



# A novel algorithm for circumventing the need to model large supercells of mismatched material interfaces



Noam Levi Hadari<sup>1</sup> & Maytal Caspary Toroker<sup>1,2,3</sup>

A longstanding challenge in materials science has been the computational modeling of interfaces between materials with different lattice parameters. Traditional approaches using plane-wave basis sets require either introducing artificial strain through unified lattice parameters or constructing prohibitively large supercells to accommodate the mismatch. These limitations have often deterred researchers from investigating large, mismatched interfaces, creating a gap in the understanding of these important systems. This work introduces an innovative algorithm that adaptively tunes the plane-wave basis sets to match the periodic structure of each material across the interface. By eliminating the need for extensive supercells or compromised lattice parameters, this new method reduces computational costs while retaining reliable results. The ability to efficiently calculate the eigen-energies of such mismatched systems, a crucial step for full density functional theory (DFT) calculations, is demonstrated with two dimensional versions of InAs/Si and SiC/Si interface potentials.

Modeling an interface between two materials is a highly valuable approach for various applications, including designing two-dimensional materials for advanced electronics, the development of battery materials for energy storage, and the integration of new materials into microelectronics<sup>1–4</sup>. However, modeling such interfaces accurately at the quantum scale in a periodic system poses significant challenges<sup>5</sup>. One source of difficulty is the need to model an interface system between materials with different lattice parameters. When two materials with different lattice parameters are joined together, the interface is usually incoherent, which means that the atomic positions do not align perfectly across the boundary. To accommodate the lattice mismatch, a strain field and a dislocation network are generated at the interface, which affect the mechanical and electronic properties of the heterostructure<sup>6</sup>.

Modeling mismatched interface systems would require a large unit cell that can capture both materials while keeping the boundary of the unit cell periodic. For example, interface systems with a 25% mismatch require a supercell consisting of three and four unit cells of each material along the interface plane, as shown schematically in Fig. 1. To represent relaxed interfaces between materials with smaller mismatches, even larger supercells are required. For instance, SiC/Si interface with ~20% mismatch requires four and five unit cells of Si and SiC, respectively, along the interface<sup>1</sup>. Similarly, the InAs/Si interface, where the mismatch is approximately 12.5%, consists of seven unit cells of InAs across from eight unit cells of Si<sup>2</sup>. The same principle applies to two-dimensional materials that do not match

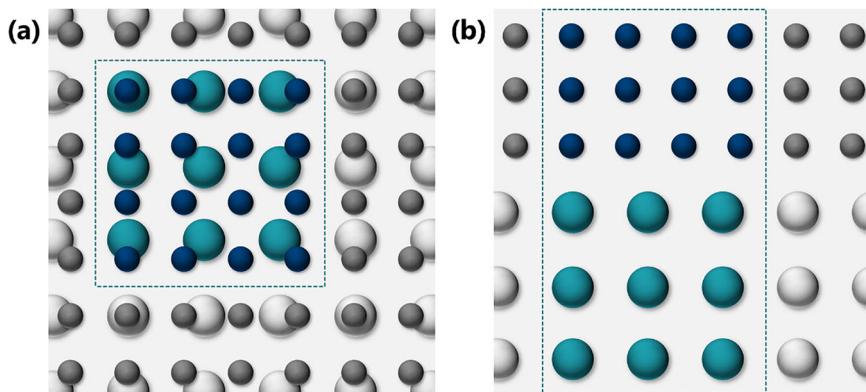
their substrate lattice parameter, such as *MoS*<sub>2</sub> on *Au* substrate, where the mismatch is ~9.1% thus require 11 unit cells of *MoS*<sub>2</sub> atop 10 *Au* unit cells<sup>3</sup>. Adding atomic layers parallel to the interface in those supercells increases the number of atoms significantly.

Such large unit cells, often containing hundreds of atoms, which approach the upper limit of typical electronic structure calculations, are computationally expensive and time-consuming to simulate. Additionally, using large supercells to represent relaxed interfaces usually imposes some amount of biaxial strain to at least one of the materials, as the ratio between the relaxed lattice parameter of two lattices is rarely an exact ratio of small numbers<sup>1–4,7,8</sup>. The forced strain is another limitation when considering simulating mismatched interfaces.

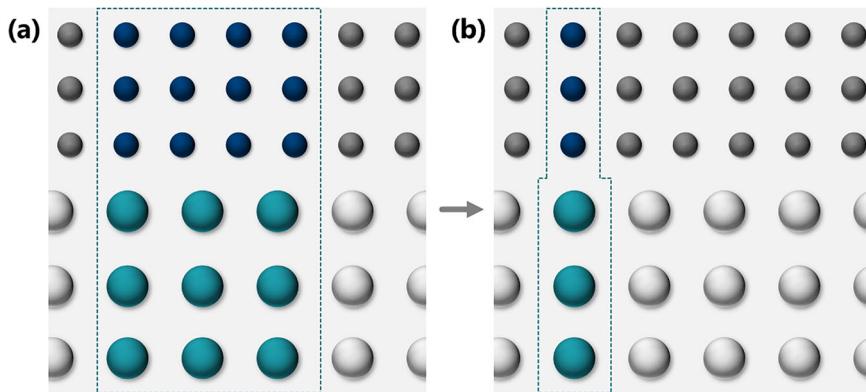
The computational challenges of modeling mismatched interfaces with large supercells have been addressed through the development of several approximation methods and numerical techniques. Effective medium theories and cluster expansion methods, for example, allow for modeling of complex systems with reduced computational demands<sup>9,10</sup>. Techniques like k-point folding and symmetry-adapted perturbation theory facilitate the use of smaller supercells while capturing essential long-range interactions<sup>11</sup>. Advanced methods, including machine learning-based structure prediction, have also enabled the extrapolation of properties from simulations with smaller supercells, balancing computational efficiency with model accuracy. Despite decades of research addressing the issue of large system sizes, the

<sup>1</sup>Department of Materials Science and Engineering, Technion-Israel Institute of Technology, Haifa, Israel. <sup>2</sup>The Nancy and Stephen Grand Technion Energy Program, Technion - Israel Institute of Technology, Haifa, Israel. <sup>3</sup>The Resnick Sustainability Center for Catalysis, Technion - Israel Institute of Technology, Haifa, Israel. e-mail: [maytalc@technion.ac.il](mailto:maytalc@technion.ac.il)

**Fig. 1 | Schematic of the relaxed interface with 25% mismatch.** **a** Top view and **b** side view, with the supercell boundaries marked by dashed lines. Large teal and small blue spheres denote the two atom types; grey spheres mark atoms outside the supercell. The mismatched interface appears as rows of three teal spheres opposite four blue spheres.



**Fig. 2 | Two types of supercells representing relaxed interface with 25% mismatch.** Large teal and small blue spheres represent the two atom species. **a** The rectangular supercell (dashed line) defines the system's periodicity. **b** The proposed smaller supercell (dashed line) redefines the periodicity; it is significantly smaller than in **a**. Grey spheres indicate atoms outside the cell.



challenge of accurately modeling interfaces with mismatches remains unresolved.

In this study, we propose an approach to simulate interfaces between materials with different lattice parameters that avoids the need for matching lattice parameters or employing very large supercells. By using a compact supercell with a single unit cell from each material and applying periodic boundary conditions with optimized basis functions, this method can achieve reliable results efficiently, minimizing artificial strain and enhancing computational efficiency.

## Results

### Novel method for modeling material interfaces

In this section, we presented a detailed description of our novel algorithm for constructing and solving the Kohn-Sham Hamiltonian matrix eigenvalue problem. Calculation of eigenvalues is a central step in performing density functional theory (DFT) calculations. We describe a solution suitable for a two-dimensional system, and only consider the  $\Gamma$  point in  $k$ -space for simplicity.

### Novel basis set with changing periodicity

In the simulation of an interface system with a mismatch, a supercell is typically employed to model the system, utilizing the periodic nature of the cell and the application of Bloch's theorem. However, when the mismatch is small, the supercell must be enlarged to accurately describe the system while maintaining the periodic boundary condition, thus making the simulation computationally expensive. According to the proposed novel method, the supercell can be reduced to contain a single unit cell of each material across the interface, as seen in Fig. 2. This modification effectively redefines the boundary, resulting in a supercell boundary that is no longer rectangular but consists of joined frames of varying sizes.

Defining the supercell in this manner implies that the potential used to construct the Hamiltonian, as well as the system's eigenvectors, are not

rectangular in shape. This non-rectangular shape is illustrated in Fig. 3, which contains the potential energy of the system with the mismatched interface, and a spatial representation of an eigen-state from this system, obtained by using the proposed method.

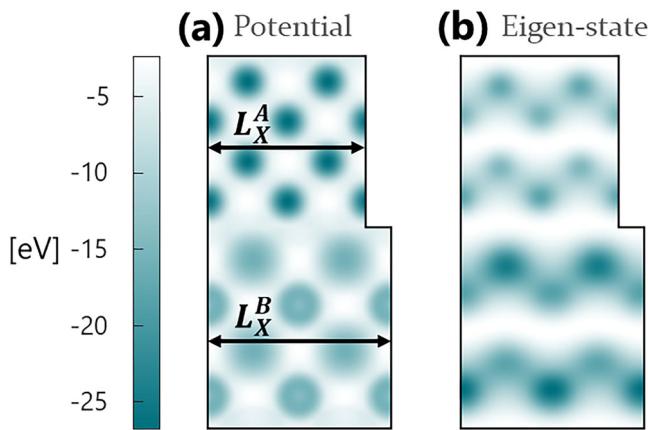
To model an interface system with a mismatch according to this method, there is a need for a novel basis set of functions that change their periodicity along the interface direction and are adapted to the lattice parameter of each material, thus mimicking the periodicity pattern represented by the corresponding supercell. The proposed basis set, which allows this change, is based on the plane waves basis set. It includes an additional factor, denoted as  $f_{(z)}$  in Eq. 1, which depends on the distance from the interface:

$$\varphi_n(z, x) = e^{\frac{2\pi i n_z z}{L_z}} * e^{\frac{2\pi i n_x x}{f_{(z)}}} \quad (1)$$

Here  $n_z$  and  $n_x$  are quantum numbers integers along the  $z$  and  $x$  directions, respectively.  $L_z$  is the length of the supercell along the  $z$  direction, while  $f_{(z)}$  is the function that defines the effective periodic boundary and the length of the supercell along the  $x$  direction. The  $f_{(z)}$  function can be chosen to be any function, where  $f_{(z)}$  equals the lattice parameter of material A at one side of the interface and equals the lattice parameter of material B at the second side of the interface as illustrated in Fig. 4. The  $f_{(z)}$  function complies with the system's periodic boundary, meaning that its period corresponds to the system's length in the  $z$ -direction (Eq. 2). It is possible to define different transition functions as  $f_{(z)}$ .

$$f(z) = f(z + L_z) \quad (2)$$

The basis set's boundary condition, defined in Eq. 3, will produce eigenstates with similar periodicity, which results from solving the eigenvalue problem with this basis set, as can be seen in Fig. 3. The boundary condition in the normal direction to the interface is similar to the periodic



**Fig. 3 | Potential energy surface and eigenstate map of the InAs/Si mismatched interface.** **a** Two-dimensional potential energy of a mismatched interface system, modeled as InAs/Si. On one side of the interface, the lattice parameter is marked  $L_X^A$ , and on the other side of the interface, the lattice parameter is  $L_X^B$ . **b** Two-dimensional spatial representation of an eigen-state, in absolute square, for the same interface system, corresponding to energy of 6.15 eV. Both the potential and the eigen-state are defined in a frame that is not rectangular.

boundary condition in the usual plane waves basis set Fig. 5.

$$\varphi_n(z, x) = \varphi_n(z + L_z, x + f(z)) \quad (3)$$

The basis set functions are orthogonal between themselves, except when the quantum number in the direction parallel to the interface,  $n_x$ , is identical (see Eq. 4):

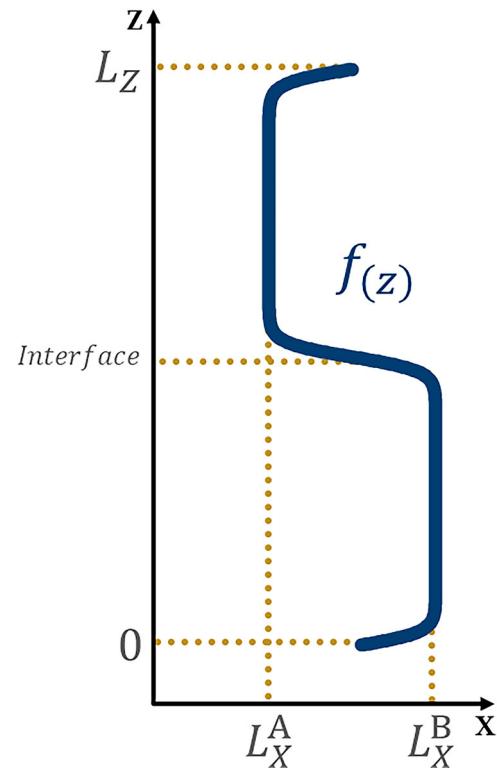
$$\begin{aligned} \langle \varphi_m | \varphi_n \rangle &= \int_0^{L_z} \int_0^{f(z)} e^{\frac{2\pi i s(n_x - m_x)z}{L_z}} * e^{\frac{2\pi i s(n_x - m_x)x}{f(z)}} dx * dz = \\ &\begin{cases} \text{for } n_x \neq m_x, \int_0^{L_z} e^{\frac{2\pi i s(n_x - m_x)z}{L_z}} * \left( \frac{f(z)}{2\pi i s(n_x - m_x)} \right) * \left( e^{2\pi i s(n_x - m_x)} - 1 \right) dz = 0 \\ \text{for } n_x = m_x, \int_0^{L_z} e^{\frac{2\pi i s(n_x - m_x)z}{L_z}} * f(z) dz \neq 0 \end{cases} \end{aligned} \quad (4)$$

Unlike plane waves, two basis functions are not orthogonal to each other if  $n_x = m_x$ , even if  $n_z = m_z$  because of the added  $f(z)$  function. Because the basis set is not completely orthogonal, the overlap matrix  $S_{mn} = \langle \varphi_m | \varphi_n \rangle$  is needed and used in the generalized eigenvalue problem.

### Numerical demonstrations and test cases

In this section we describe the results of applying our novel method to two systems, using two-dimensional test potential energies. We compare our results with those obtained from the conventional plane wave method using larger supercells. The interfaces examined involve Silicon (Si) and two other semiconductor materials - Silicon Carbide (SiC) and Indium Arsenide (InAs). These material combinations are highly relevant in the semiconductor industry, where they find applications in a diverse array of electronic and optoelectronic devices<sup>1,2</sup>. The Si/SiC and Si/InAs interfaces were specifically chosen for this study due to the significant lattice mismatch that exists between the constituent materials. This lattice mismatch is substantial enough and cannot be ignored when considering the properties and behavior of these heterojunctions. At the same time, the magnitude of the mismatch falls within a range that still allows for the use of conventional computational methods and simulation techniques to model and analyze these interfaces and therefore provide good test cases for comparison<sup>12–18</sup>.

We tested the validity of our new basis set by applying it to a system of quantum harmonic oscillators as described in the SI.



**Fig. 4 | Periodic transition function between two lattice parameters.** Function  $f(z)$  represents a smooth transition between the lattice parameter  $L_X^A$  of material A and  $L_X^B$  of material B along the z-axis. The point labeled as interface indicated the region where the transition occurs. The function  $f(z)$  is periodic with respect to the system's length  $L_z$ .

The detailed procedure for constructing two-dimensional interface models from bulk DFT potentials is provided in the SI.

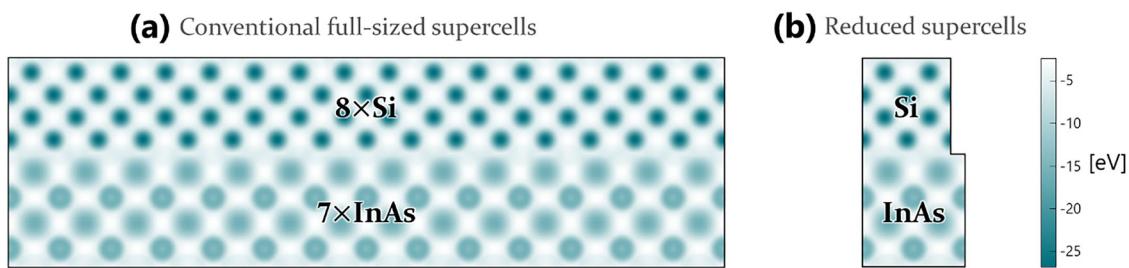
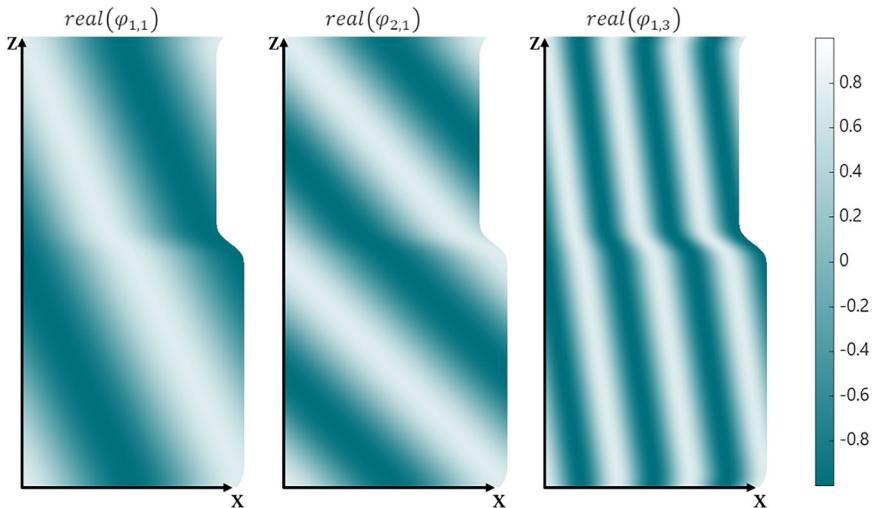
### Demonstrating with InAs/Si interface

The InAs/Si system exhibits 12.5% mismatch in lattice parameters. The bulk InAs was modeled with a zinc-blende structure and lattice parameter of 6.24 Å, while the Si bulk was modeled with a diamond structure and lattice parameter of 5.46 Å. The conventional InAs/Si supercell was created by appending the potential of eight unit cells of bulk Si, and the potential of seven unit cells of bulk InAs, with the interface oriented along the (100) plane for both materials. Both the conventional and the reduced supercells are shown in Fig. 6. Due to the periodic boundary conditions implemented in our model, both interface terminations (In and As) are inherently present, one at the center of the simulation cell and one at the boundary, as detailed in the Supplementary Information. The two supercell potentials were incorporated into the code as inputs. The corresponding eigen-states and their eigen-energies were computed using either the conventional PW method or the novel basis set functions.

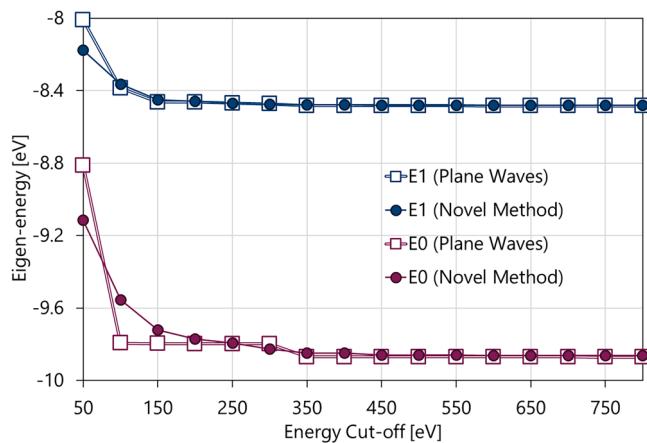
A test was performed to assess the convergence of eigen-energies within the system, utilizing both PW and the novel basis set. As depicted in Fig. 7, the test focused on the two lowest eigen-energies of the InAs/Si interface, modeled using a reduced supercell. The results demonstrate that the new basis set achieves convergence comparable to the PW basis set. A stability threshold of less than 0.01 eV was attained by both basis sets when the energy cut-off was set to at least 350 eV. Energy cut-off of 350 eV was thus used for comparing between both methods.

A comparison of the lowest eigen-energies derived from both methods validated the effectiveness of the proposed method. There was excellent agreement between the lowest eigen-energies obtained from both methods, and the spatial distribution of the eigen-states are also consistent across both methods, as depicted in Fig. 8. The first two eigenstates seen in Fig. 8 are

**Fig. 5 | Spatial representation of three basis functions showing periodicity change.** Real part of a three basis functions, denoted  $\varphi_{(nz=1, nx=1)}$ ,  $\varphi_{(nz=2, nx=1)}$  and  $\varphi_{(nz=1, nx=3)}$ . All three basis functions show periodicity change in x direction at a certain region in the z axis. The shape of the defined functions is not rectangular, as it follows the periodicity change and thus keeps the boundaries periodic.



**Fig. 6 | Potential energy of the InAs/Si interface in two supercell representations.** **a** Rectangular conventional supercell contains eight unit cells of Si across seven unit cells of stoichiometric InAs. **b** Reduced supercell includes only single unit cell of each material on both sides of the interface.



**Fig. 7 | Convergence of the lowest two eigen-energies using two basis sets.**  $E_0$  and  $E_1$  are the lowest eigen-energies of the InAs/Si interface system represented by a reduced supercell. Hollow squares mark values obtained with a plane wave basis set, and filled circles mark values obtained with the suggested novel method using a modified basis set.

spatially delocalized on the entire interface, while the next two eigenstates are located on each side of the interface. It's noteworthy that employing a larger supercell leads to the emergence of additional eigenstates, a result that can be attributed to the band folding effect of using supercells.<sup>11</sup>

We compared two different interface configurations within the supercell. As detailed in the Supplementary Information, the eigenstate spatial distributions maintain their characteristic patterns across configurations. The slight differences in eigenvalues between configurations (ranging from 0.03–0.09 eV) suggest an effect similar to small variations in k-vector sampling.

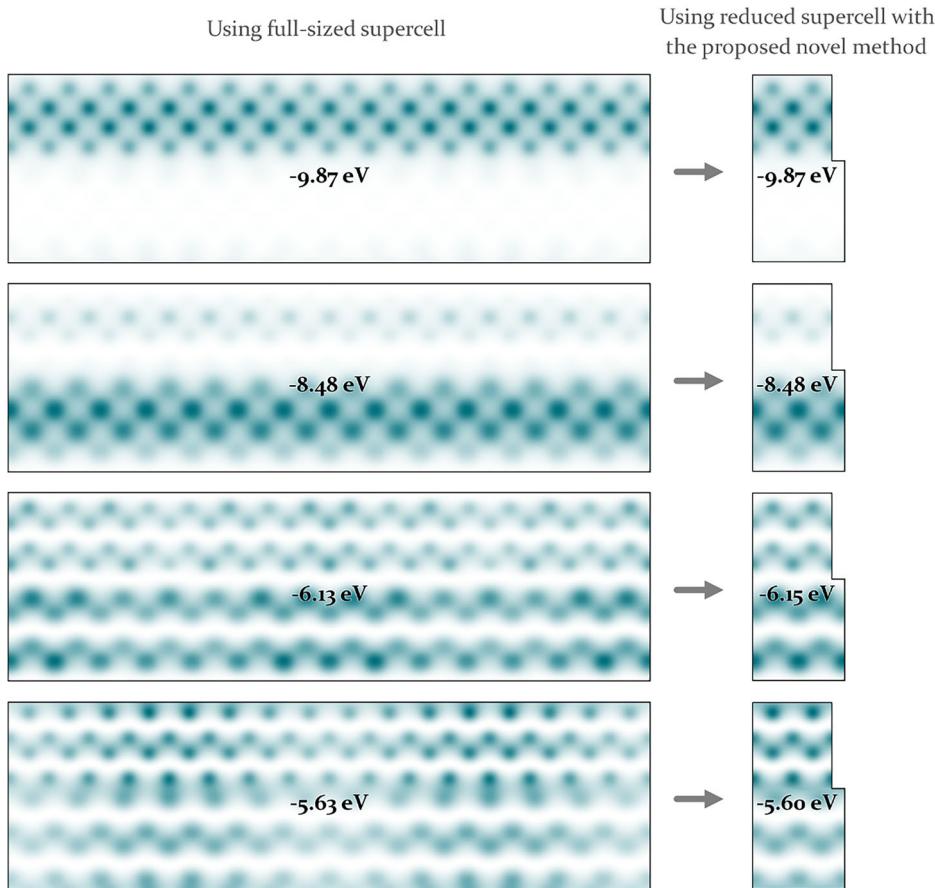
### Demonstrating with SiC/Si Interface

The SiC/Si interface is characterized by a lattice parameter mismatch of 25%. The bulk SiC was represented using a zinc-blende structure with a lattice parameter of 4.36 Å, while the Si bulk was modeled with a diamond structure and lattice parameter of 5.45 Å. To construct the conventional SiC/Si supercell, we combined the potentials from four unit cells of bulk Si, and five unit cells of SiC bulk, with the interface oriented along the (100) plane for both materials. In constructing this interface, we expanded the SiC potential unit cell to ensure that the interface consisted of a layer of carbon atoms on the SiC side and silicon atoms on the Si side. Figure 9 illustrates both the conventional and reduced supercells. The full supercells and reduced supercells were used as inputs to the code. The eigenstates and their corresponding eigen-energies were calculated using either the traditional PW method or a novel basis set functions for comparison. Energy cut-off of 350 eV was used for comparing between both methods.

A comparison of the lowest eigen-energies calculated using both methods gives an initial assessment of the effectiveness of the proposed approach. Excellent agreement was observed between the lowest eigen-energies obtained from the two methods, and the spatial distribution of the eigen-states was also consistent across both techniques, as illustrated in Fig. 10. This figure shows that the eigen-states are either delocalized over the entire interface or localized on each side of the interface, a result that holds regardless of the method employed.

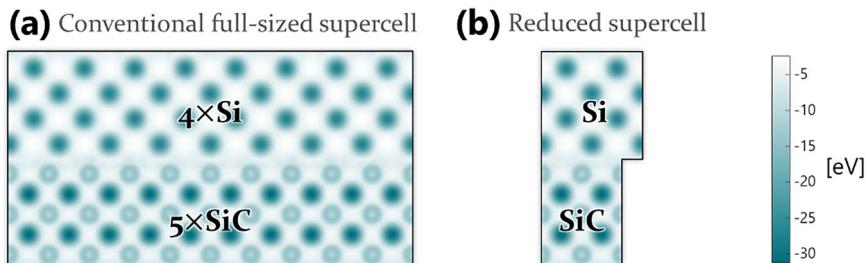
We investigated the eigen-energy full spectra for the SiC/Si interface system and compared the eigen-energies extracted using the reduced supercell method against the conventional full supercell method. Figure 11 presents a visual comparison between the eigenvalue spectra obtained from these two approaches, highlighting the correspondence between their respective energy levels and associated eigenstates. Numerous eigen-energies are consistent across both methods. However, the full supercell method introduces additional eigen-energies, which can be attributed to the

**Fig. 8 | Eigenstate distributions for the two-dimensional InAs-Si interface by two methods.** Each eigen-states shown on the left was derived using a traditional method on an extensive supercell. The corresponding eigen-states on the right pane were obtained by the novel method with a reduced unit cell. A darker color indicates a higher probability.



**Fig. 9 | Potential energy of a SiC/Si interface system in two supercell representations.**

a Rectangular conventional supercell containing four unit cells of Si across five unit cells of SiC.  
 b Reduced supercell includes only a single unit cell of each material on both sides of the interface.



band folding effect inherent in the use of supercells<sup>11</sup>. Certain eigenstates derived from the reduced supercell method do not correspond to any eigenenergy obtained from the full supercell calculation. These eigenstates could potentially be explained by performing the calculation in  $k$ -space other than the  $\Gamma$  point. To confirm that the matched eigenvalues are not coincidental, we manually compared their symmetry character and spatial representations, as illustrated in Fig. 10.

## Discussion

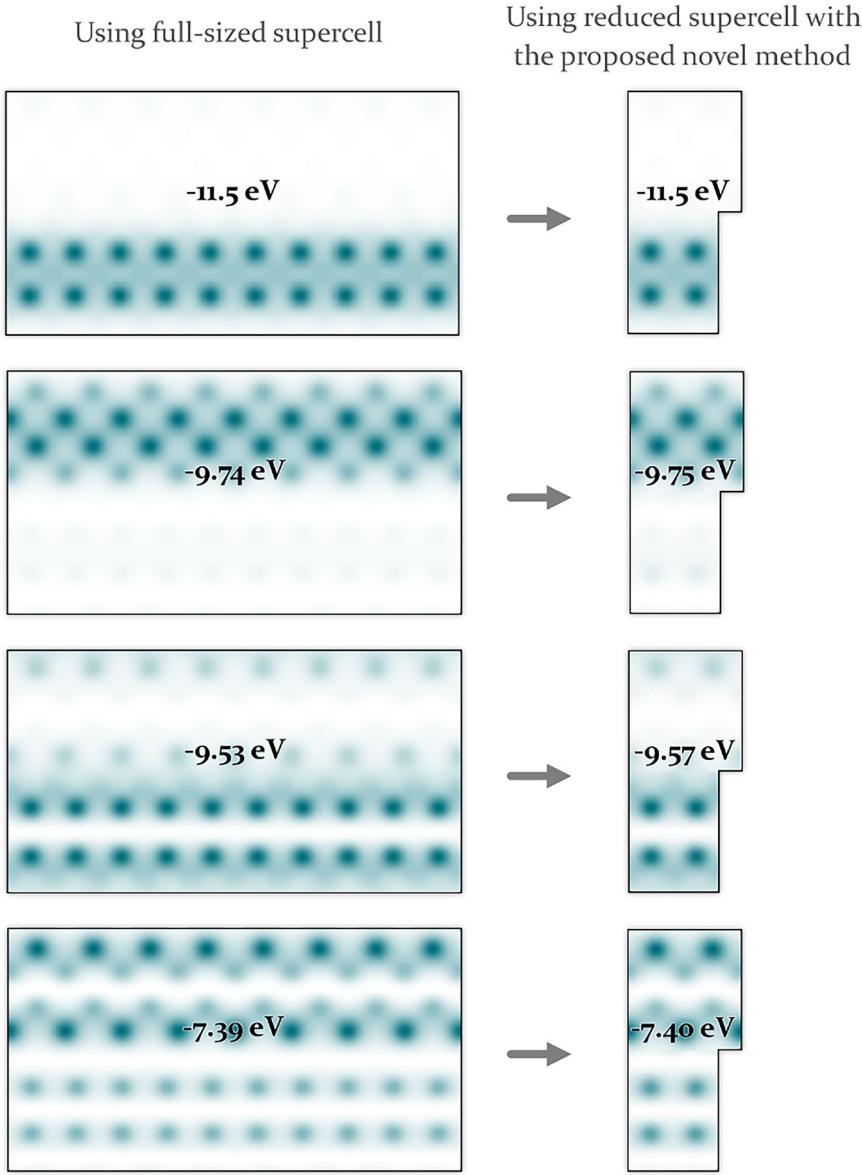
This article introduces an innovative algorithm addressing a challenge in computational materials science: accurately modeling interfaces between materials with different lattice parameters. Traditional density functional theory (DFT) methods face significant limitations in this area, typically requiring either large supercells or introducing artificial strain, which results in high computational costs and potential inaccuracies. Our novel approach overcomes these limitations by adaptively tuning plane-wave basis sets to match the periodic structure of each material across the interface, eliminating the need for extensive supercells or compromised lattice parameters.

The key innovation lies in our idea of adaptive basis sets that spatially adjust their periodicity along the interface direction, specifically tailored to the lattice parameters of each material. This method solves the eigen value problem while significantly reducing computational costs. The algorithm efficiently calculates the eigen-energies of mismatched systems, representing an important first step in the way of performing full DFT calculations with the proposed method.

Validation studies using two-dimensional potentials of InAs/Si and SiC/Si interfaces demonstrate the method's effectiveness. Despite using significantly smaller supercells, our approach shows excellent agreement with traditional basis sets in both eigen-energies and spatial distribution, though accuracy decreases at higher energies and does not perfectly reproduce the complete eigenvalue spectra. This enhanced computational efficiency makes it possible to study complex material systems that were previously impractical to model.

Our method shows clear promise for further developments and practical applications. A key planned enhancement is the extension to non-zero  $k$  points beyond the  $\Gamma$  point, which will enable comprehensive exploration of

**Fig. 10 | Eigenstate distributions for the two-dimensional SiC–Si interface by two methods.** Each eigen-state shown on the left was derived using a traditional method on an extensive supercell. The corresponding eigen-states on the right pane were obtained by the novel method. Darker color indicates higher probability.



full band structures and electronic properties. We are also actively working on a full three-dimensional implementation that will handle systems with lattice mismatches in multiple dimensions. Our approach can be extended by defining position-dependent periodicity functions for both in-plane dimensions ( $f_x(z)$  and  $f_y(z)$ ). The mathematical framework we've developed is fully compatible with this extension. Extending to more dimensions is important when dealing with lattice-mismatched materials that have different periodicities in both in-plane dimensions. We acknowledge that our current treatment of the interface region is a simplification that could be refined. Future work will explore varying the periodicity near interfaces to better capture interface-specific characteristics while maintaining computational efficiency in bulk regions. Looking forward, this approach has the potential to transform interface modeling by enabling more efficient and accurate calculations of mismatched interfaces.

## Methods

### Hamiltonian construction algorithm

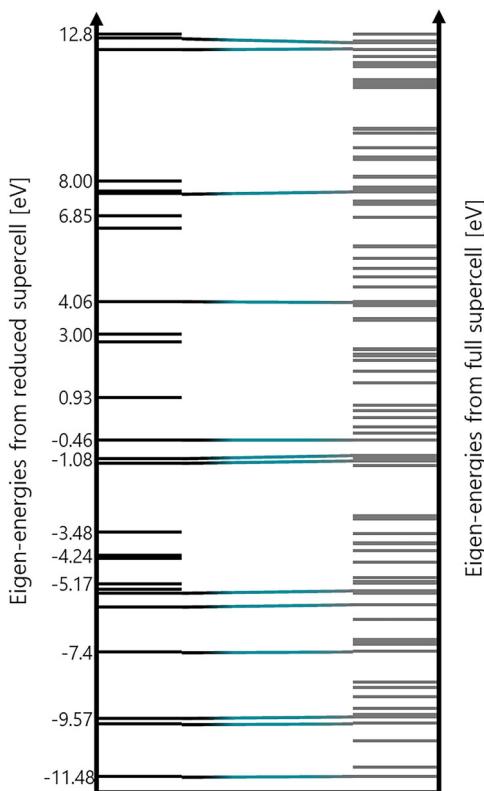
We have developed a program to construct the Kohn-Sham Hamiltonian of two-dimensional mismatched interface potential, using our novel method. The code also solves the generalized eigenvalue problem with the new basis set (Eq. 1). The code is available in the SI.

### Input requirements and preprocessing steps

The algorithm requires several inputs as detailed in section a of Fig. 12. Firstly, a potential that describes the system using a two-dimensional matrix,  $V$ , is needed. For testing the program, the total potential from calculations performed by the Vienna Ab Initio Simulation Package (VASP) is primarily used<sup>19–21</sup>. This three-dimensional potential is converted to two dimensions by averaging over one of the axes. Additionally, a transition function is required to define the perpendicular periodicity along the interface direction, represented as a vector,  $f_{(z)}$ . The length of the  $f_{(z)}$  vector should match the size of the potential matrix in the interface dimension. The sizes of the system's bounding box,  $L_z$  and  $L_x$ , are also necessary to establish the system's spatial dimensions. Finally, the cut-off energy is needed to determine the size of the basis set by defining all allowed combinations of  $n_z$  and  $n_x$ , which are the quantum numbers for the basis set.

### Basis set generation

The inputs are processed into column and row vectors, as mentioned in section b of Fig. 12. The spatial column vectors are systematically arranged in accordance with a mesh grid of spatial points, while excluding points that fulfill the condition  $X > f_{(z)}$ , thereby lying outside the supercell boundary. Conversely, the row vectors, representing quantum numbers  $n_z$  and  $n_x$ , are



**Fig. 11 | Eigenvalue spectra comparison between reduced and full supercell methods.** Black lines (left axis) for reduced supercell; grey lines (right axis) for full supercell. Matching eigenvalues are connected by bridging lines.

constructed based on a mesh grid of reciprocal space points, omitting any values that do not comply with the cut-off energy. In a similar manner to the Plane Waves basis set, it is necessary to define a cut-off energy that determines the number of basis functions used for modeling. Higher energy cut-off means more basis functions and higher accuracy of calculation, while making the calculation time longer. Cut-off energy can be determined by checking for convergence of properties up to a certain criteria. The quantum numbers values,  $n_z$  and  $n_x$ , that fulfill the cut-off criteria, define the number of basis functions and the size of the Hamiltonian matrix. Usually, the cut-off criteria is defined as shown in Eq. 5:

$$(|G + k|)^2 * \frac{\hbar^2}{2m} < E_{cut-off}, \text{ where } G = \frac{2\pi n_i}{L_i} \quad (5)$$

The proposed method and basis set require choosing the quantum numbers values that correspond to the minimal lattice parameter along the interface direction:

$$\left( \frac{4\pi^2 n_x^2}{\min(f(z))^2} + \frac{4\pi^2 n_z^2}{L_z^2} \right) * \frac{\hbar^2}{2m} < E_{cut-off} \quad (6)$$

### Block-wise construction of Hamiltonian matrix

The Hamiltonian and the basis overlap matrix are built according to the block matrix multiplication approach, as detailed in section c of Fig. 12. The matrices were calculated block after block, ordered by their rows and grouped by columns. In every loop, the matrix block is calculated by matrix

multiplication defined in Eq. 7:

$$H_{i,j} = \sum_k \varphi_i^k \left( -\frac{\hbar^2}{2m} \Delta \varphi_j^k + V^k * \varphi_j^k \right) \quad (7a)$$

$$S_{i,j} = \sum_k \varphi_i^k \varphi_j^k \quad (7b)$$

Where  $H_{i,j}$  is the Hamiltonian block, and  $S_{i,j}$  is the overlap matrix block. The potential is denoted by  $V$ , while  $\varphi_i$  and  $\varphi_j$  represent the basis functions associated with the matrix block in question. Furthermore,  $\Delta \varphi_j$  is the result of applying the Laplacian operator to the corresponding basis function. The index  $k$  indicates summation over all space points.

To optimize the computation of matrix blocks, several time-saving strategies are implemented. First, we avoid calculating blocks from the lower part of the large matrices by leveraging the Hermitian nature of the Hamiltonian and basis overlap matrices. This characteristic allows us to infer these lower blocks from their corresponding upper blocks, as detailed in section d of Fig. 12. Second, for blocks that are not from the first row of blocks, there is no need to construct the basis set matrix or calculate its Laplacian product of the block's column (the same bra  $\langle \rangle$ ). Instead, we retain these matrices in memory to enhance computational efficiency. Lastly, when processing blocks within the diagonal section, the basis matrix of the block row is identical to that of the block's column basis set, enabling us to simply replicate it rather than compute it anew.

### Laplacian operator

The construction of the Hamiltonian requires the application of the Laplacian operator to the basis functions during the computation of the kinetic term. Unlike the conventional Plane Waves basis set, which allows for the second derivative to be easily calculated through scalar multiplication (refer to Eq. 8a), the second derivative becomes more complex with our proposed basis set, as depicted in Eq. 8b. Nevertheless, our code adopts a streamlined method for computing the product of the Laplacian operator. It borrows the form from the Plane Waves basis set and omits the terms involving the derivative of  $f(z)$ , as illustrated in Eq. 8c. This method not only improves the accuracy of the results significantly due to numerical instability in calculating derivatives but also enhances computational efficiency.

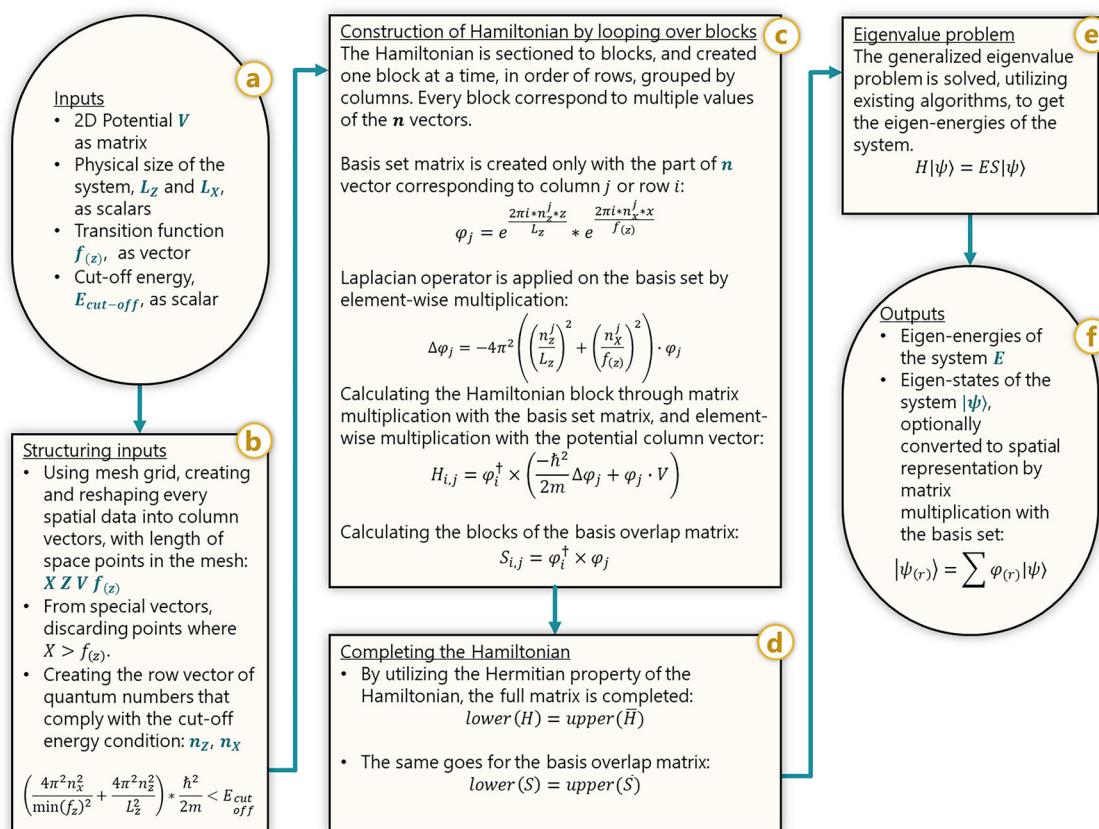
$$\Delta \varphi_n = \nabla^2 \left( e^{\frac{2\pi i * n_z * z}{L_z}} * e^{\frac{2\pi i * n_x * x}{L_x}} \right) = \left( \left( \frac{2\pi i * n_z}{L_z} \right)^2 + \left( \frac{2\pi i * n_x}{L_x} \right)^2 \right) * \varphi_n \quad (8a)$$

$$\Delta \varphi_n = \nabla^2 \left( e^{\frac{2\pi i * n_z * z}{L_z}} * e^{\frac{2\pi i * n_x * x}{f(z)}} \right) = \left( \left( \frac{2\pi i * n_z}{L_z} \right)^2 + \left( \frac{2\pi i * n_x}{f(z)} \right)^2 + \frac{2\pi i * n_x * x * f''(z)}{f_z^2} \right. \\ \left. - \frac{4\pi i * n_x * x * (f'(z))^2}{f_z^3} - \frac{8\pi^2 * n_x * n_z * x * f'(z)}{f_z^2 * L_z} + \left( \frac{2\pi i * n_x * x * f'(z)}{f_z^2} \right)^2 \right) * \varphi_n \quad (8b)$$

$$\Delta \varphi_n = \nabla^2 \left( e^{\frac{2\pi i * n_z * z}{L_z}} * e^{\frac{2\pi i * n_x * x}{f(z)}} \right) = \left( \left( \frac{2\pi i * n_z}{L_z} \right)^2 + \left( \frac{2\pi i * n_x}{f(z)} \right)^2 \right) * \varphi_n \quad (8c)$$

### Solving the generalized eigenvalue problem

The generalized eigenvalue problem of the complete Hamiltonian matrix written in section e of Fig. 12 is solved by the efficient *eig* function in MATLAB<sup>22,23</sup>. The outputs from this solution are the eigen-energies and eigen-states of the interface. The eigen-states can be optionally converted to spatial representation by matrix multiplication with the basis set.



**Fig. 12 | Flowchart of the Hamiltonian matrix construction algorithm.** Steps: a input parameters, b structuring inputs, c block-matrix multiplication to build the Hamiltonian, d completing the assembly, e solving the generalized eigenvalue problem, f generating outputs.

## Data availability

Data is provided within the manuscript or supplementary information files

## Code availability

The code used to implement the algorithm described in this study is available as supplementary information online.

Received: 15 January 2025; Accepted: 23 May 2025;

Published online: 13 June 2025

## References

1. Pizzagalli, L., Cicero, G. & Catellani, A. Theoretical investigations of a highly mismatched interface: SiC/Si(001). *Phys. Rev. B* **68**, 195302 (2003).
2. Sant, S., Luisier, M. & Schenk, A. Density functional theory based analysis of the origin of traps at the InAs/Si hetero-interface. *Appl. Phys. Lett.* **111**, 242102 (2017).
3. Tumino, F., Casari, C. S., Li Bassi, A. & Tosoni, S. Nature of point defects in single-layer MoS<sub>2</sub> supported on Au (111). *J. Phys. Chem. C* **124**, 12424–12431 (2020).
4. Lepley, N. D. & Holzwarth, N. A. W. Modeling interfaces between solids: Application to Li battery materials. *Phys. Rev. B* **92**, 214201 (2015).
5. M. Yang, M. & Mo, Y. Interfacial defect of lithium metal in solid-state batteries. *Angewandte Chemie* **133**, 21664–21671 (2021).
6. Jain, S. C., Harker, A. H. & R. A. Cowley, R. A. Misfit strain and misfit dislocations in lattice mismatched epitaxial layers and other systems. *Philos. Mag. A*, **75**, 1461–1515 (1997).
7. Tsai, A. Epitaxial formation of SiC on (100) diamond. *ACS Appl. Electron. Mater.* **2**, 2003–2009 (2020).
8. Farmanbar, M. & Brocks, G. First-principles study of van der Waals interactions and lattice mismatch at MoS<sub>2</sub>/metal interfaces. *Phys. Rev. B*, **93**, 613–622 (2016).
9. Wang, K., Cheng, D. & Zhou, B.-C. Generalization of the mixed-space cluster expansion method for arbitrary lattices. *npj Comput. Mater.* **9**, 75 (2023).
10. Gerber, E., Yao, Y., Arias, T. A. & Kim, E.-A. Ab initio mismatched interface theory of Graphene on  $\alpha$ -RuCl<sub>3</sub> : Doping and Magnetism. *Phys. Rev. Lett.* **124**, 106804 (2020).
11. Ku, W., Berlijn, T. & Lee, C.-C. Unfolding first-principles band structures. *Phys. Rev. Lett.*, **104**, 216401 (2010).
12. Reichman, Y. & Toroker, M. C. A numeric approach for investigating electron dynamics in zinc-blende semiconductor heterostructures. *Adv. Theory Simul.* **6**, 2300155 (2023).
13. Ji-Sang, P. et al. Quick-start guide for first-principles modelling of semiconductor interfaces. *J. Phys. Energy* **1**, 016001 (2019).
14. Daniele Stradi et al. Method for determining optimal supercell representation of interfaces. *J. Phys.: Condens. Matter* **29**, 185901 (2017).
15. Detz, H. & Maier, J. Atomistic modeling of interfaces in III-V semiconductor superlattices. *Phys. Status Solidi* **253**, 613–622 (2016).
16. Hine, N. D. M., Frensch, K., Foulkes, W. M. C. & Finnis, M. W. Supercell size scaling of density functional theory formation energies of charged defects. *Phys. Rev. B* **79**, 024112 (2009).
17. Marsman, M. Periodic systems, plane waves, the PAW method, and hybrid functionals. In *Computational Methods in Catalysis and Materials Science: An Introduction for Scientists and Engineers* (eds. van Santen, R. A. & Sautet, P.) ch. 4, 61–76 (Wiley-VCH Verlag GmbH & Co. KGaA, 2009).
18. Benedek, R. et al. First principles simulation of a ceramic/metal interface with misfit. *Phys. Rev. Lett.* **84**, 3362–3365 (2000).

19. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
20. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
21. Kresse, G. & Hafner, J. Ab initiomolecular dynamics for liquid metals. *Phys. Rev B* **47**, 558–561 (1993).
22. MATLAB. version 9.13.0 (R2022b). The MathWorks Inc., Natick, Massachusetts, USA (2022).
23. Moler, C. B. & Stewart, G. W. An algorithm for generalized matrix eigenvalue problems. *SIAM J. Numer. Anal.* **10**, 241–256 (1973).

## Acknowledgements

This research was conducted with the support of The Boeing Company, as part of the Boeing-Technion Sustainable Aviation Fuel Innovation Center. We sincerely thank Boeing for their valuable support and collaboration. This project was also conducted within the framework of the Guy Sella Memorial Project at Technion, established by SolarEdge Technologies LTD. Partial funding from The Israeli Sustainable Aviation Fuel Knowledge Center – iSAF, supported by The Israeli Ministry of Innovation, Science, and Technology, is gratefully acknowledged. This article is based upon work from COST IG18234 (NanoCatML), supported by COST (European Cooperation in Science and Technology) <http://www.cost.eu>.

## Author contributions

N.L.H. and M.C.T. wrote the main manuscript text, N.L.H. prepared the figures and M.C.T. supervised the project. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-025-01675-6>.

**Correspondence** and requests for materials should be addressed to Maytal Caspary Toroker.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025