Supporting information for: Applicability of Tail Corrections in the Molecular Simulations of Porous Materials

Kevin Maik Jablonka, Daniele Ongari, and Berend Smit*

Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, Ecole Polytechnique Fédérale de Lausanne (EPFL), Rue de l'Industrie 17, CH-1951 Sion, Valais, Switzerland

E-mail: berend.smit@epfl.ch

Contents

1 Additional simulation details

We used the critical temperature and pressure as well as the acentric factor tabulated in Leachman *et al.*[S1] as an input for the Peng-Robinson equation of state,[S2] to compute the chemical potential for the GCMC calculations.

1.1 Modifications to the RASPA code

We modified the variable `RDFHistogramSize[i]` in the `input.c` file of the RASPA code to increase the default limit for the cutoff for radial distribution functions from 12 Å to 20 Å.

1.2 Switching potential

By default, RASPA uses a polynomial switching function (cf. Figure S1) that smooths the truncation effect over a window (r_{on}, r_c) to ensure that both interaction potential and force (i.e., the potential's derivative) are zero at the cutoff r_c in order to avoid a discontinuity. Therefore, the order of the polynomial has to be sufficiently high. The following potential fulfills this condition[S3,S4]

$$u(r_{ij}) = \begin{cases} u(r_{ij}) & r_{ij} < r_{\mathrm{on}} \\ u(r_{ij}) \, \frac{\left(r_c^2 - r^2\right)^2 \left(r_c^2 + 2r^2 - 3r_{\mathrm{on}}^2\right)}{\left(r_c^2 - r_{\mathrm{on}}^2\right)^3} & r_{\mathrm{on}} \leq r_{ij} \leq r_c \\ 0 & r_{ij} > r_c. \end{cases} \tag{1}$$

By default, the switching potential is turned on at $r_{\mathrm{on}} = 0.9\, r_c$. Note that this correction is applied when the "truncated" general rule is used. This shifted potential is different from the "shifted" rule that, as the name suggests, shifts the whole potential by the value of the potential at the cutoff, i.e., the potential is $u_{\mathrm{shifted}}(r_{ij}) = u(r_{ij}) - u(r_c)$ for $r_{ij} < r_c$, and zero beyond the cutoff. We used the default options for the switching potential to represent the most common use-case of RASPA for adsorption studies in porous materials. Our main conclusions on the applicability of tail-corrections are anyway still valid when

using a sharp truncation of the potential at the cutoff.

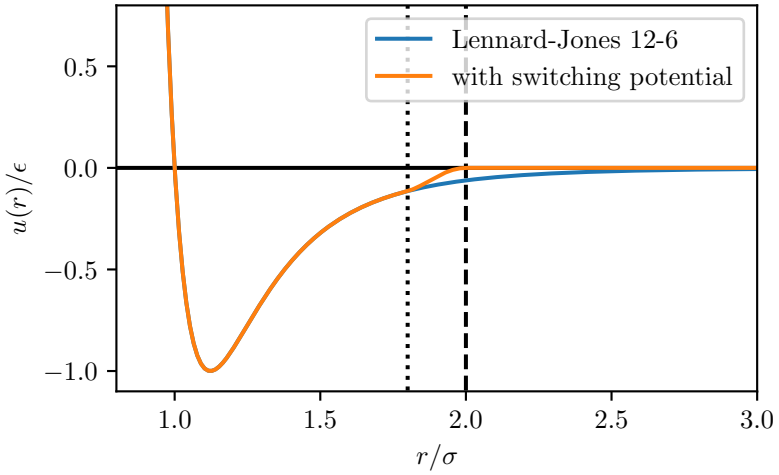

Figure S1: Lennard-Jones potential with and without switching potential for $r_{\text{cutoff}} = 2\sigma$ and $r_{\text{on}} = 0.9\, r_c$.

2 Sampling details

2.1 Preprocessing note

To avoid that some features get an unreasonable high weight due to their scale or units, the `structure_comp` package standardizes the features prior to clustering (by default).[S5]

2.2 Visualization of sampling efficiency

In the following section we visualize the spread of the *k*nn sampling in lower-dimensionality plots, to prove the diversity of the structures we selected for this study.

2.2.1 Two-dimensional representations

To better understand our sampling process, we visualized the sampling results both with principal component analysis (PCA) and metric multidimensional scaling (MDS) on a larger space of properties than the one considered for the clustering. Both PCA and MDS were performed with the `sklearn` python package,[S6] where we used the Euclidean distance metric for MDS.

For both PCA and MDS, we included the following properties to span the feature space:

The accessible surface area, the density, the largest free sphere, the largest included sphere, the largest included free sphere, the non-accessible surface area, the number of channels, the number of pockets, the pore accessible volume fraction and the pocket surface area. All these properties were computed using the Zeo++ code using a probe diameter of 1.86 Å and atomic radii corresponding to half of the Lennard-Jones's σ parameter of the force field.[S7]

A feature importance analysis also guided our decision for the initial selection of features used to span the feature space, and the skew-plot for the PCA (Figure S2) shows that seven principal components are enough to describe most of the variance in the feature space.

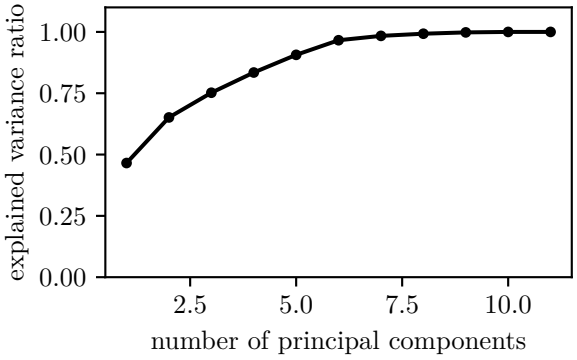

Figure S2: Skew plot for the PCA analysis.

But this also means that a two-dimensional representation, as shown in Figure S3, cannot contain all the information considered in the clustering process. Figure S3 still shows that our sampling performs well in covering most of the space of geometrical properties. Note that red points (samples considered in our studies) are mostly close in the plots because we separately sampled for MOFs, COFs and zeolites to make sure we have ten examples from each group.

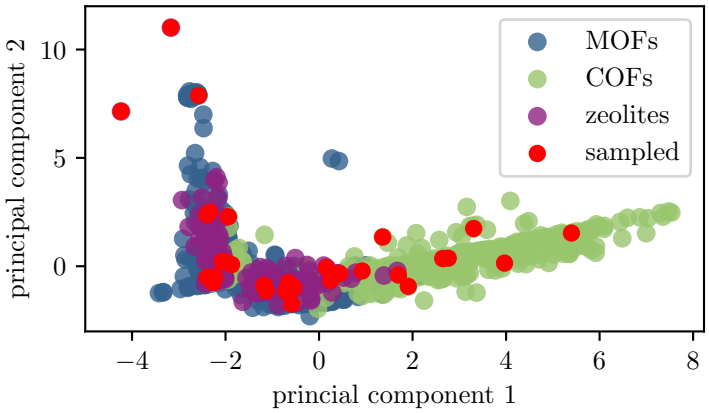

Figure S3: Projection of the feature space onto the first two principal components.

MDS is particularly interesting to compare the similarity of sample points as it tries to conserve the pairwise distances by minimizing the so-called stress function

$$S_M(z_1, z_2, \ldots, z_n) = \sum_{i \neq i'} (d_{ii'} - ||z_i - z_{i'}||)^2 \tag{2}$$

(also known as Kruskal-Shephard scaling), where the $d_{ii'}$ are the pairwise distances between samples $x_1, x_2, \ldots, x_N$ and the $z_1, z_2, \ldots, z_N \in \mathbb{R}^k$ are the points MDS tries to produce.[S8]

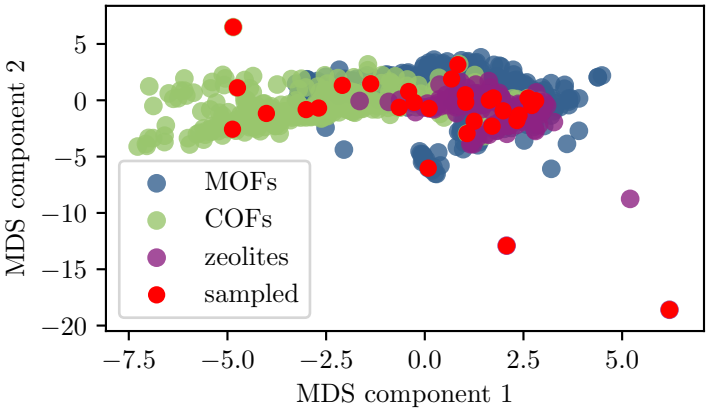

Figure S4: Two dimensional MDS projection of the feature space.

Also the MDS analysis (cf. Figure S4) shows that our sampling performs reasonably well. Red points (samples considered in this study) close in the plot are again due to the fact that we sampled separately for COFs, MOFs and zeolites.

2.2.2 Spread of Henry coefficients

Figure S5 compares the spread of Henry coefficients for a set of randomly selected structures with the set of structures obtained after $k$nn-clustering in pore-property space.

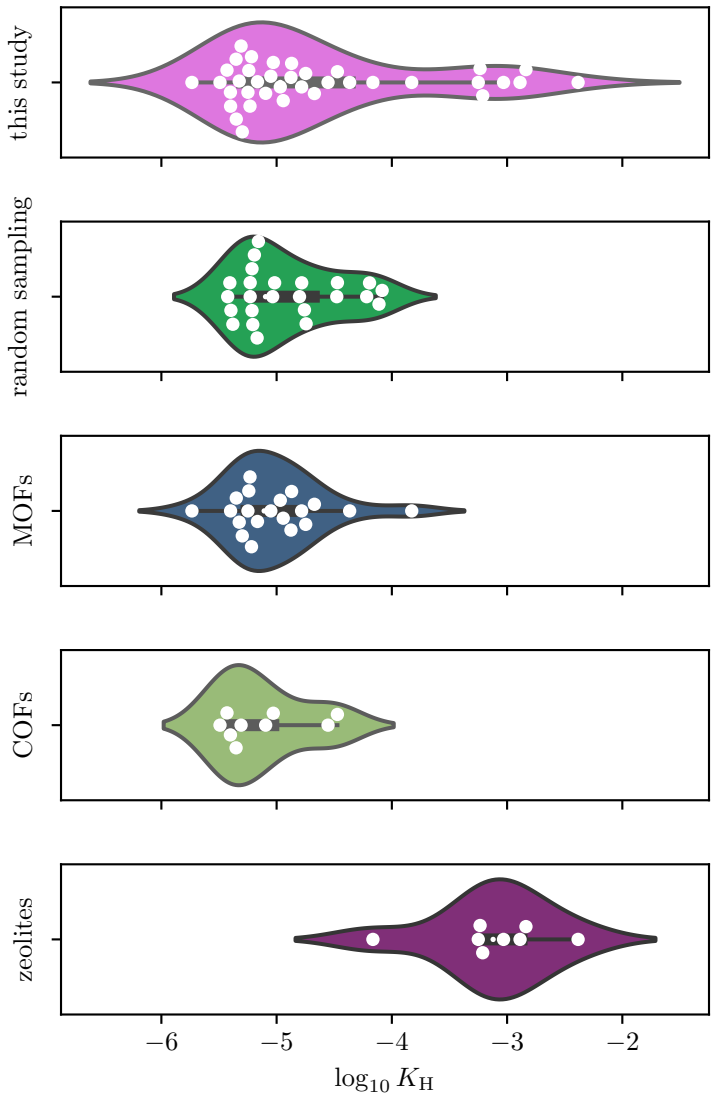

Figure S5: Distribution of the calculated Henry coefficients K_H in the different structure sets compared to random sampling.

First, one can observe that we cover a large range of Henry coefficients (already simply due to the fact that we sample from three large databases) and second, one can, as discussed above, observe that our sampling approach potentially performs better than random sampling. This is also indicated by the fact that we sampled the zeolite with the largest known cavity diameter (TSC).[S9]

3 Numeric examples

For the numeric examples we evaluated the Lennard-Jones (12-6)-potential and the analytic tail-correction on a grid for a radial distribution function described by the parametrization of Matteolia *et al.*[S10] (cf. Listing 1).

Listing 1: Python function for the parametrization of the radial distribution function used in this work

```
def rdf_model(r: float, sigma: float=1, h:float=1.070, m:float=6.0, gd:
    float=1.62, lamb:float=0.3084, alpha:float=2.817, beta:float=5.186,
    theta:float=70.90) -> float:
    """Return RDF according to a parametrization from Matteolia et al."""

    d = h * sigma
    gamma = r / d

    if gamma >= 1:
        g = gamma ** (-m) * (gd - 1 - lamb) + ((gamma - 1 + lamb) / gamma)
        g *= np.exp(-alpha * (gamma - 1)) * np.cos(beta * (gamma -1))
        g += 1

    if gamma < 1:
        g = gd * np.exp(-theta * (gamma - 1) ** 2)

    return g
```

Listing 2: Python function for the perturbed parametrization of the radial distribution function used in this work

```
def rdf_model_inhomogenous(r: float, sigma: float=1, h:float=1.070, m:float
    =6.0, gd:float=1.62, lamb:float=0.3084, alpha:float=2.817, beta:float
    =5.186, theta:float=70.90) -> float:
    """Return perturbedRDF according to a parametrization from Matteolia et
     al."""

    d = h * sigma
```

```
    gamma = r / d

    if gamma >= 1:
        g = gamma ** (-m) * (gd - 1 - lamb) + ((gamma - 1 + lamb) / gamma)
        g *= np.exp(-alpha/8 * (gamma - 1)) * (0.5 *np.cos(beta * (gamma
    -1))
        - 1 * np.cos(1 * beta * (gamma -1))
        + 0.8 *np.cos(0.1 * beta * (gamma -3))
        + 0.6 *np.cos(0.8 * beta * (gamma - 0.5) )
        + 1 *np.cos(0.8 * beta * (gamma - 2))
        + 1 *np.cos(0.3 * beta * (gamma)))
        g += 1

    if gamma < 1:
        g = gd * np.exp(-theta * (gamma - 1) ** 2)

    return g
```

The errors we provide in the plots are given by[S11]

$$\varepsilon_{\text{t.-c.}} = \left| 2 \int_{r_c}^{\infty} \mathrm{d}r\, \pi r^2 g(r) u(r) - u_{\text{t.-c.}}(r_c) \right| \tag{3}$$

and

$$\varepsilon_{\text{trun.}} = \left| 2 \int_{r_c}^{\infty} \mathrm{d}r\, \pi r^2 g(r) u(r) \right|, \tag{4}$$

where we set $N = \rho = 1$ for simplicity.

4 Convergence of loadings at different pressures

The convergence of the loading with respect to the cutoff distance at a pressure of 5.8 bar, 35 bar and 65 bar is shown in Figure S6. We can observe that the simulations with non-corrected truncation generally lead to an underestimation of the loading. On the other hand, the results from simulations with tail-correction are less sensitive to the truncation radius r_c.

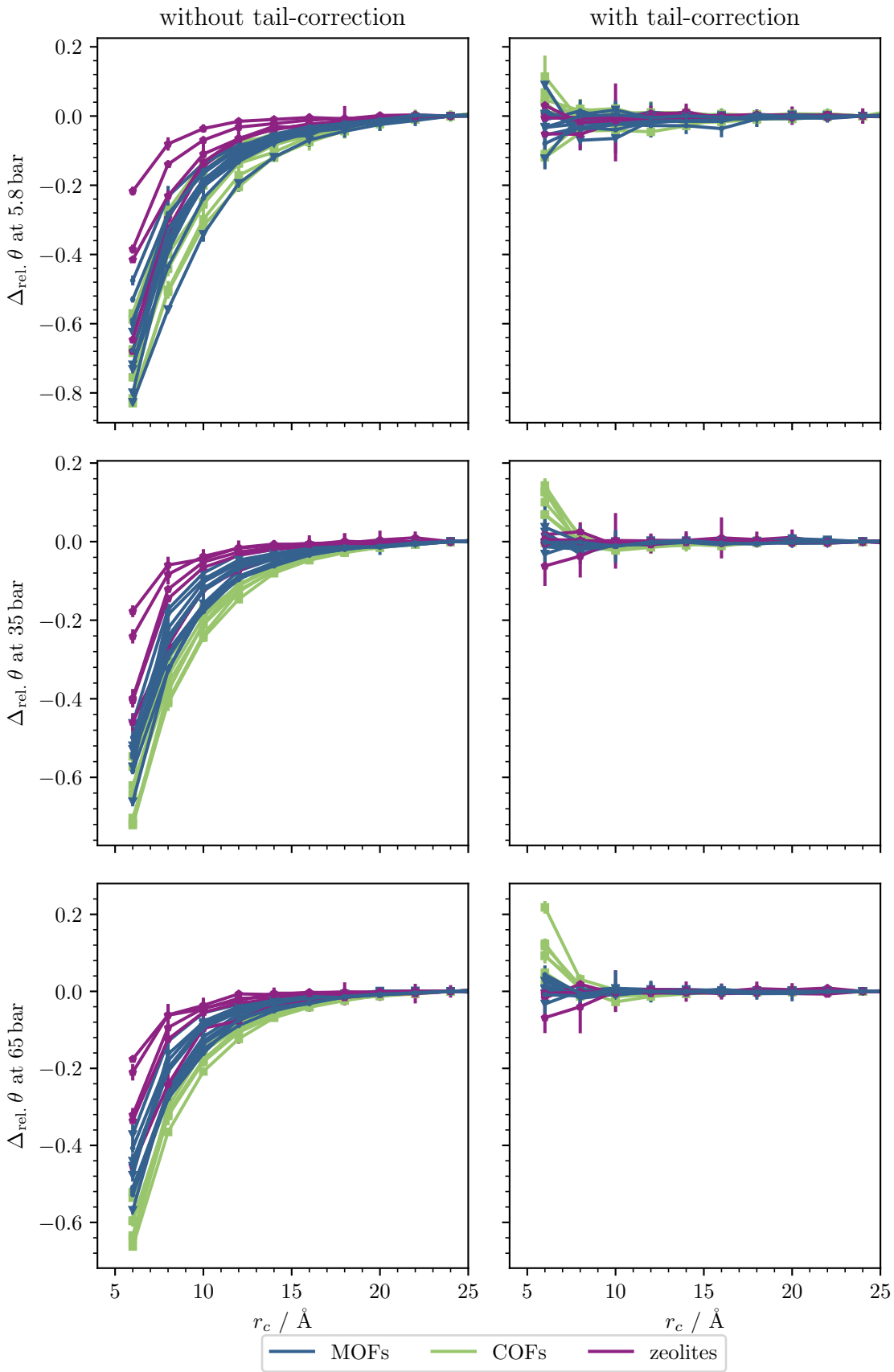

Figure S6: Convergence of loading θ at different pressures in terms of the relative error $\Delta_{\text{rel.}}\theta = \frac{\theta(r_c) - \theta(r_c = 24\,\text{Å})}{\theta(r_c = 24\,\text{Å})}$.

5 Simulations without pore blocking

Several of the structures from our sample (13030N2, 17120N2, ZIF-4, SBMOF-1, ZIF-8, PAU, AFT, TSC, GOO, UEI, XOJWID, TUDJIM, PIHNUQ, BENXUP, YUJNIB, cf. section 9 for the nomenclature) have a negligible accessible void faction (i.e., less than 1×10^{-5}). We excluded those structures from the discussion in the full text because the simulations are not relevant when compared to experiments[S12] since the methane molecules can not permeate in these pores (under the assumption of perfectly rigid structures) However, they are still interesting to analyze, in order to understand the impact of the use of tail-corrections.

The convergence of the relative error in the deliverable capacity, as shown in Figure S7, and of the Henry coefficient, as shown in Figure S8, shows the same trend as the one discussed in the full text. The same holds true for the loadings at different pressures as shown in Figure S9.

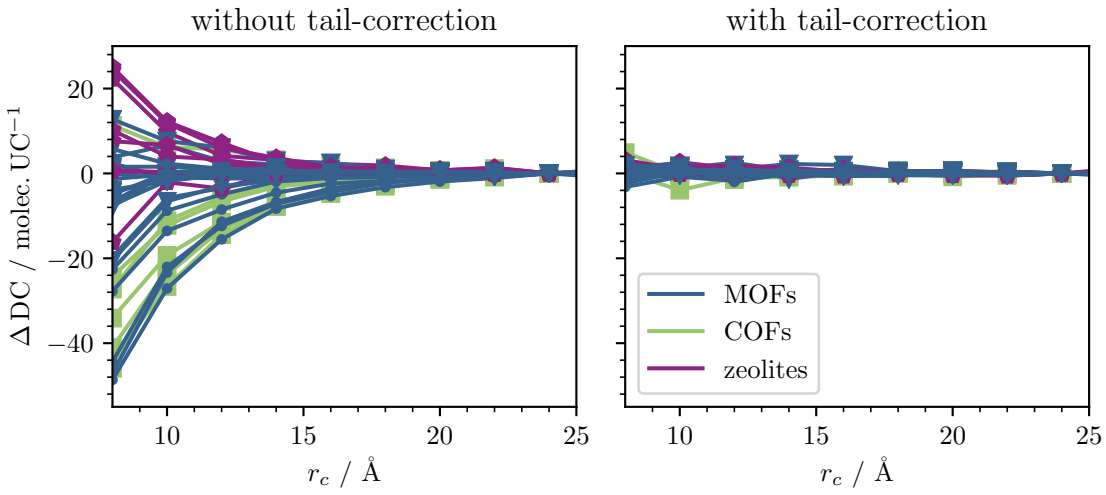

Figure S7: Convergence of deliverable capacities $\mathrm{DC}_{6.5 \times 10^1\,\mathrm{bar},\,5.8\,\mathrm{bar}}$ as a function of the cutoff distance with and without tail-corrections for simulations without pore blocking

One of the zeolites with comparatively high errors when using tail-corrections at low cutoffs (relative error > 0.25 at $r_c = 8\,\text{Å}$ in Figure S8) is TSC, whose radial distribution function is actually the only one that shows values < 0.5 at some r (cf. Figure S10). This leads to a large value for $\left|\int_{r_c}^{\infty} \mathrm{d}r\ (g(r)-1)^2\right|$. Additionally, due its large spherical pore (diameter approx. 16 Å), with more concentric cylindrical layers (cf. Figure S11), the

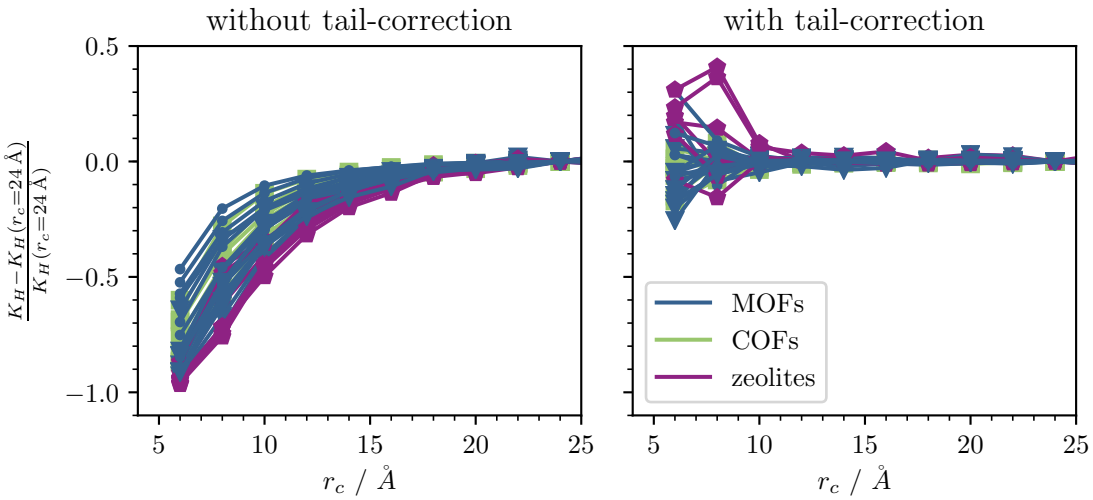

Figure S8: Convergence of the Henry coefficient K_H as a function of the cutoff distance with and without tail-corrections for simulations without pore blocking. We show the relative error due to the large spread of Henry coefficients.

potential in its center is non-interacting for cutoff distances $r_c < 8\,\text{Å}$. The other zeolite with comparatively high relative error is EMT which exhibits a similar pore structure with a large spherical pore (diameter $> 12\,\text{Å}$).

Note that only in one material, the zeolite Tschörtnerite (TSC), we observe that $g(r) < 0.5$, which is the threshold we derived in the main text. The pore geometry of this material is an extreme case: TSC is the zeolite with the largest known cavity diameter (approx. 16 Å),[S9] and this large pores are connected by small windows, with the largest free sphere diameter of the structure being 3.68 Å—i.e., to small to fit a methane bead with a Lennard-Jones parameter of $\sigma = 3.73\,\text{Å}$ (see Figure S11).

In the main text we considered blocking spheres, as computed from Zeo++, for the following structures: EMT, 15072N2, 16083N2, 18091N2. Note that these structures have both accessible and non-accessible pores and blocking sphere prevent the insertion of methane molecules in the second ones, to make them comparable with experimental data.

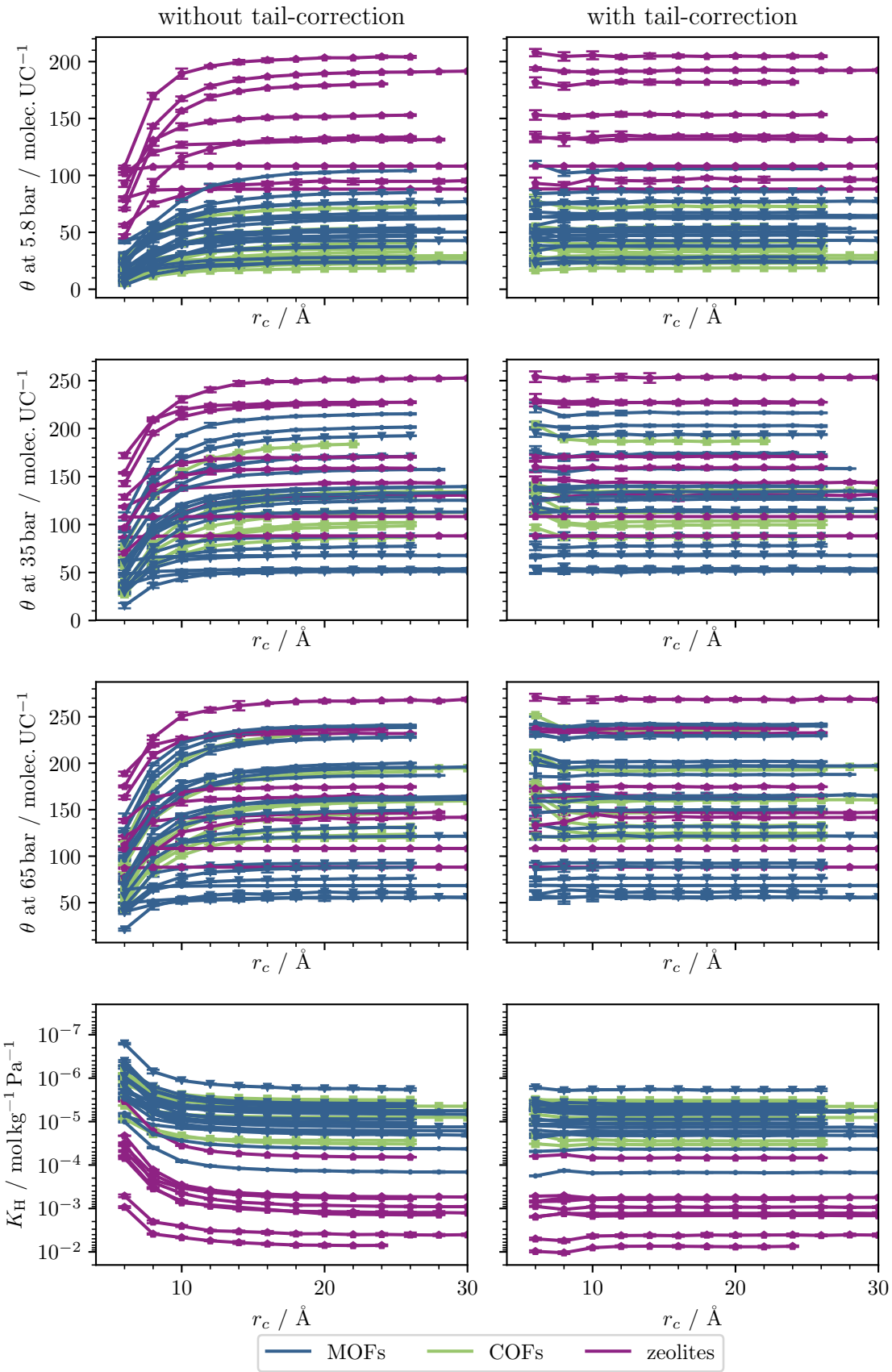

Figure S9: Absolute values for loading θ and Henry coefficients K_{H} as a function of cutoff r_c with and without tail-correction for simulations without pore blocking.

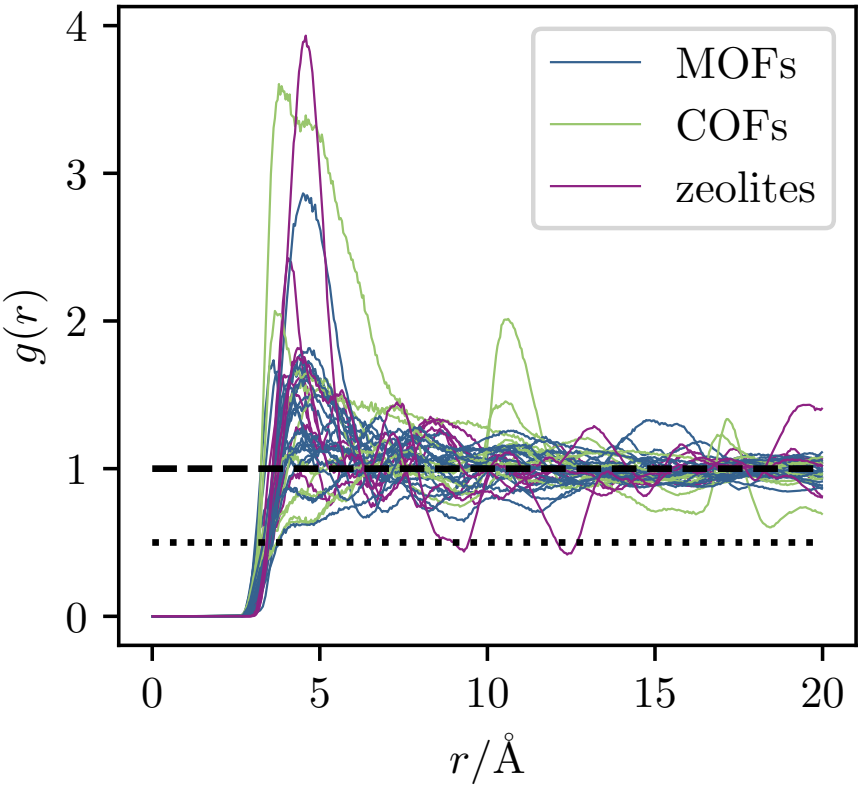

Figure S10: CH_4-framework radial distribution functions for simulations without pore blocking. Dashed line indicating $g(r) = 1$, dotted line indicating $g(r) = 0.5$.

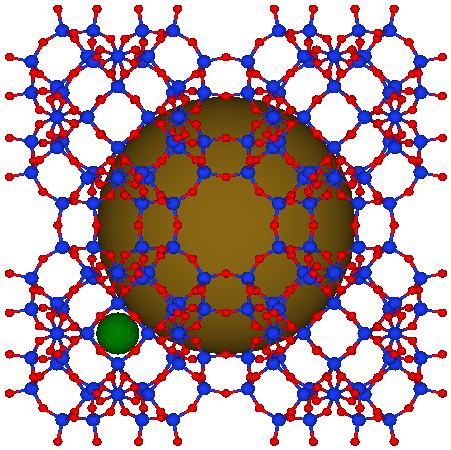

Figure S11: Visualization of the pore geometry of TSC. Golden ball in the large cavity, green ball with diameter of methane in the small cavity

6 Water adsorption in aluminum fumarate

As an additional test we considered the case of water adsorption in aluminum fumarate. For this, we performed a cell optimization using the Gaussian-plane wave method implemented in the CP2K code,[S13–S15] where we used the PBE exchange correlation functional[S16] with the DFT-D3(BJ) dispersion correction.[S17] We described the framework using the UFF force field[S18] and DDEC charges[S19] and used the TIP4P-2005 model[S20] to describe water. To calculate the Henry coefficient, we performed 6.0×10^6 Widom insertions with the RASPA code.[S21]

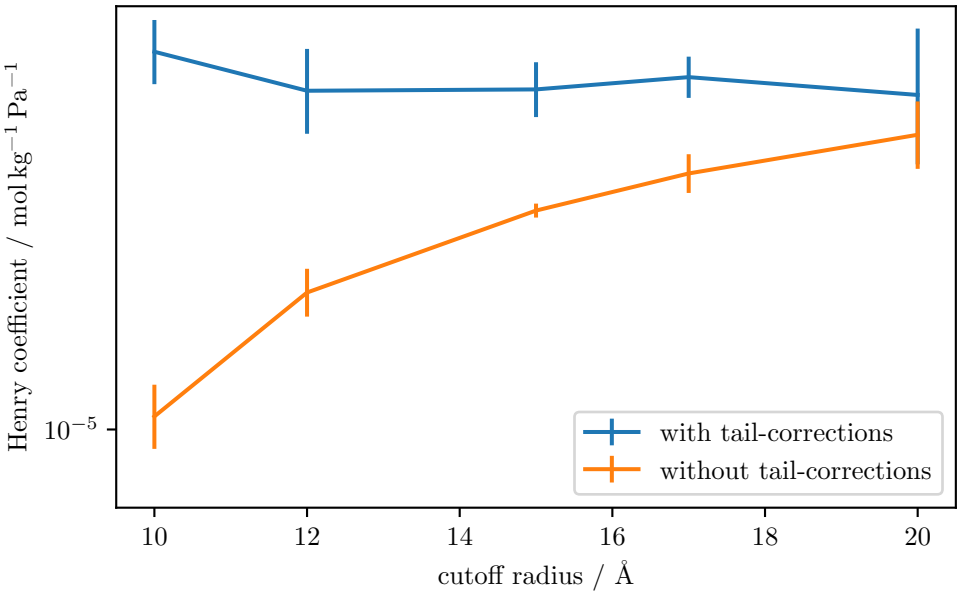

Figure S12: Henry coefficients K_H as a function of cutoff radius r_c with and without tail-correction for the case of TIP4P-2005 water in aluminium fumarate.

Also for this case, we find that the simulations with tail-corrections show a more desirable convergence behavior with respect to the cutoff radius than simulations without tail-corrections (cf. Figure S12).

7 Adsorption of alkanes

As an additional test and more direct comparison to the simulations of Macedonia and Maginn[S22] we also considered the adsorption of ethane, *n*-butane and *n*-heptane in some archetypal materials (MFI, SFS, IRMOF-1, HKUST-1, UiO-66 and 15072N2) using the TraPPE united atom model for the alkanes,[S23] the DREIDING model for the frameworks,[S24] where we added the missing parameters from the UFF force field,[S18] and Lorentz-Berthelot mixing rules for the interaction terms. We determined ideal Rosenbluth weights using 6.0×10^4 cycles in a $30\,\text{Å} \times 30\,\text{Å} \times 30\,\text{Å}$ box. To expedite the simulations of the longer chains, we calculated the Henry coefficient at 570 K using 5.0×10^4 cycles of insertions. The results are shown in Figure S13 and show the same trends as the ones discussed in the main text.

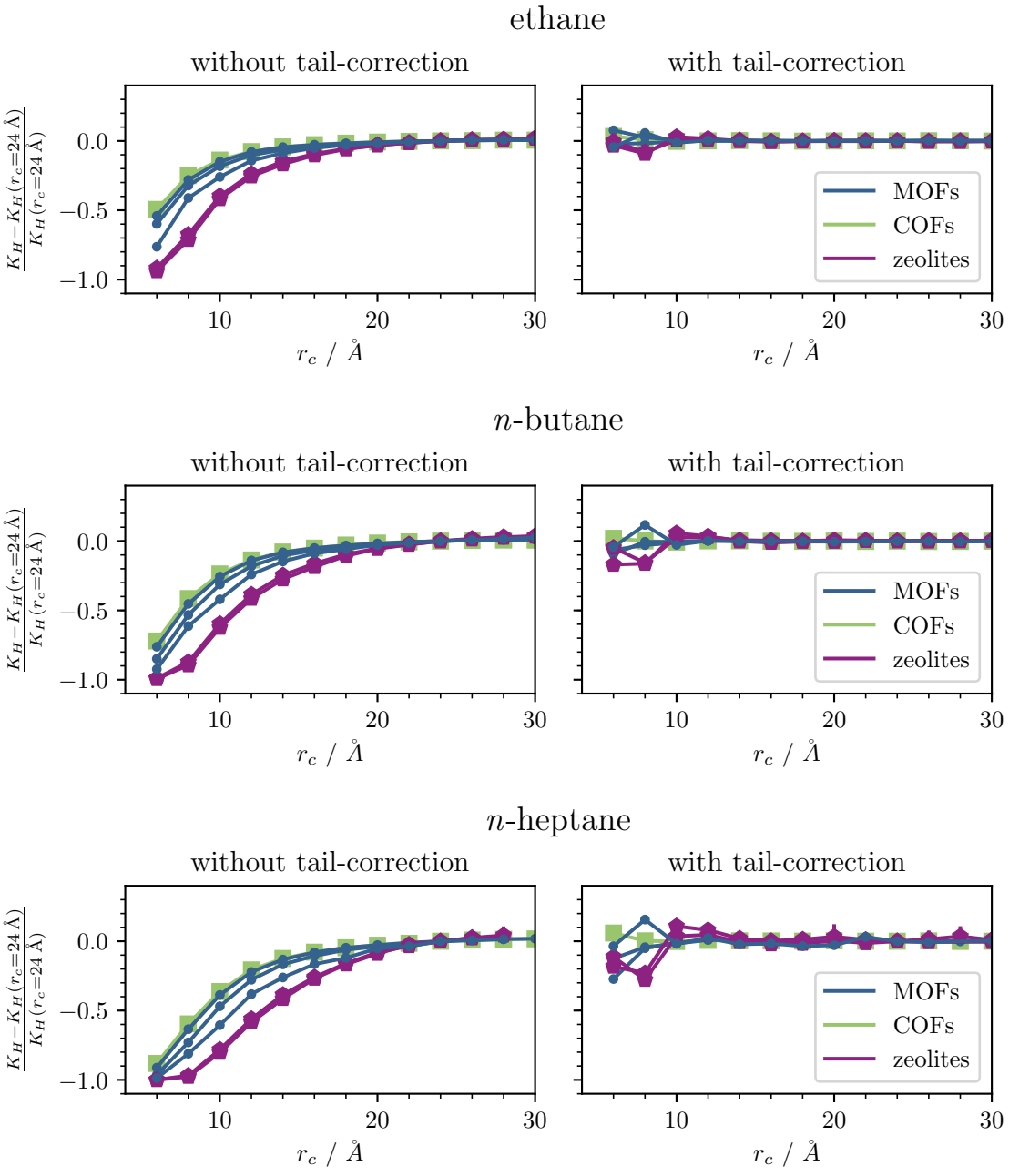

Figure S13: Relative errors (w.r.t. to the values at 24 Å) of Henry coefficients of linear alkanes in different porous materials at 570 K as a function of the cutoff radius for simulations with and without tail-corrections.

8 Generalization to summation over atom types

In general, the system consists of a collection of different pairs $p \in \mathrm{S} \times \mathrm{S}$, where S is the set of (pseudo-)atoms, for which case the error terms have the following form

$$\varepsilon_{\text{trunc.}} = \left| \sum_{p \in \mathrm{S}^2} \frac{\rho_p}{2} \int \mathrm{d}\mathbf{R}\, \tilde{g}_p^{(2)}\left(\xi, \mathbf{R}\right) \int_{|\mathbf{r}|=r_c}^{\infty} \mathrm{d}\mathbf{r}\, u_p\left(|\mathbf{r}|\right) \right| \tag{5}$$

and

$$\varepsilon_{\text{t.-c.}} = \left| \sum_{p \in \mathrm{S}^2} \frac{\rho_p}{2} \int \mathrm{d}\mathbf{R}\, \left(\tilde{g}_p^{(2)}\left(\xi, \mathbf{R}\right) - 1\right) \int_{|\mathbf{r}|=r_c}^{\infty} \mathrm{d}\mathbf{r}\, u_p\left(|\mathbf{r}|\right) \right|. \tag{6}$$

Let us rewrite the expressions by introducing

$$a_p = \int \mathrm{d}\mathbf{R}\, \tilde{g}_p^{(2}\left(\xi, \mathbf{R}\right) \tag{7}$$

and

$$b_p = \int_{|\mathbf{r}|=r_c}^{\infty} \mathrm{d}\mathbf{r}\, u_p\left(|\mathbf{r}|\right), \tag{8}$$

with which we can write

$$\begin{aligned} \varepsilon_{\text{trunc.}} &= \left| \sum_{p \in \mathrm{S}^2} a_p b_p \right| \\ &= \left| \sum_{p} a_p \sum_{p'} b_{p'} - \sum_{p,p',p \neq p'} a_p b_{p'} \right| \\ &= |AB - C| \end{aligned} \tag{9}$$

and

$$\begin{aligned} \varepsilon_{\text{t.-c.}} &= \left| \sum_{p \in \mathrm{S}^2} (a_p - 1) b_p \right| \\ &= \left| \sum_{p} a_p \sum_{p'} b_{p'} - \sum_{p'} b_{p'} - \sum_{p,p',p \neq p'} a_p b_{p'} \right|, \\ &= |AB - B - C| \end{aligned} \tag{10}$$

where we introduced

$$A = \sum a_p \quad B = \sum_{p'} b_{p'} \quad C = \sum_{p,p',p \neq p'} a_p b_{p'}. \tag{11}$$

For $B < 0$ one finds the solution set $\{A > \frac{C}{B} + \frac{1}{2}\}$. In case the covariance term C vanishes, we find the condition $\sum_p a_p > 0.5$, which is the relevant case as we are only interested in real solutions, in which case we can rewrite the inequality (eq. **??**) as

$$\sqrt{(AB - C)^2} > \sqrt{(AB - B - C)^2}. \tag{12}$$

9 Converged absolute values at higher cutoff and uncertainties

9.1 Converged absolute values for simulations with blocked pockets

In the main text we showed in all the plots a maximum value of 24 Å for the cutoff, r_c. For all the computed properties, there was a negligible change when going to higher r_c (cf. Figure S14). To demonstrate this, Tables S1, S2 and S3 list all the materials considered in this study together with the converged values (at higher r_c) for the loadings θ and the Henry coefficients K_{H}. These "converged values" are computed with the highest feasible cutoff:

$$\min\left(\max\left(\text{cutoff with tail-correction}\right), \max\left(\text{cutoff without tail-correction}\right)\right), \qquad (13)$$

which is 30 Å, or less for the calculations that could not finish at that high cutoff within the walltime-limit of 3 days.

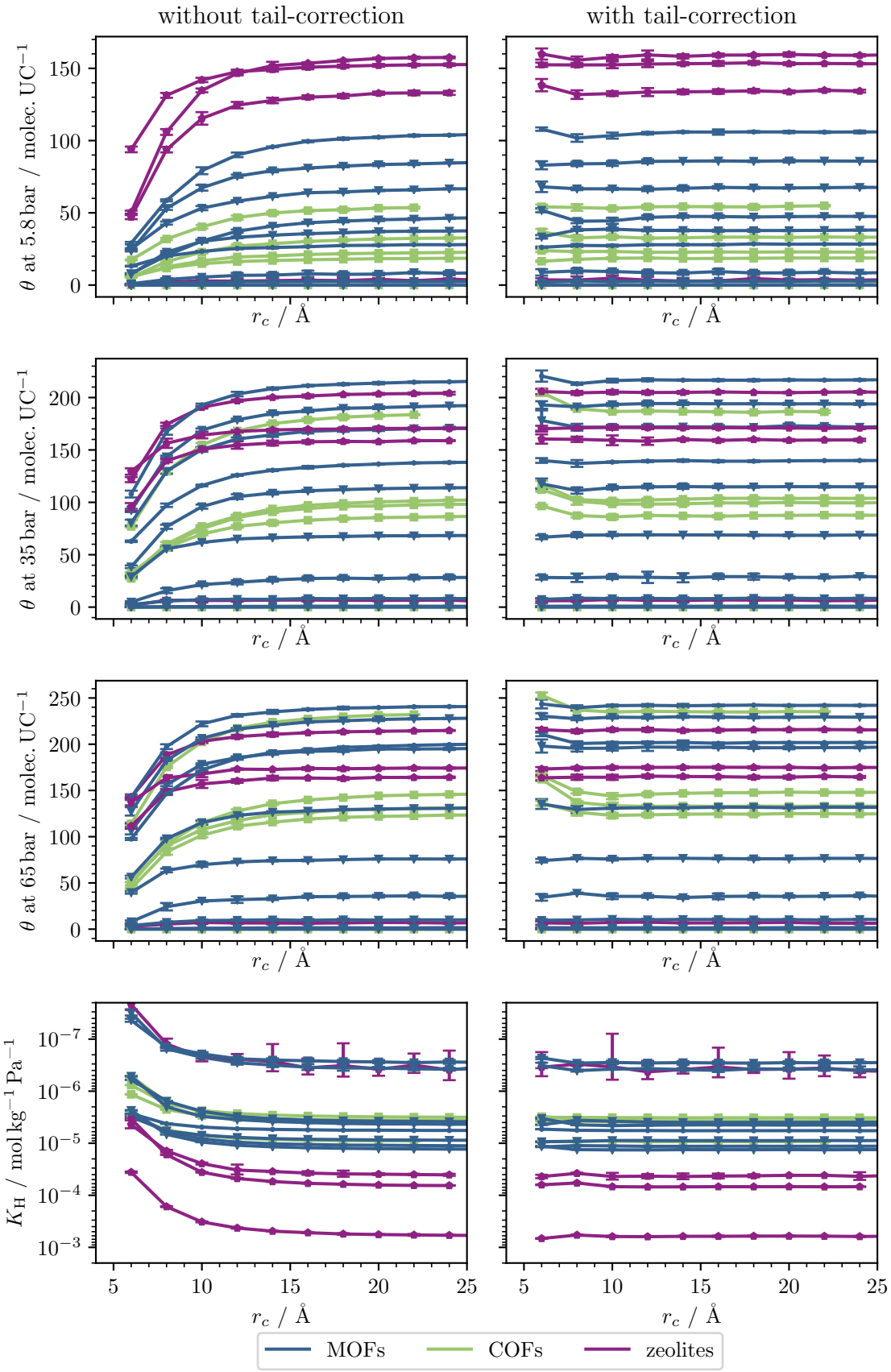

Figure S14: Absolute values for loading θ and Henry coefficients K_{H} as a function of the cutoff radius r_c with and without tail-correction for simulations without pore blocking.

In addition to that, we show in Figure S14 all the properties considered in the study in absolute values, there, one can see that the values at highest cutoff have a negligible difference from the ones at $r_c = 24\,\text{Å}$, which are plotted in the main text.

In all the computed results, the means and standard deviations (given as uncertainties in bracket notation) are obtained by block-averaging as implemented in the RASPA code.[S21]

From all of our plots we excluded the 2D COF CFT (13030N2 in Table S2, see the CURATED COFs from Ongari *et al.* for nomenclature of the COF structures[S25]), as the inter-layer distance of 3.3 Å as well as the channel diameter are smaller than the $\sigma = 3.7\,\text{Å}$ of the methane bead in our simulation (as well as the kinetic diameter[S26]). For this structure only, methane can not fit even when blocking spheres are not considered, resulting in a null uptake and extremely low Henry coefficient. For the MOFs, we use the nomenclature from the CoRE-MOF database for the ones that we sampled from the database and the common names for the ones which we chose according to their popularity.

Table S1: Converged properties for the MOFs considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_{H} in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations with blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ tc.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
UMCM-1	26.0	$5.68(3) \times 10^{-6}$	$5.74(4) \times 10^{-6}$	$2.81(2) \times 10^{1}$	$2.81(2) \times 10^{1}$	$1.380(5) \times 10^{2}$	$1.380(5) \times 10^{2}$	$2.000(3) \times 10^{2}$	$2.000(3) \times 10^{2}$
NOCKUM	26.0	$1.310(9) \times 10^{-5}$	$1.330(7) \times 10^{-5}$	$8.49(8) \times 10^{1}$	$8.49(8) \times 10^{1}$	$1.710(8) \times 10^{2}$	$1.710(8) \times 10^{2}$	$1.960(9) \times 10^{2}$	$1.960(9) \times 10^{2}$
NUTQEZ	26.0	$8.77(6) \times 10^{-6}$	$8.86(4) \times 10^{-6}$	$6.67(4) \times 10^{1}$	$6.67(4) \times 10^{1}$	$1.920(4) \times 10^{2}$	$1.920(4) \times 10^{2}$	$2.280(4) \times 10^{2}$	$2.280(4) \times 10^{2}$
RIVDEF	26.0	$4.33(5) \times 10^{-6}$	$4.47(5) \times 10^{-6}$	$3.77(2) \times 10^{1}$	$3.77(2) \times 10^{1}$	$6.83(2) \times 10^{1}$	$6.83(2) \times 10^{1}$	$7.61(4) \times 10^{1}$	$7.61(4) \times 10^{1}$
Mg-MOF-74	26.0	$1.13(1) \times 10^{-5}$	$1.140(6) \times 10^{-5}$	$1.040(6) \times 10^{2}$	$1.040(6) \times 10^{2}$	$2.160(4) \times 10^{2}$	$2.160(4) \times 10^{2}$	$2.410(5) \times 10^{2}$	$2.410(5) \times 10^{2}$
IBICIH	26.0	$3.87(3) \times 10^{-6}$	$3.97(3) \times 10^{-6}$	$4.66(3) \times 10^{1}$	$4.66(3) \times 10^{1}$	$1.140(6) \times 10^{2}$	$1.140(6) \times 10^{2}$	$1.310(5) \times 10^{2}$	$1.310(5) \times 10^{2}$
Cu-BTC	26.0	$1.78(6) \times 10^{-5}$	$1.80(5) \times 10^{-5}$	$7.60(8) \times 10^{1}$	$7.60(8) \times 10^{1}$	$2.020(6) \times 10^{2}$	$2.020(6) \times 10^{2}$	$2.390(9) \times 10^{2}$	$2.390(9) \times 10^{2}$
UIO-66	24.0	$1.610(7) \times 10^{-5}$	$1.65(2) \times 10^{-5}$	$6.37(6) \times 10^{1}$	$6.37(6) \times 10^{1}$	$1.330(8) \times 10^{2}$	$1.330(8) \times 10^{2}$	$1.63(1) \times 10^{2}$	$1.63(1) \times 10^{2}$
IRMOF-1	24.0	$4.96(2) \times 10^{-6}$	$5.03(2) \times 10^{-6}$	$3.78(3) \times 10^{1}$	$3.78(3) \times 10^{1}$	$1.720(6) \times 10^{2}$	$1.720(6) \times 10^{2}$	$2.270(6) \times 10^{2}$	$2.270(6) \times 10^{2}$
IRMOF-10	26.0	$5.59(2) \times 10^{-6}$	$5.65(4) \times 10^{-6}$	$2.340(9) \times 10^{1}$	$2.340(9) \times 10^{1}$	$1.250(6) \times 10^{2}$	$1.250(6) \times 10^{2}$	$1.950(4) \times 10^{2}$	$1.950(4) \times 10^{2}$
DIYTEM	26.0	$1.33(3) \times 10^{-5}$	$1.36(1) \times 10^{-5}$	$7.7(1) \times 10^{1}$	$7.7(1) \times 10^{1}$	$1.130(4) \times 10^{2}$	$1.130(4) \times 10^{2}$	$1.210(8) \times 10^{2}$	$1.210(8) \times 10^{2}$
MIL-125	26.0	$5.97(5) \times 10^{-6}$	$6.06(1) \times 10^{-6}$	$5.24(4) \times 10^{1}$	$5.24(4) \times 10^{1}$	$1.570(6) \times 10^{2}$	$1.570(6) \times 10^{2}$	$1.870(4) \times 10^{2}$	$1.870(4) \times 10^{2}$

Table S2: Converged properties for the COFs considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_{H} in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations with blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ t.-c.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
18091N2	22.0	$9.0(2) \times 10^{-6}$	$9.3(2) \times 10^{-6}$	$5.36(7) \times 10^{1}$	$5.36(7) \times 10^{1}$	$1.840(5) \times 10^{2}$	$1.840(5) \times 10^{2}$	$2.320(4) \times 10^{2}$	$2.320(4) \times 10^{2}$
15030N2	26.0	$3.18(2) \times 10^{-6}$	$3.24(2) \times 10^{-6}$	$2.28(2) \times 10^{1}$	$2.28(2) \times 10^{1}$	$1.020(2) \times 10^{2}$	$1.020(2) \times 10^{2}$	$1.460(3) \times 10^{2}$	$1.460(3) \times 10^{2}$
16411C2	26.0	$3.87(7) \times 10^{-6}$	$3.95(3) \times 10^{-6}$	$3.27(1) \times 10^{1}$	$3.27(1) \times 10^{1}$	$9.83(5) \times 10^{1}$	$9.83(5) \times 10^{1}$	$1.230(5) \times 10^{2}$	$1.230(5) \times 10^{2}$
17030N2	26.0	$3.66(4) \times 10^{-6}$	$3.72(2) \times 10^{-6}$	$1.85(2) \times 10^{1}$	$1.85(2) \times 10^{1}$	$8.67(4) \times 10^{1}$	$8.67(4) \times 10^{1}$	$1.310(5) \times 10^{2}$	$1.310(5) \times 10^{2}$
17040N2	24.0	$4.79(4) \times 10^{-6}$	$4.91(3) \times 10^{-6}$	$4.03(6) \times 10^{1}$	$4.03(6) \times 10^{1}$	$1.300(3) \times 10^{2}$	$1.300(3) \times 10^{2}$	$1.580(3) \times 10^{2}$	$1.580(3) \times 10^{2}$
16083N2	26.0	$4.33(4) \times 10^{-6}$	$4.42(5) \times 10^{-6}$	$2.63(2) \times 10^{1}$	$2.63(2) \times 10^{1}$	$1.110(5) \times 10^{2}$	$1.110(5) \times 10^{2}$	$1.580(6) \times 10^{2}$	$1.580(6) \times 10^{2}$
17163N3	26.0	$7.97(6) \times 10^{-6}$	$8.03(4) \times 10^{-6}$	$2.94(2) \times 10^{1}$	$2.94(2) \times 10^{1}$	$1.340(2) \times 10^{2}$	$1.340(2) \times 10^{2}$	$1.940(6) \times 10^{2}$	$1.940(6) \times 10^{2}$
15072N2	26.0	$5.6(1) \times 10^{-6}$	$5.72(4) \times 10^{-6}$	$2.46(2) \times 10^{1}$	$2.46(2) \times 10^{1}$	$1.150(2) \times 10^{2}$	$1.150(2) \times 10^{2}$	$1.690(3) \times 10^{2}$	$1.690(3) \times 10^{2}$

Table S3: Converged properties for the zeolites considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_{H} in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations with blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ t.-c.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
EMT	24.0	$4.1(2) \times 10^{-5}$	$4.3(7) \times 10^{-5}$	$1.580(6) \times 10^{2}$	$1.580(6) \times 10^{2}$	$2.04(1) \times 10^{2}$	$2.04(1) \times 10^{2}$	$2.150(5) \times 10^{2}$	$2.150(5) \times 10^{2}$
SFS	26.0	$5.91(5) \times 10^{-4}$	$6.1(2) \times 10^{-4}$	$1.530(5) \times 10^{2}$	$1.530(5) \times 10^{2}$	$1.710(7) \times 10^{2}$	$1.710(7) \times 10^{2}$	$1.74(1) \times 10^{2}$	$1.74(1) \times 10^{2}$
SFN	24.0	$6.50(4) \times 10^{-5}$	$6.84(8) \times 10^{-5}$	$1.33(1) \times 10^{2}$	$1.33(1) \times 10^{2}$	$1.590(4) \times 10^{2}$	$1.590(4) \times 10^{2}$	$1.640(8) \times 10^{2}$	$1.640(8) \times 10^{2}$
MFI	26.0	$1.41(2) \times 10^{-3}$	$1.46(3) \times 10^{-3}$	$1.31(2) \times 10^{2}$	$1.31(2) \times 10^{2}$	$1.43(1) \times 10^{2}$	$1.43(1) \times 10^{2}$	$1.470(4) \times 10^{2}$	$1.470(4) \times 10^{2}$
MTT	26.0	$5.4(1) \times 10^{-4}$	$5.61(4) \times 10^{-4}$	$9.58(7) \times 10^{1}$	$9.58(7) \times 10^{1}$	$1.30(2) \times 10^{2}$	$1.30(2) \times 10^{2}$	$1.42(2) \times 10^{2}$	$1.42(2) \times 10^{2}$

9.2 Converged absolute values for simulations without blocked pockets

Tables S4, S5 and S6 summarize the absolute values for the properties considered in this study for simulations without blocked pockets.

Table S4: Converged properties for the MOFs considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_H in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations without blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ tc.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
Cu-BTC	26.0	$1.77(3) \times 10^{-5}$	$1.78(4) \times 10^{-5}$	$7.59(5) \times 10^{1}$	$7.59(5) \times 10^{1}$	$2.020(8) \times 10^{2}$	$2.020(8) \times 10^{2}$	$2.390(7) \times 10^{2}$	$2.390(7) \times 10^{2}$
DIYTEM	30.0	$1.33(1) \times 10^{-5}$	$1.360(6) \times 10^{-5}$	$7.69(4) \times 10^{1}$	$7.69(4) \times 10^{1}$	$1.130(5) \times 10^{2}$	$1.130(5) \times 10^{2}$	$1.210(3) \times 10^{2}$	$1.210(3) \times 10^{2}$
IRMOF-10	30.0	$5.61(2) \times 10^{-6}$	$5.63(2) \times 10^{-6}$	$2.36(1) \times 10^{1}$	$2.36(1) \times 10^{1}$	$1.250(6) \times 10^{2}$	$1.250(6) \times 10^{2}$	$1.960(4) \times 10^{2}$	$1.960(4) \times 10^{2}$
MIL-125	28.0	$5.99(3) \times 10^{-6}$	$6.04(4) \times 10^{-6}$	$5.26(2) \times 10^{1}$	$5.26(2) \times 10^{1}$	$1.570(7) \times 10^{2}$	$1.570(7) \times 10^{2}$	$1.870(3) \times 10^{2}$	$1.870(3) \times 10^{2}$
PIHNUQ	30.0	$2.1(1) \times 10^{-5}$	$2.08(7) \times 10^{-5}$	$4.27(3) \times 10^{1}$	$4.27(3) \times 10^{1}$	$5.13(5) \times 10^{1}$	$5.13(5) \times 10^{1}$	$5.59(4) \times 10^{1}$	$5.59(4) \times 10^{1}$
IRMOF-1	24.0	$4.97(3) \times 10^{-6}$	$5.02(3) \times 10^{-6}$	$3.74(6) \times 10^{1}$	$3.74(6) \times 10^{1}$	$1.720(7) \times 10^{2}$	$1.720(7) \times 10^{2}$	$2.270(7) \times 10^{2}$	$2.270(7) \times 10^{2}$
YUJNIB	26.0	$1.8(1) \times 10^{-6}$	$1.9(1) \times 10^{-6}$	$2.4(1) \times 10^{1}$	$2.4(1) \times 10^{1}$	$5.3(1) \times 10^{1}$	$5.3(1) \times 10^{1}$	$6.10(8) \times 10^{1}$	$6.10(8) \times 10^{1}$
UMCM-1	26.0	$5.70(3) \times 10^{-6}$	$5.74(2) \times 10^{-6}$	$2.81(1) \times 10^{1}$	$2.81(1) \times 10^{1}$	$1.380(3) \times 10^{2}$	$1.380(3) \times 10^{2}$	$2.000(3) \times 10^{2}$	$2.000(3) \times 10^{2}$
XOJWID	26.0	$1.06(1) \times 10^{-5}$	$1.08(1) \times 10^{-5}$	$4.71(3) \times 10^{1}$	$4.71(3) \times 10^{1}$	$7.7(1) \times 10^{1}$	$7.7(1) \times 10^{1}$	$8.80(5) \times 10^{1}$	$8.80(5) \times 10^{1}$
NOCKUM	26.0	$1.300(6) \times 10^{-5}$	$1.330(9) \times 10^{-5}$	$8.49(2) \times 10^{1}$	$8.49(2) \times 10^{1}$	$1.710(9) \times 10^{2}$	$1.710(9) \times 10^{2}$	$1.960(9) \times 10^{2}$	$1.960(9) \times 10^{2}$
NUTQEZ	26.0	$8.77(7) \times 10^{-6}$	$8.92(7) \times 10^{-6}$	$6.67(6) \times 10^{1}$	$6.67(6) \times 10^{1}$	$1.930(3) \times 10^{2}$	$1.930(3) \times 10^{2}$	$2.280(8) \times 10^{2}$	$2.280(8) \times 10^{2}$
RIVDEF	26.0	$4.35(6) \times 10^{-6}$	$4.47(5) \times 10^{-6}$	$3.75(3) \times 10^{1}$	$3.75(3) \times 10^{1}$	$6.83(2) \times 10^{1}$	$6.83(2) \times 10^{1}$	$7.61(2) \times 10^{1}$	$7.61(2) \times 10^{1}$
Mg-MOF-74	26.0	$1.12(1) \times 10^{-5}$	$1.140(4) \times 10^{-5}$	$1.040(7) \times 10^{2}$	$1.040(7) \times 10^{2}$	$2.150(7) \times 10^{2}$	$2.150(7) \times 10^{2}$	$2.410(6) \times 10^{2}$	$2.410(6) \times 10^{2}$
BENXUP	26.0	$4.5(3) \times 10^{-6}$	$4.8(3) \times 10^{-6}$	$5.20(3) \times 10^{1}$	$5.20(3) \times 10^{1}$	$8.74(4) \times 10^{1}$	$8.74(4) \times 10^{1}$	$9.25(4) \times 10^{1}$	$9.25(4) \times 10^{1}$
IBICIH	26.0	$3.88(4) \times 10^{-6}$	$3.96(2) \times 10^{-6}$	$4.66(3) \times 10^{1}$	$4.66(3) \times 10^{1}$	$1.140(5) \times 10^{2}$	$1.140(5) \times 10^{2}$	$1.310(5) \times 10^{2}$	$1.310(5) \times 10^{2}$
TUDJIM	26.0	$5.77(6) \times 10^{-6}$	$5.90(3) \times 10^{-6}$	$5.39(5) \times 10^{1}$	$5.39(5) \times 10^{1}$	$1.290(7) \times 10^{2}$	$1.290(7) \times 10^{2}$	$1.490(7) \times 10^{2}$	$1.490(7) \times 10^{2}$
UIO-66	30.0	$1.63(1) \times 10^{-5}$	$1.65(1) \times 10^{-5}$	$6.42(2) \times 10^{1}$	$6.42(2) \times 10^{1}$	$1.330(7) \times 10^{2}$	$1.330(7) \times 10^{2}$	$1.640(6) \times 10^{2}$	$1.640(6) \times 10^{2}$
SBMOF-1	30.0	$1.45(2) \times 10^{-4}$	$1.49(3) \times 10^{-4}$	$5.03(2) \times 10^{1}$	$5.03(2) \times 10^{1}$	$5.38(8) \times 10^{1}$	$5.38(8) \times 10^{1}$	$5.55(8) \times 10^{1}$	$5.55(8) \times 10^{1}$
ZIF-4	30.0	$4.25(9) \times 10^{-5}$	$4.31(6) \times 10^{-5}$	$6.30(3) \times 10^{1}$	$6.30(3) \times 10^{1}$	$6.770(4) \times 10^{1}$	$6.770(4) \times 10^{1}$	$6.840(5) \times 10^{1}$	$6.840(5) \times 10^{1}$
ZIF-8	24.0	$6.65(5) \times 10^{-6}$	$6.85(7) \times 10^{-6}$	$6.15(4) \times 10^{1}$	$6.15(4) \times 10^{1}$	$1.390(4) \times 10^{2}$	$1.390(4) \times 10^{2}$	$1.600(6) \times 10^{2}$	$1.600(6) \times 10^{2}$

Table S5: Converged properties for the COFs considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_{H} in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations without blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ t.-c.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
16083N2	30.0	$4.43(5) \times 10^{-6}$	$4.47(4) \times 10^{-6}$	$2.69(2) \times 10^{1}$	$2.69(2) \times 10^{1}$	$1.130(3) \times 10^{2}$	$1.130(3) \times 10^{2}$	$1.600(4) \times 10^{2}$	$1.600(4) \times 10^{2}$
18091N2	22.0	$9.0(1) \times 10^{-6}$	$9.4(1) \times 10^{-6}$	$5.33(8) \times 10^{1}$	$5.33(8) \times 10^{1}$	$1.840(5) \times 10^{2}$	$1.840(5) \times 10^{2}$	$2.330(4) \times 10^{2}$	$2.330(4) \times 10^{2}$
13030N2	26.0	0	0	0	0	0	0	0	0
15030N2	26.0	$3.18(3) \times 10^{-6}$	$3.23(2) \times 10^{-6}$	$2.27(3) \times 10^{1}$	$2.27(3) \times 10^{1}$	$1.020(4) \times 10^{2}$	$1.020(4) \times 10^{2}$	$1.460(2) \times 10^{2}$	$1.460(2) \times 10^{2}$
16411C2	26.0	$3.87(6) \times 10^{-6}$	$3.95(7) \times 10^{-6}$	$3.28(3) \times 10^{1}$	$3.28(3) \times 10^{1}$	$9.87(5) \times 10^{1}$	$9.87(5) \times 10^{1}$	$1.240(3) \times 10^{2}$	$1.240(3) \times 10^{2}$
17030N2	26.0	$3.68(3) \times 10^{-6}$	$3.73(2) \times 10^{-6}$	$1.85(1) \times 10^{1}$	$1.85(1) \times 10^{1}$	$8.67(2) \times 10^{1}$	$8.67(2) \times 10^{1}$	$1.320(4) \times 10^{2}$	$1.320(4) \times 10^{2}$
17120N2	26.0	$3.28(4) \times 10^{-5}$	$3.36(6) \times 10^{-5}$	$7.2(1) \times 10^{1}$	$7.2(1) \times 10^{1}$	$1.130(7) \times 10^{2}$	$1.130(7) \times 10^{2}$	$1.200(4) \times 10^{2}$	$1.200(4) \times 10^{2}$
17163N3	30.0	$7.98(6) \times 10^{-6}$	$8.02(8) \times 10^{-6}$	$2.95(1) \times 10^{1}$	$2.95(1) \times 10^{1}$	$1.340(5) \times 10^{2}$	$1.340(5) \times 10^{2}$	$1.950(3) \times 10^{2}$	$1.950(3) \times 10^{2}$
15072N2	26.0	$2.7(1) \times 10^{-5}$	$2.8(1) \times 10^{-5}$	$3.55(4) \times 10^{1}$	$3.55(4) \times 10^{1}$	$1.340(1) \times 10^{2}$	$1.340(1) \times 10^{2}$	$1.910(3) \times 10^{2}$	$1.910(3) \times 10^{2}$
17040N2	26.0	$4.80(4) \times 10^{-6}$	$4.92(3) \times 10^{-6}$	$4.07(3) \times 10^{1}$	$4.07(3) \times 10^{1}$	$1.300(3) \times 10^{2}$	$1.300(3) \times 10^{2}$	$1.590(5) \times 10^{2}$	$1.590(5) \times 10^{2}$

Table S6: Converged properties for the zeolites considered in this study, with (abbreviated with "t.-c.") and without (abbreviated with "trunc.") tail-corrections. Loadings θ in units of molecules per unit cell, Henry coefficients K_H in units of $\mathrm{mol\,Pa^{-1}\,kg^{-1}}$ and cutoffs in units of Å for simulations without blocked pockets.

name	cutoff	$K_{\mathrm{H,\ trunc.}}$	$K_{\mathrm{H,\ t.-c.}}$	$\theta_{5.8,\ \mathrm{trunc.}}$	$\theta_{5.8,\ \mathrm{t.-c.}}$	$\theta_{35,\ \mathrm{trunc.}}$	$\theta_{35,\ \mathrm{t.-c.}}$	$\theta_{65,\ \mathrm{trunc.}}$	$\theta_{65,\ \mathrm{t.-c.}}$
PAU	16.0	$1.09(4) \times 10^{-3}$	$1.26(5) \times 10^{-3}$	$2.090(7) \times 10^{2}$	$2.090(7) \times 10^{2}$	$2.200(6) \times 10^{2}$	$2.200(6) \times 10^{2}$	$2.24(2) \times 10^{2}$	$2.24(2) \times 10^{2}$
UEI	30.0	$1.27(3) \times 10^{-3}$	$1.30(8) \times 10^{-3}$	$1.0800(4) \times 10^{2}$	$1.0800(4) \times 10^{2}$	$1.0800(2) \times 10^{2}$	$1.0800(2) \times 10^{2}$	$1.0800(2) \times 10^{2}$	$1.0800(2) \times 10^{2}$
MTT	30.0	$5.45(4) \times 10^{-4}$	$5.6(1) \times 10^{-4}$	$9.6(2) \times 10^{1}$	$9.6(2) \times 10^{1}$	$1.310(9) \times 10^{2}$	$1.310(9) \times 10^{2}$	$1.42(1) \times 10^{2}$	$1.42(1) \times 10^{2}$
EMT	24.0	$7.1(3) \times 10^{-3}$	$7.4(3) \times 10^{-3}$	$1.800(4) \times 10^{2}$	$1.800(4) \times 10^{2}$	$2.260(9) \times 10^{2}$	$2.260(9) \times 10^{2}$	$2.360(7) \times 10^{2}$	$2.360(7) \times 10^{2}$
SFS	26.0	$5.9(2) \times 10^{-4}$	$6.1(1) \times 10^{-4}$	$1.530(7) \times 10^{2}$	$1.530(7) \times 10^{2}$	$1.710(8) \times 10^{2}$	$1.710(8) \times 10^{2}$	$1.750(7) \times 10^{2}$	$1.750(7) \times 10^{2}$
SFN	26.0	$6.6(1) \times 10^{-5}$	$6.84(9) \times 10^{-5}$	$1.340(7) \times 10^{2}$	$1.340(7) \times 10^{2}$	$1.590(4) \times 10^{2}$	$1.590(4) \times 10^{2}$	$1.640(9) \times 10^{2}$	$1.640(9) \times 10^{2}$
AFT	26.0	$5.6(2) \times 10^{-4}$	$5.9(2) \times 10^{-4}$	$2.040(9) \times 10^{2}$	$2.040(9) \times 10^{2}$	$2.280(9) \times 10^{2}$	$2.280(9) \times 10^{2}$	$2.320(6) \times 10^{2}$	$2.320(6) \times 10^{2}$
TSC	30.0	$4.0(2) \times 10^{-3}$	$4.1(3) \times 10^{-3}$	$1.920(4) \times 10^{2}$	$1.920(4) \times 10^{2}$	$2.530(4) \times 10^{2}$	$2.530(4) \times 10^{2}$	$2.69(2) \times 10^{2}$	$2.69(2) \times 10^{2}$
GOO	30.0	$9.1(2) \times 10^{-4}$	$9.3(4) \times 10^{-4}$	$8.800(7) \times 10^{1}$	$8.800(7) \times 10^{1}$	$8.820(2) \times 10^{1}$	$8.820(2) \times 10^{1}$	$8.820(1) \times 10^{1}$	$8.820(1) \times 10^{1}$
MFI	28.0	$1.40(2) \times 10^{-3}$	$1.46(3) \times 10^{-3}$	$1.310(8) \times 10^{2}$	$1.310(8) \times 10^{2}$	$1.430(5) \times 10^{2}$	$1.430(5) \times 10^{2}$	$1.470(9) \times 10^{2}$	$1.470(9) \times 10^{2}$

References

(S1) Leachman, J. W.; Jacobsen, R. T.; Lemmon, E. W.; Penoncello, S. G. *Thermodynamic Properties of Cryogenic Fluids*; International Cryogenics Monograph Series; Springer International Publishing: Cham, 2017.

(S2) Peng, D.-Y.; Robinson, D. B. A New Two-Constant Equation of State. *Ind. Eng. Chem. Fundam.* **1976**, *15*, 59–64.

(S3) Dubbeldam, D.; Torres-Knoop, A.; Walton, K. S. On the Inner Workings of Monte Carlo Codes. *Mol. Simul.* **2013**, *39*, 1253–1292.

(S4) Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*, 2nd ed.; Interdisciplinary Applied Mathematics 21; Springer: New York, NY, 2010; ZSCC: 0001424.

(S5) Jablonka, K. M.; Ongari, D.; Smit, B. *structure_comp*; 2019; `https://github.com/kjappelbaum/structure_comp`.

(S6) Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel, V.,; and Thirion, B. and Grisel, O. and Blondel, M. and Prettenhofer, P.,; and Weiss, R. and Dubourg, V. and Vanderplas, J. and Passos, A. and,; Cournapeau, D. and Brucher, M. and Perrot, M. and Duchesnay, E., Scikit-Learn: Machine Learning in {P}ython. ZSCC: NoCitationData[s1].

(S7) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and Tools for High-Throughput Geometry-Based Analysis of Crystalline Porous Materials. *Microporous Mesoporous Mater.* **2012**, *149*, 134–141.

(S8) Tibshirani, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, 2nd ed.; Springer Series in Statistics; Springer, 2017.

(S9) Haldoupis, E.; Nair, S.; Sholl, D. S. Pore Size Analysis of \textgreater250 000 Hypothetical Zeolites. *Phys. Chem. Chem. Phys.* **2011**, *13*, 5053–5060.

(S10) Matteoli, E.; Mansoori, G. A. A Simple Expression for Radial Distribution Functions of Pure Fluids and Mixtures. *J. Chem. Phys.* **1995**, *103*, 4672–4677.

(S11) Frenkel, D.; Smit, B. In *Understanding Molecular Simulation (Second Edition)*; Frenkel, D., Smit, B., Eds.; Academic Press: San Diego, 2002; pp 23–61.

(S12) Ongari, D.; Boyd, P. G.; Barthel, S.; Witman, M.; Haranczyk, M.; Smit, B. Accurate Characterization of the Pore Volume in Microporous Crystalline Materials. *Langmuir* **2017**, *33*, 14529–14538.

(S13) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

(S14) VandeVondele, J.; Hutter, J. Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases. *J. Chem. Phys.* **2007**, *127*, 114105.

(S15) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. Cp2k: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 15–25.

(S16) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(S17) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(S18) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.

(S19) Manz, T. A.; Sholl, D. S. Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2010**, *6*, 2455–2468.

(S20) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(S21) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Mol. Simul.* **2016**, *42*, 81–101.

(S22) Macedonia, M. D.; Maginn, E. J. A Biased Grand Canonical Monte Carlo Method for Simulating Adsorption Using All-Atom and Branched United Atom Models. *Mol. Phys.* **1999**, *96*, 1375–1390.

(S23) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of n-Alkanes. *J Phys Chem B* **1998**, *102*, 2569–2577.

(S24) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A Generic Force Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(S25) Ongari, D.; Yakutovich, A. V.; Talirz, L.; Smit, B. Building a consistent and reproducible database for adsorption evaluation in Covalent-Organic Frameworks. 2019; `https://archive.materialscloud.org/2019.0034/v1`.

(S26) Ismail, A. F.; Khulbe, K. C.; Matsuura, T. *Gas Separation Membranes: Polymeric and Inorganic*; Springer International Publishing: Cham Heidelberg New York Dordrecht London, 2015.