

Supplementary Information

Discovery of Hypothetical Hetero-Interpenetrated MOFs with Arbitrarily Dissimilar Topologies and Unit Cell Shapes†

K. B. Sezginel,^a T. Feng^a and C. E. Wilmer^{a*}

Table of Contents

1. Algorithm Description	1
a. Energy map generation	3
b. Force field parameters	4
c. Algorithm parameters	5
2. Generating Unit Cell for Interpenetrating Structures	7
a. Calculating supercell parameters	7
b. Cell Distortion	8
3. High-throughput Screening	8
a. Database selection	8
b. Void fraction calculation	9
c. Analysis of candidate structures	9
4. Comparison with Literature	14
a. Comparison of candidate structures	14
b. Comparison of experimental hetero-interpenetrated structures	15

^a Department of Chemical & Petroleum Engineering, University of Pittsburgh, 3700 O'Hara St, Pittsburgh, PA 15261, USA.

* Corresponding author: wilmer@pitt.edu

1. Algorithm Description

The algorithm tests whether two given crystal structures with any topology can theoretically interpenetrate each other. For two given unit cells, different orientations of the unit cells are checked for collisions and any favorable orientations are reported. Our main aim was to screen a large database to identify new interpenetrated structures therefore we needed a method that would rapidly identify possible interpenetration candidates. Moreover, we knew most of the combinations would be unfavorable for interpenetration therefore using a high accuracy model would result in wasting computer time. For that reason, we developed the “IPMOF” algorithm to rapidly identify which MOF pairs can possibly interpenetrate each other.

Initially, two MOFs are selected and one of them is designated as *passive* and the other as *active* (see Fig. 5). For the *passive* MOF, an *energy map* is generated to calculate energy penalty associated with interpenetrating it with the *active* MOF. Secondly, different orientations of *active* MOF are tested by moving and rotating the structure. We did not consider improper rotation while testing different orientations. An *energy density* is calculated for each orientation and if the *energy density* is below the selected threshold value orientation is saved. After all orientations are tested, most favorable ones are identified according to *energy densities*.

This initial test is referred as *single unit-cell test* since it only considers interpenetration between single unit cells of two given MOFs. However, it is possible to observe atomic collisions in repeating unit cells even if there were no collisions in the single unit cells. For MOF pairs which pass the *single unit cell test*, *multiple unit-cell collision test* is performed to make sure that differences in repeating patterns of the unit cells do not cause overlap when they are extended multiple times. Next step is to generate a supercell for the interpenetrating MOFs, which allows both structures to repeat according to their symmetry. After this step is complete a structure file for the hypothetical interpenetrated MOF is generated.

The algorithm is provided here (<https://github.com/WilmerLab/IPMOF>) where detailed documentation is also included for the software. In addition, detailed examples on how to use the algorithm are also provided in Jupyter notebooks. We tried to make it as easy as possible for others to use our software in order to help the community explore new interpenetrated materials. Following the documentation it is relatively easy to perform simulations on a personal laptop with minimum programming experience. Our aim here was to make sure our algorithm can be beneficial for both experimental and computational scientists. Moreover, we provided the candidate structures discovered in this study with extensive structural and crystallographic information as well as references to publications in which the MOFs used in this study were reported.

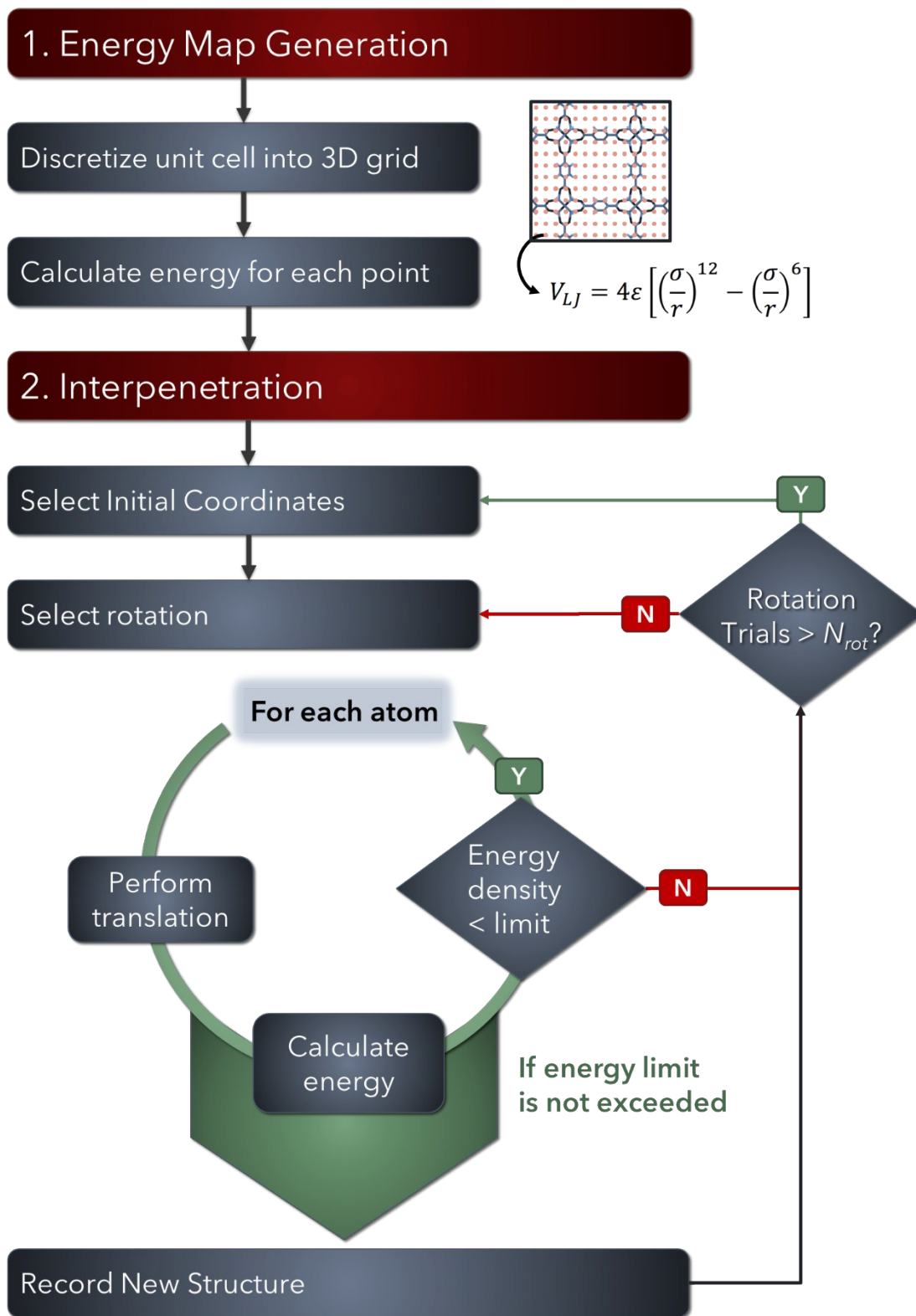


Figure 1 Flowchart for IPMOF algorithm. Algorithm and its documentation with examples are provided here: <https://github.com/WilmerLab/IPMOF>.

a. Energy map generation

The *energy map* is defined as a regular grid of points representing the interatomic potential energy of the unit cell (see Fig. 2). The potential energy for each point in the grid is calculated using Lennard-Jones potential¹ and Universal Force Field² (UFF) parameters with Lorentz-Berthelot mixing rules and a cut-off radius of 12 Å.

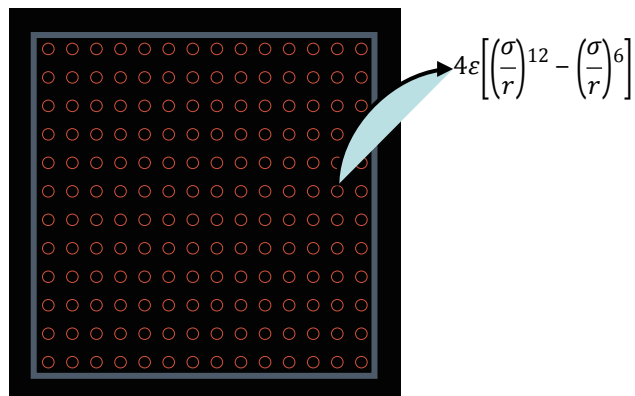


Figure 2 Energy map schematic for MOF-5. Each red point represents grid points 1 Å apart in 3 dimensions. For each point potential energy is calculated using Lennard-Jones equation.

In Figure 3 *energy map* generation process for MOF-5³ is provided. As expected, empty regions in the unit cell are lower in energy (shown in dark red) and atomically crowded regions are higher in energy (light red). The *energy map* shown below is calculated using UFF parameters of a carbon atom which provides energetic information for inserting a carbon atom to any location in the unit cell. Each type of atom results in a different energy value for the grid points of the *energy map*, therefore the *energy map* array consists of 3D coordinates as well as different energy values for each type of atom used in the calculation of the *energy map*.

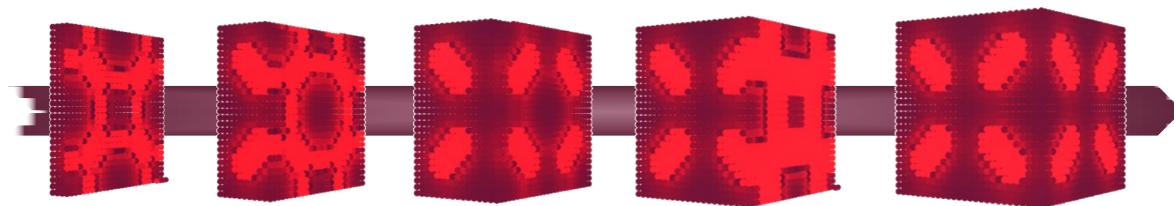


Figure 3 Energy map generation process for MOF-5. The potential energy increases from dark red to light red.

For high throughput screening energy values for various atom types are required to test interpenetration. However, this would result in having tens of energy values for each position in the *energy map*. To avoid having large *energy map* files force field parameters were simplified to only include total 11 atom types and consequently 11 energy values for each position in *energy map*.

b. Force field parameters

Lennard-Jones equation is used to estimate interatomic potentials:

$$4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right]$$

where ε and σ values used for atoms are listed below and r is the distance between atoms in Å. The resulting energy is in $K.k_B$ units (0.0083144621 kJ/mol).

As mentioned in section 1.a, for high throughput screening, energy values for various atom types are required to test interpenetration. This would be computationally expensive and the generated *energy map* files would take a lot of space. However, all MOFs are composed of organic linkers and metal cluster which suggests that almost all MOFs will have C, H, O, N atoms and one or more metal atoms. Making use of this information we decided to use a simplified force field by only considering actual UFF parameters for 10 atoms and assigning dummy force field parameters for the rest. In this way, we only needed to calculate 11 different energy values for each position in the *energy map* instead of almost a hundred. In addition to C, H, O, N atoms for organic linkers we identified the 6 most frequent metal atoms in the database (Fig. 4).

Table 1 Force field parameters used in this study.

Atom	σ (Å)	ε/k_B (K)
C	3.430851	52.87336
Cd	2.53728	114.8107
Co	2.558661	7.049781
Cu	3.113691	2.517779
H	2.571134	22.15645
Mg	2.691405	55.89469
Mn	2.637951	6.546225
N	3.260689	34.74535
Ni	2.524807	7.553337
O	3.118146	30.21335
Zn	2.461553	62.44092
Dummy	3	30

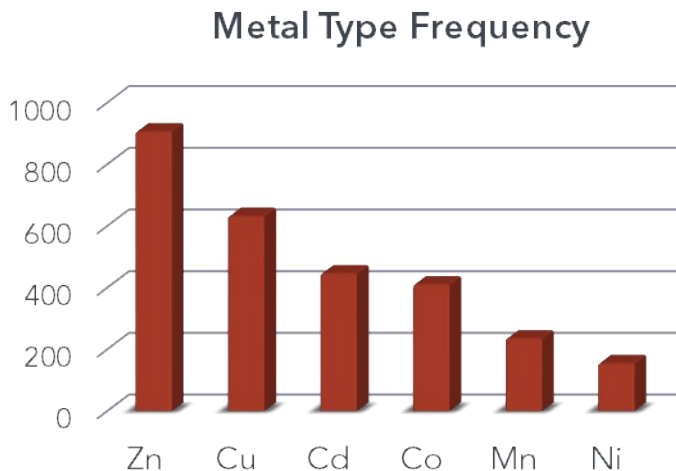


Figure 4 Number of MOFs with given metal type observed in the database used in this study

c. Algorithm parameters

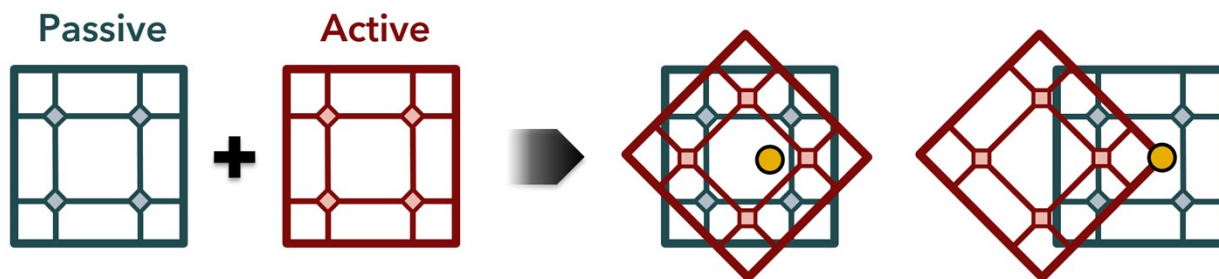


Figure 5 Interpenetration process for two cubic MOFs. Blue framework is the passive framework and red framework is the active framework. The active framework is first rotated and then translated to next point depicted by the yellow point.

We defined several algorithm parameters to introduce flexibility for various crystal structure pairs.

Grid size: The grid size is defined as the distance between grid points of the *energy map* in each dimension. The *energy map* is generated as a 3D rectangular grid that surrounds the unit cell. For high-throughput screening we used grid size of 1 Å for the generation of *energy map*.

Interpolation: When an energy value is needed at a position that does not fall exactly on a grid point in the *energy map*, which is the usual case, interpolation is used. For high-throughput screening we used trilinear interpolation which approximates the value of a point in a regular grid linearly using data in the lattice points.

Cut-off distance: The cut-off distance used in the calculation of interatomic potentials. Any atom pair that are further away from this distance were assumed to have no interaction. For high-throughput screening cut-off distance was taken as 12 Å.

Extension cut-off distance: The extension cut-off distance is used to determine the number of extension that will be performed on the unit cell for multiple unit cell collision test. For a given distance the unit cell is extended in each dimension to make sure that it covers an area at least the size of a sphere with radius of that distance. For high-throughput screening extension cut-off was taken as 50 Å.

Rotational freedom: The increments of rotation performed on the active unit cell to obtain different orientations. For high-throughput screening only 90° rotations were considered.

Rotational limit: The number of times a rotation is performed for same initial coordinate. For high-throughput screening all possible 90° rotations (24 total) were performed for each point.

Energy limit: There are three types of energy limit defined in the algorithm: atom energy limit, structure energy limit, and energy density limit. The energy limit parameter affects the allowed contact distance between the frameworks, therefore using high energy limits may result in unphysical contact distances.

- Atom energy limit is the maximum allowed energy for insertion of a single atom.
- Structure energy limit is the maximum allowed energy for insertion of a collection of atoms that constitute the active structure. It is calculated by summing the insertion energies for each atom in the unit cell.

- Energy density limit is similar to structure energy limit however it is divided by the unit cell volume to make it more comparable among different structures. For unit cells with different sizes constituting different number of atoms we chose to have a universal energy limit scaled by the volume of the cell. The energy density, ρ_{energy} , is calculated according to Eq. 2,

$$\rho_{energy} = \frac{\sum_{i=0}^{N_p} \sum_{j=0}^{N_a} V_{Lj}(i, j)}{V_{cell}} \quad \text{Eq. 2}$$

where N_p and N_a are number of atoms in active and passive MOFs; i and j are atom indices for passive and active MOFs; V_{Lj} is interatomic potential (Lennard-Jones) energy, and V_{cell} is unit cell volume in \AA^3 .

There are additional algorithm parameters available for user to manipulate how the structures are generated, how the data is reported etc. however these parameters do not affect the interpenetration.

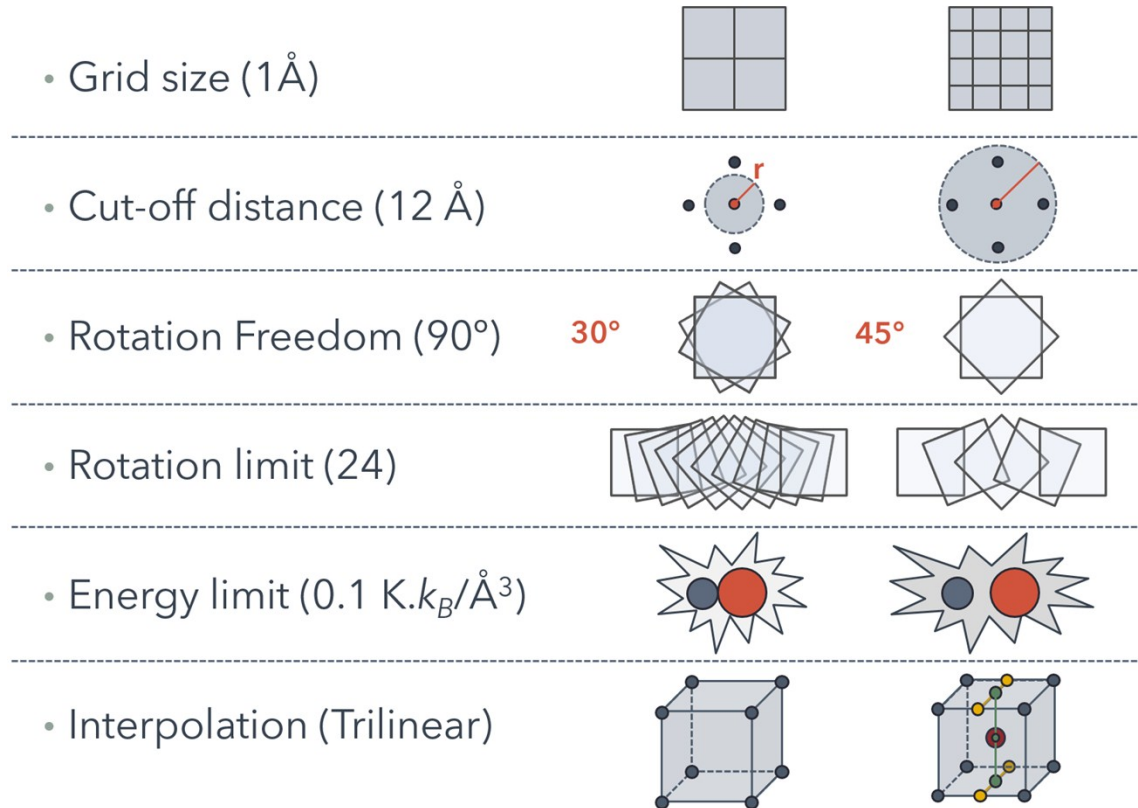


Figure 6 Algorithm parameters with values used in this study shown in parenthesis. Energy density limit of 0.1 K.k_B/ \AA^3 corresponds to 0.83144621 J/mol \AA^3 .

More information on algorithm parameters can be found in algorithm documentation.

For the high-throughput screening of the MOF database we selected a set of simulation parameters that we believe would yield a good amount of successful trials in a reasonable amount of time. In theory, it is possible to discover interpenetrated structures that were not found in this screening. Our aim in this study was to introduce a new method that allows testing whether two given periodic

structures can interpenetrate each other. The screening study performed here was performed to make an initial estimate for possible interpenetrations rather than exploring the whole possible orientations for two given MOFs. Parameters used in this study are given in excel file in Supplementary Information.

2. Generating Unit Cell for Interpenetrating Structures

a. Calculating supercell parameters

The lattice parameters for the combined unit cell are calculated by several geometric operations:

1. First $[111]$ vector for the active unit cell is converted to fractional coordinates using lattice vectors of passive unit cell. In other words, the $[111]$ vector of passive unit cell is converted to coordinate system of the passive unit cell. This vector represents how active unit cell repeats in terms of the passive unit cell. To make sure both unit cells repeat symmetrically in a supercell we need to calculate new lattice parameters.
2. The lattice parameters are calculated by finding the least common integer multiple for each dimension of the fractional $[111]$ vector. This gives the number of repeats required for the active unit cell. Multiplying these with fractional coordinates of $[111]$ vector gives number of repeats required for the passive unit cell.
3. The unit cells are aligned and repeated (extended) according to calculated parameters and the supercell is generated using the supercell parameters which can be calculated by multiplying number of repeats for passive unit cell with its original unit cell parameters.

Example case:

- Find $[111] \rightarrow [50, 20, 12]$
- Convert to fractional
 - $[5/2, 2/3, 4/5]$
- Find least common multiple
 - $[2, 3, 5]$
 - $[2, 3, 5] \times [5/2, 2/3, 4/5] \rightarrow [5, 2, 4]$
- Extend both unit cells to the common unit cell
 - **Passive** $\rightarrow [5, 2, 4] \rightarrow [100, 60, 60]$
 - **Active** $\rightarrow [2, 3, 5] \rightarrow [100, 60, 60]$

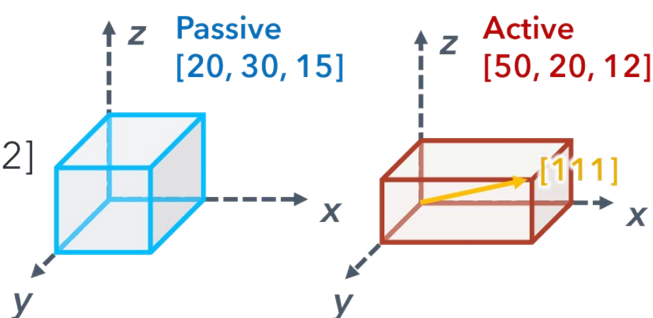


Figure 7 Example case for supercell parameters calculation

Since lattice parameters have high resolution (4,5 significant figures) the combined unit cells would generally become huge and impractical. To avoid that we allowed the active unit cell to tolerate lattice

parameter changes up to 1%. This allows us to constrain the combined unit cell to 100 repeats for any of the unit cells. Moreover, to make sure the active unit cell conserves its symmetry, we scaled the positions of the atoms according to adjustment we made to the lattice parameters. For example, if the a vector is enlarged by 0.3% then all distances in that direction are multiplied by 1.003.

b. Cell distortion

The cell distortion is defined as the scaling required to perfectly combine unit cells with a selected tolerance. Cell distortion is calculated by summing percent change in atomic coordinates for each vector of the unit cell. This means with a given 1 % tolerance the total cell distortion is less than or equal to 3 %. For the calculation of supercell parameters 1 % tolerance was used to ensure low cell distortion. However, candidate structures with very large supercells were not generated. It is possible to increase the tolerance and obtain a smaller supercell. For void fraction calculations reported in the manuscript, two of the candidate structures were generated using 2 % tolerance which resulted in a much smaller supercell with a total of ~ 4 % cell distortion. Example jupyter notebook for supercell generation is provided in the GitHub repository (github.com/WilmerLab/IPMOF).

3. High-throughput Screening

a. Database selection

For high-throughput MOF database created by Chung and co-workers⁴ has been used. In addition to the structures proposed in this study, 1005 additional MOFs were added to the final database (total 6014 MOFs). The additional MOFs have been generated using the same methodology explained in that study.

MOF Combination Selection

To eliminate MOF combinations that are physically capable of resulting in interpenetration, only MOF combinations that have a total void fraction bigger than 1.0 are considered. To explain this better, we can imagine a unit cell of a MOF as a glass of water where the atoms in the unit cell would be analogous to water in glass (Fig. 8). We can also imagine interpenetration as combining these two glasses of water into a single glass (or single unit cell). Thus, to have enough space for interpenetration the total water amount in the glasses should be less than 100 % or in the MOF case the total void fraction of the MOFs should be bigger than 1.0.

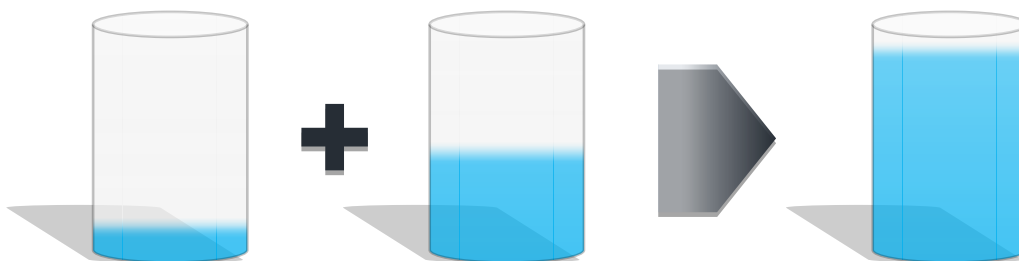


Figure 8 Analogy for MOF combination selection. The interpenetrating MOFs must have enough empty space in their unit cell so when they form a single unit cell there is enough space for both frameworks. This can be imagined as filling two glasses of water into a single glass. In order to avoid overflowing the glasses should have enough

The total possible combinations correspond to combination of two of 6014 which comes to around 18.1 million combinations. When the physically impossible combinations are eliminated the total number of combinations come down to around 4.28 million which is a little less than quarter of total possible combinations.

$$\binom{6014}{2} = \sim 18.1 \text{ million} \xrightarrow{V_{f1} + V_{f2} > 1.0} \sim 4.3 \text{ million}$$

b. Void fraction calculation

The void fraction of the MOFs in this study are calculated using RASPA 2⁵, a general purpose classical simulation package that can be used for the simulation of molecules in gases, fluids, zeolites, aluminosilicates, metal-organic frameworks, carbon nanotubes and external fields. <https://github.com/numat/RASPA2>

c. Analysis of candidate structures

Initially single unit-cell interpenetration tests were performed for ~4.3 million MOF pairs and ~6% of these pairs resulted in favorable orientations. The extended unit-cell test on these MOF pairs resulted in a total of 1045 MOF pairs which were identified as candidate interpenetrating structures (Fig. 9).

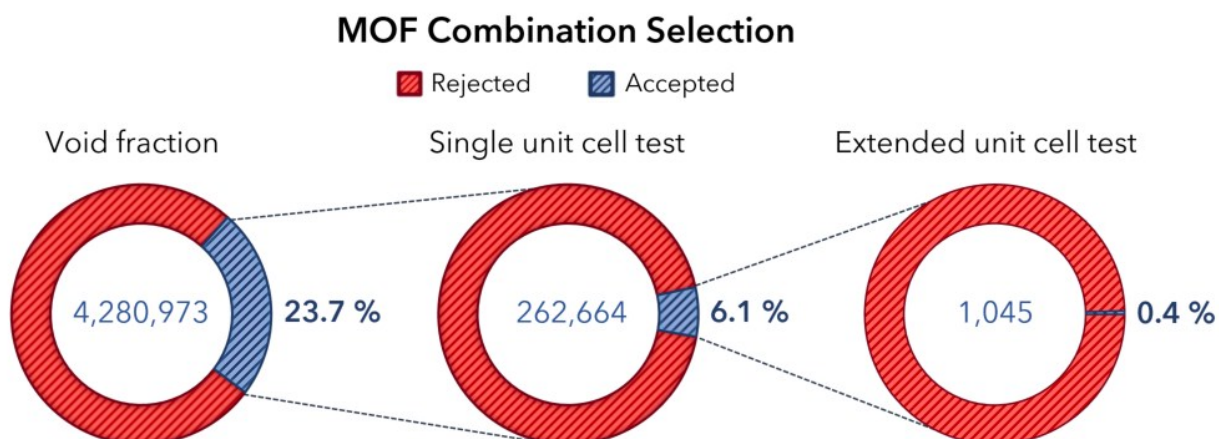


Figure 9 Number of MOF pairs in different steps of high-throughput screening

Further analysis of these MOF pairs showed that most of these structures were homo-interpenetrated due to degenerate MOFs in the database. Total 113 MOFs were identified to form 1045 candidate structures and among these 113, 46 of them were found to be MOF-5 degenerate structures. Among the 1045 candidate structures, 1012 of them were identified as homo-interpenetrating where 92 of them were determined as pairs with same CCDC reference codes and 922 of them were determined as pairs of degenerate MOFs under different reference codes. 33 hetero-interpenetrating candidates were examined and 15 of them were found to be duplicate pairs resulting in 18 unique pairs of hetero-interpenetrating MOFs (see Figure 10). Among the hetero-interpenetrating candidates following topologies were observed: *pcu+pcu*, *the+tbo*, *ftw+ftw*, *tbo+ntt*, *scu+tfe*, *reo-e+ftw*, *gee+ftw*, *ntt+pcu*, *tbo+pcu*. More information about all MOFs and all candidate structures are provided in excel file in Supplementary Information.

For candidate hetero-interpenetrated structures, supercells were calculated with 1 % distortion tolerance (given in SI excel file) and cif files were generated. The candidates which resulted in unpractically large supercells were generated with higher tolerance resulting in higher cell distortion. The candidate crystallographic information files (cif) are provided in GitHub repository (github.com/WilmerLab/IPMOF). In addition to regular structure files each candidate also has a colored version (for easy visual inspection) where each atom in the interpenetrating frameworks are replaced with carbon and oxygen atom for passive and active MOFs respectively.

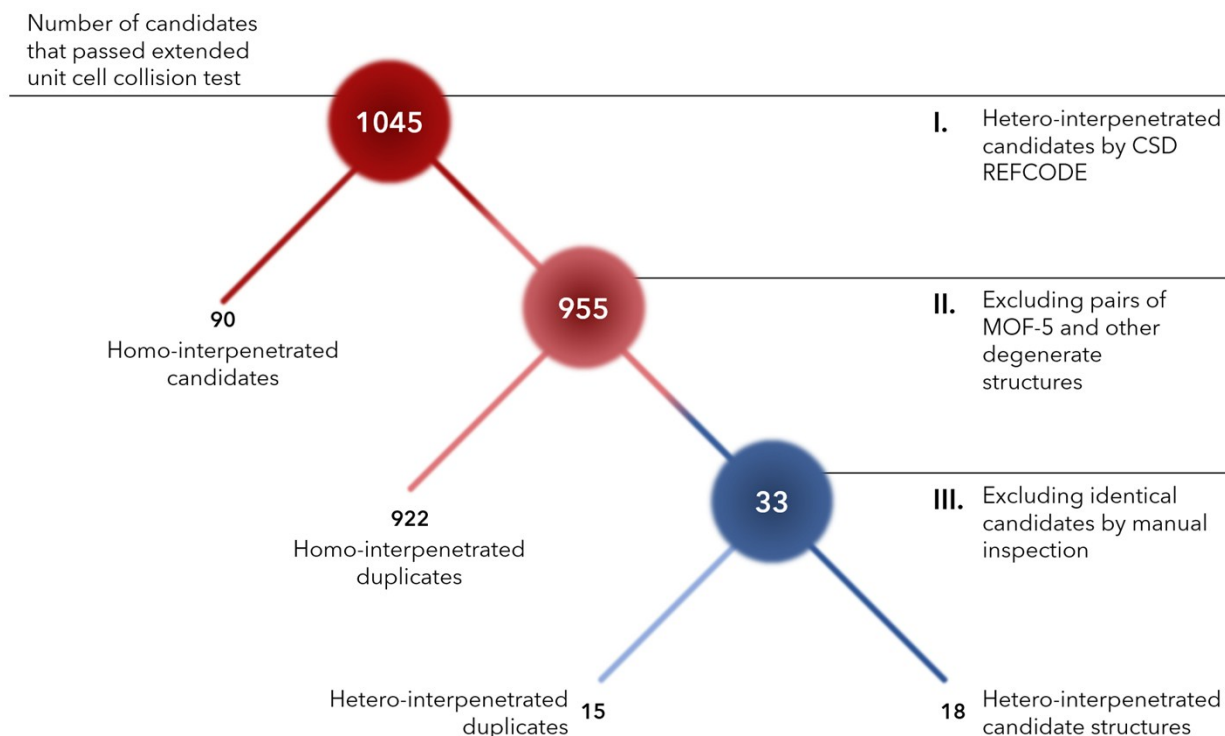
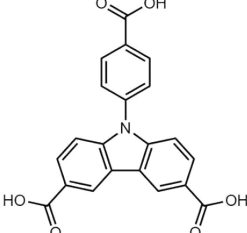
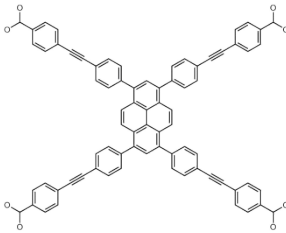
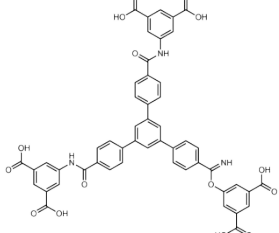
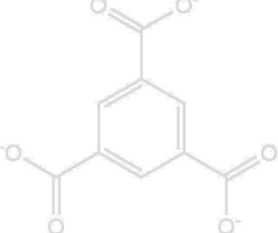
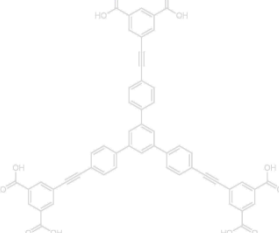
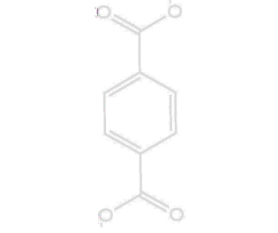
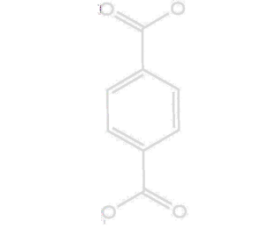
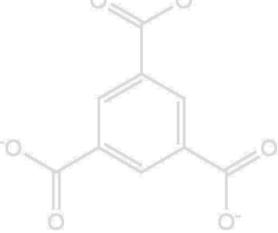
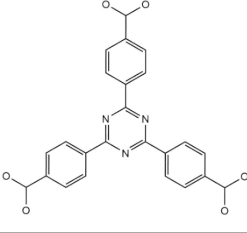
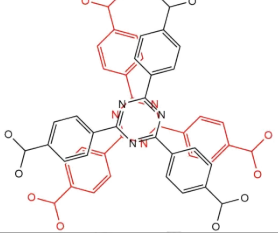
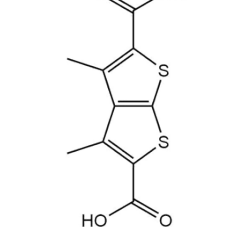
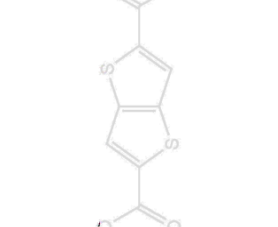


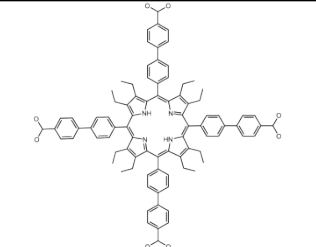
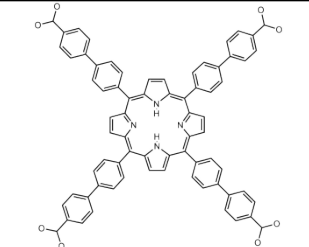
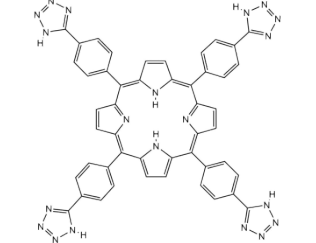
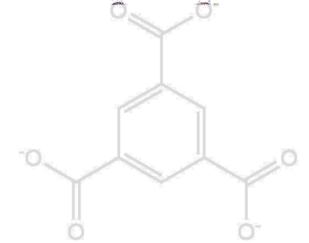
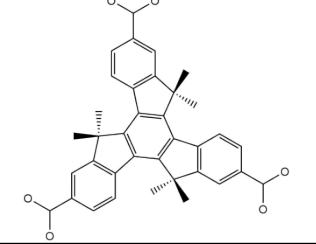
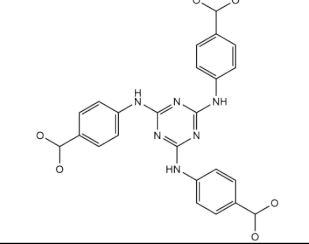
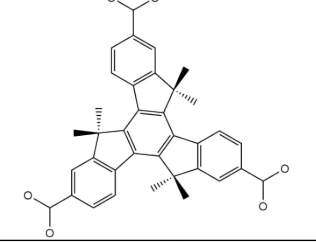
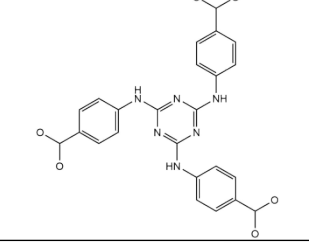
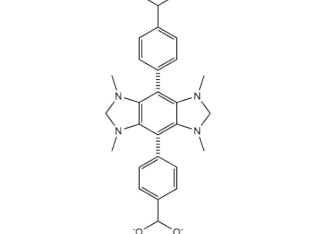
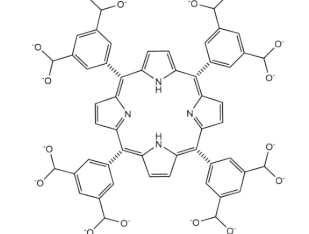
Figure 10 Analysis of candidate structures to find hetero-interpenetrated pairs (Starting from 1,045 candidates that passed the extended collision test, 90 of them were pairs of same MOFs according to their CSD reference codes **[Step I]**. From the remaining 955 candidates, pairs of MOFs that have different reference codes but are degenerate structures (922) are removed **[Step II]**. Remaining 33 candidates were inspected manually by comparing chemical formulas, crystallographic information and visually examining image renders of MOFs **[Step III]**. 15 more candidates were removed and 18 candidates were presented as hetero-interpenetrated candidate MOFs.

Additionally, experimental synthesis procedure and chemical composition of the 18 candidate structures were analyzed. Metal centers and organic linkers that compose each MOF are presented below with crystallographic information. Moreover, experimental synthesis information such as used reagents, solvents, temperature are provided in the SI excel file.

Table 2 Chemical composition information for 18 hetero-interpenetrating MOF candidates

#	Candidate	Formula*	Supercell	Linker 1	Linker 2
1	AXUBAW [IRMOF-18] + NODTIL [IRMOF-1]	ZnTMBDC [C ₇₂ H ₇₂ O ₂₆ Zn ₈] + ZnBDC [C ₄₈ H ₂₄ O ₂₆ Zn ₈]	[1 1 1] + [1 1 1] D: 1.1072		
2	BAZGAM [MOF-399] + SEMNIJ [NU-110]	CuBBC [C ₃₆₀ H ₂₁₆ Cu ₁₂ O ₄₈] + Cu L [C ₆₂₄ H ₂₈₈ Cu ₂₄ O ₉₆]	[1 1 1] + [1 1 1] D: 1.7241		
3	HABQUY [PCN-610] + VEBHUG [IRMOF-20]	CuTTEI [C ₄₈₀ H ₁₉₂ Cu ₂₄ O ₉₆] + ZnTTDC [C ₄₈ H ₁₂ O ₂₆ S ₁₂ Zn ₈]	[8 8 22] + [49 49 45] D: 2.3537		
4	HABRAF [PCN-68] + HANHAH [ZnBTC]	CuPTEI [C ₄₃₂ H ₁₉₂ Cu ₂₄ O ₉₆] + ZnBTC [C ₇₂ H ₂₄ O ₄₈ Zn ₁₂]	[1 1 1] + [2 2 2] D: 0.8743		
5	HABRAF [PCN-68] + XADDIR [CuBTC]	CuPTEI [C ₄₃₂ H ₁₉₂ Cu ₂₄ O ₉₆] + CuBTC [C ₇₂ H ₂₄ Cu ₁₂ O ₄₈]	[1 1 1] + [2 2 2] D: 0.2717		
6	IZEPAF [MIL-hypo-1] + WUHDAG [NU-1104]	CrBDC [C ₂₈₈ H ₁₄₄ Cr ₃₆ O ₁₅₆] + ZrPor-PTP [C ₂₄₀ H ₁₃₀ N ₁₂ O ₃₂ Zr ₆]	[43 43 43] + [42 42 42] D: 2.8926		
7	JUTCUW + PIYZAZ	[Cd(CN) ₂] + [Cd(C ₂) ₂]	[1 1 1] + [1 1 1] D: 1.9544	$C \equiv N$	$C \equiv C$

8	KOZQEX [DUT-75] + WUHCUZ [NU-1103]	CuPCDC [C ₂₅₂ H ₁₂₀ Cu ₁₈ N ₁₂ O ₇₂] + ZrPy-PTP [C ₂₂₈ H ₁₁₈ O ₃₂ Zr ₆]	[1 1 1] + [1 1 1] D: 0.7444		
9	LEHXUT + XAMDUM02 [CuBTC]	CuBBC [C ₄₀₈ H ₂₁₆ Cu ₂₄ N ₂₄ O ₁₂₀] + CuBTC [C ₇₂ H ₂₄ Cu ₁₂ O ₄₈]	[1 1 1] + [2 2 2] D: 0.1275		
10	LURRIA [PCN-68] + MIBQAR06 [IRMOF-1]	CuPTEI [C ₄₃₂ H ₁₉₂ Cu ₂₄ O ₉₆] + ZnBDC [C ₄₈ H ₂₄ O ₂₆ Zn ₈]	[1 1 1] + [2 2 2] D: 2.3519 D: 0.1149		
11	PEVPUD [IRMOF-1] + XAMDUM04 [CuBTC]	ZnBDC [C ₄₈ H ₂₄ O ₂₆ Zn ₈] + CuBTC [C ₇₂ H ₂₄ Cu ₁₂ O ₄₈]	[1 1 1] + [1 1 1] D: 2.3791 D: 2.9419		
12	PIBNUK [ZnTATB] + PIBPIA [CdTATB]	ZnTATB [C ₁₉₂ H ₉₆ N ₂₄ O ₄₈ Zn ₁₂] + CdTATB [C ₁₉₂ H ₉₆ Cd ₁₂ N ₂₄ O ₄₈]	[43 43 43] + [42 42 42] D: 2.8962		
13	TEQPEN [IFMC-28] + VEBHUG [IRMOF-20]	ZnDMTDC [C ₆₀ H ₃₆ O ₂₆ S ₁₂ Zn ₈] + ZnTTDC [C ₄₈ H ₁₂ O ₂₆ S ₁₂ Zn ₈]	[1 1 1] + [3 3 1] D: 0.1337		

14	TOVGAO [PCN-228] + JOZWIG02 [NU-1102]	ZrTCP-1 [C ₂₁₆ H ₁₂₄ N ₁₂ Ni ₃ O ₃₂ Zr ₆] + ZrPor-PP [C ₂₁₆ H ₁₃₀ N ₁₂ O ₃₂ Zr ₆]	[1 1 1] + [1 1 1] D: 0.1516		
15	UQOFOX [PCN-528] + VEHJUP	MnTTPP [C ₉₆ H ₅₆ Mn ₆ N ₄₀] + Zn3,4 BTC [C ₁₀₈ H ₃₆ O ₇₂ Zn ₂₂]	[23 1 1] + [22 23 1] D: 1.5778 D: 1.1833		
16	VETSUK [POST-65(Co)] + HEXVEM [mesoMOF-1]	CoHMTT [C ₅₇₆ H ₄₃₂ Cl ₁₂ Co ₂₆ O ₁₄₄] + CuTATAB [C ₁₉₂ H ₁₂₀ Cu ₁₂ N ₄₈ O ₄₈]	[1 1 1] + [1 1 1] D: 1.2227		
17	VETTAR [POST-65(Ni)] + HEXVEM [mesoMOF-1]	NiHMTT [C ₅₇₆ H ₄₃₂ Cl ₁₂ Ni ₂₆ O ₉₆] + CuTATAB [C ₁₉₂ H ₁₂₀ Cu ₁₂ N ₄₈ O ₄₈]	[1 1 1] + [1 1 1] D: 0.7567		
18	RUTNOK [IRMOF-76] + VEJHEZ [MMPF-4]	Zn LO [C ₆₉ H ₄₅ N ₆ O ₁₃ Zn ₄] + ZnTDCPP [C ₁₅₆ H ₆₀ N ₁₂ O ₅₀ Zn ₁₉]	[1 1 1] + [1 1 1] D: 0.7355		

Notes:

- The red color represents the MOFs which can also *homo-interpenetrate*. The candidates where both MOFs are black in color are more likely to form *hetero-interpenetration* since *in theory* both MOFs don't interpenetrate with themselves.
- The formulas are just simplification of metal types and linkers. Even though two MOFs have the same metal type and linker (e.g. UQOFOX and HANHAH) they could have different structures.
- Supercell column has information on number of extensions required to form supercell for each of the MOFs. (e.g. [1 1 1] + [2 2 2] means MOF1 needs to be extended to [1 1 1] and MOF needs to be extended to [2 2 2] in x, y, z dimensions to form the supercell). The D value gives the distortion for supercell in percentage.
- Hashed bonds imply 90 degree rotations around the bonding axis.

4. Comparison with Literature

a. Comparison of candidate structures

An example comparison of results between our method and another method proposed by Kwon et al.⁶ is provided. In their work, Kwon and co-workers developed an algorithm that can identify interpenetration of MOFs with orthogonal unit cells. They used their algorithm to screen the CoRE⁴ database and provided a list of 113 candidates in Table S2 of their Supplementary Information. Comparing this list with our candidates list, we found out that 16 candidates included MOFs that weren't present in our database of 6014 MOFs. From the rest of 97 candidates, 45 of them were also present in the candidate list we provided. For the 52 candidates remaining, our calculations didn't find any interpenetrating pairs due to our selection of simulation parameters.

In the database used in this study, almost all MOFs are reported in their primitive unit cell. For cases where we are testing interpenetration of MOFs with non-orthogonal unit cells, 90-degree rotations may be insufficient to find the interpenetration. In the study by Kwon et al. they first converted all primitive unit cells of orthogonal MOFs to their conventional orthogonal cells and tested interpenetration after that. If we were to do the same before running interpenetration tests, 90-degree rotations would be sufficient to find the same pairs as well. However, using primitive unit cells, smaller rotational increments are necessary to find the interpenetrating configurations. It is important to point out that the frameworks obtained depend not just on the method, but also the parameters used. In this paper, we have described a rather general method, and as a demonstration we show results we obtained from choosing a *particular* set of parameters. Any reader can use our code, which we make freely available, and change the parameters to better suit their interests. Below we show how some additional structures were found by adding using smaller rotational increments.

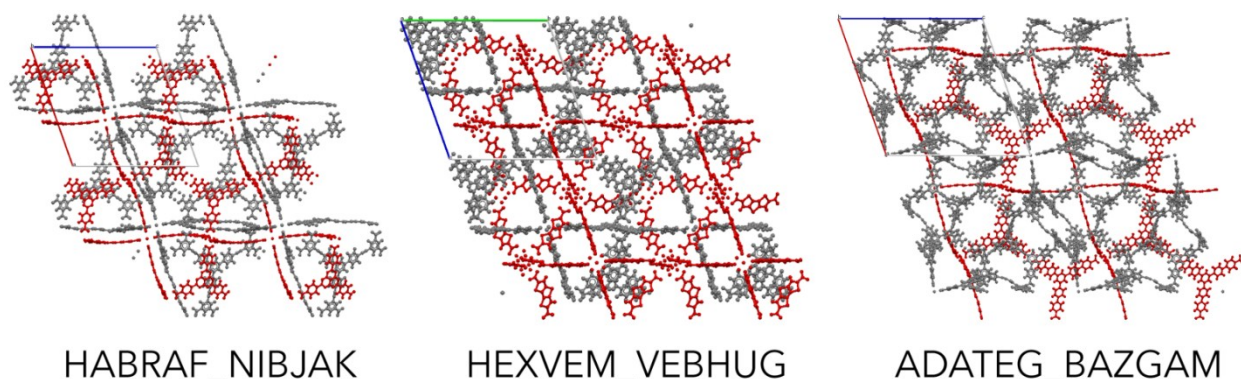


Figure 11 Candidate interpenetrating MOF pairs identified using 30-degree rotations. Same pairs are also present in the list provided by Kwon et al.⁶ CCDC⁷ reference codes for the MOF pairs are given below structures separated by an underscore. The first name corresponds the gray framework and second name corresponds to the red framework.

We performed interpenetration tests for three MOF pairs from the list by Kwon et al. and we obtained interpenetrated structures for [HABRAF_NIBJAK], [HEXVEM_VEBHUG], and [ADATEG_BAZGAM] using 30-degree rotations (see Fig. 11). These MOF pairs were not identified in the original screening since they have trigonal unit cells where the interpenetrated configuration cannot be obtained using 90-degree rotations without orthogonalizing the unit cells.

b. Comparison of experimental hetero-interpenetrated structures

In addition to WEBZEK⁸ provided in the article, hetero-interpenetration test was performed for 5 more MOFs using the algorithm. The crystal structure files were downloaded from the CCDC⁷ database and the interpenetrating frameworks were separated. Then for each MOF the algorithm was ran using interpenetrating layers and resulting structures with lowest energy density were selected. Using the cif files provided by the algorithm X-ray diffraction (XRD) patterns were generated using Mercury⁹ software. Moreover, the experimental XRD patterns were gathered from each study and results were compared. Both XRD patterns and visual inspection confirmed the high similarity between predicted and experimental structure. Here unit cell images and XRD patterns for these 5 MOFs are provided (Fig. 10 – Fig. 14). In each figure, the XRD pattern with black line at the bottom is for prediction and the other color at the top is for experimental result. The yellow and blue boxes point out the subtle differences between structures and arrows show the directional difference. The interpenetrating frameworks are colored in gray and red for visual aid.

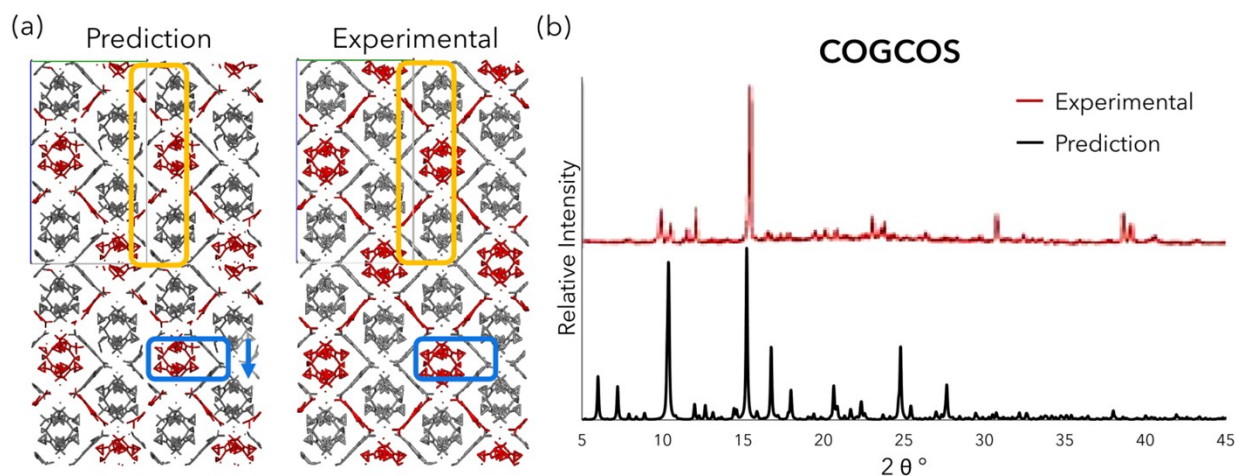


Figure 12 Hetero-interpenetration test for COGCOS¹⁰ (a) Structure comparison (b) XRD pattern comparison

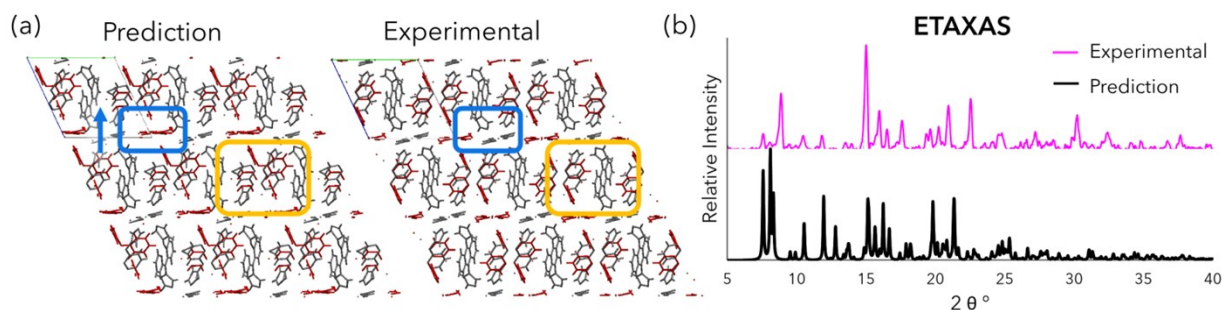


Figure 13 Hetero-interpenetration test for ETAXAS¹¹ (a) Structure comparison (b) XRD pattern comparison

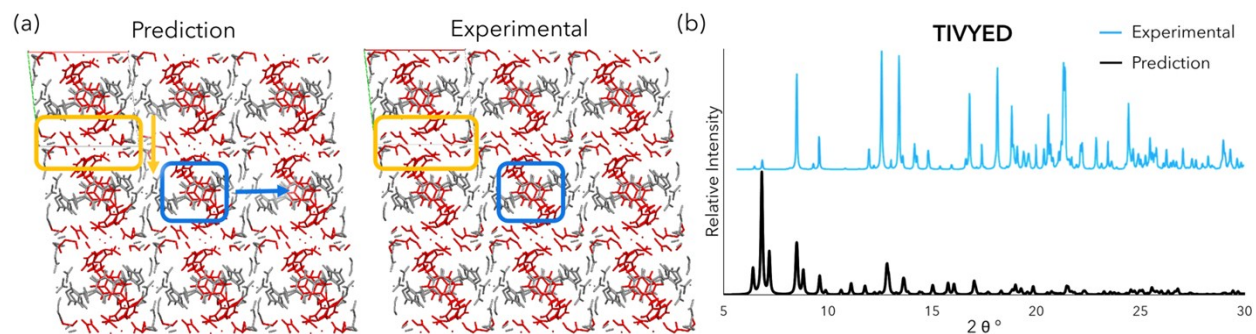


Figure 14 Hetero-interpenetration test for TIVYED¹² (a) Structure comparison (b) XRD pattern comparison

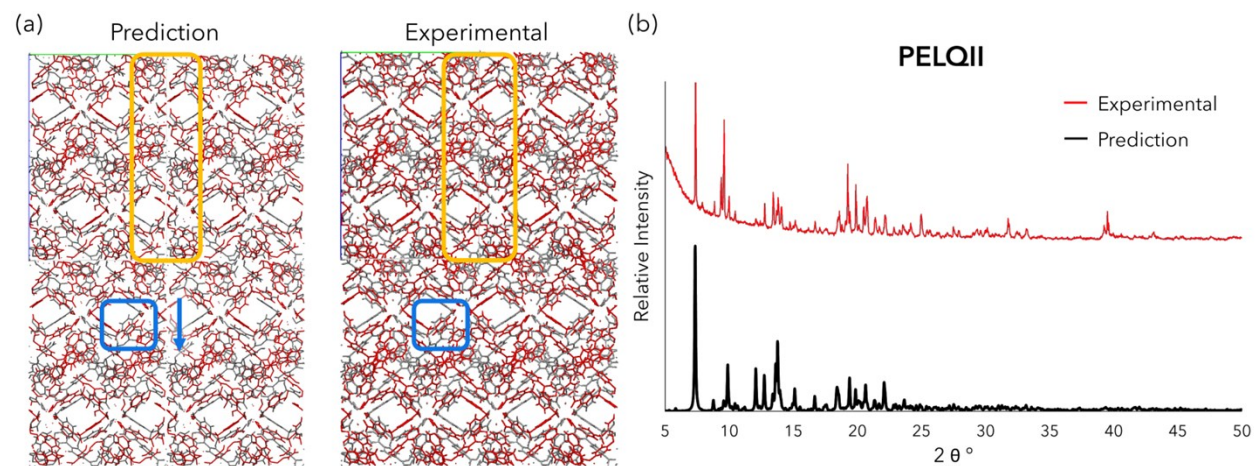


Figure 15 Hetero-interpenetration test for PELQII¹³ (a) Structure comparison (b) XRD pattern comparison

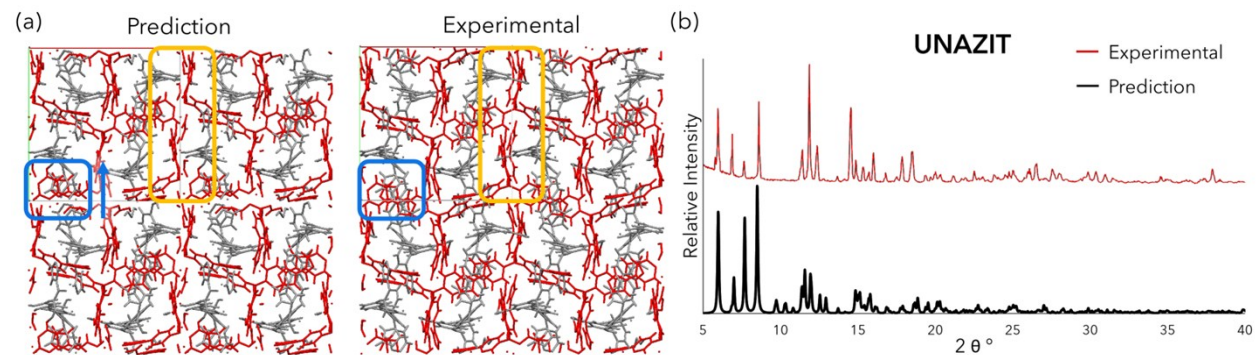


Figure 16 Hetero-interpenetration test for UNAZIT¹⁴ (a) Structure comparison (b) XRD pattern comparison

References

- 1 J. E. Lennard-Jones, *Proc. Phys. Soc.*, 1931, **43**, 461.
- 2 A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
- 3 H. Li, M. Eddaoudi, M. O’Keeffe and O. M. Yaghi, *Nature*, 1999, **402**, 276–279.
- 4 Y. G. Chung, J. Camp, M. Haranczyk, B. J. Sikora, W. Bury, V. Krungleviciute, T. Yildirim, O. K. Farha, D. S. Sholl and R. Q. Snurr, *Chem. Mater.*, 2014, **26**, 6185–6192.
- 5 D. Dubbeldam, S. Calero, D. E. Ellis and R. Q. Snurr, *Mol. Simul.*, 2016, **42**, 81–101.
- 6 O. Kwon, S. Park, H.-C. Zhou and J. Kim, *Chem. Commun.*, 2017, **53**, 1953–1956.
- 7 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.*, 2016, **72**, 171–179.
- 8 H. Xu, W. Bao, Y. Xu, X. Liu, X. Shen and D. Zhu, *CrystEngComm*, 2012, **14**, 5720–5722.
- 9 C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466–470.
- 10 X.-G. Guo, W.-B. Yang, X.-Y. Wu, L. Lin and C.-Z. Lu, *Eur. J. Inorg. Chem.*, 2014, **2014**, 2481–2485.
- 11 H.-J. Lee, P.-Y. Cheng, C.-Y. Chen, J.-S. Shen, D. Nandi and H. M. Lee, *CrystEngComm*, 2011, **13**, 4814–4816.
- 12 L. Carlucci, G. Ciani, S. Maggini and D. M. Proserpio, *Cryst. Growth Des.*, 2008, **8**, 162–165.
- 13 M.-D. Zhang, C.-M. Di, L. Qin, Q.-X. Yang, Y.-Z. Li, Z.-J. Guo and H.-G. Zheng, *CrystEngComm*, 2012, **15**, 227–230.
- 14 X.-Q. Yao, D.-P. Cao, J.-S. Hu, Y.-Z. Li, Z.-J. Guo and H.-G. Zheng, *Cryst. Growth Des.*, 2011, **11**, 231–239.