

Supplementary Information

Integrated classical/quantum–classical neural
networks, DFT, and experiments for predicting
bandgap and color of $\text{Cr}_x\text{Sb}_x\text{Ti}_{1-2x}\text{O}_2$
yellow-orange pigments

S. Y. Vaselnia¹, M. Khajeh Aminian^{1*}, R. Dehghan Banadaki¹

^{1*}Nano Pigments and Coatings Laboratory, Department of Physics,
Yazd University, Yazd, 89195-741, Iran.

*Corresponding author(s). E-mail(s): kh.aminian@yazd.ac.ir;
Contributing authors: vaselnia@stu.yazd.ac.ir;
re.dehghan@stu.yazd.ac.ir;

Contents

Supplementary Note 1. Quantum circuits and feature encoding	3
Supplementary Note 2. SHAP dependence analysis	4
Supplementary Figures	7
Supplementary Fig. 1	7
Supplementary Fig. 2	8
Supplementary Fig. 3	8
Supplementary Fig. 4	9
Supplementary Fig. 5	10
Supplementary Fig. 6	12
Supplementary Fig. 7	14
Supplementary Fig. 8	14
Supplementary Fig. 9	15
Supplementary Fig. 10	17
Supplementary Fig. 11	18
Supplementary Fig. 12	19
Supplementary Fig. 13	20
Supplementary Fig. 14	21
Supplementary Fig. 15	22
Supplementary Fig. 16	23
Supplementary Fig. 17	24
Supplementary Fig. 18	25
Supplementary Fig. 19	26
Supplementary Tables	27
Supplementary Table 1	27
Supplementary Table 2	27
Supplementary Table 3	27
Supplementary Table 4	28
Supplementary Table 5	28
Supplementary Table 6	29

Supplementary Note 1. Quantum circuits and feature encoding

We provide detailed descriptions of the quantum circuits employed in our hybrid quantum–classical neural network models. All quantum circuits were simulated using a noiseless statevector backend in PennyLane. Classical features $\mathbf{x} = (x_1, \dots, x_n) \in [0, \pi]^n$ were encoded on n qubits as described below.

Feature encoding

The HQCNN-SE is preceded by an angle embedding that applies data-dependent, non-trainable X -rotations to each qubit:

$$U_{\text{enc}}^{\text{HQCNN-SE}}(\mathbf{x}) = \bigotimes_{i=1}^n R_X(x_i), \quad R_X(\theta) = e^{-i\theta X/2}. \quad (1)$$

For the HQCNN-HEA and HQCNN-QI, each input feature x_i is encoded via a trainable affine rotation:

$$U_{\text{enc}}^{\text{train}}(\mathbf{x}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \bigotimes_{i=1}^n R_Y(\alpha_i x_i + \beta_i), \quad (2)$$

where α_i and β_i are trainable parameters.

Ansatz architectures

The HQCNN-SE contains L layers, each consisting of single-qubit rotations followed by entangling operations. The trainable parameters form a tensor $\boldsymbol{\theta} \in \mathbb{R}^{L \times n \times 3}$. At layer l , the local rotations acting on the n qubits are

$$U_{\text{rot}}^{(l)} = \bigotimes_{i=1}^n \text{Rot}(\theta_{l,i}^{(1)}, \theta_{l,i}^{(2)}, \theta_{l,i}^{(3)}), \quad (3)$$

where the single-qubit rotation gate is defined as

$$\text{Rot}(\phi, \theta, \omega) = R_Z(\omega) R_Y(\theta) R_Z(\phi). \quad (4)$$

Entanglement is introduced through a ring-connectivity pattern with layer-dependent range r_l :

$$U_{\text{ent}}^{(l)} = \prod_{i=1}^n \text{CNOT}_{i, (i+r_l) \bmod n}. \quad (5)$$

Each layer applies the local rotations followed by the entangling block, yielding the overall HQCNN-SE unitary

$$U_{\text{HQCNN-SE}} = \prod_{l=1}^L U_{\text{ent}}^{(l)} U_{\text{rot}}^{(l)}. \quad (6)$$

Each HQCNN-HEA layer consists of single-qubit rotations followed by full entanglement:

$$U_{\text{rot}}^{(l)} = \bigotimes_{i=1}^n \left(R_X(\theta_{l,i}^{(1)}) R_Y(\theta_{l,i}^{(2)}) R_Z(\theta_{l,i}^{(3)}) \right), \quad (7)$$

$$U_{\text{ent}} = \prod_{i=1}^{n-1} \prod_{j=i+1}^n \text{CNOT}(i, j), \quad (8)$$

and the total circuit is

$$U_{\text{HQCNN-HEA}} = \prod_{l=1}^L U_{\text{ent}} U_{\text{rot}}^{(l)}. \quad (9)$$

Also, each HQCNN-QI layer begins with full entanglement

$$U_{\text{ent}} = \prod_{i=1}^{n-1} \prod_{j=i+1}^n \text{CNOT}(i, j), \quad (10)$$

followed by trainable mixing rotations:

$$U_{\text{mix}}^{(l)} = \bigotimes_{i=1}^n R_X(\gamma_{l,i}) R_Y(\delta_{l,i}) R_Z(\beta_{l,i}), \quad (11)$$

with the full unitary

$$U_{\text{HQCNN-QI}} = \prod_{l=1}^L U_{\text{mix}}^{(l)} U_{\text{ent}}. \quad (12)$$

Supplementary Note 2. SHAP dependence analysis

To further interpret the predictive behavior of the models, we examined SHAP dependence plots for the most relevant features. The plots illustrate how variations in each feature influence the predicted target while also revealing interactions between features through the color coding of the data points.

E_g prediction – Scenario 1

Supplementary Fig. 18 presents the SHAP dependence plots of all models for E_g prediction in Scenario 1. For the C-NN model, CBM, VBM, and efermi have the strongest impact on the prediction. The CBM dependence plot displays a positive relationship with the target. As CBM increases, the SHAP values shift from negative to positive, showing that higher conduction band energies tend to increase the predicted E_g . The color scale is related to efermi, and the gradual color change from lower to higher values along the CBM axis shows that efermi tends to increase alongside CBM. In contrast, VBM shows the opposite trend. Increasing VBM results in progressively more negative SHAP contributions, indicating that higher valence band energies reduce

the predicted E_g . Here, the points are colored by CBM, and the color distribution indicates how changes in the conduction band position accompany variations in the valence band edge. The dependence plot for e_{fermi} reveals a negative trend across its range. As the e_{fermi} increases, SHAP values decrease from positive to negative contributions, suggesting that higher Fermi energies tend to reduce the predicted E_g in this model. CBM is mapped to the point colors, and the color gradient suggests a correlated distribution between e_{fermi} and CBM in the dataset.

In the HQCNN-SE model, CBM and VBM are the main factors shaping the prediction. The CBM dependence plot exhibits a strong positive relationship between CBM and its SHAP contribution. Higher CBM values correspond to a transition from negative to positive SHAP contributions. Point colors represent VBM, and the gradual change in color along the CBM axis reflects the correlated variation between these features. VBM behaves in the opposite way. Increasing VBM leads to progressively more negative SHAP contributions, indicating that higher valence band energies reduce the predicted E_g . The color scale corresponds to e_{fermi} , and its gradient shows how e_{fermi} and VBM change together. The dependence plot for the lattice angle γ shows relatively small SHAP magnitudes compared with the CBM and VBM, indicating a weaker influence on the prediction. There is a slight downward trend, where increasing γ leads to decreasing SHAP values. The points are colored by α , and no clear color pattern is visible, suggesting limited interaction between these two structural parameters in the model.

The SHAP dependence plots for both HQCNN-HEA and HQCNN-QI show similar trends. The CBM dependence plots show a strong positive relationship with the target. As CBM increases, SHAP values rise steadily from negative to positive contributions, indicating that higher conduction band energies lead to larger predicted E_g . The color scale corresponds to VBM, and the color gradient along the CBM axis reflects the correlated distribution of these features in the dataset. In contrast, increasing VBM results in progressively more negative SHAP contributions. In these plots, the points are colored by e_{fermi} , and the color transition suggests a correlated variation between the VBM and the e_{fermi} . The dependence plots for e_{fermi} show a relatively weak effect in both models. With increasing e_{fermi} , the SHAP values exhibit a decreasing trend, indicating a negative contribution. Nevertheless, the narrow SHAP range suggests that its influence is considerably smaller than the dominant contributions of CBM and VBM. In the HQCNN-HEA model, CBM is mapped to the point colors, while in the HQCNN-QI model, they are colored by VBM. The color gradients chiefly reflect the correlated variation between e_{fermi} and these features in the dataset.

E_g prediction – Scenario 2

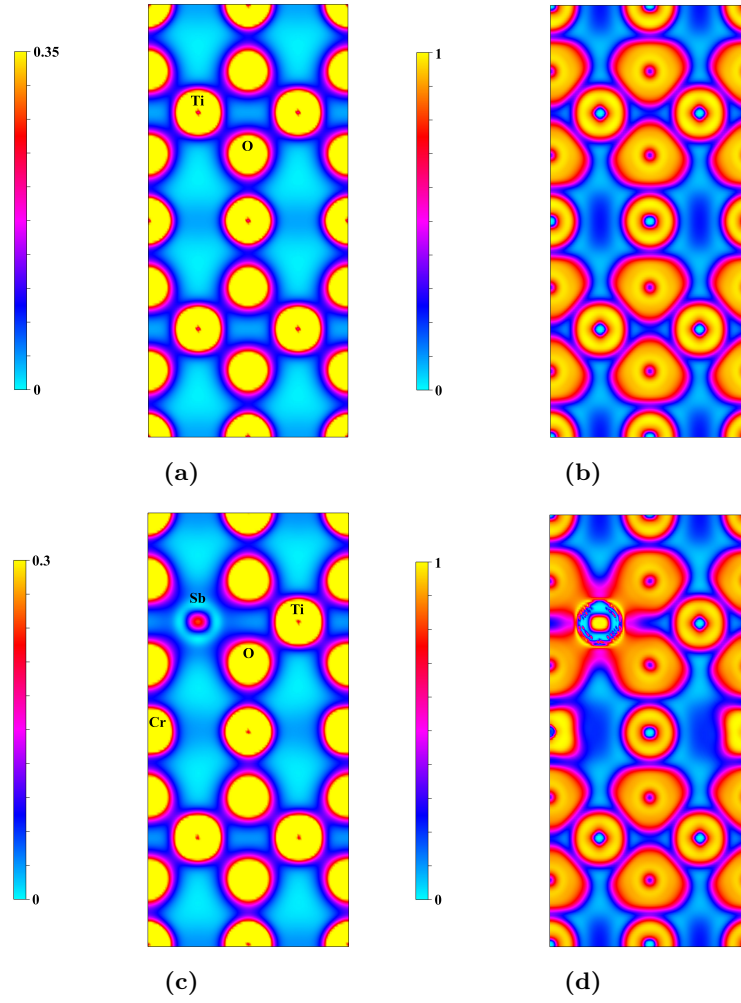
Supplementary Fig. 19 illustrates the SHAP dependence plots of all models for the prediction of E_g in Scenario 2. For the C-NN, the dependence plot of a_{sym} shows a clear nonlinear pattern. For small a_{sym} values, SHAP values are strongly negative, meaning the model predicts a lower E_g . With increasing a_{sym} , the SHAP values gradually become positive, and for $a_{\text{sym}} \approx 2\text{--}3$ the slope becomes steeper, meaning that larger a_{sym} values correspond to higher predicted E_g . Coloring by b shows no pattern, implying that feature b contributes mainly through weak interactions with

asym. For the feature a , a weak positive relationship is observed. Increasing a leads to slightly higher SHAP values, consistent with a secondary effect on the E_g . The color scale corresponds to asym, and the gradual color transition along the a -axis reflects the correlated variation between a and the asym in the dataset. In contrast, feature b exhibits a mild negative dependence. Larger values of b result in decreasing SHAP contributions. Coloring by asym reveals a gradient, indicating that part of the variation in SHAP originates from its interaction with asym, rather than its direct influence.

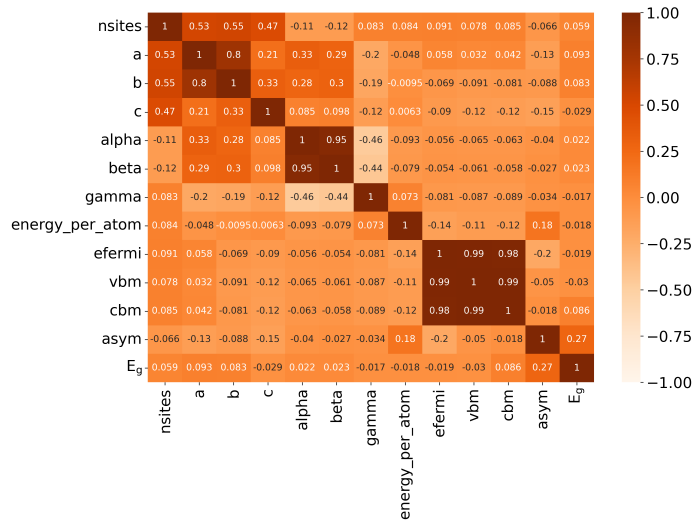
The HQCNN-SE and HQCNN-HEA models show a dependence pattern for asym very similar to the C-NN. The points are colored by b , but the absence of a clear pattern suggests that b interacts only weakly with asym in this model. The feature a again displays a moderate positive correlation with its SHAP values. Point colors represent nsites, and the color distribution indicates that the nsites does not significantly interact with a . The dependence plot for b shows a negative slope, with points colored by asym. The color gradient reflects that the variation in SHAP partially follows changes in asym.

The HQCNN-QI model exhibits nearly identical dependence behavior. The asym plot keeps the same nonlinear pattern, with strong positive SHAP contributions for large asym values. Coloring by b again shows no clear interaction pattern. Feature a continues to display a mild positive relationship with its SHAP values, and the points are colored by the structural parameter c , though no clear pattern emerges. This implies that c does not significantly modulate the contribution of a . The feature b shows a decreasing trend in SHAP contributions with increasing values. Color encodes the c values, and the points remain scattered, suggesting weak interaction between b and c .

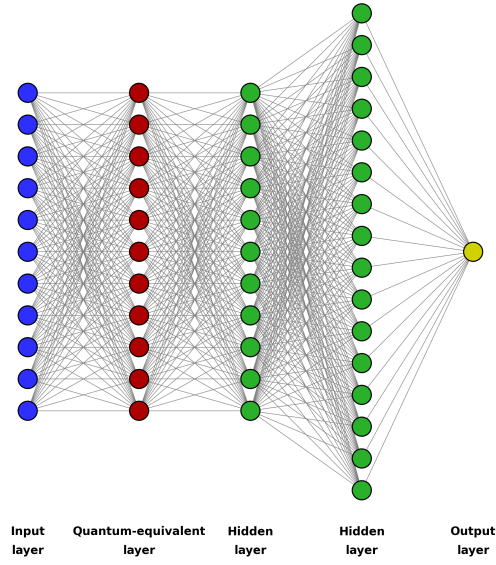
Supplementary Figures



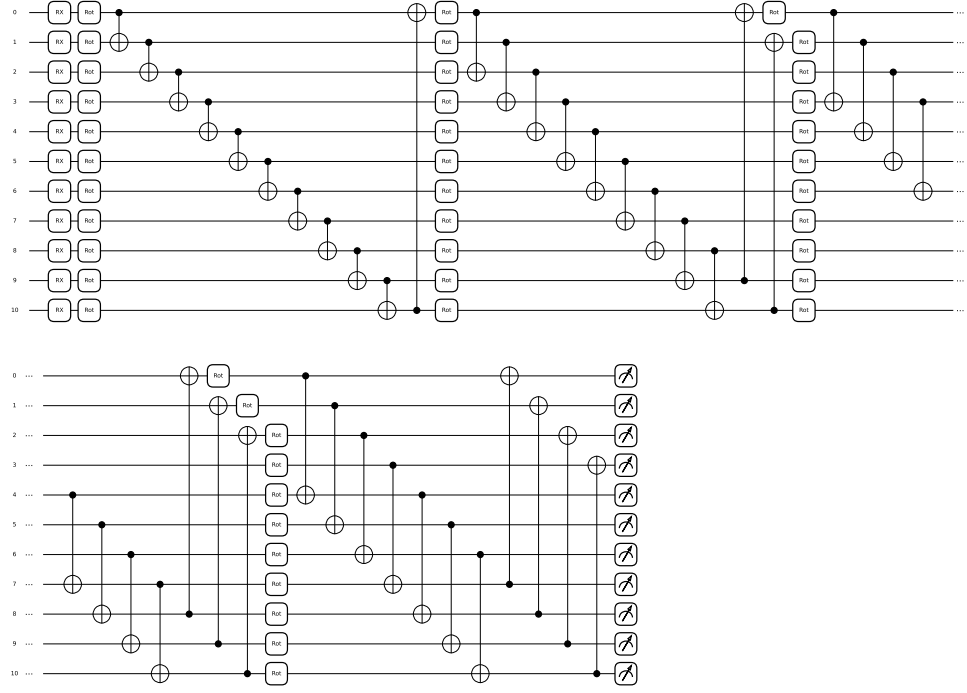
Supplementary Fig. 1: Charge density (left panels) and electron localization function (right panels) of undoped rutile TiO_2 (a, b) and TiO_2 co-doped with one Cr atom and one Sb atom (c, d).



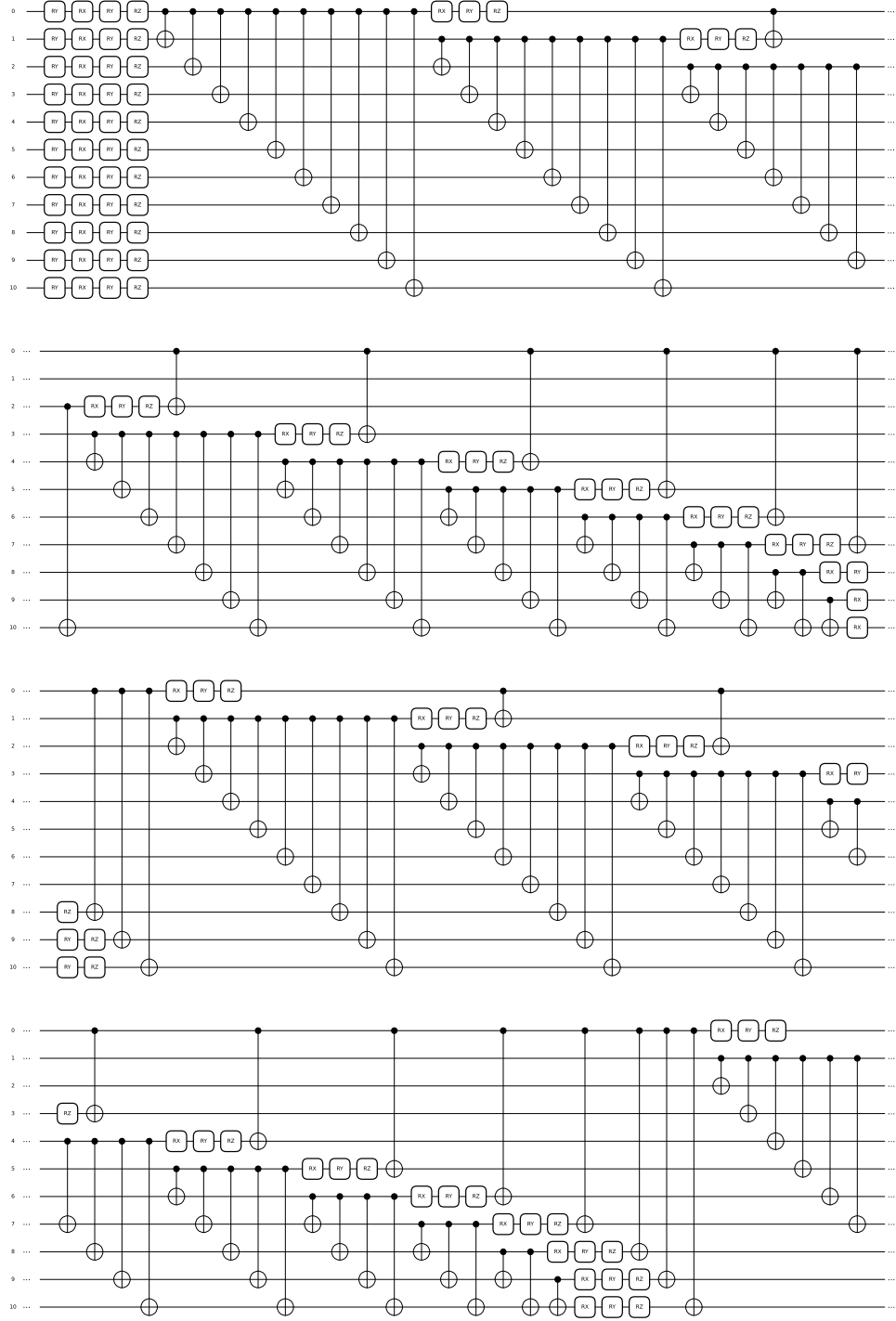
Supplementary Fig. 2: The heat map of the Pearson correlation coefficient matrix for the relation between features and the target of E_g .



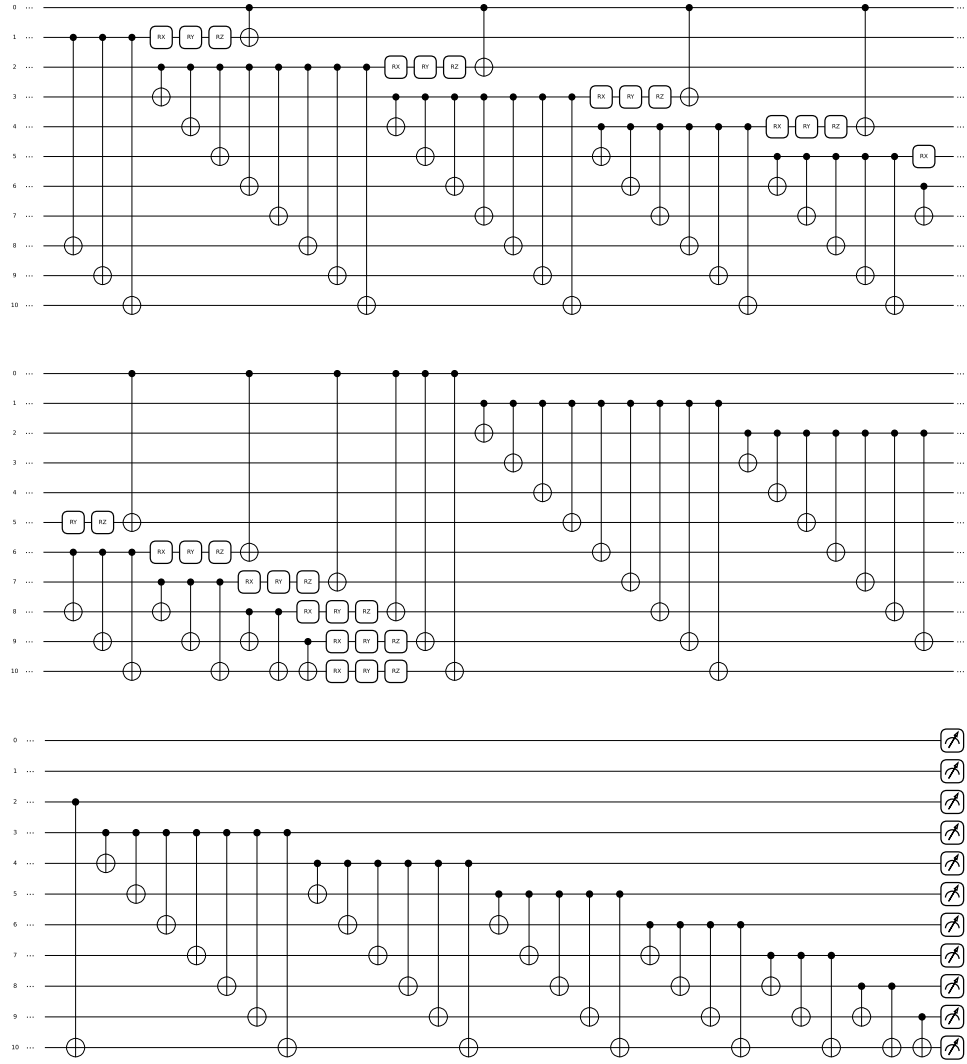
Supplementary Fig. 3: Schematic of the C-NN diagram used to predict the E_g in Scenario 1.



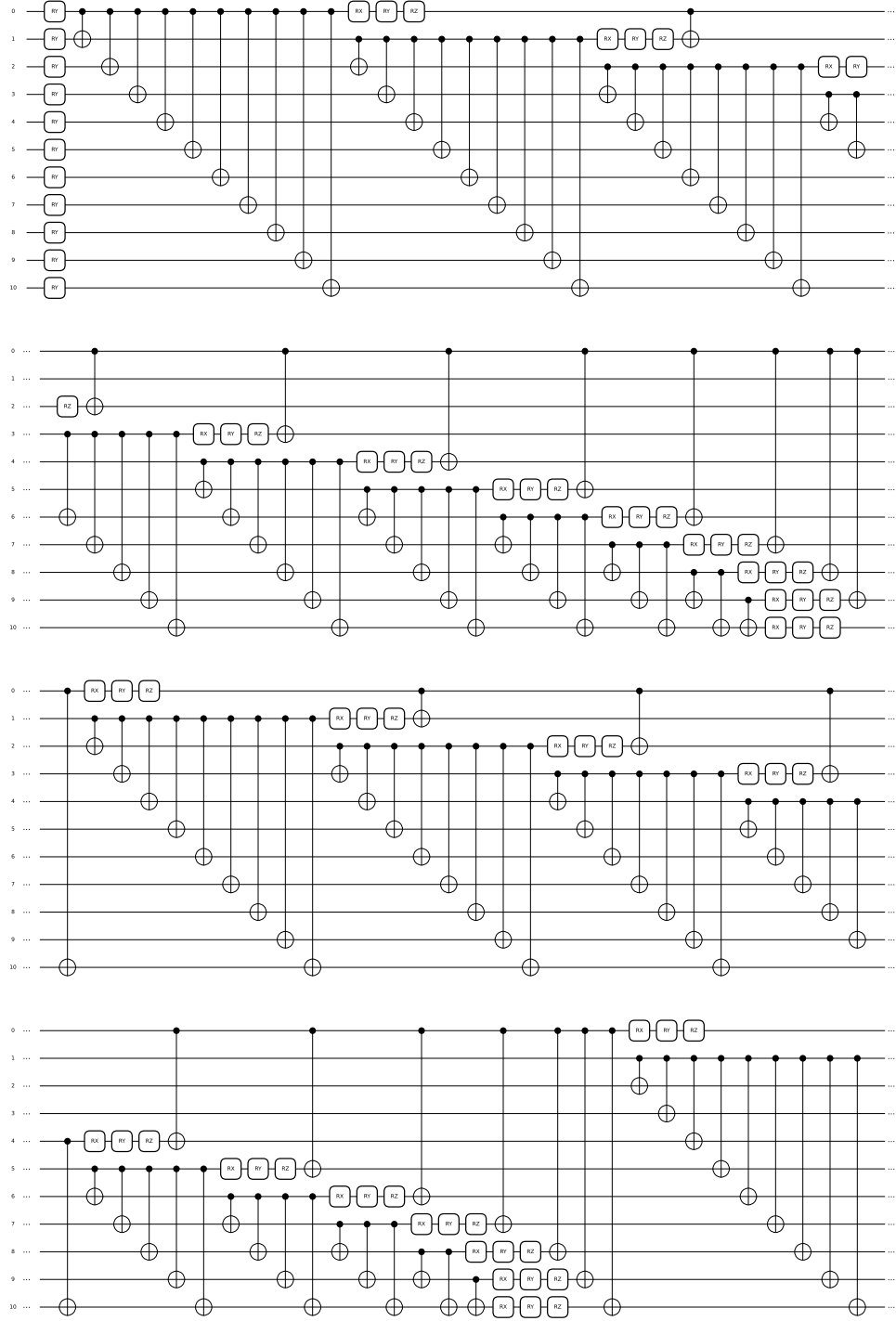
Supplementary Fig. 4: Schematic of the HQCNN-SE quantum circuit showing feature encoding and strongly entangling layers, used to predict the E_g in Scenario 1.



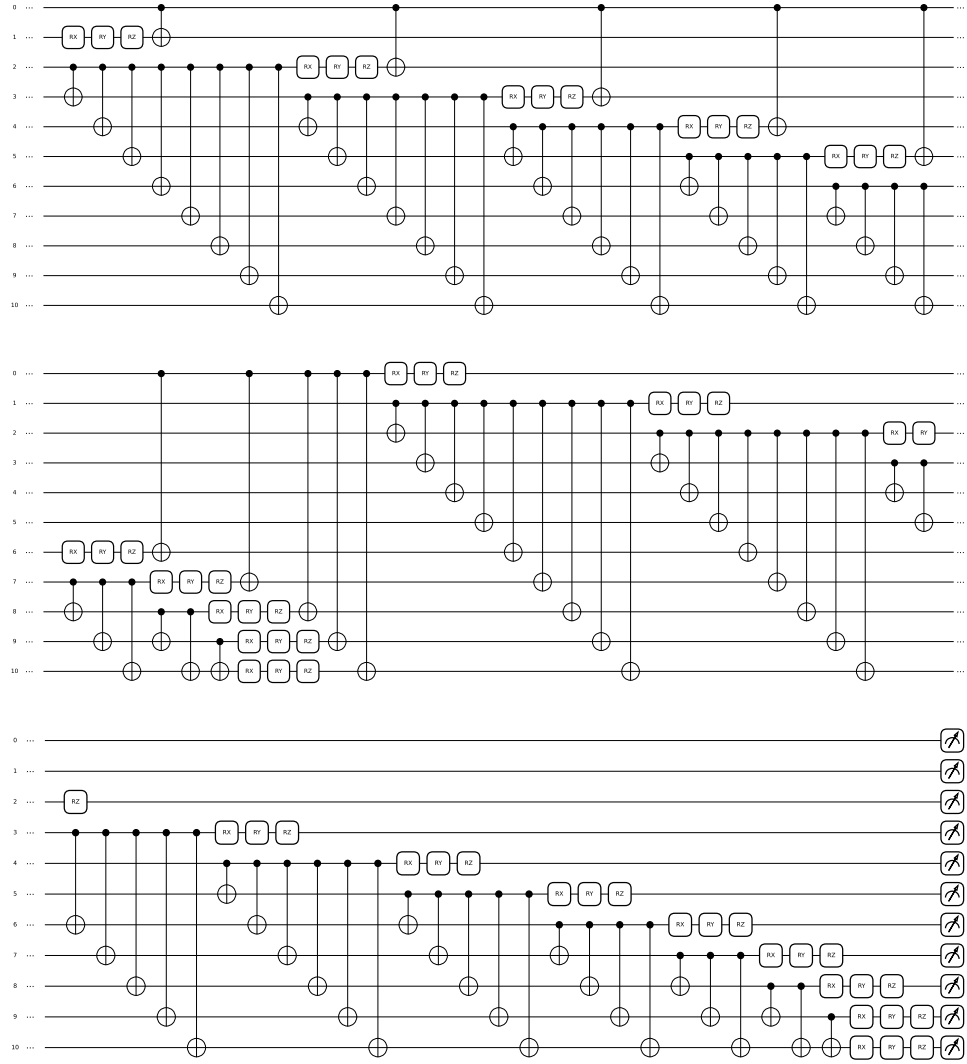
Supplementary Fig. 5: Schematic of the HQCNN-HEA quantum circuit showing feature encoding and hardware-efficient ansatz layers with full entanglement, used to predict the E_g in Scenario 1.



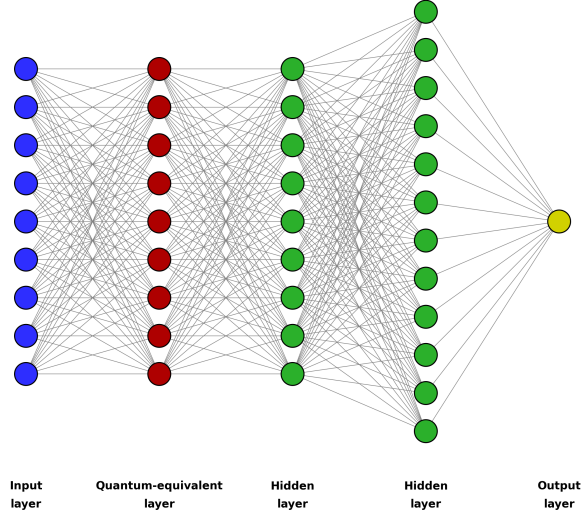
Supplementary Fig. 5: Schematic of the HQCNN-HEA quantum circuit showing feature encoding and hardware-efficient ansatz layers with full entanglement, used to predict the E_g in Scenario 1 (continued).



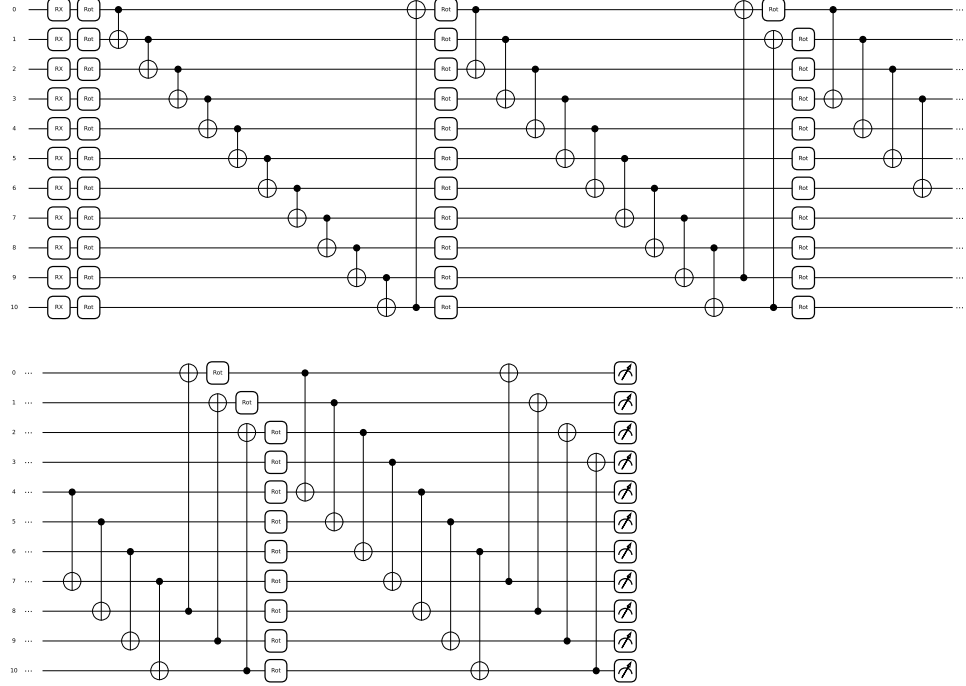
Supplementary Fig. 6: Schematic of the HQCNN-QI quantum circuit showing feature encoding and QAOA-inspired ansatz layers with full entanglement, used to predict the E_g in Scenario 1.



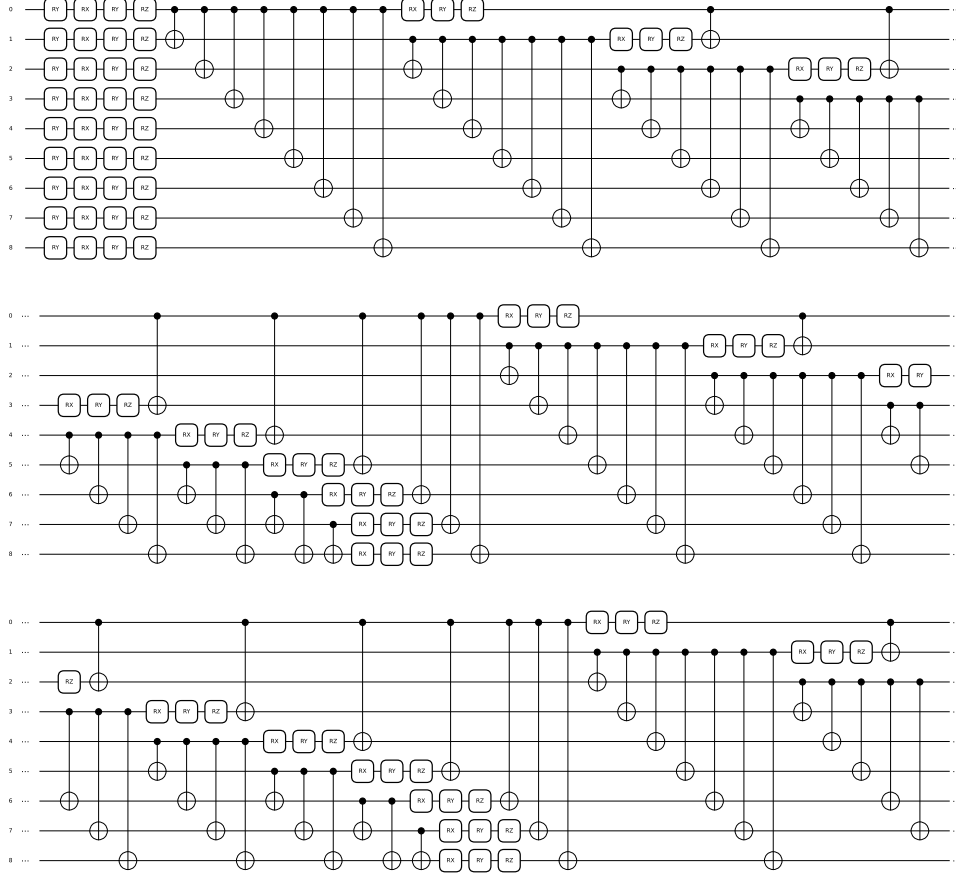
Supplementary Fig. 6: Schematic of the HQCNN-QI quantum circuit showing feature encoding and QAOA-inspired ansatz layers with full entanglement, used to predict the E_g in Scenario 1 (continued).



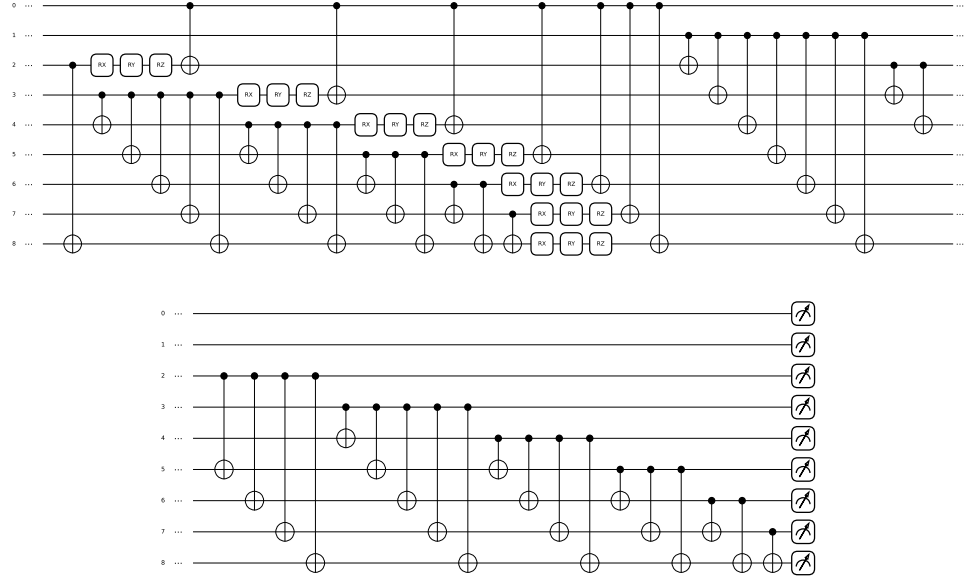
Supplementary Fig. 7: Schematic of the C-NN diagram used to predict the E_g in Scenario 2.



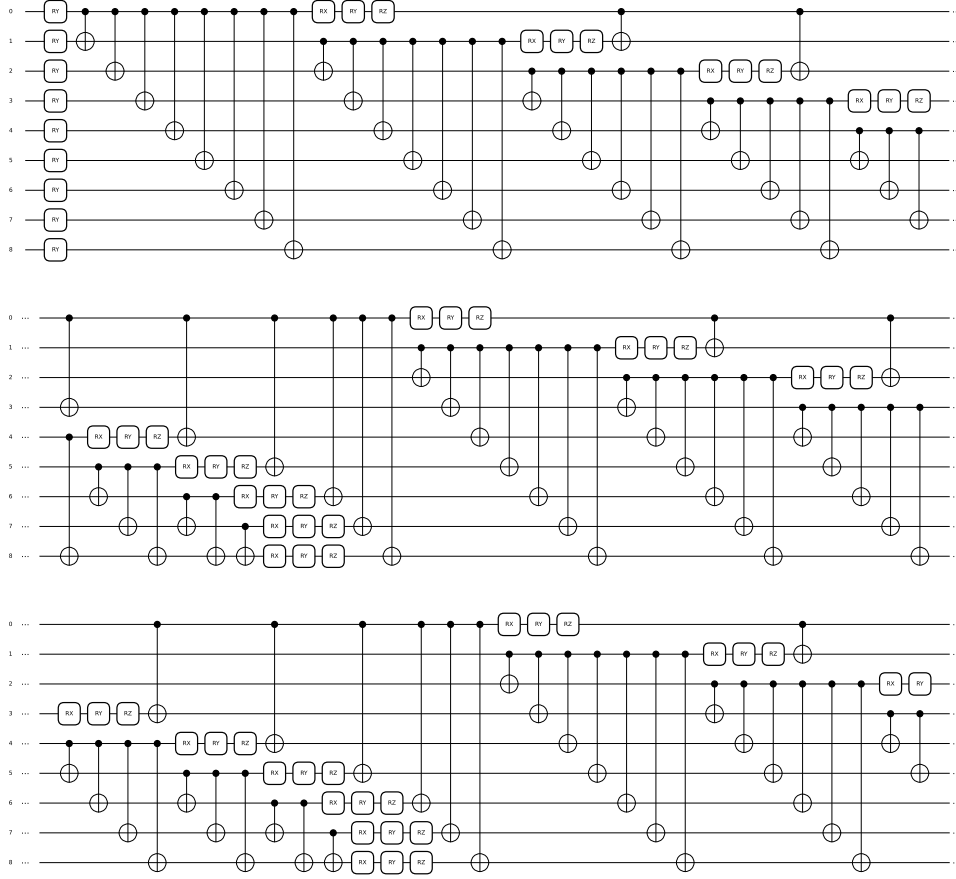
Supplementary Fig. 8: Schematic of the HQCNN-SE quantum circuit showing feature encoding and strongly entangling layers, used to predict the E_g in Scenario 2.



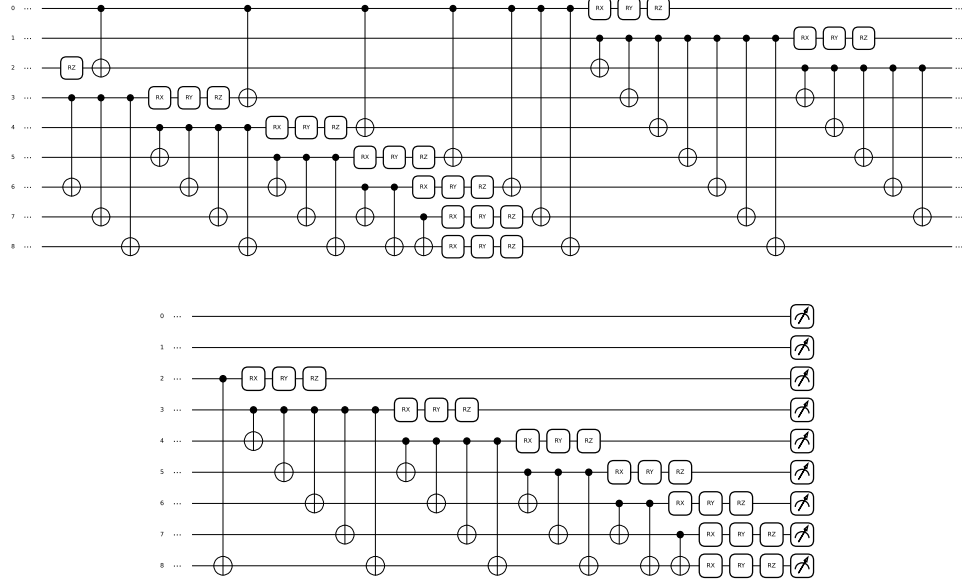
Supplementary Fig. 9: Schematic of the HQCNN-HEA quantum circuit showing feature encoding and hardware-efficient ansatz layers with full entanglement, used to predict the E_g in Scenario 2.



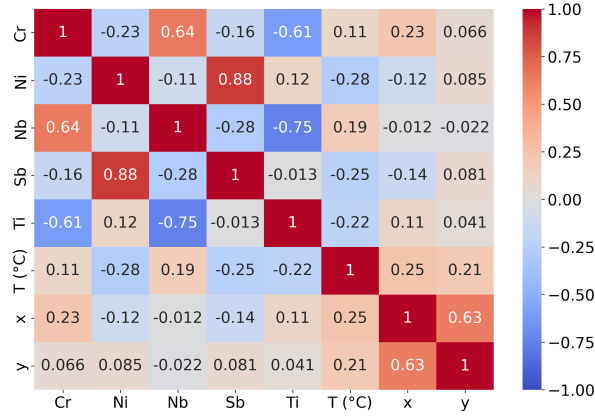
Supplementary Fig. 9: Schematic of the HQCNN-HEA quantum circuit showing feature encoding and hardware-efficient ansatz layers with full entanglement, used to predict the E_g in Scenario 2 (continued).



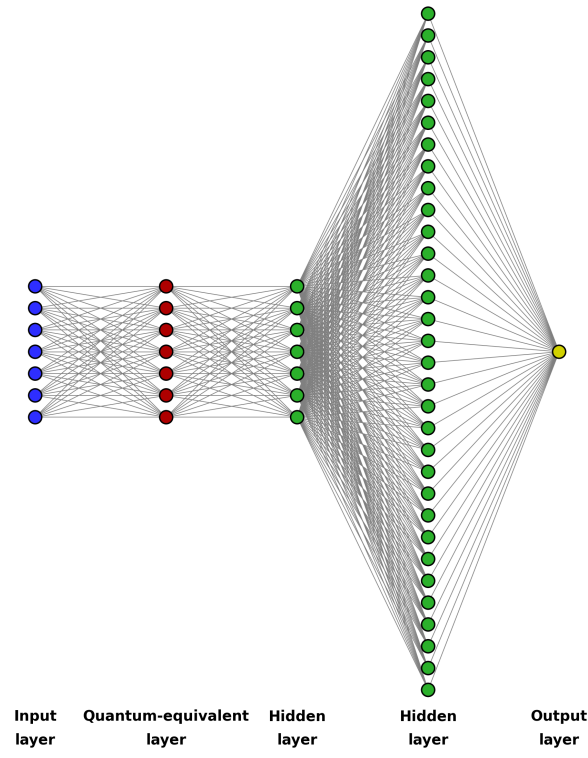
Supplementary Fig. 10: Schematic of the HQCNN-QI quantum circuit showing feature encoding and QAOA-inspired ansatz layers with full entanglement, used to predict the E_g in Scenario 2.



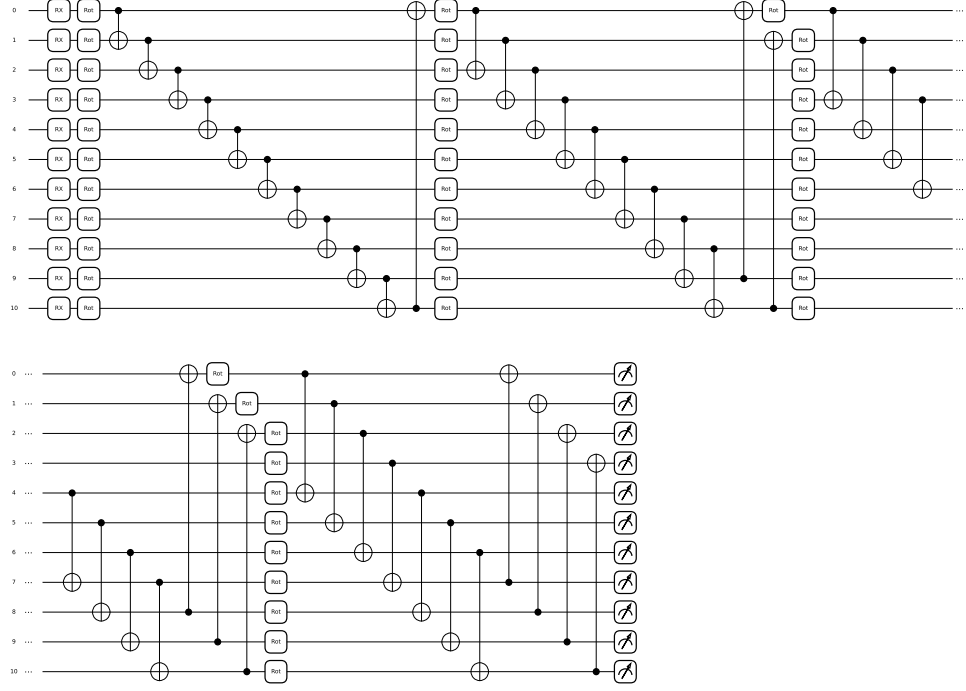
Supplementary Fig. 10: Schematic of the HQCNN-QI quantum circuit showing feature encoding and QAOA-inspired ansatz layers with full entanglement, used to predict the E_g in Scenario 2 (continued).



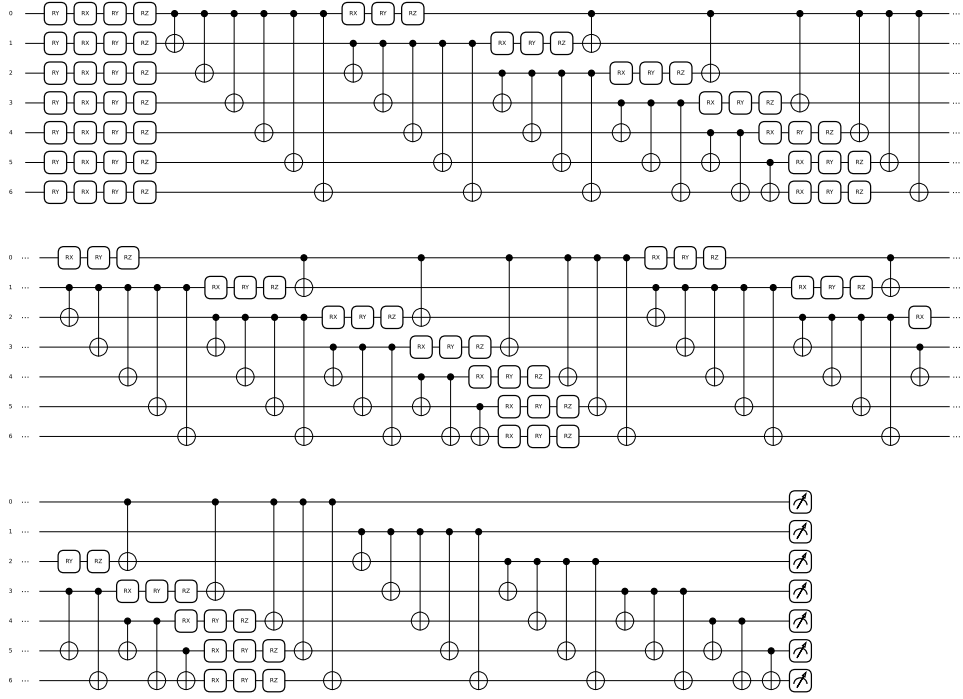
Supplementary Fig. 11: Heat map of the Pearson correlation coefficient matrix for the features and the color coordinates x and y .



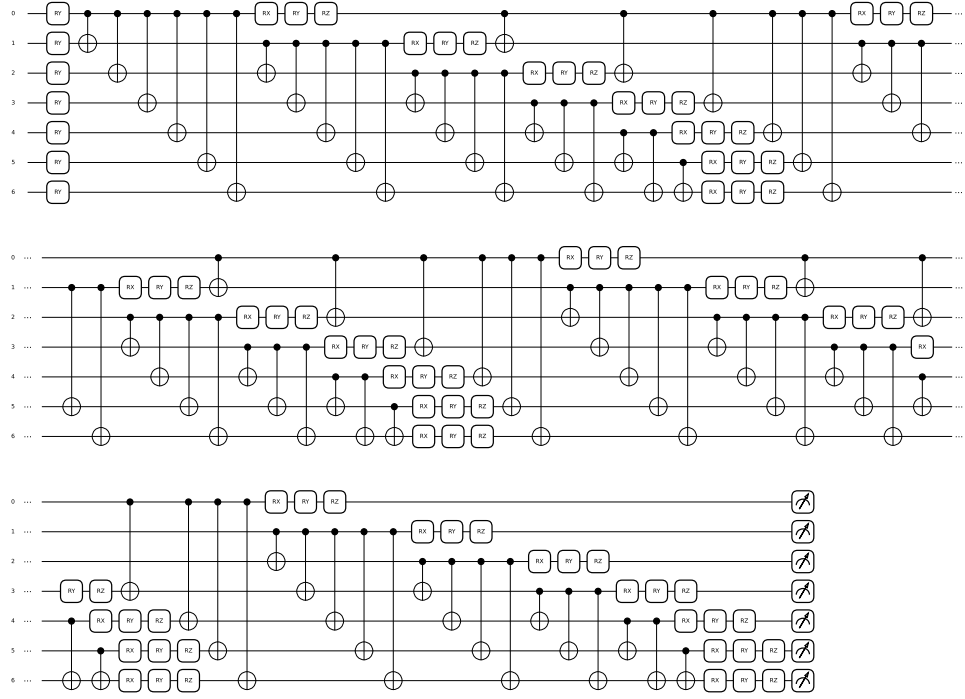
Supplementary Fig. 12: Schematic of the C-NN diagram used for color prediction.



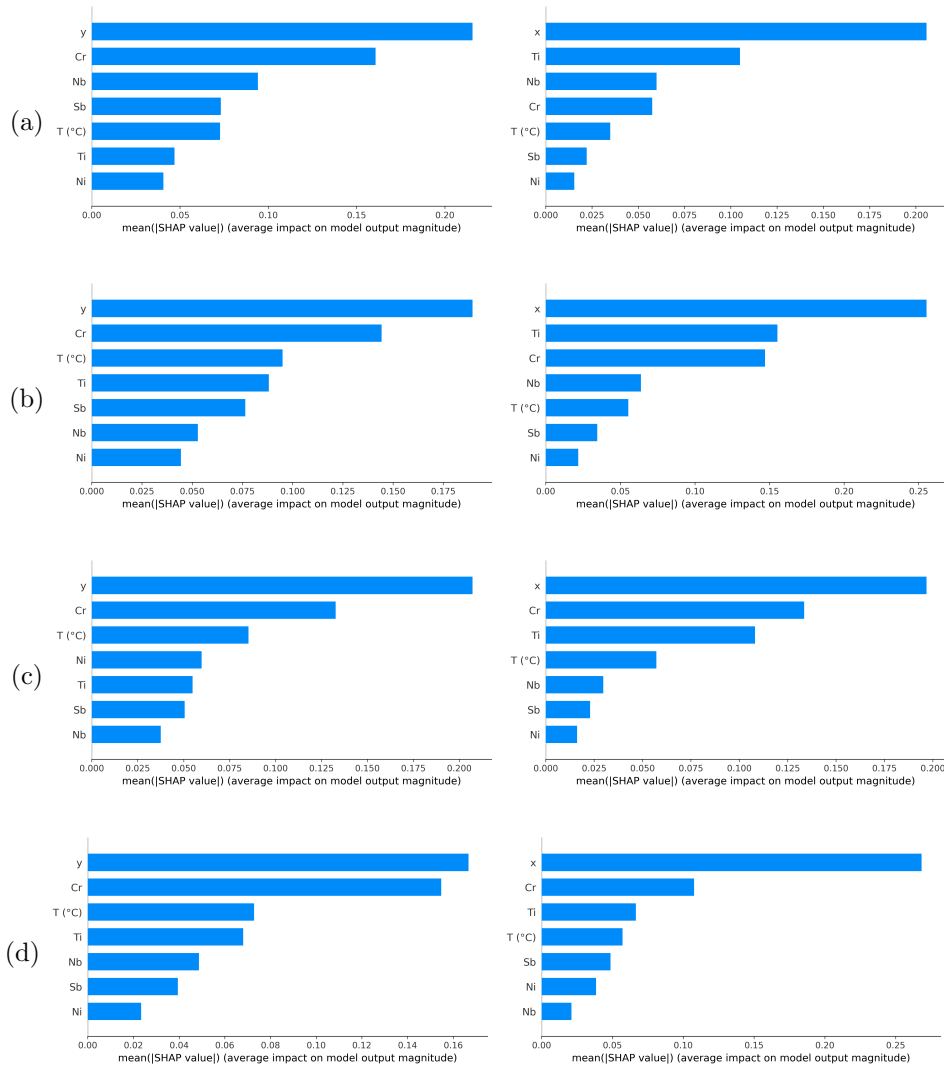
Supplementary Fig. 13: Schematic of the HQCNN-SE quantum circuit, showing feature-encoding and strongly entangling layers for color prediction.



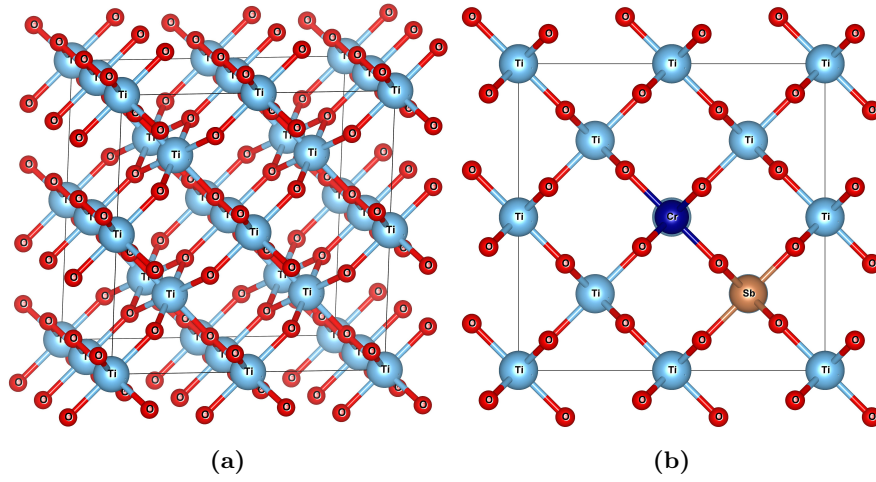
Supplementary Fig. 14: Schematic of the HQCNN-HEA quantum circuit, showing feature-encoding and hardware-efficient ansatz layers with full entanglement for color prediction.



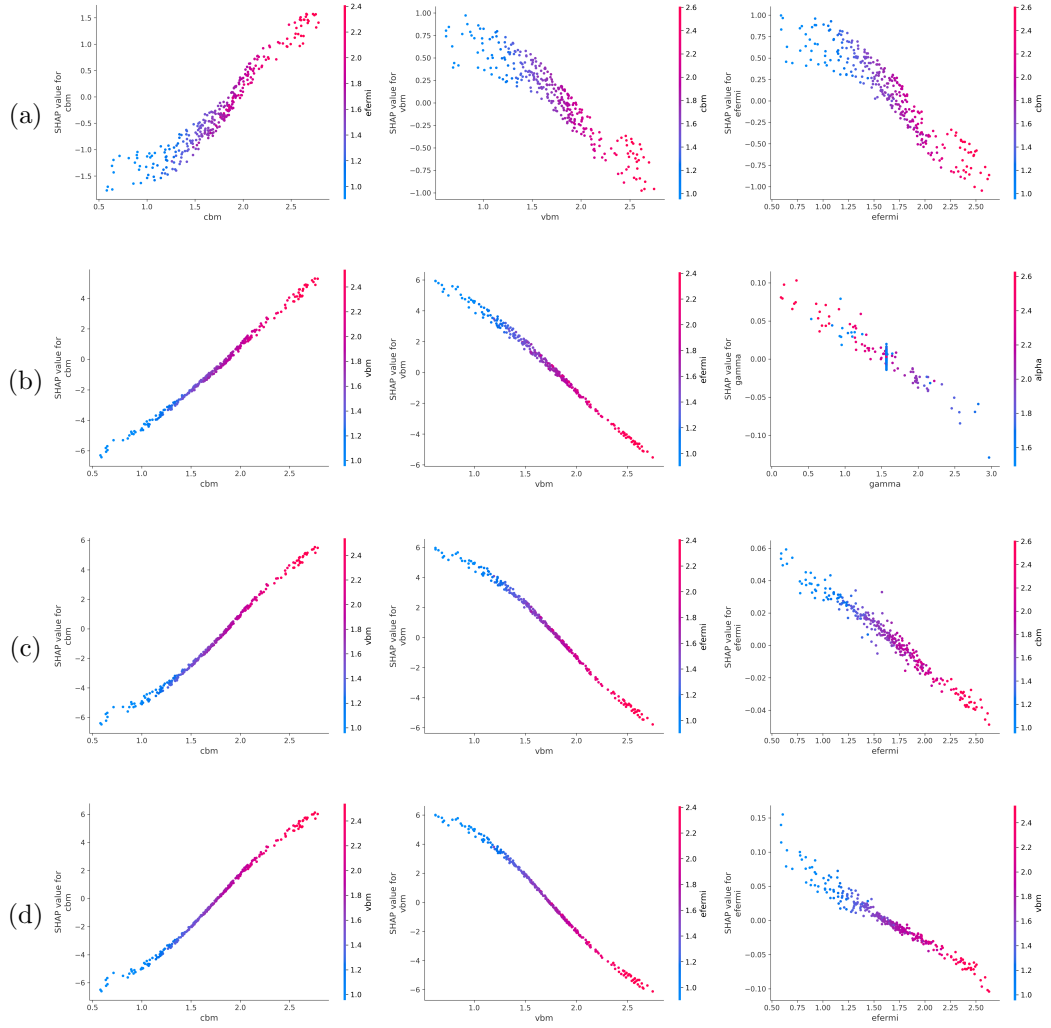
Supplementary Fig. 15: Schematic of the HQCNN-QI quantum circuit, showing feature-encoding and QAOA-inspired ansatz layers with full entanglement for color prediction.



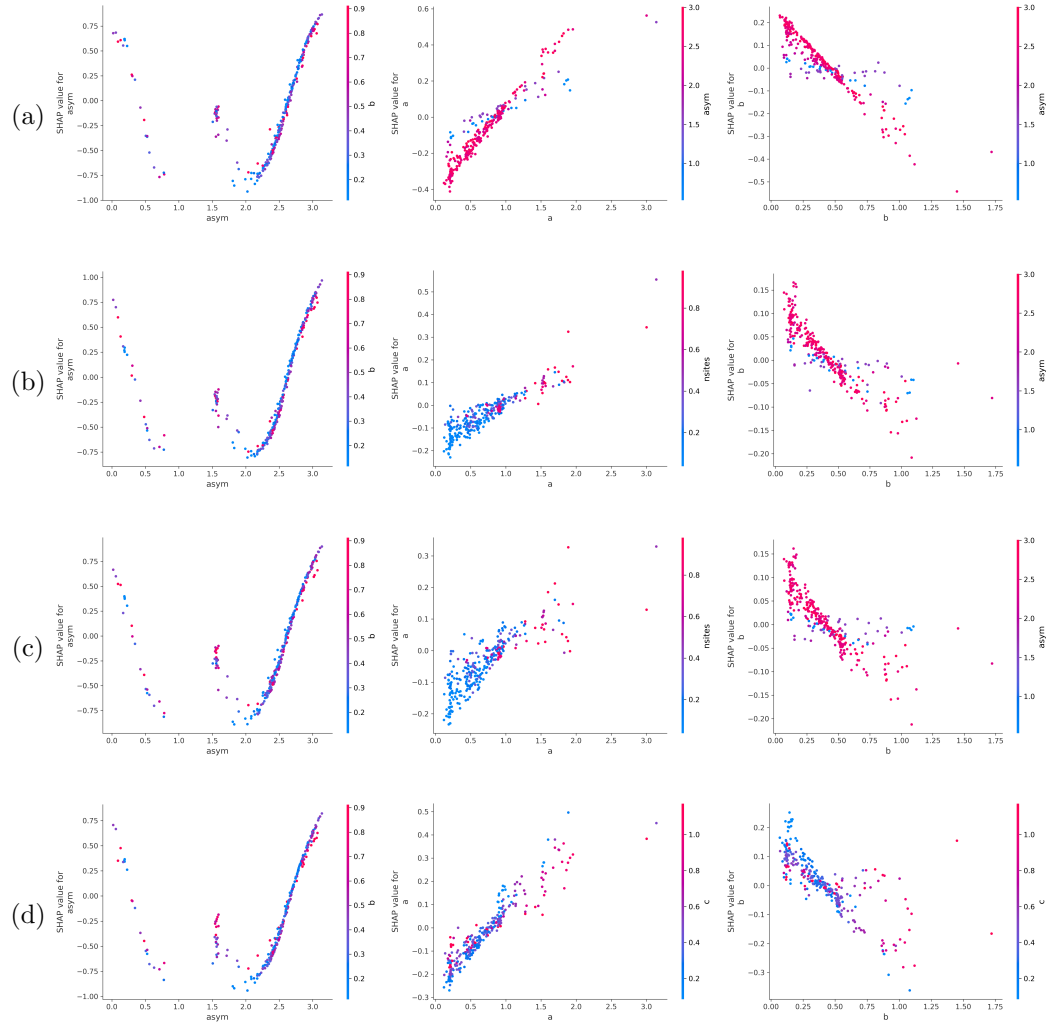
Supplementary Fig. 16: Mean absolute SHAP value bar plots of the C-NN (a), HQCNN-SE (b), HQCNN-HEA (c), and HQCNN-QI (d) models for the prediction of the x (left panels) and y (right panels) colorimetric coordinates.



Supplementary Fig. 17: Crystal structures of the $2 \times 2 \times 2$ rutile TiO_2 supercell: (a) undoped and (b) co-doped with one Cr and one Sb atom.



Supplementary Fig. 18: SHAP dependence plots of the C-NN (a), HQCNN-SE (b), HQCNN-HEA (c), and HQCNN-QI (d) models for E_g prediction in Scenario 1.



Supplementary Fig. 19: SHAP dependence plots of the C-NN (a), HQCNN-SE (b), HQCNN-HEA (c), and HQCNN-QI (d) models for E_g prediction in Scenario 2.

Supplementary Tables

Supplementary Table 1: Evaluation metrics of hybrid quantum–classical models at different variational circuit depths for E_g prediction in Scenario 1.

Model	Variational circuit depth	MAE (eV)	MSE (eV)	R^2
HQCNN-SE	1	0.0588	0.0041	0.9490
	2	0.0499	0.0029	0.9646
	3	0.0582	0.0039	0.9526
	4	0.0295	0.0013	0.9835
HQCNN-HEA	1	0.0239	0.0010	0.9879
	2	0.0253	0.0012	0.9855
	3	0.0190	0.0007	0.9915
	4	0.0244	0.0008	0.9902
HQCNN-QI	1	0.0498	0.0038	0.9531
	2	0.0375	0.0017	0.9787
	3	0.0786	0.0074	0.9086
	4	0.0651	0.0045	0.9450

Supplementary Table 2: Evaluation metrics of all models for E_g prediction in Scenario 1.

Model	MAE (eV)	MSE (eV)	R^2
C-NN	0.0360	0.0015	0.9816
C-NN (cross-validation)	0.0312	0.0016	0.9805
HQCNN-SE	0.0295	0.0013	0.9835
HQCNN-SE (cross-validation)	0.0296	0.0016	0.9807
HQCNN-HEA	0.0244	0.0008	0.9902
HQCNN-QI	0.0651	0.0045	0.9450

Supplementary Table 3: Evaluation metrics of all models for E_g prediction in Scenario 2.

Model	MAE (eV)	MSE (eV)	R^2
C-NN	0.1124	0.0230	0.7177
C-NN (cross-validation)	0.1165	0.0249	0.6899
HQCNN-SE	0.1059	0.0222	0.7273
HQCNN-SE (cross-validation)	0.1112	0.0245	0.6964
HQCNN-HEA	0.0986	0.0200	0.7543
HQCNN-QI	0.1074	0.0206	0.7473

Supplementary Table 4: Evaluation metrics of hybrid quantum–classical models at different variational circuit depths for color prediction (x coordinate).

Model	Variational circuit depth	MAE (eV)	MSE (eV)	R^2
HQCNN-SE	1	0.0186	0.0006	0.6357
	2	0.0159	0.0004	0.7635
	3	0.0157	0.0005	0.7322
	4	0.0156	0.0004	0.7440
HQCNN-HEA	1	0.0166	0.0005	0.7205
	2	0.0199	0.0008	0.5642
	3	0.0169	0.0005	0.6985
	4	0.0155	0.0004	0.7544
HQCNN-QI	1	0.0168	0.0005	0.6925
	2	0.0144	0.0004	0.7739
	3	0.0154	0.0004	0.7647
	4	0.0164	0.0005	0.7009

Supplementary Table 5: Evaluation metrics of all models for predicting the colorimetric coordinates x and y .

Model	Target coordinate	MAE (eV)	MSE (eV)	R^2
C-NN	x	0.0155	0.0004	0.7799
	y	0.0065	0.0001	0.8028
C-NN (cross-validation)	x	0.0150	0.0004	0.7377
	y	0.0075	0.0001	0.8043
HQCNN-SE	x	0.0156	0.0004	0.7440
	y	0.0055	0.0001	0.8731
HQCNN-SE (cross-validation)	x	0.0143	0.0004	0.7542
	y	0.0072	0.0001	0.8311
HQCNN-HEA	x	0.0155	0.0004	0.7544
	y	0.0065	0.0001	0.8217
HQCNN-QI	x	0.0164	0.0005	0.7009
	y	0.0062	0.0001	0.7788

Supplementary Table 6: List of features used (indicated by a checkmark) for E_g prediction in Scenarios 1 and 2.

Symbol	Scenario 1 features	Scenario 2 features	Description
nsites	✓	✓	Atomic sites per unit cell
a	✓	✓	Lattice parameter a
b	✓	✓	Lattice parameter b
c	✓	✓	Lattice parameter c
alpha	✓	✓	Lattice parameter α
beta	✓	✓	Lattice parameter β
gamma	✓	✓	Lattice parameter γ
energy_per_atom	✓	✓	Energy per atom
efermi	✓	✗	Fermi energy
vbm	✓	✗	Valence band maximum
cbm	✓	✗	Conduction band minimum
asym	✗	✓	Fermi level asymmetry relative to bandgap center ($\frac{VBM+CBM}{2} - E_F$)