Supplemental Materials for "The study of hydrogen Adsorption-induced Topological Surface State In-Out Hop in MgB_2 Nodal-Line Semimetals via Physics-Informed Bayesian Optimization"

Qinchi Yue,[1,*] Kun Bu,[2,†] Ruzhi Wang,[1] and Changhao Wang[1,‡]

[1] *State Key Laboratory of Materials Low-Carbon Recycling, College of Material Science and Engineering, Beijing University of Technology, Beijing 100124, China.*
[2] *School of Physics and Optoelectronic Engineering, Shandong University of Technology, Zibo 255000, China*

I. THE ROBUSTNESS OF THE TSSS UNDER THE INFLUENCE OF SPIN-ORBIT COUPLING(SOC)

To examine the robustness of the TSSs under the influence of SOC, we have calculated the band structure of the MgB_2 (010) 20-layered slab both with SOC and without SOC, as shown in Fig. S1. It is shown that the SOC does not induce a non-trivial energy gap in the TSSs, and the TSSs continue to be primarily contributed by the surface B atoms.

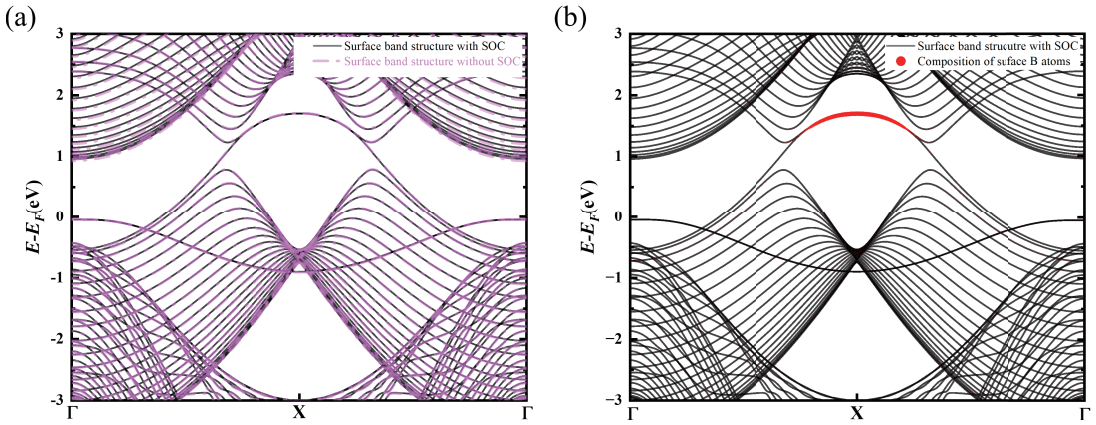

FIG. S1: Projected band structure calculations for the MgB_2 (010) 20-layered slab. (a) Band structures of the MgB_2 (010) 20-layered slab with and without SOC. The black solid line denotes the band calculation incorporating SOC, while the pink dashed line represents the calculation without SOC. (b) Band projection for the surface B atoms of the MgB_2 (010) 20-layered slab structure. The black solid line shows the band structure considering SOC, and the red projection highlights the contribution from the surface B atoms.

II. THE SURFACE RECONSTRUCTION FROM ZIGZAG TO BEARDED TERMINATION

The structural transformation of the MgB_2 (010) 20-layered slab from zigzag to beard type under hydrogen adsorption-induced conditions, as shown in Fig. S2. The definition of surface structure in this article is based on the definitions of Zigzag and Bearded in the boundary states of graphene[1, 2].

*First author
†Corresponding author
‡Corresponding author; Electronic address: wangch33@bjut.edu.cn

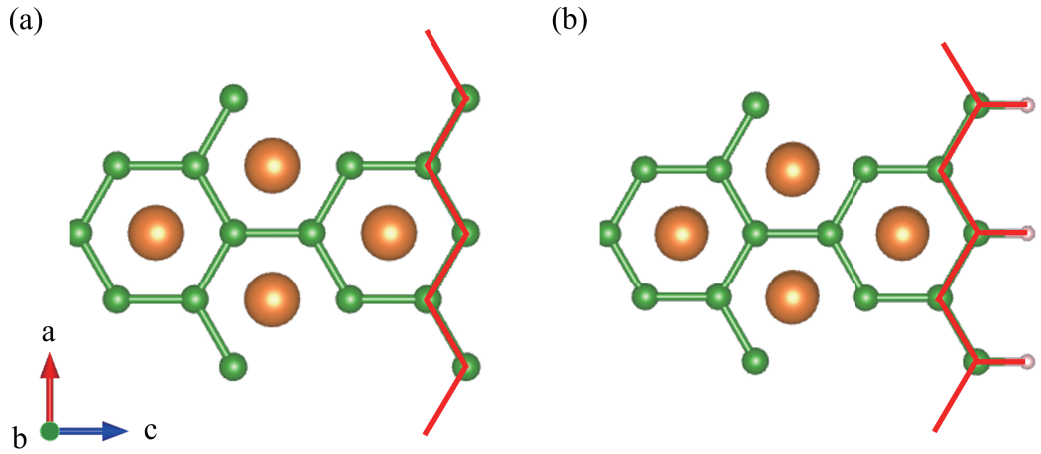

FIG. S2: The structures of MgB_2 (010) 20-layered slab before and after H atom adsorption (a) zigzag structure (b) bearded structure.

III. THE APPROXIMATION OF THE DISPERSION NEAR HIGH-SYMMETRY POINTS BY SECOND-ORDER POLYNOMIALS

To examine the accuracy of the approximation of the dispersion near high-symmetry points by second-order polynomials, we have compared the results of the dispersion between two Dirac points obtained from second-order polynomials and those obtained through DFT-computed, respectively, as shown in Fig. S3. It can be observed that the TSSs are concentrated in a very small area near the highly symmetrical point (approximately one-tenth of the complete Γ-X-Γ path), and the second-order polynomial ($\varepsilon_a(\boldsymbol{k}) = -4.261\boldsymbol{k}^2 + 8.6919\boldsymbol{k} - 1.9542$) can accurately describe the energy dispersion of the TSSs.

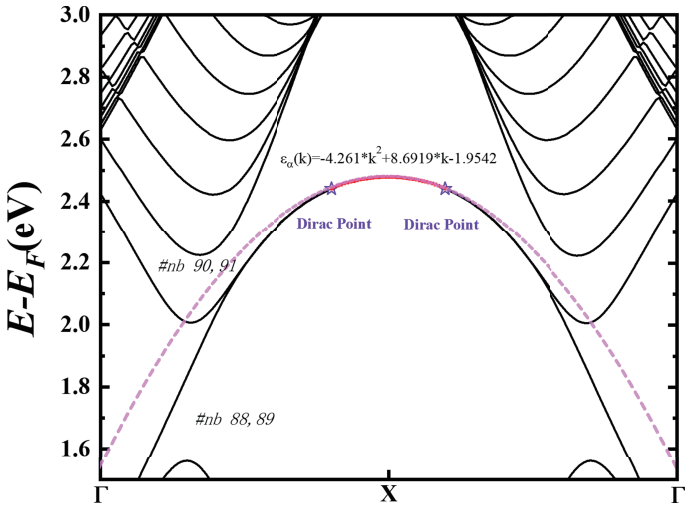

FIG. S3: The approximation of the dispersion near high-symmetry points by second-order polynomials. The pink dotted line represents the quadratic function curve, and the black line represents the energy band of MgB_2 (010) 20-layered slab. The positions of the Dirac points are marked with purple pentagrams in the energy band.

IV. THE COMPARISON OF DATA FITTING USING DIFFERENT FITTING METHODS AND PIBO

To verify the advantages of PIBO optimization, we respectively used the following three methods to fit the TSSs in the most physically meaningful Process a). (1) We used a Bayesian optimization model without physics-informed constraints for fitting to prove the necessity of physical embedding. (2) We replaced Δ(d) in the $\mathcal{H}(\mathbf{k})$ with a third-order polynomial. Then, we fitted the eigenvalues of $\mathcal{H}(\mathbf{k})$ to prove the rationality of constructing the coupling

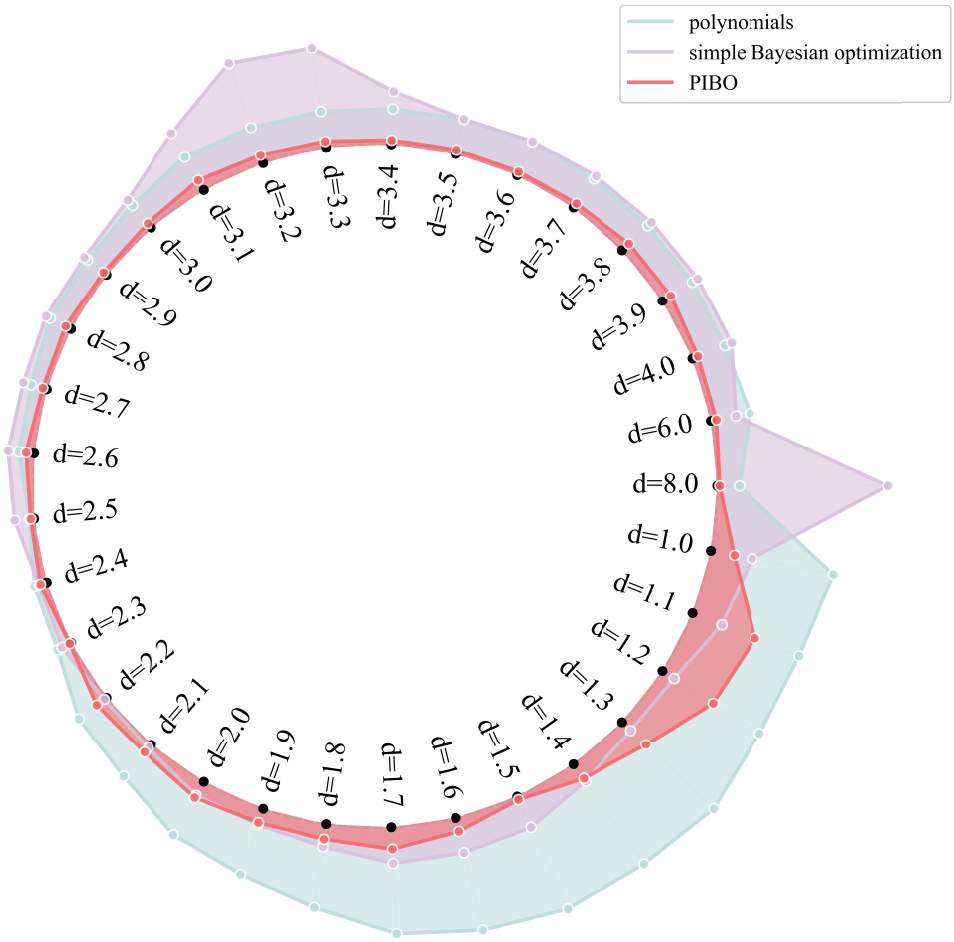

FIG. S4: The energy error values of TSSs (ΔE) obtained using different fitting methods. The green, purple, and red lines represent the energy error values of TSSs obtained through polynomial fitting, Bayesian optimization, and PIBO optimization, respectively.

function Δ(d). (3) We also attempted to use various fitting methods to calculate the TSSs energy values to prove the superiority of Bayesian optimization in hyperparameter tuning. We define ΔE to quantify the degree of data restoration by different fitting methods,

$$\Delta E = |E_{pred}| - |E_{actrual}|. \tag{S1}$$

In Eq. S1, E_{pred} represents the energy value of TSSs obtained by fitting, while $E_{actrual}$ represents the energy calculation value obtained through DFT-computed. By comparing the magnitude of Delta, the superiority and inferiority of different fitting methods can be intuitively compared.

First, we fit the TSSs using a Bayesian optimization model without physics-informed constraints. The results are shown as the purple line in Fig. S4. The maximum energy error obtained by this method is 0.1928 eV, which is significantly larger than the 0.07614 eV energy error of PIBO. Furthermore, during the entire dynamic fitting process, we observe irregular spikes in the error values at many positions (e.g., at 8 Å and 3.3 Å). This indicates that the hyperparameter tuning characteristics of a Bayesian optimization model without physics-informed constraints are ineffective when fitting the eigenvalues of $\mathcal{H}(\mathbf{k})$. Additionally, an error approaching 0.2 eV is unacceptable when analyzing the band structure of the system. Therefore, it is necessary to insert $\mathcal{P}_{\text{phys}}$ into the loss function $\mathcal{L}(\boldsymbol{\theta})$.

In addition, we replaced the $\Delta(d)$ in $\mathcal{H}(\mathbf{k})$ with a simple polynomial,

$$y(d) = c_3 d^3 + c_2 d^2 + c_1 d + c_0. \tag{S2}$$

where,$c_3 = -0.02492$,$c_2 = 0.03755$,$c_1 = -1.73392$,$c_0 = 2.48512$. The fitting results are presented as a green dashed line in Fig. S4. We can see that due to the lack of the nonlinear part, the effective Hamiltonian composed of $y(d)$

exhibits a larger energy deviation than PIBO near the In-Out Hop mutation point. Furthermore, the energy error is also larger at other positions, with a maximum energy error of 0.142 eV. We also used four common methods, Nelder-Mead[3], L-BFGS-B[4], Powell[5], and CG[6], for $\Delta(d)$ fitting. The Current function value of these methods was all greater than 100. Due to the excessively large data fitting deviation and error, they could not be effectively compared with the aforementioned methods, so they are not shown in the figure. In conclusion, PIBO is an irreplaceable data fitting method when studying the physical system in this paper.

[1] Nakada, K., et al., *Edge state in graphene ribbons: Nanometer size effect and edge shape dependence.* Physical Review B, 1996. **54**(24): p. 17954-17961.

[2] Su, X., et al., *Edge State Engineering of Graphene Nanoribbons.* Nano Letters, 2018. **18**(9): p. 5744-5751.

[3] Nelder, J.A. and R. Mead, *A Simplex Method for Function Minimization.* The Computer Journal, 1965. **7**(4): p. 308-313.

[4] Morales, J.L. and J. Nocedal, *Remark on algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound constrained optimization.* ACM Transactions on Mathematical Software, 2011. **38**(1): p. 1-4.

[5] Marumo, N. and A. Takeda, *Universal heavy-ball method for nonconvex optimization under Hölder continuous Hessians.* Mathematical Programming, 2024. **212**: p. 147-175.

[6] Pan, Y. and J. Pan, *roptim: An R package for general purpose optimization with C++.* R package version 0.1 6, 2022, 2022.