Machine learning transferable atomic forces for large systems from underconverged molecular fragments

S-I Fragment Radii for Atomic IRMOF Positions

The range of atomic interactions for the different atomic sites with the surrounding atoms present in the IRMOF bulk structures (Fig. 2) can be analyzed employing the Hessian[47]. The Hessian contains the derivatives of the atomic forces with respect to all atomic positions, which can be used to determine environment radii for all atomic sites (Tab. SI). For each atomic site, i.e., the central atom, a molecular fragment can be constructed based on the fragment radius. Within the molecular fragments, the specific central atoms are embedded in a bulk like environment up to the given fragment radius. Consequently, the atomic forces of the central atoms are converged to the bulk force values. To ensure the atomic force convergence of every atomic site and an equivalent description of the atomic environments, the largest fragment radius as a worst case scenario forms the basis of the non-redundant set of molecular fragments (Fig. 4).

Table SI Fragment radii R_{frag} in Å obtained by the Hessian-based determination of the atomic interactions[47] for the atomic sites present in the IRMOF bulk structures. Dashes indicate atomic sites, which are not present in the respective IRMOFs. For site H2 in IRMOF-10, as well as for sites C5, H2 and H3 in IRMOF-16, the considered fragment (biphenyl) is already size-converged. For this reason, the fragment radii are small. Thus, a further reduction of the fragment size during the analysis is not required. The given fragment radii for these cases are given in italic.

A	IRMOF-1	IRMOF-10	IRMOF-16
Zn1	4.333	4.328	4.333
O1	5.165	5.160	5.158
O2	2.379	2.376	2.375
C1	8.502	8.718	6.316
C2	7.304	4.328	4.327
C3	3.817	3.830	3.827
C4	–	3.816	4.811
C5	–	4.364	*1.492*
C6	–	–	4.389
C7	–	–	3.830
H1	2.725	2.731	2.730
H2	–	*1.962*	*1.975*
H3	–	–	*1.975*

S-II Atom-Centered Symmetry Functions (ACSFs)

Table SII Compilation of the radial atom-centered symmetry functions (ACSFs)[64] for the different element combinations employed for the construction of HDNNP1 and HDNNP2. A defines the element of the atom at which the ACSF is centered and B the element of the neighboring atom. The η parameter defining the width of the Gaussian is given in a_0^{-2} with a_0 being the Bohr radius. The cutoff radius $R_{\mathrm{c}} = 4.359\,$Å has been used.

	carbon			oxygen			zinc			hydrogen		
no.	A	B	η	A	B	η	A	B	η	A	B	η
1	C	H	0.000	O	H	0.000	Zn	C	0.000	H	H	0.000
2	C	C	0.000	O	C	0.000	Zn	O	0.000	H	C	0.000
3	C	O	0.000	O	O	0.000	Zn	Zn	0.000	H	O	0.000
4	C	Zn	0.000	O	Zn	0.000	Zn	Zn	0.001	H	C	0.007
5	C	Zn	0.002	O	O	0.003	Zn	C	0.002	H	O	0.010
6	C	Zn	0.003	O	Zn	0.004	Zn	Zn	0.002	H	H	0.014
7	C	Zn	0.005	O	C	0.006	Zn	C	0.003	H	C	0.020
8	C	C	0.006	O	O	0.007	Zn	Zn	0.003	H	O	0.041
9	C	O	0.006	O	Zn	0.009	Zn	O	0.004	H	C	0.053
10	C	H	0.007	O	H	0.010	Zn	Zn	0.004	H	H	0.079
11	C	Zn	0.007	O	O	0.013	Zn	C	0.005	H	C	0.182
12	C	C	0.015	O	C	0.016	Zn	C	0.007			
13	C	O	0.016	O	Zn	0.017	Zn	O	0.009			
14	C	H	0.020	O	O	0.021	Zn	O	0.017			
15	C	C	0.034	O	Zn	0.031	Zn	O	0.031			
16	C	O	0.039	O	C	0.039						
17	C	H	0.053	O	H	0.041						
18	C	C	0.084	O	C	0.106						
19	C	O	0.106									
20	C	H	0.182									

Table SIII Compilation of the angular atom-centered symmetry functions (ACSFs)[64] for the different element combinations employed in the construction of HDNNP1 and HDNNP2. A defines the element of the atom at which the ACSF is centered, while B and C specify the elements of the neighboring atoms. The η parameter defining the width of the Gaussian has been set to zero, and the cutoff radius is $R_c = 4.359$ Å. The set of angular ACSFs of each element combination consists of all permutations of the parameter $\zeta \in \{1, 2, 4, 16\}$, defining the width of the angular part (see Ref. 64) and the parameter $\lambda \in \{-1, 1\}$ controlling the positions of the maxima. For the element combinations C-Zn-Zn, Zn-C-C, Zn-C-Zn, Zn-O-Zn and Zn-Zn-Zn the angular ACSF with the parameter combination of $\zeta = 16$ and $\lambda = -1$ are not included.

	carbon			oxygen			zinc			hydrogen		
no.	*A*	*B*	*C*	*A*	*B*	*C*	*A*	*B*	*C*	*A*	*B*	*C*
1	C	H	H	O	H	C	Zn	C	C	H	H	C
2	C	H	C	O	H	O	Zn	C	O	H	C	C
3	C	H	O	O	C	C	Zn	C	Zn	H	C	O
4	C	H	Zn	O	C	O	Zn	O	O	H	O	O
5	C	C	C	O	C	Zn	Zn	O	Zn			
6	C	C	O	O	O	O	Zn	Zn	Zn			
7	C	C	Zn	O	O	Zn						
8	C	O	O	O	Zn	Zn						
9	C	O	Zn									
10	C	Zn	Zn									

Table SIV Compilation of the radial atom-centered symmetry functions (ACSFs) for the different element combinations of the ACSF set for HDNNP3. A defines the element of the atom at which the ACSF is centered and B the element of the neighboring atom. The η parameter defining the width of the Gaussian is given in a_0^{-2} with the cutoff radius $R_c = 8.718\,\text{Å} = 16.475\,a_0$, a_0 being the Bohr radius.

no.	A	B	η	A	B	η	A	B	η	A	B	η
1	C	H	0.000	O	H	0.000	Zn	C	0.000	H	H	0.000
2	C	C	0.000	O	C	0.000	Zn	O	0.000	H	C	0.000
3	C	O	0.000	O	O	0.000	Zn	Zn	0.000	H	O	0.000
4	C	Zn	0.000	O	Zn	0.000	Zn	C	0.001	H	C	0.001
5	C	H	0.001	O	C	0.001	Zn	O	0.001	H	C	0.002
6	C	C	0.001	O	O	0.001	Zn	Zn	0.001	H	C	0.003
7	C	O	0.001	O	Zn	0.001	Zn	C	0.002	H	C	0.004
8	C	Zn	0.001	O	C	0.002	Zn	O	0.002	H	C	0.006
9	C	H	0.002	O	O	0.002	Zn	Zn	0.002	H	C	0.007
10	C	C	0.002	O	Zn	0.002	Zn	C	0.003	H	H	0.010
11	C	O	0.002	O	C	0.003	Zn	O	0.003	H	C	0.010
12	C	Zn	0.002	O	O	0.003	Zn	Zn	0.003	H	C	0.012
13	C	H	0.003	O	Zn	0.003	Zn	C	0.004	H	O	0.013
14	C	C	0.003	O	C	0.004	Zn	O	0.004	H	C	0.016
15	C	O	0.003	O	O	0.004	Zn	Zn	0.004	H	C	0.021
16	C	Zn	0.003	O	Zn	0.004	Zn	C	0.005	H	C	0.027
17	C	H	0.004	O	C	0.005	Zn	O	0.005	H	C	0.036
18	C	C	0.004	O	O	0.005	Zn	Zn	0.005	H	C	0.049
19	C	O	0.004	O	Zn	0.005	Zn	C	0.006	H	C	0.070
20	C	Zn	0.004	O	O	0.006	Zn	O	0.006	H	H	0.082
21	C	C	0.005	O	Zn	0.006	Zn	Zn	0.006	H	C	0.102
22	C	O	0.005	O	C	0.007	Zn	C	0.007	H	C	0.161
23	C	Zn	0.005	O	O	0.007	Zn	O	0.007	H	O	0.194
24	C	H	0.006	O	Zn	0.007	Zn	Zn	0.007			
25	C	C	0.006	O	O	0.008	Zn	C	0.008			
26	C	Zn	0.006	O	Zn	0.008	Zn	O	0.008			
27	C	H	0.007	O	C	0.009	Zn	Zn	0.008			
28	C	O	0.007	O	O	0.009	Zn	C	0.009			
29	C	Zn	0.007	O	O	0.010	Zn	Zn	0.009			
30	C	C	0.008	O	Zn	0.010	Zn	C	0.010			
31	C	Zn	0.008	O	C	0.011	Zn	O	0.010			
32	C	O	0.009	O	O	0.011	Zn	Zn	0.010			
33	C	Zn	0.009	O	O	0.012	Zn	C	0.011			
34	C	H	0.010	O	Zn	0.012	Zn	Zn	0.011			
35	C	C	0.010	O	H	0.013	Zn	C	0.012			
36	C	Zn	0.010	O	C	0.014	Zn	O	0.012			
37	C	O	0.011	O	O	0.015	Zn	Zn	0.012			
38	C	Zn	0.011	O	Zn	0.015	Zn	C	0.013			
39	C	H	0.012	O	C	0.018	Zn	Zn	0.013			
40	C	Zn	0.012	O	O	0.018	Zn	C	0.014			

Table SV Continuation of Table SIV.

no.	A	B	η	A	B	η	A	B	η	A	B	η
41	C	C	0.013	O	Zn	0.018	Zn	Zn	0.014			
42	C	Zn	0.013	O	O	0.021	Zn	O	0.015			
43	C	O	0.014	O	Zn	0.022	Zn	Zn	0.015			
44	C	Zn	0.014	O	C	0.023	Zn	C	0.016			
45	C	H	0.016	O	O	0.025	Zn	Zn	0.016			
46	C	C	0.016	O	Zn	0.028	Zn	C	0.018			
47	C	Zn	0.016	O	C	0.030	Zn	O	0.018			
48	C	O	0.018	O	Zn	0.034	Zn	O	0.022			
49	C	Zn	0.018	O	C	0.039	Zn	O	0.028			
50	C	H	0.021	O	Zn	0.044	Zn	O	0.034			
51	C	C	0.021	O	C	0.052	Zn	O	0.044			
52	C	O	0.023	O	C	0.073						
53	C	H	0.027	O	C	0.105						
54	C	C	0.027	O	H	0.194						
55	C	O	0.030									
56	C	C	0.035									
57	C	H	0.036									
58	C	O	0.039									
59	C	C	0.046									
60	C	H	0.049									
61	C	O	0.052									
62	C	C	0.062									
63	C	H	0.070									
64	C	O	0.073									
65	C	C	0.086									
66	C	H	0.102									
67	C	O	0.105									
68	C	H	0.161									

Table SVI Compilation of the angular atom-centered symmetry functions (ACSFs) for the different element combinations in HDNNP3. A defines the element of the atom at which the ACSF is centered, B and C the element of the neighboring atoms. The η parameter defining the width of the Gaussian is defined as $\eta = 0.000\,\mathrm{a_0^{-2}}$. There are two shells of angular ACSF: one shell is similar to HDNNP1/HDNNP2 (Tab. SIII) with the cutoff radius $R_\mathrm{c} = 4.359\,\text{Å} = 8.237\,\mathrm{a_0}$ and a second shell with the cutoff radius $R_\mathrm{c} = 8.718\,\text{Å} = 16.475\,\mathrm{a_0}$. The set of angular ACSF of each element combination is expanded by all different combinations of the parameter $\zeta \in \{1, 2, 4, 16\}$, defining the width of the angular part and the parameter $\lambda \in \{-1, 1\}$. For the element combinations C-Zn-Zn, O-O-O, Zn-C-C, Zn-C-Zn, Zn-O-Zn, Zn-Zn-Zn and H-O-O the angular ACSFs with the parameter combination of $\zeta = 16$ and $\lambda = -1$ are not included.

no.	A	B	C	A	B	C	A	B	C	A	B	C
1	C	H	H	O	H	C	Zn	C	C	H	H	C
2	C	H	C	O	H	O	Zn	C	O	H	C	C
3	C	H	O	O	C	C	Zn	C	Zn	H	C	O
4	C	H	Zn	O	C	O	Zn	O	O	H	O	O
5	C	C	C	O	C	Zn	Zn	O	Zn			
6	C	C	O	O	O	O	Zn	Zn	Zn			
7	C	C	Zn	O	O	Zn						
8	C	O	O	O	Zn	Zn						
9	C	O	Zn									
10	C	Zn	Zn									