

Cite this: *Digital Discovery*, 2025, 4, 3351

Deep learning reveals key predictors of thermal conductivity in covalent organic frameworks

Prakash Thakolkaran,^{†a} Yiwen Zheng,^{†b} Yaqi Guo,^a Aniruddh Vashisth^{†*b} and Siddhant Kumar^{†*a}

The thermal conductivity of covalent organic frameworks (COFs), an emerging class of nanoporous polymeric materials, is crucial for many applications, yet the link between their structure and thermal properties remains poorly understood. Analysis of a dataset containing over 2400 COFs reveals that conventional features such as density, pore size, void fraction, and surface area do not reliably predict thermal conductivity. To address this, an attention-based machine learning model was trained, accurately predicting thermal conductivities even for structures outside the training set. The attention mechanism was then utilized to investigate the model's success. The analysis identified dangling molecular branches as a key predictor of thermal conductivity, leading us to define the dangling mass ratio (DMR), a descriptor that quantifies the fraction of atomic mass in dangling branches relative to the total COF mass. Feature importance assessments on regression models confirm the significance of DMR in predicting thermal conductivity. These findings indicate that COFs with dangling functional groups exhibit lower thermal transfer capabilities. Molecular dynamics simulations support this observation, revealing significant mismatches in the vibrational density of states due to the presence of dangling branches.

Received 27th March 2025
Accepted 21st September 2025

DOI: 10.1039/d5dd00126a

rsc.li/digitaldiscovery

Introduction

Covalent organic frameworks^{1,2} (COFs) are an emerging class of nanoporous polymeric materials. Compared to metal-organic frameworks (MOFs)^{3,4} and zeolites,⁵ the crystalline backbone of COFs is composed of organic building blocks – known as knots and linkers – connected by strong covalent bonds, thereby offering higher stability.⁶ The crystalline nature of COFs, along with their high surface areas, tunable pore sizes, and functionalizable organic linkers, make them exceptionally suited for a wide range of promising applications. These applications include (but are not limited to) photoconductivity,^{7,8} chemosensing,^{9,10} catalysis,¹¹ drug delivery,^{12,13} thermoelectrics,^{14,15} semiconductors,¹⁶ and gas storage and separation.^{17,18}

A key property dictating the applications of COFs is thermal transfer. For example, low thermal conductivity is desired for thermoelectrics to maintain large internal thermal gradients and increase efficiency.¹⁹ On the other hand, a high thermal conductivity is desired for gas adsorption and separation where efficient thermal transfer is important for the longevity and stability of the nanoporous membranes.²⁰ COFs are promising

nanoporous candidates for such applications as the ability to design their crystalline topology (*e.g.*, lattice type, pore size) and chemistry *via* the choice of knots and linkers opens up a large and diverse space of thermal conductivities.^{1,2,21–23} Therefore, it is of great interest to gain a deeper understanding of the thermal transfer mechanism and to obtain strong structure–property trends, which in turn enables the application-specific design of COFs.

Two approaches can be used to elucidate the thermal structure–property relationships of COFs: experimental trial-and-error methods which involve synthesis and characterization,^{24,25} and computational high-throughput screening which relies on first-principles calculations.^{26–28} With COFs offering practically an unlimited design space, an experimental trial-and-error approach to screen new COF candidates (including developing new synthesis routes for each candidate) is prohibitively inefficient. Moreover, experimentally synthesized COFs will always contain crystalline defects,^{29,30} which greatly influence the thermal conductivity. This makes it difficult to relate the thermal transfer mechanisms to the geometrical and chemical make-up of a COF structure. In contrast, molecular dynamics (MD) simulations³¹ offer a high-throughput virtual screening alternative to lab-based experiments. However, despite recent advances in computing hardware, even virtual screening can be prohibitively inefficient for a large design space. This is highlighted by the example that generating a dataset of just 2471 two-dimensional COFs and their thermal

^aDepartment of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands. E-mail: Sid.Kumar@tudelft.nl^bDepartment of Mechanical Engineering, University of Washington, Seattle, WA, USA. E-mail: Vashisth@uw.edu

† These authors contributed equally to this work.



conductivity for this study required 1.3 million CPU-hours in cloud computing (translating into approximately four months and \$70 000 in time and cost, respectively). This underscores the need for efficient structure–property maps that are accurate, generalizable, and interpretable.

Recently, Islamov *et al.*³² performed a high-throughput screening study of over 10 000 MOF structures and their thermal conductivity (κ) using MD. They found that the majority of the MOFs possess $\kappa < 1 \text{ W m}^{-1} \text{ K}^{-1}$ with a few exceptions possessing ultra-high thermal conductivity ($\kappa > 10 \text{ W m}^{-1} \text{ K}^{-1}$). While COFs are generally more thermally stable (owing to the strong covalent bonding of the building blocks), a recent study by Thakur *et al.*³³ on a high-throughput screening study on over 10 000 COFs demonstrated that, generally, COFs also possess a similar range of κ , with the majority of the structures exhibiting $\kappa < 1 \text{ W m}^{-1} \text{ K}^{-1}$. Interestingly, MOFs and COFs possess similar structure–property trends. (i) Increasing the pore size is mostly sufficient to achieve a low thermal conductivity. (ii) Additionally, one can increase the void fraction, the surface area, or include heavy atoms to further decrease thermal conductivity. (iii) However, to increase the thermal conductivity, there are many factors that need to align. Islamov *et al.*³² demonstrated that while lowering the pore size and increasing the density results in a higher ceiling of attainable thermal conductivity, other factors, such as the topology, mass-mismatch, and linker length also need to be accounted for. Thakur *et al.*³³ came to similar conclusions alongside the observation that aligning the polymeric chains in the structure to the heat flow direction raises the ceiling of attainable thermal conductivity. However, these hand-crafted guidelines and ad hoc correlations are insufficient for developing an accurate and generalizable predictive model that captures the thermal conductivity structure–property relationships of COFs. Our study shows that no combination of commonly used descriptors consistently predicts thermal conductivity. Consequently, a definitive method for identifying COFs with tailor-made thermal properties for various applications remains elusive.

To address this knowledge gap, we turn to machine learning (ML) with an emphasis on interpretability and explainability. Previous efforts in ML-assisted design for thermally conductive organic materials include the discovery of polymers with high κ through a hierarchical feature selection process,³⁴ a reinforcement learning approach using SMILES-based representations,³⁵ and methods leveraging molecular fingerprints as input features, such as fine-tuning a pre-trained regression model followed by screening,³⁶ training a neural network,³⁷ and employing an active learning approach.³⁸ Hu *et al.*³⁹ provide a review of recent efforts and outlook on ML for thermally conductive organic materials. For the ML modeling of structure–property maps of porous polymers (including but not limited to COFs) there are primarily two routes:

- high interpretability, limited accuracy: using a hand-crafted and pre-extracted high-level of the crystalline network (*e.g.*, pore size, density, surface area, atomic composition) as input to classical regression algorithms;^{40–42}
- high accuracy, limited interpretability: using information-rich but raw graph representation of the crystalline network

(where nodes denote atoms with features such as position and atom type, and edges denote bonds with the bond order as the feature) as input to end-to-end deep learning-based regression algorithms.^{43–46}

Here, we bridge the two routes to develop an accurate predictive model using deep learning while also combining interpretability insights from deep learning with prior knowledge and descriptors to explain the thermal conductivity structure–property relationships of COFs (see Fig. 1 for an overview).

In the following, we first conduct a large-scale data analysis of the thermal conductivity structure–property relations of COFs and identify the deficiencies in commonly used descriptors. Next, we provide a deep learning model that accurately predicts the thermal conductivity of COFs. Subsequently, an analysis of the attention scores of the deep learning model uncovers the presence of dangling atoms in the crystalline network of COFs as a strong and so-far missing key predictor for thermal conductivity. Additional physics-based analysis sheds light on the important role of dangling atoms in lowering the thermal conductivity of COFs by disrupting heat transfer pathways. We close by utilizing the ML model for efficient high-throughput screening and identifying COFs with extreme thermal conductivities.

Results

High-throughput data analysis

We create a labelled dataset of COFs and their thermal conductivities by selecting 2471 two-dimensional COFs from the unlabelled dataset of Mercado *et al.*⁴⁷ The COF structures in this design space are made up of one type of nodal linker (also referred to as a knot) and one type of connecting linker. The sampled subset consists of COFs with 104 different linkers arranged in 25 different topologies (*e.g.*, honeycomb, kagome, square lattice, *etc.*), and bonded by four different types of covalent linkages (*i.e.*, carbon–carbon, amine, amide, and imine). We conduct non-equilibrium molecular dynamics (NEMD) simulations to calculate thermal conductivities in two orthogonal in-plane directions (see SI Section 1.1 for more details). The resulting dataset shows a strong skew toward low thermal conductivity values, with the majority of COFs exhibiting $\kappa < 1 \text{ W m}^{-1} \text{ K}^{-1}$. In SI Section 1.2, we show that there is minimal in-plane anisotropy in thermal conductivity across the dataset. Therefore, we utilize average thermal conductivity in both directions, denoted as κ , as the quantity of interest to explore the structure–property relationships of COFs. In the following and in Fig. 2, we present the observations from data analysis of the key descriptors of COFs and thermal conductivities in our labelled dataset. The distribution of all features and their correlations are detailed in Section 1.3 of the SI.

Density. Fig. 2a illustrates the distribution of κ vs. density. There is a large spread of densities within the dataset and the majority of the COFs possess a density of 0.45–0.55 g cm^{-3} . Moreover, in accordance with previous findings,³³ we observe an increasing trend in κ with increasing density. The Pearson correlation coefficient between κ and density is $r = 0.594$,



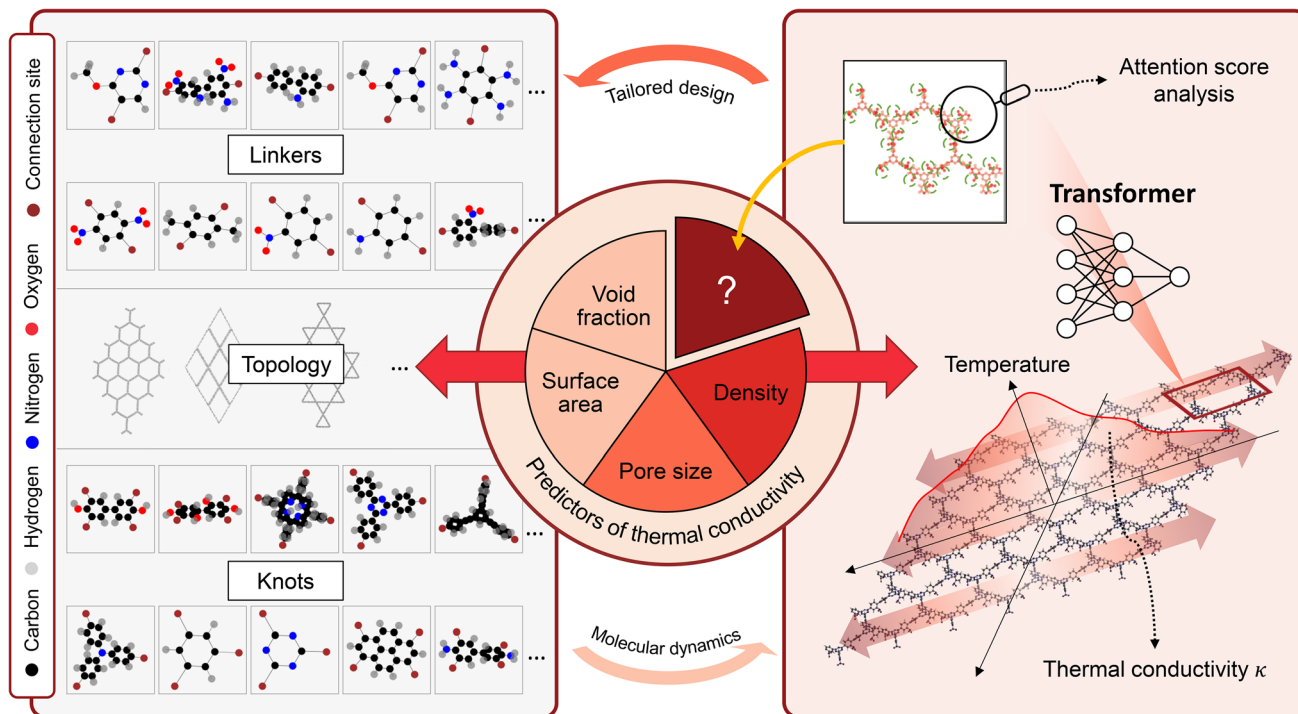


Fig. 1 Overview of this study. A COF structure consists of molecular substructures called linkers and knots arranged in a periodic pattern described by the topology. Through molecular dynamics, we construct a large dataset of 2471 COFs (with diverse linkers, knots, and topologies) and their corresponding thermal conductivities. We demonstrate that conventional descriptors, such as density, pore size, void fraction, and surface area, fail to predict the thermal conductivity reliably. We employ a machine learning framework based on an attention mechanism and transformer architecture to uncover a novel predictor, thus enhancing our understanding of the structure–property relationship of thermal conductivity of COFs.

suggesting a moderate positive relationship. However, increasing the density increases the ceiling of attainable κ , but does not guarantee a high value.

Pore size. The pore size of any COF is represented by the largest pore diameter (LPD). Fig. 2b shows that the LPD in the dataset is spread over a large range between 5 Å and 80 Å and a concentration of structures in the 15–25 Å region. Moreover, there is an inverse relationship between the pore size and the thermal conductivity, as previously discovered by Freitas *et al.*,⁴⁸ where the range of achievable κ decreases significantly with increasing pore sizes. This is evidenced by the correlation coefficient $r = -0.508$, indicating a moderate negative correlation. Here again, a small pore size does not guarantee a high thermal conductivity. On the other hand, the pore size can be taken as the sole design parameter to effectively reduce the thermal conductivity range (as mentioned in ref. 32 and 33).

Void fraction. Fig. 2c shows that the distribution of the void fractions of COFs across the dataset is within the 0.5–0.96 range, with the biggest share of COFs possessing a void fraction of around 0.85 and higher. Furthermore, a high void fraction of around 0.85 delivers both the largest values and range of attainable κ within our dataset. This finding is counterintuitive. Void fraction is typically inversely correlated with density (see SI Section 1.3). Although low density is generally associated with a low κ value, as shown in Fig. 2a, the unexpected observation is that a high void fraction (which corresponds to low density)

surprisingly results in a high κ value, as also depicted in Fig. 2c. Beyond a favorable range of void fractions, no distinct trend emerges in the relationship with κ , as the correlation coefficient $r = -0.213$ suggests only a weak negative relationship.

Surface area. Gravimetric surface area (GSA), *i.e.*, surface area per unit mass, varies highly across all the COF structures (see Fig. 2d) with a larger cluster of points in the 7000–8000 m² g⁻¹ range. We also observe that around an intermediate GSA of 7000 m² g⁻¹, we have the widest spread and highest values of κ . The correlation coefficient of $r = -0.147$ indicates a very weak negative relationship, suggesting that GSA is not a significant predictor.

From the above initial analysis, we notice that the geometrical descriptors used here are insufficient to provide solid trends with respect to thermal conductivity. For example, with all the optimal descriptor values, *i.e.*, a large density, a low pore size, a large void fraction, and an intermediate GSA, it is still not guaranteed that we will obtain a COF with a high κ . To elucidate this further, we choose four COFs – denoted by (i)–(iv) – and compare them in Fig. 2a–d.

COF (i) has the highest thermal conductivity in the dataset with $\kappa = 4.025$ W m⁻¹ K⁻¹. The structure has a relatively small pore size of 13.71 Å (see Fig. 2a), a medium density of 0.928 g cm⁻³ (see Fig. 2b), a relatively high void fraction of 0.844 (see Fig. 2c), and an intermediate surface area of 7381 m² g⁻¹. In contrast, the second COF (ii) possesses a similar pore size of



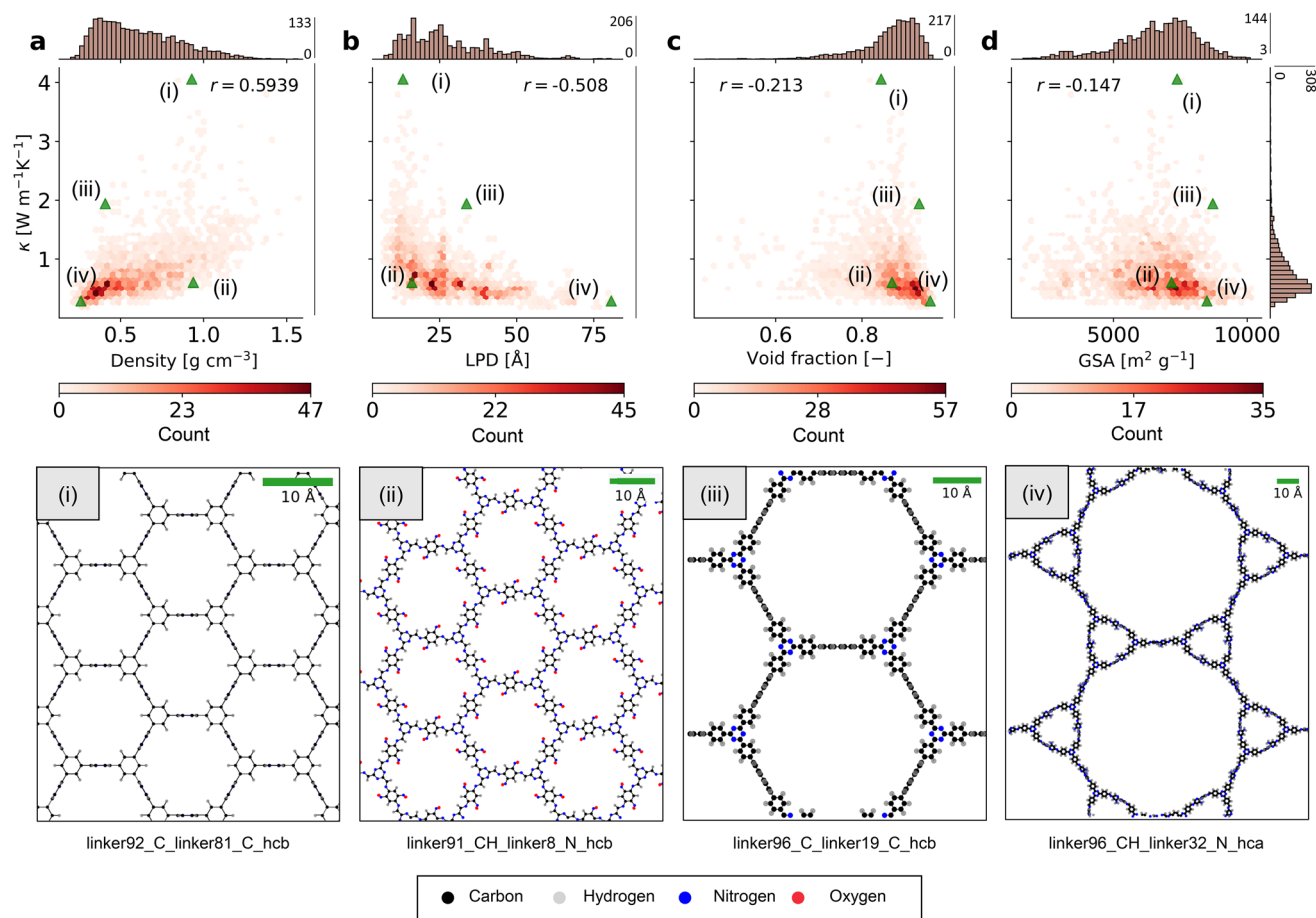


Fig. 2 Distribution of (a) κ versus density (with the color indicating data count per bin), (b) κ versus largest pore diameter (LPD), (c) κ versus void fraction, and (d) κ versus gravimetric surface area (GSA). Also shown are four COF structures (i–iv) with contrasting properties, marked by green triangles in the plots above. The r -value indicates the Pearson correlation coefficient.

9.88 Å, a similar density of 0.938 g cm⁻³, a similar void fraction of 0.869, and a similar surface area of 7171 m² g⁻¹, but exhibits thermal transfer capabilities almost seven times lower than that of (i) with $\kappa = 0.603$ W m⁻¹ K⁻¹.

The third COF (iii) has a pore size of 33.73 Å in the intermediate-to-high range, a low density at 0.408 g cm⁻³, a high void fraction of 0.932, and a slightly larger surface area of 8741 m² g⁻¹. Following the trends described in the literature,^{33,48} the geometrical descriptors are not optimal for a high κ . Nonetheless, COF (iii) has rather high $\kappa = 1.96$ W m⁻¹ K⁻¹. Lastly, COF (iv) has an exceptionally high pore size of 80.68 Å and a low density of 0.262 g cm⁻³ but high void fraction of 0.958. Due to the extreme geometry, the COF (iv) has a very low thermal conductivity of $\kappa = 0.289$ W m⁻¹ K⁻¹.

Furthermore, we trained classical ensemble regression models using these descriptors to predict thermal conductivity (detailed results provided in SI Section 2) and observed that the models yield poor prediction accuracy. In summary, the correlation analysis with the previously introduced descriptors and the presented examples illustrate the fact that the thermal conductivity structure–property relationships are complex and call for further examination.

Predicting thermal conductivity using deep learning

Our starting point is the deep learning-based Porous Material Transformer (PMTransformer) developed by Park *et al.*⁴⁶ The PMTransformer is a multi-modal transformer⁴⁹ model (see Fig. 3 for a schematic) designed for universal transfer learning for predicting the properties of porous materials, including COFs, MOFs, and zeolites. It receives two distinct sets of inputs: local and global features. Local features, which reflect the chemistry of building blocks and bonds, are derived from a crystal graph convolutional neural network⁵⁰ (CGCNN). The CGCNN directly processes the crystal graph, where atoms and bonds are nodes and edges, respectively, and generates embeddings that capture essential chemical details. These embeddings serve as local features in the PMTransformer. Conversely, global features describe crystalline characteristics, including topological and geometric descriptors such as pore size and surface area. These features are derived from three-dimensional energy grids, which are generated by calculating the interaction energy between the material structure and a methane gas molecule at each grid point using the package GRIDAY.⁵¹ Similar to images in Vision Transformers,⁵² the energy grids are divided into patches and flattened through



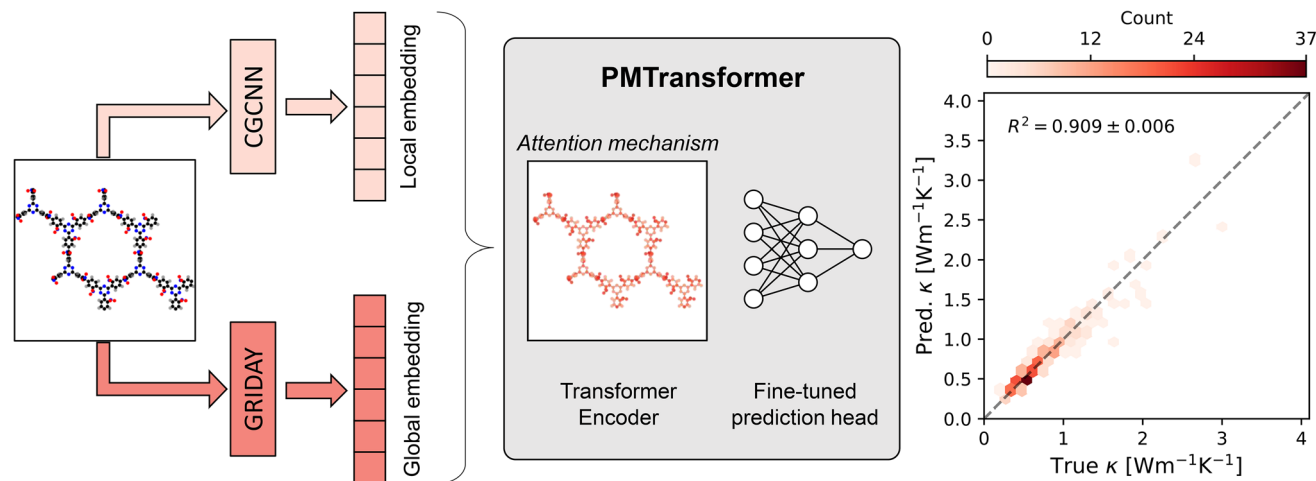


Fig. 3 Schematic of the PMTransformer model. A sample COF structure is shown on the left. A crystal graph convolutional neural network computes local embeddings of the COF graph, while GRIDAY computes a three-dimensional energy grid of the structure, which becomes the global embeddings. These embeddings are combined and input into the PMTransformer (which includes a transformer encoder with an attention mechanism and a prediction head) to predict κ . On the right, a parity plot compares κ predictions with ground truth values. We report the mean and standard deviation of the goodness-of-fit across five random seeds. The random seeds affect the initialization of the prediction head, while the rest of the model is initialized from the pre-trained PMTransformer weights. The dashed line represents the ideal line with zero intercept and unit slope.

linear projections. Eventually, both local and global embeddings are input into a transformer encoder, where the attention mechanism⁴⁹ helps the model focus on the most relevant parts of the molecular structure for predicting (in our case) the thermal conductivity.

To enable universal transfer learning, the transformer encoder in PMTransformer is pre-trained on an extremely large dataset of 1.9 million hypothetical porous materials to predict easily obtainable yet essential properties such as topology, void fraction, and building block prediction. This approach ensures the encoder captures critical information necessary for accurately predicting other, more complex properties in downstream tasks with much smaller datasets. For more details regarding the pre-training of PMTransformer, refer to ref. 46.

In this study, we leveraged the pre-trained transformer encoder of the PMTransformer model to perform transfer learning for the prediction of the thermal conductivity of COFs. We fine-tune the transformer model using the pre-trained weights as a starting point and jointly train a prediction head based on a multi-layer perceptron architecture to map the output of the transformer encoder to the thermal conductivity of COFs. We used the mean squared error (MSE) as a loss function for training. The dataset of 2471 COFs is randomly split into two subsets: 90% of the data is used for training (out of which 10% for validation), and the remaining data is used for testing. We repeat the training across five different random seeds, where each seed initializes the prediction head randomly while keeping the transformer initialized with the same pre-trained weights. Additional details of the training protocol and an ablation study are provided in SI Section 3.

We demonstrate the model's prediction accuracy on the test dataset in Fig. 3, where it achieves a goodness-of-fit R^2 of 0.909 ± 0.006 and a mean absolute error (MAE) of $0.075 \text{ W m}^{-1} \text{ K}^{-1}$

for predicting κ . The model's strong performance suggests that there may be gaps in the feature sets typically used to describe COFs. It also highlights the possibility that additional key descriptors, beyond those commonly used, could play an important role in predicting thermal conductivity, thus warranting further exploration of structure–property relationships.

Discovering novel thermal transfer mechanisms for two-dimensional COFs *via* self-attention

We leverage the multi-head self-attention mechanism⁴⁹ of the transformer architecture to interpret the predictions made by the deep learning model. The attention mechanism in transformers is designed to identify and assign weights, known as attention scores, to the significance of different parts of the input data/encodings relative to all the other parts. When applied to molecular structures, such as COFs, the attention mechanism enables the model to dynamically focus attention on key molecular substructures in relation to all the other substructures present based on their relevance to the property being predicted, in this case, thermal conductivity. In multi-head attention, multiple attention heads are needed to capture different aspects or relationships within the data simultaneously. Specifically, in the context of the PMTransformer model applied to a COF to predict its thermal conductivity, we calculate the attention for each atom in the COF by averaging the attention scores from all attention heads and computing the joint attention *via* multiplicative aggregation across all layers of the transformer encoder.

Additionally, we identify the main branch of a COF as the shortest continuous path connecting the boundary points necessary for periodicity. We then classify all the atoms extending and excluded from the main branch as dangling



mass. More details on how we classify each atom as either part of the main branch or part of a dangling side branch are presented in SI Section 4. We observe distinct patterns in the attention assigned to various atomic sites within a COF's graph representation, which correspond to the location of the dangling masses. Two representative examples are shown in Fig. 4 and 5.

Consider the first example (see Fig. 4a and b). The top row illustrates a COF structure (Fig. 4a) from the test dataset with a thermal conductivity of $1.105 \text{ W m}^{-1} \text{ K}^{-1}$ and no dangling masses except hydrogen atoms. Additionally, the atom-wise attention scores are uniformly distributed and low for all the carbon atoms with some elevated attention to the nitrogen atoms. In the second row, we illustrate a COF structure (Fig. 4b) with the same topology as the previous one, but with a much lower thermal conductivity of $0.533 \text{ W m}^{-1} \text{ K}^{-1}$. Its attention profile reveals that certain atom groups are being paid special attention to by the transformer, *i.e.*, specifically the $-\text{NO}_2$ functional group on the benzene ring that is dangling from the main branch of the COF. Notably, the same groups of atoms that exhibit higher attention scores are also classified as dangling masses (see Fig. 4b attention profile & dangling mass).

In the second example (Fig. 5a and b), the COF structure on the top does not contain any dangling mass except hydrogen

atoms and has a thermal conductivity of $2.715 \text{ W m}^{-1} \text{ K}^{-1}$ with a highly uniform attention profile (see Fig. 5a). Analogously, we pick a second COF structure with the same topology and comparable geometrical descriptors. Once more, we observe that the second structure contains a significant amount of dangling mass (*i.e.*, $-\text{CN}$ branches extending from the rings) with corresponding elevated attention scores and has a lower thermal conductivity of $1.611 \text{ W m}^{-1} \text{ K}^{-1}$.

The same trend is observed for numerous examples, with additional ones presented in the SI, Section 5. We hypothesize that this increased attention is indicative of the deep learning model's understanding that these dangling masses are significant in predicting thermal conductivity. Furthermore, it suggests that the presence of dangling masses disrupts the heat transfer pathways and thereby reduces the thermal conductivity through the material. A similar effect of dangling mass lowering the thermal conductivity has been reported previously in singular polymer chains⁵³ and amorphous polymers.⁵⁴ This effect, while known in disordered systems, has not been examined in crystalline COFs. Here, we demonstrate that it remains relevant even in ordered, porous frameworks, inviting future investigation with high-fidelity modeling methods (such as density functional theory) to examine the underlying mechanism in more detail.

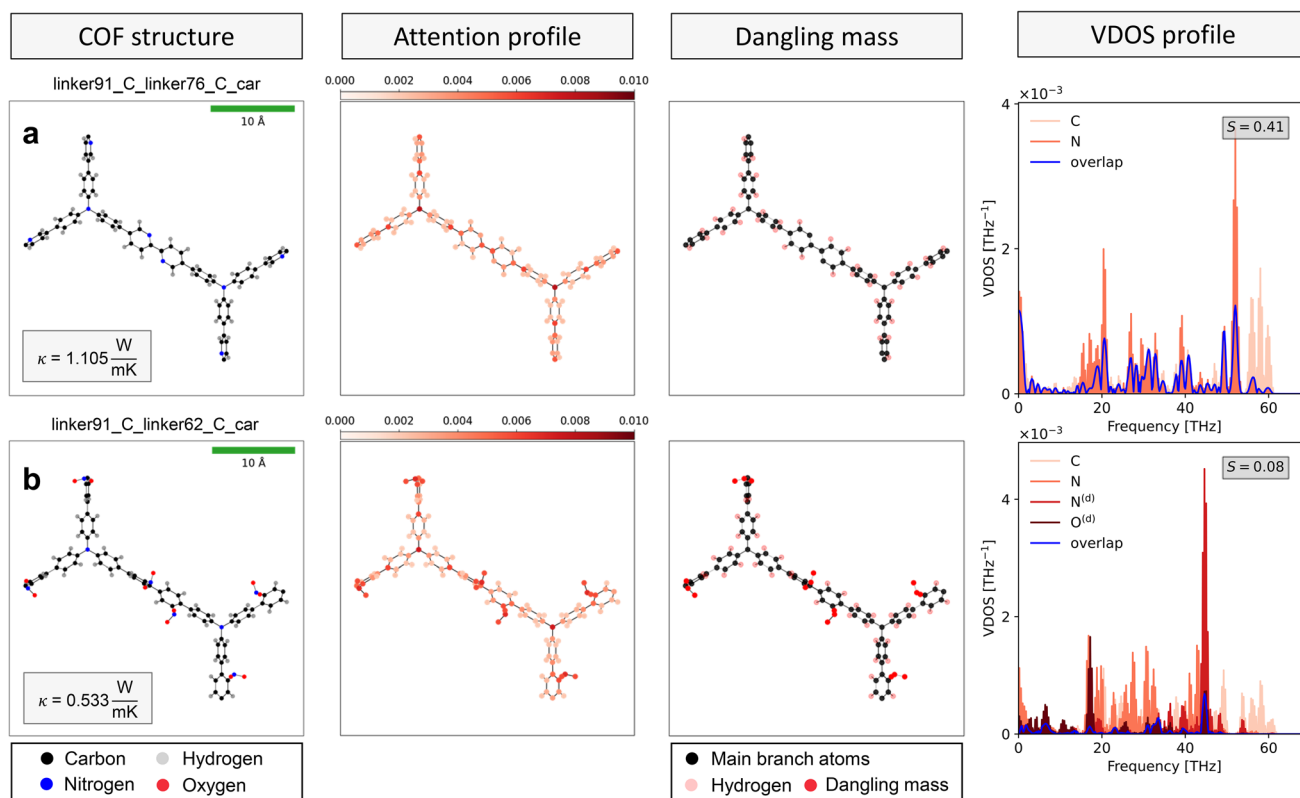


Fig. 4 An example pair of COFs with the same topologies and similar geometric descriptors, but contrasting thermal conductivities. The first column illustrates the COF structures without (a) and with (b) dangling mass. The second column shows the atom-level attention score profile computed by the attention mechanism. The third column shows the same COF structure distinguishing atoms on the main branch and dangling atoms (with separate distinction for hydrogen atoms). The fourth column shows the VDOS profiles of various groups of atoms within the corresponding COF structure with the overlap metric S . The legend indicates the VDOS profile for main branch atoms (.) and dangling atoms (.)^(d). The reported thermal conductivities are obtained from NEMD simulations, rather than being predicted by the PMTransformer model.



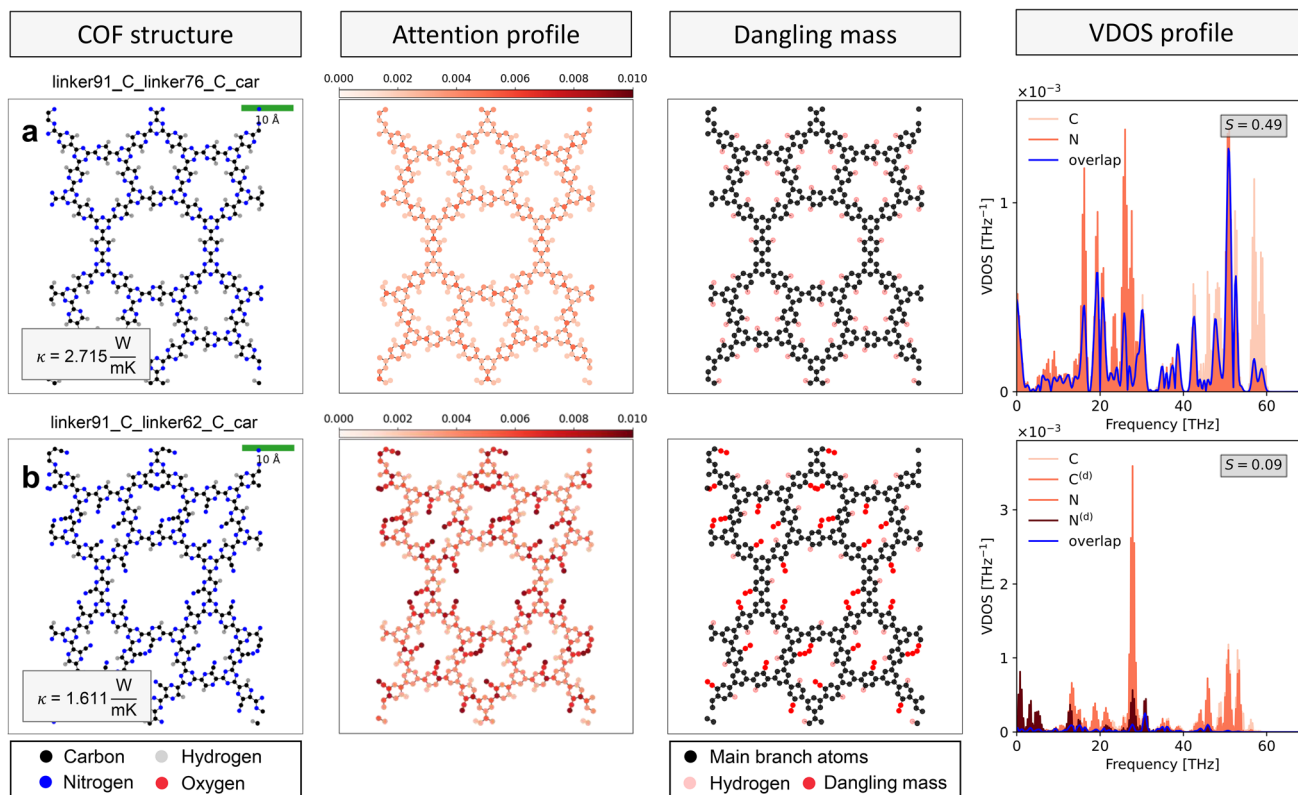


Fig. 5 An example pair of COFs with the same topologies and similar geometric descriptors, but contrasting thermal conductivities. The first column illustrates the COF structures without (a) and with (b) dangling mass. The second column shows the atom-level attention score profile computed by the attention mechanism. The third column shows the same COF structure distinguishing atoms on the main branch and dangling atoms (with separate distinction for hydrogen atoms). The fourth column shows the VDOS profiles of various groups of atoms within the corresponding COF structure with the overlap metric S . The legend indicates the VDOS profile for main branch atoms (.) and dangling atoms (.)^d. The reported thermal conductivities are obtained from NEMD simulations, rather than being predicted by the PMTransformer model.

To further validate the structure–property relationship interpreted from the deep learning model, we examine the impact of dangling atoms on the vibrational density of states (VDOS), which characterizes the distribution of vibrational modes in a system as a function of frequency, is known to impact thermal transfer properties. Overlaps in VDOS profiles between different atoms have been shown to affect these properties in both COFs⁵⁵ and MOFs.⁵⁶

Through MD simulations, the VDOS is calculated using the Fourier transform of the normalized velocity autocorrelation function of specific groups of atoms (see SI Section 6 for details). Fig. 4 and 5 (third column) show the VDOS profile of carbon, nitrogen, and oxygen atoms in the representative COF examples. For any COF, we define a VDOS overlap metric S as the ratio of the area under the curve (AUC) of the minimum VDOS across all atom types (at each frequency) and the AUC of the maximum VDOS across all atom types (at each frequency), *i.e.*,

$$S = \frac{\int_{\omega} \min\{f_C^V(\omega), f_N^V(\omega), f_O^V(\omega), f_B^V(\omega), f_C^{V,d}(\omega), f_N^{V,d}(\omega), f_O^{V,d}(\omega), f_B^{V,d}(\omega)\} d\omega}{\int_{\omega} \max\{f_C^V(\omega), f_N^V(\omega), f_O^V(\omega), f_B^V(\omega), f_C^{V,d}(\omega), f_N^{V,d}(\omega), f_O^{V,d}(\omega), f_B^{V,d}(\omega)\} d\omega} \quad (1)$$

Here, ω denotes the frequency; $f_C^V(\omega)$, $f_N^V(\omega)$, $f_O^V(\omega)$, and $f_B^V(\omega)$ denote the VDOS at frequency ω for the carbon, nitrogen, oxygen and boron atoms, respectively. Additionally, we differentiate between atoms in the main branch and dangling atoms, where the VDOS profiles for dangling atoms are denoted with a (.)^d superscript. A large overlap in VDOS profiles of different atom types would result in $S \approx 1$, whereas a small overlap would yield $S \approx 0$. A high overlap indicates that phonon waves (*i.e.*, quantized modes of vibrations responsible for thermal energy transfer within the crystal lattice) can propagate freely throughout the structure. This would lead to minimized phonon scattering and facilitate efficient heat transfer across the material.⁵⁷

In the COF structures, which lack or have minimal dangling atoms, the VDOS profiles of the atoms are broad and overlap significantly, *e.g.*, $S \approx 0.41$ and 0.49 for examples in Fig. 4a and 5a, respectively. This results in an even distribution of vibrational modes across a wide frequency spectrum for all atoms. In contrast, the COFs with substantial dangling atoms exhibit a VDOS profile with minimal overlap, *e.g.*, $S \approx 0.08$ and 0.09 for



examples in Fig. 4b and 5b, respectively. Notably, dangling atoms significantly reduce the VDOS overlap in both low-frequency (0 to 20 THz) and high-frequency (40 to 60 THz) modes. The vibrational modes of individual atoms in these COFs are confined to much narrower frequency bands, which leads to a less harmonized VDOS profile. The lack of overlap suggests that the vibrations are highly localized around the dangling atoms, which causes the phonon waves to be scattered, disrupting their ability to transfer heat efficiently through the material.⁵⁸ We interpret the mismatch in the VDOS profiles as an energy barrier for phonon transport, where phonons encounter resistance, preventing smooth energy transfer across atoms. This suggests that the absence of dangling mass is crucial for enhancing thermal conductivity in COFs. Dangling atoms create these mismatches in VDOS, which in turn hinders phonon transport and reduces thermal conductivity. A similar trend is observed in additional pairs of COFs (Fig. S9, SI) and the correlation between κ and S for a selected representative set of eight COFs is presented in Fig. S6c (SI).

To further confirm the effect of dangling atoms on phonon dispersion in COFs, we calculate the phonon spectral energy density (pSED) of the example COFs and their dangling counterparts, as presented in Section 5 in the SI. The pSED of COFs containing more dangling atoms exhibits higher magnitudes and broadening bands, indicating stronger anharmonicity and vibrational scattering in the vibrational modes, consequently resulting in reduced thermal conductivity.

To quantify the role of dangling masses in the structure–property relationship for thermal conductivity of COFs, we introduce the dangling mass ratio (DMR). The $\text{DMR} \in [0, 1]$ quantifies the ratio between the mass of dangling atoms and the total mass of the COF, *i.e.*,

$$\text{DMR} = \frac{\sum_{i \in \mathcal{D}} m_i}{\sum_{i \in \mathcal{D}} m_i + \sum_{i \in \mathcal{M}} m_i}, \quad (2)$$

where m_i is the mass of the i th atom belonging to the set of either dangling atoms or of non-dangling main branch atoms – denoted by \mathcal{D} and \mathcal{M} , respectively. Higher DMR values indicate a greater proportion of dangling masses relative to the COF's total mass. In SI Section 1.3, we qualitatively show the relationship between DMR, the VDOS overlap S , and κ .

We then introduce DMR as an additional feature, alongside pore size, density, void fraction, and surface area, and train classical ensemble regression models, such as the Random Forest,⁵⁹ Gradient Boosting,⁶⁰ XGBoost,⁶¹ and AdaBoost⁶² and outline their performances in Table 1. These simpler models are

Table 1 Performances of standard ensemble regression models at predicting κ by including DMR as an input feature. The table reports R^2 scores obtained through 10-fold cross-validation

Regression model	R^2
Random Forest	0.728 ± 0.069
Gradient Boosting	0.688 ± 0.070
XGBoost	0.691 ± 0.035
AdaBoost	0.416 ± 0.125

employed not to compete with the PMTransformer in predictive accuracy, but to provide directly interpretable feature importance scores and help in understanding the relevance of DMR. In contrast, extracting meaningful and quantitative feature importance from the PMTransformer is non-trivial due to its complex model architecture and the absence of an explicit correspondence between input features and learned representations. To assess the importance of individual features, including DMR, we analyzed their Gini importance and conducted a permutation feature importance analysis.⁶³ These values quantify the contribution of each feature to the overall predictive power of the model. The Gini importance measures how much each feature reduces the Gini impurity or randomness when making predictions in the ensemble model.⁵⁹ On the other hand, permutation feature importance measures the impact of a feature by assessing how much the model's performance decreases when the values of that feature are randomly shuffled. A greater drop in performance indicates that the feature is more important in predicting the target variable.⁵⁹ We note that for the best performing regressor, the Random Forest, DMR emerges as the second most influential predictor with 27.8% Gini importance, closely following behind density (see Table 2). Similarly, the permutation feature importance analysis also highlights DMR as a key feature, further supporting its significant role in predicting thermal conductivity, with its importance consistently ranking just below density. Finally, we conducted SHAP analysis⁶⁴ on the Random Forest to better understand the impact of each feature on the individual test data points. As shown in Fig. 6, the analysis reveals that higher DMR values are associated with lower predicted κ , confirming the trend we observed. The consistency of these findings across different methods—Gini, permutation feature importance, and SHAP—reinforces the importance of DMR as a key factor in predicting thermal conductivity. We note that the computed importance scores are inherently linked to the design space and dataset introduced by Mercado *et al.*⁴⁷ Specifically, if the dataset did not contain structures with dangling masses, the DMR feature would naturally show little to no importance. Thus, the relevance of DMR as a descriptor depends on its variability within the dataset under consideration.

Generalizability and high-throughput screening for discovering two-dimensional monolayer COFs with tailored properties

To evaluate the generalization capabilities of the deep learning model for high-throughput screening, we screen previously

Table 2 Feature importances for the best-performing regression model, *i.e.*, Random Forest. The table reports both Gini and Permutation Feature Importance (PFI) values for all features

Feature	Gini	PFI
Density	0.397	0.670
DMR	0.278	0.511
LPD	0.185	0.455
Void Fraction	0.070	0.093
GSA	0.069	0.022



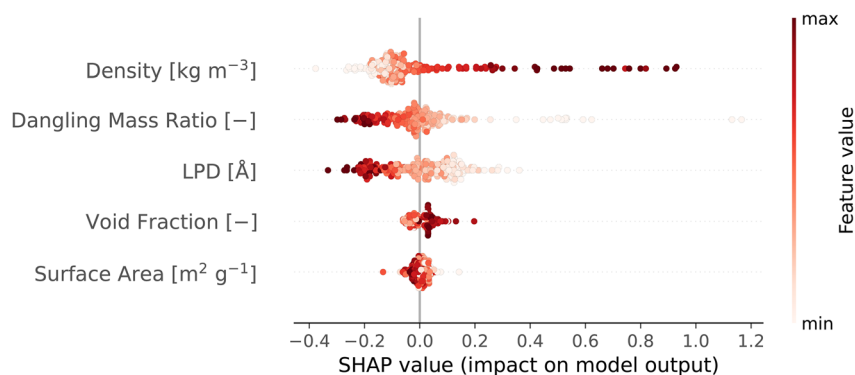


Fig. 6 Distribution of the effect of each feature on the model's predictions. Each point represents a SHAP value for a feature across all samples, with the color intensity indicating feature values (ranging between minimum and maximum values). The position along the x-axis represents the magnitude of the feature's impact.

unseen 6170 two-dimensional COFs in the unlabelled dataset by Mercado *et al.*⁴⁷ from which our training and test sets were originally subsampled. Moreover, we screen 35 638 two-dimensional COFs in the unlabelled dataset by Lan *et al.*⁶⁵ This dataset covers a vastly different design space with a variety of linkers and knots previously unseen by our trained model.

The screening is performed with the objective of facilitating the identification COFs with extremely high or low thermal conductivity without resorting to computationally expensive simulations. We observe an average screening rate of 0.07 seconds per COF structure (see SI Section 7 for details on computational runtimes), thereby enabling a speedup of almost seven orders of magnitude compared to MD-based screening.

Once potential candidates with extreme thermal conductivities are identified using the PMTransformer, we then calculate their thermal conductivities through MD simulations for these selected few, rather than for the entire datasets. This approach circumvents reliance on the PMTransformer's absolute predictions by validating promising candidates with high-fidelity MD calculations. The results are summarized in Fig. 7. Our screening process on the dataset by Mercado *et al.*⁴⁷ revealed COFs exhibiting thermal conductivities up to twice the maximum value observed in the training set, but stemming from the same design space. By screening the dataset by Lan *et al.*,⁶⁵ we discover COFs with thermal conductivities five times as high as the ones in our training set. Additionally, we discovered several COFs with thermal conductivities lower than any values previously recorded in our dataset. A visual inspection reveals that these COFs feature a combination of large pore sizes, low density, and a substantial presence of dangling masses. For all the identified COFs, we observe that thermal conductivity shows a strong trend with both dangling mass ratio (DMR) and VDOS overlap ratio (S), which is in agreement with the previous observations.

Discussion

In this study, we compiled an extensive dataset of thermal conductivities for 2471 two-dimensional COFs, computed using NEMD. Despite observing some structure–property trends with conventional descriptors such as density, pore size, void

fraction, and surface area, no single descriptor or combination thereof consistently predicted thermal conductivity with high accuracy, highlighting the complexity of COF structure–property relationships. To enhance prediction accuracy, we trained a transformer-based deep learning model incorporating a multi-head attention mechanism. This model significantly outperformed traditional ensemble regression models, achieving an R^2 of 0.909 ± 0.006 .

Further analysis using the transformer's attention mechanism revealed that COFs with higher amounts of dangling atoms exhibited lower thermal conductivities due to disrupted heat transfer pathways, identifying a novel and significant predictor of thermal conductivity. The random forest regressor, identified as the best-performing regression ensemble model, was used primarily as a simple tool to assess feature importance rather than to compare predictive performance with the PMTransformer. It was analyzed using Gini importance, permutation feature importance, and SHAP values. These analyses consistently highlighted the dangling mass ratio as the second most important predictor of thermal conductivity. The vibrational density of states analysis further supported these findings, showing that dangling masses introduce mismatched vibrational modes that hinder thermal transfer. While our current analysis focuses on two-dimensional monolayer COFs, additional mechanisms may emerge in three-dimensional architectures and warrant a separate investigation.

We further leveraged the deep learning model to efficiently screen thousands of COFs, both within and beyond the design space we analyzed, to identify candidates with extreme thermal conductivities. This approach provides a rapid and reliable method as well as valuable insights for designing COFs with tailored thermal characteristics. To support further research, we release the dataset of COF thermal conductivities and encourage the research community to utilize and expand upon it. Finally, we emphasize the need to bridge high interpretability with high accuracy in machine learning models to navigate the complex structure–property space of COFs and other nanoporous materials, thereby enabling the design of optimal materials for applications such as gas separation and thermal management.



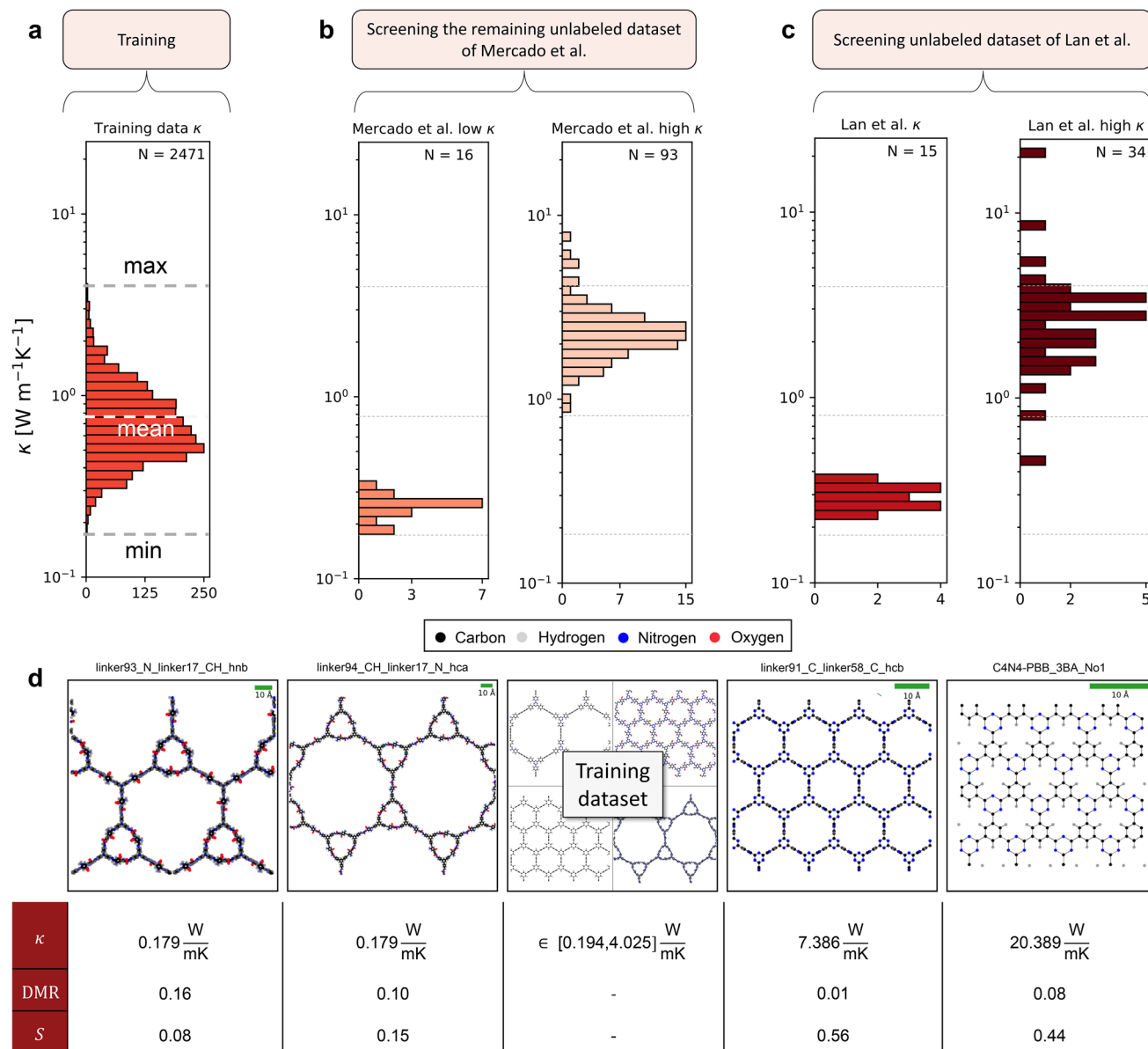


Fig. 7 High-throughput screening. (a) Distribution of κ in the training dataset. Dashed lines indicated minimum, mean, and maximum κ across the training dataset. (b and c) Distribution of κ of selected COF candidates identified in high-throughput screening for low and high κ in the unlabeled datasets of (b) Mercado *et al.*⁴⁷ and (c) Lan *et al.*⁶⁵ While the former dataset shares the same design space as the training dataset, the latter has a different design space. All thermal conductivities shown here are computed using MD. (d) Representative COFs identified through high-throughput screening with κ lower and higher than the minimum and maximum, respectively, across the training dataset. Also indicated are their corresponding DMR and S values.

Methods

The molecular dynamics setup, the distribution of in-plane thermal conductivities, and feature correlations (including DMR) (Section 1); implementation details of the ensemble regression models and regression performances excluding DMR (Section 2); the deep learning setup (Section 3); the dangling mass computation (Section 4); additional comparisons of COFs with and without dangling mass, including phonon dispersion maps for all examples (Section 5); details on the VDOS and pSED computation (Section 6); and estimated runtimes (Section 7) are summarized in the SI.

Author contributions

P. T. and Y. Z.: conceptualization, data curation, formal analysis, investigation, writing – original draft. P. T., Y. Z., Y. G.: methodology, software, validation, visualization. A. V. and S. K.: conceptualization, supervision, funding acquisition, writing – review & editing.

Conflicts of interest

The authors declare no competing interests.



Data availability

The code and data generated during this study are available in the following repository: <https://doi.org/10.4121/5866cc9a-78bf-4a0c-9280-ae526da86ac9>. The code and data are also available at: <https://github.com/mmc-group/deep-learning-thermal-conductivity-of-COFs>. The data includes unit-cell CIF files obtained from the dataset of Mercado *et al.*⁴⁷ available at <https://doi.org/10.24435/materialscloud:2018.0003/v2>. We also used the dataset provided by Lan *et al.*,⁶⁵ which is available at <https://figshare.com/s/c7e3b7610a71b9d64210> with DOI: <https://doi.org/10.1038/s41467-018-07720-x> for the associated paper. The codes are built upon the open-source MOFTransformer/PMTransformer repository by Park *et al.*,⁴⁶ which can be accessed at <https://github.com/hspark1212/MOFTransformer> with DOI: <https://doi.org/10.1021/acsami.3c10323> for the associated paper.

Supplementary information is available. See DOI: <https://doi.org/10.1039/d5dd00126a>.

Acknowledgements

P. T., Y. G., and S. K. acknowledge that this material is based upon work supported by the Air Force Office of Scientific Research under award number FA8655-23-1-7020. P. T., Y. G., and S. K. acknowledge that this work was supported in part by Oracle Cloud credits and related resources provided by Oracle for Research. A. V. and Y. Z. acknowledge the Microsoft Climate Research Initiative and University of Washington for funding. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript. We thank Prof. Marcel Sluiter of TU Delft for his helpful discussions and insights during the course of this research. Additionally, we thank Eric Vo of the University of Washington for his help in preparing simulations of COF structures.

Notes and references

- H. Wang, H. Wang, Z. Wang, L. Tang, G. Zeng, P. Xu, M. Chen, T. Xiong, C. Zhou, X. Li, D. Huang, Y. Zhu, Z. Wang and J. Tang, *Chem. Soc. Rev.*, 2020, **49**, 4135–4165.
- X. Zhao, P. Pachfule and A. Thomas, *Chem. Soc. Rev.*, 2021, **50**, 6871–6913.
- C. Liu, J. Wang, J. Wan and C. Yu, *Coord. Chem. Rev.*, 2021, **432**, 213743.
- N. Stock and S. Biswas, *Chem. Rev.*, 2011, **112**, 933–969.
- E. Pérez-Botella, S. Valencia and F. Rey, *Chem. Rev.*, 2022, **122**, 17647–17695.
- H. R. Abuzeid, A. F. EL-Mahdy and S.-W. Kuo, *Giant*, 2021, **6**, 100054.
- M. Dogru, M. Handloser, F. Auras, T. Kunz, D. Medina, A. Hartschuh, P. Knochel and T. Bein, *Angew. Chem., Int. Ed.*, 2013, **52**, 2920–2924.
- L. Stegbauer, S. Zech, G. Savasci, T. Banerjee, F. Podjaski, K. Schwinghammer, C. Ochsenfeld and B. V. Lotsch, *Adv. Energy Mater.*, 2018, **8**, 1703278.

- Z. Ali, T. Huo, Y. Zhang and G. Wang, *Microchem. J.*, 2024, **200**, 110340.
- Y. Li, M. Chen, Y. Han, Y. Feng, Z. Zhang and B. Zhang, *Chem. Mater.*, 2020, **32**, 2532–2540.
- Y. Yusran, H. Li, X. Guan, Q. Fang and S. Qiu, *EnergyChem*, 2020, **2**, 100035.
- M. C. Scicluna and L. Vella-Zarb, *ACS Appl. Nano Mater.*, 2020, **3**, 3097–3115.
- H. Guo, Y. Liu, N. Wu, L. Sun and W. Yang, *ChemistrySelect*, 2022, **7**, e202202538.
- L. Wang, B. Dong, R. Ge, F. Jiang and J. Xu, *ACS Appl. Mater. Interfaces*, 2017, **9**, 7108–7114.
- Y. Chumakov, F. Aksakal, A. Dimoglo, A. Ata and S. A. Palomares-Sánchez, *J. Electron. Mater.*, 2016, **45**, 3445–3452.
- S. Wang, X. Xu, Y. Yue, K. Yu, Q. Shui, N. Huang and H. Chen, *Small Struct.*, 2020, **1**, 2000021.
- H. Fan, A. Mundstock, A. Feldhoff, A. Knebel, J. Gu, H. Meng and J. Caro, *J. Am. Chem. Soc.*, 2018, **140**, 10094–10098.
- M.-X. Wu and Y.-W. Yang, *Chin. Chem. Lett.*, 2017, **28**, 1135–1143.
- X. Zhou, Y. Yan, X. Lu, H. Zhu, X. Han, G. Chen and Z. Ren, *Mater. Today*, 2018, **21**, 974–988.
- A. M. Evans, M. R. Ryder, W. Ji, M. J. Strauss, A. R. Corcos, E. Vitaku, N. C. Flanders, R. P. Bisbey and W. R. Dichtel, *Faraday Discuss.*, 2021, **225**, 226–240.
- X. Zou, H. Ren and G. Zhu, *Chem. Commun.*, 2013, **49**, 3925.
- Z.-F. Pang, T.-Y. Zhou, R.-R. Liang, Q.-Y. Qi and X. Zhao, *Chem. Sci.*, 2017, **8**, 3866–3870.
- B. Gui, G. Lin, H. Ding, C. Gao, A. Mal and C. Wang, *Acc. Chem. Res.*, 2020, **53**, 2225–2234.
- X. Xu, J. Chen, J. Zhou and B. Li, *Adv. Mater.*, 2018, **30**, 1705544.
- C. Choy, *Polymer*, 1977, **18**, 984–1004.
- K. Utimula, T. Ichibha, R. Maezono and K. Hongo, *Chem. Mater.*, 2019, **31**, 4649–4656.
- L. Lindsay, D. A. Broido and T. L. Reinecke, *Phys. Rev. Lett.*, 2012, **109**, 095901.
- A. H. Romero, E. K. U. Gross, M. J. Verstraete and O. Hellman, *Phys. Rev. B*, 2015, **91**, 214310.
- H. Li and J.-L. Brédas, *Chem. Mater.*, 2021, **33**, 4529–4540.
- Z. Li, Z. Liu, Z. Li, T. Wang, F. Zhao, X. Ding, W. Feng and B. Han, *Adv. Funct. Mater.*, 2020, **30**, 1909267.
- D. C. Rapaport, *The Art of Molecular Dynamics Simulation*, Cambridge University Press, 2004.
- M. Islamov, H. Babaei, R. Anderson, K. B. Sezginel, J. R. Long, A. J. H. McGaughey, D. A. Gomez-Gualdrón and C. E. Wilmer, *npj Comput. Mater.*, 2023, **9**, 11.
- S. Thakur and A. Giri, *Small*, 2024, **20**, 2401702.
- X. Huang, S. Ma, C. Y. Zhao, H. Wang and S. Ju, *npj Comput. Mater.*, 2023, **9**, 191.
- R. Ma, H. Zhang and T. Luo, *ACS Appl. Mater. Interfaces*, 2022, **14**, 15587–15598.
- S. Wu, Y. Kondo, M.-a. Kakimoto, B. Yang, H. Yamada, I. Kuwajima, G. Lambard, K. Hongo, Y. Xu, J. Shiomi, *et al.*, *npj Comput. Mater.*, 2019, **5**, 66.
- K. Ishikiriyama, *Thermochim. Acta*, 2022, **708**, 179135.



- 38 J. Xu and T. Luo, *npj Comput. Mater.*, 2024, **10**, 74.
- 39 Y. Hu, Q. Wang and H. Ma, *J. Appl. Phys.*, 2024, **135**, 120701.
- 40 P. Yang, H. Zhang, X. Lai, K. Wang, Q. Yang and D. Yu, *ACS Omega*, 2021, **6**, 17149–17161.
- 41 X. Cao, Z. Zhang, Y. He, W. Xue, H. Huang and C. Zhong, *Ind. Eng. Chem. Res.*, 2022, **61**, 11116–11123.
- 42 J. S. De Vos, S. Ravichandran, S. Borgmans, L. Vanduyfhuys, P. Van Der Voort, S. M. J. Rogge and V. Van Speybroeck, *Chem. Mater.*, 2024, **36**, 4315–4330.
- 43 G. Zhao and Y. G. Chung, *J. Chem. Theory Comput.*, 2024, **20**, 5368–5380.
- 44 V. Korolev and A. Mitrofanov, *J. Chem. Inf. Model.*, 2024, **64**, 1919–1931.
- 45 Y. Kang, H. Park, B. Smit and J. Kim, *Nat. Mach. Intell.*, 2023, **5**, 309–318.
- 46 H. Park, Y. Kang and J. Kim, *ACS Appl. Mater. Interfaces*, 2023, **15**, 56375–56385.
- 47 R. Mercado, R.-S. Fu, A. V. Yakutovich, L. Talirz, M. Haranczyk and B. Smit, *Chem. Mater.*, 2018, **30**, 5069–5086.
- 48 S. K. S. Freitas, R. S. Borges, C. Merlini, G. M. O. Barra and P. M. Esteves, *J. Phys. Chem. C*, 2017, **121**, 27247–27252.
- 49 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, *arXiv*, 2017, preprint, arXiv:1706.03762, DOI: [10.48550/arXiv.1706.03762](https://doi.org/10.48550/arXiv.1706.03762).
- 50 T. Xie and J. C. Grossman, *Phys. Rev. Lett.*, 2018, **120**, 145301.
- 51 GitHub, Sangwon91/GRIDAY: Energy shape calculator for the porous materials—github.com, <https://github.com/Sangwon91/GRIDAY>, accessed 14-06-2024.
- 52 A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit and N. Houlsby, *arXiv*, 2020, preprint, arXiv:2010.11929v2, DOI: [10.48550/arXiv.2010.11929](https://doi.org/10.48550/arXiv.2010.11929).
- 53 D. Luo, C. Huang and Z. Huang, *J. Heat Transfer*, 2017, **140**, 031302.
- 54 T. Feng, J. He, A. Rai, D. Hun, J. Liu and S. S. Shrestha, *Phys. Rev. Appl.*, 2020, **14**, 044023.
- 55 D. Feng, Y. Feng, Y. Liu, W. Zhang, Y. Yan and X. Zhang, *J. Phys. Chem. C*, 2020, **124**, 8386–8393.
- 56 P. Ying, J. Zhang and Z. Zhong, *J. Phys. Chem. C*, 2021, **125**, 12991–13001.
- 57 S. Wang, L. Ren, M. Han, W. Zhou, C. Wong, X. Bai, R. Sun and X. Zeng, *Nanoscale*, 2023, **15**, 8706–8715.
- 58 P. Ying, J. Zhang, X. Zhang and Z. Zhong, *J. Phys. Chem. C*, 2020, **124**, 6274–6283.
- 59 L. Breiman, *Mach. Learn.*, 2001, **45**, 5–32.
- 60 J. H. Friedman, *Ann. Stat.*, 2001, **29**, 1189–1232.
- 61 T. Chen and C. Guestrin, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- 62 Y. Freund and R. E. Schapire, *J. Comput. Syst. Sci.*, 1997, **55**, 119–139.
- 63 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
- 64 S. Lundberg, *arXiv*, 2017, preprint, arXiv:1705.07874v2, DOI: [10.48550/arXiv.1705.07874](https://doi.org/10.48550/arXiv.1705.07874).
- 65 Y. Lan, X. Han, M. Tong, H. Huang, Q. Yang, D. Liu, X. Zhao and C. Zhong, *Nat. Commun.*, 2018, **9**, 5274.

