THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Chimie ParisTech

Développements méthodologiques pour la simulation de l'adsorption des gaz simples dans les matériaux nanoporeux

Methodology developments for the simulation of simple gas adsorption in nanoporous materials

Soutenue par
Lionel ZOUBRITZKY
Le

École doctorale nº388
Chimie Physique et Chimie Analytique de Paris Centre

Spécialité
Chimie Physique

Composition du jury :

François-Xavier COUDERT
Directeur de Recherche CNRS, Chimie ParisTech   *Directeur de thèse*

REMERCIEMENTS

Bonjour, merci.

TABLE DES MATIÈRES

NOTATIONS AND ABBREVIATIONS

r triplet of coordinates in real space.

k triplet of coordinates in Fourier space.

p configuration of a system.

$\mathcal{X}$ system or sub-system.

$\|x\|$ Euclidean norm of vector x.

$\|x\|_{\mathrm{p}}$ periodic distance, *i.e.* the smallest Euclidean norm of a periodic image of vector x.

pcu topology from the RCSR.

FAU topology from IZA-SC.

MC Monte-Carlo

1

CATION PLACEMENT IN ZEOLITES

Among the different families of adsorbents, zeolites are among the most used in industry. Small gases tend to adsorb preferentially on cations when they are present, which makes the study of cationic zeolites particularly relevant to the understanding of industrial adsorption processes. This chapter presents the chemical nature of cationic zeolites and explain the challenges and methods used to obtain representative models of them for subsequent adsorption simulations. The last part of the chapter focuses on the newly developed "shooting star" methodology, which significantly reduces the amount of resource needed to compute the repartition of cations in zeolites.

1.1 Cationic zeolites

In this section, we detail the general structure of cationic zeolites and highlight the challenges associated with the simulation of such systems.

1.1.1 Structure and properties

Zeolites are a family of aluminosilicates, first discovered by Axel Fredrik Cronstedt REF in 1756. Similar to glass, their atomic structure is made of tetrahedra where the central atom, called a T-atom, is usually silicon, bridged together by oxygens. They are crystalline however, with typical unit cell lengths varying between 10 Å and 100 Å [REF], and have nanopores.

The different kinds of zeolites are separated by topology: each known zeolite topology is given a capital three-letter code by the International Zeolite Association Structure Committee (IZA-SC) [REF]. As explained in ??, the topology designates the information extracted from the network of chemical bonds: in the case of zeolites, this network can be simplified by taking one vertex per T-atom, and abstracting each oxygen bridge as a simple edge between the two corresponding T-atoms. This simple protocol makes it clear that the net of a zeolite is very close to the structure it represents – in other words, simplifying the crystal into its net does not lose much information on the initial structure – which explains why a classification by topology is the primary way to distinguish different zeolite structures. It is noteworthy that such a classification is vastly insufficient in the case of MOFs, where the chemistry of metal nodes and the nature of the ligands is thus used first; while for other kinds of materials such as glasses, it is downright impossible to establish because of the absence of cell periodicity.

While millions of hypothetical zeolite topologies have been identified by numerical methods [REF], only 256 have been experimentally observed. This discrepancy may stem from the metastability of zeolites, whose porosity make them thermodynamically less favourable than glass [10.1021/acs.cgd.3c00893]. On the other hand, new experimental zeolite topologies have been discovered each year in the last twenty years, which indicate that accessing new structures is also limited by the current synthetic methods used. Moreover, new synthetic routes can allow new chemistry on existing topologies [REF].

Indeed, each given topology may give rise to several zeolites that differ by their exact chemical composition. Isoelectronic metal substitution consists in the replacement of silicon by germanium or, in some cases, titanium, zinc or tin [REF]. A much more common kind of substitution consists in replacing some T-atoms by aluminium, or, more anecdotally, boron and gallium [REF]: each such substitution introduces one extra electron in the system, hence these zeolites contain cations to maintain electroneutrality. The nature of the cation itself leads to some more variability: sodium cations are usually the ones used during synthesis [REF], and can then be substituted by other cations [REF]. For simplicity, all zeolites discussed will be considered having only either Si or Al as T-atoms.

Ion exchange is actually the most prevalent industrial use case for cationic zeolites, especially as water softener for laundry agents [REF]. Their considerable internal surface and the accessibility of active sites such as the cations or the substituted metals make them also useful for catalysis [REF]. Other applications include nuclear waste storage, building material additives, water absorbents, soil treatment, and others. Finally, they can be used as adsorbents for gas separation or storage, which is also of capital industrial interest.

Zeolite structures are usually rigid structures, although some exceptions exist [REF]. In addition, T-atom substitution tends to modify only slightly the T-O bond length (Si-O: REF while Al-O: REF) and does not affect the O-T-O angles [REF]. As a consequence, the structure of a substituted zeolite can usually be well approximated by taking the reference structure of the topology

(provided as a pure silicate by the IZA-SC) and replacing the relevant silicon atoms. However, not all placements of the substituents may correspond to existing zeolites

1.1.2 Aluminium placement

Cationic zeolites have some aluminiums as part of their structure, however their repartition among the T-atoms is not random.

Rules

The first rule, which is obeyed by almost all known zeolites, was discovered by Löwenstein [REF]. It states that no two aluminiums may be neighbours, in the topological sense of being bridged by an oxygen. This hard rule limits the maximum number of aluminiums to half the number of T-atoms in general, and less for topologies that contain odd cycles: for instance, if there is a cycle of 5 T-atoms, only two among them can be Al.

The amount of aluminiums in a zeolite is usually discussed in terms of Si/Al ratio. Löwenstein's rule thus limits the minimal Si/Al ratio to 1 in general, which corresponds to the maximum number of Al and thus of cations. Zeolites with high Si/Al ratios, usually called high-silica zeolites, contain few cations and tend to be hydrophobic [REF]; in general, they also tend to adsorb fewer molecules since cations are privileged adsorption sites.

While Löwenstein's rule is widely accepted as an axiom, and is mainly sustained by the absence of experimental synthesis of any zeolite with Si/Al ratio lower than 1, it is not currently proven. This comes from the cost of the few spectroscopic methods that allow identifying the precise locations of Al atoms: ^{27}Al and ^{29}Si magic-angle spinning NMR can be used to provide some limited information, while [10.1039/B203966B, 10.1039/B301634J] used UV-Visible-NIR spectroscopy on Co^{2+}-exchanged ZSM-5 with high Si/Al ratio to retrieve their Al distribution. But this method cannot be generalized to low-silica zeolites. DFT computation confirms the rule in general, but also show potential violations for some zeolite topologies such as HEU [10.1016/j.mtcomm.2021.102028].

The second rule, proposed by [REF Dempsey et al, 10.1021/j100722a020] on the basis of FAU zeolites with varying Si/Al ratios, states that when possible, the number of next-neighbours Al atoms should be minimized to decrease the number of unfavourable interactions between close aluminiums. This rule is actually known to be violated in many real zeolites, including dealuminated FAU. Speculatively, this may come from the fact that the [REF Dempsey et al] made their observations on zeolites whose synthesis lead them to reach their global energy minimum, while many other synthetic routes can create structures trapped in a local energy minimum with respect to their Si/Al ordering.

Simulation methodologies

Overall, the lack of experimental evidence attesting to the precise position of aluminiums in zeolites make the theoretical prediction of such placement delicate, since it lacks any verifiable comparison point. In particular, a crucial question is whether Si/Al ordering is truly guided by thermodynamical consideration or not. The latter case would appear for zeolites whose structure is in a metastable state, which is likely to be the case for many real zeolites as evidenced by [10.1039/B301634J] in one of the few experimental studies tackling that very issue.

Even if the aluminium placement can be obtained as that which minimizes the global energy of the structure, this does not necessarily mean that the placement itself is periodic, nor that it follows the same periodicity as the unit cell. Yet, as it is difficult to provide a concrete result on the localisation of aluminiums without making such hypotheses, the studies in the literature only focus on the study of the placement of aluminiums in a single unit cell, supposing that the rest is periodic.

In these conditions, the last remaining question is the simulation methodology to use to find the global energy minimum of the structure. The most basic method consists in generating a few random configurations, computing their energy, and choosing that which minimizes the energy. [REF Marie Jeffroy] proposed a more refined Monte-Carlo (MC) protocol for this problem: starting from a configuration with the target Si / Al ratio under Löwenstein's rule, each MC step consists in exchanging the positions of a Si and an Al while maintaining Löwenstein's rule, and accepting or refusing the step according to the Metropolis criterion (see [infra]). For low Si / Al ratios, this algorithm requires adding some "restart" option that entirely repopulates the aluminiums at random under Löwenstein's rule, because that very rule can lead to blocked situation where no single Si / Al exchange is possible, see [FIGURE].

[REF Findley et al., 10.1021/acs.jpcc.8b03475] proposed an innovative method to assess the validity of theoretically-obtained aluminium placements, by studying CO_2 adsorption isotherms. They found that different Si / Al orderings lead to different isotherms, which, conversely, means that obtaining a precise experimental isotherm can allow retrieving the degree of Al ordering. In accordance to other results of Marie Jeffroy [REF], they conclude that some topologies (REF) are more likely to have their aluminiums placed at random compared to others (REF).

Hence, given the uncertainty that any aluminium placements obtained by thermodynamical simulation actually reflects those observed in real zeolites, we choose to consider that the aluminium placement in our zeolites is taken at random. In practise, each zeolite with a given Si / Al ratio will be modelled as the average of six zeolites of that topology with the same Si / Al ratio, all following Löwenstein's rule but having different Al repartition (taking symmetry into account as well). The results will be computed on each of these six models, and averaged to obtain the result for the zeolite.

To obtain the zeolite structures with random aluminium placements obeying Löwenstein's rule and the target Si / Al ratio, we use the following algorithm:

- The starting point is either the zeolite filled with Si if the required Si / Al ratio is above 3, else a zeolite filled with as many Al as possible using a greedy algorithm. That greedy technique consists in doing a graph traversal of the T-atoms, transforming each atom into an Al if it is not the neighbour of an Al, or making it a Si otherwise: if all cycles in the graph are even, this is guaranteed to provide an aluminium placement with the minimal Si / Al of 1.

- The main algorithm then alternates between adjusting the number of Al and changing their positions, for a few steps each, back and forth.

  - The position update follows the principle of Marie Jeffroy's MC scheme but disregarding the energy change: one Si and one Al are exchanged at random if the resulting structure still follows Löwenstein's rule.

- The number of Al atoms is modified if the current Si /Al ratio does not correspond to the target: if it is too high, an Al is converted into a Si, and otherwise a Si that has no Al neighbour is converted into an Al.

- Since the situation can be blocked (see [FIGURE]), the algorithm sometimes restart from scratch, or sometimes does a partial restart by converting multiple Al atoms into Si.

- Configurations where the correct Si /Al ratio is reached are kept in a set. Each configuration is assigned a hash, which is simply a sequence of 0 if the T-atom is a Si or 1 otherwise, in a fixed order of the T-atom. To avoid considering symmetry-equivalent configurations as different, the signature of a configuration is computed as the lexicographically minimum of all the hashes corresponding to the application of a symmetry operation on the configuration. For each unique signature, a corresponding configuration is then retained, and the six models are taken at random among these configurations.

One last consideration is the choice of the supercell. It appears that, for a large number of zeolites topologies, the smallest unit cell cannot be filled with aluminiums so that it both obeys Löwenstein's rule and reaches the minimal Si /Al ratio for that topology. In the case of LTA for instance, the smallest compatible cell is a $2 \times 2 \times 2$ supercell of the smallest unit cell. As a consequence, the greedy algorithm used before is actually attempted on several supercells, and the one leading to the smallest Si /Al ratio is retained.

The full list of minimum Si /Al ratio found for each zeolite topology (except interrupted ones) and the respective supercell is presented in Table 1.1. The supercell is considered with respect to the idealized framework available for each topology on the IZA-SC database[1]. This table presents the lowest Si /Al ratio found, but apart for the cases where it is 1, the values here are not proven to be minimal, since the algorithm relies on random attempts. Similarly, the supercell used could be not minimal.

It is worth mentioning that a previous screening study of adsorption in zeolite by [10.1021/acsami.0c20892] claimed to use the lowest possible Si /Al ratios to screen through zeolites, but failed to reach the ones presented here, systematically in the case where a supercell needed to be used. Another [10.1021/acs.jpcc.3c06115, 10.1021/acs.jpcc.4c00066] alluded to that issue in the case of LTA, yet failed to identify some other topologies (ATT, JBY, LEV, LTL, MRT, OFF) which could reach an Si /Al ratio of 1. The choice of the supercell during aluminium placement is thus an easily overlooked, yet critical, point to consider.

1.1.3 Cations

The electroneutrality of the framework requires that each negative introduced by an Al substitution be compensated by the introduction of a cation in the framework. In many cases, Na^+ or H^+ are used during synthesis; they are sometimes replaced by other cations in later steps.

These cations have a primordial role for adsorption processes, because they constitute privileged adsorption sites for many small gases. The precise mechanism by which they contribute to the overall adsorption capacity of a material depends on their nature, their location, and the nature of the guest molecule however. Moreover, in addition to simple one-to-one adsorption cases of a

---

[1]accessible at https://europe.iza-structure.org/IZA-SC/ftc_table.php

Topology	Si/Al	Supercell	Topology	Si/Al	Supercell	Topology	Si/Al	Supercell	Topology	Si/Al	Supercell	Topology	Si/Al	Supercell
ABW	1	1×1×1	CHA	1	1×1×1	ITH	1.33	1×2×1	MVY	1	1×1×1	SBT	1	1×1×1
ACO	1	1×1×1	CON	1.33	1×1×1	ITR	1.33	1×1×1	MWF	1	1×1×1	SEW	1.36	1×2×1
AEI	1	1×1×1	CSV	1.22	1×2×1	ITT	1.56	2×2×3	MWW	1.53	2×2×1	SFE	1.33	1×1×1
AEL	1	1×1×1	CZP	1	1×1×1	ITW	1.4	1×1×1	NAB	1.5	2×2×1	SFF	1.46	2×1×2
AEN	1	1×1×1	DAC	1.57	1×3×2	IWR	1.33	2×2×2	NAT	1.22	2×2×4	SFG	1.47	1×2×2
AET	1	1×1×1	DDR	1.73	1×2×1	IWS	1.27	1×1×2	NES	1.52	1×1×1	SFH	1.29	5×1×2
AFG	1	1×1×1	DFO	1	2×2×2	IWV	1.62	1×1×2	NON	1.44	1×1×1	SFN	1.29	1×5×2
AFI	1	1×1×1	DFT	1	4×4×3	IWW	1.33	1×1×1	NPO	2	4×4×5	SFO	1	2×2×4
AFN	1	1×1×1	DOH	1.62	1×2×1	JBW	1	5×4×3	NPT	2	2×2×2	SFS	1.67	1×1×1
AFO	1	1×1×1	DON	1.29	1×1×1	JNT	1	1×1×1	NSI	1.4	2×5×3	SFW	1	1×1×1
AFR	1	2×2×4	EAB	1	1×1×1	JOZ	1.5	2×1×1	OBW	1.92	1×1×1	SGT	1.49	2×2×1
AFS	1	1×1×1	EDI	1	4×4×4	JRY	1	1×1×1	OFF	1	3×3×4	SIV	1	1×1×1
AFT	1	1×1×1	EEI	1.44	1×1×1	JSN	1	3×4×2	OKO	1.62	1×1×1	SOD	1	1×1×1
AFV	1	3×3×2	EMT	1	1×1×1	JSR	1.67	1×1×1	OSI	1	1×1×1	SOF	1.5	1×1×1
AFX	1	1×1×1	EON	1.4	2×1×1	JST	2	1×1×1	OSO	2	1×1×1	SOR	1.18	1×1×2
AFY	1	1×1×1	EOS	1.18	1×2×2	JSW	1	1×1×1	OWE	1	2×4×3	SOS	1.4	2×4×3
AHT	1	1×1×1	EPI	1.74	4×2×3	JSY	1	1×1×1	PAU	1	1×1×1	SOV	1.29	1×1×1
ANA	1	1×1×1	ERI	1	1×1×1	JZO	1.37	1×1×1	PCR	1.5	1×1×1	SSF	1.25	2×2×2
ANO	1	1×1×1	ESV	1.4	1×1×1	JZT	1.35	1×1×1	PHI	1	1×1×1	SSY	1.33	1×1×1
APC	1	1×1×1	ETL	1.48	2×1×1	KFI	1	1×1×1	PON	1	1×1×1	STF	1.46	1×1×2
APD	1	1×1×1	ETR	1	1×1×1	LAU	1	2×2×4	POR	1	1×1×1	STI	1.25	1×1×1
AST	1	1×1×1	ETV	1.33	1×1×1	LEV	1	3×3×2	POS	1.29	1×1×1	STT	1.46	1×1×1
ASV	1	1×1×1	EUO	1.49	1×1×1	LIO	1	1×1×1	PSI	1	1×1×1	STW	1.5	1×1×1
ATN	1	1×1×1	EWF	1.49	2×1×1	LOS	1	1×1×1	PTF	1.5	1×1×1	SVV	1.33	1×1×1
ATO	1	1×1×1	EWO	1.4	2×2×5	LOV	1.25	2×2×1	PTO	1.4	1×1×1	SWY	1	1×1×1
ATS	1	1×1×1	EWS	1.53	1×1×1	LTA	1	2×2×2	PTT	1	1×1×1	SZR	1.25	1×1×2
ATT	1	3×4×3	EZT	1	1×1×1	LTF	1.27	1×1×4	PTY	1.5	1×1×1	TER	1.35	1×1×1
ATV	1	1×1×1	FAR	1	1×1×1	LTJ	1	1×1×1	PUN	1.25	1×1×1	THO	1	2×4×4
AVE	1	1×1×1	FAU	1	1×1×1	LTL	1	2×2×4	PWN	1	1×1×1	TOL	1	1×1×1
AVL	1	3×3×2	FER	1.57	1×1×3	LTN	1	1×1×1	PWO	1.5	1×1×1	TON	1.4	2×2×5
AWO	1	1×1×1	FRA	1	1×1×1	MAR	1	1×1×1	PWW	1.5	1×1×1	TSC	1	1×1×1
AWW	1	2×2×4	GIS	1	1×1×1	MAZ	1.4	1×1×2	RFE	1.4	1×1×1	TUN	1.53	1×1×1
BCT	1	1×1×1	GIU	1	1×1×1	MEI	1.43	1×1×1	RHO	1	1×1×1	UEI	1	1×1×1
BEC	1.29	2×2×2	GME	1	1×1×1	MEL	1.67	1×1×1	RRO	1.25	2×1×1	UFI	1.29	2×2×1
BIK	1.4	4×2×5	GON	1.29	1×1×1	MEP	1.71	1×1×1	RSN	1.59	4×1×4	UOS	1.18	2×4×3
BOF	1.4	2×1×1	GOO	1	1×1×1	MER	1	1×1×1	RTE	1.4	1×1×2	UOV	1.38	1×1×1
BOG	1.4	1×1×1	HEU	1.25	1×1×2	MFI	1.67	1×1×1	RTH	1.53	2×1×3	UOZ	1	1×1×1
BOZ	1.88	1×1×1	IFO	1	1×1×1	MFS	1.45	3×1×1	RUT	1.57	2×2×3	USI	1	1×1×1
BPH	1	1×1×1	IFR	1.29	1×1×2	MON	1.67	2×2×1	RWR	1.6	4×4×1	UTL	1.62	1×2×2
BRE	1.29	4×2×4	IFW	1.29	1×1×1	MOR	1.67	1×1×2	RWY	3	1×1×1	UWY	1.37	1×2×3
BSV	1	1×1×1	IFY	1	1×1×2	MOZ	1	1×1×4	SAF	1	1×1×1	VET	1.43	1×1×1
CAN	1	1×1×1	IHW	1.55	1×1×1	MRT	1	2×4×2	SAO	1	1×1×1	VFI	1	1×1×1
CAS	1.4	1×1×1	IMF	1.55	1×1×1	MSE	1.55	1×1×1	SAS	1	1×1×1	VNI	1.73	3×3×1
CDO	1.62	4×2×2	IRN	1.19	1×1×1	MSO	1	2×2×2	SAT	1	1×1×1	VSV	1.78	4×4×1
CFI	1.46	2×5×1	IRR	1.6	2×2×2	MTF	1.44	1×1×2	SAV	1	1×1×1	WEI	1.5	1×1×1
CGF	1	2×2×4	ISV	1.31	2×2×1	MTN	1.72	1×1×1	SBE	1	1×1×1	YFI	1.4	1×1×1
CGS	1	1×1×1	ITE	1.56	1×2×1	MTT	1.4	1×1×1	SBN	1.5	1×2×1	YUG	1.56	3×2×4
			ITG	1.33	1×1×2	MTW	1.33	1×5×3	SBS	1	1×1×1	ZON	1	4×2×2

Table 1.1: Minimal Si/Al obtained with corresponding supercell for all non-interrupted topologies

small molecule on a cation, X-ray spectroscopy evidences the presence of adsorption complexes involving several cations and guest molecules at the same time [10.1039/C2CP23237E]. Finally, the very nature of the cations present in the structure can be quite diverse, ranging from monovalent alkali to polyvalent metal ions, and even molecular ions such as many of the organic structure directing agents (OSDAs) used during zeolite synthesis.

In order to perform numerical simulation of the adsorption process, it is thus necessary to have a model of the zeolite that places its cations in the correct positions. Fortunately, and opposite to the previous case of aluminiums, X-ray diffraction is a convenient technique which can be used to obtain the crystallographic sites in which the cations are located, as well as the respective occupancy of each such site.

1.2 PREDICTION OF CATION PLACEMENT

While experimental cation maps are available on some zeolite structures, two seminal problems remain. First, when screening through a large amount of zeolite topologies and Si/Al ratio, only a small fraction of structures have their experimental cation map available. Second, even with the crystallographic sites, not all cation repartitions that obey the given occupancies have the same energy. For example, the site I and I' of FAU being very close to one another, the occupation of a site I prevents either of the neighbouring sites I' from being occupied at the same time. Deciding on the exact location of cations in a zeolite model is thus no trivial matter.

1.2.1 Current methodologies

To obtain equilibrium data, such as average positions, on molecular systems, there are two general simulation frameworks which can be used: molecular dynamics and Monte-Carlo (MC) simulation. We only used MC and its variations, because they are the only strategy available to study a variable number of molecules, which is necessary for the simulation of adsorption as we will explain in more details in **??**. The possible use of molecular dynamics and its variants will be discussed later in section 1.4.

CANONICAL MONTE-CARLO

The Monte-Carlo (MC) method designates a large variety of algorithms which all converge towards their result by performing multiple random samplings. A wide subset of these form the Markov Chain Monte-Carlo (MCMC) schemes, which are a class of algorithms that aim at sampling a target probability distribution. In the context of statistical physics, such methods provide practical ways of obtaining macroscopic observable values from the simulation of the microscopic behaviour of particles. Indeed, an observable $\mathcal{O}$ can be computed as the ensemble average of its microscopic counterpart $\langle o \rangle$, which can be computed as an integral over the configuration space. For example, in the canonical ensemble where the number of particles, the external temperature and the volume of the system is fixed,

$$\mathcal{O} = \langle o \rangle = \frac{1}{Z} \int \mathrm{d}\boldsymbol{p}\, o(\boldsymbol{p}) \exp\left(-\beta \mathrm{U}(\boldsymbol{p})\right) \tag{1.1}$$

where $\mathrm{U}(\boldsymbol{p})$ is the energy of configuration $\boldsymbol{p}$ and Z is the partition function. In other words, $\mathcal{O}$ can be computed as an integral of values of o on a probability distribution of $\boldsymbol{p}$ given by $p \mapsto \frac{1}{Z} \exp\left(-\beta \mathrm{U}(\boldsymbol{p})\right)$.

The direct evaluation of this probability distribution is impossible in general, because of the partition function Z whose value cannot be computed. Instead, the MCMC method gives a way to directly evaluate the observable $\mathcal{O}$ by providing a sequence of configurations $\boldsymbol{p}$ that follows the same probability distribution as the target. The integral can then be computed as simply the average value of o on this specific set of configurations $\boldsymbol{p}$.

The Metropolis-Hastings algorithm is the most common MCMC scheme used to provide such a sequence of configurations. It consists in starting from an initial state, *i.e.* an initial position for all the mobile particles of the system, and making it evolve through a sequence of steps. At each step, a trial move, *i.e.*, a modification of the state, is attempted and may, or may not, be accepted with a certain probability. In the case of a simulation in the canonical ensemble, the difference of energy ΔE between after and before the move gives the probability $\min\left(1, \exp\left(-\beta\Delta E\right)\right)$ that the move be accepted. Intuitively, this means that any move that decreases the energy of the system is accepted, while the others are accepted with a probability that decreases as the energy difference increases, sharply at low temperature and more slowly at high temperature. In the limit of infinite temperature, all the moves are accepted; at zero temperature, only the moves that decrease the energy are. The details on how to compute energy are deferred to section 1.3.

To avoid drift, the algorithm obeys detailed balance which, conceptually, represents the microreversibility of the simulated physical processes. In practise, this means that, for each MC move that brings the system from configuration $\boldsymbol{p}_{\mathcal{A}}$ to $\boldsymbol{p}_{\mathcal{B}}$, there must exist a "converse" MC moves that can bring the system from $\mathcal{B}$ to $\mathcal{A}$ such that $\pi_{\mathcal{A}}\mathrm{P}_{\mathcal{A}\to\mathcal{B}} = \pi_{\mathcal{B}}\mathrm{P}_{\mathcal{B}\to\mathcal{A}}$ where $\pi_{\mathcal{X}}$ is the stationary probability of state $\boldsymbol{p}_{\mathcal{X}}$ and $\mathrm{P}_{\mathcal{X}\to\mathcal{Y}}$ is the probability of the move from $\boldsymbol{p}_{\mathcal{X}}$ to $\boldsymbol{p}_{\mathcal{Y}}$. In the canonical ensemble, $\pi_{\mathcal{X}}$ is equal to to $\frac{1}{Z}\exp\left(-\beta\mathrm{U}\left(\boldsymbol{p}_{\mathcal{X}}\right)\right)$, while the probability of transition is $\min\left(1, \exp\left(-\beta\Delta E\right)\right)$. Hence, detailed balance simply translates to the algorithmic condition that the probability of choosing the trial move $\mathcal{A} \to \mathcal{B}$ must be equal to that of choosing the trial move $\mathcal{B} \to \mathcal{A}$, irrespective of the probability of accepting these moves.

The MC moves which are used for simulations in the canonical ensemble are usually the following:

- translation: displace each atom of the particle by the same vector, taken uniformly at random in a sphere.

- rotation: rotate the particle around one of its atoms (called the "bead") along three uniformly random Euler angles, possibly constrained to some values only.

- reinsertion: remove and reintroduce the particle uniformly at random in the system.

To satisfy detailed balance, the sphere in which the translation vector is taken must have a fixed radius, the bead must be fixed and the angle constraints must be symmetric as well. For the particular case of the movement of mono-atomic species, such as most cations in zeolites, the rotation MC move is useless and thus not used.

Overall, the algorithm thus consists in a sequence of steps, each of which is an attempt to displace a cation either locally (translation) or not (reinsertion) and is accepted or not depending on the difference of energy the move generates. Configurations corresponding to neighbour steps are thus either identical if the move was refused, or only separate by one atom move, so

they are strongly correlated. To reduce computational cost, the computation of the microscopic state function o is thus only realized once every cycle, each cycle corresponding to a number of steps. We use either 20 or the number of mobile atoms in the system, whichever is the highest.

The Metropolis-Hastings algorithm thus provides a way to walk through the configuration space. By privileging moves that decrease energy, the system spends more steps in low energy regions; by yet allowing opposite moves sometimes, the scheme ensures that all the configuration space is eventually visited. Hence, the algorithm guarantees ergodicity: in an infinitely long simulation, the fraction of steps spent in a part of the configuration space is proportional to its volume. Finding the space with the lowest energy, and thus the largest probability and volume, should then simply be a matter of running the simulation long enough for the state to fall in the deepest energy pit.

SIMULATED ANNEALING

Unfortunately, simply doing canonical MC at room temperature does not provide the cations with enough energy to overcome the barriers inherent to the structure of the zeolitic framework in any reasonable time. For example, a cation trapped in the site I of FAU will never manage escape into a site III in a room-temperature simulation of reasonable length without a reinsertion move: this simply comes from the high activation barrier it needs to overcome while physically moving out of the sodalite cage. Reinsertion moves are thus useful to allow large configuration movements to occur without having to pass through physical energy barriers.

However, a reinsertion can only move one particle at a time, which is sometimes not enough. For instance, the population of two sites I' from an initial configuration where only the central site I is occupied, requires that both the central cation move to the edge of the cage and that an external cation be set in the opposite window. Since both movements require passing through an activation barrier, the probability of it happening naturally is low, and thus requires particularly long simulation times to be observed.

To circumvent this issue, the temperature of the simulation can be raised: this mechanically makes all MC steps more likely to be accepted, so higher-energy regions become more reachable. At the same time, a simulation at a hot temperature is much less likely to explore the details of the energy landscape enough to fall into the energy minima of interest, since the accessible configuration space became so much larger. A compromise thus consists in performing simulated annealing, which is simply a simulation in which the temperature varies between cold and hot, alternating between increasing and decreasing phases, possibly interspersed with plateaus.

PARALLEL TEMPERING

Simulated annealing explores the configuration space by alternating between cold phases, where the exploration is detailed but localized, and hot phases, where the exploration is fast but superficial. Parallel tempering, also called "replica exchange" proposes doing both phases at the same time.

In more detail, a parallel tempering simulation consists in running several MC simulations, the "replicas", in parallel, each at a different temperature. The temperatures are taken between room temperature and a hot temperature. An extra MC step is introduced, called the "exchange" step, which consists in exchanging the configurations of two replicas with neighbour temperatures.

As with any MC step, the Metropolis-Hastings algorithm ensures that the exchange is accepted only if the difference of energy resulting from this exchange is not too high: as a consequence, for the exchange step to be accepted regularly, the temperatures must be taken so that the typical energies of the system between neighbour temperatures overlap, as shown on [FIGURE].

In this strategy, the coldest simulation is constantly exploring a local free energy minimum, while the hottest simulation easily accepts MC moves and thus walks through the entire configuration space. By exchanging configurations, new configurations regularly trickle down from the hot simulations to the colder ones, while the general Metropolis-Hastings strategy ensures convergence towards a global minimum. It is also more efficient than simulated annealing since the exploration of the configuration space is not constrained by the small duration of the hot phase. Parallel tempering is therefore the method of choice used to obtain convergence when facing difficult energy landscapes such as the one encountered for cation placement.

Naturally, its main downside is its computational cost, proportional to the number of simulations and thus effectively related to the difference between the highest and the lowest temperatures. The choice of intermediate temperatures must be guided by an initial discovery of the typical energies encountered by system at each temperature, which requires a few dedicated simulations prior to the parallel tempering. Finding the optimal number of simulations requires striking a balance with the proportion of accepted MC exchange steps, whose impact on the overall simulation convergence can be difficult to measure. Additionally, simulations with neighbour temperatures must regularly synchronize to attempt an exchange, which strongly limits the amount of parallelization usable in the implementation of parallel tempering.

Combined aluminium and cation placement

Specifically in the case cationic zeolites, [Marie Jeffroy REF] proposes a simulation method to obtain both the placement of cations in the structure and the location of aluminium atoms across the T-sites. This consists in running an MC simulation where the cations are allowed all previously mentioned MC steps, while the aluminiums are only allowed exchanging positions with siliciums, or a restart step that randomly positions aluminiums and siliciums, both while obeying Löwenstein's rule as explained in **??**.

The statistical ensemble simulated in this setting includes both aluminium and cations, which means that the distribution sampled through such a simulation method represents the equilibrium for both species. From a chemical point of view, this situation is representative of a framework whose synthesis was done with that very cation, and was thermodynamically-driven. While this does correspond to reality in some cases, there is also evidence that the synthesis is kinetically-driven in many other occurences: for such situation, obtaining the aluminium positions from statistical equilibrium is not representative of the actual framework. And in yet other situations, the cation is exchanged post-synthesis, which decorrelates the aluminium placement in the framework from the cation used for adsorption purposes.

Nonetheless, even in those cases, there is no less value in the aluminium placement obtained from such a hybrid strategy than from random sampling of the T-atoms. In fact, by virtue of the added dimensions of the free energy landscape, a kinetically trapped cation may move more freely following the rearrangement of the aluminiums in the structure, which can make

the entire simulation converge more efficiently. If not for the hope of obtaining more relevant aluminium placements, this combined strategy remains valuable for its diminished computationally cost compared to parallel tempering.

Unfortunately, the fact that the framework is not strictly constant as before prevents the use of precomputed energy grid, later discussed in section 1.3.4. This was not a problem in the context of [Marie Jeffroy]'s study because the force field they used was such that a silicium-aluminium swap would not change the van der Waals interactions, but this is not the case of most force fields, including the one we use in our study [BoulfelfelSholl2021]. Being unable to precompute the short-term energy contribution of the framework would thus make the simulations significantly slower, tipping the balance against the use of this method.

1.2.2 Shooting star methodology

Previously exposed methods such as canonical MC, simulated annealing or parallel tempering focus on making the system (or at least, one of its replicas) reach equilibrium after a certain number of initial steps, so that the rest of the simulation occurs at equilibrium. This is crucial for the problem of sampling the equilibrium distribution of states in general. For the case of cation placement in zeolites however, the observed crystallographic positioning of the cations indicates that this distribution is sharply peaked: only a discrete number of sites can be occupied, and the question is to find their positions and occupancy. As a consequence, we can use another, simpler, strategy, which is less efficient for the actual sampling of the distribution, but can uncover relevant local minima faster.

PRINCIPLE

This new methodology takes the same simulation elements as the ones previously detailed, but arranges them in a different fashion. We call it "shooting star" because it can be decomposed into one hot simulation, like the bright head of the meteor, and a series of cold simulations that stem from it, like the cooling incandescent debris left in its wake.

The first element is the exploration of the phase space. In order to fully explore it, a naive approach consists in doing steps, each of which randomly reinserts a particle in the system. Without the possibility for step rejection, this leads to sampling extremely unlikely states almost all the time however. We propose to perform global exploration by simply running a canonical MC simulation at a very high temperature, for instance $2000\,\mathrm{K}$. Such an approach makes most kinetic obstacles irrelevant, but still prevents absurd configurations like when two atoms collide. This MC simulation is simply called the hot simulation.

The second element is the local minima exploration, which is grafted onto the previous one: every Δ_{spawn} cycles of the hot simulation, a snapshot of its current configuration is taken and used as the starting point for a cold simulation, *i.e.* another canonical MC simulation that runs at room temperature. Δ_{spawn} is chosen so that two consecutive starting points should be uncorrelated. The higher the temperature of the hot simulation, the less correlated its consecutive steps are, so Δ_{spawn} can actually be very low. Each cold simulation runs for M_{init} initial steps, that allow the system to cool down, and M_{prod} production steps at the target room temperature.

Similar to the previous accelerated methodologies, only a fraction of the explored states can actually be used to perform statistics. In the case of simulated annealing, it corresponded to

the cold regions of the simulation; for parallel tempering, only the coldest replica is useful to obtain statistics at that temperature; for the current methodology, the relevant steps are the production of the cold simulation. The main advantage of the shooting star resides in the efficient exploration of the configuration space, rooted in the hot simulation that is independent from the cold ones. Conversely, it lacks the physical meaning of simulated annealing – which is the numerical counterpart to the annealing process used in metallurgy – or the time-reversibility of parallel tempering and bare canonical MC.

In terms of parameters, the shooting star method requires choosing T_{hot} the hot temperature, T_{cold} the cold one, N_{cold} the total number of cold simulations to launch, the previously mentioned M_{init}, M_{prod} and Δ_{spawn}, as well as the number N_{init} of initial steps of the hot simulation before launching the first cold simulation. Of all these parameters, T_{cold}, N_{init} and $N_{prod} = M_{prod} \times N_{cold}$ must also be chosen for a single canonical MC simulation and Δ_{spawn} can always be taken around 100 (its value has little effect on overall performance and quality as long as it is not too low nor too high). The extra parameters T_{hot} is equivalent to the hottest temperature used in simulated annealing and parallel tempering. Finally, M_{init} must be evaluated by an initial run, similar to how the intermediate temperatures used in parallel tempering are obtained or the slope and plateau lengths of simulated annealing are usually chosen. The complexity of the overall setup is thus similar to that of simulated annealing.

Extraction of sites from densities

To perform gas adsorption prediction, what is needed is a model of the cationic zeolite. With this perspective, the previous simulations can be used to simply provide their cation configurations with the lowest energy, in order to obtain such a model.

It is interesting however to evaluate the shooting star methodology as a new general meta-algorithm to simulate systems in sharp energy landscapes. In order to do so, we need a comparison point that can assess the validity of the pseudo-simulation obtained by gluing together all the production steps of the cold simulations. Such a comparison point is the experimental cation localization that can be obtained by X-ray diffraction on zeolite samples with known topology and Si / Al ratio.

The shooting star (or other) simulation method yields a sequence of cation configurations sampled at room temperature. This sequence can be flattened into a map of the zeolite framework, with one additional point at each position occupied by a cation in one of the steps. Looking at this map allows qualitatively identifying the crystallographic sites as the zones where the density of points is the highest. In order to get quantitative information on the site localization and population, we devise a clustering algorithm that extract the sites from the previous density map.

The first step consists in dividing the space by a regular grid, whose voxels have the same angles as the unit cell but possibly different lengths, taken to be closest to a target value around 0.15 Å. To each voxel is then attributed the number of cations encountered within across the recorded simulation steps. This first step thus bins the density onto a regular grid.

In the second step, the bins are sorted by decreasing order of their value. Each voxel is then taken in order: if it is closer than a set distance to a previously recorded site, then the location of the site is updated as the average of its current position and that of the voxel, weighted by the respective density of both; the density of the voxel is then aded to that of the site. Otherwise,

if it is far enough from all existing sites, it is added to the (initially empty) list of recorded sites, along with its density. This yields a series of sites and their respective occupancy. To be more efficient, the main loop early stops when the density of the inquired voxels go below a threshold value, taken as the maximum density of all bins divided by 100: as a consequence, the occupancies are corrected by a multiplicative factor so that the sum of the occupancies of all sites correspond to the number of cations. We note that this should not be required in theory, but it is the necessary fix to a double discretization artifact, once inherent to the MC scheme, and one due to our binning.

We observe slightly better results by sorting the bins according the the decreasing order their smoothed counterparts, where smoothing is done by convolving the density map by a small Gaussian, of standard deviation the size of the diagonal of a voxel. Indeed, this erases some small statistical fluctuations that may change the precise order of the voxels.

VALIDATION

1.3 ENERGY COMPUTATION

All of the previously exposed methods rely on Monte-Carlo steps which may, or may not, be accepted, depending on the difference of energy between the two configurations. By far, the greatest computational cost of any of these methods lies in these energy computations; it is therefore crucial to be able to compute it efficiently.

THE RASPA2 SOFTWARE

RASPA2 is an open-source software developed by [REF] for classical simulations of molecular systems. Although general purpose, it is mainly used with nanoporous materials and in particular for the study of adsorption and diffusion within. RASPA2 implements Monte-Carlo schemes with the Metropolis-Hastings algorithm to perform canonical simulations, but also grand canonical for adsorption (see **??**), as well as parallel tempering. Furthermore, it includes Gibbs ensemble simulation, flexible guest species and frameworks, and many other functionalities which were not needed here.

RASPA2 is written in C, a low-level programming language which allows reaching optimal performance, but is tedious to program in. It cannot be used on GPU and is single-threaded. In order to develop new simulation strategies in an easier setting, I re-implemented the algorithms used for both canonical and grand-canonical simulation of rigid species in rigid frameworks using the Julia programming language (see **??**).

The energy computation algorithms detailed herein are those which were implemented in this new code, but most of them follow that implemented in RASPA2. Moreover, RASPA2 was used to validate the new code, by checking that the simulations yielded the correct physical values (energies, adsorption capacities, and such). By contrast, the shooting star methodology presented before does not exist in RASPA2.

1.3.1 Force fields

The energy of a molecular structure results from the interaction of its constituent atoms, which are intrinsically quantum. At the molecular level, a popular method used is the quantum Density Functional Theory (DFT), which includes a variety of particular methods, each presenting its own compromise between precision and performance. Some alternatives include methods

derived from molecular orbital theory. The number of energy computations for a single molecular simulation can be of the order of 10 000 however, which would be much too costly with such methods.

What is commonly used instead are classical force fields, which approximate quantum interactions as a sum of classical interactions between each pair of atoms. These pairwise interactions are defined by one formula per pair of nature of atoms; in the case of non-bonded atoms, each formula is simply a function of the distance between the two atoms. Zeolites are mostly rigid materials, which means that their deformation due to their vibration modes, or subsequent to the organization of cations or adsorbed gas inside, is negligible for the purpose of the molecular simulation. As a consequence, the interactions between framework atoms can be skipped, the only ones to consider for cation placement are those between the cation and the different atoms of the zeolites, which are not bonded.

The pairwise interaction formula is usually divided between the energy contribution of the charges of the atoms, and a short-range van der Waals term. Designing a force fields thus customarily consists in choosing the atomic charges and the formula for the van der Waals term. Several options have been used in the literature for the latter: a very common option is the Lennard-Jones potential

$$r \mapsto 4\varepsilon \left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right) \tag{1.2}$$

where r is the distance between atoms. One alternative is the Buckingham potential

$$r \mapsto A \exp\left(-B \times r\right) - \frac{C}{r^6} \tag{1.3}$$

The potential parameter ε, σ, A, B or C are part of the design of the force field.

1.3.2 Short-term interactions

Since the van der Waals term decays quickly towards zero, it is computationnally useful to set it to exactly zero beyond a cutoff distance. To do so while avoiding the introduction of unphysical forces, one possibility consists in shifting the formula by a constant, so that the function remains continuous at the cutoff. If no force is required, such as in the MC methods, it is possible to leave the the potential unshifted and discontinuous at the cutoff, but a constant tail correction can be added to the total energy to account for the missing interactions. In any case, knowing that the interaction becomes zero after the cutoff allows some computational optimizations, such as using cell lists in order to iterate over the relevant atom pairs only, instead of all atom pairs.

For simplicity of the implementation, the dimensions of the simulated unit cell must be taken so that for any atoms A and B, there is at most one periodic image of B which is lower than the cutoff distance from atom A. This translates to the constraint that the orthogonal lengths of the unit cell must be as long as at least twice the cutoff. We take this cutoff as 12 Å, as is common in the literature [REF] and is the default in RASPA2.

Given this constraint, the total energy of a configuration due to the van der Waals interaction can be computed by summing the contribution of each the pair of atoms which are closer than the cutoff distance. This can done naively with a double loop over all atoms, computing the distance of each pair and rejecting that above the threshold. Here, the distance considered

is the periodic distance, *i.e.* the smallest distance between periodic images of the two atoms, respective to the periodicity of the crystal. Computing it is up to 6 times more costly than without periodicity, and can become a computational bottleneck for simple enough force fields.

To avoid unecessary distance computations, a possibility consists in binning the space of the unit cell, and skipping the pairs of atoms that belong to bins further than the cutoff distance. This idea underlies the principle of neighbour list which can be used to obtain the list of relevant pairs of atoms efficiently. We use the CellListMap.jl [10.1016/j.cpc.2022.108452] julia package for this purpose, although we only enable it when the cell is large enough since the naive approach can be more efficient in many cases where the orthogonal lengths of the cell are close to twice the cutoff distance.

1.3.3 Ewald summation

ENERGY DECOMPOSITION

In opposition to the van der Waals term, charge interactions do not have a cutoff, because Coulomb's law does not decay fast enough towards zero. The interaction between atoms A and B, of respective charge q_A and q_B, is:

$$E_{A,B}^{Coulomb} = \frac{q_A q_B}{4\pi\varepsilon_0 \times \|\boldsymbol{r}_B - \boldsymbol{r}_A\|} = \frac{q_n q_m}{4\pi\varepsilon_0 \times \|\boldsymbol{r}_m - \boldsymbol{r}_n + \boldsymbol{l}\|} \tag{1.4}$$

where n and m are the unique atom numbers in a reference unit cell of respectively A and B, and $\boldsymbol{l}$ is the lattice vector between the origin of the unit cell of A and that of B. Therefore the total contribution of charge interactions in a crystal with N atoms per unit cell is:

$$E^{Coulomb} = \frac{1}{2} \sum_{A \neq B} E_{A,B}^{Coulomb} = \frac{1}{8\pi\varepsilon_0} \sum_{\boldsymbol{l}} \sum_{\substack{n=1 \\ n \neq m \text{ if } \boldsymbol{l}=0}}^{N} \sum_{m=1}^{N} \frac{q_n q_m}{\|\boldsymbol{r}_n - \boldsymbol{r}_m + \boldsymbol{l}\|} \tag{1.5}$$

where $\boldsymbol{l}$ iterates over the lattice vectors.

The semi-convergence of the sum over $\boldsymbol{l}$ makes it impossible to compute directly. Hence, several methods have been developed to obtain the result such as particle mesh Ewald (PME) and particle-particle-particle-mesh (P^3M). We focus on the initial Ewald summation technique underlying them, which is used in RASPA2 and which we also implemented.

Since the crystal is a periodic system, any function of the system that depends on a position in space can be decomposed into a simple sum by Fourier transform. This is in particular true of the electric potential, hence the problematic sum above could be computed in Fourier space. However, the charges are located on the atoms, and are thus discrete, which causes singularities that prevent the computation. The idea behind Ewald summation consists in fictitiously smoothing the point charges into Gaussians: these smooth charges yield a potential computable in Fourier space, while the potential resulting from the real point charge minus their Gaussian counterpart is computed in real space. Electric potential is additive, so the total electric potential $\Phi(\boldsymbol{r})$ is then obtained by adding the real space term $\Phi^{direct}(\boldsymbol{r})$ and the Fourier

term $\Phi^{\text{reciprocal}}(\boldsymbol{r})$. Finally, the resulting energy is simply the action of the point charges on the potential:

$$E^{\text{Coulomb}} = \frac{1}{2} \sum_{n=1}^{N} q_n \left(\Phi^{\text{direct}}(\boldsymbol{r}_n) + \Phi^{\text{reciprocal}}(\boldsymbol{r}_n) \right) - E^{\text{self}} \tag{1.6}$$

E^{self} designates a spurious self-interaction energy term which must be substracted.

DIRECT TERM

To each atom A of the crystal is associated a Gaussian charge density

$$\boldsymbol{r} \mapsto q_A \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 \exp\left(-\alpha^2 \|\boldsymbol{r}_A - \boldsymbol{r}\|^2 \right) \tag{1.7}$$

Thus, the direct term is the potential resulting from the point charge minus this Gaussian. This can be obtained by solving Poisson's law:

$$-\nabla \Phi(\boldsymbol{r}) = 4\pi \rho(\boldsymbol{r}) \tag{1.8}$$

which yields, after some simplification steps not detailed here:

$$\Phi_A^{\text{direct}}(r) = q_A \frac{\text{erfc}(\alpha r)}{r} \tag{1.9}$$

where r is the distance to atom A and erfc is the complementary error function $\text{erfc} = 1 - \text{erf}$, with $\text{erf} : x \mapsto \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t$. Since erfc quickly decays towards 0, the total direct potential at a given point can be computed by only considering the contribution of the closest charges around, so

$$\Phi^{\text{direct}}(\boldsymbol{r}) \approx \sum_{n=1}^{N} \frac{q_n}{\|\boldsymbol{r} - \boldsymbol{r}_n\|_{\text{p}}} \left(\text{erfc}\left(\alpha \|\boldsymbol{r} - \boldsymbol{r}_n\|_{\text{p}} \right) \right) \tag{1.10}$$

where $\|\boldsymbol{x}\|_{\text{p}}$ designates the distance to the origin of the closest periodic image of $\boldsymbol{x}$.

RECIPROCAL TERM

For the reciprocal term, we start by writing the total charge density coming from the Gaussians given in Equation 1.7:

$$\rho^{\text{reciprocal}}(\boldsymbol{r}) = \sum_{\boldsymbol{l}} \sum_{n=1}^{N} \frac{q_n}{4\pi\varepsilon_0} \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 \exp\left(-\alpha^2 \|\boldsymbol{r}_n + \boldsymbol{l} - \boldsymbol{r}\|^2 \right) \tag{1.11}$$

Its Fourier decomposition over the periodic unit cell of volume V is:

$$\begin{aligned}
\rho^{\text{reciprocal}}(\boldsymbol{k}) &= \frac{1}{V} \int_V e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \rho^{\text{reciprocal}}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} \\
&= \frac{1}{V} \int_\Omega e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{n=1}^{N} \frac{q_n}{4\pi\varepsilon_0} \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 \exp\left(-\alpha^2 \|\boldsymbol{r}_n + \boldsymbol{r}\|^2 \right) \mathrm{d}\boldsymbol{r} \\
&= \frac{2\pi\alpha^3}{V\pi^{3/2}} \sum_{n=1}^{N} \frac{q_n}{4\pi\varepsilon_0} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n} \int_0^\infty \|\boldsymbol{r}'\|^2 \exp\left(-\alpha^2 \|\boldsymbol{r}'\|^2 \right) \int_0^\pi e^{-i\boldsymbol{k}\cdot\boldsymbol{r}'} \sin(\theta) \mathrm{d}\theta \mathrm{d}\boldsymbol{r}' \\
&= \frac{1}{4\pi\varepsilon_0 V} \exp\left(\frac{-k^2}{4\alpha^2} \right) \sum_{n=1}^{N} q_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n}
\end{aligned}$$

$$\tag{1.12}$$

after several simplification steps not detailed here. Using Poisson's law (Equation 1.8), the electric potential due to the Gaussians can then finally be expressed as:

$$\Phi^{\text{reciprocal}}(\boldsymbol{r}) = \frac{1}{\varepsilon_0 V} \sum_{\boldsymbol{k} \neq \boldsymbol{0}} \sum_{n=1}^{N} \frac{q_n}{k^2} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}_n)} \exp\left(\frac{-k^2}{4\alpha^2}\right) \tag{1.13}$$

The term $\kappa_{\boldsymbol{k}} = \frac{1}{k^2} \exp\left(-k^2/(4\alpha^2)\right)$ decays fast, so the sum over $\boldsymbol{k} \neq \boldsymbol{0}$ is actually computed as a sum over a finite number of values of $\boldsymbol{k}$. These values are chosen so that for any $\boldsymbol{k}$ not retained, $\kappa_{\boldsymbol{k}}$ is below the target precision of the computation.

SELF-INTERACTION TERM

The last remaining term is the self-interaction, which must be removed. Its existence comes from the way the reciprocal term is computed: $\Phi^{\text{reciprocal}}(\boldsymbol{r})$ includes the contribution of all charges on point $\boldsymbol{r}$, but there should not be any contribution of atom A at position $\boldsymbol{r}_A$.

Moreover, in the case of rigid polyatomic molecules, like the small gases we will discuss in **??**, there should not be any contribution of atom A at the positions $\boldsymbol{r}_B$ of other atoms B part of the same molecule.

Each atom $1 \leq n \leq N$ thus contributes a spurious potential term:

$$\Phi_n^{\text{self}} = \frac{2\alpha q_n}{\sqrt{\pi}} + \frac{1}{2} \sum_{\substack{m \\ m \smallsmile n,\ m \neq n}} \frac{q_m \text{erf}\left(\|\boldsymbol{r}_n - \boldsymbol{r}_m\|_{\text{p}}\right)}{4\pi\varepsilon_0 \|\boldsymbol{r}_n - \boldsymbol{r}_m\|_{\text{p}}} \tag{1.14}$$

where $m \smallsmile n$ is true when both atoms m and n belong to the same molecule.

TOTAL CHARGE ENERGY

Combining the last three terms, we obtain the overall Coulomb energy:

$$\begin{aligned}
E^{\text{Coulomb}} &= \sum_{\boldsymbol{l}} \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m \not\smallsmile n}}^{n} \frac{q_n q_m}{8\pi\varepsilon_0} \frac{\text{erfc}\left(\alpha \|\boldsymbol{r}_n - \boldsymbol{r}_m + \boldsymbol{l}\|\right)}{\|\boldsymbol{r}_n - \boldsymbol{r}_m + \boldsymbol{l}\|} \\
&+ \frac{1}{2\varepsilon_0 V} \sum_{\boldsymbol{k} \neq \boldsymbol{0}} \sum_{n=1}^{N} \sum_{m=1}^{N} \frac{q_n q_m}{k^2} e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_m-\boldsymbol{r}_n)} \exp\left(\frac{-k^2}{4\alpha^2}\right) \\
&- \sum_{n=1}^{N} \frac{2\alpha q_n}{\sqrt{\pi}} - \frac{1}{2} \sum_{n=1}^{N} \sum_{\substack{m=1 \\ m \smallsmile n,\ m \neq n}}^{N} \frac{q_n q_m}{4\pi\varepsilon_0} \frac{\text{erf}\left(\|\boldsymbol{r}_n - \boldsymbol{r}_m\|_{\text{p}}\right)}{\|\boldsymbol{r}_n - \boldsymbol{r}_m\|_{\text{p}}}
\end{aligned} \tag{1.15}$$

1.3.4 Precomputation

Different algorithms that solve a particular problem generally differ in the specific space-time tradeoff they choose. Obtaining the energy of a configuration presents the same possibility: while entirely computing the energy from scratch requires a set amount of computation, it is possible to make subsequent computations faster at the cost of a greater memory footprint of the algorithm, by storing some intermediate results which can be reused afterwards.

ENERGY GRID

In the case of simulations involving a rigid framework, the most obvious precomputation that can be done is that of the short-term interactions of the guest species with the framework atoms. Indeed, such interactions only depend on the nature and the position of the guest atom, so it is possible to store them on an energy grid.

More precisely, fore each possible guest atom, the unit cell of the framework is divided into a regular grid, with a set spacing between consecutive points. We use 0.15 Å, which is fine enough to be accurate but not too fine to avoid using too much memory. For each grid point, the sum of the interactions of all framework atoms with a guest atom in that position is stored at that point. The grid is then written to disk.

During the simulation, in order to compute the short-term interactions between the framework atoms and a guest atom at a given position, the value is interpolated from the surrounding grid points. Using a simple linear interpolation of the values does not yield very satisfactory results however, since derivative discontinuity would push that atoms towards the grid points, which would result in simulation artifacts. To prevent this, RASPA2 stores not only the value of the energy, but also the three first derivatives, three values of the second derivatives and one of the third, following a scheme by [REF]. Moveover, the grid is forced to be orthorombic: this means that it may not be aligned with the unit cell of the framework, but it does not matter as long as each point in the unit cell can be surrounded by recorded grid points. I reimplemented the same grid algorithm, which ensures a smooth interpolation everywhere in space.

EWALD PRECOMPUTATION

Short-term interactions that can be stored in grids include the short-term direct part of the Ewald summation. The reciprocal term cannot however be simply stored as such.

Yet, it is still possible to precompute a large part of the reciprocal interaction which is due to the framework. Looking back at Equation 1.13, the list of relevant $\boldsymbol{k}$, associated to their factor $\exp\left(-k^2/\left(4\alpha^2\right)\right)/k^2$, is computed once and for all. For each atom n, the factors $\exp\left(\boldsymbol{k}\cdot\boldsymbol{r}_n\right)$ are also computed once. Hence, Equation 1.13 can be rewritten as

$$\Phi^{\text{reciprocal}}(\boldsymbol{r}) = \frac{1}{\varepsilon_0 V} \sum_{\boldsymbol{k}\neq\boldsymbol{0}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \frac{1}{k^2} \exp\left(\frac{-k^2}{4\alpha^2}\right) \underbrace{\left(\sum_{n\text{ framework}} q_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n} + \sum_{n\text{ guest}} q_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n} \right)}_{\text{precomputed part}}$$

$$= \frac{1}{\varepsilon_0 V} \sum_{\boldsymbol{k}\neq\boldsymbol{0}} \kappa_{\boldsymbol{k}} F_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{1}{\varepsilon_0 V} \sum_{\boldsymbol{k}\neq\boldsymbol{0}} \kappa_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{n\text{ guest}} q_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n}$$

$$(1.16)$$

where $\kappa_{\boldsymbol{k}} = \frac{1}{k^2}\exp\left(-k^2/(4\alpha^2)\right)$ and $F_{\boldsymbol{k}} = \sum\limits_{n\text{ framework}} q_n e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n}$ are precomputed terms.

Finally, the entire self-interaction term can be computed once for all, since it does not depend on the configuration. This is only the case here because the simulated moleculses are rigid bodies; for deformable structures, the self-interaction term depends on the actual conformation of the guests, which may depend on the state.

BLOCKING POCKETS

Some MC trial moves attempt to displace an atom to a location that would lead to unphysically high energies – typically, too close to a framework atom. Such positions can be determined once and for all, and the relevant MC moves automatically refused without the need for any energy computation.

To do so, the simplest strategy consists in making a boolean grid, aligned with the energy grids used for short-term interactions, which stores a 1 when the voxel is blocked or 0 if it is allowed. The first step of any MC trial move then consists in checking whether each atom is in an allowed voxel, or refusing the move otherwise. In practise, the simplest implementation consists in making one such boolean grid per atom kind (since the energy threshold depends on the atom), and for each atom, checking the closest grid point of its kind.

In addition to its use as computational optimization, such boolean grids can also be used to mark some parts of the framework as inaccessible, irrespective of the corresponding energy. This is useful when small window size prevents a species – for instance a gas molecule – from entering a pore – for instance the sodalite cages in FAU zeolites. Indeed, without explicitly blocking the pocket, an MC move could make a species appear in the pore, whereas it should be kinetically impossible to enter. This strategy will thus be used when studying the adsorption of gases in zeolites (see **??**), but for the small cations under study here, there is no need to additionally block any part of the frameworks.

RASPA2 only implements blocking for the latter case – *i.e.* explicit pocket removal. It does not use a boolean grid: instead, the blocked space is described as a combination of blocking spheres. For each trial move, the distance between the atoms of the moved guest and the center of each sphere is computed: if one is below its radius, the move is rejected. This approach requires less precomputation and thus less memory than storing the boolean grids, but its cost is much greater, and increases with the number of blocking spheres, preventing it from being used as an optimization like before. It is thus less efficient than the approach I implemented.

OPTIMIZATIONS FOR COMPUTING THE DIFFERENCE OF ENERGIES

A simulation that follows the Metropolis-Hastings algorithm walks across configurations, each pair of contiguous states being separated by a single MC move. In this context, computing the energy of each trial configuration to compare it to the previous one is wasteful: the only part of the energy impacted by the MC move is that which stems from the moved species. Instead of computing the energy of the trial configuration, it is thus better to directly compute the energy difference due to the move.

For the short-term interactions, this is straightforward: the energy difference is simply the sum of the interactions of the species at the new position minus that at its previous position. If there are N species, moving one thus costs $O(N)$ computations for these terms, instead of $O(N^2)$ required to compute this energy the first time. The direct part of the Ewald summation follows the same principle.

For the reciprocal part, Equation 1.16 can be revisited by removing the first constant term, and singling out the atoms n for the moved species:

$$\Delta\Phi^{\text{reciprocal}}(\boldsymbol{r}) = \frac{1}{\varepsilon_0 V} \sum_{\boldsymbol{k}\neq\boldsymbol{0}} \kappa_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \sum_{n \text{ moved}} q_n \left(e^{-i\boldsymbol{k}\cdot\boldsymbol{r}'_n} - e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_n} \right) \tag{1.17}$$

where r' is the position after the move and r before.

In order to allow computing this efficiently, each term $e^{-i k \cdot r_n}$ is thus stored in memory (instead of recomputing it each time). The new values $e^{-i k \cdot r'_n}$ computed for the trial move are also temporarily kept in memory, so they can be used to replace the previous ones if the trial move is accepted. RASPA2 does not implement this optimization, and simply recomputes the $e^{-i k \cdot r_n}$ terms each time, which is slower but requires less memory.

1.4 POSSIBLE ALTERNATIVES

Molecular dynamics (MD) could also be used to place cations in zeolites. Generally speaking, the principle of MD consists in running a simulation where the particles follow Newton's laws of motion, with forces based on force fields like for MC schemes. The simulation is divided in time steps, between which the forces are computed on each particle to derive their position at the next time step. Hence, the trajectories of the particles in MD follow a physical path.

Temperature can be included in different ways in MD, all of which use an element of randomness that eventually impacts the velocities of particles, in order to ensure that the average values extracted from the simulation correspond to those which would be obtained at that temperature. This means that the precedent discussion on the inaccessibility of high-energy regions to the system at room temperature still applies in MD. A simple MD simulation is thus doomed to only explore local energy minima, failing to reach ergodicity like with the simple canonical MC scheme.

To circumvent this issue, it is necessary to bias the simulation in order to force the system to explore regions of high energy. One such strategy consists in doing metadynamics [REF], which is a variation of MD in which the potential energy landscapes is modified during the simulation. The effective free energy of a point in configuration space is computed as the sum of the real free energy of that point without bias, added to the sum of some Gaussian functions centred on previously explored points. Every few steps, another Gaussian function centred on the current configuration point is added to the bias, so that the effective free energy of explored regions increases through the simulation, which allows the system to eventually escape local minima and explore other regions.

A Gaussian is a function of the form $f : x \mapsto A e^{-(x-\mu)^\top \Sigma (x-\mu)}$. In practise, the variable x of the Gaussians used in metadynamics are not directly the 3N coordinates of the N particles, but a small-dimensional vector whose coordinates are called "collective variables". These collective variables are themselves function of the 3N initial coordinates. The parameter A and Σ of the Gaussians being usually fixed, each Gaussian is stored by its centre μ, whose coordinates are thus the collective variables of the current configuration point. Therefore, compared to regular MD, running a metadynamics simulation incurs the additional cost of computing the collective variables x of each visited configuration point, storing the Gaussians' centres, and computing their value at each step.

For cation placement the main issue is to find a small set of relevant collective variables, since the cost of metadynamics grows with their number. We did not explore this venue because we expected it to be similar in difficulty to the MC alternatives, with no clear reason why it should be more efficient. MD in general is also known to converge more slowly than MC, because

large configuration change made possible by MC steps are prohibited by the physics-driven approach of MD, and the same argument could be made of metadynamics compared to other accelerated MC schemes. Exploring metadynamics or its variants could nonetheless be of interest in the search for more alternative, and potentially more efficient, algorithms.

LIST OF PUBLICATIONS

Peer-reviewed papers

Preprint

Publication from a previous project

RÉSUMÉ EN FRANÇAIS

INTRODUCTION

RÉSUMÉ

ABSTRACT