Classical Density Functional Theory as a Fast and Accurate Method for Adsorption Property Prediction of Porous Materials

Vincent Dufour-Décieux,[1] Philipp Rehner,[1] Johannes Schilling,[1]
Elias Moubarak,[2] Joachim Gross,[3] and André Bardow[1,*]

[1]*Energy & Process Systems Engineering, ETH Zürich, Zürich 8092, Switzerland*

[2]*Laboratory of Molecular Simulation, Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Switzerland*

[3]*Institute of Thermodynamics and Thermal Process Engineering, University of Stuttgart, 70569 Stuttgart, Germany*

(Dated: July 18, 2024)

Abstract

Abstract

Physical adsorption separation is vital for many industrial processes, prompting researchers to develop new materials for energy-efficient processes. Porous adsorbent materials are of particular interest due to their diverse design possibilities and computational screening has accelerated the search for optimal materials. Classical density functional theory (cDFT) has recently been used as a faster alternative to state-of-the-art computational methods for screening of porous materials. However, extensive validation of cDFT predictions has not been performed for many materials, in a wide range of conditions, and with guest molecules exhibiting strong Coulombic interactions. In this paper, we validate the cDFT predictions by calculating the adsorption properties for more than 500 Metal-Organic Frameworks with three adsorbate molecules (CH_4, N_2, and CO_2) and comparing them to state-of-the-art results from Grand Canonical Monte Carlo (GCMC) simulations. For CO_2, we introduce the computation of Coulombic interactions between the MOF and the molecule, which are necessary to accurately describe this system. Our results demonstrate cDFT's ability to accurately replicate GCMC adsorption isotherms and enthalpies of adsorption while needing a median time of only 6 minutes per material. These features position cDFT as a serious candidate for adsorption properties estimations of porous materials for a wide range of physical adsorption-based processes.

I. INTRODUCTION

Physical adsorption plays an important role across various industrial processes including separating chemicals (e.g., O_2/N_2 [1–3] or xenon from air [4]), radioactive off-gases treatment [5], climate change mitigation [6], cooling processes [7], and purification [8–10]. To lower cost, a key lever is to reduce the energy demand, leading to the ongoing development of new materials.

Porous materials, such as Metal-Organic Frameworks (MOFs), Covalent-Organic Frameworks (COFs), and zeolites, exhibit extensive design possibilities and have received considerable interest for physical adsorption [11–13]. By systematically exploring this vast design

* abardow@ethz.ch

space, researchers can tailor materials with favorable properties for their specific application. However, experimentally studying a wide variety of porous materials can be time-consuming. To mitigate the laborious process of synthesizing an extensive array of materials, researchers perform computational screenings to find the most promising materials for their application [12–16]. These computational screenings usually rely on Monte Carlo simulations, such as configurational-bias Monte Carlo (CBMC) and Grand Canonical Monte Carlo (GCMC) [12, 14, 17, 18] or, more recently, machine learning (ML) models trained on GCMC data [16, 19, 20].

Monte Carlo simulations are state-of-the-art for computing adsorption properties and have been used for computational screenings for more than a decade [21]. In initial screenings, Monte Carlo simulations are often used to compute Henry coefficients and enthalpies of adsorption of porous materials for a few thousand materials [22, 23]. These properties are subsequently employed to evaluate the suitability of these materials for applications, e.g., carbon capture, employing metrics such as selectivity and parasitic energy. However, these simple material metrics often exhibit weak correlation with more comprehensive process metrics [12, 24], prompting contemporary investigations to incorporate an additional layer of process considerations into their screening methodologies [12, 21, 25]. For process modeling and optimization, adsorption properties of the MOFs, such as the adsorption isotherms and enthalpies of adsorption, are needed in a wide range of temperatures. To accurately extrapolate data across temperatures using the Clausius-Clapeyron equation and to model mixture behavior with the Ideal Adsorbed Solution Theory (IAST), it is essential to compute adsorption isotherms and enthalpies of adsorption across the entire range of relevant pressures [26]. This range spans pressures from the Henry's law regime to the saturation pressure. However, computing full adsorption isotherms with GCMC simulations can take up to several days [27], which limits the scope of materials screened [12, 25]. GCMC simulations can also display numerical uncertainties in the estimation of the enthalpy of adsorption [26], propagating to the temperature extrapolation of adsorption isotherms with Clausius-Clapeyron equation.

As a faster alternative to Monte Carlo simulations, ML-based models have recently been developed and are able to accurately reproduce the Monte Carlo results [19, 20, 28]. However, training these models requires a substantial amount of data, usually coming from GCMC simulations, involving a time-intensive phase of training data generation. Furthermore,

as the ML-based methods are not physics-based, these models are usually trained on one property in specific conditions, such as the loading at a defined set of pressures for fixed temperature [19, 28] or the enthalpy of adsorption at one pressure and one temperature [20]. New training data must be generated when other properties or the same properties in new conditions are needed, e.g., in process optimization.

As a fast, accurate, physics-based alternative, classical Density Functional Theory (cDFT) has recently emerged to compute adsorption properties of porous materials. Screenings of MOFs have already been performed for applications separating or storing noble gas [29], toxic gases [30], dihydrogen [31, 32], and carbon isotopes [33]. The computational cost of cDFT was shown to be much lower compared to GCMC simulations, on the order of tens of seconds for the computation of one loading [32]. To test the reliability of cDFT predictions, validation has been performed by comparing cDFT predictions with GCMC results or experiments. Fu *et al.* [32, 34] compare Monte Carlo and cDFT results for the loading of methane and dihydrogen for 1200 MOFs and show very good agreement. However, in both of these studies, the molecules are considered a single Lennard-Jones site and the comparisons were done for two pressures at one temperature. Other validations compare full adsorption isotherms predicted by cDFT with GCMC simulations or experiments, but they are limited to fewer than 10 MOFs [28–31, 33, 35, 36]. Therefore, a comprehensive validation of cDFT predictions with GCMC results is needed on a wide range of pressures and temperatures for single and non-single Lennard-Jones site molecules to test the reliability of cDFT predictions for screenings of thousands of MOFs. In addition, the studies mentioned above do not include validation of the cDFT predictions of the enthalpies of adsorption, which is a crucial property for process modeling and optimization as it is used to extrapolate isotherms in temperatures and to study heat exchanges.

In this research paper, we perform this comprehensive validation of cDFT by comparing cDFT predictions of adsorption isotherms and enthalpies of adsorption with GCMC results for more than 500 MOFs, using CH_4, N_2, and CO_2 as guest molecules. N_2 and CO_2 are important molecules to study as they appear in important separation processes such as carbon capture and no extensive validation for these molecules have been performed. To improve cDFT predictions accuracy for CO_2, Coulombic interactions between the porous material and the guest molecule are introduced in the framework. We confirm the low computational time of cDFT by computing adsorption isotherms in a median time of 6

CPU-minutes on one node of an AMD Ryzen Threadripper PRO 3975WX, while closely reproducting GCMC simulation results for the three adsorbate molecules. We demonstrate the ability of cDFT to faithfully replicate GCMC predictions of enthalpies of adsorption while exhibiting low numerical error. These comprehensive comparisons underscore the potential of cDFT as a relevant option in the computational assessment of porous materials for physical adsorption.

In Sec. II, we present the cDFT method and the key developments to generalize this method. We then present the GCMC calculations that are used as a benchmark in our study and how the cDFT setup is designed to match GCMC results (Sec. III). In Sec. IV, we compare the cDFT predictions to the GCMC benchmark for a set of adsorption isotherms, enthalpy of adsorption curves, and their computational cost. In Sec. V, we explore the underlying reasons for cDFT deviations from GCMC results. We finally conclude in Sec. VI on the potential of cDFT in the context of materials screening for physical adsorption-based processes.

II. VERSATILE AND EFFICIENT: CLASSICAL DENSITY FUNCTIONAL THEORY

In this section, we present a succinct overview of the computational procedure to calculate the adsorbate molecule loading within a porous material using cDFT. We emphasize key advancements that broaden the scope of the cDFT framework used here [37], specifically its applicability to 1) solids featuring non-orthorhombic unit cells and 2) adsorbate molecules that exhibit strong Coulombic interactions with the material. We also detail how we incorporate enthalpies of adsorption calculations.

A. Overview of classical Density Functional Theory based on PC-SAFT

The cDFT framework describes systems in the grand canonical ensemble μVT where the grand potential Ω represents the thermodynamic potential. In cDFT, the grand potential Ω is minimized with respect to the density profiles $\rho_k(\mathbf{r})$ of every adsorbate k to find the equilibrium state of the system. The grand potential is described with the following

relationship:

$$\Omega\left[\{\rho_k(\mathbf{r})\}\right] = F\left[\{\rho_k(\mathbf{r})\}\right] + \sum_i^{N_c} \int \rho_i(\mathbf{r})(V_i^{\text{ext}}(\mathbf{r}) - \mu_i)d\mathbf{r}.$$

$\rho_k(\mathbf{r})$ is the density profile of adsorbate k which depends on the position $\mathbf{r}$. N_c is the number of adsorbates, and V_k^{ext} is the external potential used to describe the interactions between the adsorbate k and the porous material. The external potential V_k^{ext} includes the van der Waals interactions and the Coulombic interactions:

$$V_k^{\text{ext}}(\mathbf{r}) = V_k^{\text{VdW}}(\mathbf{r}) + V_k^{\text{Coulomb}}(\mathbf{r}). \tag{1}$$

The van der Waals term V_k^{VdW} is computed as described by Kessler *et al.* [38] and the Coulombic term V_k^{Coulomb} is described in Sec. II B. μ_k is the chemical potential of the adsorbate k. F is the Helmholtz energy functional describing interactions within the fluid phase and is described here by the functional developed by Sauer and Gross [39] based on the PC-SAFT equation of state [40].

PC-SAFT calculates the Helmholtz energy as the sum of contributions from intermolecular interactions. In this study, we consider the following contributions:

$$F[\{\rho(\mathbf{r})\}] = F^{\text{ig}}[\{\rho(\mathbf{r})\}] + F^{\text{hs}}[\{\rho(\mathbf{r})\}] + F^{\text{hc}}[\{\rho(\mathbf{r})\}] + F^{\text{disp}}[\{\rho(\mathbf{r})\}] + F^{\text{mp}}[\{\rho(\mathbf{r})\}], \tag{2}$$

where the individual contributions are the ideal gas reference (ig), hard-sphere repulsion (hs), hard-chain formation (hc), dispersive attraction (disp), and multipolar interactions (mp) [39].

The evaluation of the Helmholtz energy functional and the minimization of the grand potential is provided by the FeO_s Python package [37]. Once the grand potential is minimized, the equilibrium density is integrated over a unit cell to calculate the loading. Repeating the calculations for multiple bulk pressures yields adsorption isotherms, i.e., the loading as a function of the pressure at constant temperature. The enthalpy of adsorption at each pressure gives the enthalpy of adsorption curves as detailed in Sec. II B.

We also compute the Henry coefficients and zero-loading enthalpy using a pragmatic approach. The lowest pressure of the isotherms is chosen to be at the end of the Henry regime [26], ensuring that any lower pressure will also be in the Henry regime. We calculate the loading and enthalpy of adsorption at half of this lowest pressure to provide an additional point within the Henry regime. To determine the Henry coefficient, we perform a linear

regression using three points: zero pressure, half of the lowest pressure, and the lowest pressure. The slope of this regression gives us the Henry coefficient. To determine the zero-loading enthalpy, we first perform a linear regression using the enthalpy of adsorption values at half of the lowest pressure and at the lowest pressure. By extrapolating this regression to zero pressure, we obtain the zero-loading enthalpy.

B. Key developments for generalizable cDFT

To generalize the cDFT framework used here, we extend it by three key developments: 1) extension for non-orthorhombic unit cells, 2) adding Coulombic interactions in the external potential, and 3) calculation of the enthalpy of adsorption. Here, we describe these developments succinctly and give more details in the Supplementary Materials, Sec. S1.

To generalize cDFT to more porous materials, the cDFT method is extended to non-orthorhombic unit cells. The unit cell is discretized along its skewed axes and a coordinate transform is applied to obtain Cartesian coordinates, which are required to calculate distances for the external potential. In cDFT, the skewness of the coordinate system also needs to be accounted for when evaluating the (non-local) Helmholtz energy functional. The applied coordinate transform and rapid evaluation of convolution integrals in skewed coordinates are included in the Supplementary Information, Sec. S1A.

Previous work [30] only considers the Van der Waals interactions between the MOF and the guest molecules, even if the guest molecules could exhibit Coulombic interactions with the MOF. As guest molecules of high interest, such as CO_2, can have strong Coulombic interactions with porous materials, considering these interactions is crucial to obtain accurate predictions of adsorption properties. To describe these Coulombic interactions, we employ a method similar to that used by Hong *et al.* [41] in a different context. To account for the orientation of the guest molecule within the external potential, they sample the Van der Waals interactions between a guest molecule and the solid at various orientations at a given position $\mathbf{r}$. We extend this approach by including the Coulombic interactions V^{Coulomb} between the guest molecule and the MOF. To minimize the number of orientations sampled and consequently the computational cost, preliminary calculations of likely orientations are estimated, and then 10 orientations are computed. The minimum energy is used for the external potential at this position $\mathbf{r}$. More details on the initial estimate of optimal directions

are included in the Supplementary Materials, Sec. S1B.

In addition to the adsorption isotherm, an important property for modeling and designing adsorption processes is the enthalpy of adsorption. The enthalpy of adsorption is needed, for example, to calculate heat requirements in adsorption processes or to extrapolate adsorption isotherms with the Clausius-Clapeyron relation. In this work, we calculate the enthalpy of adsorption analytically from cDFT. The enthalpy of adsorption for every adsorbate Δh_k^{ads} is calculated from:

$$0 = \sum_j \left(\frac{\partial N_k}{\partial \mu_j}\right)_T \Delta h_j^{\text{ads}} + T \left(\frac{\partial N_k}{\partial T}\right)_{p,x_i}, k = 1, ..., N_c \tag{3}$$

with N_k the number of adsorbed molecules of adsorbate k, μ_k the chemical potential, temperature T, (bulk) pressure p, composition of the bulk phase x_i, and the number of components N_c. The resulting evaluation of the enthalpy of adsorption is numerically robust and fast relative to the calculation of the density profile itself. Details on the derivation and the calculation of partial derivatives of density profiles are included in the Supplementary Material, Sec. S1C.

The setup of cDFT provides several degrees of freedom (including, for example, the force field for the external potential). Since the goal of this research paper is to show that cDFT calculations can reproduce GCMC predictions, the choice of these parameters is partially given by the parameters used in GCMC. Therefore, in the next section, we present the GCMC data that is used as a benchmark and then define the choice of the degrees of freedom for cDFT.

III. GCMC AND CDFT SETUP

A. GCMC database description

A database containing adsorption isotherms, Henry coefficients, enthalpies of adsorption and zero-loading enthalpies of adsorption was computed with GCMC in the work of Moubarak *et al.* [26]. This database contains the adsorption properties of N_2 and CO_2 at 298.15 K for 500 MOFs. For 50 of these MOFs and with both N_2 and CO_2, the adsorption properties were computed at 4 additional temperatures (323.15 K, 348.15 K, 373.15 K, and 398.15 K). Another database [42] contains adsorption properties of CH_4 at 298.15 K for

1050 MOFs (partially different from the N_2 and CO_2 set). The adsorption properties were computed for 10 to 30 pressures between the end of the Henry regime and the saturation pressure (see Ref. [26] for more explanation). The enthalpies of adsorption computed with RASPA are corrected to consider non-ideal gas behavior as described in Moubarak *et al.* [42].

All the details associated with these GCMC simulations are described by Moubarak *et al.* [26, 42], and here we report the choices most relevant for our study. The GCMC simulations used the Universal Force field (UFF) [43] to describe the framework atoms and the Transferable Potentials for Phase Equilibria (TraPPE) force field [44] to describe the adsorbates. The porous materials' atomic charges were computed using either the Density Derived Electrostatic and Chemical (DDEC) theory [45], or the extended charge equilibration (EQeq) method [46] for the experimental and in silico structures respectively. Pore blocking, i.e., regions of the porous medium that are inaccessible to the adsorbate molecules, is done with the software Zeo++ [47] using a radius of 1.492 Å for CH_4, 1.324 Å for N_2, and 1.22 Å for CO_2, a number of samples of 100 $Å^{-3}$, and the UFF radii as atomic radius.

To obtain an accurate comparison of computational costs, GCMC simulations for three MOFs with different unit cell sizes were re-run with the three adsorbates considered here (CH_4, N_2, and CO_2) in the same conditions as the database with the same working machine as cDFT calculations (AMD Ryzen Threadripper PRO 3975WX).

B. Deviation metric definition

To compare two adsorption isotherms or two enthalpy of adsorption curves, we compute the mean absolute relative deviation (MARD) between the two curves:

$$\mathrm{MARD} = \frac{1}{N_{\mathrm{pressures}}} \sum_{i}^{N_{\mathrm{pressures}}} \frac{|y_i^{\mathrm{test}} - y_i^{\mathrm{ref}}|}{|y_i^{\mathrm{ref}}|}, \tag{4}$$

where $N_{\mathrm{pressures}}$ is the number of pressures in the two curves (same pressures for both curves), and $y_i^{\mathrm{test/ref}}$ is the loading or the enthalpy of adsorption for the test/ref case at pressure i. When cDFT and GCMC predictions are compared, the reference is always GCMC. In cases where two cDFT or two GCMC curves are compared, the reference is defined when these comparisons occur. An absolute error was chosen as a non-absolute one would lead to canceling errors if an overestimation and underestimation happens within one comparison.

TABLE I. Parameters used for cDFT calculations.

Adsorbate	CO_2	N_2	CH_4
Coulombic interactions ($V^{Coulomb}$ in Eq. 1)	Yes	No	
Solid	UFF [43]		
Fluid description for external potential (for V^{ext} in Eq. 1)	TraPPE [44]	PC-SAFT Non-polar [40]	
Fluid description for Helmholtz functional (for F in Eq. 2)	PC-SAFT Polar [48]	PC-SAFT Non-polar [40]	
Grid density ($Å^{-1}$)	2		

Additional metrics are shown in the Supplemental Materials to give insights into the question of overestimation and underestimation.

C. Parameters for cDFT

cDFT calculations need the definition of five main degrees of freedom per adsorbate molecule. (see Table III C): 1) the inclusion of Coulombic interactions in the calculation of the external potential, 2) the parameters to describe the MOFs, 3) the description of the fluid for the calculation of the external potential, 4) the description of the fluid in the Helmholtz energy functional, and 5) the grid point density. All these five parameters are summarized in Table III C. Other minor degrees of freedom need to be defined, which are discussed in the Supplementary Materials Sec. S2A.

Coulombic interactions, as described briefly in Sec. II B and more thoroughly in the Supplementary Materials, Sec. S2B, can be included in the calculation of the external potential at the cost of additional computations. As CH_4 is known to have negligible Coulombic interactions, these interactions are not included, which corresponds to setting the V^{Coulomb} term of Eq. 1 to 0. For N_2 and CO_2, an analysis of the accuracy of adsorption isotherms and enthalpy of adsorption predictions compared with GCMC for 100 MOFs is detailed in

the Supplementary Materials Sec. S2B. This analysis shows that including the Coulombic interactions for CO_2 makes the cDFT predictions more closely match the GCMC calculations while the effect for N_2 is slightly unfavorable. The observation for N_2 comes from the fact that cDFT predictions slightly overestimate GCMC predictions without the Coulombic interactions. This overestimation increases when adding Coulombic interactions. The initial overestimation without Coulombic interaction might be attributed to PC-SAFT parameters compensating for their limited consideration of Coulombic interactions by enhancing dispersion interactions. Therefore, Coulombic interactions V^{Coulomb} are included for the calculations of the external potential when CO_2 is the adsorbate molecule but not for CH_4 and N_2.

The atoms of the MOFs are usually described with the UFF force field [43] or the DREIDING force field [49]. Following the GCMC database discussed in Sec. III, we opt to apply the UFF force field for characterizing the atoms in the MOFs when computing the external potential for cDFT.

The description of the fluid for the external potential depends on the inclusion of the Coulombic interactions: if these interactions are included, the parameters used for describing the adsorbate molecule need to resolve the orientations, which is not the case for PC-SAFT parameters. Therefore, we use the TraPPE force field [44] to describe CO_2 in the calculation of the external potential. For the other adsorbates, we consider two options: the set of non-polar PC-SAFT parameters by Gross and Sadowski [40] and the set of polar PC-SAFT parameters by Gross [48]. The non-polar set of parameters neglects the multipolar term of Eq. 2, whereas it is included for the polar set of parameters. As CH_4 is a non-polar molecule, we use the set of non-polar parameters. For N_2, we want to be consistent in choosing PC-SAFT parameters for calculating the external potential and for the Helmholtz energy functional. In the Supplementary Materials, Sec. S2B, we compare these two sets of parameters and show that the two options produce similar predictions. Therefore, for simplicity, we choose the set of non-polar parameters.

For consistency, CH_4 and N_2 are also described with the set of non-polar parameters for the Helmholtz functional calculation. For CO_2, we compare the two sets of parameters in the Supplementary Materials, Sec. S2B, and we show that using the polar set of parameters delivers predictions matching the GCMC data more closely, which is logical as CO_2 has some quadrupole moment.

To numerically minimize the grand potential Ω, the unit cell is discretized into a set of equidistant grid points. The number of grid points is a key parameter for the trade-off between accuracy and speed of computation, as more grid points result in more accurate results but more expensive computations. The density of grid points has been studied to find a compromise between accuracy, computational effort, and convergence (see Supplementary Materials, Sec. S2C). The analysis shows that a grid point density of $2\,\text{Å}^{-1}$ is an appropriate value to have high accuracy while keeping low computational costs. Increasing the grid point density to $2.5\,\text{Å}^{-1}$ results in an average of 0.5 % MARD change for the adsorption isotherms predictions and always less than 5.5 %, which is relatively low. In the interest of computational cost, the grid point density is set to $2\,\text{Å}^{-1}$. However, some enthalpies of adsorption display significant deviations at $2\,\text{Å}^{-1}$ that we attribute to numerical errors (see Supplementary Materials, Sec. S2C). When notable numerical errors arise in the enthalpy of adsorption computed with cDFT (occurring in approximately 7 % of materials), further calculations are conducted at higher grid densities, with an increment of $0.5\,\text{Å}^{-1}$ reaching up to $4\,\text{Å}^{-1}$. Despite these adjustments, some enthalpies of adsorption continue to exhibit notable numerical errors even at a grid density of $4\,\text{Å}^{-1}$. These significant deviations are identified by assessing whether the enthalpy of adsorption varies by more than $10\,\text{kJ mol}^{-1}$ between two consecutive pressures. The computational expenses shown in Sec. IV C include all the examined grid densities.

IV. COMPARISON OF CDFT AND GCMC PREDICTIONS

A. Adsorption isotherms

To validate the cDFT calculations, we compare the predictions of cDFT with GCMC predictions, the state-of-the-art for computing adsorption properties. In Fig. 1, the cumulative curves of the MARDs between adsorption isotherms predicted by cDFT and GCMC are shown for the three adsorbate molecules CH_4, N_2, and CO_2. To offer a comparison with a common practice among the community, the cumulative curves of the MARD of switching from the UFF [43] to the DREIDING force field [49] are also shown. The data is taken for around 700 MOFs for N_2 and CO_2 from the CRAFTED database calculated by Oliveira *et al.* [50]. We computed the MARD using UFF as a reference. The data is at 298.15 K, and

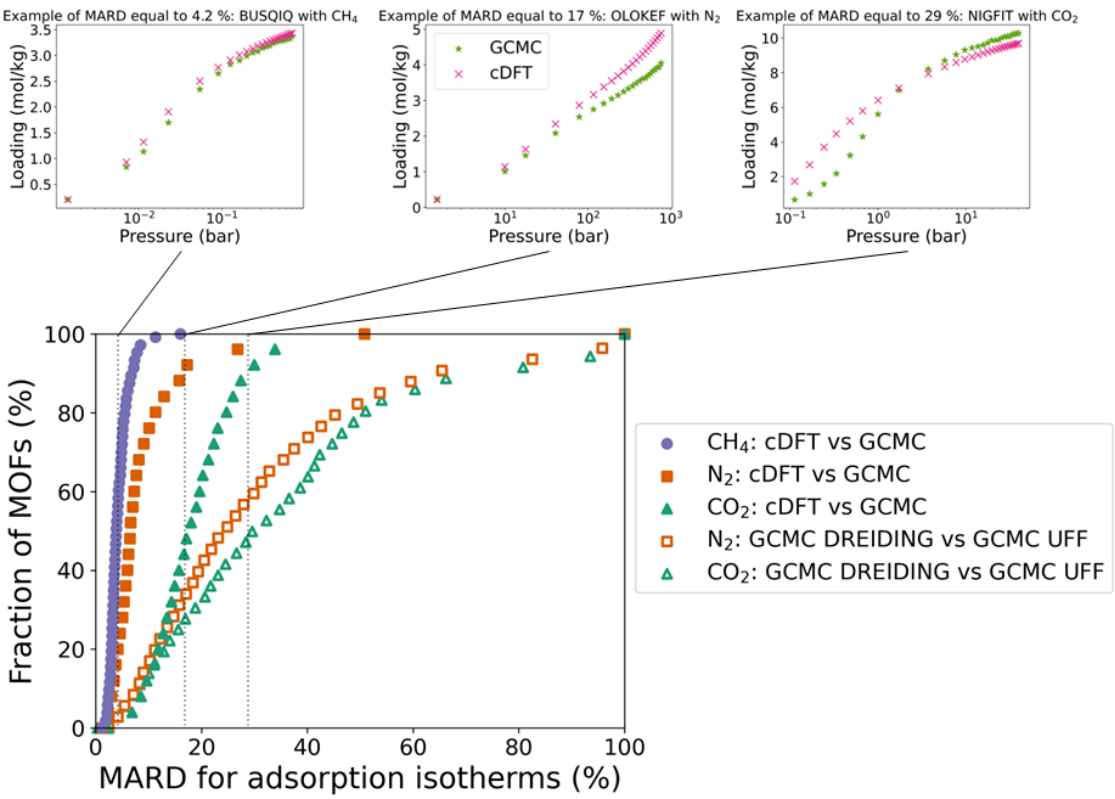


FIG. 1. Cumulative curves of the MARDs between cDFT and GCMC computations for adsorption isotherms of three adsorbate molecules CH_4, N_2, and CO_2. The top row of plots serves to illustrate MARD values of 4.2 %, 17 %, and 29 %, offering an intuitive grasp of the meaning of MARD between two adsorption isotherms. As a point-of-comparison, we include MARD comparisons between adsorption isotherms computed using the UFF [43] and DREIDING [49] force fields from another dataset recently published by Oliveira *et al.* [50]. The maximum MARD shown, for clarity, is 100 %. However, for the MARD between using UFF and DREIDING force field for N_2 and CO_2, there are instances (38/721 and 19/705, respectively) surpassing the 100 % threshold and 1/499 case for the comparison between cDFT and GCMC for CO_2. All isotherms were calculated at 298.15 K.

the DDEC charges are used. Note that the set of MOFs studied in the CRAFTED database differs from those studied here. Switching from the UFF to the DREDING forcefields and switching from GCMC calculations to cDFT calculations are two independent decisions and their deviations would be additive. The inclusion of the deviations for switching from UFF to DREIDING offers a comparison with common practice in the community.

For CH_4, the MARD has a median of 3.9 % and is below 16 % for all materials, which

can be regarded as an almost perfect agreement. As the methods of GCMC and cDFT are intrinsically different, small inherent deviations always occur. CH_4 is the case where the underlying molecular model in GCMC and cDFT exhibit the closest agreement because, in both models, CH_4 is described as a single Lennard-Jones site without Coulombic interactions. We provide a reference point for the order of magnitude of the MARD between cDFT and GCMC: Gowers *et al.* [27] compare the adsorption isotherms predicted by five implementations of GCMC (Cassandra [51], DL Monte [52], MuSiC [53], RASPA [54], and Towhee [55]) and show that in the case where no electrostatic interactions are considered, the MARD between the implementations is on average of 3.1 %, and can be up to 5.2 %. Therefore, the median MARD between cDFT and GCMC for CH_4 is within the MARD of different implementations of GCMC and thus very accurate. With CH_4, cDFT calculations did not converge for 22/1050 MOFs for at least one pressure and were removed from the analysis.

For N_2, the MARD has a median of 6.5 % whilst consistently staying under 51 % across all materials. Illustrated in Figure 1, the use of the DREIDING force field [49] as an alternative to UFF [43]——a common choice for GCMC-based adsorption isotherm computations—— yields a median MARD of 24 % with MARD values going up to 588 %. This observation shows that opting for cDFT instead of GCMC when N_2 is the adsorbate molecule results in a smaller change of the results. N_2 exhibits higher MARD than CH_4, possibly because the orientations of N_2 molecules can play a role in GCMC that is currently not resolved with cDFT as N_2 is not described as a single Lennard-Jones site, contrary to CH_4. With N_2, 1/500 MOF had at least one pressure where cDFT calculations did not converge and was removed from the analysis.

For CO_2, Coulombic interactions are considered. The MARD has a median of 17.5 % and is always below 127 %. For this scenario, switching from UFF force field to DREIDING results in a median MARD of 30 % with MARD values going up to 1339 %. Consequently, the deviations between cDFT and GCMC predictions for CO_2 are much lower than uncertainties in the GCMC force field decision. cDFT predictions deviate more with CO_2 as the adsorbate molecule than with CH_4 or N_2. A possible explanation is that the non-consideration of the orientations of CO_2 molecules has a greater effect as CO_2 exhibits stronger Coulombic interactions than CH_4 and N_2 or that the Coulombic interactions sampling could be made finer. With CO_2, 1/500 MOF had at least one pressure where cDFT calculations did not

converge and was removed from the analysis.

Calculations at temperatures other than 298.15 K show similar deviations between cDFT and GCMC predictions in the available range 323 K to 393 K (see the Supplementary Materials, Sec. S3). Overall, cDFT almost perfectly reproduces adsorption isotherms of GCMC for CH_4 and closely matches them for N_2 and CO_2. As both models are inherently different, deviations are expected, and we show that these deviations are below those observed when switching force fields.

The MARD on the full adsorption isotherms does not give a complete picture of the different regimes observed in an adsorption isotherm: in the Supplementary Materials Sec. S4, more details are given on the behavior at low and high pressure. At low pressure, in the Henry regime, the loading mostly depends on the adsorbate-MOF interactions, whereas at higher pressure, the loading mostly depends on the pore shape of the MOF. In the Supplementary Materials Sec. S4, these different regimes are more detailed. We show, for example, that the Henry coefficients calculated by cDFT deviate from GCMC calculations by less than 10 % for 99% of the MOFs for CH_4 and is within a factor 2 from GCMC for 95% and 84% of the MOFs for N_2 and CO_2 respectively. At high pressure, cDFT predictions closely match the loading computed by GCMC for CH_4 and N_2 with deviations lower than 10% in most cases. However, a small underestimation of the loading at high pressure is observed for CO_2.

B. Enthalpy of adsorption

Fig. 2 displays the results of a similar analysis to Fig. 1, but for the enthalpy of adsorption instead of the loading. Due to its non-stochastic nature, enthalpies of adsorption computed by cDFT display much smaller numerical uncertainties than the ones computed by GCMC. In order to decouple the deviations that come from the numerical uncertainties in GCMC enthalpies of adsorption and the deviations coming from the differences between cDFT and GCMC, we also show the comparison of the zero-loading enthalpies of adsorption. These values can be precisely computed with GCMC and are important as they give the temperature dependence of the Henry coefficient. A comparison of the zero-loading enthalpies of adsorption is shown in Fig. 3.

The deviations between cDFT and GCMC are low for the enthalpy of adsorption (Fig.

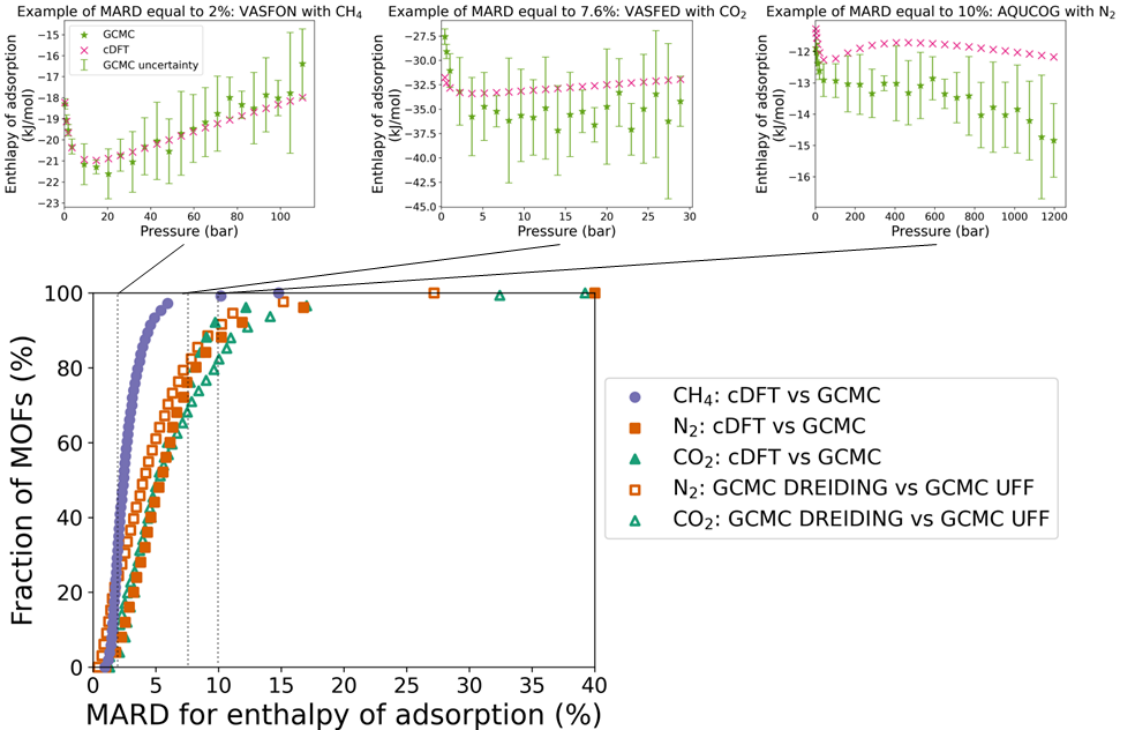


FIG. 2. Cumulative curves of the MARD between cDFT and GCMC computations for the enthalpy of adsorption of three adsorbate molecules CH_4, N_2, and CO_2. The top row of plots serves to illustrate MARD values of 2 %, 7.6 %, and 10 %, offering an intuitive grasp of the meaning of MARD between two enthalpies of adsorption curves. The numerical uncertainty of the GCMC calculations is shown. As a reference, we include MARD comparisons between enthalpies of adsorption computed using the UFF [43] and DREIDING [49] force fields, from another dataset recently published by Oliveira *et al.* [50]. The maximum MARD shown, for clarity, is 40 %. However, for the MARD between using cDFT and GCMC for N_2 and CO_2, there are instances (1/499 and 5/499 respectively) surpassing the 40 % threshold. All enthalpies of adsorption were calculated at 298.15 K.

2): the median MARD is 2.4 % for CH_4 as the adsorbate molecule, 5.3 % for N_2 and 5.2 % for CO_2. For N_2 and CO_2, these MARD values are close to the deviation occurring when switching from UFF to DREIDING force field, whereas the values for CH_4 are lower. However, some of the deviations are due to the numerical uncertainties in GCMC calculations. The zero-loading enthalpy of adsorption has a very small numerical uncertainty when calculated with GCMC and allows for more precise comparison. The zero-loading enthalpy of adsorption computed with cDFT matches very closely the one computed with GCMC for CH_4 ($R^2 = 0.9987$), and matches reasonably close for N_2 and CO_2 (respectively, $R^2 = 0.9011$

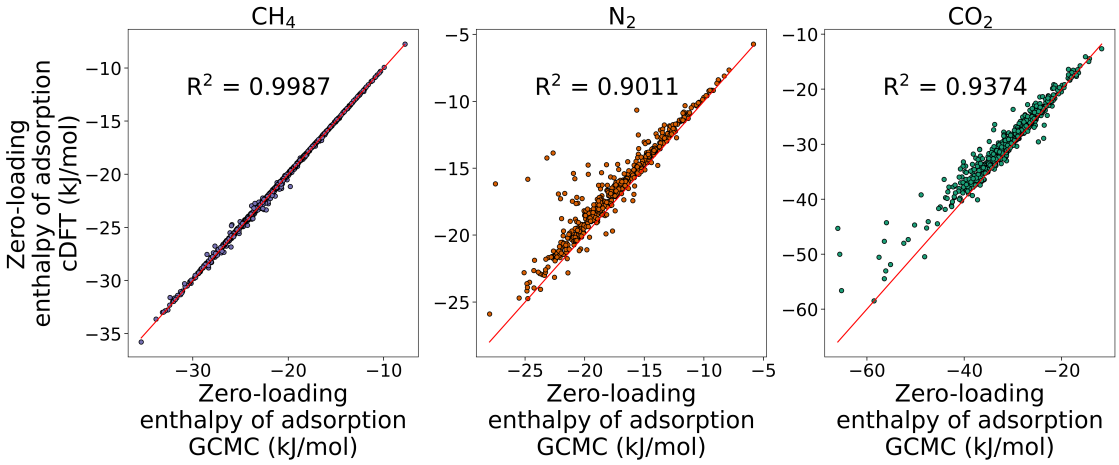


FIG. 3. Log-log parity plot for the zero-loading enthalpy of adsorption predicted by cDFT versus by GCMC for the three molecules (CH_4, N_2, CO_2). Each point corresponds to one MOF. The cDFT predictions of the zero-loading enthalpy of adsorption perfectly matches the GCMC results for CH_4 and are in good agreement for N_2 and CO_2.

and $R^2 = 0.9374$) (see Fig. 3).

As mentioned in Sec. III C, the enthalpies of adsorption computed by cDFT sometimes display unreasonable values attributed to numerical errors (see Supplementary Sec. S3C for more details). These numerical errors can be partially removed by increasing the grid density. In this work, the grid density is increased by $0.5\,\text{Å}^{-1}$ to a maximum $4\,\text{Å}^{-1}$ to avoid these errors. At a grid density of $4\,\text{Å}^{-1}$, still 6/499 materials for CO_2 and 1/499 for N_2 exhibit these large numerical errors at specific pressures. In Fig. 2, the material for N_2 and the 5 materials for CO_2 that exhibit MARD higher than 40 % for the enthalpies of adsorption show large numerical errors at specific pressures.

For the enthalpy of adsorption, the MARD between cDFT and GCMC also does not depend on temperature, as shown in the Supplementary Materials Sec S3.

C. Computational cost comparison

In materials screening, the computational cost of the property estimation method determines the number of materials that can be screened within a specific time. cDFT calculations take a median run time of 6 min per adsorption isotherm, with a minimum of 8 s and a maximum of 40 h (see Fig. 4). The computational time increases exponentially with

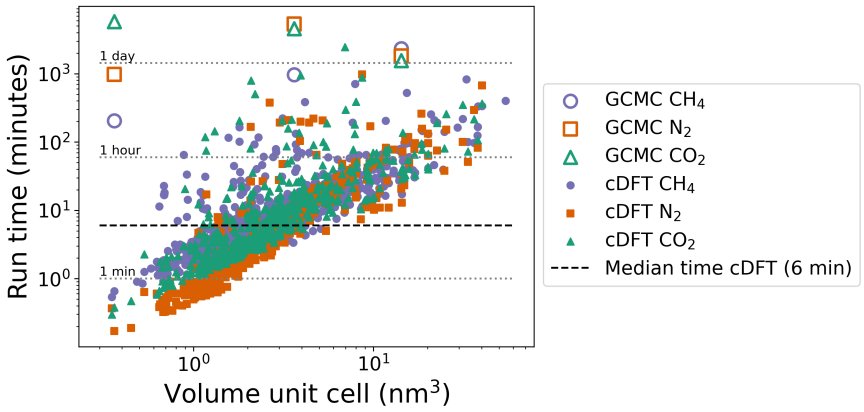


FIG. 4. Run time of cDFT and GCMC simulations as a function of the volume of the unit cell. The GCMC adsorption isotherms were re-run for 3 MOFs covering a significant range of unit cell volume, and the precise run time was measured (empty symbols).

the volume of the unit cell as a larger unit cell requires more grid points for a fixed grid point density. When the volume is multiplied by 10, the computational cost multiplies by 33. Computational costs that are particularly high for a given unit cell volume come from the re-calculation of the cDFT predictions at higher grid density when the enthalpies of adsorption exhibit large numerical errors as mentioned in Sec. III C.

For the three MOFs (OLOKEF, WUSSEK, and XAKZAL in the database) where precise computational costs were calculated, GCMC calculations took between 3.4 h and 96 h on one node of an AMD Ryzen Threadripper PRO 3975WX, the same machine used for the cDFT calculations. At an equal volume of the unit cell, cDFT is almost always faster, ranging from 15 000 times faster (small volume unit cell MOF (0.36 nm^3) for CO_2) to 42 times faster (large volume unit cell MOF (14 nm^3)), resulting in a significant amount of time saved that can allow researchers to screen around 100 times more materials within the same time. Note that for both GCMC and cDFT, the computational cost of computing the enthalpies of adsorption is negligible, once the density profiles have been computed.

V. ASSESSMENT OF CDFT FEATURES AND LIMITATIONS

In this paper, we aim to validate cDFT as a potential tool for adsorption property prediction of porous materials. We demonstrate the generalization of two primary features of cDFT previously mentioned in the literature: its low computational cost and minimal

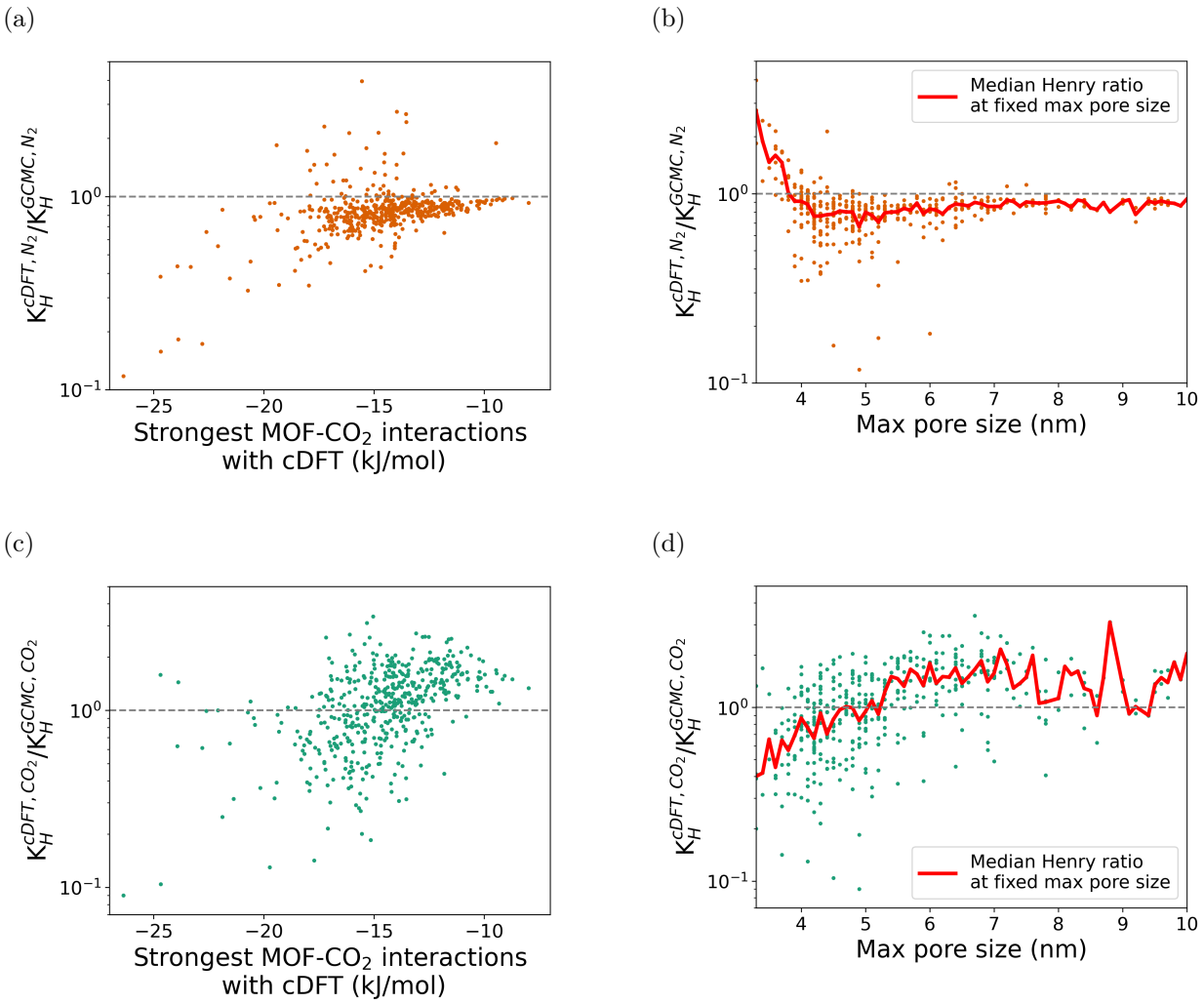


FIG. 5. Scatter plot of the strongest MOF-CO_2 interactions computed with cDFT ((a) and (c)) and the maximal pore size ((b) and (d)) with the ratio of the Henry coefficients computed with cDFT K_H^{cDFT} over the ones computed with GCMC K_H^{GCMC} for N_2 ((a) and (b)) and CO_2 ((c) and (d)). The y axes are in logarithm scale. Each point corresponds to one MOF.

numerical uncertainty in the computation of the enthalpy of adsorption. These attributes position cDFT as a relevant tool for efficiently screening materials. The cost-effectiveness facilitates the screening of a large number of materials, while the low numerical uncertainty in enthalpies at high pressure enables robust temperature extrapolation of adsorption isotherms using the Clausius-Clapeyron relation, which is necessary for gradient-based optimization for process optimizations. However, these advantages come with a trade-off in accuracy, stemming from the discretization of adsorbate positions or the exclusion of adsorbate ori-

entation as a degree of freedom during density profile optimization. Our study reveals that this loss in accuracy is negligible for CH_4 as the adsorbate but becomes more pronounced for N_2 and CO_2. Understanding the materials prone to a larger deviation between cDFT and GCMC is crucial when applying cDFT to specific materials.

In Fig. 5, we present insights into instances where cDFT exhibits significant deviations from GCMC. We employ the ratio of the Henry coefficient calculated using cDFT (K_H^{cDFT}) to that obtained with GCMC (K_H^{GCMC}). This ratio is compared to two descriptors: the strongest MOF-CO_2 interactions and the maximal pore size. The strongest MOF-CO_2 interactions are the minimum energy interactions at a grid point between a MOF and CO_2. Strong MOF-CO_2 interactions generally mean strong Coulombic interactions. The maximal pore size is the size of the biggest pore in the MOF structure. The pore size distribution is computed with the Zeo++ software [56]. A small maximal pore size means that all the pores in the MOF are quite small.

For N_2 as the adsorbate in Fig. 5a, stronger interactions between the MOF and CO_2 correspond to underestimation of the Henry coefficients by cDFT. If we consider that strong MOF-CO_2 interactions come from strong Coulombic interactions, the underestimation of the Henry coefficients by cDFT for N_2 could be corrected by the inclusion of the Coulombic interactions for N_2 for the cases where they seem non-negligible. In Fig. 5a, we can also observe that some Henry coefficients are overestimated by cDFT, which are not explained by the strength of the MOF-CO_2 interactions. In Fig. 5b, we can see that the points that are overestimated always occur when the maximal pore size in the MOF is smaller than 5 nm. When the maximal pore size is small, all the pores of the MOF are small. Therefore, small pores in a MOF lead to an overestimation of the Henry coefficients for N_2. A plausible explanation is the following: the orientation of the adsorbate probably plays a larger role when the adsorbate is close to the atoms of the MOF and in smaller pores, the adsorbate is more often close to the MOF atoms. An overestimation may then result from grid points where the external potential calculated with PC-SAFT predicts a grid point where adsorbate molecules are allowed but where no orientation would allow N_2.

In Fig. 5c and 5d, we can observe that similar conclusions can be drawn for CO_2: the Henry coefficient is mispredicted by cDFT for strong MOF-CO_2 interactions and small maximal pore sizes. cDFT underestimates the Henry coefficient for stronger MOF-CO_2 interactions, whereas the Coulombic interactions are included for CO_2. A possible explanation

is that the stronger the Coulombic interactions, the more likely the sampling algorithm for the Coulombic interactions (described in the Supplementary Materials S1B) is to miss the optimal orientation, leading to interactions that are less strong than they should be. Concerning the effect of the small pores, a similar explanation could be given here as for N_2 with smaller pores leading to most adsorbate molecules being closer to MOF surfaces where orientations of adsorbates matter more, leading in this case to underestimation of the Henry coefficients. Plots similar to Fig. 5 for the comparison at high pressure are discussed in the Supplementary Materials Sec. S5.

The non-inclusion of the orientations in cDFT seems to be one of the main causes of the discrepancies with GCMC. Taking into account the orientations of the molecules is possible for the calculation of the external potential $V_k^{\text{ext}}(\mathbf{r})$ with the method used for CO_2 and the calculations of Coulombic interactions. However, as shown in Fig. S2a does not lead to better predictions overall and comes at the cost of computational power. Including the orientations of the molecules during the minimization of the grand potential, where the external potential is fixed, is doable [57] but would require additional efforts to implement and is out-of-the scope of this paper. The discrepancy encountered when Coulombic interactions are strong is also problematic because in applications such as Direct Air Carbon Capture and Storage (DACCS), strong Coulombic interactions may be necessary to capture the diluted CO_2. For this case, we can see in Fig. 3, that even if strong Coulombic interactions are mispredicted, the trend is still captured by cDFT. However, if correcting these deviations is necessary, a finer sampling of orientations in the calculation of the external potential could be performed, or GCMC calculations could be run for these particular cases.

In a screening framework, a tradeoff between accuracy and computational cost is necessary and we think that the cDFT framework presented here is an interesting option with a small loss of accuracy and a significant decrease in computational cost. A first preselection could be performed with cDFT and a refined estimation of adsorption properties could be made subsequently with GCMC. However, to study specific adsorption mechanisms of an adsorbate in a porous material (e.g., [58]), GCMC simulations would remain the preferable option as cDFT lacks the resolution to study these mechanisms.

VI. CONCLUSIONS

In this paper, we validate that classical density functional theory (cDFT) is an accurate and efficient method for adsorbent property prediction in the screening of porous materials for a wide range of MOFs with CH_4, N_2 and CO_2 as the adsorbates. For CO_2, we introduce the calculations of Coulombic interactions to accurately describe the interactions between the MOF and the guest molecule in the cDFT framework. We show that with this cDFT framework, GCMC calculations of adsorption isotherms of over 500 MOFs can closely be reproduced when CH_4 is the adsorbate, and reasonably well when N_2 and CO_2 are the adsorbates. The enthalpy of adsorption is also well reproduced with the added benefit that cDFT calculations exhibit very low numerical uncertainty. In addition to reproducing GCMC results, cDFT runs between 42 and 15 000 times faster, predominantly depending on the volume of the unit cell of the porous material.

These properties pave the way for cDFT to be used in process optimization, where the loading is often computed using empirical extrapolation models to new conditions of pressures, temperatures, and compositions. cDFT could also be used to design porous materials since the computation time for obtaining adsorption properties is low.

VII. ACKNOWLEDGEMENTS

V. D.-D. acknowledges the Rütli Stiftung and the ETH Zurich Foundation. P. R. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 497566159. J. S., E. M. and A. B. acknowledge the PrISMa Project (No 299659) funded through the ACT programme (Accelerating CCS Technologies, Horizon2020 Project No 294766) and the USorb-DAC Project supported by a grant from The Grantham Foundation for the Protection of the Environment to RMI's climate tech accelerator program, Third Derivative. J. G. acknowledges funding by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2075-390740016. The authors would like to thank Berend Smit for helpful discussions and early access to their data. The Table of Content graphic was created with the help of the software OVITO [59].

VIII. SUPPLEMENTARY INFORMATION

The codes developed for this work are available on GitLab:
https://gitlab.ethz.ch/epse/molecular-design-public/paper-3d_dft

A. New feature description

1. *Non-orthorhombic cells*

The non-orthorhombic unit cell can be defined by the angles α, β, and γ between the skewed axes with unit vectors $\vec{e}_u$, $\vec{e}_v$, and $\vec{e}_w$. A point $\mathbf{s} = \begin{pmatrix} u & v & w \end{pmatrix}^\intercal$ in the skewed coordinate system can be transformed into Cartesian coordinates using the transform [60]:

$$\underbrace{\begin{pmatrix} x \\ y \\ z \end{pmatrix}}_{\mathbf{r}} = \underbrace{\begin{pmatrix} 1 & \cos\gamma & \cos\beta \\ 0 & \sin\gamma & \zeta \\ 0 & 0 & \sqrt{1-\cos^2\beta-\zeta^2} \end{pmatrix}}_{\mathbf{H}} \underbrace{\begin{pmatrix} u \\ v \\ w \end{pmatrix}}_{\mathbf{s}} \quad \text{with} \quad \zeta = \frac{\cos\alpha - \cos\beta\cos\gamma}{\sin\gamma}. \tag{5}$$

Here, u, v, and w are not scaled by the length of the unit cell in the respective direction but use the same metric as the Cartesian coordinates x, y, and z. Fast Fourier transform (FFT) is used to rapidly evaluate a convolution integral of the form:

$$f \otimes g = \int f(\mathbf{r}')g(\mathbf{r}-\mathbf{r}')\mathrm{d}\mathbf{r}' = \mathcal{F}^{-1}\left(\mathcal{F}(f)\mathcal{F}(g)\right), \tag{6}$$

with the Fourier transform

$$\mathcal{F}(f) = \int f(\mathbf{r})e^{-2\pi i\mathbf{k}\cdot\mathbf{r}}\mathrm{d}\mathbf{r}, \tag{7}$$

and its inverse

$$\mathcal{F}^{-1}(f) = \int f(\mathbf{k})e^{2\pi i\mathbf{k}\cdot\mathbf{r}}\mathrm{d}\mathbf{k}. \tag{8}$$

The FFT is crucial for an efficient implementation of cDFT, because it reduces the complexity of the convolution integrals from $\mathcal{O}(n^2)$ to $\mathcal{O}(n\log n)$. However, the frequency values $\mathbf{k}$ at which the transformation is evaluated can not be chosen arbitrarily. If an FFT is evaluated in the skewed coordinate system $\mathbf{s}$, the resulting $\mathbf{k}_s$ values are also skewed. To calculate a convolution with an analytic weight function, the axes in frequency space must be transformed into a Cartesian coordinate system. This transformation is done by aligning

the phases $\mathbf{k}_r \cdot \mathbf{r} = \mathbf{r}^\intercal \mathbf{k}_r$ and $\mathbf{k}_s \cdot \mathbf{s} = \mathbf{s}^\intercal \mathbf{k}_s$ of the Fourier transform. Using the coordinate transform in real space $\mathbf{r} = \mathbf{H}\mathbf{s}$ or $\mathbf{r}^\intercal = \mathbf{s}^\intercal \mathbf{H}^\intercal$, the following relation is obtained:

$$\mathbf{s}^\intercal \mathbf{k}_s = \mathbf{r}^\intercal \mathbf{k}_r = \mathbf{s}^\intercal \mathbf{H}^\intercal \mathbf{k}_r \quad \Rightarrow \quad \mathbf{k}_s = \mathbf{H}^\intercal \mathbf{k}_r \quad \Rightarrow \quad \mathbf{k}_r = \left(\mathbf{H}^\intercal\right)^{-1} \mathbf{k}_s. \tag{9}$$

The values $\mathbf{k}_s = \begin{pmatrix} k_u & k_v & k_w \end{pmatrix}^\intercal$ are determined by the FFT. The values of $\mathbf{k}_r = \begin{pmatrix} k_x & k_y & k_z \end{pmatrix}^\intercal$ are then determined from Eqs. (6) and (9) as:

$$k_x = k_u \tag{10}$$

$$k_y = \frac{1}{\sin\gamma}\left(k_v - k_u \cos\gamma\right) \tag{11}$$

$$k_z = \frac{1}{\sqrt{1 - \cos^2\beta - \zeta^2}}\left(k_u\left(\zeta\frac{\cos\gamma}{\sin\gamma} - \cos\beta\right) - k_v\frac{\zeta}{\sin\gamma} + k_w\right). \tag{12}$$

Integrating in a skewed coordinate system also introduces a constant factor that stems from the Jacobian determinant:

$$\mathrm{d}\mathbf{r} = \det\left(\mathbf{H}\right)\mathrm{d}\mathbf{s} \quad \text{with} \quad \det\left(\mathbf{H}\right) = \sin\gamma\sqrt{1 - \cos^2\beta - \zeta^2}. \tag{13}$$

For the calculation of convolution integrals, the constant factor can be ignored because it cancels out during the inverse transform in which:

$$\mathrm{d}\mathbf{k}_r = \det\left(\left(\mathbf{H}^\intercal\right)^{-1}\right)\mathrm{d}\mathbf{k}_s = \frac{1}{\det\left(\mathbf{H}\right)}\mathrm{d}\mathbf{k}_s. \tag{14}$$

is used. When integrating distributed properties over the unit cell (evaluating a Fourier transform at $\mathbf{k} = \mathbf{0}$), however, the Jacobian determinant must be accounted for.

2. Coulombic interactions in external potential

Here, we describe the algorithm that is used to obtain the external potential at a position $\mathbf{r}$ when Coulombic interactions are considered. The detailed algorithm can be found in the code provided in the GitHub repository given above.

for *each MOF atom a* **do**

$dir_{atoms}[a] = pos_a - \mathbf{r}$;

compute energy adsorbate centered on r pointing to a for Lennard-Jones $E^{LJ}_{r,a}$ and Coulombic $E^{Coul}_{r,a}$ interactions;

end

$dir^{LJ}_1 = dir_{atoms}[argsort_a(E^{LJ}_{r,a})[0]]$;
$dir^{LJ}_2 = dir_{atoms}[argsort_a(E^{LJ}_{r,a})[1]]$;
$dir_{LJ} = cross(dir^{LJ}_1, dir^{LJ}_2)$;

$dir^{Coul}_{pos} = WeightedPCA(dir_{atoms}[where_a(E^{Coul}_{a,r} > 0)], boltzmann(E^{Coul}_{a,r}[where_a(E^{Coul}_{a,r} > 0)]))$;
$dir^{Coul}_{neg} = WeightedPCA(dir_{atoms}[where_a(E^{Coul}_{a,r} < 0)], boltzmann(E^{Coul}_{a,r}[where_a(E^{Coul}_{a,r} < 0)]))$;
$frac_{pos} = sum(|E^{Coul}_{a,i}[where_a(E^{Coul}_{a,r} > 0)]|)/sum(|E^{Coul}_{a,r}|)$;
$frac_{neg} = sum(|E^{Coul}_{a,r}[where_a(E^{Coul}_{a,r} < 0)]|)/sum(|E^{Coul}_{a,r}|)$;
$proj\_neg\_dir\_on\_pos\_plane = projection(dir^{Coul}_{neg}, dir^{Coul}_{pos})$;
$dir_{Coul} = frac_{neg} * dir^{Coul}_{neg} + frac_{pos} * proj\_neg\_dir\_on\_pos\_plane$;

$energies_{temp} = zeros(10)$;
for *j in range(10)* **do**

$dir_{temp} = j/10 * dir_{LJ} + (1 - j/10) * dir_{Coul}$;

compute total energy adsorbate E_j molecule in direction dir_{temp};

$energies_{temp} = E_j$

end

Algorithm 1: Algorithm to compute external potential when Coulombic interactions are considered.

For the case where Coulombic interactions are included for CO_2, the following method is used: at a position $\mathbf{r}$, we search for the orientation of the adsorbate molecule that would give the most favorable interactions. For this purpose, we noticed that the Coulombic interactions between an adsorbate molecule and an atom of the MOF are, in general, larger in magnitude than the Lennard-Jones interactions, except when one atom of the adsorbate

molecule is too close to the atom of the MOF. From this observation, we concluded that the Lennard-Jones interactions forbid some orientations while the Coulombic interactions push to a favorable orientation. Therefore, we compute two directions: an approximate optimal direction for Lennard-Jones interactions which are away from the 'most forbidden directions' and an approximate optimal direction for Coulombic interactions. We then compute the sum of Lennard-Jones and Coulombic interactions in 10 directions between the two computed optimal directions and retain the minimum energy for these 10 calculations.

For the optimal direction for Lennard-Jones interactions, we compute the Lennard-Jones interaction of the adsorbate molecule pointing toward each of the atoms of the MOFs. The two highest interactions define the two directions that are the less favorable. We approximate a non-favorable direction to a favorable plane, meaning that the adsorbate molecule will be attracted to the plane perpendicular to the non-favorable direction. By taking the cross product of the two directions that are the less favorable, we obtain the intersection of the two 'favorable' planes, which is set to be the approximate optimal direction for Lennard-Jones.

For the optimal direction for Coulombic interactions, we compute the Coulombic interaction of the adsorbate molecule pointing toward each of the atoms of the MOFs. The Boltzmann average of all the favorable orientations is calculated with a weighted Principal Component Analysis (wPCA). The Boltzmann average of all the unfavorable orientations is also calculated with a wPCA. The approximated optimal direction for Coulombic interactions is then the weighted average of the favorable direction and the projection of this favorable direction on the plane perpendicular to the unfavorable direction. The weight of the favorable and unfavorable directions is the fraction of the total Boltzmann energy of, respectively, the favorable and unfavorable interactions.

To save computational power, this algorithm is only computed when the position $\mathbf{r}$ is at a distance higher than $2\,\text{Å}^{-1}$ from any point of the MOF, since if it is closer than this distance, the interaction energy at this position would be too high for a molecule to be adsorbed at this position. For positions at a distance of less than $2\,\text{Å}^{-1}$, a value of $50\,\text{kJ}\,\text{mol}^{-1}$ is given to the external potential, which is high enough that it does not affect the loading.

To demonstrate the accuracy of our algorithm, we compare the values of the external potential of 3 MOFs obtained either with our algorithm or by using the Boltzmann average of the energy sampled every 20° for the two angles in spherical coordinates (called 'External

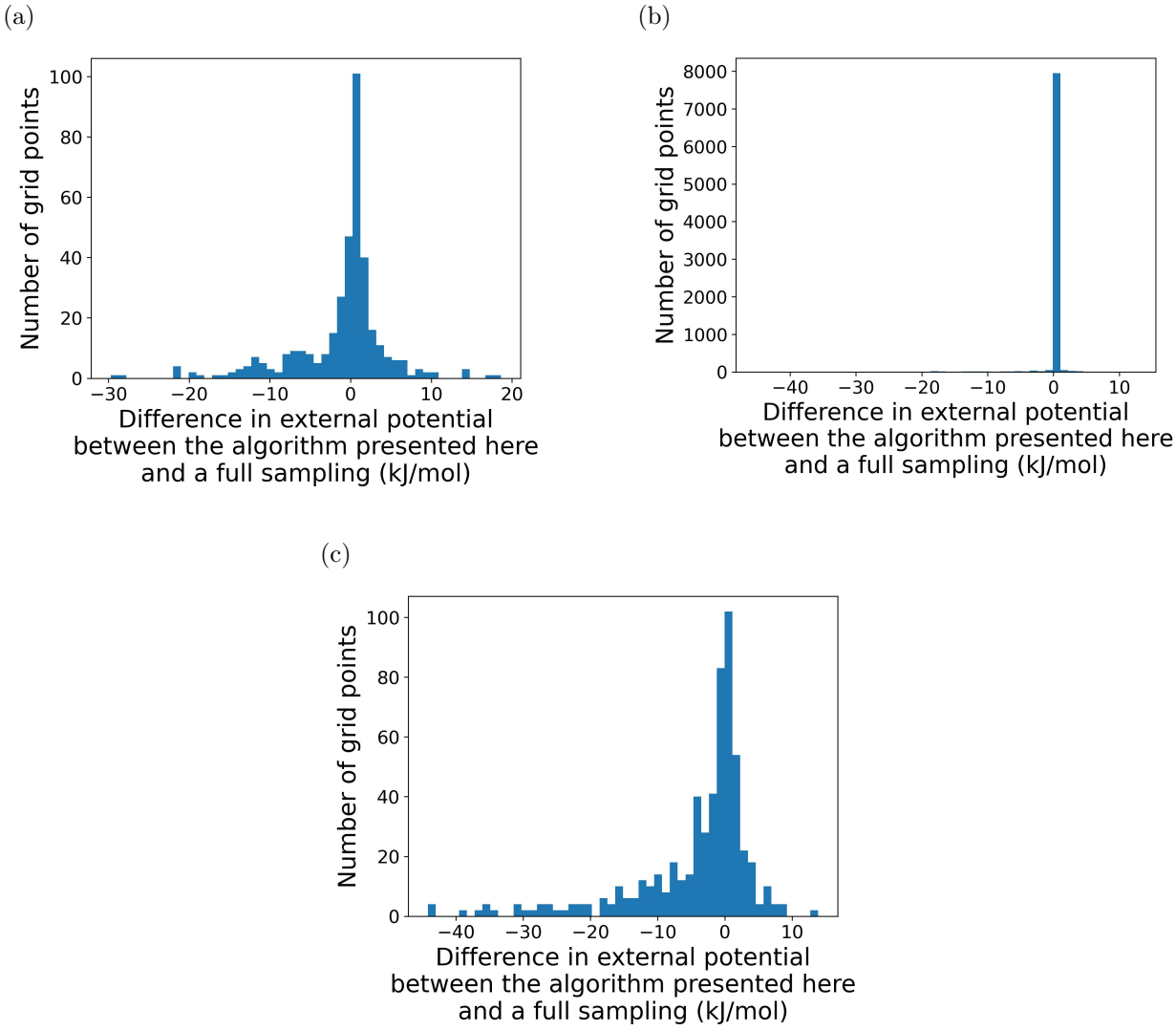


FIG. 6. Comparison of the external potential calculated at each grid point with the algorithm proposed here and the full sampling of molecule's orientations. The adsorbate molecule is always CO_2 and the MOFs are MAGVOG02 (a), GUKXIU (b), and BUSQIQ (c). For clarity, only grid points where at least one of the calculations had a value below 35 kJ/mol are shown. A potential above this value means that the grid point is close to an atom of the MOF and the external potential can have very high values.

potential with full sampling' in Fig. 6). We observe in Fig. 6 that these computations match reasonably well, while our algorithm is between 30 and 170 times faster for the MOFs considered here. Our algorithm tends to overestimate slightly the external potential, which is logical as it samples less directions for the gas molecule as the full sampling. Therefore,

it can miss directions with strong interactions with the MOF. It should be pointed out that the algorithm presented here has only been tested for linear adsorbates but could show some limitations for non-linear adsorbates. A full sampling approach might be necessary for non-linear adsorbates which would come at the cost of computational power.

3. Enthalpy of adsorption calculation with cDFT

The isosteric enthalpy of adsorption of component j is defined as:

$$\Delta h_j^{\mathrm{ads}} = h_j^{\mathrm{b}} - \left(\frac{\partial U}{\partial N_j}\right)_{T,N_{i\neq j}}, \tag{15}$$

with the partial molar enthalpy of component j in the bulk phase h_j^{b}, the composition of the bulk phase x_j, the internal energy of the fluid in the porous medium U, the number of particles in the porous medium N_j. The solid material is assumed to be rigid and therefore, the volume of the porous medium is not considered a variable. Using $h_j^{\mathrm{b}} = Ts_j^{\mathrm{b}} + \mu_i$ and $\mathrm{d}U = T\mathrm{d}S + \sum_i \mu_i \mathrm{d}N_i$ the chemical potential can be canceled out from the contributions of the bulk phase and the inhomogeneous system, leading to:

$$\Delta h_j^{\mathrm{ads}} = T\left(s_j^{\mathrm{b}} - \left(\frac{\partial S}{\partial N_j}\right)_{T,N_{i\neq j}}\right). \tag{16}$$

The partial derivative of the entropy S can then be replaced using a Maxwell relation based on the Helmholtz energy F as follows:

$$\left(\frac{\partial S}{\partial N_i}\right)_{T,N_j} = -\left(\frac{\partial^2 F}{\partial T \partial N_i}\right) = -\left(\frac{\partial \mu_i}{\partial T}\right)_{N_k}, \tag{17}$$

which leads to an expression for the enthalpy of adsorption without caloric properties of the adsorbed fluid:

$$\Delta h_i^{\mathrm{ads}} = T\left(s_i^{\mathrm{b}} + \left(\frac{\partial \mu_i}{\partial T}\right)_{N_k}\right). \tag{18}$$

To evaluate the enthalpy of adsorption with classical DFT, which uses temperatures and chemical potentials to define states, the derivative at constant number of particles must be replaced. This is done starting from the total differential of the number of particles:

$$\mathrm{d}N_i = \sum_j \left(\frac{\partial N_i}{\partial \mu_j}\right)_T \mathrm{d}\mu_j + \left(\frac{\partial N_i}{\partial T}\right)_{\mu_k} \mathrm{d}T. \tag{19}$$

Calculating the derivative with respect to T at constant N_i leads to:

$$0 = \sum_j \left(\frac{\partial N_i}{\partial \mu_j}\right)_T \left(\frac{\partial \mu_j}{\partial T}\right)_{N_k} + \left(\frac{\partial N_i}{\partial T}\right)_{\mu_k}, \tag{20}$$

from which the unknown derivative $\left(\frac{\partial \mu_i}{\partial T}\right)_{N_k}$ can be calculated. In practice, the expression has the disadvantage that $\left(\frac{\partial N_i}{\partial T}\right)_{\mu_k}$ depends on the (sometimes unknown) thermal de Broglie wavelength, which later cancels with s_i^{b}. This can be remedied by calculating the derivative of Eq. (19) with respect to T at constant (bulk) pressure and composition:

$$\left(\frac{\partial N_i}{\partial T}\right)_{p,x_k} = \sum_j \left(\frac{\partial N_i}{\partial \mu_j}\right)_T \left(\frac{\partial \mu_j}{\partial T}\right)_{p,x_k} + \left(\frac{\partial N_i}{\partial T}\right)_{\mu_k}. \tag{21}$$

From classical bulk thermodynamics, we know $\left(\frac{\partial \mu_j}{\partial T}\right)_{p,x_k} = -s_j^{\mathrm{b}}$ and therefore, Eq. (21) can be used in Eq. (20) to give:

$$0 = \sum_j \left(\frac{\partial N_i}{\partial \mu_j}\right)_T \left(s_j^{\mathrm{b}} + \left(\frac{\partial \mu_j}{\partial T}\right)_{N_k}\right) + \left(\frac{\partial N_i}{\partial T}\right)_{p,x_k}. \tag{22}$$

After multiplying with T, the expression shown in the main manuscript remains:

$$0 = \sum_j \left(\frac{\partial N_i}{\partial \mu_j}\right)_T \Delta h_j^{\mathrm{ads}} + T\left(\frac{\partial N_i}{\partial T}\right)_{p,x_k}. \tag{23}$$

Due to $\left(\frac{\partial N_i}{\partial \mu_j}\right)_T = -\left(\frac{\partial^2 \Omega}{\partial \mu_i \partial \mu_j}\right)_T$, Eq. (23) is a symmetric linear system of equations. The derivatives of the particle numbers are obtained by integrating over the respective derivatives of the density profiles, which are discussed in the following section.

4. Derivatives of density profiles

For converged density properties, equilibrium properties can be calculated as partial derivatives of thermodynamic potentials analogous to classical (bulk) thermodynamics. The density profiles are calculated implicitly from the Euler-Lagrange equation, which can be simplified to:

$$\Omega_{\rho_i}(T, \{\mu_k\}, [\{\rho_k(\mathbf{r})\}]) = F_{\rho_i}(T, [\{\rho_k(\mathbf{r})\}]) - \mu_i + V_i^{\mathrm{ext}}(\mathbf{r}) = 0. \tag{24}$$

The derivatives of the density profiles can then be calculated from the total differential of Eq. (24), leading to:

$$\mathrm{d}\Omega_{\rho_i}(\mathbf{r}) = \left(\frac{\partial \Omega_{\rho_i}(\mathbf{r})}{\partial T}\right)_{\mu_k,\rho_k} \mathrm{d}T + \sum_j \left(\frac{\partial \Omega_{\rho_i}(\mathbf{r})}{\partial \mu_j}\right)_{T,\mu_k,\rho_k} \mathrm{d}\mu_j + \int \sum_j \left(\frac{\delta \Omega_{\rho_i}(\mathbf{r})}{\delta \rho_j(\mathbf{r}')}\right)_{T,\mu_k,\rho_k} \delta\rho_j(\mathbf{r}')\mathrm{d}\mathbf{r}' = 0.$$

Using Eq. (24) and the shortened notation for derivatives of functionals in their natural variables, e.g., $F_T = \left(\frac{\partial F}{\partial T}\right)_{\rho_k}$, the expression can be simplified to:

$$F_{T\rho_i}(\mathbf{r})\mathrm{d}T - \mathrm{d}\mu_i + \int \sum_j F_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}')\delta\rho_j(\mathbf{r}')\mathrm{d}\mathbf{r}' = 0. \tag{25}$$

Similar to the Gibbs-Duhem relation for bulk phases, Eq. (25) shows how temperature, chemical potentials, and density profiles in an inhomogeneous system cannot be varied independently. The derivatives of the density profiles with respect to the intensive variables can be directly identified as:

$$\int \sum_j F_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial T}\right)_{\mu_k} \mathrm{d}\mathbf{r}' = -F_{T\rho_i}(\mathbf{r}), \tag{26}$$

and

$$\int \sum_j F_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial \mu_k}\right)_{T} \mathrm{d}\mathbf{r}' = \delta_{ik}. \tag{27}$$

Both of these expressions are implicit (linear) equations for the derivatives. In practice, it is useful to explicitly cancel out the (often unknown) thermal de Broglie wavelength Λ_i from the expression, where it has no influence. This is done by splitting the intrinsic Helmholtz energy into an ideal gas and a residual part:

$$F = k_\mathrm{B}T \int \sum_i m_i\rho_i(\mathbf{r}) \left(\ln\left(\rho_i(\mathbf{r})\Lambda_i^3\right) - 1\right) \mathrm{d}\mathbf{r} + \hat{\mathcal{F}}^\mathrm{res}. \tag{28}$$

Then, $F_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') = m_i\frac{k_\mathrm{B}T}{\rho_i(\mathbf{r})}\delta_{ij}\delta(\mathbf{r}-\mathbf{r}') + \hat{\mathcal{F}}^\mathrm{res}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}')$ and Eq. (27) can be rewritten as:

$$m_i\frac{k_\mathrm{B}T}{\rho_i(\mathbf{r})}\left(\frac{\partial \rho_i(\mathbf{r})}{\partial \mu_k}\right)_T + \int \sum_j \hat{\mathcal{F}}^\mathrm{res}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial \mu_k}\right)_T \mathrm{d}\mathbf{r}' = \delta_{ik}. \tag{29}$$

In practice, the division by the density should be avoided for numerical reasons and the energetic properties are reduced with the factor $\beta = \frac{1}{k_\mathrm{B}T}$. The final expression is:

$$m_i\left(\frac{\partial \rho_i(\mathbf{r})}{\partial \beta\mu_k}\right)_T + \rho_i(\mathbf{r}) \int \sum_j \beta\hat{\mathcal{F}}^\mathrm{res}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial \beta\mu_k}\right)_T \mathrm{d}\mathbf{r}' = \rho_i(\mathbf{r})\delta_{ik}. \tag{30}$$

For the temperature derivative, it is more convenient to express Eq. (25) in terms of the pressure of a bulk phase that is in equilibrium with the inhomogeneous system. In the

following, only paths along constant bulk composition are considered. With this constraint, the total differential of the chemical potential simplifies to:

$$\mathrm{d}\mu_i = -s_i \mathrm{d}T + v_i \mathrm{d}p, \tag{31}$$

which can be used in Eq. (25) to give:

$$\left(F_{T\rho_i}(\mathbf{r}) + s_i\right) \mathrm{d}T + \int \sum_j F_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}')\delta\rho_j(\mathbf{r}')\mathrm{d}\mathbf{r}' = v_i \mathrm{d}p. \tag{32}$$

It is useful at this point to rewrite the partial molar entropy as:

$$s_i = v_i \left(\frac{\partial p}{\partial T}\right)_{V,N_k} - F^{\mathrm{b}}_{T\rho_i}. \tag{33}$$

Then, the intrinsic Helmholtz energy can be split into an ideal gas and a residual part again, and the de Broglie wavelength cancels:

$$\left(m_i k_\mathrm{B} \ln\left(\frac{\rho_i(\mathbf{r})}{\rho_i^\mathrm{b}}\right) + F^{\mathrm{res}}_{T\rho_i}(\mathbf{r}) - F^{\mathrm{b,res}}_{T\rho_i} + v_i \left(\frac{\partial p}{\partial T}\right)_{V,N_k}\right) \mathrm{d}T + m_i \frac{k_\mathrm{B}T}{\rho_i(\mathbf{r})}\delta\rho_i(\mathbf{r}) + \int \sum_j \hat{\mathcal{F}}^{\mathrm{res}}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}')\delta\rho_j(\mathbf{r}')\mathrm{d}\mathbf{r}' = v_i \tag{34}$$

Finally, the expressions for the derivatives with respect to pressure:

$$m_i \left(\frac{\partial \rho_i(\mathbf{r})}{\partial \beta p}\right)_{T,x_k} + \rho_i(\mathbf{r}) \int \sum_j \beta\hat{\mathcal{F}}^{\mathrm{res}}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial \beta p}\right)_{T,x_k} \mathrm{d}\mathbf{r}' = v_i\rho_i(\mathbf{r}), \tag{35}$$

and temperature

$$\begin{aligned} m_i \left(\frac{\partial \rho_i(\mathbf{r})}{\partial T}\right)_{p,x_k} &+ \rho_i(\mathbf{r}) \int \sum_j \beta\hat{\mathcal{F}}^{\mathrm{res}}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') \left(\frac{\partial \rho_j(\mathbf{r}')}{\partial T}\right)_{p,x_k} \mathrm{d}\mathbf{r}' \\ &= -\frac{\rho_i(\mathbf{r})}{k_\mathrm{B}T}\left(m_i k_\mathrm{B} \ln\left(\frac{\rho_i(\mathbf{r})}{\rho_i^\mathrm{b}}\right) + F^{\mathrm{res}}_{T\rho_i}(\mathbf{r}) - F^{\mathrm{b,res}}_{T\rho_i} + v_i \left(\frac{\partial p}{\partial T}\right)_{V,N_k}\right) \end{aligned} \tag{36}$$

follows. All derivatives X_i shown here can be calculated from the same linear equation:

$$m_i X_i + \rho_i(\mathbf{r}) \int \sum_j \beta\hat{\mathcal{F}}^{\mathrm{res}}_{\rho_i\rho_j}(\mathbf{r},\mathbf{r}') X_i \mathrm{d}\mathbf{r}' = Y_i \tag{37}$$

by just replacing the right-hand side Y_i.

B. Definition of cDFT parameters

1. *Other parameters to define for cDFT calculations*

Following the setup used for the GCMC simulations, a cutoff of 12 Å is set for the Lennard-Jones interactions, Lorentz-Berthelot combining rules are used for solid-fluid interactions,

TABLE II. Derivatives of the density profile ρ with different variables.

derivative	right hand side
$\left(\frac{\partial \rho_i(\mathbf{r})}{\partial \beta \mu_k}\right)_T$	$\rho_i(\mathbf{r})\delta_{ik}$
$\left(\frac{\partial \rho_i(\mathbf{r})}{\partial \beta p}\right)_{T,x_k}$	$\rho_i(\mathbf{r})v_i$
$\left(\frac{\partial \rho_i(\mathbf{r})}{\partial T}\right)_{p,x_k}$	$-\frac{\rho_i(\mathbf{r})}{k_\mathrm{B}T}\left(m_i k_\mathrm{B}\ln\left(\frac{\rho_i(\mathbf{r})}{\rho_i^\mathrm{b}}\right)+F^\mathrm{res}_{T\rho_i}(\mathbf{r})-F^{\mathrm{b,res}}_{T\rho_i}+v_i\left(\frac{\partial p}{\partial T}\right)_{V,N_k}\right)$

TABLE III. Table with the solver setups used for cDFT. For parameters not explicitly given, the defaults in the software FeOs are used [37].

Order	Sequence of algorithms	Parameters
1	Picard Iteration	Damping coeff = 0.01, Max iteration = 50
	Picard Iteration	Damping coeff = 0.1, Tolerance = 1e-5
	Anderson Mixing	Mmax = 10, Tolerance = 1e-8
2	Newton	Tolerance = 1e-8
3	Picard line search	Log=True, Max iter = 2000
4	Picard Iteration	Damping coeff = 0.005, Max iteration = 2000
	Picard Iteration	Damping coeff = 0.01, Tolerance = 1e-5
	Anderson Mixing	Mmax = 10, Tolerance = 1e-8

and tail corrections are included. Ewald summation [61] is used for Coulombic interactions in GCMC, whereas a cutoff of 30 Å is used for cDFT. Four solver setups are considered to minimize the grand potential. These solver setups are used in a hierarchical order, if the previous solver setup failed to converge. These four solver setups are described in Table VIII B 1 and are used in the order they appear in the table. It is possible that the 4 solver setups fail. In this event, these cases are reported and these failures are discussed.

2. *Coulombic and PC-SAFT parameters definitions*

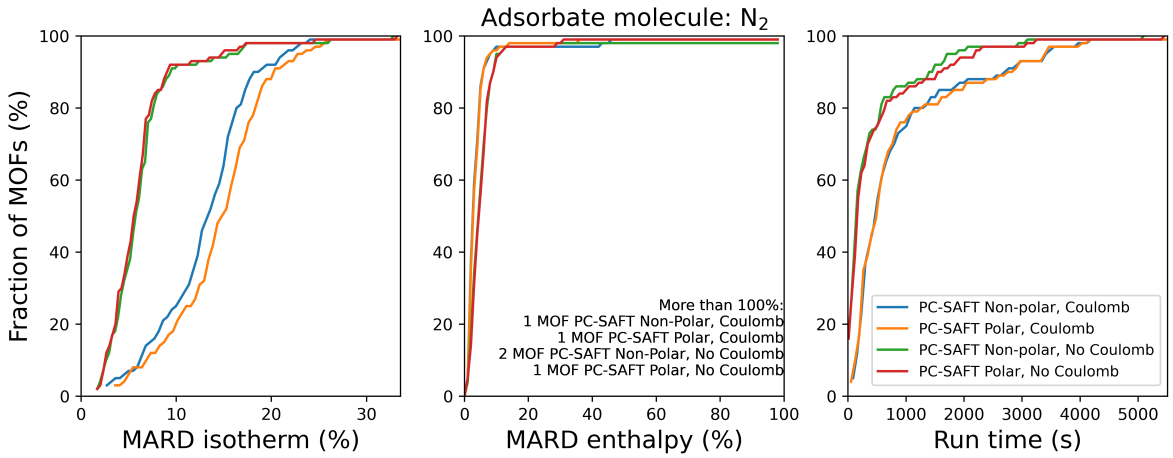
As discussed in the manuscript, the adsorbate molecules can be described with different parameters, the two most important ones being the inclusion or not of Coulombic interactions and the PC-SAFT parameters used. We consider here two sets of PC-SAFT parameters:

the initial set of PC-SAFT parameters developed in 2001 by Gross and Sadowski [40], which do not include the multipolar term of Eq. 1 in the main manuscript (called here 'PC-SAFT Non-polar'), and the polar set of PC-SAFT parameters developed in 2005 by Gross [48], which include the multipolar term of Eq. 1 in the main manuscript (called here 'PC-SAFT Polar').

As CH_4 is not a polar molecule, no Coulombic interactions will be considered. In addition, the non-polarity also means that no polar set of PC-SAFT parameters were developed for CH_4. Therefore, CH_4 is described with the 'PC-SAFT Non-polar' set of parameters throughout the manuscript [40]. As N_2 and CO_2 can both have some polarity effects, through a quadrupole moment, the inclusion or not of Coulombic interactions and the choice of PC-SAFT parameters will be studied by comparing the MARD (as defined in the manuscript by Eq. 2) for adsorption isotherms and enthalpy of adsorption, the runtime to get these values, and the number of times the solvers fail. This study is performed on 100 MOFs randomly selected in the database, with the 100 MOFs being the same for N_2 and CO_2. The results are shown in Fig. 7.

For N_2, the set of PC-SAFT parameters does not make a difference for any of the metrics, probably since the quadrupolarity of N_2 is negligible, so the non-polar parameters [40] are chosen for simplicity. Adding the Coulombic interactions increases the MARD for the adsorption isotherms and slightly decreases the MARD for the enthalpy of adsorption. The increase in MARD for the adsorption isotherms may be surprising as adding Coulombic interactions should be a more precise description. To understand better what is happening, we show the Mean Relative Deviation (MRD) of the cDFT calculations with the GCMC reference in Fig. 8. This MRD corresponds to the MARD of Eq. 4 without the absolute value at the numerator and can give insights if cDFT is overestimating or underestimating GCMC at the cost of a possible error cancellation. In Fig. 8, we observe that even without Coulombic interactions, cDFT predictions slightly overestimate the loading compared to GCMC calculations. Adding Coulombic interactions, increases the overprediction and make the MARD and MRD worse. The fact that cDFT slightly overpredicts GCMC without Coulombic interactions, may come from the PC-SAFT parameters already containing some Coulombic interactions implicitly by increasing the dispersion term. The reduction of the MARD for the enthalpy of adsorption can be explained by the fact that the external potential calculations match more closely the energies in GCMC as the TraPPE potential is used

(a)



(b)

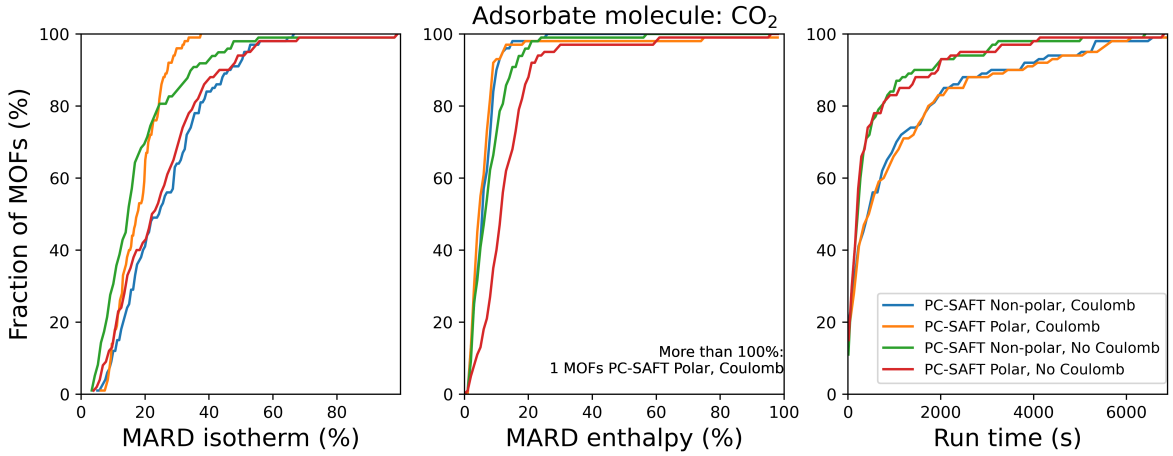


FIG. 7. Cumulative plots for the MARD for the adsorption isotherms (left) and the enthalpy of adsorption (middle) and runtime (right) for 4 parameter sets for N_2 (a) and CO_2 (b). 100 MOFs are used here. For the MARD of the enthalpy of adsorption, some deviations are much higher than 100 % and are not shown here for clarity of the visualization. This high deviation will be discussed in Section VIII B 3. For N_2, all the pressures converged when no Coulombic interactions were used, and only 0.01 % of them did not converge when no Coulombic interactions were used. For CO_2, all calculations converged and 'PC-SAFT Polar, Coulomb' whereas only 0.01 %, 0.01 %, and 0.02 % of all calculations did not converge for, respectively, the 'PC-SAFT Non-Polar, Coulomb', the 'PC-SAFT Polar, no Coulomb' and 'PC-SAFT Non-polar, No Coulomb' cases.

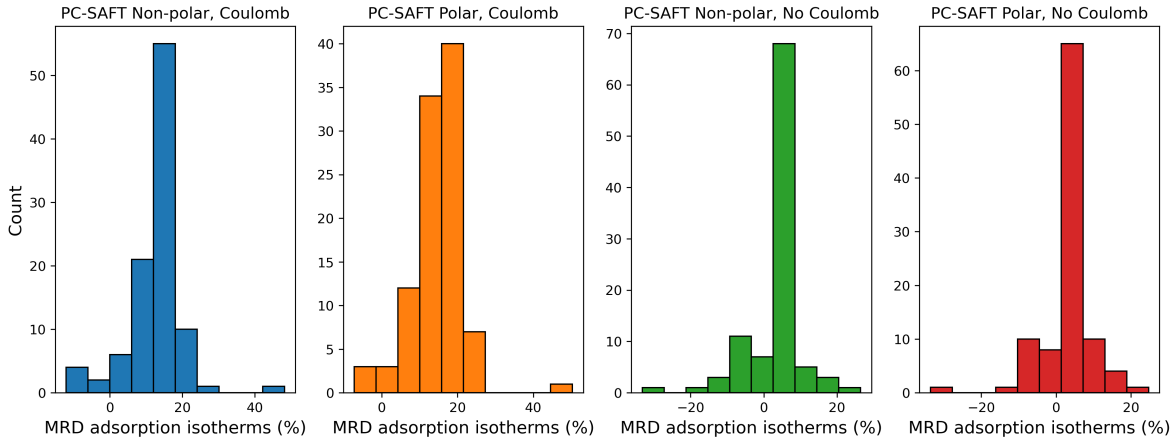


FIG. 8. Histogram of the MRD of cDFT calculations using GCMC as the reference for the adsorption isotherms of N_2 for the four parameters sets. Without the Coulombic interactions, the cDFT calculations overpredicts slightly the GCMC calculations and adding Coulombic interactions increases this overprediction.

to calculate the external potential. As less computation time is needed for the calculation of the external potential, the cases without Coulombic interactions are faster. We choose to use no Coulombic interactions for N_2 as it runs faster, converges more often, and gives a reduced MARD for adsorption isotherms, and only slightly greater MARD for enthalpy of adsorption.

We choose to describe CO_2 with Coulombic interactions and the 'PC-SAFT polar' set of parameters as this case has reduced MARD for adsorption isotherms and enthalpy of adsorption, and all calculations converged. These observations occur because CO_2 can have strong Coulombic interactions that change the adsorption properties. The run time is higher but not unreasonable and this additional computational cost increases the accuracy. The case with Coulombic interactions and 'PC-SAFT polar' set of parameters for CO_2 displays 1 MOFs with MARD for the enthalpy of adsorption higher than 100 % and these cases will be discussed in Sec. VIII B 3.

3. Grid density study

The main parameter of cDFT that defines the trade-off between accuracy and computational cost is the density of grid points: more grid points result in higher accuracy but also

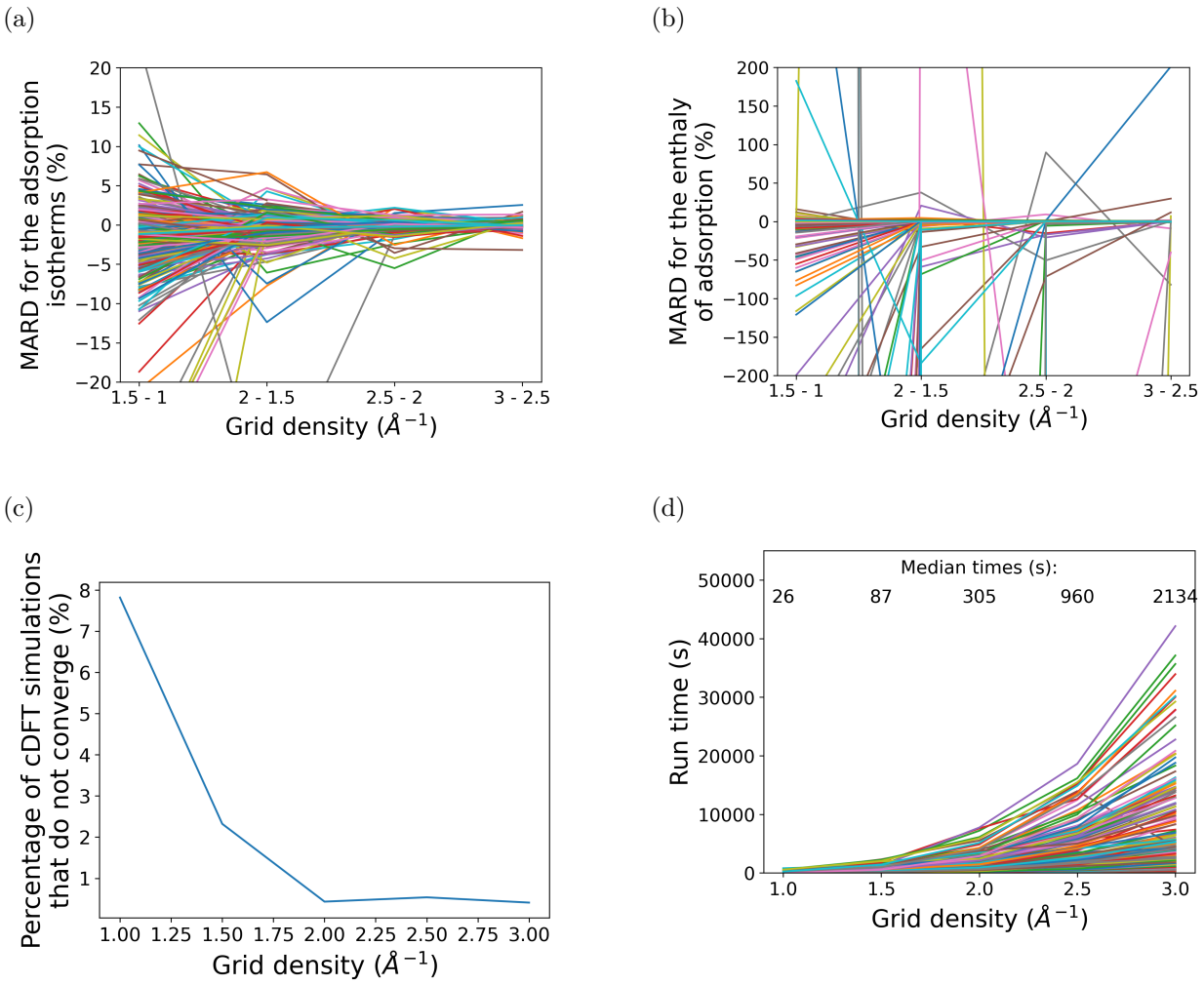


FIG. 9. Study of the effect of a change in grid density on (a) the adsorption isotherms, (b) enthalpy of adsorption, (c) percentage of converged cDFT calculations, and (d) run time of the cDFT calculations. For (a) and (b), the MARD of cDFT with GCMC calculations of the curves (adsorption isotherms or enthalpy of adsorption) between two subsequent values of grid density is shown. A MARD of 0 would mean that no evolution of values is observed when increasing the grid density.

in higher computational cost. To find the best balance, we study the effect of various grid densities (between 1 and 3 Å^{-1}) on the run time of an adsorption isotherm, the accuracy of adsorption isotherm and enthalpy of adsorption predictions, and the percentage of converged optimizations. This study is performed on the same set of 100 MOFs as in Section VIII B 2, with the 3 adsorbate molecules, at 298 K, with the parameters mentioned in Sec. VIII B 2. The objective of this study is to determine the minimum number of grid points required

to effectively prevent any subsequent changes in the outcomes as the grid point density increases. The evolution of the MARD between cDFT and GCMC of two subsequent values of grid points will be computed with Eq. 2 with the reference being the higher grid point density. The results are presented in Fig. 9.

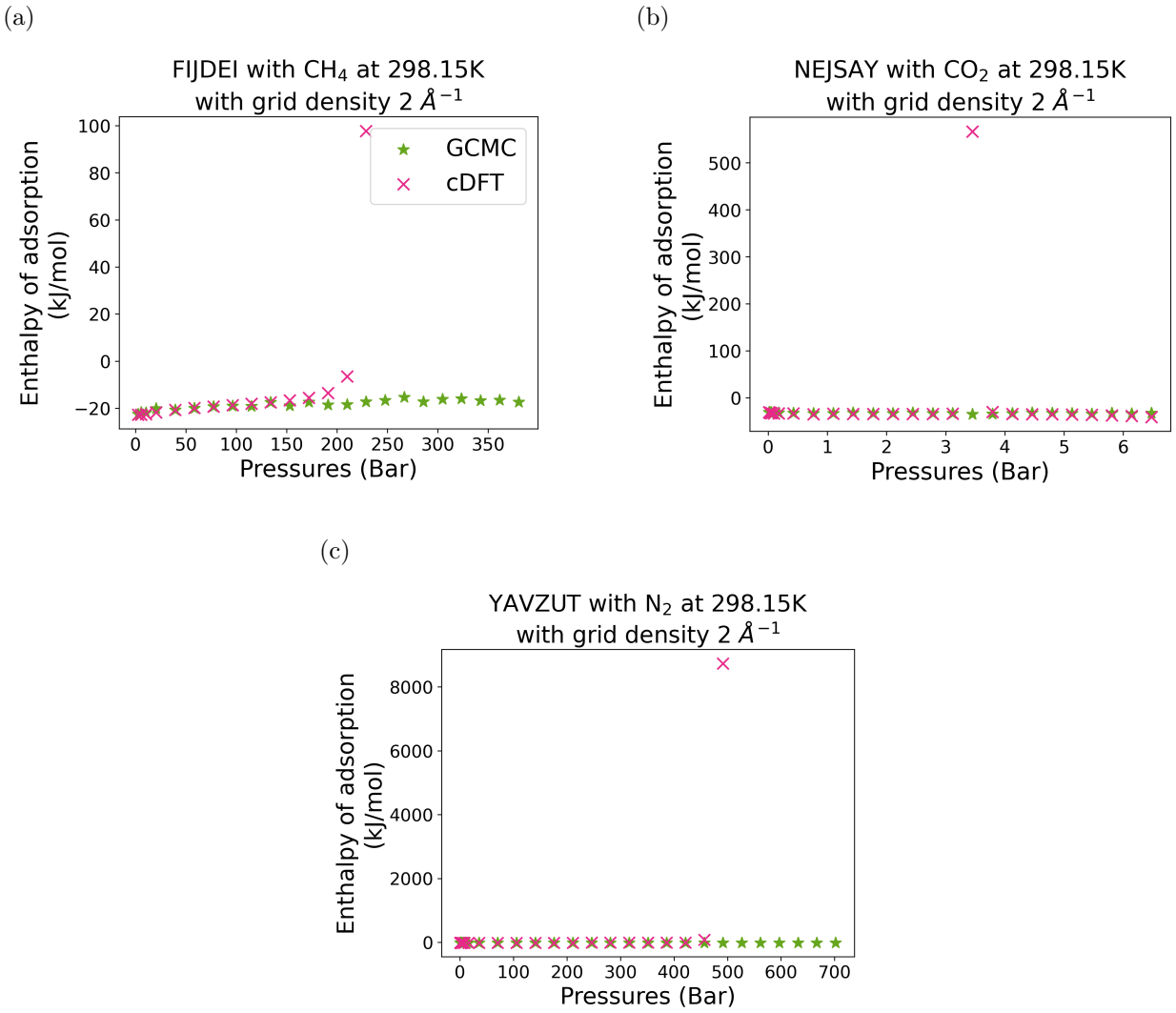


FIG. 10. Examples of enthalpy of adsorption curves where cDFT suddenly displays a substantial error. As these changes are sudden and of significant magnitude, we assume that they are due to numerical errors during the cDFT optimization. Increasing the grid density reduces the probability of these sudden deviations to appear.

For the adsorption isotherms in Fig. 9a, the MARD changes by less than 5.5 % from 2 $Å^{-1}$ to 2.5 $Å^{-1}$ and by less than 3.2 % from 2.5 $Å^{-1}$ to 3 $Å^{-1}$. The maximum evolution between 2 $Å^{-1}$ and 3 $Å^{-1}$ is 6.2 %. This observation makes 2 $Å^{-1}$ a reasonable density to offer good

accuracy for adsorption isotherms in most cases. For the case of the enthalpy of adsorption in Fig. 9b, there are several systems that show substantial evolutions between $2\,\text{Å}^{-1}$ and $2.5\,\text{Å}^{-1}$. These cases actually have unreasonable values for the enthalpy of adsorption at specific pressures (see Fig. 10). As these large deviations are sudden and abrupt, we assume that they are due to numerical errors during the optimization phase of cDFT. However, these significant deviations can easily be detected and corrected by increasing the density of grid points in most cases.

Fig. 9c shows the percentage of cDFT simulations that do not converge. At a grid density of $2\,\text{Å}^{-1}$, only $0.4\,\%$ of simulations do not converge. The simulations that do not converge are not equally shared between systems; out of the 300 systems, 297 of them have convergence for all the pressures of the adsorption isotherms. Further studies of optimization computations could help increase the number of convergence cases.

In Fig. 9d, the run times of whole adsorption isotherms at different grid densities are shown. Doubling of the grid density approximately multiply by 10 the computational time. In addition, the run time also depends on the system: the run time is proportional to the MOF unit cell volume as for larger unit cells, more grid points are needed, which increases the computational time.

Overall, a grid density of $2\,\text{Å}^{-1}$ was selected because, at this value, the deviation with GCMC does not change by more than $5.5\,\%$ for the adsorption isotherms, requires a relatively short run time (median time of approximately $5\,\text{min}$), and experiences a low percentage of convergence failure. However, as this grid density may still display substantial numerical errors in the calculations of the enthalpies of adsorption, we set a criterion to detect these substantial numerical errors and re-run the whole isotherm with a grid density $0.5\,\text{Å}^{-1}$ higher until a grid density of $4\,\text{Å}^{-1}$. At this grid density, the majority of the enthalpies of adsorption display extpected behavior, and the cases where it is not the case will be discussed in the manuscript (Sec. 4C). A criterion of a change of less than $10\,\text{kJ}\,\text{mol}^{-1}$ of the enthalpy of adsorption between two consecutive pressures was implemented.

C. Additional temperatures

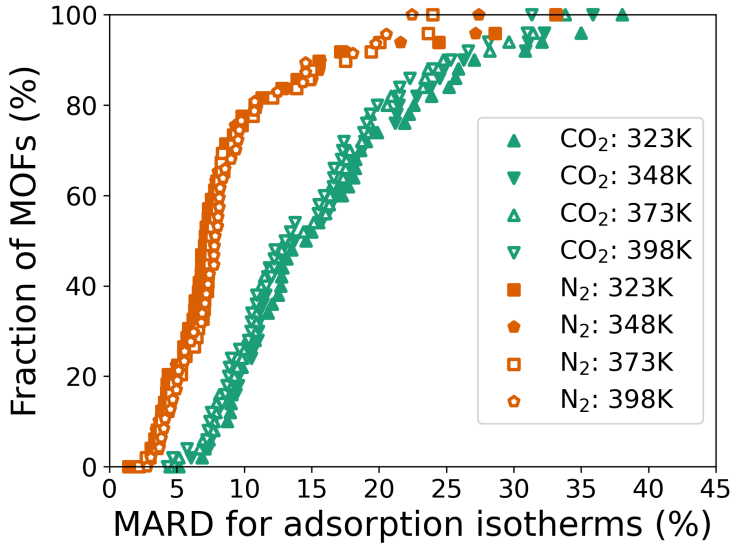


FIG. 11. MARD for the adsorption isotherms of CO_2 and N_2 at four temperatures. No significant effect of temperature can be observed on the deviation of the cDFT calculations of the adsorption isotherms compared with the ones computed with GCMC.

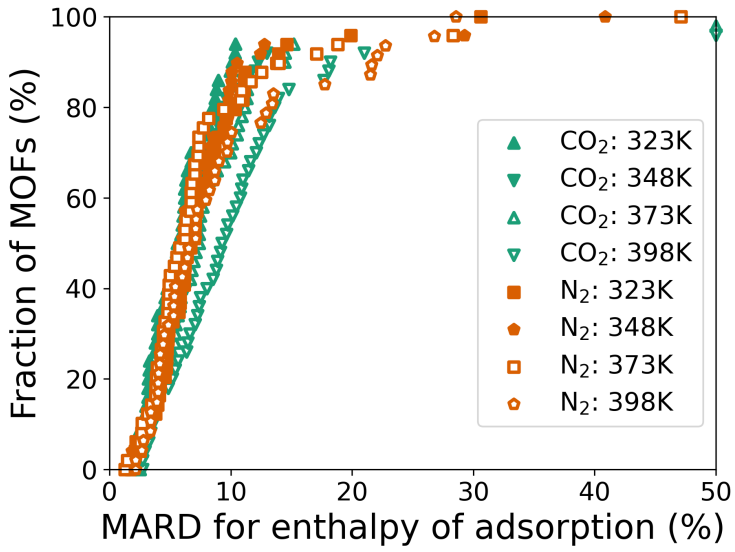


FIG. 12. MARD for the enthalpy of adsorption of CO_2 and N_2 at four temperatures. No significant effect of temperature can be observed onthe deviation of the cDFT calculations of the enthalpies of adsorption compared with the ones computed with GCMC. For CO_2 and N_2 respectively, 6/200 and 1/195 cases have MARD higher than 50%.

D. Henry coefficients and high pressure regime

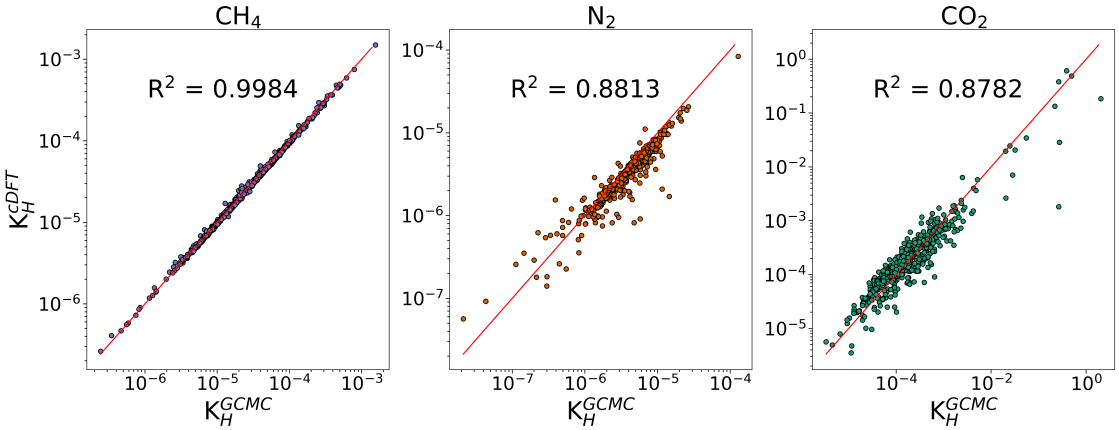


FIG. 13. Log-log parity plot for the Henry coefficients predicted by cDFT K_H^{cDFT} versus by GCMC K_H^{GCMC} for the three molecules (CH_4, N_2, CO_2). Each point corresponds to one MOF. The cDFT predictions of the Henry coefficient perfectly matches the GCMC results for CH_4 and are in good agreement for N_2 and CO_2.

Henry coefficients computed by cDFT K_H^{cDFT} are compared with the ones obtained with

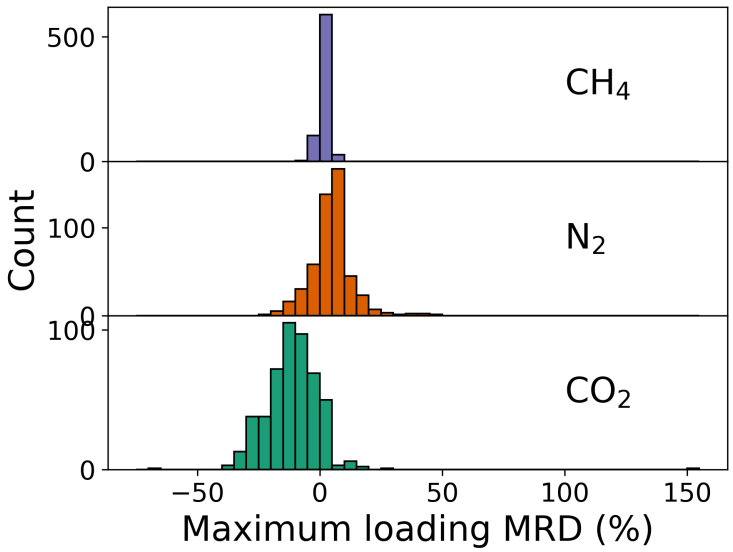


FIG. 14. Histogram of the maximum loading MRD predicted by cDFT versus by GCMC for the three molecules (CH_4, N_2, CO_2). The cDFT predictions of high pressure regime perfectly matches the GCMC results for CH_4 and are in good agreement for N_2 and CO_2.

GCMC K_H^{GCMC}. In this case, the deviation metric is the ratio of the Henry coefficients K_H^{cDFT}/K_H^{GCMC}. Fig. 13 compares of the Henry coefficients calculated with cDFT with the ones computed with GCMC. For CH_4, cDFT Henry coefficients almost perfectly match the ones from GCMC with 97% of the Henry coefficients deviating by less than 10%. For N_2 and CO_2, cDFT calculations differ a bit more, with respectively 5% and 16% of the Henry coefficients deviating by more than a factor of 2. These deviations should, however, be compared to the large range of values that Henry coefficients can display, as they span 4 orders of magnitude for N_2 and 6 orders of magnitude for CO_2.

The Henry regime is dominated by the MOF-adsorbate interactions, whereas the high-pressure regime is dominated by the pore geometry of the MOF. In order to check if the cDFT predictions follow the same behavior as GCMC at high pressure, we define a 'Maximum loading MRD', which is the same as the MARD for adsorption isotherms defined in Eq. 4 of the main manuscript, but only considering the points at the highest five pressures and without taking absolute values. We do not take absolute values here so we can know if the maximum loading is overestimated or underestimated. As this behavior is unlikely to change in the high-pressure regime, error cancellation does not occur. In Fig. 14, we can see that even at high pressure, the cDFT loading predictions match very closely the GCMC

calculations, with all the deviations within 10%. For N_2, the cDFT predictions are good too, with 91% of the maximum loading MRD with 15%. For CO_2, it can be observed that cDFT tends to underestimate the loading at high pressure by a median value of -11%, which is still reasonable.

E. High-pressure regime predictor of deviation between cDFT and GCMC

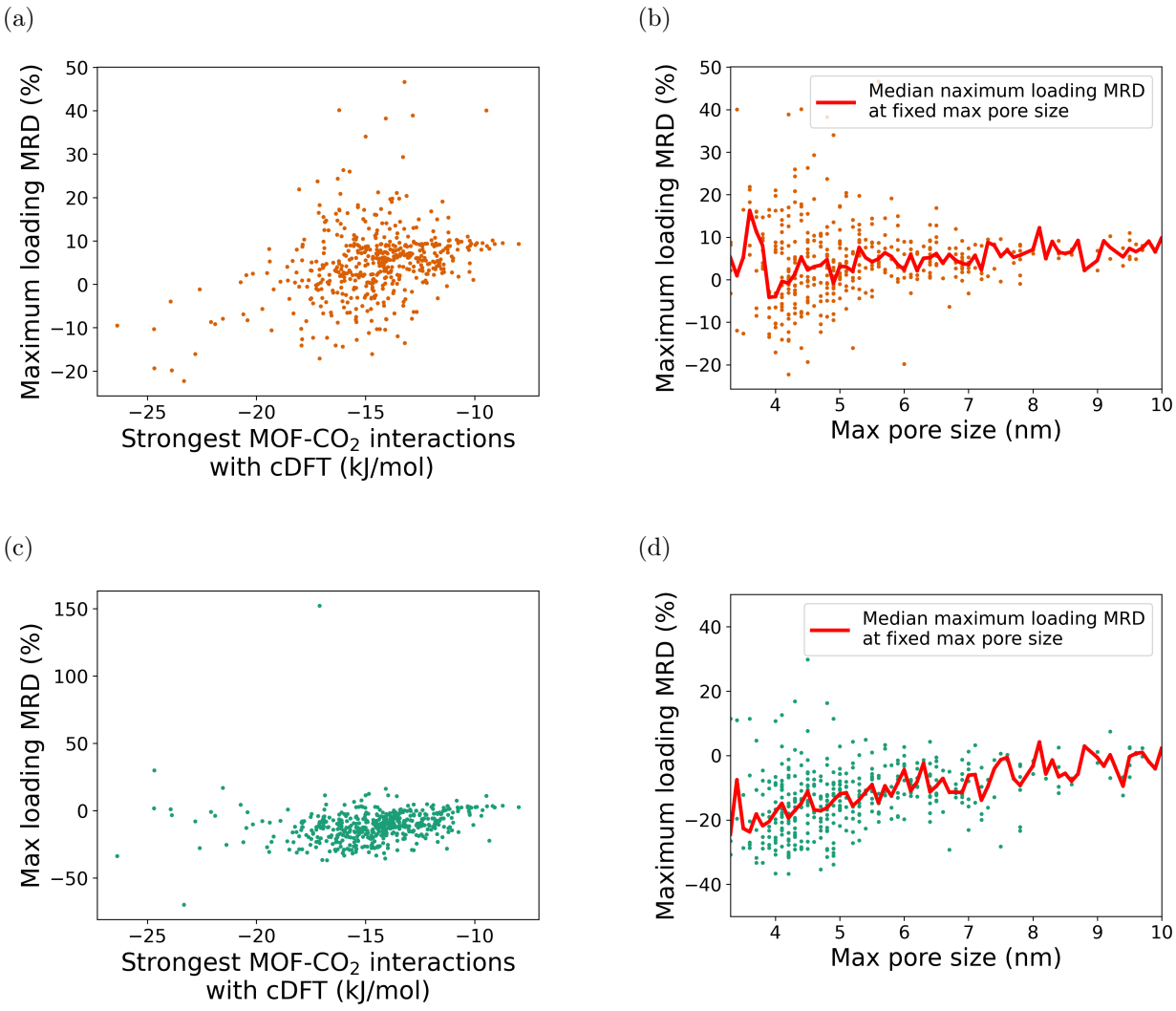


FIG. 15. Scatter plot of the strongest MOF-CO_2 interactions computed with cDFT ((a) and (c)) and the maximal pore size ((b) and (d)) with the maximum loading MRD between cDFT and GCMC for N_2 ((a) and (b)) and CO_2 ((c) and (d)). The y axes are in logarithm scale. Each point corresponds to one MOF.

In Fig. 15, similar plots to Fig. 4 of the main manuscript are shown, but here for the maximum loading MRD, showing the trends for the high-pressure regime. From Fig. 15a, it also seems that including Coulombic interactions for N_2 would be beneficial for the few materials where the maximum loading MRD is underestimated. However, the maximal pore size does not seem to play an important role here. For CO_2, no clear trend of the maximum loading MRD with the strongest MOF-CO_2 interactions calculated with cDFT can be observed in Fig. 15c. A possible explanation is that over the whole MOF, the Coulombic interactions are, on average, correctly calculated and at high pressures, the adsorbates are loaded over the whole MOF. In contrast, in the Henry regime, only a few grid points would exhibit a significant loading and the deviation of cDFT would depend on how accurately the Coulombic interactions were calculated at these few points. However, a decreasing trend of the maximum loading MRD with the maximum pore size is observed, similar to the one observed in Fig. 4c of the main manuscript.

IX. REFERENCES

[1] Daniel Ferreira, Marta Boaventura, Patrick Barcia, Roger D Whitley, and Adelio Mendes. Two-stage vacuum pressure swing adsorption using aglilsx zeolite for producing 99.5+% oxygen from air. *Industrial & Engineering Chemistry Research*, 55(3):722–736, 2016.

[2] Masoud Mofarahi, Jafar Towfighi, and Leila Fathi. Oxygen separation from air by four-bed pressure swing adsorption. *Industrial & Engineering Chemistry Research*, 48(11):5439–5444, 2009.

[3] Ehsan Javadi Shokroo, Danial Jafari Farsani, Hadiseh Khalilpour Meymandi, and Nadia Yadollahi. Comparative study of zeolite 5a and zeolite 13x in air separation by pressure swing adsorption. *Korean Journal of Chemical Engineering*, 33:1391–1401, 2016.

[4] Debasis Banerjee, Cory M Simon, Sameh K Elsaidi, Maciej Haranczyk, and Praveen K Thallapally. Xenon gas separation and storage using metal-organic frameworks. *Chem*, 4(3):466–494, 2018.

[5] Nick R Soelberg, Troy G Garn, Mitchell R Greenhalgh, Jack D Law, Robert Jubin, Denis M Strachan, Praveen K Thallapally, et al. Radioactive iodine and krypton control for nuclear fuel reprocessing facilities. *Science and Technology of Nuclear Installations*, 2013, 2013.

[6] Berend Smit, Jeffrey A Reimer, Curtis M Oldenburg, and Ian C Bourg. *Introduction to carbon capture and sequestration*, volume 1. World Scientific, 2014.

[7] Franz Lanzerath, Uwe Bau, Jan Seiler, and André Bardow. Optimal design of adsorption chillers based on a validated dynamic object-oriented model. *Science and Technology for the Built Environment*, 21(3):248–257, 2015.

[8] Akriti Sarswat, David S Sholl, and Ryan P Lively. Achieving order of magnitude increases in CO_2 reduction reaction efficiency by product separations and recycling. *Sustainable Energy & Fuels*, 6(20):4598–4604, 2022.

[9] Miguel Palomino, Avelino Corma, Fernando Rey, and Susana Valencia. New insights on CO_2-methane separation using LTA zeolites with different Si/Al ratios and a first comparison with MOFs. *Langmuir*, 26(3):1910–1917, 2010.

[10] Tea-Hoon Kim, Seo-Yul Kim, Tae-Ung Yoon, Min-Bum Kim, Wanje Park, Hyug Hee Han, Chang-in Kong, Chae-Young Park, Jeong-Hoon Kim, and Youn-Sang Bae. Improved methane/nitrogen separation properties of zirconium-based metal–organic framework by incorporating highly polarizable bromine atoms. *Chemical Engineering Journal*, 399:125717, 2020.

[11] Liangying Li, Lidong Guo, Zhiguo Zhang, Qiwei Yang, Yiwen Yang, Zongbi Bao, Qilong Ren, and Jing Li. A robust squarate-based metal–organic framework demonstrates record-high affinity and selectivity for xenon over krypton. *Journal of the American Chemical Society*, 141 (23):9358–9364, 2019.

[12] Charithea Charalambous, Elias Moubarak, Johannes Schilling, Eva Sanchez Fernandez, Jin-Yu Wang, Laura Herraiz, Fergus Mcilwaine, Kevin Maik Jablonka, Seyed Mohamad Moosavi, Joren Van Herck, et al. Shedding Light on the Stakeholders' Perspectives for Carbon Capture. *ChemRxiv, This content is a preprint and has not been peer-reviewed.*, 2023. doi: 10.26434/chemrxiv-2023-sn90q.

[13] Weizhen Sun, Li-Chiang Lin, Xuan Peng, and Berend Smit. Computational screening of porous metal-organic frameworks and zeolites for the removal of SO_2 and NO_x from flue gases. *AIChE Journal*, 60(6):2314–2323, 2014.

[14] Patrick Ryan, Omar K Farha, Linda J Broadbelt, and Randall Q Snurr. Computational screening of metal-organic frameworks for xenon/krypton separation. *AIChE Journal*, 57(7): 1759–1766, 2011.

[15] Ki Chul Kim, Peyman Z Moghadam, David Fairen-Jimenez, and Randall Q Snurr. Computational screening of metal catecholates for ammonia capture in metal–organic frameworks. *Industrial & Engineering Chemistry Research*, 54(13):3257–3267, 2015.

[16] Yaling Yan, Zenan Shi, Huilin Li, Lifeng Li, Xiao Yang, Shuhua Li, Hong Liang, and Zhiwei Qiao. Machine learning and in-silico screening of metal–organic frameworks for O_2/N_2 dynamic adsorption and separation. *Chemical Engineering Journal*, 427:131604, 2022.

[17] Rajamani Krishna and Jasper M van Baten. In silico screening of zeolite membranes for CO_2 capture. *Journal of Membrane Science*, 360(1-2):323–333, 2010.

[18] Rodrigo Neumann Barros Ferreira, Breanndan O'Conchuir, Tonia Elengikal, Binquan Luan, Ricardo Luis Ohta, Felipe Lopes de Oliveira, Ashish Mhadeshwar, Jayashree Kalyanaraman, Anantha Sundaram, Joseph Falkowski, et al. Cloud-based, high-throughput, end-to-end computational screening of solid sorbent materials for carbon capture. In *Greenhouse Gas Control Technologies*, 2022.

[19] Jake Burner, Ludwig Schwiedrzik, Mykhaylo Krykunov, Jun Luo, Peter G Boyd, and Tom K Woo. High-performing deep learning regression models for predicting low-pressure CO_2 adsorption properties of metal–organic frameworks. *The Journal of Physical Chemistry C*, 124 (51):27996–28005, 2020.

[20] Hyunsoo Park, Sauradeep Majumdar, Xiaoqi Zhang, Jihan Kim, and Berend Smit. Inverse design of metal–organic frameworks for direct air capture of CO_2 via deep reinforcement learning. *Digital Discovery*, 3(4):728–741, 2024.

[21] Amir H Farmahini, Shreenath Krishnamurthy, Daniel Friedrich, Stefano Brandani, and Lev Sarkisov. Performance-based screening of porous materials for carbon capture. *Chemical Reviews*, 121(17):10666–10741, 2021.

[22] Li-Chiang Lin, Adam H Berger, Richard L Martin, Jihan Kim, Joseph A Swisher, Kuldeep Jariwala, Chris H Rycroft, Abhoyjit S Bhown, Michael W Deem, Maciej Haranczyk, et al. In silico screening of carbon-capture materials. *Nature materials*, 11(7):633–641, 2012.

[23] Song Li, Yongchul G Chung, and Randall Q Snurr. High-throughput screening of metal–organic frameworks for co_2 capture in the presence of water. *Langmuir*, 32(40):10368–10376, 2016.

[24] Arvind Rajendran, Sai Gokul Subraveti, Kasturi Nagesh Pai, Vinay Prasad, and Zukui Li. How Can (or Why Should) Process Engineering Aid the Screening and Discovery of Solid

Sorbents for CO_2 Capture? *Accounts of Chemical Research*, pages 16730–16740, 2023.

[25] Conor Cleeton, Felipe Lopes de Oliveira, Rodrigo F Neumann, Amir H Farmahini, Binquan Luan, Mathias Steiner, and Lev Sarkisov. A process-level perspective of the impact of molecular force fields on the computational screening of mofs for carbon capture. *Energy & Environmental Science*, 16(9):3899–3918, 2023.

[26] Elias Moubarak, Seyed Mohamad Moosavi, Charithea Charalambous, Susana Garcia, and Berend Smit. A robust framework for generating adsorption isotherms to screen materials for carbon capture. *Industrial & Engineering Chemistry Research*, 62(26):10252–10265, 2023.

[27] Richard J Gowers, Amir H Farmahini, Daniel Friedrich, and Lev Sarkisov. Automated analysis and benchmarking of GCMC simulation programs in application to gas adsorption. *Molecular Simulation*, 44(4):309–321, 2018.

[28] Xuanjun Wu, Zijian Cao, Xiuyang Lu, and Weiquan Cai. Prediction of methane adsorption isotherms in metal–organic frameworks by neural network synergistic with classical density functional theory. *Chemical Engineering Journal*, 459:141612, 2023.

[29] Fangyuan Guo, Yu Liu, Jun Hu, Honglai Liu, and Ying Hu. Fast screening of porous materials for noble gas adsorption and separation: a classical density functional approach. *Physical Chemistry Chemical Physics*, 20(44):28193–28204, 2018.

[30] Fangyuan Guo, Yu Liu, Jun Hu, Honglai Liu, and Ying Hu. Screening of porous materials for toxic gas adsorption: classical density functional approach. *Industrial & Engineering Chemistry Research*, 59(32):14364–14373, 2020.

[31] Yu Liu, Shuangliang Zhao, Honglai Liu, and Ying Hu. High-throughput and comprehensive prediction of h 2 adsorption in metal-organic frameworks under various conditions. *AIChE Journal*, 61(9):2951–2957, 2015.

[32] Jia Fu, Yu Liu, Yun Tian, and Jianzhong Wu. Density functional methods for fast screening of metal–organic frameworks for hydrogen storage. *The Journal of Physical Chemistry C*, 119 (10):5374–5385, 2015.

[33] Yun Tian, Weiyang Fei, and Jianzhong Wu. Separation of carbon isotopes in methane with nanoporous materials. *Industrial & Engineering Chemistry Research*, 57(14):5151–5160, 2018.

[34] Jia Fu, Yun Tian, and Jianzhong Wu. Classical density functional theory for methane adsorption in metal-organic framework materials. *AIChE Journal*, 61(9):3012–3021, 2015.

[35] Lucas J dos Santos, Elvis do A Soares, Amaro G Barreto, and Frederico W Tavares. Com-

parison of 3d-cdft and gcmc simulations for fluid–structure analysis in amorphous carbon nanoporous materials. *Adsorption*, pages 1–12, 2024.

[36] Elvis do A Soares, Amaro G Barreto Jr, and Frederico W Tavares. Classical density functional theory reveals structural information of H_2 and CH_4 fluids adsorbed in MOF-5. *Fluid Phase Equilibria*, 574:113887, 2023.

[37] Philipp Rehner, Gernot Bauer, and Joachim Gross. FeOs: An Open-Source Framework for Equations of State and Classical Density Functional Theory. *Industrial & Engineering Chemistry Research*, 62(12):5347–5357, 2023.

[38] Christopher Kessler, Johannes Eller, Joachim Gross, and Niels Hansen. Adsorption of light gases in covalent organic frameworks: comparison of classical density functional theory and grand canonical Monte Carlo simulations. *Microporous and Mesoporous Materials*, 324:111263, 2021.

[39] Elmar Sauer and Joachim Gross. Classical density functional theory for liquid–fluid interfaces and confined systems: A functional for the perturbed-chain polar statistical associating fluid theory equation of state. *Industrial & Engineering Chemistry Research*, 56(14):4119–4135, 2017.

[40] Joachim Gross and Gabriele Sadowski. Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules. *Industrial & Engineering Chemistry Research*, 40 (4):1244–1260, 2001.

[41] Dandan Hong, Yu Liu, Honglai Liu, and Ying Hu. Development of dual-model classical density functional theory and its application to gas adsorption in porous materials. *AIChE Journal*, 67(4):e17120, 2021.

[42] Elias Moubarak, Susana Garcia, and Berend Smit. Methane adsorption isotherms for mof structures from the prisma database. *ChemRxiv*, 2024. doi:10.26434/chemrxiv-2024-vx710. URL http://dx.doi.org/10.26434/chemrxiv-2024-vx710.

[43] Anthony K Rappé, Carla J Casewit, KS Colwell, William A Goddard III, and W Mason Skiff. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, 114(25):10024–10035, 1992.

[44] Jeffrey J Potoff and J Ilja Siepmann. Vapor–liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AIChE journal*, 47(7):1676–1682, 2001.

[45] Thomas A Manz and David S Sholl. Chemically meaningful atomic charges that reproduce

the electrostatic potential in periodic and nonperiodic materials. *Journal of Chemical Theory and Computation*, 6(8):2455–2468, 2010.

[46] Christopher E Wilmer, Ki Chul Kim, and Randall Q Snurr. An extended charge equilibration method. *J. Phys. Chem. Lett.*, 3(17):2506–2511, 2012. doi:10.1021/jz301439a.

[47] Thomas F Willems, Chris H Rycroft, Michaeel Kazi, Juan C Meza, and Maciej Haranczyk. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous and Mesoporous Materials*, 149(1):134–141, 2012.

[48] Joachim Gross. An equation-of-state contribution for polar components: Quadrupolar molecules. *AIChE journal*, 51(9):2556–2568, 2005.

[49] Stephen L Mayo, Barry D Olafson, and William A Goddard. DREIDING: a generic force field for molecular simulations. *Journal of Physical chemistry*, 94(26):8897–8909, 1990.

[50] Felipe Lopes Oliveira, Conor Cleeton, Rodrigo Neumann Barros Ferreira, Binquan Luan, Amir H Farmahini, Lev Sarkisov, and Mathias Steiner. CRAFTED: An exploratory database of simulated adsorption isotherms of metal-organic frameworks. *Scientific Data*, 10(1):230, 2023.

[51] Jindal K Shah and Edward J Maginn. A general and efficient Monte Carlo method for sampling intramolecular degrees of freedom of branched and cyclic molecules. *The Journal of chemical physics*, 135(13), 2011.

[52] JA Purton, Jenny C Crabtree, and SC Parker. DL_MONTE: a general purpose program for parallel Monte Carlo simulation. *Molecular Simulation*, 39(14-15):1240–1252, 2013.

[53] Amit Gupta, Shaji Chempath, Martin J Sanborn, Louis A Clark, and Randall Q Snurr. Object-oriented programming paradigms for molecular modeling. *Molecular Simulation*, 29(1):29–46, 2003.

[54] David Dubbeldam, Sofía Calero, Donald E Ellis, and Randall Q Snurr. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Molecular Simulation*, 42(2):81–101, 2016.

[55] Marcus G Martin. MCCCS Towhee: a tool for Monte Carlo molecular simulation. *Molecular Simulation*, 39(14-15):1212–1222, 2013.

[56] Marielle Pinheiro, Richard L Martin, Chris H Rycroft, Andrew Jones, Enrique Iglesia, and Maciej Haranczyk. Characterization and comparison of pore landscapes in crystalline porous materials. *Journal of Molecular Graphics and Modelling*, 44:208–219, 2013.

[57] Philipp Rehner, Thijs van Westen, and Joachim Gross. Equation of state and helmholtz energy functional for fused heterosegmented hard chains. *Physical Review E*, 105(3):034110, 2022.

[58] Yann Magnin, Estelle Dirand, Guillaume Maurin, and Philip L Llewellyn. Abnormal CO_2 and H_2O Diffusion in CALF-20 (Zn) Metal–Organic Framework: Fundamental Understanding of CO_2 Capture. *ACS Applied Nano Materials*, 6(21):19963–19971, 2023.

[59] Alexander Stukowski. Visualization and analysis of atomistic simulation data with ovito–the open visualization tool. *Modelling and simulation in materials science and engineering*, 18(1): 015012, 2009.

[60] David Dubbeldam, Ariana Torres-Knoop, and Krista S. Walton. On the inner workings of Monte Carlo codes. *Mol. Simul.*, 39(14-15):1253–1292, December 2013. ISSN 0892-7022. doi: 10.1080/08927022.2013.819102.

[61] Paul P Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annalen der physik*, 369(3):253–287, 1921.

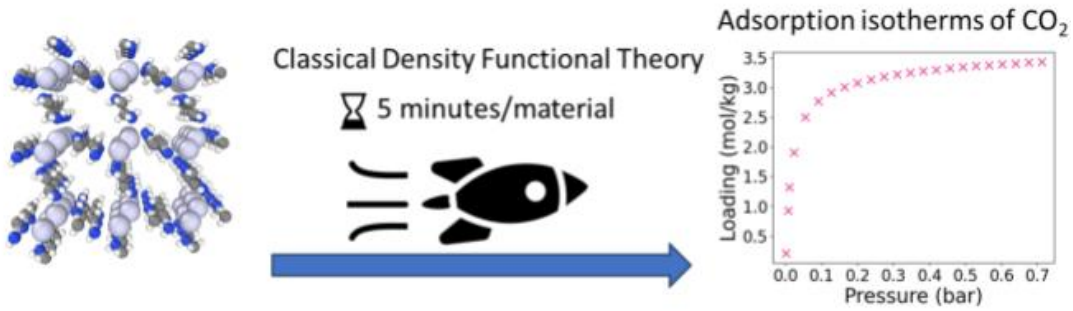


FIG. 16. Table Of Content Graphic