3 Results

In this section, we present the results obtained from two baseline optimization approaches for discovering lattice configurations of $CsSnI_3$ that yield a desired fundamental bandgap of $1.5eV$. These results will later serve as a comparative benchmark for evaluating the performance of the proposed physics-inspired reinforcement learning-based algorithm. Throughout these methods, we have used an exact GPAW-based bandgap calculator for our optimization. The loss function that we minimize through our approach is given by:

$$\mathcal{F}(\mathbf{q}) = |\Delta_{fun}(\mathbf{q}) - \Delta^*|^2 \quad (99)$$

Where $\Delta_{fun}(\mathbf{q})$ is the fundamental bandgap of the lattice (equation 29), $\mathbf{q}$ is the vector representation of our structure, and Δ^* is the choice for our target bandgap (which is $1.5eV$).

For the bandgap calculations, we have used GPAW in python in the Plane-Wave mode. The exchange correlation potential we have chosen for these calculations is the GLLB-sc exchange correlation [43]. For our calculations we have taken the kinetic energy cutoff to be $500eV$ and a (8,8,8) Monkhorst-Pack of k-points shifted by $\frac{1}{16}$. The density mixing in the SCF step is done via Pulay mixing (section 2.1.1.4) with linear mixing parameter $\beta = 0.5$, the number of densities mixed $n_{max} = 5$, and the weight parameter $w = 50$. The gradient calculation in these approaches has been done via a 1st-order (one-sided) finite-differences scheme [94].

3.1 Case 1: Cubic Lattices

The first baseline algorithm only searches over various strained cubic configurations of $CsSnI_3$. For this, we have used an ADAM-based Stochastic Gradient Descent optimizer of our loss functional to optimize the bandgap of the material. The lattice structure in this optimization is represented by a single parameter, which is the lattice side length. The atomic positions in this optimization have been fixed - Cs is at the body center, Sn is at the vertices, and I are at the face centers. For this optimization, we chose a learning rate of 0.01. Performing this optimization, we obtained good convergence in around 50 iterations with the optimal lattice parameter of $6.481\text{\AA}$, which gives a

deviation from target of $8.062 \times 10^{-5} eV$. The convergence of the results with iterations has been plotted in figure 5.

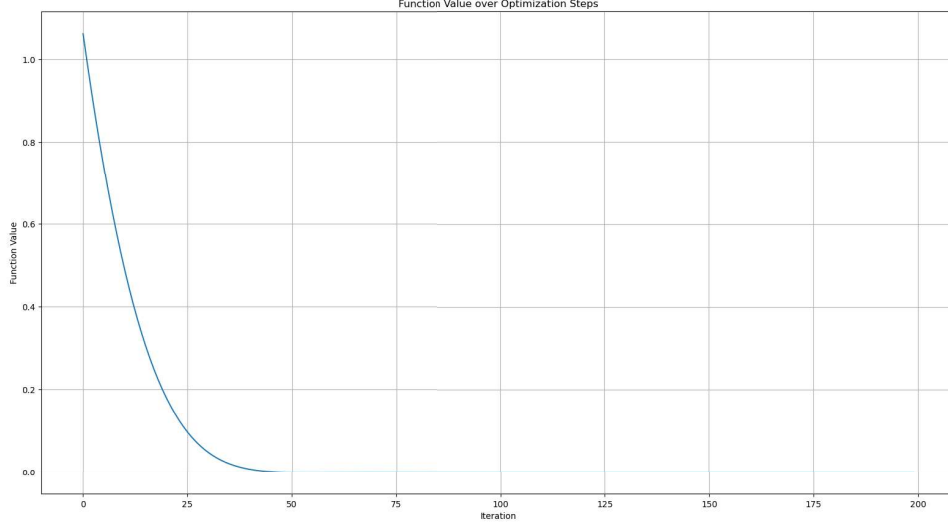


Figure 5: Evolution of $\mathcal{F}(\mathbf{q})$ for search over cubic lattices

3.2 Case 2: Cuboidal Lattices

An extension to the first baseline approach allows us to extend our search to cuboidal lattices (cubic/tetragonal/orthorhombic). In this discovery scheme, we first perform fixed cell relaxations by minimizing the Kohn-Sham energy via ADAM-based stochastic gradient descent. 4 iterations of this relaxation are done per step with a learning rate of 0.05. After relaxing the structure, the lattice parameters are updated using ADAM-based stochastic gradient descent to minimize the bandgap loss function with a learning rate of 0.05. 12 iterations of the whole pipeline (relaxation + update of lattice) are done and the results are outputted. Figure 6 plots the evolution of the Kohn-Sham energy throughout the relax+search dynamics while figure 7 plots the evolution of the loss function during the search. As it can be seen from the 2 figures, we obtain good convergence within 6 search iterations. The deviation from the target in this search is $8.808 \times 10^{-3} eV$. The structure of $CsSnI_3$ that gives this output is given by $a = 6.620 \text{\AA}$, $b = 6.284 \text{\AA}$, $c = 6.336 \text{\AA}$, $Cs \equiv (0.578, 0.502, 0.505)$ lattice units, $Pb \equiv (0.988, 0.989, 0.994)$ lattice units, $I_1 \equiv (0.423, 0.980, 0.013)$ lattice units, $I_2 \equiv$

$(0.937, 0.497, 0.003)$ lattice units, and $I_3 \equiv (0.944, 0.996, 0.520)$ lattice units (here lattice units are coordinates in units of lattice parameters).

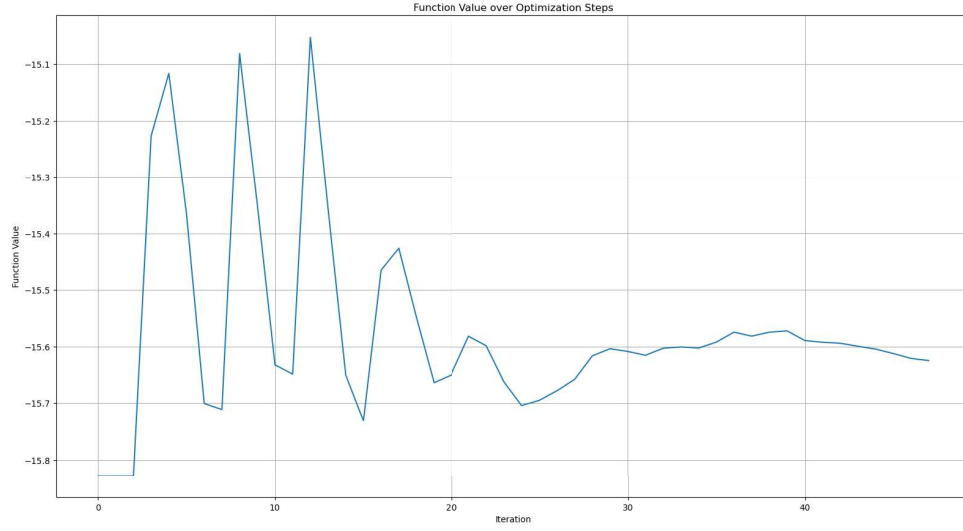


Figure 6: Evolution of Kohn-Sham Energy for search over cuboidal lattices

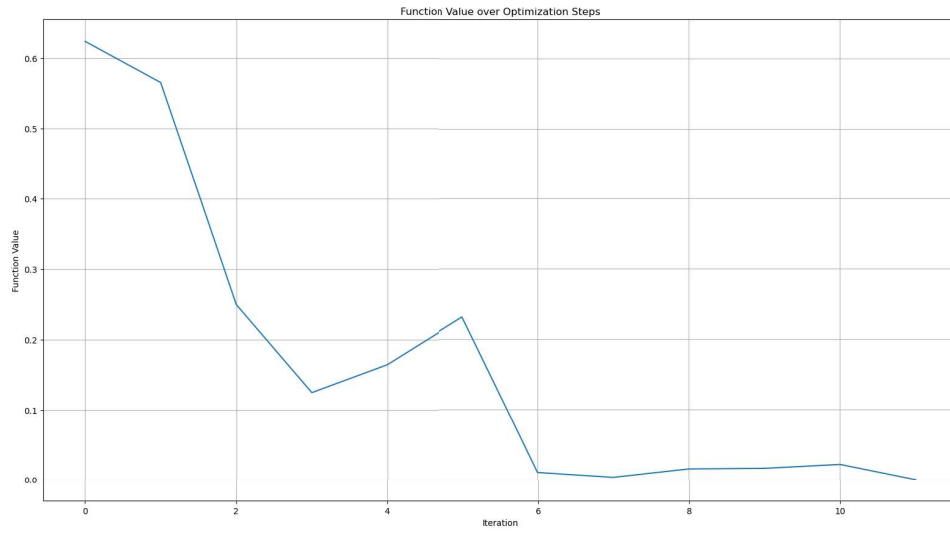


Figure 7: Evolution of $\mathcal{F}(\mathbf{q})$ for search over cuboidal lattices

4 Discussion

In this thesis, we provided a reinforcement learning way to optimally discover new materials and evaluated two traditional baseline algorithms for optimizing lattice configurations of $CsSnI_3$ to achieve a desired fundamental bandgap of $1.5eV$. These baseline results demonstrate significant challenges and limitations inherent in traditional optimization approaches, particularly when extending the search beyond low-dimensional spaces to high-dimensional search scenarios.

4.1 Challenges in High-Dimensional Discovery via Traditional Approaches

The baseline experiments clearly illustrate the computational challenges of performing high-dimensional search and optimization using traditional gradient-based methods like ADAM stochastic gradient descent (SGD). Although the first baseline, constrained solely to optimizing lattice parameters with fixed atomic positions, reached convergence within 50 steps, the approach was inherently limited by its restricted search space.

The second baseline, which extended the search to cuboidal lattice configurations and incorporated fixed-cell relaxations to optimize atomic positions, further highlights the computational limitations for high-dimensional discovery. Despite setting a relatively high learning rate (0.05), along with the search still being restricted to just search in the structural space, the algorithm demonstrated slow convergence and required extensive computational resources, again taking approximately two days to yield results.

These baseline experiments demonstrate fundamental limitations of traditional gradient-based optimization methods for high-dimensional materials search spaces - they are computationally expensive, relatively inflexible, and prone to getting trapped in local minima. Importantly, the baseline experiments have thus far only explored variations in the structural, without even touching upon the complexity introduced by chemical substitutions — an area that vastly expands the dimensionality and complexity of the search space, which would make the search through traditional methods even more unfeasible.

In contrast, the physics-inspired reinforcement learning-based Hamiltonian optimization framework we propose promises substantial improvements over these traditional methods. By integrating local surrogate modeling, our proposed method reduces the computational burden associated with expensive Kohn-Sham calculations, enabling faster property evaluations. This surrogate-based approach is crucial because traditional methods’ reliance on direct DFT computations severely limits their scalability, especially when exploring high-dimensional structural and chemical spaces.

4.2 Advantages of Our Physics-Inspired Reinforcement Learning Approach

To address the challenges imposed by traditional approaches, we develop a novel, physics-inspired reinforcement learning framework grounded in Stochastic Hamiltonian Dynamics. Our methodology not only addresses the computational challenges but also significantly expands the potential for comprehensive exploration in high-dimensional materials discovery. The benefits of our approach are explained below:

- **Local Surrogate Modeling for Computational Efficiency:** We incorporate local surrogate models into our optimization approach to substantially reduce computational demands by approximating expensive Kohn-Sham calculations. The use of surrogate models allow us to perform rapid property evaluations, thus accelerating the optimization process considerably. Additionally, by making these surrogate models inherently differentiable, we seamlessly integrated efficient gradient-based optimization strategies into our discovery process.
- **Chemical Space Exploration via Gumbel-Softmax Reparameterization:** Recognizing the complexity inherent in exploring chemical space, we implement the Gumbel-Softmax reparameterization trick. This approach converts discrete, categorical variables that represent the chemical composition into continuous, differentiable variables, thus enabling effective gradient-based optimization across chemical configurations. This novel integration allows us to thoroughly and efficiently explore a broader chemical landscape.

- Stochastic Hamiltonian Dynamics for Stable Convergence with Exploration:** Drawing inspiration from Hamiltonian mechanics, our optimization framework employed physically meaningful dynamics characterized by structured momentum and energy considerations. This physics-based approach ensured stable, consistent convergence behaviors, systematically guiding the optimization trajectory toward globally optimal solutions while minimizing inefficient random explorations. Moreover, by introducing adaptive stochastic terms, we enabled the algorithm to effectively navigate and avoid local minima, significantly improving our exploration capabilities. This stochastic framework provided a robust and systematic approach to global optimization.
- Continually Accelerated Discovery through Reinforcement Learning via Adaptive Learning of Optimization Parameters:** By embedding reinforcement learning principles into our discovery framework, we ensure continual acceleration in the optimization process. Each iteration informs subsequent cycles, with newly acquired data dynamically refining our surrogate models and discovery strategy by iteratively learning the dynamic parameters such as inertia tensors, damping constants, and stochasticities. This iterative feedback loop progressively enhances our model’s predictive capability and significantly increases the efficiency and effectiveness of our material discovery process.
- Flexibility, Scalability, and Extensibility:** The framework we introduce readily accommodates multiple objectives, additional constraints, and external data inputs, making it highly versatile for various materials discovery tasks. Its scalability ensures practicality and effectiveness as the dimensionality and complexity of the search spaces expands, bridging theoretical optimization with real-world applications.

In summary, by continual learning of the Optimaltonian during our physics-inspired reinforcement learning approach, we effectively overcome the limitations faced by traditional optimization methods. By combining surrogate modeling, adaptive and differentiable optimization, stochastic exploration, and reinforcement learning, we establish a robust, accelerating, and flexible platform for future materials discovery initiatives.

5 Conclusion and Future Outlook

In conclusion, this thesis presents a novel physics-inspired reinforcement learning framework that significantly advances the computational efficiency and effectiveness of materials discovery. By leveraging sophisticated techniques such as Stochastic Hamiltonian Dynamics integrated with local surrogate modeling, the proposed approach effectively addresses the challenges posed by high-dimensional search spaces, inherent computational complexity, and the limitations of traditional optimization methods.

The proposed methodology not only demonstrates enhanced stability and convergence properties but also exhibits significant flexibility and scalability, accommodating both structural and chemical degrees of freedom through differentiable optimization facilitated by the Gumbel-Softmax reparameterization trick. The computational advantages of surrogate models have been particularly highlighted, enabling rapid property evaluations and iterative improvement through reinforcement learning principles, leading to progressively accelerated discovery cycles.

Looking ahead, we can also incorporate entropy and free-energy principles into this discovery framework which would allow us to better balance between exploration and exploitation in our discovery framework [95], thus ensuring even more robust global optimization and efficient navigation of complex energy landscapes. Additionally, we could also add learning symmetries, symmetry reductions, and renormalizing group flows on the Hamiltonians in our methodology to accelerate the discovery even further [96]. Also, coupling these methods with advanced probabilistic modeling and Bayesian inference may enable more sophisticated uncertainty quantification and adaptive exploration strategies, further broadening the applicability and impact of the developed framework in materials science and beyond. In addition to this, instead of just targeting a single material property (like bandgap), we can opt to search for good materials using multi-objective functions (balancing multiple properties like stability, bandgap, excited state lifetimes, charge transfer, and electron-hole overlap) to find overall good materials. Additionally, our approach of using Hamiltonians for learning can also be extended to quantum algorithms by learning quantum Hamiltonians instead of classical Hamiltonians [97].

References

1. Sun, J. *et al.* Accelerating the discovery of acceptor materials for organic solar cells by deep learning. *npj Computational Materials* **10**. ISSN: 2057-3960. <http://dx.doi.org/10.1038/s41524-024-01367-7> (Aug. 2024).
2. Choubisa, H. *et al.* Interpretable discovery of semiconductors with machine learning. *npj Computational Materials* **9**. ISSN: 2057-3960. <http://dx.doi.org/10.1038/s41524-023-01066-9> (June 2023).
3. Mannodi-Kanakkithodi, A. A guide to discovering next-generation semiconductor materials using atomistic simulations and machine learning. *Computational Materials Science* **243**, 113108. ISSN: 0927-0256. <https://www.sciencedirect.com/science/article/pii/S092702562400329X> (2024).
4. Williams, T., McCullough, K. & Lauterbach, J. A. Enabling Catalyst Discovery through Machine Learning and High-Throughput Experimentation. *Chemistry of Materials* **32**, 157–165. eprint: <https://doi.org/10.1021/acs.chemmater.9b03043>. <https://doi.org/10.1021/acs.chemmater.9b03043> (2020).
5. Huang, G., Huang, F. & Dong, W. Machine learning in energy storage material discovery and performance prediction. *Chemical Engineering Journal* **492**, 152294. ISSN: 1385-8947. <https://www.sciencedirect.com/science/article/pii/S1385894724037811> (2024).
6. *Material Discovery: NREL* <https://www2.nrel.gov/materials-science/materials-discovery>. Accessed: 2025-03-11.
7. Lookman, T., Balachandran, P. V., Xue, D. & Yuan, R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Computational Materials* **5**, 1–17. <https://api.semanticscholar.org/CorpusID:186802303> (2019).
8. Merchant, A. *et al.* Scaling deep learning for materials discovery. *Nature* **624**, 1–6 (Nov. 2023).
9. Vasudevan, R., Pilania, G. & Balachandran, P. V. Machine learning for materials design and discovery. *Journal of Applied Physics* **129**, 070401. ISSN: 0021-8979. eprint: https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/5.0043300/13928931/070401_1_online.pdf. <https://doi.org/10.1063/5.0043300> (Feb. 2021).
10. Cheetham, A. K. & Seshadri, R. Artificial Intelligence Driving Materials Discovery? Perspective on the Article: Scaling Deep Learning for Materials Discovery. *Chemistry of Materials* **36**, 3490–3495. eprint: <https://doi.org/10.1021/acs.chemmater.4c00643>. <https://doi.org/10.1021/acs.chemmater.4c00643> (2024).
11. Liu, Y., Zhao, T., Ju, W. & Shi, S. Materials discovery and design using machine learning. *Journal of Materiomics* **3**. High-throughput Experimental and Modeling Research toward Advanced Batteries, 159–177. ISSN: 2352-8478. <https://www.sciencedirect.com/science/article/pii/S2352847817300515> (2017).

12. Kim, H., Choi, H., Kang, D., Lee, W. B. & Na, J. Materials discovery with extreme properties via reinforcement learning-guided combinatorial chemistry. *Chem. Sci.* **15**, 7908–7925. <http://dx.doi.org/10.1039/D3SC05281H> (21 2024).
13. Manna, S. *et al.* Learning in continuous action space for developing high dimensional potential energy models. *Nature Communications* **13**. ISSN: 2041-1723. <http://dx.doi.org/10.1038/s41467-021-27849-6> (Jan. 2022).
14. Ramesh, A. & Ravindran, B. *Physics-Informed Model-Based Reinforcement Learning* 2023. arXiv: [2212.02179](https://arxiv.org/abs/2212.02179) [cs.LG]. <https://arxiv.org/abs/2212.02179>.
15. Bajaj, C. & Nguyen, M. *DPO: Differential reinforcement learning with application to optimal configuration search* 2024. arXiv: [2404.15617](https://arxiv.org/abs/2404.15617) [cs.LG]. <https://arxiv.org/abs/2404.15617>.
16. Zhang, Y., Apley, D. W. & Chen, W. Bayesian Optimization for Materials Design with Mixed Quantitative and Qualitative Variables. *Scientific Reports* **10**. ISSN: 2045-2322. <http://dx.doi.org/10.1038/s41598-020-60652-9> (Mar. 2020).
17. Jang, E., Gu, S. & Poole, B. *Categorical Reparameterization with Gumbel-Softmax* 2017. arXiv: [1611.01144](https://arxiv.org/abs/1611.01144) [stat.ML]. <https://arxiv.org/abs/1611.01144>.
18. Kohn, E. S. Effect of the Bandgap on the Elasto-Optic and Electro-Optic Properties of Hexagonal Semiconductors. *Journal of Applied Physics* **40**, 2608–2613. ISSN: 0021-8979. eprint: https://pubs.aip.org/aip/jap/article-pdf/40/6/2608/18350449/2608_1_online.pdf. <https://doi.org/10.1063/1.1658041> (May 1969).
19. Sahni, V. in *Quantal Density Functional Theory* 99–123 (Springer Berlin Heidelberg, 2004). ISBN: 9783662096246. http://dx.doi.org/10.1007/978-3-662-09624-6_4.
20. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **136**, B864–B871. <https://link.aps.org/doi/10.1103/PhysRev.136.B864> (3B 1964).
21. Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, A1133–A1138. <https://link.aps.org/doi/10.1103/PhysRev.140.A1133> (4A 1965).
22. Fujita, S. & Ito, K. in *Quantum Theory of Conducting Matter* 85–95 (Springer New York, 2007). ISBN: 9780387741031. http://dx.doi.org/10.1007/978-0-387-74103-1_7.
23. Hart, G. L. W., Jorgensen, J. J., Morgan, W. S. & Forcade, R. W. A robust algorithm for k-point grid generation and symmetry reduction. *Journal of Physics Communications* **3**, 065009. ISSN: 2399-6528. <http://dx.doi.org/10.1088/2399-6528/ab2937> (June 2019).
24. Miguel A. L. Marques, E. K. U. G. *Time-Dependent Density Functional Theory* https://www-old.mpi-halle.mpg.de/mpi/publi/br/gross_persoendlich/MG03.pdf. Accessed: 2025-02-10.
25. Tolba, S. A., Gameel, K. M., Ali, B. A., Almossalami, H. A. & Allam, N. K. in *Density Functional Calculations* (ed Yang, G.) chap. 1 (IntechOpen, Rijeka, 2018). <https://doi.org/10.5772/intechopen.72020>.

26. *The Atoms Object* <https://wiki.fysik.dtu.dk/ase/ase/atoms.html>. Accessed: 2025-02-10.
27. Janssen, J., Makarov, E., Hickel, T., Shapeev, A. V. & Neugebauer, J. Automated optimization and uncertainty quantification of convergence parameters in plane wave density functional theory calculations. *npj Computational Materials* **10**. ISSN: 2057-3960. <http://dx.doi.org/10.1038/s41524-024-01388-2> (Nov. 2024).
28. Basiuk, V. A. Electron smearing in DFT calculations: A case study of doxorubicin interaction with single-walled carbon nanotubes. *International Journal of Quantum Chemistry* **111**, 4197–4205. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.23003>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.23003> (2011).
29. Kanungo, B., Zimmerman, P. M. & Gavini, V. Exact exchange-correlation potentials from ground-state electron densities. *Nature Communications* **10**. ISSN: 2041-1723. <http://dx.doi.org/10.1038/s41467-019-12467-0> (Oct. 2019).
30. Hodgson, M. J. P., Wetherell, J. & Fromager, E. Exact exchange-correlation potentials for calculating the fundamental gap with a fixed number of electrons. *Phys. Rev. A* **103**, 012806. <https://link.aps.org/doi/10.1103/PhysRevA.103.012806> (1 2021).
31. Fritsch, D. & Schorr, S. Climbing Jacob’s ladder: A density functional theory case study for Ag₂ZnSnSe₄ and Cu₂ZnSnSe₄. *Journal of Physics: Energy* **3**, 015002. <https://dx.doi.org/10.1088/2515-7655/abc07b> (2020).
32. Woods, N. *On the Nature of Self-Consistency in Density Functional Theory* 2018. arXiv: 1803.01763 [cond-mat.other]. <https://arxiv.org/abs/1803.01763>.
33. Liu, X., Wang, X., Wen, Z. & Yuan, Y. On the Convergence of the Self-Consistent Field Iteration in Kohn–Sham Density Functional Theory. *SIAM Journal on Matrix Analysis and Applications* **35**, 546–558. eprint: <https://doi.org/10.1137/130911032>. <https://doi.org/10.1137/130911032> (2014).
34. Lin, L., Lu, J. & Ying, L. Numerical methods for Kohn–Sham density functional theory. *Acta Numerica* **28**, 405–539 (2019).
35. Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters* **73**, 393–398. ISSN: 0009-2614. <https://www.sciencedirect.com/science/article/pii/0009261480803964> (1980).
36. Mortensen, J. J. *et al.* GPAW: An open Python package for electronic structure calculations. *The Journal of Chemical Physics* **160**, 092503. ISSN: 0021-9606. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0182685/19717263/092503_1_5.0182685.pdf. <https://doi.org/10.1063/5.0182685> (Mar. 2024).
37. Banerjee, A. S., Suryanarayana, P. & Pask, J. E. Periodic Pulay method for robust and efficient convergence acceleration of self-consistent field iterations. *Chemical Physics Letters* **647**, 31–35. ISSN: 0009-2614. <https://www.sciencedirect.com/science/article/pii/S0009261416000464> (2016).

38. Castelli, I. E., García-Lastra, J. M., Thygesen, K. S. & Jacobsen, K. W. Bandgap calculations and trends of organometal halide perovskites. *APL Materials* **2**, 081514. ISSN: 2166-532X. eprint: https://pubs.aip.org/aip/apm/article-pdf/doi/10.1063/1.4893495/13863163/081514_1_online.pdf. <https://doi.org/10.1063/1.4893495> (Aug. 2014).
39. Yang, W., Cohen, A. J. & Mori-Sánchez, P. Derivative discontinuity, bandgap and lowest unoccupied molecular orbital in density functional theory. *The Journal of Chemical Physics* **136**. ISSN: 1089-7690. <http://dx.doi.org/10.1063/1.3702391> (May 2012).
40. Blase, X., Duchemin, I., Jacquemin, D. & Loos, P.-F. The Bethe–Salpeter Equation Formalism: From Physics to Chemistry. *The Journal of Physical Chemistry Letters* **11**. PMID: 32787315, 7371–7382. eprint: <https://doi.org/10.1021/acs.jpcclett.0c01875>. <https://doi.org/10.1021/acs.jpcclett.0c01875> (2020).
41. Borlido, P. *et al.* Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning. *npj Computational Materials* **6**. ISSN: 2057-3960. <http://dx.doi.org/10.1038/s41524-020-00360-0> (July 2020).
42. Tran, F. *et al.* Semilocal exchange-correlation potentials for solid-state calculations: Current status and future directions. *Journal of Applied Physics* **126**, 110902. ISSN: 0021-8979. eprint: https://pubs.aip.org/aip/jap/article-pdf/doi/10.1063/1.5118863/19994257/110902_1_1.5118863.pdf. <https://doi.org/10.1063/1.5118863> (Sept. 2019).
43. Kuisma, M., Ojanen, J., Enkovaara, J. & Rantala, T. T. Kohn-Sham potential with discontinuity for band gap materials. *Phys. Rev. B* **82**, 115106. <https://link.aps.org/doi/10.1103/PhysRevB.82.115106> (11 2010).
44. Verma, P. & Truhlar, D. G. HLE17: An Improved Local Exchange–Correlation Functional for Computing Semiconductor Band Gaps and Molecular Excitation Energies. *The Journal of Physical Chemistry C* **121**, 7144–7154. eprint: <https://doi.org/10.1021/acs.jpcc.7b01066>. <https://doi.org/10.1021/acs.jpcc.7b01066> (2017).
45. Adhikari, S. *et al.* Accurate Prediction of HSE06 Band Structures for a Diverse Set of Materials Using -Learning. *Chemistry of Materials* **35**, 8397–8405. eprint: <https://doi.org/10.1021/acs.chemmater.3c01131>. <https://doi.org/10.1021/acs.chemmater.3c01131> (2023).
46. Lebègue, S., Klintonberg, M., Eriksson, O. & Katsnelson, M. I. Accurate electronic band gap of pure and functionalized graphene from GW calculations. *Phys. Rev. B* **79**, 245117. <https://link.aps.org/doi/10.1103/PhysRevB.79.245117> (24 2009).
47. Cardona, M. & Pollak, F. H. Energy-Band Structure of Germanium and Silicon: The k-p Method. *Phys. Rev.* **142**, 530–543. <https://link.aps.org/doi/10.1103/PhysRev.142.530> (2 1966).
48. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979. <https://link.aps.org/doi/10.1103/PhysRevB.50.17953> (24 1994).

49. Marsman, M. & Kresse, G. Relaxed core projector-augmented-wave method. *The Journal of Chemical Physics* **125**, 104101. ISSN: 0021-9606. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2338035/15387791/104101\1_online.pdf. <https://doi.org/10.1063/1.2338035> (Sept. 2006).
50. Grubmüller, H., Heller, H., Windemuth, A. & Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Mol. Sim.* **6**, 121– (Mar. 1991).
51. Verstraete, M. & Zanolli, Z. in (Mar. 2014). ISBN: 978-3-89336-912-6. https://zeilazanolli.wordpress.com/wp-content/uploads/2019/04/2013_dfpt_c2_verstraete_zanolli.pdf.
52. Shang, H., Carbogno, C., Rinke, P. & Scheffler, M. Lattice dynamics calculations based on density-functional perturbation theory in real space. *Computer Physics Communications* **215**, 26–46. ISSN: 0010-4655. <https://www.sciencedirect.com/science/article/pii/S0010465517300437> (2017).
53. Zeller, J. *Evaluation of the Self-Consistent Sternheimer Method in a Localized Basis* Available at <https://research.tue.nl/en/studentTheses/evaluation-of-the-self-consistent-sternheimer-method-in-a-localiz>. MA thesis (Eindhoven University of Technology, 2020).
54. Wu, Z. *et al.* OpenCL-accelerated first-principles calculations of all-electron quantum perturbations on HPC resources. *Frontiers in Chemistry* **11**, 1156891 (May 2023).
55. Friesecke, G. & Kniely, M. New Optimal Control Problems in Density Functional Theory Motivated by Photovoltaics. *Multiscale Modeling & Simulation* **17**, 926–947. eprint: <https://doi.org/10.1137/18M1207272>. <https://doi.org/10.1137/18M1207272> (2019).
56. Shockley, W. & Queisser, H. J. Detailed Balance Limit of Efficiency of p-n Junction Solar Cells. *Journal of Applied Physics* **32**, 510–519. ISSN: 0021-8979. eprint: https://pubs.aip.org/aip/jap/article-pdf/32/3/510/18323491/510\1_online.pdf. <https://doi.org/10.1063/1.1736034> (Mar. 1961).
57. Kim, S. & Walsh, A. Ab initio calculation of the detailed balance limit to the photovoltaic efficiency of single p-n junction kesterite solar cells. *Applied Physics Letters* **118**, 243905. ISSN: 0003-6951. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/5.0049143/14551731/243905\1_online.pdf. <https://doi.org/10.1063/5.0049143> (June 2021).
58. Vos, A. D. Detailed balance limit of the efficiency of tandem solar cells. *Journal of Physics D: Applied Physics* **13**, 839. <https://dx.doi.org/10.1088/0022-3727/13/5/018> (1980).
59. Chantana, J., Takeguchi, K., Mavlonov, A., Kawano, Y. & Minemoto, T. Correlation between detailed balance limit and actual environmental factors for perovskite/crystalline Si tandem solar cells with different structures. *Materials Science in Semiconductor Processing* **152**, 107085. ISSN: 1369-8001. <https://www.sciencedirect.com/science/article/pii/S1369800122006114> (2022).

60. Miah, M. H., Khandaker, M. U., Rahman, M. B., Nur-E-Alam, M. & Islam, M. A. Band gap tuning of perovskite solar cells for enhancing the efficiency and stability: issues and prospects. *RSC Advances* **14**, 15876–15906. ISSN: 2046-2069. <http://dx.doi.org/10.1039/D4RA01640H> (2024).
61. Kim, H.-S. & Park, N.-G. Importance of tailoring lattice strain in halide perovskite crystals. *NPG Asia Materials* **12**. ISSN: 1884-4057. <http://dx.doi.org/10.1038/s41427-020-00265-w> (Dec. 2020).
62. Baumann, F. *et al.* Strain in Halide Perovskites and Solar Cell Stability: Accelerated Stress Tests under Bias Voltage. *ACS Energy Letters* **10**, 476–483. eprint: <https://doi.org/10.1021/acsenenergylett.4c02822>. <https://doi.org/10.1021/acsenenergylett.4c02822> (2025).
63. Wang, Y. *et al.* Octahedral units in halide perovskites. *Nature Reviews Chemistry*. ISSN: 2397-3358. <http://dx.doi.org/10.1038/s41570-025-00687-6> (Feb. 2025).
64. Reynaud, M. *et al.* Si-Integrated BaTiO₃ for Electro-Optic Applications: Crystalline and Polarization Orientation Control. *ACS Applied Electronic Materials* **5**, 4605–4614. eprint: <https://doi.org/10.1021/acsaelm.3c00747>. <https://doi.org/10.1021/acsaelm.3c00747> (2023).
65. Meng, W. *et al.* Revealing the strain-associated physical mechanisms impacting the performance and stability of perovskite solar cells. *Joule* **6**, 458–475. ISSN: 2542-4351. <https://www.sciencedirect.com/science/article/pii/S2542435122000435> (2022).
66. Zhang, J. *et al.* The Role of Grain Boundaries in Organic–Inorganic Hybrid Perovskite Solar Cells and its Current Enhancement Strategies: A Review. *ENERGY & ENVIRONMENTAL MATERIALS* **7**. e12696 EEM-2023-0437.R1, e12696. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/eem2.12696>. <https://onlinelibrary.wiley.com/doi/abs/10.1002/eem2.12696> (2024).
67. Liu, Y., Stradins, P., Deng, H., Luo, J. & Wei, S.-H. Suppress carrier recombination by introducing defects: The case of Si solar cell. *Applied Physics Letters* **108**, 022101. ISSN: 0003-6951. eprint: https://pubs.aip.org/aip/apl/article-pdf/doi/10.1063/1.4939628/14474901/022101_1_online.pdf. <https://doi.org/10.1063/1.4939628> (Jan. 2016).
68. Choon, S. L. & Lim, H. N. Impact of vacancies in halide perovskites for batteries and supercapacitors. *Materials Today Energy* **43**, 101577. ISSN: 2468-6069. <https://www.sciencedirect.com/science/article/pii/S2468606924000893> (2024).
69. Kasai, H. & Matsumura, H. Study for improvement of solar cell efficiency by impurity photovoltaic effect. *Solar Energy Materials and Solar Cells* **48**, 93–100. ISSN: 0927-0248. <https://www.sciencedirect.com/science/article/pii/S0927024897000755> (1997).
70. Ruder, S. *An overview of gradient descent optimization algorithms* 2017. arXiv: [1609.04747](https://arxiv.org/abs/1609.04747) [cs.LG]. <https://arxiv.org/abs/1609.04747>.
71. Danilova, M. *et al.* *Recent Theoretical Advances in Non-Convex Optimization* 2021. arXiv: [2012.06188](https://arxiv.org/abs/2012.06188) [math.OC]. <https://arxiv.org/abs/2012.06188>.

72. Kingma, D. P. & Ba, J. *Adam: A Method for Stochastic Optimization* 2017. arXiv: [1412.6980](https://arxiv.org/abs/1412.6980) [cs.LG]. <https://arxiv.org/abs/1412.6980>.
73. Malkiel, I. & Wolf, L. *MTAdam: Automatic Balancing of Multiple Training Loss Terms* 2020. arXiv: [2006.14683](https://arxiv.org/abs/2006.14683) [cs.LG]. <https://arxiv.org/abs/2006.14683>.
74. Gidas, B. in *Disordered Systems and Biological Organization* 321–326 (Springer Berlin Heidelberg, 1986). ISBN: 9783642826573. http://dx.doi.org/10.1007/978-3-642-82657-3_31.
75. Borysenko, O. & Byshkin, M. CoolMomentum: a method for stochastic optimization by Langevin dynamics with simulated annealing. *Scientific Reports* **11**. ISSN: 2045-2322. <http://dx.doi.org/10.1038/s41598-021-90144-3> (May 2021).
76. Aarts, E., Korst, J. & Michiels, W. in *Search Methodologies* 187–210 (Springer US). ISBN: 9780387234601. http://dx.doi.org/10.1007/0-387-28356-0_7.
77. Shao, W. & Guo, G. Multiple-Try Simulated Annealing Algorithm for Global Optimization. *Mathematical Problems in Engineering* **2018**, 1–11. ISSN: 1563-5147. <http://dx.doi.org/10.1155/2018/9248318> (July 2018).
78. Sastry, K., Goldberg, D. & Kendall, G. in *Search Methodologies* 97–125 (Springer US). ISBN: 9780387234601. http://dx.doi.org/10.1007/0-387-28356-0_4.
79. Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **52**, 2893–2906. <https://link.aps.org/doi/10.1103/PhysRevE.52.2893> (3 1995).
80. Goedecker, S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics* **120**, 9911–9917. ISSN: 0021-9606. eprint: <https://pubs.aip.org/aip/jcp/article-pdf/120/21/9911/19031818/9911\1\online.pdf>. <https://doi.org/10.1063/1.1724816> (June 2004).
81. Boscain, U., Sigalotti, M. & Sugny, D. Introduction to the Pontryagin Maximum Principle for Quantum Optimal Control. *PRX Quantum* **2**, 030203. <https://link.aps.org/doi/10.1103/PRXQuantum.2.030203> (3 2021).
82. Vieillard, N., Scherrer, B., Pietquin, O. & Geist, M. *Momentum in Reinforcement Learning* 2020. arXiv: [1910.09322](https://arxiv.org/abs/1910.09322) [cs.LG]. <https://arxiv.org/abs/1910.09322>.
83. Bettiol, P. & Vinter, R. B. Sensitivity Interpretations of the Costate Variable for Optimal Control Problems with State Constraints. *SIAM Journal on Control and Optimization* **48**, 3297–3317. eprint: <https://doi.org/10.1137/080732614>. <https://doi.org/10.1137/080732614> (2010).
84. Pata, V. in *Fixed Point Theorems and Applications* 75–80 (Springer International Publishing, Cham, 2019). ISBN: 978-3-030-19670-7. https://doi.org/10.1007/978-3-030-19670-7_14.
85. Peng, S. A General Stochastic Maximum Principle for Optimal Control Problems. *SIAM Journal on Control and Optimization* **28**, 966–979. eprint: <https://doi.org/10.1137/0328054>. <https://doi.org/10.1137/0328054> (1990).

86. Annunziato, M., Borzì, A., Nobile, F. & Tempone, R. On the Connection between the Hamilton-Jacobi-Bellman and the Fokker-Planck Control Frameworks. *Applied Mathematics* **05**, 2476–2484 (Jan. 2014).
87. Falcone, M. in *Encyclopedia of Systems and Control* (eds Baillieul, J. & Samad, T.) 1585–1590 (Springer International Publishing, Cham, 2021). ISBN: 978-3-030-44184-5. https://doi.org/10.1007/978-3-030-44184-5_209.
88. Jordan, R., Kinderlehrer, D. & Otto, F. The Variational Formulation of the Fokker–Planck Equation. *SIAM Journal on Mathematical Analysis* **29**, 1–17. eprint: <https://doi.org/10.1137/S0036141096303359>. <https://doi.org/10.1137/S0036141096303359> (1998).
89. Hou, B., Wu, J. & Qiu, D. Y. Unsupervised representation learning of Kohn–Sham states and consequences for downstream predictions of many-body effects. *Nature Communications* **15**. ISSN: 2041-1723. <http://dx.doi.org/10.1038/s41467-024-53748-7> (Nov. 2024).
90. Samuels, J. *One-Hot Encoding and Two-Hot Encoding: An Introduction* Jan. 2024. <https://rgdoi.net/10.13140/RG.2.2.21459.76327>.
91. *Organometal Halide Perovskites* <https://cmr.fysik.dtu.dk/organometal/organometal.html>. Accessed: 2025-03-10.
92. Hu, Y., Yin, J., Gao, X., Liu, X. & Song, H. Projected gradient descent algorithm for ab initio crystal structure relaxation under a fixed unit cell volume. *Physical Review B* **109**. ISSN: 2469-9969. <http://dx.doi.org/10.1103/PhysRevB.109.224108> (June 2024).
93. Cheng, X., Wang, L. & Cao, Y. Quadrature Based Neural Network Learning of Stochastic Hamiltonian Systems. *Mathematics* **12**. ISSN: 2227-7390. <https://www.mdpi.com/2227-7390/12/16/2438> (2024).
94. Sengupta, B., Friston, K. & Penny, W. Efficient gradient computation for dynamical models. *NeuroImage* **98**, 521–527. ISSN: 1053-8119. <https://www.sciencedirect.com/science/article/pii/S1053811914003097> (2014).
95. Rahme, J. & Adams, R. P. *A Theoretical Connection Between Statistical Physics and Reinforcement Learning* 2021. arXiv: [1906.10228](https://arxiv.org/abs/1906.10228) [cs.LG]. <https://arxiv.org/abs/1906.10228>.
96. Bachtis, D., Aarts, G. & Lucini, B. Adding machine learning within Hamiltonians: Renormalization group transformations, symmetry breaking and restoration. *Phys. Rev. Res.* **3**, 013134. <https://link.aps.org/doi/10.1103/PhysRevResearch.3.013134> (1 2021).
97. Kruse, G., Coehlo, R., Rosskopf, A., Wille, R. & Lorenz, J. M. *Hamiltonian-based Quantum Reinforcement Learning for Neural Combinatorial Optimization* 2024. arXiv: [2405.07790](https://arxiv.org/abs/2405.07790) [quant-ph]. <https://arxiv.org/abs/2405.07790>.

A Derivation of Deterministic PMP Through Lagrange Multipliers

Consider that we wish to extremize the functional given by equation 47. We consider the dynamics described by equation 48 as the constraint and put that in with a Lagrange multiplier. Then, we define an unconstrained action, given by:

$$\begin{aligned} S &= \lambda J + \int_0^T \langle \mathbf{p}(t), (\mathbf{f}(\mathbf{q}, \mathbf{u}, t) - \dot{\mathbf{q}}) \rangle dt \\ &= \int_0^T (\lambda L(\mathbf{q}, \mathbf{u}, t) + \langle \mathbf{p}(t), (\mathbf{f}(\mathbf{q}, \mathbf{u}, t) - \dot{\mathbf{q}}) \rangle) dt + \lambda h(T, \mathbf{q}(T)) \end{aligned} \quad (100)$$

Where, $\mathbf{p}$ is the Lagrange multiplier for the dynamics constraint, and λ is a non-zero constant which is added because extremizing J is same as extremizing λJ . Simplifying this, we get:

$$\begin{aligned} S &= \int_0^T (\lambda L(\mathbf{q}, \mathbf{u}, t) + \langle \mathbf{p}, \mathbf{f} \rangle + \langle \dot{\mathbf{p}}, \mathbf{q} \rangle) dt + \lambda h(T, \mathbf{q}(T)) - \langle \mathbf{p}(T), \mathbf{q}(T) \rangle + \langle \mathbf{p}(0), \mathbf{q}(0) \rangle \\ &= \int_0^T (\mathcal{H} + \langle \dot{\mathbf{p}}, \mathbf{q} \rangle) dt + \lambda h(T, \mathbf{q}(T)) - \langle \mathbf{p}(T), \mathbf{q}(T) \rangle + \langle \mathbf{p}(0), \mathbf{q}(0) \rangle \end{aligned} \quad (101)$$

Now, perturbing the control $\mathbf{u} \rightarrow \mathbf{u} + \delta \mathbf{u}$, we get:

$$\delta S = \int_0^T \nabla_{\mathbf{u}} \mathcal{H} \cdot (\delta \mathbf{u}) dt \quad (102)$$

For this to be optimal, δS should 0 for arbitrary $\delta \mathbf{u}$, which gives us the optimal control condition as:

$$\nabla_{\mathbf{u}} \mathcal{H} = 0 \quad (103)$$

Now, perturbing the control $\mathbf{q} \rightarrow \mathbf{q} + \delta \mathbf{q}$, we get:

$$\delta S = \int_0^T (\nabla_{\mathbf{q}} \mathcal{H} + \dot{\mathbf{p}}) \cdot (\delta \mathbf{q}) dt + (\lambda \nabla_{\mathbf{q}} h - \mathbf{p}(T)) \cdot \delta \mathbf{q}(T) - \mathbf{p}(0) \cdot \delta \mathbf{q}(0) \quad (104)$$

As the initial condition for $\mathbf{q}$ is already predetermined, we have $\delta \mathbf{q}(0) = 0$. Also, for the optimal trajectory, we again need $\delta S = 0$ for arbitrary $\delta \mathbf{q}$. Substituting these, we get:

$$\dot{\mathbf{p}} = -\nabla_{\mathbf{q}} \mathcal{H} \quad (105)$$

$$\lambda \nabla_{\mathbf{q}} h = \mathbf{p}(T) \quad (106)$$

Moreover, perturbing the control $\mathbf{p} \rightarrow \mathbf{p} + \delta \mathbf{p}$, we get:

$$\delta S = \int_0^T (\nabla_{\mathbf{p}} \mathcal{H} - \dot{\mathbf{q}}) \cdot (\delta \mathbf{p}) dt \quad (107)$$

As this should be 0 for arbitrary $\delta \mathbf{p}$, we have:

$$\dot{\mathbf{q}} = \nabla_p \mathcal{H} \quad (108)$$

Finally, the dynamics of $\mathcal{H}$ are given as:

$$\frac{d\mathcal{H}}{dt} = \nabla_q \mathcal{H} \cdot \dot{\mathbf{q}} + \nabla_p \mathcal{H} \cdot \dot{\mathbf{p}} + \frac{\partial \mathcal{H}}{\partial t} = \nabla_q \mathcal{H} \cdot \nabla_p \mathcal{H} - \nabla_p \mathcal{H} \cdot \nabla_q \mathcal{H} + \frac{\partial \mathcal{H}}{\partial t} = \frac{\partial \mathcal{H}}{\partial t} \quad (109)$$

B Proof of Uniqueness of Optimal Control in Stochastic Black Box Optimization

Consider that $\exists$ another optimal control $\mathbf{v}^*$. We can then write $\mathbf{v}^*$ as:

$$\mathbf{v}^* = \mathbf{u}^* + \mathbf{u}_1 = -e^{-\gamma t} \overset{\leftrightarrow}{M}^{-1} \mathbf{p} + \mathbf{u}_1 \quad (110)$$

Where $\mathbf{u}_1$ is a non-zero vector. By equation 69, we should have:

$$\langle e^{\gamma t} \overset{\leftrightarrow}{M} \mathbf{v}^* + \mathbf{p}, (\mathbf{u} - \mathbf{v}^*) \rangle \geq 0 \quad \forall \mathbf{u} \quad (111)$$

Simplifying this, we get:

$$\langle e^{\gamma t} \overset{\leftrightarrow}{M} \mathbf{u}_1, (\mathbf{u} - \mathbf{v}^*) \rangle \geq 0 \quad \forall \mathbf{u} \quad (112)$$

As this should be true $\forall \mathbf{u}$, it should also be true for $\mathbf{u} = \mathbf{u}^*$. Thus, we get:

$$-e^{\gamma t} \langle \overset{\leftrightarrow}{M} \mathbf{u}_1, \mathbf{u}_1 \rangle \geq 0 \quad (113)$$

As $\overset{\leftrightarrow}{M}$ is symmetric positive definite, this inequality is only satisfied as an equality, which only happens at:

$$\mathbf{u}_1 = 0 \quad (114)$$

As we had assumed that $\mathbf{u}_1$ is a non-zero vector, we get a contradiction. Thus, the optimal control for this case is unique.

C Defining the PAW Transformation

The electronic states for real materials have different properties in different regions of space: they are highly oscillating near the ionic nuclei, but are fairly smooth far away from the nuclei. Based on this property of the electronic states, we can select augmentation regions around the nuclei where the electronic states are highly oscillatory and need to be smoothened, and the interstitial region, where the transformation keeps the Kohn-Sham states invariant. Usually, these augmentation regions are taken to be spherical and the C^1 smoothness constraint is applied to the boundaries of the augmentation spheres. The PAW transformation is then defined as:

$$\hat{\mathcal{T}} = I + \sum_a \hat{\mathcal{T}}_a \quad (115)$$

Here, a is the index of the nuclei. Consider Ω_a to be the augmentation region containing the atom a , then $\hat{\mathcal{T}}_a$ satisfies:

$$\hat{\mathcal{T}}_a \psi(\mathbf{r}) = 0 \quad \forall \mathbf{r} \notin \Omega_a \quad (116)$$

This makes the transformation local, and each $\hat{\mathcal{T}}_a$ can be defined based on what nucleus is in the augmentation region. In the interstitial region, we can just normally expand the electronic states into any suitable basis (plane wave or Wannier) while in the interstitial region, we use expansions into a partial-wave basis (consider $|\phi_i\rangle$ to be the partial wave basis and $|\tilde{\phi}_i\rangle$ to be their smoothened counterparts). A natural choice for the partial wave basis is the radial Hydrogen-like basis for the isolated atom in the augmentation region, and can be orthogonalized to the core states if necessary. The smoothened partial wave basis is chosen such that the smoothening map preserves the desired smoothness (which is C^1 in most cases).

Consider the augmentation sphere Ω_a . Inside Ω_a , we can expand any state $|\psi_i\rangle$ as:

$$|\psi_i(\mathbf{r})\rangle = \sum_j c_{ij}^{(a)} |\phi_j^{(a)}(\mathbf{r})\rangle \quad \forall \mathbf{r} \in \Omega_a \quad (117)$$

Also, any smoothened state can be written as:

$$|\tilde{\psi}_i(\mathbf{r})\rangle = \sum_j c_{ij}^{(a)} |\tilde{\phi}_j^{(a)}(\mathbf{r})\rangle \quad \forall \mathbf{r} \in \Omega_a \quad (118)$$

Both expansions have identical coefficients. Adding equations 117 and 118, and rearranging the terms, we get:

$$|\psi_i(\mathbf{r})\rangle = |\tilde{\psi}_i(\mathbf{r})\rangle + \hat{\mathcal{T}}_a |\tilde{\psi}_i(\mathbf{r})\rangle = |\tilde{\psi}_i(\mathbf{r})\rangle - \sum_j c_{ij}^{(a)} |\tilde{\phi}_j^{(a)}(\mathbf{r})\rangle + \sum_j c_{ij}^{(a)} |\phi_j^{(a)}(\mathbf{r})\rangle \quad \forall \mathbf{r} \in \Omega_a \quad (119)$$

This means, we get:

$$\hat{\mathcal{T}}_a |\tilde{\psi}_i(\mathbf{r})\rangle = \sum_j c_{ij}^{(a)} (|\phi_j^{(a)}(\mathbf{r})\rangle - |\tilde{\phi}_j^{(a)}(\mathbf{r})\rangle) \quad \forall \mathbf{r} \in \Omega_a \quad (120)$$

Moreover, as $\hat{\mathcal{T}}_a$ are always linear transformations, we can write $c_{ij}^{(a)}$ as:

$$c_{ij}^{(a)} = \langle p_j^{(a)} | \tilde{\psi}_i \rangle \quad (121)$$

The states $|p_j^{(a)}\rangle$ are the projector states and are the quantity that actually define the PAW transformation. These projector states depend on the atom type and are stored in the PAW datasets. Also, in order to satisfy equation 118, these projector states must satisfy:

$$\sum_j |\tilde{\phi}_j^{(a)}\rangle \langle p_j^{(a)}| = I \quad (122)$$

Moreover, inserting equation 118 into equation 121, and comparing coefficients, we see that the projector states should also satisfy the property that:

$$\langle p_i^{(a)} | \tilde{\phi}_j^{(a)} \rangle = \delta_{ij} \quad (123)$$