Università degli Studi di Sassari
Dipartimento di Chimica

Dottorato di Ricerca in Scienze Chimiche
Ciclo XXI

Cellular Automata for the Mesoscopic Simulation of Adsorption and Diffusion in Zeolites

Tesi di Dottorato di
Federico Giovanni Pazzona

Il Coordinatore:
Prof. Giampaolo Giacomelli

Il Supervisore
Prof. Pierfranco Demontis

Anno Accademico 2007/2008

Contents

1 Introduction **5**
- 1.1 The problem of confinement 5
- 1.2 Brief history of Cellular Automata 11
- 1.3 The LTA framework 15

2 Lattice structure **19**
- 2.1 Defining the ThPCA statistics 19
- 2.2 Spatial structure of the lattice 22
  - 2.2.1 Cells and real pores: a parallel 25
- 2.3 Local and Global Occupancies 26
- 2.4 Site Thermodynamics 28
- 2.5 Cell Thermodynamics 29

3 Time evolution **31**
- 3.1 Time step 32
- 3.2 Randomization operator 33
  - 3.2.1 Memoryless randomization 33
  - 3.2.2 Jump randomization 35
- 3.3 Local propagation operator 38
  - 3.3.1 Thermodynamic and kinetic barriers 41
- 3.4 Global propagation operator 43
  - 3.4.1 Non-interacting case 43
  - 3.4.2 Interacting case: partitioning scheme 44

4 Equilibrium and transport properties **49**
- 4.1 Probability distribution of occupancies 49
- 4.2 Reduced variance 50
- 4.3 Evaluation of chemical potential 51
  - 4.3.1 Non-interacting case 52
  - 4.3.2 Interacting case 53
- 4.4 Diffusion coefficients 55

4.4.1 Self-diffusion coefficient . . . 55
4.4.2 Collective diffusion coefficient . . . 56
4.4.3 Chemical diffusion coefficient . . . 57
4.5 Other time-dependent properties . . . 58

5 Non-interacting case: numerical simulations 61
5.1 Static equilibrium properties . . . 62
5.1.1 Equilibrium distributions . . . 62
5.1.2 Diffuse phase transition . . . 65
5.2 Transport properties . . . 68
5.2.1 Diffusion coefficients . . . 69
5.2.2 Local density . . . 74

6 Interacting case: Numerical Simulations 77
6.1 Cellular Automata from static models . . . 77
6.2 Numerical simulations to fit reference data . . . 83
6.2.1 Application: xenon in zeolite NaA . . . 84
6.2.2 Application: methane in zeolite ZK4 . . . 89
6.2.3 Application: ethylene in zeolite NaA . . . 90
6.2.4 Application: xenon in zeolite CaA . . . 91

7 Exploring correlation effects 93
7.1 Correlated motion . . . 93
7.2 Correlations in local density fluctuations . . . 99

8 Conclusions 107
8.1 Acknowledgements . . . 109

A Mean field theory 111
A.1 Single cell at equilibrium: grand-canonical formulation . . . 111
A.1.1 Non-interacting case . . . 113
A.2 Mean-field displacement autocorrelation function . . . 115
A.2.1 Basic probability distributions and averages . . . 117
A.2.2 Mean-field evaluation of the displacement autocorrelation function . . . 120
A.2.3 Discussion of the mean-field results. . . . 126
A.3 Local density fluctuations . . . 130

B Theory of the non-interacting case 135
B.1 Computation of the partition functions . . . 135
B.1.1 Limiting Distributions for $T \to \infty$ . . . 137
B.1.2 Limiting Distributions for $T \to 0$ . . . 138

C From a structured to a less-structured cell **141**
C.1 Less-structured cell of level $j = 1$ 146
C.2 Corrected energy function 148
C.3 Less-structured cell of level $j = 2$ 150
C.4 Test of the reduction procedure 151
C.5 Test cell # 1 151
C.6 Test cell # 2 154

D Collective diffusion coefficient **155**
D.1 Collective diffusion coefficient 155

E Sampling schemes and detailed balance **157**
E.1 Mixed Arrhenius-Metropolis acceptance probability for jump randomization 157
E.2 Global detailed balance for randomization 159
E.3 Block propagation: partitioning scheme in details 161
E.4 Global detailed balance for propagation 165

Chapter 1

Introduction

1.1 The problem of confinement

Microporous materials such as zeolites have found many different uses in many different areas of application, like heterogeneous catalysis, separations, molecular sieving,[1] oil recovery, ion-exchange, and other industrial processes. Zeolites consist of AlO_4 and SiO_4 tetrahedra linked together in such a way as to form cages of 3-13 Å diameter, connected in a continuous network of pores extending in one, two, or three dimensions. The commercial interest for this kind of materials is constantly increasing and new fields of utilization have been predicted in the near future.[2] Application of fundamental scientific principles to the key technological issues involved has been difficult, however, and much more progress has been attained through exploitation of empirical processing parameters than through a deep understanding of the chemical and physical mechanisms that control catalytic activity, sorption and diffusion.

Advanced experimental research in this field is often very expensive and time-consuming; moreover the interpretation of the results is not always unambiguous. All above-mentioned applications of zeolites depend on a large number of microscopic processes, among which are adsorption and diffusion. The geometrical restrictions imposed by the confining host at the level of molecular length scale strongly influence the thermodynamic and transport properties of the guest molecules.[3] Not only a full understanding of the role the confined geometry plays on the dynamic behavior of molecules in microporous solids would be of remarkable relevance to the technological application, but also represents a challenge in the range of the still unsolved many-body problems in statistical mechanics. An important subject where researchers in this field have focused their attention is the mobility of molecules adsorbed

on the inner surface of the micropore. The physics of this phenomenon is influenced by the nature of the guest-host interactions[3] which provide the energy landscape for the transport process, and therefore influences the aptitude of a guest molecule to migrate from pore to pore. Increasing the number and type of molecules in the accessible volume deeply modifies this energy landscape destroying or creating effective paths for molecular motion. As a consequence geometric effects as pore size and pore shape along with connectivity and tortuosity should be recognized among the factors playing a major role.

Despite a significant experimental and theoretical research effort, many of the diverse physical phenomena associated with diffusion in tight confinement, such as heterogeneous catalysis, biological transport, percolation, and even a dramatic change of the phase diagram,[4] are still far from being understood.[5] This motivates the search for a further simplification of the atomistic models of micropores, to reach a simpler description of the effective interactions such that they can be easily computed and at the same time able to capture the essential features of the real physical systems.[6–9] In essence, the diffusivity of a guest molecule will depend on the size and the shape of the channels and cages of the internal pores, related to the size and shape of the molecule itself. In addition, the number of the adsorbed molecules in each of the channels and cages will deeply modify diffusion by both chemical and steric effects.

Molecular simulations. Many properties of fluids in porous media become inaccessible to experimental measurement when the pore size approaches molecular dimensions.[10] Thus, molecular simulations commonly take the role of experiment in testing approximate theoretical models.

Recent developments in algorithms combined to the use of increasingly faster computers are extending the realm of applicability of first principles atomistic calculations for predicting performance and properties of microporous materials.[11] The starting point for any model at molecular level is a proper description of the microstructure of the zeolite that is generally provided by experimental crystal structure determination.[12,13] Then this collection of atoms and ions (fixed in space or moving according to a suitable interaction potential) is linked to a description of the intermolecular forces through which adsorbed molecules interact with the inner surface and with each other. This is the conventional approach based on empirical potentials.

So far Monte Carlo (MC)[14] and Molecular Dynamics (MD)[15] techniques are the most widely used tools in the molecular modeling of adsorption and

diffusion in zeolites. MC simulations are used mainly to study the thermodynamics of molecules adsorbed in zeolites. Predictions of adsorption isotherms and heats of adsorption are generally in good agreement with experimental results. The simulations also provide interesting details about the molecular-level arrangement of the molecules within the zeolite pores. This information has proven very useful in explaining the complex behavior of molecules in tight confinement.[16]

When studying the problem of molecular confinement with a simulation approach, a proper representation of dynamic events has the same importance as a proper sampling of equilibrium properties, since in this context adsorption and diffusion are so tightly bounded that it's impossible to study the two problems separately. Anyway, while setting up a molecular simulation one has to face the problem of selecting which method is more appropriate to obtain meaningful data at the timescale of interest. Actually, in a molecular investigation of adsorption/diffusion in microporous material the range covered by the involved timescales is somewhat wide, going from bond vibrations at the femtosecond level to long-range diffusion at the microsecond or longer timescale depending on the specific conditions at which the phenomenon occurs.[17] Theoretically, phenomena with the fastest motions are studied by different methods, including first principles molecular dynamics[18] and classical molecular dynamics simulations.[19] These studies provide insight into the atomistic structure and interactions with the adsorbed molecules of the microporous framework. Other events at slower timescales, such as ionic exchange or long-range diffusion are rarely suitable for studies with current classical Molecular Dynamics simulations (MD) as they span scales that are at least of an order of magnitude larger than the largest studies to date (hundreds of nanoseconds).[20]

The quest for an efficient reductionistic approach. In principle, MD simulations could investigate any classical problem at a molecular level, but the wide range of diffusional time scales encountered by molecules in zeolites limits the applicability of MD to processes that occur on the order of femto- to nano-seconds.[17] Even using models based on simple effective pairwise atom-atom potentials, investigations of these phenomena are still computationally very costly. In some cases they can turn out to be quite impossible: if the sorption sites of the zeolite are separated by large free energy barriers compared to thermal energies the resulting slowdown of molecular motion could be so marked as to make not amenable to direct MD simulation using present computer technology. Then, Transition State Theory (TST) and related methods must be used to simulate the temperature dependence of

site-to-site jump rate constants.

Because most of the practical applications of these materials involve shape selective catalysis and separation processes, a field where the transport properties of adsorbed molecules play a central role, a question arises: what are the fundamental interactions that control the dependence of diffusion in microporous materials and can they be represented in a coarse-grained fashion able to reproduce the main features of transport phenomena on very long timescale at very long distances? Answering this question means to find out the essence of adsorption/diffusion phenomena.

In this context, properly formulated thermodynamic models (ThM) with very local interactions has been proved to be very useful in the description of equilibrium thermodynamic properties of adsorption systems (see Cheung[21] and Ayappa[22]), while providing no informations about dynamics of transfer processes associated to adsorption. In some models, the thermodynamics of a lattice is represented through the properties of a single, open cell treated as a small grand-canonical system and able to host a limited number of guest molecules. The approaches to obtain the cellular statistical properties range from the development of a grand-canonical partition function formalism[21–23] to the formulation of conditional probability distribution functions of hard spheres.[24,25] In both cases, interactions between the adsorbed molecules are represented by a mean field potential, which can be a fixed mean interaction,[24,25] or a function of the *occupancy* of the single cell[21,23] (the number of molecules adsorbed which is a discrete, *local* observable) or a function of the *loading*[22] (average occupancy, a continuous *global* observable). The strength of such models stands in their very simple way to take account of molecular interactions by introducing *local* potentials in a simple parametric form. Some models can be used to fit adsorption isotherms,[22,23] some others to fit occupancy distribution functions.[21,24,25] The main drawback is that they contain no kinetic information.

Dynamic long-time properties can be acquired through Kinetic Monte Carlo (KMC) models, where the crystal lattice is replaced with a three-dimensional lattice of binding sites, and migration rates from site to site are modeled in order to reproduce the loading dependence of activated diffusion in zeolites.[7,26–31] When improved with the introduction of the correct jump rates and transmission factors obtained from MD simulations, such models can be coupled with TST and then become quantitatively able to reproduce the self-diffusivity profiles of adsorbed species in zeolites.[32,33] However, KMC models are not aimed at the accurate reproduction of equilibrium properties.

Anyway, a properly weighted coupling of static thermodynamic models and lattice kinetic models together can be thought to provide the best simultaneous representation of adsorption and diffusion properties.

Cellular Automata for confined systems. By applying methods like ThM or KMC the main question becomes: can a simple, space-time discrete model properly describe at the macroscopic scale the features of diffusion in microporous materials such as zeolites? The answer should be yes if the parameters governing the thermodynamic and kinetic behaviors are properly set. This is because a lattice model works with a drastically reduced number of degrees of freedom, in order to cover large scales on space and time. The main task is then to find how to transfer the essential physical features of molecules adsorbed inside micropores (as they are obtained from atomistic simulations or from experiments) to the parameters defining the local properties of each cage or channel, and to use such a coarse-grained description as a tool to perform large-scale simulations of the zeolitic material of interest.

Cellular Automata (CA)[34–36] are discrete dynamical systems where space is represented by a uniform grid with each cell containing a few bits of data, and local and homogeneous laws allow the system to evolve in discrete time steps of fixed duration. Due to their discreteness and the absence of truncations or approximations in their dynamics, CA evolve without being affected by any numerical instability. Moreover, on the analogy of MD, a CA model is an implementation of a N-body system where all correlations are taken into account, fluctuations arise spontaneously, and due to time homogeneity CA observables can be monitored in the same way as in MD simulations. However it is important to note that a proper care of correlations must be taken into account because in discrete models they may be stronger than in the continuous counterpart.[37] CA are widely used as a modeling tool in natural science, combinatorial mathematics, computer science. With their simple structure, they represent a natural way of studying the evolution of large physical systems under both equilibrium and non-equilibrium conditions.[38,39]

Construction of a reductionistic model. The structure itself of a zeolite inspires a coarse-grained description of its internal void space:[40] the channels and cages connected to one another define a system of communicating cells that can contain and exchange a limited number of guest molecules. Such a reductionistic representation can be used as the starting point to set up a mesoscopic simulation. The next step is to identify the relevant timescales. As a first approximation, all involved timescales can be grouped into two main regimes: intra-cell motion at short times, and inter-cell migration at long times.[41] Intra-cell motion is a source of complexity in the guest motion, since it causes the trajectory of the diffusant to be affected by correlations which, in turn, propagate in time thus affecting the long-time diffusivity.

In fact, although one could be solely interested in modeling the long-range diffusion, the reductionistic model would be complete only if correlations on the intracell scale are properly taken into account. Two routes are proposed: in the first one, effective transmission coefficients taking account of the backscattering effect (in which after migration, the guest molecule returns back to its departure location) are introduced as corrections to the diffusion coefficient to reproduce the reduced long-range mobility due to correlations[32] (simulations in Section 6 will be discussed where this route has been followed to model dynamic properties). In the second one, correlations due to the intracell motion can be mimicked through a sequence of correlated jumps of the guests inside of each cell (simulations performed following this strategy will be discussed in Section 7). The goal of the present thesis has been to develop a possible framework for such a coarse-grained approach.

The heart of the model. The basic ingredient in our model dynamics, allowing it to effectively capture the essential phenomenology of adsorption and diffusion of simple molecules in zeolites, is a collective migration mechanism in which the guest molecules in the same pore compete to gain access to the adsorption sites. Such a competition is ruled essentially by a differentiation between the adsorption strengths of *inner sites* (from which is not possible for a guest to jump outside of the host cell) and *exit sites* (allowing intercell transfers), representing the guest-host interactions which play a major role in problems of diffusion in tight confinement. A differentiation among several types of adsorption sites, mutually differing in adsorption energy, has been shown to be a feature of most nanoporous materials,[31,42–44] giving rise to a strong dependence of both equilibrium and transport properties on loading and temperature.[32,45–47] For example, experimental evidences[48] have shown that in NaY zeolite benzene molecules appear to occupy two types of sites, cationic sites on the walls of the supercage and less favorable window sites in apertures between adjacent supercages, a behavior consistent with results from molecular dynamics simulations.[49]

The actual dynamic state of each cell (including its tendency to molecule transfer) will be determined by the energies of its occupied sites, according to a local MC sampling scheme of the allowed configuration space. In this way one can introduce an elementary time-scale larger than the time-scale of the slowest processes occurring in the cell, and the equilibrium properties will depend on the value of the few parameters that completely define the thermodynamics of the model.

1.2 Brief history of Cellular Automata

A Cellular Automaton is an idealization of a physical system in which space and time are discrete, and the physical quantities take only a finite set of values.

Arising from an idea of John von Neumann (who, at the end of 1940s, was involved in the design of the first digital computer), the concept of CA constitutes the first applicable model of massively parallel computation. The original idea of its inventor was actually the construction of a machine able to imitate the behavior of the human brain, in order to solve very complex problems. In particular, von Neumann was attempting to contruct a system able to be self-controlling, self-repairing, and self-replicating — in other words, a form of artificial life.

The first abstraction of such an idea was a fully discrete system made up of cells, each one characterized by an internal state represented by a finite number of information bits and evolving in discrete time step, like a simple automata which needs only a simple recipe to update its state from time to time. The updating rule was a function of the states of the neighboring cells (that is, it was *local*) and was the same for all cells (i.e. it is also *homogeneous*). On the analogy with biological systems, all cells were updated simultaneously, i.e. they were subjected to the same clock.

The research von Neumann carried in that direction produced a discrete structure of cells bearing in themselves the recipe to generate identical individuals by means of a complex rule (know as von Neumann's rule) which has the so-called property of *universal computation*, i.e. the ability of the CA to simulate any computer circuit.

Besides the use of CA formalism to study artificial life, CA have been brought to the wide audience by the John Conway's *Game of Life*,[50] an ecological model where a very simple evolution rule generates an unexpectedly complex behavior, and a large number of self-organizing structures emerge and move across space.

As they are discrete models, CA exhibit many of the behaviors encountered in continuous systems but in a much simpler framework. For that reason they have been widely used as a tool to study the basic principles of statistical mechanics. A remarkable (and also the first) example of such an use of CA is given in the Stephen Wolfram's study of self-organization phenomena in statistical mechanics.[51] In his work and in its subsequent research activity (which led in 2002 to publication of his controversal book *A New Kind of Science*[52]), Wolfram proposed and analyzed a wide variety of CA rules generating complex behaviors in spite of their simplicity, with the

aim of demonstrating that discrete systems can be used to find abstraction of complex structures and behaviors in nature.

One of the most remarkable properties of many CA rules (like the von Neumann's rule or the Conway's Game of Life) is the already mentioned *universal computation*, according to which there exists an initial configuration of the cellular automaton which leads to the solution of any computer algorithm — in other words, anything that can be elaborated with algorithms can be implemented in the context of such rules. That stimulated many researchers to the development of artificial universes, and with this purpose CA machines were successfully designed (by T. Toffoli and N. Margolus[35]) to easily simulate the time evolution of such models.

Lattice-Gas Cellular Automata. The use of CA to model real systems under a coarse-grained approach finds its justification in the fact that often systems with a large number of degrees of freedom exhibit macroscopic behavior where the details of the microscopic dynamics are relatively unimportant. In fact the flow of a gas, a fluid or even a granular material are very similar at a macroscopic scale: this suggests that systems with different microscopic characteristics can produce similar macroscopic behaviors, and that a properly designed CA *microdynamics* (such term indicates the discrete-time evolution of the observables produced in each CA cell) can reproduce a given macroscopic behavior in spite of its fully discrete nature. Of course a discrete microdynamics cannot be realistic in the most strict sense, but the observables it produces become statistically meaningful once they are averaged over a sufficiently long observation time.

From that idea many automata were constructed which are made up of particles (representing the guests) moving across nodes of a regular lattice. CA simulations of particle-based models are often referred to as Lattice-Gas Cellular Automata (LGCA), to distinguish them from traditional CA as meant by, e.g., von Neumann, Conway and Wolfram — in fact, such a distinction regards only their underlying philosophy, since from a mathematical point of view LGCA *are* cellular automata. A deep analysis of such models is presented in the J.-P. Rivet and J. P. Boon's book, *Lattice Gas Hydrodynamics*.

The first example of LGCA is the lattice-gas model developed in the 1970s by Hardy, Pomeau and de Pazzis (known as the HPP model) to study the fundamental statistical properties of a gas of interacting particles via a simple and fully discrete dynamics of colliding, momentum-conserving particles. The discovering (in the 1980s) that the HPP model is in fact a cellular automaton pushed researchers towards the use of CA rules to study hydro-

dynamic problems. The model proposed in 1986 by U. Frisch, B. Hasslacher and Y. Pomeau (the FHP model[53]) is an example of a local discrete system exhibiting a macroscopic behavior which (within some appropriate limits) obeys the Navier-Stokes equation of hydrodynamics. In the FHP model particles diffuse in a two-dimensional triangular lattice, where each node is surrounded by six neighboring nodes and is build up of six channels: each channel is able to host at most one particle, and connects the node to one of the six neighbors thus imparting the guest particle a *discrete velocity*. At each time step, the evolution algorithm is made up of two operations: (1) a *collision* event where particles occupying the same node change their respective velocity according to some local algorithm, and (2) a *streaming* event where particles move from node to node in the lattice.

Although many improvements to the original model have been implemented during the years, the discrete CA approach did not surpass the traditional numerical methods of hydrodynamics, whereas it has been much more successful in the modeling of complex situations for which traditional computing tecniques are not applicable, like flows in porous media, immiscible flows ad instabilities, spreading of liquid droplets, wetting phenomena, transport, fluid dynamics, pattern formation, reaction-diffusion processes, nucleation-growth phenomena. An exhaustive survey of CA models as abstraction of physical systems can be found in the B. Chopard and M. Droz's book *Cellular Automata Modelling of Physical Systems*.[36]

Cellular Automata for adsorption and diffusion in zeolites. As in zeolites the space is partitioned in connected channels and cages, the analogy with the discrete nature of the CA space is straightforward. The statistical agreement obtained with a real host-guest system on a certain scale, after a proper setting of the CA constitutive parameters, often can be achieved on several different scales. For instance, when considering the problem of adsorption/diffusion of simple adsorbates in zeolites, connection with the largest space-time scales is established once agreement of adsorption isotherm and diffusivity profile has been obtained. By using a hierarchical approach, the statistical CA realism can be further improved by attaining agreement between experimental and emulated stationary density distributions, while (in addition) agreement of local rates of transfer processes allows emulation to be efficient on even smaller scales. As in any other coarse-graining framework, the smaller are the scales one wants to recover, the more are the hierarchical levels to be introduced in the CA structure.

In the present thesis a microdynamics will be associated to a thermodynamic model to build up a Thermodynamic Partitioning Cellular Automaton

(ThPCA)[45,54–56] able to model in the same simulation both equilibrium and transport properties of a system of weakly interacting particles moving inside of a regular network of cells, by incorporating both the easy-to-tune cellular statistics as made available by thermodynamic lattice models, and the correct long-time kinetic behavior obtainable by MD and KMC.

In the definition of ThPCA:

- the word *partitioning* denotes an automated partitioning scheme according to which the system evolves by alternating independent sublattices; this allows an interacting probabilistic model to reach thermodynamic equilibrium while evolving with a high degree of synchronicity without introducing any extra-correlation between adjacent cells.

- the word *thermodynamic* emphasizes that the configuration space of each cell is given by a purely thermodynamic model.

With its simple and flexible parametric structure, the ThPCA approach will be exploitable in the wide range of problems connected to molecular adsorption, diffusion and related processes in microporous materials. The main input for the present model may come from data like: (i) local density dependent mean-field potentials and transition probabilities obtained from atomistic simulations that could be used as the starting point to derive adsorption and diffusion properties. (ii) Thermodynamic and kinetic data obtained from experiments and/or other simulation methods, used as the starting point of a fit procedure aimed to trace which local behaviors cause the system to produce them, or aimed to further enlarge the scales. Tecniques will be extensively illustrated which allow to efficiently manipulate local parameters in order to make the ThPCA able to fit very well given adsorption and diffusion data. Besides a parametric study of the model aimed to fully understand the response of the automaton to different input parameters, numerical results shall be presented for real systems. After a digression in Section 1.3 about the particular zeolite topology investigated here, in Chapter 2 the spatial structure and the basic thermodynamics of the model shall be presented. In Chapter 3 the evolution rule will be illustrated. Two possible schemes will be presented: one producing an uncorrelated self-motion of the guests, and another one (more realistic) producing a highly correlated motion. Next, in Chapter 4 the basic macroscopic quantities (both equilibrium and transport properties) which can be evaluated directly from the output of one ThPCA simulation will be discussed. Chapters 5 and 6 will be devoted to the discussion of numerical ThPCA simulations of systems of both non-interacting and interacting guests in the absence of correlations, while the effect of correlations introduced by the jump randomization procedure will be extensively

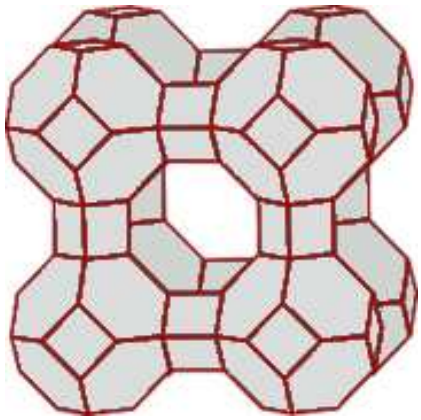

Figure 1.1: Three-dimensional schetch of one α-cage.

illustrated in Section 7.

In Appendix A a mean-field formulation of the model will be proposed and discussed: this will provide a better comprehension of the nature of the model. Several times in the main body of the thesis formulas shall be used which will be derived within the illustration of the mean-field theory.

Next, in Appendix C a strategy to transfer the essential information of a structured cell into a less-structured one is proposed.

Finally, proofs of detailed balance, sampling schemes and other mathematical details will be found in the Appendixes D and E.

1.3 The LTA framework

In the present thesis a CA formalism will be developed to emulate adsorption and diffusion of several (simple) chemical species inside a grid of cells mimicking the topology of *Linde Type A* (LTA) zeolites,[12] where nearly spherical cavities are arranged as a cubic lattice. Sizes of all chosen species are large enough to be able to diffuse only via the α-cages (see Figure 1.1), each having an internal cavity radius about 5.7 Å being connected to six neighboring cavities by six eight-ring (i.e. eight oxygens arranged in a plane) nearly circular windows of ~ 4.2 Å in diameter. Such an ordered matrix can trap clusters of molecules in its void space. Smaller cavities (called β-cages) are also present in the sodalite units, which can be entered only through a six-ring window which is far too small to be entered by the selected guests.

One of the most studied LTA zeolites (both experimentally and by numerical simulations) is the zeolite *A*, with a framework Si/Al ratio of 1.0, and a simple enough structure for all ions to be locatable. NMR studies are present in the literature[10] which have been performed by detecting small

xenon clusters trapped in zeolite NaA, where 12 Na^+ ions are present in each pseudounit cell (which has formula $Al_{12}Si_{12}O_{48}^{12-}$) and partly block the windows, and in zeolite CaA,[57] where the number of Ca^{2+} counterions is 6 and no windows are blocked. Such experiments allowed to study how the effect of confinement causes the distribution of guests inside the cavities to deviate from the purely statistical distribution (the hypergeometric distribution) derived by considering uniquely excluded volumes.[24,25,58]

The ZK4 is a cation-free zeolite extensively used as host in MD simulations of diffusion of methane under confinement.[15,41,59–66] As a result of the absence of extra-framework cations, all windows are free for diffusion.

A convenient choice for the construction of a discrete model. There are several reasons why such a zeolite framework has been chosen to be mapped onto a CA. Besides the fact that the cubic symmetry simplifies the computational handling, the very closed LTA structure causes the confined molecules to spend a relatively long time inside of the cage, so that even a model which totally neglects correlations in motion can capture the qualitative adsorption/diffusion behavior of the guest species (this will be the topic of Sections 5 and 6).

Secondly, a free-energy barrier centered at the windows[61] allows an ideal partitioning of an LTA unit cell into eight α-cages sharing no void space: in terms of a discrete representation of such a framework, each cage can be assumed to be reduced to a finite set of K adsorption sites with no sites in common with neighboring cages, where K corresponds to the saturation capacity of the zeolite under consideration. In the language of CA, such a cage becomes a CA cell. Its occupancy is exactly determined by counting how many of its sites are occupied, and the sum of the occupancies of all cells gives the total number of guests in the system.

Moreover, adsorption and diffusion of guest species in LTA zeolites is of interest since the loading dependence of the diffusivity in such systems is somewhat controversal. While a direct connection between the shape of the inverse thermodynamic factor (which will be discussed in Section 4) and the diffusion profile *vs.* concentration can be assessed for a variety of guests adsorbed in more open zeolite structures[67] (e.g. faujasite-type zeolites), the same is not true for LTA where a knowledge of the inner-cage free-energy profile at various loadings[32,68,69] is required in order to correctly interpret the loading dependence of diffusion.

The model proposed in the present thesis provides a flexible, fully reliable computational framework that can be used to performing coarse-grained simulations of adsorption and diffusion in LTA zeolites. Due to its paramet-

ric structure and hierarchic evolutionary scheme, it can be used as a tool to interpret data obtained from experiments or molecular simulations. It rests on a minimal representation of the competition mechanism the guests are involved in while being adsorbed and accessing the windows connecting the host cavity to the neighboring ones. Although presented as a mapping of an LTA zeolite on a discrete parameter space, the computational scheme developed here easily extends to other spatial arrangements, provided an appropriate representation of the topology is given.

Chapter 2

Lattice structure

2.1 Defining the ThPCA statistics

In this section the most essential properties of the configuration space sampled by the ThPCA model space shall be illustrated.

CA are an alternative route to deal with the *emulation* of a physical system under a coarse-grained approach. Due to its discrete nature, its state space contains a large but *finite* number of configurations. This means that, at least in principle, its global partition function pcan be calculated. Of course this is a feature of discrete models in general, not an exclusive of CA. Nevertheless, a CA can be preferred because of the *local* character of interactions, which allows to factorize the *global state space* of the system into easier to handle *local state spaces*. To clarify this point, the partition function of a generic lattice will be compared with the partition function of a CA with very local interactions (that is, with interactions restricted to each single cell as in a purely thermodynamic cell model[21,22]).

Partition Function of a generic lattice. First of all, the system is supposed to be placed in a large heat bath at temperature T, so that the diffusing guests (which are represented as equivalent and structureless particles, having constant kinetic energy at constant temperature) perform a random walk in the canonical ensemble. The canonical partition function of a system of N guests distributed in a lattice $\mathscr{A}$ of ν structured sites can be written as[70]

$$\mathcal{Q}^{\mathscr{A}}(N,\nu,T) = \sum_{\boldsymbol{\eta}^{\mathscr{A}}}^{(N)} e^{-\beta\mathcal{E}(\boldsymbol{\eta}^{\mathscr{A}})}, \tag{2.1}$$

where $\beta = (k_B T)^{-1}$ (with k_B as the Boltzmann's constant), $\boldsymbol{\eta}^{\mathscr{A}} = \{\eta_1^{\mathscr{A}}, \ldots, \eta_\nu^{\mathscr{A}}\}$ is a configuration of guests in the sites ($\eta_i^{\mathscr{A}}$ is assumed to be 0 or 1 in case

the i-th site to be empty or occupied by one particle, respectively) having total energy $\mathcal{E}(\boldsymbol{\eta}^{\mathscr{A}})$. In Eq. (2.1) the compact notation has been introduced

$$\sum_{\boldsymbol{\eta}^{\mathscr{A}}}^{(N)} = \sum_{\eta_1^{\mathscr{A}},\ldots,\eta_\nu^{\mathscr{A}}=0,1} \delta_{\mathrm{kr}}\left(\eta_1^{\mathscr{A}} + \cdots + \eta_\nu^{\mathscr{A}}, N\right), \tag{2.2}$$

where δ_{kr} is a Krönecker delta restricting the sum to the configurations preserving the total occupancy N. Such a notation shall be used throughout the entire thesis for restricted sums.

ThPCA Partition Function. Now let us suppose to partition the lattice $\mathscr{A}$ of ν sites into a new lattice $\mathscr{L}$ of M connected, equal groups (named *cells*) of K sites, so that $MK = \nu$. Cell are assumed to be distinguished one from each other through the coordinates $\mathbf{r}$ of their respective center. The state space remains the same, but now each configuration can be expressed as

$$\boldsymbol{\eta}^{\mathscr{L}} = \bigcup_{\mathbf{r}\in\mathscr{L}} \boldsymbol{\eta}(\mathbf{r}), \tag{2.3}$$

with $\boldsymbol{\eta}(\mathbf{r}) = \{\eta_1, \ldots, \eta_K\}(\mathbf{r})$. Following the typical approach of thermodynamic models[21,22] $\mathcal{E}$ is replaced with the local energy function $E(\boldsymbol{\eta})$ applying to each *cell* configuration, so that the energy function of the total lattice $\mathscr{L}$ is given by the sum of the energy functions calculated on the single cells:

$$\mathcal{E}(\boldsymbol{\eta}^{\mathscr{L}}) = \sum_{\mathbf{r}\in\mathscr{L}} E\left[\boldsymbol{\eta}(\mathbf{r})\right]. \tag{2.4}$$

Due to the independence of the energy function from cell to cell, the statistics of the lattice $\mathscr{L}$ can be computed over the cells rather than the sites. Therefore it becomes useful to introduce the *occupancy distribution* $\boldsymbol{a} = \{a_0, a_1, \ldots, a_K\}$ relative to each lattice configuration $\boldsymbol{\eta}^{\mathscr{L}}$, where a_0 is the number of cells containing no guests, a_1 is the number of cells containing 1 guest, and so on. The occupancy distributions can be used to formulate the lattice partition function as

$$\mathcal{Q}^{\mathscr{L}}(N, \nu, T) = M! \sum_{\boldsymbol{a}} \prod_{n=0}^{K} \frac{[Q(n, K, T)]^{a_n}}{a_n!}, \tag{2.5}$$

where the sum extends over all the possible occupacy distributions, and a central role is played by the partition function of a closed n-occupied cell:

$$Q(n, K, T) = \sum_{\boldsymbol{\eta}}^{(n)} e^{-\beta E(\boldsymbol{\eta})}. \tag{2.6}$$

which under such a form does not include the degrees of freedom of each guest molecule in each site — in Section 2.4 it will be shown that inclusion of them causes the energy function $E(\boldsymbol{\eta})$ to be replaced by the *free energy* function $F(\boldsymbol{\eta}; T)$.

Two important features of $Q(n, K, T)$ are:

i) Its *hierarchical structure*: each cell of the lattice contains a local lattice of *sites* with its own (local) state space. In the spirit of coarse-grained models, the level of detail in the specification of the local lattice structure may not necessarily be high.

ii) Due to the absence of guest-guest interactions residing in different cells, the local partition function $Q(n, K, T)$ becomes the heart of the model. It contains all the statistical informations about the sites it contains, and its evaluation is much easier than the partition function of the global system. The importance of $Q(n, K, T)$ in the determination of the properties of the whole system becomes more clear in the limit of $M \to \infty$, where the global partition function $\mathcal{Q}^{\mathscr{L}}$ depends only on the ratio $\langle n \rangle = N/M$ (which is the average occupancy, also called the *loading*) and on the temperature, and the equilibrium occupancy distribution $\boldsymbol{a}^{\mathrm{eq}}(\langle n \rangle, T)$ becomes largely the most important contribution among all the possible distributions in the sum of Eq. (2.5), giving

$$\frac{\ln \mathcal{Q}^{\mathscr{L}}(\langle n \rangle, T)}{M} \approx \sum_{n=0}^{K} p(n; \langle n \rangle, T) \ln \frac{Q(n, K, T)}{p(n; \langle n \rangle, T)}, \tag{2.7}$$

where $p(n; \langle n \rangle, T) = a_n^{\mathrm{eq}}(\langle n \rangle, T)/M$ is the equilibrium probability of a cell to have occupancy n when the loading is $\langle n \rangle$ and the temperature is T.

As can be seen, the thermodynamic approach largely reduces the complexity of the partition function in Eq. (2.1).

The strict locality of interactions allows a certain degree of synchronicity in the update of the cells. The ThPCA evolves in time by means of an evolution rule consisting of a combination of two kinds of operations: (i) those involving no guest exchanges between neighboring cells, and (ii) those which do involve such transfers. Operations of the first kind (i) can be performed on all cells simultaneously, while operations of the second kind (ii) can be performed simultaneously on independent group of cells called *blocks* in which the lattice can be partitioned.

The modeling of $Q(n, K, T)$ allows a full control of the structural equilibrium properties of the system. Such a modeling can be used, as example,

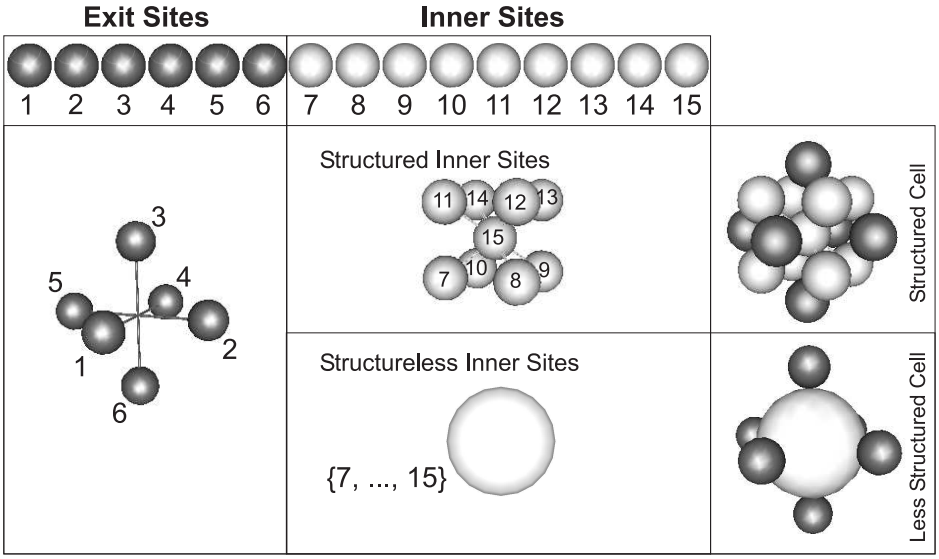

Figure 2.1: Different representations of a three-dimensional cell. On top a 'string representation' is sketched where the sites are identified through an index $j = 1, \ldots, K = 15$. The exit sites (the first 6 values of j) are arranged as the sides of a cube. The inner sites (the remaining values of j) could have a well defined structure (and then the internal structure of the cell would be completely defined) or not (in this case the cell would be only *partially structured*).

to coarse-grain a more complex system so as to gain access to wider space and time scales. Moreover, due to the independence of the energy function from cell to cell it is possible to calculate *on the fly* during simulations in the canonical ensemble thermodynamic properties (such as the chemical potential) which in general need separate simulations in the grand-canonical ensemble to be evaluated. This will be the topic of Section 4.3.2.

2.2 Spatial structure of the lattice

The system is constituted by a periodic (hyper)cubic lattice $\mathscr{L}$ of $M = L^d$ cells (where d is the dimensionality of the system), where N guest molecules can diffuse at constant temperature T. The ratio $\langle n \rangle = N/M$ defines the *loading* of the system. For the sake of simplicity, the model shall be illustrated in three-dimensions, $d = 3$. Extension to higher (or lower) dimensions is straightforward.

The discrete coordinates of each cell shall be denoted as $\mathbf{r}$, the lattice spacing as λ (equal to the distance between the centres of two neighboring zeolite cages), and the basis set of orthogonal unit vectors of the lattice as $\mathbf{e}_j, \ldots, \mathbf{e}_d$, satisfying $\mathbf{e}_{j+d} = -\mathbf{e}_j$, $j = 1, \ldots, d$. Each cell $\mathbf{r} \in \mathscr{L}$ communi-

cates with its $2d$ first-neighboring cells. A superscript j on the coordinates of a cell will indicate its first-neighbor along the direction of $\mathbf{e}_j$:

$$\mathbf{r}^j = \mathbf{r} + \lambda \mathbf{e}_j, \qquad j = 1, \ldots, 2d \tag{2.8}$$

while the subscript will indicate its j-th component:

$$r_j = \mathbf{r} \cdot \mathbf{e}_j, \qquad j = 1, \ldots, d \tag{2.9}$$

For each $j = 1, \ldots, d$ the component r_j can take a discrete value of the form $a\lambda$, with $a \in [0, L-1] \in \mathbb{N}$. The system evolves in discrete time steps of duration τ (details shall be given in Section 3.1).

Exit sites and inner sites. Each cell is made up of a sublattice of K sites. An exclusion principle applies such that each site can be either empty or occupied by only one guest at the same time, therefore the sum K defines the *maximum occupancy* of the cell. The first K_{ex} sites are *exit sites* (allowing intercell transfers of guests) and the next $K_{\text{in}} = K - K_{\text{ex}}$ are *inner sites* (not allowing such transfers). From now on, the notation $(\mathbf{r}, j)$ shall indicate the j-th site of the cell $\mathbf{r}$ (with $j = 1, \ldots, K$ and $\mathbf{r} \in \mathscr{L}$).

In principle the internal structure of the cell could be completely specified by defining exactly a spatial arrangement of all its K sites. Using this structured-cell approach, every site-to-site particle jump would represent one *migration event* and contribute to the short-time self-diffusivity. As shortcoming, the resulting detailed model would be very specific to the particular reference system under emulation. That is, the sites positions should correspond exactly to the positions of potential energy minima in the real pore, so that a complete re-parametrization of the entire cell structure would be needed in order to switch the modeling to different types of pores/adsorbates. When the arrangement of the inner sites is not specified the cell becomes less-structured. In Figure 2.1 a comparison between a structured and a less-structured cell is shown. The structured cell pictured in Figure 2.1 mimics the arrangement of the potential minima found in an α-cage of a ZK4 zeolite[63] when loaded with a number of methane molecules from 1 up to 15. Such an idealized adsorption unit is characterized by: (i) Inner sites arranged as a central bcc cell, and (ii) Exit sites arranged as the centers of the 6 bcc cells neighboring to the central one.

In the present thesis the CA modeling of the reference host-guest systems will be carried out by using the less-structured cell of Figure 2.1. In Appendix C it is shown how the essential thermodynamic properties of such a structured cell can be transferred to a less-structured cell where the topology of inner sites is unimportant.

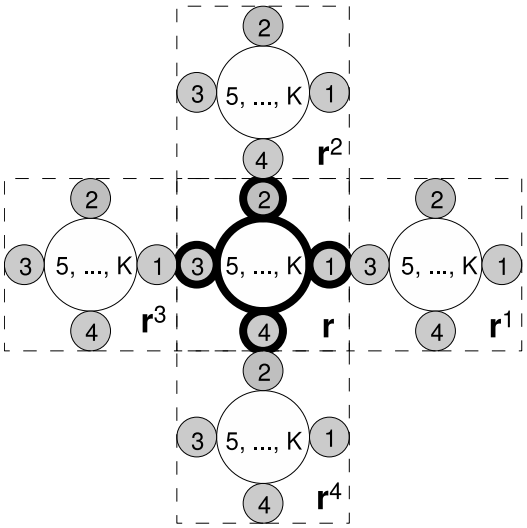

Figure 2.2: Structure of connections in a portion of a two-dimensional system. Each cell is sketched as a white circle (representing the set of inner sites) and four gray circles (the exit sites). Numbers are used to identify the sites inside of each cell. Coordinates of neighboring cells of cell $\mathbf{r}$ are expressed according to the notation introduced in Eq. (2.8).

Connections between neighboring cells. Let us now describe the structure of connections between adjacent cells.

(i) In each cell, say $\mathbf{r}$, the j-th exit site is associated to the particular direction of motion $\mathbf{e}_j$. That is, a guest in the site $(\mathbf{r}, j)$ is allowed to access the window setting the connection between the cell $\mathbf{r}$ and the cell $\mathbf{r}^j$. In case of successful jumps, the guest will travel the distance $\lambda\mathbf{e}_j$ to move from $\mathbf{r}$ to $\mathbf{r}^j$. In the left part of Figure 2.1 the exit sites of a three-dimensional cell are shown. It follows that $K_{\text{ex}} = 2d$.

(ii) As a consequence of (i), for each direction $j = 1, \ldots, d$ the exit site $(\mathbf{r}, j)$ is the only site of $\mathbf{r}$ pointing towards $\mathbf{r}^j$, and the exit site $(\mathbf{r}^j, j + d)$ is the only site of $\mathbf{r}^j$ pointing on $\mathbf{r}$. Therefore, a jump from the exit site $(\mathbf{r}, j)$ to the exit site $(\mathbf{r}^j, j + d)$ is the only way a guest has to migrate from cell $\mathbf{r}$ to cell $\mathbf{r}^j$. The structure of connections is made clear in Figure 2.2 for a two-dimensional system.

(iii) Statements in (i) and (ii) together with the exclusion principle represent a *constraint on particle traffic*: Two adjacent cells can exchange at most one guest particle at each time step.

(iv) The exit sites *topology* defines the connections among cells, therefore it must be specified apart from the level of detail in the cell structure. In

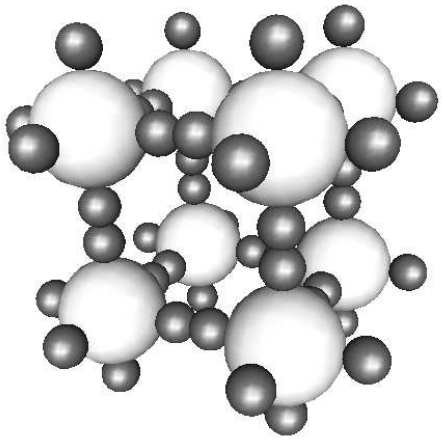

Figure 2.3: A small portion of a three-dimensional lattice of cubically connected cells.

each cell, the inner sites arrangement may be specified or not, depending on the required level of detail. In the next Section some important consequences of this assumption will be discussed.

2.2.1 Cells and real pores: a parallel

Besides the assumption in Eq. (2.4) of restricting *inside of each cell* the guest-guest interactions, along with the coarse-grained spirit of the model the inner sites arrangement is intentionally not fully specified. The main task then becomes to properly set the cell energy function $E(\boldsymbol{\eta})$ in order that the model will reproduce the main thermodynamic and long-range transport properties of some selected real host-guest systems. The low level of detail of the ThPCA cell allows the cell partition function to be very easy to compute, giving the model a wide flexibility.

The analogy between real pore and ThPCA cell stands on the following points:

(i) A real pore translates directly into a ThPCA cell.

(ii) Each real pore region close to the interface with one of the neighboring pores translates into one exit site of the ThPCA cell; if K_{ex} is the number of such interfaces then each cell has exactly K_{ex} exis sites.

(iii) The real pore space not close to interfaces with neighboring pores translates into the set of inner sites of the ThPCA cell.

(iv) While each molecule in the real pore has always its own identity and spatial position, in a ThPCA cell instead there is no information about the exact locations of exit and inner sites, therefore all the guests of each cell, say $\mathbf{r}$, are assumed to have the same spatial coordinates $\mathbf{r}$ — in other words, an indetermination in the position of the guests exists which is equal to the volume of the cell, and is unimportant for the long-time diffusion. As a consequence, only transfers from cell to cell will contribute to the total self-diffusivity D_s while displacements of the guests inside of each cell will not contribute. This is equivalent to say that the inter-cell migration scale becomes the only relevant scale.

A note about the dynamics inside of each cell. While in the real pore the configuration of the guest molecules changes in time through a sequence of simultaneous displacements in the real space, in the ThPCA cell a new configuration will be produced by the action of a probabilistic operator on the current configuration. In this context, the distinction between exit and inner sites makes the tendency of a cell to transfer particles to its neighboring cells a strongly temperature- and occupancy-dependent property. Therefore, just like in a real microporous system, intercell diffusion turns out to be a temperature-activated process, with a strongly loading-dependent diffusivity.

2.3 Local and Global Occupancies

Local occupancies. Local occupancies are defined inside of each cell. The *occupancy* of the j-th site of the cell $\mathbf{r}$ is denoted as $\eta_j(\mathbf{r}, t)$, which is 0 if empty or 1 if occupied. Each possible set of occupancies of the K sites represents a *cell configuration*,

$$\boldsymbol{\eta}(\mathbf{r}, t) = \{\eta_1, \ldots, \eta_K\}(\mathbf{r}, t), \tag{2.10}$$

which defines completely the instantaneous state of $\mathbf{r}$. The occupancy of a cell with configuration $\boldsymbol{\eta}$ is defined as

$$n(\boldsymbol{\eta}) = \sum_{j=1}^{K} \eta_j. \tag{2.11}$$

It shall be useful for the treatise in the next Sections to introduce the *partial configurations* of exit and inner sites,

$$\begin{aligned} \boldsymbol{\eta}_{\text{ex}} &= \{\eta_1, \ldots, \eta_{K_{\text{ex}}}\}, \\ \boldsymbol{\eta}_{\text{in}} &= \{\eta_{K_{\text{ex}}+1}, \ldots, \eta_K\}, \end{aligned} \tag{2.12}$$

and their relative *partial occupancies*,

$$n_{\mathrm{ex}}(\boldsymbol{\eta}) = \sum_{j=1}^{K_{\mathrm{ex}}} \eta_j, \qquad n_{\mathrm{in}}(\boldsymbol{\eta}) = \sum_{j=K_{\mathrm{ex}}+1}^{K} \eta_j. \tag{2.13}$$

A few words about the notation:

- Symbols with parentheses shall indicate operators. For instance, $n(\cdot)$ will indicate the occupancy operator, which evaluates the occupancy of a given configuration according to Eq. (2.11).
- If the argument of the operator (e.g. $\boldsymbol{\eta}$ in $n(\boldsymbol{\eta})$) is replaced by the cell position vector $\mathbf{r}$ it is assumed that the operator is acting on the proper argument evaluated at the cell $\mathbf{r}$. That is, $n(\mathbf{r}) \equiv n(\boldsymbol{\eta}(\mathbf{r}))$.
- Quantities represented through symbols without parentheses, e.g. n, are generic observables.
- In case the quantities appearing in equations are evaluated at the same instant of time, the time t will be kept as an implicit variable.
- A pair of neighboring cells will be indicated by enclosing the two cell coordinates into braces, i.e. $\{\mathbf{r}, \mathbf{r}^j\}$, where $\mathbf{r}^j$ is given by Eq. (2.8). Configuration and occupancy of the pair will be indicated respectively as:

$$\boldsymbol{\eta}\left(\mathbf{r}, \mathbf{r}^j\right) = \left\{\boldsymbol{\eta}(\mathbf{r}), \boldsymbol{\eta}(\mathbf{r}^j)\right\}, \tag{2.14}$$

$$n\left(\mathbf{r}, \mathbf{r}^j\right) = \left\{n(\mathbf{r}), n(\mathbf{r}^j)\right\}. \tag{2.15}$$

- The index $\alpha =$ 'ex', 'in' will designate the type of site (exit or inner, respectively).

Global occupancies Global occupancies are defined over the whole system $\mathscr{L}$. The *lattice configuration* (or *lattice state*, denoted $\boldsymbol{\eta}(\mathscr{L})$) is the set of all cells' configurations, and the *lattice occupancy* is the set of all cells' occupancies:

$$\boldsymbol{\eta}(\mathscr{L}) = \bigcup_{\mathbf{r}\in\mathscr{L}} \boldsymbol{\eta}(\mathbf{r}), \qquad n(\mathscr{L}) = \bigcup_{\mathbf{r}\in\mathscr{L}} n(\mathbf{r}). \tag{2.16}$$

The *loading* and the *partial* (*exit* or *inner*) *loadings* are defined respectively as:

$$\langle n \rangle = \frac{1}{M} \sum_{\mathbf{r} \in \mathscr{L}} n(\mathbf{r}), \qquad \langle n_\alpha \rangle = \frac{1}{M} \sum_{\mathbf{r} \in \mathscr{L}} n_\alpha(\mathbf{r}). \tag{2.17}$$

The CA simulates a random walk in the canonical ensemble, therefore the loading $\langle n \rangle$ (as well as temperature and chemical potential) is also a global invariant of the CA.

2.4 Site Thermodynamics

The site potential energy shall be indicated as

$$\varepsilon_\alpha(n) = \varepsilon_\alpha^o + \varepsilon_\alpha^{\mathrm{pp}}(n;T), \tag{2.18}$$

where the first term, $\varepsilon_\alpha^o \leq 0$, is the *sorption energy* of an α-site, and the second term (the superscript $^{\mathrm{pp}}$ stands for particle-particle (i.e., guest-guest) interactions) is the interaction potential between a guest in a type α-site and the remaining $n-1$ guests in the cell. (It is assumed $\varepsilon_\alpha^{\mathrm{pp}}(1;T) = 0$). The temperature-dependence is due to the fact that $\varepsilon_\alpha^{\mathrm{pp}}$ is a *local, mean-field effective interaction potential*: in case it has been derived by averaging the guest-guest interactions in a more structured cell it could depend on the (temperature-dependent) distribution of the adsorbed particles in the structured sites.

A mean-field *effective volume* $v_\alpha(n;T)$ is associated to each site, and assumed to depend on the number of guests in the cell and on the temperature. Ideally, the effective volume can be split into a fixed part and an occupancy-temperature-dependent part:

$$v_\alpha(n;T) = v_\alpha + \delta v_\alpha(n;T), \tag{2.19}$$

where v_α is the effective volume of one α-site when $n = 1$, while $v_\alpha(n;T)$ (which is null for $n < 2$) is the modification to the space available to one guest occupying one α-site due to the presence of other guests. The entropy associated to the effective volume of each α-site results

$$\begin{aligned} s_\alpha(n;T) &= k_B \ln \frac{v_\alpha}{\Lambda^d} + k_B \ln \left[\frac{\delta v_\alpha(n;T)}{v_\alpha} + 1 \right] \\ &= s_\alpha(T) + s_\alpha^{\mathrm{pp}}(n;T), \end{aligned} \tag{2.20}$$

where $\Lambda = h/\sqrt{2\pi m k_B T}$ is the thermal de Broglie wavelength. In the last line of Eq. (2.20) the entropy is splitted into non-occupancy-dependent ($s_\alpha(T)$) and occupancy-dependent ($s_\alpha^{\mathrm{pp}}(n;T)$) contribution.

The *site free energy* can therefore be defined as

$$f_\alpha(n;T) = f^o_\alpha(T) + \phi_\alpha(n;T), \tag{2.21}$$

where $f^o_\alpha(T) = \varepsilon_\alpha(n) - Ts_\alpha(T)$, and $\phi_\alpha(n;T) = \varepsilon^{\mathrm{PP}}_\alpha(n;T) - Ts^{\mathrm{PP}}_\alpha(n;T)$. The parameter f^o_α plays the role of an effective guest-host free energy associated to a site of type α, while $\phi_\alpha(n)$ contains all the occupancy-dependent terms. As long as the wall-guest interactions in microporous material determines the most of the adsorption and diffusion properties, the site free-energies f^o_{ex} and f^o_{in} play a primary role in the present model. Together with the competition of guests for occupying the sites, the heterogeneity introduced by assigning different values to f^o_{ex} and f^o_{in} represents the heart of the model since it mimics the process in which guests in real micropores compete to gain access to the windows. The *site partition function* results[71]

$$\begin{aligned} q_\alpha(n;T) &= \frac{v_\alpha(n;T)}{\Lambda^d}\, e^{-\beta\varepsilon_\alpha(n)} \\ &= e^{-\beta f_\alpha(n)}. \end{aligned} \tag{2.22}$$

2.5 Cell Thermodynamics

In this Section the energy structure and partition function of a generic cell of occupancy n shall be described. For simplicity of notation, from now on the temperature dependence of all free energies, partition functions and probabilities will be considered as implicit, e.g. $Q(n, K, T)$ will be replaced by $Q(n)$ and so on.

The cellular locality of interactions allows a *cellular partition function* to be defined exactly, and the single cell to be statistically equivalent to a small grand-canonical system. As it will be made clear in Section 4.3.2, this will enable us to calculate the chemical potential *on the fly* during the course of the simulation in the canonical ensemble, instead of requiring separate simulations in the grand-canonical ensemble.

As each configuration $\boldsymbol{\eta}$ of guests in the cell defines a cell *microstate*, together with the temperature the pair $(n_{\mathrm{ex}}, n_{\mathrm{in}})$ defines a cell *level* of free energy

$$F(n_{\mathrm{ex}}, n_{\mathrm{in}}) = \sum_{\alpha=\mathrm{ex,in}} n_\alpha f_\alpha(n). \tag{2.23}$$

It will be useful for later purposes to formulate the free energy as a sum of a F_0 (containing no n-dependent terms) and a Φ (containing the n-dependent

terms):

$$F(n_{\rm ex}, n_{\rm in}) = \sum_{\alpha={\rm ex,in}} n_\alpha f_\alpha + \sum_{\alpha={\rm ex,in}} n_\alpha \phi_\alpha(n) \equiv F_0(n_{\rm ex}, n_{\rm in}) + \Phi(n_{\rm ex}, n_{\rm in}). \tag{2.24}$$

When referred to a particular configuration, the free energy will be denoted as $F(\boldsymbol{\eta})$, where

$$F(\boldsymbol{\eta}) = F(n_{\rm ex}(\boldsymbol{\eta}), n_{\rm in}(\boldsymbol{\eta})). \tag{2.25}$$

The same notation applies also for F_0 and Φ.

A generic level $(n_{\rm ex}, n_{\rm in})$ contains a number of microstates (configurations) equal to the degeneracy $\binom{K_{\rm ex}}{n_{\rm ex}}\binom{K_{\rm in}}{n_{\rm in}}$. Therefore the *level partition function* is

$$Q_c(n_{\rm ex}, n_{\rm in}) = \prod_{\alpha={\rm ex,in}} \binom{K_\alpha}{n_\alpha} [q_\alpha(n)]^{n_\alpha} \tag{2.26}$$

$$= \binom{K_{\rm ex}}{n_{\rm ex}}\binom{K_{\rm in}}{n_{\rm in}} e^{-\beta F(n_{\rm ex}, n_{\rm in})}. \tag{2.27}$$

The *cellular partition function* turns out to be

$$Q(n) = \sum_{n_{\rm ex}, n_{\rm in}}^{(n)} Q_c(n_{\rm ex}, n_{\rm in}). \tag{2.28}$$

The conditional probability of a level for a given occupancy results

$$P(n_{\rm ex}|n) = \frac{Q_c(n_{\rm ex}, n - n_{\rm ex})}{Q(n)}, \tag{2.29}$$

where is implicitly defined by $n - n_{\rm ex} = n_{\rm in}$. The conditional probability of a particular configuration $\boldsymbol{\eta}$ given the occupancy n is therefore

$$P(\boldsymbol{\eta}|n) = \delta_{\rm kr}(n(\boldsymbol{\eta}), n)\, P(n_{\rm ex}(\boldsymbol{\eta})|n) \left\{ \binom{K_{\rm ex}}{n_{\rm ex}(\boldsymbol{\eta})} \binom{K_{\rm in}}{n - n_{\rm ex}(\boldsymbol{\eta})} \right\}^{-1}, \tag{2.30}$$

where the Krœneker delta $\delta_{\rm kr}$ selects only the configurations compatible with the occupancy n. ($\delta_{\rm kr}(n, m)$ takes value 1 if $n = m$ and 0 otherwise.)

Chapter 3

Time evolution

Basically, the evolution rule follows the randomization-propagation approach typical of LGCA models,[34,38] i.e. at each time step:

- A *randomization* changes the internal configuration of each cell (treated as a closed system) according to its present state only.
- A *propagation* opens every cell to its respective neighborhood allowing guest migrations from cell to cell.

However, the details of both operations as implemented in the present model differ from Refs.[34,38] in order to (i) be consistent with the physical problem considered here, and to (ii) take into account of local interactions.

In particular: The system evolves in time through the cyclic application of a time-independent lattice *randomization* operator, denoted $\mathcal{R}_{\mathcal{L}}$, mixing the internal configuration of each cell while preserving its occupancy, and a lattice *propagation* operator allowing guests in the exit sites to migrate to the neighboring cells, which in the presence of local interactions (i.e. for $\phi_\alpha(n) \neq 0$ in Eq. (2.24)) is a time-*dependent* operator and therefore denoted as $\mathcal{P}^t_{\mathcal{L}}$, whereas it can be formulated as a time-independent operator (therefore being denoted as $\mathcal{P}_{\mathcal{L}}$) if the guests are non-interacting (i.e. for $\phi_\alpha(n) = 0$).

Cyclic application means that:

- At each time t, the configuration resulting from randomization is used as input by the propagation;
- The output of the propagation at time t is used as input for randomization at time $t + \tau$ (where τ is the time step), and so on.

Both $\mathcal{R}_{\mathcal{L}}$ and $\mathcal{P}^t_{\mathcal{L}}$ are *global* operators, i.e. they work on the entire lattice configuration $\boldsymbol{\eta}(\mathscr{L})$. In Sections 3.2 and 3.3 it will be shown that they are

constructed as compositions of the *local* operators $\mathcal{R}$ and $\mathcal{P}$ (both time-independent) working respectively on one cell and one pair of neighboring cells.

The sequential application of $\mathcal{R}_\mathcal{L}$ and $\mathcal{P}_\mathcal{L}$ on the lattice configuration at time t produces the lattice configuration at time $t + \tau$:

$$\boldsymbol{\eta}(\mathscr{L}, t + \tau) = \mathcal{P}^t_\mathcal{L} \circ \mathcal{R}_\mathcal{L} \left(\boldsymbol{\eta}(\mathscr{L}, t)\right). \tag{3.1}$$

In the next Section a brief digression about the time step shall be made. After that, both operators will be extensively illustrated.

3.1 Time step

Each time step has a duration τ which can *ideally* be subdivided into a randomization and a propagation time, i.e. $\tau = \tau^R + \tau^P$. *Ideally* means that only the entire time step τ is required to make contact with the characteristic time of a real system simulated by the ThPCA, whereas the partial time steps τ^R and τ^P are not associated to the time-scale of any physical process occurring in the system.

The propagation time. In turn, the propagation time τ^P can be split up into propagation substeps having the same duration τ^P_*. Since each cell line is parallel to one particular direction of motion, as it will be described in details in the Section 3.3, it is possible to design the lattice partitions in such a way that the propagation on the entire lattice will be performed by first propagating along one direction, then along another one, and so on, until all d directions have been accounted once. Since 2 lattice partitions are constructed for each direction of motion, there are $2d$ ideal propagation substeps of duration $\tau^P_* = \tau^P/2d$.

Then, $\tau = \tau^R + 2d\tau^P_*$. Such a formalism shall be very useful in the formalization of the Block-partitioning scheme to be described in Section E.3.

Connection with a reference host-guest system. For direct coarse-graining purposes, the time step duration is given by equating the diffusivity of the molecular system one wants to emulate to the diffusivity of the automaton, both at infinite dilution. The resulting formula is

$$\tau = \frac{\lim_{\langle n\rangle\to 0}\left\{\langle \delta\mathbf{r}(0)\cdot\delta\mathbf{r}(0)\rangle + 2\sum_{z=1}^{\infty}\langle \delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle\right\}}{2\lim_{\langle n\rangle\to 0}\left\{\int_0^\infty \langle \mathbf{v}(t)\cdot\mathbf{v}(0)\rangle \mathrm{d}t\right\}}, \tag{3.2}$$

where $\langle \mathbf{v}(t)\cdot\mathbf{v}(0)\rangle$ is the velocity-autocorrelation function of the guest in the atomistic system, and $\langle \delta\mathbf{r}(t)\cdot\delta\mathbf{r}(0)\rangle$ is the displacement autocorrelation

function of the automaton (both quantities will be discussed in Section 4. In Eq. (3.2) the quantity $\langle \delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0) \rangle$ does not depend on τ but only on z).

3.2 Randomization operator

The lattice operator $\mathcal{R}_\mathcal{L}$ changes the internal configuration of the cells while preserving their respective occupancies. That is, during this operation each cell is treated as a *closed* canonical system. No intercell transfers are allowed.

The action of the operator $\mathcal{R}_\mathcal{L}$ is:

$$\mathcal{R}_\mathcal{L} : \boldsymbol{\eta}(\mathscr{L}, t) \rightarrow \quad \boldsymbol{\eta}(\mathscr{L}, t + \tau^R) \equiv \boldsymbol{\eta}^R(\mathscr{L}, t). \tag{3.3}$$

The operator $\mathcal{R}_\mathcal{L}$ uses the local operator $\mathcal{R}$, which works *independently on each cell* by mapping the input configuration $\boldsymbol{\eta}$ of the cell onto the output configuration $\boldsymbol{\eta}^R$ (having the same occupancy, i.e. $n(\boldsymbol{\eta}) = n(\boldsymbol{\eta}^R)$) according to some probabilistic algorithm. That is,

$$\mathcal{R}_\mathcal{L}\left(\boldsymbol{\eta}(\mathcal{L})\right) = \bigcup_{\mathbf{r} \in \mathscr{L}} \mathcal{R}\left(\boldsymbol{\eta}(\mathbf{r})\right). \tag{3.4}$$

Since only equilibrium systems are treated in the present thesis, only randomization schemes which satisfy detailed balance will be considered. A number of such algorithms (i.e. ways to transform $\boldsymbol{\eta}$ into $\boldsymbol{\eta}^R$) is possible. Two main types of randomization shall be described: the first one, called *memoryless randomization*, destroys the correlations in the motion of the guests, and the second one instead, called *jump randomization*, produces time-correlations in the dynamic properties. In Appendix E.2 the general proof of detailed balance for the randomization operation over the whole system is given.

3.2.1 Memoryless randomization

In this case in each cell the randomization is performed by a local operator $\mathcal{R}^{\mathrm{pf}}$ (the superscript 'pf' stands for 'partition function', so as to emphasize that such an operator determines the output configuration merely on the basis of the partition function Q_c), whose action is

$$\mathcal{R}^{\mathrm{pf}} : \boldsymbol{\eta} \rightarrow \boldsymbol{\eta}^R \text{ with probability } p^R\big(\boldsymbol{\eta} \rightarrow \boldsymbol{\eta}^R\big) \tag{3.5}$$

or, in a more compact form,

$$\mathcal{R}^{\mathrm{pf}} : \boldsymbol{\eta} \underset{p^R}{\longrightarrow} \boldsymbol{\eta}^R \tag{3.6}$$

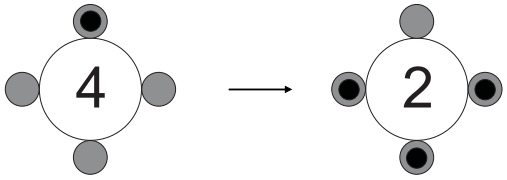

Figure 3.1: Schematic representation of the Memoryless Randomization applied on a two-dimensional cell. Occupied exit sites are pictured with a black dot inside, while the number of occupied inner sites is given by the number inside of the white circle. Such a transformation occurs with probability $P(3|5)$, according to Eq. (3.7). The cell occupancy ($n = 5$) is preserved during randomization.

where $p^R\big(\boldsymbol{\eta} \to \boldsymbol{\eta}^R\big)$ is defined as

$$p^R\big(\boldsymbol{\eta} \to \boldsymbol{\eta}^R\big) = P(\boldsymbol{\eta}^R|n(\boldsymbol{\eta})). \tag{3.7}$$

The formulation of p^R requires only the knowledge of the level and the local partition functions given in Eqs. (2.27) and (2.28).

It should be noted that $P(\boldsymbol{\eta}^R|n(\boldsymbol{\eta}))$ does not depend from the previous configuration except from the total number of guests, $n(\boldsymbol{\eta})$, which must be conserved during such randomization. With such a choice, $\mathcal{R}^{\mathrm{pf}}$ causes the complete loss of correlations between the input and the output cell configuration, and pushes the system towards the equilibrium state much more strongly than through a correlated process. The resulting self-diffusivity will only depend (linearly) on the probability of a guest to escape the host cell, therefore it will be called *memoryless diffusivity*. Every time a randomization with $\mathcal{R}^{\mathrm{pf}}$ has to be performed on a cell, the most general procedure would consists in the construction of a matrix $[\xi_{\boldsymbol{\eta},\boldsymbol{\eta}'}]$ of Boolean random variables such that:

(i) For each configuration $\boldsymbol{\eta}$ there is one and only one configuration $\boldsymbol{\eta}'$ such that $\xi_{\boldsymbol{\eta},\boldsymbol{\eta}'} = 1$.

(ii) The probability of the event $\xi_{\boldsymbol{\eta},\boldsymbol{\eta}'} = 1$ is given by $p^R(\boldsymbol{\eta} \to \boldsymbol{\eta}')$.

Then, each site occupancy of the post-randomization configuration is defined by

$$\eta_i^R = \sum_{\boldsymbol{\eta}'} \xi_{\boldsymbol{\eta},\boldsymbol{\eta}'} \eta_i'. \tag{3.8}$$

Of course such an approach is suitable only for small values of the maximum occupancy K. As example, for $K = 15$ and $n = 7$ the number of available configurations is $K!/n!(K-n)!$ and one would have to scan a 6435×6435 matrix to determine the output configuration. Such a drawback does not

exist in simulations with less-structured cells, where all the configurations of occupancy n can be grouped into *levels* according to their values of $n_{\mathrm{ex}}, n_{\mathrm{in}}$. Then a post-randomization configuration can be generated in an efficient way in the following way:

(i) Extract a new *level* $(n_{\mathrm{ex}}^R, n-n_{\mathrm{ex}}^R)$ according to its conditional probability $P(n_{\mathrm{ex}}^R|n)$ defined in Eq. (2.29).

(ii) Re-dispose randomly n_{ex} of the n guests in the exit sites.

An example of memoryless randomization is sketched in Figure 3.1. The above described sampling procedure would be unsuitable in structured-cell ThPCA, where in principle it is not possible to group equally probable configurations into large sets, thus the randomization becoming a very expensive operation.

3.2.2 Jump randomization

In the present Section an alternative randomization shall be described which is defined as a sequence of jumps, where each jumping guest preserves its own identity. Let us call $\mathcal{R}^{\mathrm{jump}}$ a local operator which transforms the cell configuration $\boldsymbol{\eta}$ in the following way:

i) Selection of the guests in a random sequence;

ii) For each selected guest, a target site is chosen among all the K sites of the cell.

iii) The selected guest will jump from its departure site, say j, to the target site, say k, with probability $p_{\mathrm{jump}}^R(\boldsymbol{\eta} \to \boldsymbol{\eta}')$ where $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$ are the cell configurations before and after the jump, respectively. If the transition from $\boldsymbol{\eta}$ to $\boldsymbol{\eta}'$ is produced due to a jump of a guest from its departure site, j, to some destination site, say k, then the acceptance probability is

$$p_{\mathrm{jump}}^R(\boldsymbol{\eta} \to \boldsymbol{\eta}') = C\, e^{\beta f_j^o} e^{\beta\left\{\Phi(\boldsymbol{\eta})-\max\left[\Phi(\boldsymbol{\eta}),\Phi(\boldsymbol{\eta}')\right]\right\}}, \tag{3.9}$$

where $C \in (0,1] \in \mathbb{R}$ can be set as a constant or as a function $C(\boldsymbol{\eta}, \boldsymbol{\eta}')$, provided that $C(\boldsymbol{\eta}, \boldsymbol{\eta}') = C(\boldsymbol{\eta}', \boldsymbol{\eta})$, and Φ is the interaction free-energy defined in Eq. (2.24).

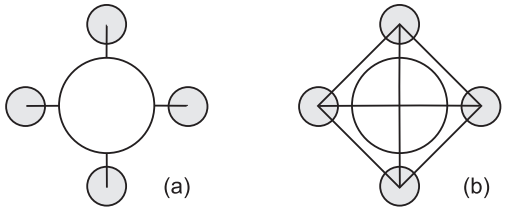

Figure 3.2: Schematic representation of two possible inner topologies in a two-dimensional cell that can be specified in the Jump Randomization scheme. Connections between sites are represented with straight lines. In (a) jumps from an exit site to different exit sites are forbidden according to the matrix A_{jk} defined in Eq. (3.12), whereas in (b) all sites can be chosen as targets independently of the departure site, according to the matrix A'_{jk} defined in Eq. (3.13).

Details about the derivation of Eq. (3.9) can be found in Appendix E.1. The jump randomization operator can then be defined as

$$\mathcal{R} : \boldsymbol{\eta} \underset{p^{(1)}_{\text{jump}}}{\longrightarrow} \boldsymbol{\eta}^{(1)} \underset{p^{(2)}_{\text{jump}}}{\longrightarrow} \boldsymbol{\eta}^{(2)} \underset{p^{(3)}_{\text{jump}}}{\longrightarrow} \cdots \underset{p^{(n)}_{\text{jump}}}{\longrightarrow} \boldsymbol{\eta}^{(n)} \tag{3.10}$$

where $\boldsymbol{\eta}^{(0)} \equiv \boldsymbol{\eta}$ and $\boldsymbol{\eta}^{(n)} \equiv \boldsymbol{\eta}^R$, while $p^{(n)}_{\text{jump}} = p^R_{\text{jump}}(\boldsymbol{\eta}^{(j-1)} \to \boldsymbol{\eta}^{(j)})$ which follows the definition given in Eq. (3.9).

As can be seen in Eq. (3.9), the first exponential term in the probability p^R_{jump} is an Arrhenius-like jump probability (for the adsorption barrier), whereas the second one is a the Metropolis acceptance rule for the cellular interactions. It is direct to obtain the non-interacting Arrhenius jump probability from Eq. (3.9) by putting $\Phi = 0$.

In the pure Metropolis formulation, the transition probability only takes into account a difference in energy between initial and final configuration. A consequence of this fact is that the ability of a guest to jump from a site with $\epsilon_j \ll 0$ to a site with the same adsorption energy, equals the acceptance of a jump between two zero-energy sites, i.e. the barrier due to adsorption disappears. This causes the Metropolis formulation to be unsuitable to obtain meaningful kinetic properties. The probability defined in Eq. (3.9) instead takes into account the presence of the adsorption barrier f^o_i.

After one randomization step in a cell with $\mathcal{R}^{\text{jump}}$, the output configuration, $\boldsymbol{\eta}^R$, may be equal, very similar, or totally different from the input one depending on the particular path chosen. Therefore, since the operator $\mathcal{R}^{\text{jump}}$ transforms $\boldsymbol{\eta}$ into $\boldsymbol{\eta}^R$ following a path while preserving the identities of the guests, correlations are introduced in the diffusion process.

Not allowed ex-ex jumps. The correlations can be made stronger or weaker through a proper manipulation of the prefactor C of Eq. (3.9). As

example, the adsorption barriers can be scaled by setting C as a constant between 0 and $\exp(-\beta \max[f^o_{\text{ex}}, f^o_{\text{in}}])$, where the latter value would give the full optimization of the Arrhenius adsorption barrier. That would raise the mobility of the guest, causing correlations to be weakened. To illustrate an example, let us introduce a *site type function* α_i defined as

$$\alpha_i = \begin{cases} \text{'ex'} & \text{for } i = 1, \ldots, K_{\text{ex}} \\ \text{'in'} & \text{for } i = K_{\text{ex}} + 1, \ldots, K_{\text{in}} \end{cases} \tag{3.11}$$

Now the jumps between different exit sites can be forbidden by assuming $C = A_{jk}$ where

$$A_{jk} = \begin{cases} 0 & \text{if } \alpha_j = \alpha_k = \text{'ex'} \\ \gamma & \text{otherwise} \end{cases} \tag{3.12}$$

with $0 < \gamma \leq \exp(-\beta \max[f^o_{\text{ex}}, f^o_{\text{in}}])$. The structure of such inner site-to-site connections is sketched in Figure 3.2a. With this approach, it is not possible anymore for a guest adsorbed in an exit site to reach a different exit site during the same randomization event. At least two randomizations are then required for a guest to perform such a displacement. This will considerably slow down the diffusion process.

Allowed ex-ex jumps. If instead one assumes $C = A'_{jk}$ where

$$A'_{jk} = \gamma \quad \text{for any } j, k \tag{3.13}$$

then each guest can choose any of the K cell sites to jump in. Such a situation is sketched in Figure 3.2b. Of course the mobility of guests undergoing randomization with A'_{jk} will be higher than A_{jk}, giving a slightly higher diffusivity. More important, during one randomization step, due to the less constrained dynamics arising from the use of A'_{jk} the configuration of guests in the cell will in general change more than the case of A_{jk}, thus producing weaker correlations. Examples will be given in Section 7.1.

Moreover by choosing $\gamma < 1$ not only the entire diffusion process slows down, but also correlations increase due to the backscattering of guests during propagation: for example when during propagation a guest migrates from an exit site of a cell, say $\mathbf{r}_1$, to the adjacent exit site of the neighboring cell, say $\mathbf{r}_2$, due to the low value of p^R_{jump} the probability of the guest to come back to cell $\mathbf{r}_1$ at the next propagation step is very high. The correlations introduced by $\mathcal{R}^{\text{jump}}$ mimic the memory effects caused by the backscattering phenomenon occurring in real zeolites.[32,63]

The uncorrelated limit for randomization. The correlated randomization converges to the uncorrelated one when at each time step each cell undergoes a large number of jump randomizations, that is

$$\lim_{z\to\infty}\left[\mathcal{R}^{\mathrm{jump}}\right]^{z}(\boldsymbol{\eta}) = \mathcal{R}^{\mathrm{pf}}(\boldsymbol{\eta}) \tag{3.14}$$

Along with this criterion, a finer tuning of correlations can be accomplished by introducing a dependence of the number of cell randomizations at each time step and the value of γ on the cell occupancy.

The introduction of backscattering correlations in the self-motion by means of the jump randomization operator $\mathcal{R}^{\mathrm{jump}}$ causes the value of D_s obtained from numerical simulations to strongly deviate from the value obtained by using the memoryless randomization. In Section 7 such correlation effects will be studied by means of numerical simulations under different conditions, while in Appendix A.2 the autocorrelation function $\langle\delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle$ (embedding the backscattering effects) will be analyzed under a mean-field approach to trace back to the dependence of correlations on the main macroscopic observables.

3.3 Local propagation operator

The propagation operator $\mathcal{P}$ applies on independent pairs of cells. It allows the two communicating exit sites to exchange one guest according to a propagation probability satisfying detailed balance.

Let us consider a pair of adjacent cells $\mathbf{r}$ and $\mathbf{r}^j$. They communicate through the respective exit sites $(\mathbf{r}, j)$ and $(\mathbf{r}^j, j+d)$. If only one of these sites is occupied, then the occupying guest can try to migrate to the other site. In case of successful jump, the local observables $n_{\mathrm{ex}}(\cdot)$, $n(\cdot)$ would be transformed into $n^P_{\mathrm{ex}}(\cdot)$, $n^P(\cdot)$ given by:

$$\begin{aligned}
n^P_{\mathrm{ex}}(\mathbf{r}) &= n_{\mathrm{ex}}(\mathbf{r}) - \eta_j(\mathbf{r}) + \eta_{j+d}(\mathbf{r}^j),\\
n^P_{\mathrm{ex}}(\mathbf{r}^j) &= n_{\mathrm{ex}}(\mathbf{r}^j) - \eta_{j+d}(\mathbf{r}^j) + \eta_j(\mathbf{r}),\\
n^P(\mathbf{r}) &= n^P_{\mathrm{ex}}(\mathbf{r}) + n_{\mathrm{in}}(\mathbf{r}),\\
n^P(\mathbf{r}^j) &= n^P_{\mathrm{ex}}(\mathbf{r}^j) + n_{\mathrm{in}}(\mathbf{r}^j).
\end{aligned} \tag{3.15}$$

The $n_{\mathrm{in}}(\cdot)$s would remain unchanged, that is

$$\begin{aligned}
n^P_{\mathrm{in}}(\mathbf{r}) &= n_{\mathrm{in}}(\mathbf{r}),\\
n^P_{\mathrm{in}}(\mathbf{r}^j) &= n_{\mathrm{in}}(\mathbf{r}^j).
\end{aligned} \tag{3.16}$$

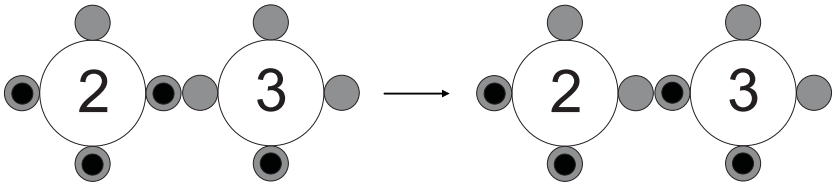

Figure 3.3: Schematic representation of the propagation operation between two communicating two-dimensional cells. Such a transformation is expected to happen with probability p^P given in Eq. (3.22).

The individual output configurations in case of successful jump shall be denoted $\boldsymbol{\eta}^P(\mathbf{r})$ and $\boldsymbol{\eta}^P(\mathbf{r}^j)$ (according to Eq. (2.14). The abbreviate notation $\boldsymbol{\eta}^P(\mathbf{r},\mathbf{r}^j)$ will indicate the simultaneous pair of configurations, i.e. $\boldsymbol{\eta}^P(\mathbf{r},\mathbf{r}^j) = \boldsymbol{\eta}^P(\mathbf{r}) \cup \boldsymbol{\eta}^P(\mathbf{r}^j)$. They are equivalent to the input configurations except for the fact that the occupancies of the communicating exit sites are swapped, that is

$$\mathcal{P} : \left\{ \begin{array}{l} \eta_j(\mathbf{r}) \\ \eta_{j+d}(\mathbf{r}^j) \end{array} \right. \rightarrow \left\{ \begin{array}{l} \eta^P_j(\mathbf{r}) = \eta_{j+d}(\mathbf{r}^j) \\ \eta^P_{j+d}(\mathbf{r}^j) = \eta_j(\mathbf{r}) \end{array} \right. \text{ with probability } p^P\big(\boldsymbol{\eta}(\mathbf{r},\mathbf{r}^j) \rightarrow \boldsymbol{\eta}^P(\mathbf{r},\mathbf{r}^j)\big)\,. \quad (3.17)$$

With the above definition of $\mathcal{P}$, the post-propagation configuration of a pair of neighboring cells reads

$$\boldsymbol{\eta}^P(\mathbf{r},\mathbf{r}^j) = \mathcal{P}\left(\boldsymbol{\eta}(\mathbf{r},\mathbf{r}^j)\right). \quad (3.18)$$

The propagation probability. To complete the construction of the local propagation operator $\mathcal{P}$, it is required to formulate the propagation probability p^P. As for the randomization case, this can be done by making use of local partition functions: When the propagation operator acts on a pair of cells there are only two possible output configurations of the pair: that is, (1) the input configuration itself, with weight

$$q(\mathbf{r},\mathbf{r}^j) = \prod_{\boldsymbol{\rho}=\mathbf{r},\mathbf{r}^j} \prod_{\alpha=\text{ex,in}} \Big[q_\alpha\big(n(\boldsymbol{\rho})\big)\Big]^{n_\alpha(\boldsymbol{\rho})}, \quad (3.19)$$

and (2) the new one obtainable by swapping the occupancies of the two communicating exit sites having weight

$$q^P(\mathbf{r},\mathbf{r}^j) = \prod_{\boldsymbol{\rho}=\mathbf{r},\mathbf{r}^j} \prod_{\alpha=\text{ex,in}} \Big[q_\alpha\big(n^P(\boldsymbol{\rho})\big)\Big]^{n^P_\alpha(\boldsymbol{\rho})}, \quad (3.20)$$

where the $n^P(\boldsymbol{\rho})$ values are given in Eq. (3.15). Since inside of the cells permutations of guests are not allowed during propagation, the level degeneracies $\binom{K_\alpha}{n_\alpha(\boldsymbol{\rho})}$ and $\binom{K_\alpha}{n_\alpha^P(\boldsymbol{\rho})}$ are not included in the products in Eqs. (3.19) and (3.20).

Now the propagation probability can be defined as

$$p^P(\mathbf{r},\mathbf{r}^j) = \frac{q^P(\mathbf{r},\mathbf{r}^j)}{q(\mathbf{r},\mathbf{r}^j)+q^P(\mathbf{r},\mathbf{r}^j)}\, e^{-\beta\epsilon_{\mathrm{ki}}(\mathbf{r},\mathbf{r}^j)}\delta(\mathbf{r},\mathbf{r}^j) \tag{3.21}$$

$$= \frac{e^{-\beta\epsilon_{\mathrm{ki}}(\mathbf{r},\mathbf{r}^j)}}{1+e^{\beta\Delta F(\mathbf{r},\mathbf{r}^j)}}\delta(\mathbf{r},\mathbf{r}^j), \tag{3.22}$$

where the quantity $\Delta F(\mathbf{r},\mathbf{r}^j)$ introduced in Eq. (3.22) is the difference between the free energy after and before the successful jump, given by

$$\Delta F(\mathbf{r},\mathbf{r}^j) = \sum_{\boldsymbol{\rho}=\mathbf{r},\mathbf{r}^j}\Big[\Phi\left(\boldsymbol{\eta}^P(\boldsymbol{\rho})\right)-\Phi\left(\boldsymbol{\eta}(\boldsymbol{\rho})\right)\Big], \tag{3.23}$$

and the exclusion term $\delta(\mathbf{r},\mathbf{r}^j)$ is defined as

$$\delta(\mathbf{r},\mathbf{r}^j) = \eta_j(\mathbf{r})\left[1-\eta_{j+d}(\mathbf{r}^j)\right]+\eta_{j+d}(\mathbf{r}^j)\left[1-\eta_j(\mathbf{r})\right], \tag{3.24}$$

or, equivalently,

$$\delta(\mathbf{r},\mathbf{r}^j) = \begin{cases} 1, & \text{if } \eta_j(\mathbf{r})+\eta_{j+d}(\mathbf{r}^j)=1 \\ 0, & \text{otherwise} \end{cases} \tag{3.25}$$

Before to proceed with the description of the kinetic barrier $\epsilon_{\mathrm{ki}}(\mathbf{r},\mathbf{r}^j)$ (see Section 3.3.1) the general strategy for simulating a propagation event will be illustrated. Formally, a propagation between two communicating cells $\mathbf{r}$ and $\mathbf{r}^j$ is performed by picking a random Boolean ζ such that $\zeta=1$ with probability p^P given in Eq. (3.22). Then, the post-propagation occupancies of the two communicating exit sites are determined through the following relations:

$$\eta_j^P(\mathbf{r}) = \eta_j(\mathbf{r})+\left[\eta_{j+d}(\mathbf{r}^j)-\eta_j(\mathbf{r})\right]\zeta(\mathbf{r},\mathbf{r}^j), \tag{3.26}$$

$$\eta_{j+d}^P(\mathbf{r}^j) = \eta_{j+d}(\mathbf{r}^j)+\left[\eta_j(\mathbf{r})-\eta_{j+d}(\mathbf{r}^j)\right]\zeta(\mathbf{r},\mathbf{r}^j). \tag{3.27}$$

It should be noted that all quantities defined here depend only *on the local configurations* and the temperature, not on the space (the 'space notation' $\mathbf{r}$ is used only for simplicity).

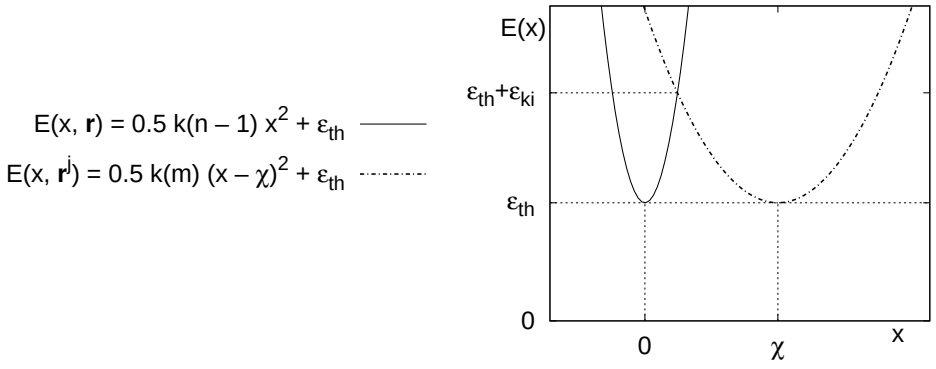

Figure 3.4: The intercell kinetic barrier between the cell $\mathbf{r}$ (with occupancy n) and the cell $\mathbf{r}$ (with occupancy m).

3.3.1 Thermodynamic and kinetic barriers

Eq. (3.22) can be rearranged as

$$p^P(\mathbf{r}, \mathbf{r}^j) = \delta(\mathbf{r}, \mathbf{r}^j) e^{-\beta\epsilon_{\rm th}(\mathbf{r},\mathbf{r}^j)} e^{-\beta\epsilon_{\rm ki}(\mathbf{r},\mathbf{r}^j)} \tag{3.28}$$

in order to highlight the thermodynamic barrier $\epsilon_{\rm th}(\mathbf{r}, \mathbf{r}^j) = k_{\rm B}T \ln(1 + e^{\beta\Delta F(\mathbf{r},\mathbf{r}^j)})$.

The kinetic barrier $\epsilon_{\rm ki}(\mathbf{r}, \mathbf{r}^j) \geq 0$ adds up to the thermodynamic barrier in the migration process. It may be set up to be constant, or dependent on some local observables with the detailed balance constraint that $\epsilon_{\rm ki}$ must remain unchanged while reversing the propagation process. A suitable choice is to model $\epsilon_{\rm ki}$ as a function of the occupancies of departure and target cells, i.e. $\epsilon_{\rm ki}(n, m)$, defined for $n = 1, \ldots, K$ and $m = 0, \ldots, K-1$, where $n = n(\mathbf{r})$ and $m = n(\mathbf{r}^j)$, provided

$$\epsilon_{\rm ki}(n, m) = \epsilon_{\rm ki}(m + 1, n - 1). \tag{3.29}$$

Anyway, in order to model the local kinetic barrier with such an approach, one should manage a $K \times K$ matrix. The number of parameters to model can be reduced if one observes that the detailed balance is automatically satisfied if $\epsilon_{\rm ki}$ is formulated as a function of some invariant (local) observables which does not change during propagation. A trivial example of such an observable would be the sum of the occupancies of the cells involved in the propagation operation, i.e. $\epsilon_{\rm ki} = \epsilon_{\rm ki}(n + m)$, which reduces the $K \times K$ matrix into a $2K$-array.

Another way would be to picture the propagation event as a process involving three local configurations denoted I (initial configuration), II (tran-

sition configuration), and III (target configuration):

$$\mathbf{n}_{\mathrm{J}} = \left\{ n_{\mathrm{J}}(\mathbf{r}),\, n^*_{\mathrm{J}}(\mathbf{r}, \mathbf{r}^j),\, n_{\mathrm{J}}(\mathbf{r}^j) \right\} \tag{3.30}$$

where J= I, II, III, and:

- $n_{\mathrm{J}}(\mathbf{r})$ is the occupancy of the cell $\mathbf{r}$ for the configuration J;
- $n^*_{\mathrm{J}}(\mathbf{r}, \mathbf{r}^j)$ is the occupancy of the transition state for the configuration J, that is it has value 1 if, according to configuration J, one guest is on the top of the barrier between the cells $\mathbf{r}$ and $\mathbf{r}^j$, and 0 otherwise;
- $n_{\mathrm{J}}(\mathbf{r}^j)$ is the occupancy of the cell $\mathbf{r}^j$ for the configuration J.

For the event of jump of one guest from $\mathbf{r}$ to $\mathbf{r}^j$, one has:

$$\begin{array}{llllll} n_{\mathrm{I}}(\mathbf{r}) & = n, & n^*_{\mathrm{I}}(\mathbf{r}, \mathbf{r}^j) & = 0, & n_{\mathrm{I}}(\mathbf{r}^j) & = m \\ n_{\mathrm{II}}(\mathbf{r}) & = n-1, & n^*_{\mathrm{II}}(\mathbf{r}, \mathbf{r}^j) & = 1, & n_{\mathrm{II}}(\mathbf{r}^j) & = m \\ n_{\mathrm{III}}(\mathbf{r}) & = n-1, & n^*_{\mathrm{III}}(\mathbf{r}, \mathbf{r}^j) & = 0, & n_{\mathrm{III}}(\mathbf{r}^j) & = m+1 \end{array}$$

The final configuration will be III or I if the jump is successful or not, respectively. The reverse propagation process would be

$$\begin{array}{llllll} n_{\mathrm{I}}(\mathbf{r}^j) & = m+1, & n^*_{\mathrm{I}}(\mathbf{r}, \mathbf{r}^j) & = 0, & n_{\mathrm{I}}(\mathbf{r}) & = n-1 \\ n_{\mathrm{II}}(\mathbf{r}^j) & = m, & n^*_{\mathrm{II}}(\mathbf{r}, \mathbf{r}^j) & = 1, & n_{\mathrm{II}}(\mathbf{r}) & = n-1 \\ n_{\mathrm{III}}(\mathbf{r}^j) & = m, & n^*_{\mathrm{III}}(\mathbf{r}, \mathbf{r}^j) & = 0, & n_{\mathrm{III}}(\mathbf{r}) & = n \end{array}$$

As can be seen the transition configuration is the same in both the forward and the reverse process. If it is assumed that the two neighboring cells will contribute to the kinetic barrier with two harmonics related to their respective occupancies in the transition configuration, that is

$$\begin{array}{ll} E(x, \mathbf{r}) & = \frac{1}{2} k(n-1) x^2 \\ E(x, \mathbf{r}^j) & = \frac{1}{2} k(m)(x-\chi)^2, \end{array} \tag{3.31}$$

then the energy $\epsilon_{\mathrm{ki}}(n, m)$ will be given by the intersection between $E(x, \mathbf{r})$ and $E(x, \mathbf{r}^j)$ for $0 \leq x \leq \chi$. In Eq. (3.31) the parameters $k(n-1)$ and $k(m)$ are force constant, while the variable x and the parameter χ are 'distances' in configuration space. In particular, χ can be related to the distance between the centers of two cells or alternatively be treated simply as another adjustable parameter which has the effect to raise or decrease all the kinetic barriers. With such a strategy the $K \times K$ kinetic barriers can be set up in a very simplified form through the modeling of the K-vector $\{k(n)\}_{n=0,\ldots,K-1}$ (plus χ if it is used as a free parameter). The construction of the kinetic barrier among two cells is schematized in Figure 3.4.

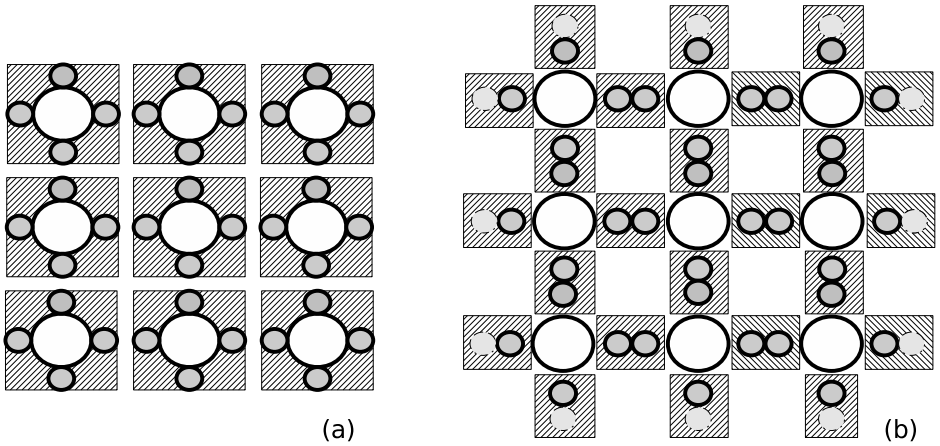

Figure 3.5: Lattice partitions switching during the time evolution of the system in the case of non-interacting guests. (a) Randomization subset (that is valid also for the interacting case), and (b) Propagation subset.

3.4 Global propagation operator

The global propagation operator extends the local propagation operation to the entire lattice. It can be *fully synchronous* (i.e. acting synchronously on all pairs of neighboring cells) or *block synchronous* (i.e. acting on alternating, independent partitions of the whole set of neighboring pairs), depending on whether the guest-guest interactions are absent, or present, respectively.

3.4.1 Non-interacting case

If the local interaction term in Eq. (2.24) $\phi_\alpha(n)$ is null for all values of n and the kinetic intercell barrier ϵ_{ki} introduced in Eq. (3.29) is constant, then the guests are non-interacting, and the local propagation probability reads

$$p^P(\mathbf{r}, \mathbf{r}^j) = \frac{1}{2} e^{-\beta\epsilon_{\mathrm{ki}}} \delta(\mathbf{r}, \mathbf{r}^j). \tag{3.32}$$

When the kinetic barrier ϵ_{ki} is fixed, the exponential term is indicated as κ^o:

$$\kappa^o = e^{-\beta\epsilon_{\mathrm{ki}}}. \tag{3.33}$$

Since the only time-dependent variable in Eq. (3.32) is $\delta(\mathbf{r}, \mathbf{r}^j)$, which depends on the occupancy of the two adjacent exit sites (i.e. $\eta_j(\mathbf{r})$ and $\eta_{j+d}(\mathbf{r}^j)$), then the non-interacting propagation operator can be applied synchronously on all the pairs of adjacent exit sites in the whole lattice. Therefore the global

propagation operator for the non-interacting case can be written as

$$\mathcal{P}^{\mathrm{ni}}_{\mathscr{L}}(\boldsymbol{\eta}_{\mathscr{L}}) = \bigcup_{j=1}^{d} \bigcup_{\mathbf{r}\in\mathscr{L}} \mathcal{P}\left(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)\right) \tag{3.34}$$

As a result, the whole automaton evolves in time by alternating two lattice partitions: the first one is the set of all cells, each taken as a closed system (see Figure 3.5a); the second one is the set of all pairs of adjacent exit sites.

3.4.2 Interacting case: partitioning scheme

If the interaction term $\phi_\alpha(n)$ is nonzero, then it is no longer possible to apply the above defined propagation *synchronously* on all pairs of cells. This is because of the dependence of the interaction contribution $\phi_\alpha(n)$ on the occupancy of the other exit sites of the same cell (since $n = n_{\mathrm{ex}} + n_{\mathrm{in}}$). A partitioning scheme is needed to make the model able to perform a propagation without losing the synchronicity of the updates. Such a scheme is possible under the Block Cellular Automata (BCA) approach introduced by Toffoli and Margolus,[35,72] originally designed to construct reversible CA and/or conserve selected physical quantities during each local operation. The grid is subdivided into different partitions of non-overlapping blocks. The key idea is that, due to the locality of the evolution laws, all blocks belonging to the same partition can undergo the propagation synchronously. Then, the partitions are switched in a random sequence in order each cell to communicate with all its 6 neighbors at each time step without introducing memory effects.

The need of a partitioning scheme. As mentioned in the introduction of the present Section 3.4, an operation on the lattice is *fully synchronous* when it can be applied simultaneously on all the cells $\mathbf{r} \in \mathscr{L}$. The randomization operation (described in Section 3.2) represents an example of fully synchronous operation.

Although strictly local at the cellular level, the ThPCA interactions introduced in Sections 2.4 and 2.5 do not allow all the exit sites to be independent one of each other: indeed, guests located in exit sites of the same cell will interact, so that every change in occupancy caused by gain or loss of guests in the exit sites due to the local propagation along some direction will cause the energy of the whole cell to change. Now, since the instantaneous cell energy enters the propagation probability formulation in Eq. (3.22), the propagation probabilities centered on a generic cell $\mathbf{r}$, that is $p^P(\mathbf{r}, \mathbf{r}^1), \ldots, p^P(\mathbf{r}, \mathbf{r}^{2d})$, will have different values depending on which sequence the $2d$ pairs are invoked

for propagation. For this reason the propagation cannot be fully synchronous in the presence of interactions.

Nevertheless, the entire lattice $\mathscr{L}$ can be partitioned into *independent* sublattices (called *partitions*) of non-overlapping groups of communicating cells (called *blocks*), so that the propagation can be performed *simultaneously* on all the blocks of each sublattice. Such an operation is said to be *block-synchronous*. Therefore, with the opportune partitioning scheme also non-synchronous operations (provided they are local) on the lattice can be performed with a certain degree of synchronicity. Along with the previous statements, a BCA can be defined as a CA containing one or more block-synchronous operations.

The Margolus approach. In BCA models, detailed balance of the local map within each partition translates directly into detailed balance of the whole automaton. Before to illustrate in details the partitioning scheme adopted in the present model, the procedure of construction of a generic BCA will be summarized following Toffoli and Margolus:[35,72]

- The array of cells is partitioned into a collection of finite, disjoint and uniformly arranged pieces called *blocks*.
- A block rule is given that looks at the contents of a block and updates the whole block (rather than a single cell as in an ordinary cellular automaton). The same rule is applied to every block. Since blocks do not overlap, no information is exchanged between adjacent blocks.
- After a partition of blocks has been updated, the partition is switched so as to have some overlap between the blocks of the new partition and the blocks of the old one. A time step is completed when all cells have exchanged information one time with each of their neighbors.

Following this general scheme, various partitioning strategies can be designed *ad hoc* to optimize CA operations on different systems. The crucial point is to choose the smallest of all possible independent block of cells, in order to maximize the number of synchronous operations.

In the ThPCA,

$$(\text{one block}) \equiv (\text{one pair of adjacent cells}).$$

In three dimensions, the host system is a cube of L^3 cells (where L is the side of the cube, properly chosen as an even number). For each direction x, y and z there are L^2 cell lines spanning the cube from one side perpendicularly to the opposite side. Including periodic boundary conditions, two

alternating partitions are therefore possible for each line, each consisting of $L/2$ non-overlapping pairs of adjacent cells. Then at each propagation step the partitions are constructed as follows: (i) a sequence of the three directions x, y and z is randomly chosen, and (ii) for each direction, the sequence of the two partitions is randomly chosen for each cell line.

In Appendix E.3 the partitioning scheme is described in details, while in Appendix E.4 the proof of detailed balance for the propagation operation on the whole system is given.

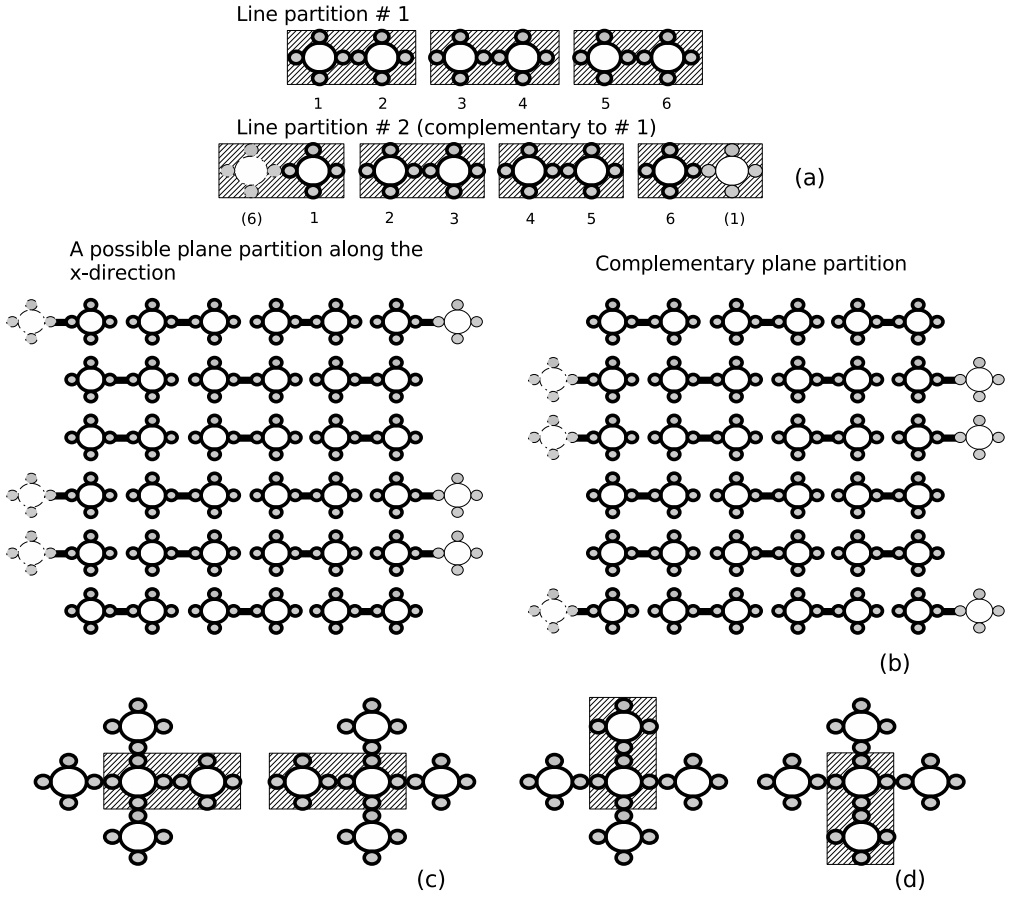

Figure 3.6: A two-dimensional schematic picture of the block partitions adopted in the ThPCA to perform the propagation operation in the presence of intracell interactions. Periodic boundary conditions are included. (a) The two possible partitions of a line of cells going from one side of the lattice to the opposite one. (b) Two possible complementary partitions of a plane of cells into blocks, as arising from the random choice of the sequence of partitions of each line. The partitioning scheme into blocks allows to update all blocks in each partition independently of each other while keeping detailed balance preserved in the presence of interactions. (c,d) The partitioning scheme under the point of view of a single cell in two dimensions: first of all, a random choice determines whether the cell will try to exchange guests first in the x (c) or in the y (d) direction; then for each direction a random choice will establish the updating sequence into each direction (that is, whether to update first the left or the right direction in the x case, and first the up or the down direction in the y case.)

Chapter 4

Equilibrium and transport properties

Actually, sorption and diffusion in zeolites are closely related: adsorption strongly affects diffusion and vice-versa.[67] A simultaneous understanding of both phenomena is needed to formulate models which can be useful to the interpretation of experimental results, and to properly design zeolite-based separation and/or catalytic processes. In the present Chapter the main equilibrium and transport properties that can be evaluated directly from numerical simulations of the ThPCA shall be introduced. The equilibrium properties of major interest are the probability distribution of occupancies (since from that it is possible to compute a number of very relevant thermodynamic quantities, like the reduced variance), and the chemical potential. Then, the following transport properties shall be discussed: diffusion coefficients, and local density fluctuation, where 'local density' here stands for 'cell occupancy'.

4.1 Probability distribution of occupancies

A fluid adsorbed in a zeolite shows a spatially varying density. This is the key to characterize intracrystalline mass transport and adsorption. In order to develop the thermodynamics of this strongly inhomogeneous fluid, a theory to predict how the sorbed particles distribute themselves into the inner void space that make up the zeolitic host is needed. This probability distribution of guests over independent subvolumes[73] is a quantity that can be related to different interesting information about the system.[21] In a comprehensive theory, factors like the finite size of the guests (which gives rise to the excluded volume effects) and the interaction forces between the guests and the

zeolitic host must be taken into account. Furthermore, at this molecular level of resolution, averaged properties and fluctuations turn out to be intimately related and equally important. In this field Xenon-129 NMR spectroscopy has been used under a variety of experimental conditions as a versatile tool to directly measure the fluctuations and the spatial correlations typical of guest particles in confined systems.[74,75] It has been possible to determine many different features of molecular dispersion within zeolite cavities and it has been noted that care must be taken in assuming a specific type of distribution.

The probability distribution of guest molecules plays a central role among all the ThPCA equilibrium properties, since from that one is possible to compute a wide variety of thermodynamic quantities, such as e.g. the chemical potential. The distribution of the adsorbate in the lattice has been studied in terms of the probability finding exactly n guests within a cell, which can be easily averaged from a numerical N-constant simulation in the long-time limit. Such a distribution shall be denoted as

$$\mathbf{p}_N = \{p_N(n), n = 0, \ldots, K\}\,, \tag{4.1}$$

where the subscript N emphasizes that such a distribution has been calculated from an N-constant simulation. This is done in order to distinguish $\mathbf{p}_N$ from the theoretical (mean-field) occupancy distribution $\mathbf{p}$ (which will be introduced in Section A) derived from a grand-canonical formalism.

4.2 Reduced variance

A remarkable (static) equilibrium property is the *reduced variance*, that is the ratio between the variance of the occupancy probability distribution and the loading. That can be interpreted as a measure of the the thermodynamic tendency of a cell to accept a new guest.[41] It is related to the isothermal compressibility, $\kappa_V = -V^{-1}(\partial V/\partial p)_{N,T}$, according to[70]

$$\frac{\langle n^2\rangle - \langle n\rangle^2}{\langle n\rangle} = \frac{\langle n\rangle k_B T \kappa_V}{V}, \tag{4.2}$$

where V is the volume of one cavity. This point deserves some comments. Indeed, it is appropriate to remember that a real zeolitic framework allows the adsorbed molecules to move in a potential field of varying magnitude and sign. The molecules are localized by the repulsive part of this potential field, and the extent of this localization depends on the temperature and other thermodynamic variables. There is, therefore, no exactly defined volume, and a

strict statistical mechanical connection between the reduced variance and a local isothermal compressibility related to the occupancy fluctuations in the cavity would be desirable but, unfortunately, difficult to assess rigorously. Anyway, the reduced variance gives a rough estimate of a local microscopic compressibility that acts as a suitable order parameter to describe the transition to an even more decreasing mobility regime. As it will be made clear in Section 4.4.3, such a quantity is exactly the reciprocal of the thermodynamic factor (usually denoted Γ) connecting the collective diffusivity with the Fickian diffusivity.[40] It can be evaluated directly from the knowledge of $p_N(n)$, the j-th moment of which follows the standard definition

$$\langle n^j \rangle = \sum_{n=0}^{K} n^j p_N(n). \tag{4.3}$$

4.3 Evaluation of chemical potential

The locality of the ThPCA interactions allows to easily evaluate the chemical potential, μ, from a N-constant simulation. An adsorption isotherm, that is the plot of μ *vs.* the loading $\langle n \rangle$ (or the inverse plot), is a valuable source of information with regard to the peculiar features of a host-guest system, since it describes quantitatively the amount of gas adsorbed by a porous material at a fixed temperature as a function of fugacity f (the notation f for the fugacity will have has no superscript nor arguments, in order to avoid any confusion with the free energies f_α^o and $f_\alpha(n)$), which is related to the chemical potential through

$$\mu = \mu^o + \beta^{-1} \ln f, \tag{4.4}$$

where μ^o is the standard chemical potential of the guest. As in microporous materials the filling of micropores is ruled by the stronger interactions between the adsorbate molecules and the pore walls, the shape of an isotherm can be exploited to extract information about these forces. Connection at the macroscopic scale between the ThPCA and a reference host-guest system one wants to emulate lays on agreement between the reference adsorption isotherm and theone produced by the ThPCA parameterization. Two strategies can be used to evaluate the chemical potential in the present model. The first one, which will be illustrated in Section 4.3.1, is valid only if guest-guest interactions are null, and requires the knowledge of the equilibrium values of the total number of filled exit and inner sites. With a more general strategy (see Section 4.3.2) one can exploit the cellular locality of interactions to evaluate the chemical potential as a function of the set of canonical partition functions of a closed cell, and of the occupancy probability distribution, p_N.

4.3.1 Non-interacting case

In the absence of guest-guest interactions, the system can be subdivided into two ideal subsystem: (i) the subsystem of exit sites, and (ii) the subsystem of inner sites. According to fluctuation theory,[70] the canonical partition function of each subsystem can be written as

$$Q_\alpha(N_\alpha^{\mathrm{eq}}) = \binom{M_\alpha}{N_\alpha^{\mathrm{eq}}} \exp\left(-\beta N_\alpha^{\mathrm{eq}} f_\alpha^o\right), \tag{4.5}$$

where $\alpha = \mathrm{ex}, \mathrm{in}$ respectively for the subsystem of exit and inner sites. Therefore $M_\alpha = MK_\alpha$ is the total number of α-type sites, in the lattice, while N_α^{eq} shall indicate the number of occupied sites of the α-subsystem when the entire system is *at equilibrium*. The free energy can be computed of the α-subsystem through

$$F_\alpha(N_\alpha^{\mathrm{eq}}) = -k_B T \ln Q_\alpha(N_\alpha^{\mathrm{eq}}), \tag{4.6}$$

and the chemical potential through

$$\mu_\alpha(N_\alpha^{\mathrm{eq}}) = \frac{\partial F_\alpha(N_\alpha^{\mathrm{eq}})}{\partial N_\alpha^{\mathrm{eq}}}. \tag{4.7}$$

Using Stirling's approximation for large systems and noting that $\rho_\alpha = N_\alpha^{\mathrm{eq}}/M_\alpha$ is the fraction of occupied sites in the α subsystems, one obtains

$$F_\alpha(N_\alpha^{\mathrm{eq}}) = k_B T M_\alpha \left[\frac{\rho_\alpha f_\alpha^o}{k_B T} + \rho_\alpha \ln \rho_\alpha + (1-\rho_\alpha)\ln(1-\rho_\alpha)\right] \tag{4.8}$$

for the free energy and

$$\mu_\alpha(N_\alpha^{\mathrm{eq}}) = f_\alpha^o + k_B T \ln\left(\frac{\rho_\alpha}{1-\rho_\alpha}\right) \tag{4.9}$$

for the chemical potential. Since the N_α^{eq} are equilibrium values, the relation

$$\mu_{\mathrm{ex}}(N_{\mathrm{ex}}^{\mathrm{eq}}) = \mu_{\mathrm{in}}(N_{\mathrm{in}}^{\mathrm{eq}}) \tag{4.10}$$

is satisfied. From Eq. (4.9) the equilibrium condition (4.10) implies

$$\frac{\rho_{\mathrm{ex}}(1-\rho_{\mathrm{in}})}{\rho_{\mathrm{in}}(1-\rho_{\mathrm{ex}})} = \exp\left(-\frac{\Delta f^o}{k_B T}\right), \tag{4.11}$$

where $\Delta f^o = f_{\mathrm{ex}}^o - f_{\mathrm{in}}^o$. Therefore, evaluating the chemical potential is possible from the knowledge of ρ_{ex} and ρ_{in} and using Eq. (4.9).

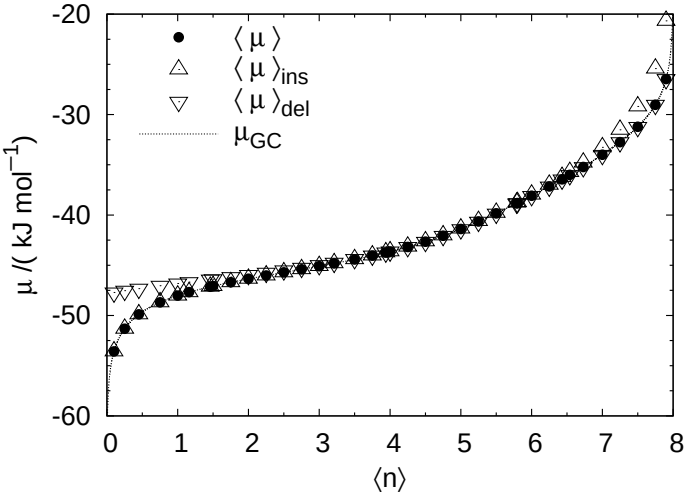

Figure 4.1: Modeled chemical potential (units of kJ mol^{-1}) for adsorption of xenon in NaA zeolite at 300 K.[56] Average chemical potentials extracted from numerical simulations are plotted together with the chemical potential expected from the grand-canonical formulation of the single cell (see Section A.1), indicated as μ_{GC}.

4.3.2 Interacting case

In this case the approach illustrated in Section 4.3.1 cannot be used since, due to the presence of guest-guest interactions, the energy of exit and inner sites are no longer independent. Formally, the chemical potential of a ThPCA obeying to Eq. (A.1) can be evaluated through the *insertion formula*

$$e^{-\beta\mu} = \frac{Q(n+1)}{Q(n)} \frac{p(n)}{p(n+1)}, \qquad n = 0, \ldots, K-1 \tag{4.12}$$

or, equivalently through the *deletion formula*

$$e^{-\beta\mu} = \frac{Q(n)}{Q(n-1)} \frac{p(n-1)}{p(n)}, \qquad n = 1, \ldots, K \tag{4.13}$$

where the partition function Q is calculated as described in Section 2.5, and the distribution $\mathbf{p}$ is *exact* (e.g. as derived from the mean-field relation in Eq. (A.1)). That makes the two equations (4.12) and (4.13) perfectly equivalent. However, if instead the output of a (long but finite) N-constant simulation is used to compute the occupancy distribution $\mathbf{p}_N$, then due to numerical error $\mathbf{p}_N \approx \mathbf{p}$ and not only the chemical potential estimations calculated through Eqs. (4.12) and (4.13) will be different, but also both formulas will return different values depending on the occupancy n. Although the difference is small for a sufficiently long simulation of a large system (e.g. 10^5 time steps for a $16 \times 16 \times 16$ system), a proper criterion is needed

to obtain one average value of chemical potential. By averaging Eq. (4.12) over all allowed occupancies one obtains

$$\left\langle e^{-\beta\mu}\right\rangle_{\text{ins}} = \sum_{n=1}^{K} \frac{Q(n)}{Q(n-1)} p_N(n-1), \tag{4.14}$$

where 'ins' stands for an hypothetical 'insertion' move, while the expression resulting from averaging of Eq. (4.13) is

$$\left\langle e^{-\beta\mu}\right\rangle_{\text{del}} = \sum_{n=0}^{K-1} \frac{Q(n+1)}{Q(n)} \frac{p_N^2(n)}{p_N(n+1)}, \tag{4.15}$$

where 'del' stands for 'deletion'. The average chemical potential estimations related with Eqs. (4.14) and (4.14) are $\langle\mu\rangle_{\text{ins}} = -\beta^{-1}\ln\langle e^{-\beta\mu}\rangle_{\text{ins}}$ and $\langle\mu\rangle_{\text{del}} = -\beta^{-1}\ln\langle e^{-\beta\mu}\rangle_{\text{del}}$. Both formulas will produce the correct chemical potential for intermediate loadings. For small loadings instead Eq. (4.15) becomes unappropriate since empty cells (allowing no deletion moves) become frequent, whereas for high loadings the probability of saturated cells (allowing no insertion moves) increases making Eq. (4.14) inadequate. Therefore the proper choice is

$$\left\langle e^{-\beta\mu}\right\rangle = \sum_{n=1}^{K} \frac{n}{K} \frac{Q(n)}{Q(n-1)} p_N(n-1) + \sum_{n=0}^{K-1} \left(1 - \frac{n}{K}\right) \frac{Q(n+1)}{Q(n)} \frac{p_N^2(n)}{p_N(n+1)}, \tag{4.16}$$

giving the average chemical potential $\langle\mu\rangle = -\beta^{-1}\ln\langle e^{-\beta\mu}\rangle$, which embeds weighted insertion and deletion together. The prefactors $\frac{n}{K}$ and $(1-\frac{n}{K})$ take account of the fact that a less occupied cell would preferably undergo a deletion rather than an insertion, and viceversa for a more occupied cell. For example, in an hypothetical deletion mechanism a site would be randomly picked so that the probability of the deletion move to be performed would be $\frac{n}{K}$. Likewise, in an hypothetical insertion move the probability to randomly pick an empty site to put one additional guest in would be $(1-\frac{n}{K})$.

A result which will be extensively discussed in Chapter 6 is anticipated here in order to show the difference among Eqs. (4.14), (4.15), and (4.16): in Figure 4.1 the chemical potential from ThPCA simulations of a system emulating the equilibrium properties of xenon adsorption in NaA zeolite taken from measurements by Jameson *et al.*[76,77] is shown as derived by using Eqs. (4.14), (4.15), and (4.16), and compared with the mean-field value (obtained from Eq. (A.1) which will be introduced in Appendix A).

4.4 Diffusion coefficients

In numerical simulations of the ThPCA, diffusion coefficients which are of most common interest in the study of the intercell migration in microporous systems have been evaluated. They are: (i) the *self* diffusivity, related to the motion of a tagged guest particle in the system; (ii) the *collective* (or Maxwell-Stefan) diffusivity, related to the motion of the center-of-mass of the diffusing species,[31] and (iii) the *chemical* (or Fick) diffusivity, which measures the transport of mass and the decay of density fluctuations in the system. The latter two coefficients are related one to each other by means of a purely thermodynamic factor.

4.4.1 Self-diffusion coefficient

The *self diffusion coefficient* measures the diffusive motion of a single guest. For a continuous system, it reads

$$D_s = \frac{1}{2d} \lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t} \left\langle [\mathbf{r}_I(t) - \mathbf{r}_I(0)]^2 \right\rangle , \tag{4.17}$$

where $\mathbf{r}_I(t)$ is the position of the I-th tagged guest at time t, and d is the dimensionality of the system. The Green-Kubo relation links the self-diffusivity with the velocity autocorrelation function (VACF) according to:

$$D_s = \frac{1}{d} \int_0^\infty \langle \mathbf{v}_I(t) \cdot \mathbf{v}_I(0) \rangle \mathrm{d}t, \tag{4.18}$$

where $\mathbf{v}_I(t)$ indicates the instantaneous velocity of the I-th guest molecule at time t. It will result that, for a discrete system Eq. (4.17) is still valid (i.e. the long-time slope of the mean-square displacement is still $2d$ times the self-diffusivity) while the discrete counterpart of the Green-Kubo relation will appear slightly different from Eq. (4.18). In order to define the trajectory of a guest in the discrete space of the automaton, it will be very useful to introduce the *instantaneous cell-to-cell displacement* of a single guest:

$$\delta\mathbf{r}_I(t) = \lambda \sum_{j=1}^{2d} \varphi_{Ij}(t)\mathbf{e}_j, \tag{4.19}$$

where $\boldsymbol{\varphi}_I(t) = \{\varphi_{Ij}(t)\}_{j=1,\ldots,2d}$ is a boolean vector satisfying the following property:

> $\varphi_{Ij}(t) = 1, 0$ whether the guest I undergoes or not propagation from its current cell, say $\mathbf{r}$, to its j-th neighboring cell, i.e. $\mathbf{r}^j$, at time t.

Its total displacement at time t (i.e. after $Z = t/\tau$ time steps) is given by

$$\Delta\mathbf{r}_I(t) = \sum_{z=0}^{Z-1} \delta\mathbf{r}_I(z\tau), \tag{4.20}$$

and its mean square displacement is given by

$$\langle\,[\Delta\mathbf{r}_I(t)]^2\rangle = \left\langle \sum_{z=0}^{Z-1}\sum_{z'=0}^{Z-1} \delta\mathbf{r}_I(z\tau)\cdot\delta\mathbf{r}_I(z'\tau)\right\rangle\,, \tag{4.21}$$

where z, z' denote time steps. By expanding the sum in Eq. (4.21) and using the Einstein relation for the self-diffusivity, which for a discrete-time evolution reads

$$D_s = \frac{1}{2d}\lim_{t\to\infty}\frac{\langle\,[\Delta\mathbf{r}_I(t+\tau)]^2\rangle - \langle\,[\Delta\mathbf{r}_I(t)]^2\rangle}{\tau}, \tag{4.22}$$

one obtains the self-diffusion coefficient for the ThPCA:

$$D_s = \frac{1}{2d\tau}\left[\langle\delta\mathbf{r}_I(0)\cdot\delta\mathbf{r}_I(0)\rangle + 2\sum_{z=1}^{\infty}\langle\delta\mathbf{r}_I(z\tau)\cdot\delta\mathbf{r}_I(0)\rangle\right], \tag{4.23}$$

where the contribution at equal times has been separated from the correlations between different times. The quantity $\langle\delta\mathbf{r}(t)\cdot\delta\mathbf{r}(0)\rangle$ in Eq. (4.23), termed *displacement autocorrelation function* (DACF), is used as the discrete conterpart of the continuous VACF. The DACF is extremely important in the present model since the entity of correlations in the self-motion caused by the jump randomization changes remarkably the shape of $\langle\delta\mathbf{r}(t)\cdot\delta\mathbf{r}(0)\rangle$ *vs.* t.

4.4.2 Collective diffusion coefficient

The collective diffusivity for a continuous system is defined as

$$D_c = \frac{1}{2d}\lim_{t\to\infty}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle\left\{\sum_{I=1}^{N}[\mathbf{r}_I(t)-\mathbf{r}_I(0)]\right\}^2\right\rangle, \tag{4.24}$$

and the related Green-Kubo formula reads

$$D_c = \frac{1}{Nd}\int_0^{\infty}\left\langle\left[\sum_{I=1}^{N}\mathbf{v}_I(t)\cdot\sum_{I=1}^{N}\mathbf{v}_I(0)\right]\right\rangle\mathrm{d}t, \tag{4.25}$$

where by expanding the sums the velocity crossed-correlation function (VCCF), $\langle \mathbf{v}_I(t) \cdot \mathbf{v}_J(0) \rangle$ plays the same role the VACF played in Eq. (4.18) for the self-diffusivity. Also in that case it is possible to formulate the same quantities for a discrete system: the Einstein relation for collective diffusion becomes

$$D_c = \frac{N}{2d} \lim_{t \to \infty} \frac{\langle \left[\Delta \mathbf{r}_{\mathrm{CM}}(t+\tau) \right]^2 \rangle - \langle \left[\Delta \mathbf{r}_{\mathrm{CM}}(t) \right]^2 \rangle}{\tau}, \tag{4.26}$$

where $\Delta \mathbf{r}_{\mathrm{CM}}(t) = \mathbf{r}_{\mathrm{CM}}(t) - \mathbf{r}_{\mathrm{CM}}(0)$, with $\mathbf{r}_{\mathrm{CM}}(t)$ as the coordinates of the center-of-mass at the time t. Eq. (4.26) can be expanded as a function of the *displacement crossed-correlation function* (DCCF) $\langle \delta \mathbf{r}_I(t) \cdot \delta \mathbf{r}_J(t') \rangle$ where $\delta \mathbf{r}_I(t)$ is the instantaneous cell-to-cell displacement function, Eq. (4.19), for the i-th guest, obtaining

$$\begin{aligned} D_c =& D_0 + \frac{1}{d\tau} \sum_{z=1}^{\infty} \langle \delta \mathbf{r}(z\tau) \cdot \delta \mathbf{r}(0) \rangle \\ &+ \frac{1}{2d\tau N} \sum_{\substack{1 \le I,J \le N \\ I \ne J}} \left\{ \langle \delta \mathbf{r}_I(0) \cdot \delta \mathbf{r}_J(0) \rangle + 2 \sum_{z=1}^{\infty} \langle \delta \mathbf{r}_I(z\tau) \cdot \delta \mathbf{r}_J(0) \rangle \right\}, \end{aligned} \tag{4.27}$$

where the zero-time diffusivity has been indicated as

$$D_0 = \frac{1}{2d\tau} \langle \delta \mathbf{r}(0) \cdot \delta \mathbf{r}(0) \rangle. \tag{4.28}$$

In Eq. (4.27) all contributions arising from both self- and mixed-correlations are present: the sum of the first two terms gives the self-diffusivity, and the last sum incorporates all contributions arising from correlations between the displacements of all possible pairs of guests in the system. Details about the derivation of Eq. (4.27) can be found in Appendix D.1.

4.4.3 Chemical diffusion coefficient

Finally the chemical diffusion coefficient D_{chem} (sometimes reported as *transport diffusivity*) measures the transport of mass and the decay of density fluctuations in the system, and it is related to the collective diffusivity by[78]

$$D_{\mathrm{chem}} = \frac{\partial \ln f}{\partial \ln \langle n \rangle} D_c \tag{4.29}$$

where the derivative is called *thermodynamic factor*, and f is the fugacity as given in Eq. (4.4). In Section A.1 the loading will be related to the chemical

potential via the following relation in the grand-canonical ensemble:

$$\langle n \rangle = \frac{\sum_{n=0}^{K} nQ(n)e^{\beta\mu n}}{\sum_{n'=0}^{K} Q(n')e^{\beta\mu n'}} \tag{4.30}$$

where $Q(n)$ is the canonical partition function of a closed, n-occupied cell. Eq. (4.30) can be used to show that the thermodynamic factor is indeed equal to

$$\frac{\partial \ln f}{\partial \ln\langle n \rangle} = \frac{\langle n \rangle}{\langle n^2 \rangle - \langle n \rangle^2}, \tag{4.31}$$

(where f is the fugacity) i.e. the reciprocal of the reduced variance.

4.5 Other time-dependent properties

From the point of view of the guest species, the dynamical evolution of a host-guest system is described by the migration process, where each guest molecule travels inside the confining system, leaving a pore to enter a neighboring one and so on. Diffusivities and DACF provide an exhaustive description of such a process.

From the point of view of the host, the state of each adsorption unit evolves by exchanging guest molecules with the surrounding units. As a consequence, the cell occupancy will change with time. Although such a process is stochastic (which is obvious since the evolution rules are probabilistic) the system heterogeneity and the hierarchic character of the involved processes will cause it to be affected by correlations. Moreover, the statistical properties related to the time evolution of the occupancy of the cells will turn out to be non-trivial. Therefore, a statistical analysis of the time-dependency of cell occupancies will be very useful to supplement the picture of the migration process given in the previous section. With this aim, the *mean life times* and the decay of fluctuations in the local density will be used as statistical tests to investigate such stochastic process.

Mean life time. In units of time steps, the *mean life time* (MLT) of the occupancy n is defined as the average number of consecutive time steps during which a cell persists in the occupancy n. It should be noted from the above definition that MLT is *not* the average extent of time a cell retains the same n guests inside of it, since it does not embed any information about guests' identity.

Just as the diffusivity, the mean life time of occupancies at different conditions of loading is very closely related to the adsorption properties of the

host-guest system and is very much affected by the system heterogeneity, which leads some occupancies to persist for a considerably higher time than some other ones.

Decay of local density fluctuations. In the present model, the local density fluctuation time-autocorrelation function (LDFACF)[79] can be defined as

$$C(t) = \langle \delta n(0)\, \delta n(t) \rangle \,, \tag{4.32}$$

where $\delta n(t) = n(t) - \langle n \rangle$. Such a correlation function gives a measure of the way a portion of the system loses memory of its previous states. Although local, it is a collective property since it depends on the moves of all guests while they are entering and leaving every cell of the system. Therefore, correlations in the self-migration processes of each guest will enter $\langle \delta n(0)\, \delta n(t) \rangle$ in some extent, together with the correlations between different guests occupying the same cell. As a result, the LDFACF may show an unexpected complex decay. In this sense, the ThPCA represents a very easy-to-hand and reliable tool to investigate the nature of collective correlation phenomena on a purely statistical mechanical approach.

Chapter 5

Non-interacting case: numerical simulations

In the present Chapter the possible behaviors of the model in the absence of guest-guest interactions shall be illustrated. In fact, even in this special case the flexibility of the ThPCA parameter space allows to effectively capture the qualitative phenomenology of adsorption and diffusion in zeolites. Here the heterogeneity of the system is represented only through the differentiation in statistical weight between exit and inner sites (lattice-gases with two non-equivalent sites were extensively studied by Chvoj *et al.*,[46,80] Tarasenko *et al.*,[47,81,82] and by Bhide *et al.*[28,29] through KMC simulations). As a result of choosing $f^o_{\text{ex}} \neq f^o_{\text{in}}$, different configurations will be differently weighted depending on the temperature T. Non-trivially loading dependent diffusion profiles will emerge as a consequence of the fact that the temperature controls the accessibility of the exit sites, therefore influencing the frequency of transfers.

In this Chapter a test cell shall be considered which is characterized by $K_{\text{ex}} = 6$ exit and $K_{\text{in}} = 10$ inner sites. Such a cell is modeled as coarse-grained representation of a α-cage of the zeolite ZK4 (which has a saturation limit of about $15 - 16$ molecules per cage[32,41,59,63,64]). The simulations described in this Chapter have been performed by using the memoryless randomization operator, extensively described in Section 3.2.1. In the absence of guest-guest interactions, the global propagation operation is performed according to the procedure illustrated in Section 3.4.1. Unless specified otherwise, the factor κ_0 introduced in Eq. (3.33) is given a value of 1. Due to the memoryless randomization, the random walks of the guests can be treated, with a good approximation, as independent Markov processes. Time correlations in the self-motion and between different guests are so weak that all types of diffusivity, i.e. D_s, D_c, and D_{chem} scale linearly with κ_0.

5.1 Static equilibrium properties

The competitive adsorption of guest molecules at different sites fully characterizes the system's properties. Following the lines of Bhide *et al.*[28,29] the energy parameters has been set up as $f^o_{\rm ex} = -10$ kJ mol^{-1} and $f^o_{\rm in} = -20$ kJ mol^{-1}. As mentioned in the Introduction, the choice of differentiating two types of site into each cell arises makes contact with the experimental evidence of two different types of adsorption locations in various zeolites,[43] or more generally, of n types of sites differing in their ability to bind a guest species.[42] To describe equilibrium adsorption in zeolites the Langmuir model[83] is often used, in which the guests do not interact except by excluding each other from the adsorption sites, each being capable of holding at most one molecule. Moreover, while in the simple Langmuir model all sites are equal, the introduction of different types of sites leads to different equilibria corresponding to adsorption processes in pores of different binding energy. Since each equilibrium can be represented by a single Langmuir isotherm, the adsorption isotherm for n simultaneous equilibria is well reproduced by a n-site Langmuir isotherm.[42] Such an adsorption model has been widely used to successfully interpretate adsorption data.[84] Its efficiency has been also validated by MD and Grand-Canonical Monte Carlo (GCMC) simulations,[44] and used in KMC models[31] to take into account the loading dependence of the self-diffusion coefficient in the study of diffusion in zeolites. Along with the different abilities assigned to the two kinds of sites in moving guests in and out of the confining cell, the different adsorption energies mimic the real situation in which different adsorption locations in the cavity influence the ways a molecule has access to a window (to reach an exit site, in the Th-PCA language) and then migrates into the adjacent cavity. Simulations have been performed at several temperatures. The data for the case of $T \to \infty$ refer to the homegeneous situation where exit and inner sites are statistically equivalent $f^o_{\rm ex} = f^o_{\rm in}$.

5.1.1 Equilibrium distributions

The equilibrium probability distributions $\{P(n_{\rm ex}|n)\}$ and $\{p_N\}$, defined respectively in Eqs. (2.29) and (4.1), have been studied. In Figure 5.1 the distribution $\mathbf{p}_N$ is plotted for several temperatures at integer loadings, together with a strictly statistical distribution (hypergeometric distribution, $\mathbf{p}^{\rm hyp}$)[58] which assumes that the sorbed species occupy *equivalent*, mutually exclusive lattice sites in the cells:

$$p^{\rm hyp}(n) = \binom{K}{n}\binom{K(M-1)}{N-n}\Big/\binom{MK}{N}. \tag{5.1}$$

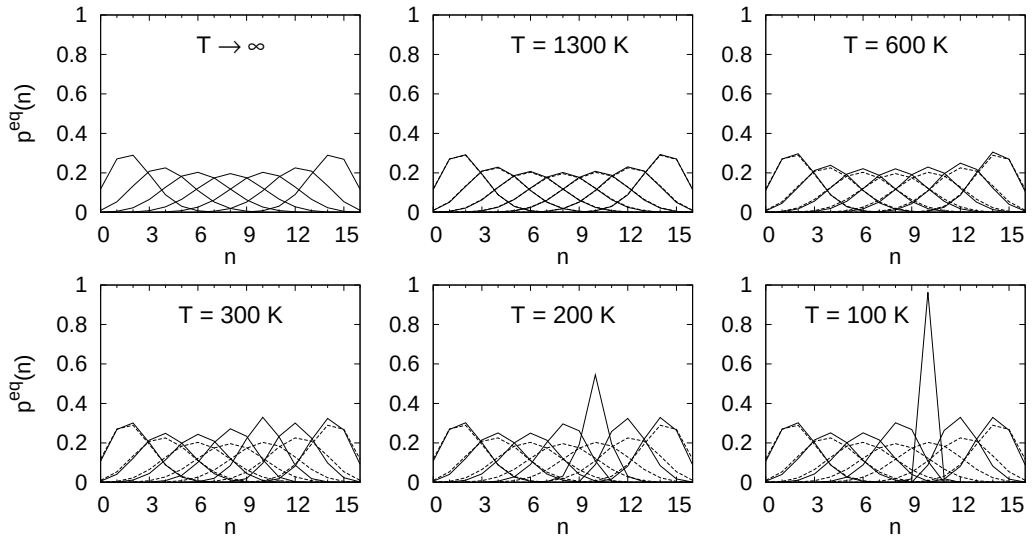

Figure 5.1: The occupancy distributions f^{eq} (straight lines) computed from Eq. (4.1) for loadings $\langle n \rangle = 1, \ldots, 15$ at various temperatures are shown in comparison with the hypergeometric distributions f^{hyp} (dashed lines).

Due to the presence of two energetically different sorption sites, the computed distributions differ from the hypergeometric distribution. Although the maximum of each curve is centered around the loading value, the fluctuations of the occupancy n of each cell around this value become temperature-sensitive; accordingly, by increasing the temperature the distinction between the hypergeometric distributions and the calculated ones becomes less evident, because the energy difference between inner and exit sites becomes less important and the excluded volume effect tends to prevail (details about the convergence of the occupancy distribution $\mathbf{p}_N$ to the hypergeometric $\mathbf{p}^{\mathrm{hyp}}$ at high temperature can be found in Appendix B.1). The same energy difference is responsible of the behavior of the system at very low temperature.

As T increases, accessibility of exit sites also increases and higher values of n_{ex} become possible. From observation of Figure 5.1 it is interesting to observe on what grounds the measured distributions cannot be reproduced by the hypergeometric ones. A symmetry relation between the hypergeometric curves exists. In the present case, this means that the distributions $\mathbf{p}^{\mathrm{hyp}}$ calculated at the loadings $\langle n \rangle$ and $K - \langle n \rangle$ are mirror images. The calculated distribution deviates from the hypergeometric in the following ways: decreasing the temperature causes the curves no longer to be related by symmetry, and an increasing peak for $\langle n \rangle = 10$ appears. In the limit of zero Kelvin two well-defined regions emerge. The first one is a set of hyper-

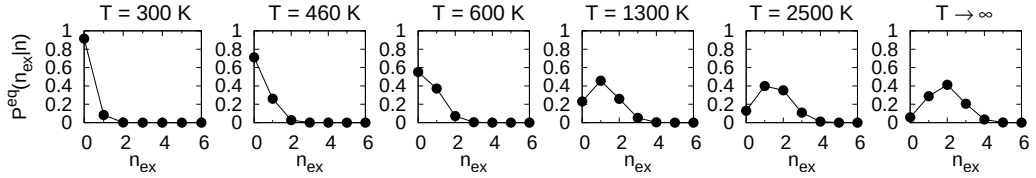

Figure 5.2: Conditional probability $P^{\mathrm{eq}}(n_{\mathrm{ex}}|n)$ for $n = 5$ at various temperatures. For low temperatures the guests tend to occupy preferably the inner sites, so $P^{\mathrm{eq}}(n_{\mathrm{ex}}|n)$ is significant only for the lowest possible values of n_{ex}. Therefore the accessibility of the exit sites is low. The temperature-sensitiveness of the single cell is well represented by the behavior of the conditional probability $P(n_{\mathrm{ex}}|n)$ w.r.t. temperature.

geometric distributions going from loading zero up to ten, while the second is again a set of hypergeometric distributions going from loading ten up to sixteen (details about the convergence of the occupancy distribution $\mathbf{p}_N$ to two hypergeometrics $\mathbf{p}_{\mathrm{in}}^{\mathrm{hyp}}$ and $\mathbf{p}_{\mathrm{ex}}^{\mathrm{hyp}}$ in the limit of zero Kelvin can be found in Appendix B.1). The boundary between the two sets is located at loading 10 because $f_{\mathrm{in}}^{o} < f_{\mathrm{ex}}^{o}$: In the limit of zero Kelvin the first region includes the set of the equilibrium distributions on the ten equivalent inner sites while the second includes the distributions on the six energetically higher exit sites. Though the interaction forces between the guest particles have been completely neglected, this example proven how an asymmetry in the energy of the sorption sites is enough to deeply modify the molecular distribution over the zeolitic cavities markedly expanding the temperature effects.

In Figure 5.2 the distribution $P(n_{\mathrm{ex}}|n)$ (defined in Eq. (2.29)) is plotted at various temperatures for the specific occupancy $n = 5$. At $T = 300$ K, $P(n_{\mathrm{ex}}|n)$ is significant only for the lowest possible values of n_{ex} because the guests tend to occupy preferably the inner sites so that the states with high values of n_{ex} are rarely populated. Under such conditions the exit sites are poorly accessible. The situation will change by increasing the temperature, until $T \to \infty$ where the distribution $P(n_{\mathrm{ex}}|n)$ reaches the hypergeometric form $\binom{K_{\mathrm{ex}}}{n_{\mathrm{ex}}}\binom{K_{\mathrm{in}}}{n_{\mathrm{in}}}/\binom{K}{n}$.

As can be seen from Figure 5.3, by lowering the temperature one reduces the thermodynamical tendency of a cell to accept a new guest. The non-linear trend of the reduced variance for $T < \infty$ reflects the difference in thermodynamic properties between the two types of site in the system.

Negligible correlations in space. A remarkable property of the model is that, on sufficiently large grids and far from critical conditions, the spatial correlations are negligible. To evaluate the degree of static correlations among neighboring cells, away from the phase transition (see next Sec-

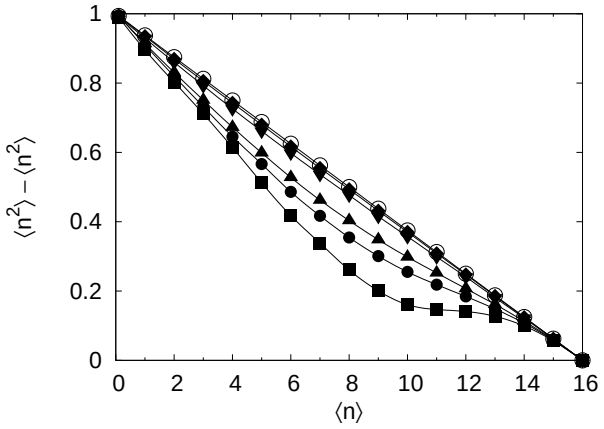

Figure 5.3: Reduced variance as a function of loading at various temperatures.

tion), the static structure factor S (not reported) defined as[85] $S(\mathbf{r}, \mathbf{r}') = \langle \delta n(\mathbf{r}) \delta n(\mathbf{r}') \rangle$, where the occupancies $n(\mathbf{r})$ and $n(\mathbf{r}')$ are evaluated at the same time step, has been calculated from simulations. Following Jameson[86] another simple way to keep track of spatial correlations is to compute the probability $g(m, n)$ of finding a cell of occupancy n in the neighborhood of a cell of occupancy m. Evaluating S and the distribution g for a large number of configurations near equilibrium, it has been found that $S(\mathbf{r}, \mathbf{r}') \approx 0$ for $\mathbf{r} \neq \mathbf{r}'$, and $g(m, n) \approx p_N(n)$ for all values of m, n. Therefore, spatial correlations are negligible and the occupancies are distributed approximately in the same way both locally and globally.

5.1.2 Diffuse phase transition

When $f^o_{\text{in}} \neq f^o_{\text{ex}}$ the system can be formally divided in two subsystems.[46] The first one is the subsystem of exit sites, characterized by the energy f^o_{ex}. The second one is the subsystem of inner sites, characterized by the energy f^o_{in}. When the term $|f^o_{\text{ex}} - f^o_{\text{in}}|/k_B T$ increases, then the exchanges between the two subsystems become increasingly difficult. Let us consider the case in which the inner sites are the most binding, i.e. $f^o_{\text{in}} < f^o_{\text{ex}}$. At sufficiently low temperature and below some particular loading $\langle n \rangle^*$, the guests will mostly occupy the inner sites, causing the system to behave predominantly as an isolated subsystem of inner sites. Around some particular loading (indicated as $\langle n \rangle^*$) almost all the inner sites are occupied. Above $\langle n \rangle^*$ the guests begin to fill the exit sites and, because the subsystem of inner sites is almost saturated, a transition between the two phases occurs such that the system starts to behave predominantly as the isolated subsystem of exit sites.

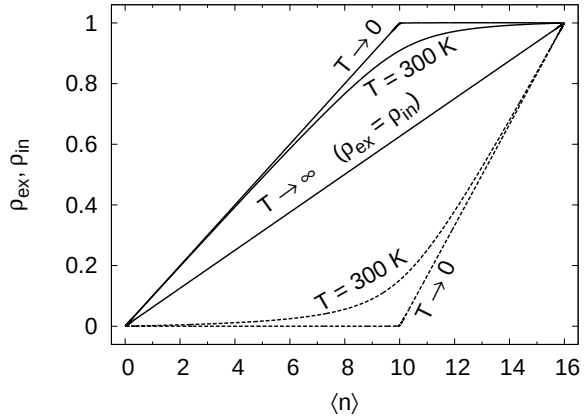

Figure 5.4: Partial loading ρ_{ex} of the subsystem of exit sites (solid line), and partial loading ρ_{in} (dashed line) of the subsystem of inner sites, plotted together w.r.t. the total loading $\langle n \rangle$ at various temperatures. For very high temperatures ($T \to \infty$) they increase in the same way because the two subsystems are equivalent. For low temperatures the two subsystems become very different, therefore the partial loadings increase with $\langle n \rangle$ in different ways.

The sharp change in the properties of the system denotes a phase transition. At *very* low temperature ($T \to 0$) the phase transition occurs exactly at the loading $\langle n \rangle^* = K_{\text{in}}$, where the thermodynamic properties of the system undergo a net change: in such a case the transition can be defined as a *point transition*.[46] But, if the temperature is *not* very low then the change in properties around $\langle n \rangle^*$ is smooth and one cannot locate the value of $\langle n \rangle^*$ except that it falls within a loading interval

$$K_{\text{in}} - \delta < \langle n \rangle^* < K_{\text{in}} + \delta,$$

with $\delta > 0$. Such a phenomenon is reported as *diffuse phase transition*.[46] Decreasing the temperature, δ diminishes and the transition becomes more neat. It has been observed how the phase transition takes place first calculating the chemical potential at two temperatures $T = 300$ K and $T = 100$ K.

Such observables are reported in Figure 5.4 as functions of the loading at various temperatures, in order to visualize how the system splits in two subsystems: (i) for $T \to \infty$ the fractions ρ_{ex} and ρ_{in} increase in the same way (i.e., as $\rho = \langle n \rangle / K$); (ii) for $T < \infty$, the two fractions are equal only at $\langle n \rangle = 0$ and K, while for all the other loadings $\rho_{\text{ex}} < \rho_{\text{in}}$ since the guests occupy preferably inner sites; (iii) for $T \to 0$, we have $\rho_{\text{ex}} = 0$ until ρ_{in} reaches the value of 1 at loading K_{in}, then ρ_{in} stops varying and ρ_{ex} increases from 0 to 1.

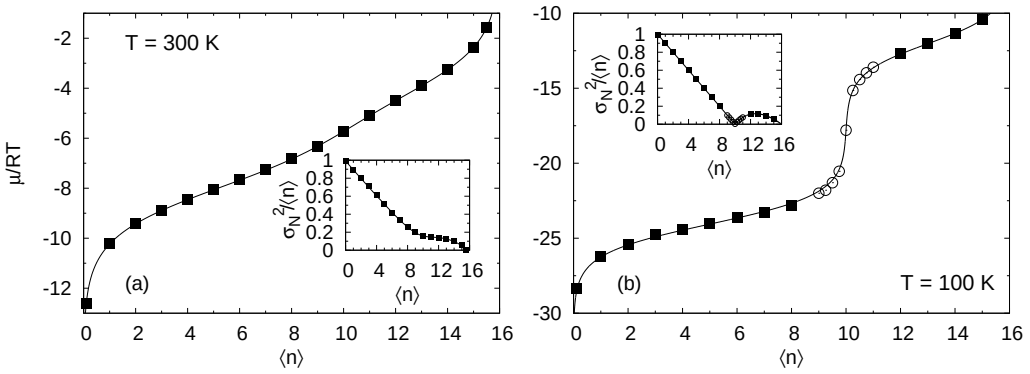

Figure 5.5: The chemical potential and (in the insets) the reduced variance for $T = 300$ K (top) and $T = 100$ K (bottom). The simulations were performed on a grid of 32^3 cells for running times of 10^5 time steps (black squares) and of 10^8 time steps (white circles). Solid lines are fits through Eq. (A.12). μ is in units of kJ mol^{-1}.

As can be seen in Figure 5.5, the phase transition at low temperature produces a singularity in the plot of μ *vs* $\langle n \rangle$ (i.e., the inverse plot of the adsorption isotherm). Since the variance and the chemical potential are connected by Eq. (B.9) (see Appendix B.1), in the same conditions the reduced variance of f^{eq} exhibits a cusp.

This behavior is a feature of all systems having two (or more) distinct adsorption locations. Of course, the inverse plot of Figure 5.5 is the dual-site Langmuir isotherm[87] of Eq. (A.12). In general, a system having n different adsorption sites will separate, at low temperatures, into n subsystems, and increasing the loading from zero to the saturation limit it will undergo $n - 1$ diffuse phase transitions.

In Figure 5.6 the energy distribution $P(E^{\text{sys}})$ (that is, the probability of finding the entire system in the energy E^{sys}) is reported for differently long simulations at low temperature. As can be seen, in the proximity of a phase transition the ThPCA manifests its inherently noisy nature, and averaging over this statistical noise requires massive calculations.[37]

Since the global system is canonical, one obtains the molar specific heat per cell at constant volume, $C_{V,m}$, from the energy distribution $P(E^{\text{sys}})$. Indicating the variance of $P(E^{\text{sys}})$ per cell as

$$\sigma_E^2 = \frac{1}{M} \sum_{E^{\text{sys}}} \left(E^{\text{sys}} - \langle E^{\text{sys}} \rangle\right)^2 P(E^{\text{sys}}), \tag{5.2}$$

if E^{sys} is in units of J mol^{-1} then $C_{V,m} = \sigma_E^2 / RT^2$ where R is the constant of gas in units of J mol^{-1}K^{-1}. The shape of the specific heat as a function of loading is shown in Figure 5.7. At $T = 100$ K a sharp $C_{V,m}$ peak appears

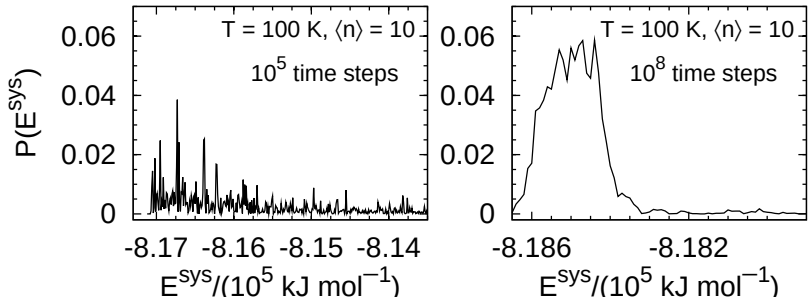

Figure 5.6: The energy distributions around the phase transition ($T = 100$ K, $\langle n \rangle = 10$) for a system with $M = 16^3$ cells. The distribution was computed averaging over 10^5 time steps in (a) and over 10^8 time steps in (b) (see text for further details).

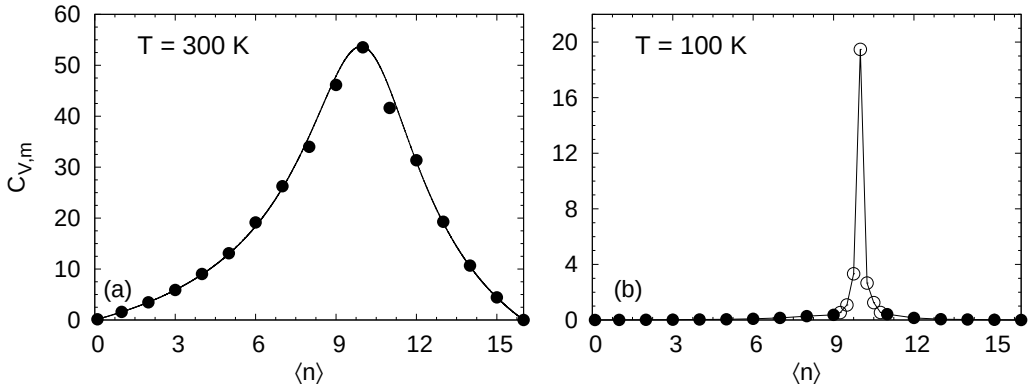

Figure 5.7: The molar specific heat per cell for $T = 300$ K (top) and $T = 100$ K (bottom). Black squares and white circles indicate simulations 10^5 and 10^8 time steps long respectively.

at $\langle n \rangle = K_{\text{in}}$. Its width is expected to further decrease and its height to increase if $T \to 0$. This behavior of the specific heat confirms the existence of a first-order diffuse phase transition in the system.

5.2 Transport properties

In each computer simulation, starting from a random distribution of guests, after 20 000 time steps of equilibration the evolution of the system has been observed during a time interval ranging from 10^5 to 10^8 time steps, depending on the statistical accuracy required to compute the averages of interest. The data at temperature $T \to \infty$ are obtained from simulations with $f^o_{\text{ex}} = f^o_{\text{in}}$.

5.2.1 Diffusion coefficients

A useful quantity in the discussion of the diffusion properties of the model is the *exit sites accessibility* defined for a cell of occupancy $n \geq 1$ as the ratio between $\langle n_{\mathrm{ex}}(n) \rangle$ (the average number of filled exit sites in an n-occupied cell) and the occupancy n. Such a quantity is independent of the loading $\langle n \rangle$, and measures the thermodynamic tendency of a guest to reach an exit site according to the number of guests which occupy the same host cell. In formula,

$$\langle n_{\mathrm{ex}}(n) \rangle = \sum_{n_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}} n_{\mathrm{ex}} P(n_{\mathrm{ex}}|n), \tag{5.3}$$

where $P(n_{\mathrm{ex}}|n)$ has been defined in Eq. (2.29) as the conditional probability of a cell to have n_{ex} filled exit sites given that its occupancy is n. The average accessibility is defined as $K_{\mathrm{ex}}\rho_{\mathrm{ex}}/\langle n \rangle$. The accessibility has been plotted for several temperatures in Figure 5.8. Because the inner sites are the most binding, at finite T it will increase with n.

At $T = 300$ K, for low values of n the probability of a guest to reach an exit site is very low and increases not much with increasing the occupancy. Instead, it increases rapidly in the range $K_{\mathrm{in}} < n < K$ because in such cases at least $n - K_{\mathrm{in}}$ exit sites must be occupied. In the limit of maximum occupancy $n = K$ the cell is saturated, and obviously this gives an accessibility value of K_{ex}/K, which is independent of temperature.

Increasing the temperature, the accessibility becomes less occupancy-sensitive and in the limit of $T \to \infty$ (which is equivalent to the case of

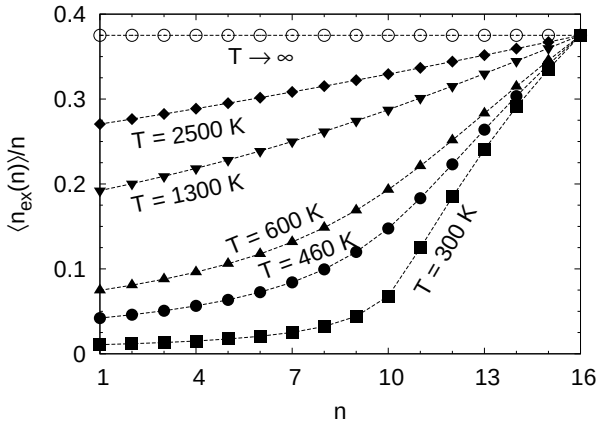

Figure 5.8: Exit sites accessibility for several temperatures.

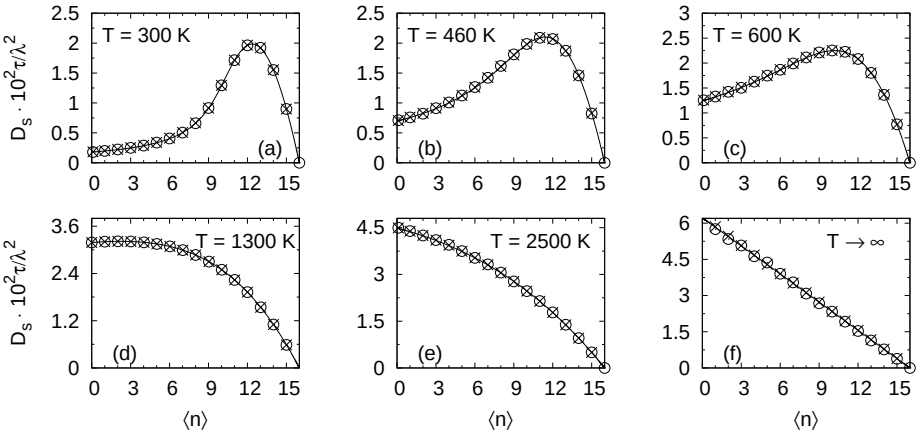

Figure 5.9: Self diffusivity (white circles) obtained from simulations at various temperatures, plotted together with D_0 obtained from measurement of the mean residence time of a guest in a cell (crosses, see Eqs. (A.9), (A.8) and (A.10) in Appendix A, devoted to the mean-field formulation of the model) and through Eq. (A.15) (solid line).

$f^o_{\rm ex} = f^o_{\rm in}$) the accessibility is constant at $K_{\rm ex}/K$, because the sites are all equivalent.

The use of the memoryless randomization operator has two important consequences on the diffusivity:

- The self diffusivity D_s approximates well the self-diffusivity of an uncorrelated random walk D_0 defined in Eq. (4.28). This can be seen in the good overlapping of the profiles of D_s and D_0 *vs* $\langle n \rangle$ for various temperatures as shown in Figure 5.9.

- As a consequence of this fact, if correlations *among different guests* exist they are *instantaneous*, that is, relative to the same time step. Due to the memoryless randomization D_c and D_s are very similar, meaning that the correlations between the random walks of different guests are negligible.

In Figure 5.10 several profiles of (a) D_s, (inset of (b)) D_c, and (b) $D_{\rm chem}$ are shown. They will be discussed in the next Section.

Self-diffusivity. In Figure 5.10a the behavior of D_s *vs* loading is reported for various temperatures. Changes in temperature lead to different profiles of the self-diffusivity, corresponding to the I, II, IV and V types observed by Kärger and Pfeifer[40] (see Figure 5.11) in the PFG-NMR measurements of intracristalline self-diffusion coefficient depending on sorbate concentration.

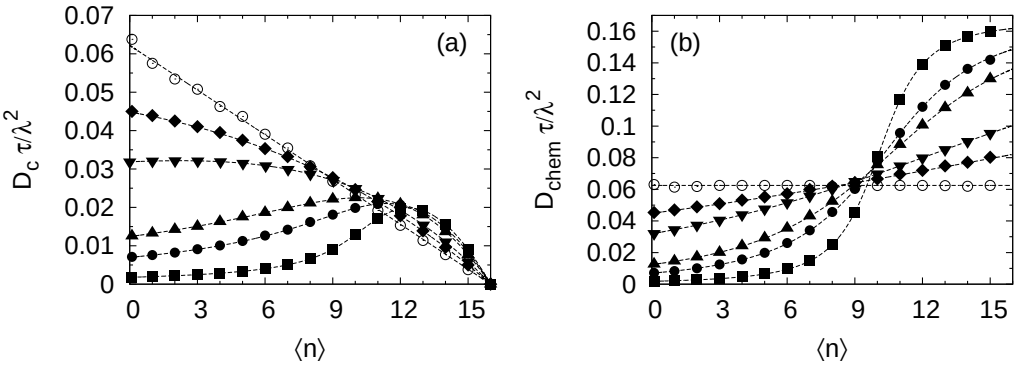

Figure 5.10: (a) Collective diffusion coefficients for various loadings and temperatures. (b) Chemical diffusion coefficients obtained from the collective diffusivity D_c and the thermodynamic factor using Eq. (4.29). $T = 300$ K (squares), $T = 460$ K (black circles), $T = 600$ K (up triangles), $T = 1300$ K (down triangles), $T = 2500$ K (diamonds), $T \to \infty$ K (white circles).

Here two basic events which contribute to the intercell migration process will be listed:

- Event 1: the event in which a guest reaches an exit site during randomization.

- Event 2: the event in which two adjacent exit sites are simultaneously occupied during propagation.

The probabilities of Events 1 and 2 *increase* as the average accessibility of the exit sites increases. If by increasing the loading $\langle n \rangle$ their increasing trend are different, then in general the resulting loading dependence of the diffusivity will not be linear. The curves in Figure 5.10a will be discussed in terms of the balance between Events 1 and 2.

i) The self-diffusivity trend at $T = 300$ K (black squares) will be considered first. From low to intermediate loading the curve shows the increasing-like behavior of type V, as reproduced by the model of Tunca and Ford,[33] because in this range of loadings the probability of Event 1 increases more rapidly than the probability of Event 2. Further increases of $\langle n \rangle$ will cause the diffusivity D_s to increase, until it reaches a maximum. At this point, the probabilities of Events 1 and 2 are balanced: if few guests are removed from the system the diffusivity will decrease due to a decrease of the probability of Event 1; if instead a few guests are added then the diffusivity will decrease because the

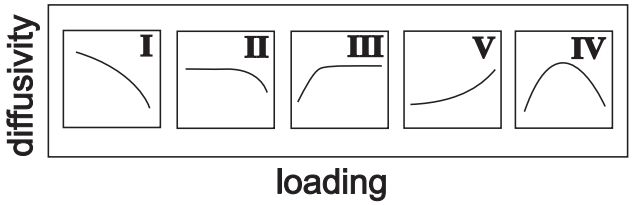

Figure 5.11: The five different profiles of self-diffusivity *vs* loading observed by Kärger & Pfeifer.[40]

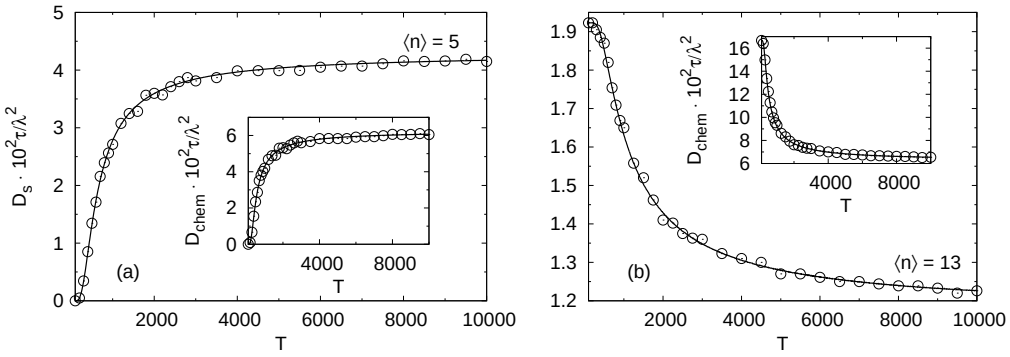

Figure 5.12: Self diffusivity D_s and collective diffusivity D_c (insets) *vs* the absolute temperature T from $T = 100$ K to $T = 10\,000$ K. In (a) the loading is $\langle n \rangle = 5$ and in (b) is $\langle n \rangle = 13$ (see text for details).

average accessibility will become large enough to enhance the probability of Event 2. Therefore, at intermediate-high loadings the diffusivity trend will be of type IV (D_s reaching a maximum and then decreasing to zero). It should be noted that at $T = 300$ K the behavior of D_s *vs* $\langle n \rangle$ is qualitatively analogous to the trend obtained by Demontis *et al.*[15] and by Dubbeldam *et al.*[32] in MD simulations of diffusion of methane in ZK4, and by Coppens *et al.*[9] through Dynamic Monte Carlo simulations of diffusion on a lattice with multiple types of sites.

ii) The same considerations are valid for the self-diffusivity trends at $T = 460$ and 600 K. Since the self-diffusion is ruled by Event 1 at low loadings, and by Event 2 at high loadings, for a fixed value of $\langle n \rangle$ any increase in temperature will cause D_s to increase considerably if $\langle n \rangle$ is low, while if $\langle n \rangle$ is high then D_s will slightly decrease. In order to see a more detailed picture of the effect of temperature on the diffusivity at constant loading, in Figure 5.12 curves of diffusivity *vs* temperature are shown for $\langle n \rangle = 5$ and 13 (the loading which corresponds to the maximum diffusivity falls between these two loadings). Not only the

response of D_s to the increasing temperature is reversed when passing from a condition of low loading to high loading, but one can also see from the trends of Figure 5.10a and from the diffusivity scales of Figure 5.12 that D_s becomes less temperature-sensitive when the system goes toward saturation, where the properties of the model become independent of temperature.

iii) Further increases in temperature will change the diffusion profile. From $T = 600$ K to $T = 2500$ K the trends change from the type discussed before to the type I of Figure 5.11, passing through the type II at $T = 1300$ K, where the probabilities of Events 1 and 2 are balanced for low loadings. At high temperature, say $T = 2500$ K, the system behaves in a way very similar to the case of $T \to \infty$, at which D_s decreases linearly with $\langle n \rangle$ because the sites are all equivalent, so that the accessibility is constant with n (see Figure 5.8) and the only effect controlling the migration process is mutual exclusion.

One can see by direct comparison that the D_s trends presented in Figure 5.10 of this work are similar to the trends found by Bhide *et al.* (see Figure 5 of Ref.[29]) in the study of diffusion of *interacting* guests in a lattice with two non-equivalent sites, although these two models are quite different: indeed, there traditional Monte Carlo lattice-gases are used to study the random walk of guests between neighboring positions in a fully structured lattice of adsorption sites, while in the present model the diffusion phenomenon is produced by random walk of guests *from cell to cell*, where each cell contains a number of adsorption sites without a fully defined spatial arrangement. Even working with non-interacting guests and neglecting both correlation and intracell motion timescale, the distinction between exit and inner sites is a sufficient assumption for reproducing an effect of confinement, giving rise to the typical diffusivity profiles of confined systems.

Collective and chemical diffusion coefficient. Collective and chemical diffusion coefficient, which have been defined in Eqs. (4.26) and (4.29) are reported in Figure 5.10a and 5.10b. It should be noted that, due to the absence of correlations the collective coordinate $N\Delta\mathbf{r}_{\mathrm{CM}}(t)$ (introduced in Eq. (4.26)) diffuses in the same way as the single guests, i.e. $D_c \approx D_s$. As discussed in Section 4.4.3, dividing D_c by the reduced variance produces the chemical diffusivity; therefore the curves in Figure 5.10b shall be discussed while keeping in mind also Figure 5.10b and Figure 5.3.

At infinite temperature the reduced variance and D_c respond to an increase of density exactly in the same way (i.e. they are both linearly decreasing with $\langle n \rangle$). This results in a D_{chem} constant with loading.

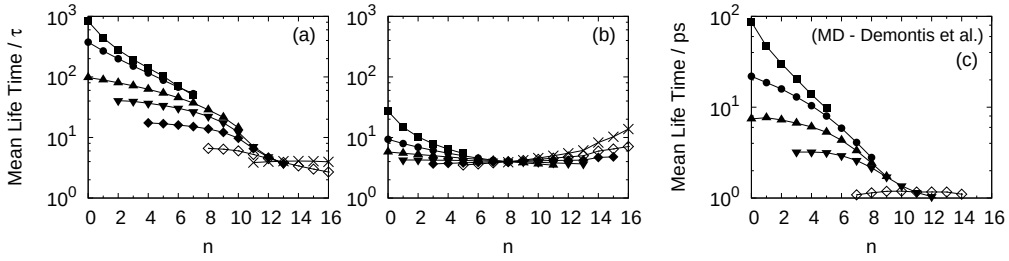

Figure 5.13: Curves of the mean life time of a cell as a function of the occupancy at (a) $T = 300$ K, and (b) $T \to \infty$. Each curve corresponds to a particular loading. $\langle n \rangle = 1$ (squares), $\langle n \rangle = 3$ (circles), $\langle n \rangle = 5$ (up triangles), $\langle n \rangle = 7$ (down triangles), $\langle n \rangle = 9$ (diamond), $\langle n \rangle = 11$ (left triangles), $\langle n \rangle = 13$ (right triangles), $\langle n \rangle = 15$ (crosses). For each loading, data are plotted only for the most probable occupancies (i.e., the occupancies appearing with a probability $> 10^{-3}$). In (c) the mean life times from MD simulations of methane in ZK4 at 360 K performed by Demontis *et al.*[41] are also reported for loadings 1, 3, 5, 7, and 11.

This balance between fluctuations and diffusive properties is broken when the difference in binding ability of the two types of site becomes non-negligible. Lowerings of temperature will reduce the density fluctuations, and the D_{chem} profile will show two distinct diffusive regimes:[46] low diffusivity for low loadings, and high diffusivity for high loadings. Similarly to the case of D_s, increasing the temperature the difference between βf^o_{ex} and βf^o_{in} will become less relevant and D_{chem} will change less with temperature. For high temperatures D_{chem} will increase almost linearly with $\langle n \rangle$, with a slope decreasing with T until it reaches the above discussed profile at $T \to \infty$.

Diffusivity versus temperature. At constant loading the trend of D_{chem} w.r.t. T is similar to D_s (see Figure 5.12). The difference is that, while D_s is very temperature-sensitive at low loadings but less sensitive at high loadings, D_{chem} is very temperature-sensitive for both low and high loadings.

5.2.2 Local density

Mean life time. In Figure 5.13 curves of the Mean Life Time (MLT) are shown for a relatively low temperature, $T = 300$ K, and for infinite temperature. For each loading, data are plotted only for the most probable occupancies (i.e. the occupancies n appearing with a proability $p_N(n) > 10^{-3}$).

Observing the behavior of this function for $T = 300$ K (Figure 5.13), it should be noted how a non-negligible difference between βf^o_{ex} and βf^o_{in} can introduce a marked separation in the mean life times of different occupancies,

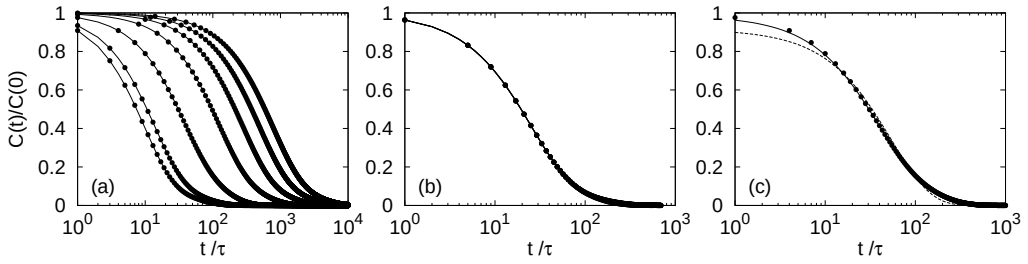

Figure 5.14: Normalized time-autocorrelation function $C(t)/C(0)$ for various loadings. Dots are numerical data, lines are fitting curves. (a) $T = 300$ K. Loadings are $\langle n \rangle = 1, 3, 5, 7, 9, 11, 15$ for curves from right to left. (b) $T \to \infty$. Loadings are $\langle n \rangle = 1, 7, 10, 15$, curves are overlapped. (c) $C(t)/C(0)$ for $\langle n \rangle = 9$ at $T = 300$ K: the dashed line is a single exponential fit, the solid line is a double exponential fit.

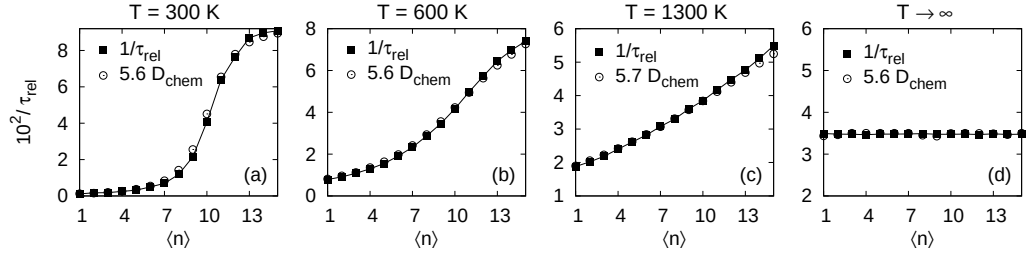

Figure 5.15: Inverse of the instantaneous relaxation time, τ_{rel}, obtained through the double exponential fit of $C(t)/C(0)$, compared with the chemical diffusion coefficient D_{chem}.

causing the curves to cover several orders of magnitude on the time axis. (i) For low and intermediate loadings, the less filled cells live longer because they host all or almost all of the guests in the inner sites, therefore the particle transfers are little frequent and so the variations of the lowest occupancies. (ii) If the loading increases, then $\langle \alpha_n \rangle$ will increase, therefore the transfers will become more frequent and this will reduce the MLT of all the occupancies. Moreover, the mean life times will be less sensitive to n. It is worth noting that this is the same kind of behavior shown by the MLT of the occupancies in the α-cages in MD simulations of diffusion of methane in ZK4,[41] as can be seen in Figure 5.13. This confirms that the ThPCA approach effectively captures the essential features of the process of diffusion in zeolites. (iii) For the highest loadings (e.g., $\langle n \rangle = 15$ in Figure 5.13), the more filled cells will release particles *slightly* more slowly because almost all of the exit sites will be occupied, therefore the mean life times of the most occupied cells will be slightly longer.

The separation in MLT of cells with different occupancies becomes less evident by increasing the temperature. In the limit of $T \to \infty$ (see Fig-

ure 5.13) the sites are all equivalent, therefore the mean life times are all of the same order of magnitude, and the MLT curves are more symmetrical when compared to the case of diffusion in presence of the energy effects discussed in the previous point.

Time-autocorrelation of density fluctuations. The study of the LD-FACF function $C(t)$ defined in Eq. (4.32) provides informations about how the local density fluctuations relax in time towards equilibrium.

In Figs. 5.14a, 5.14b and 5.14c the normalized time-autocorrelation function $C(t)/C(0)$ is reported. The best fit of $C(t)/C(0)$ is provided by a double exponential function

$$C(t)/C(0) \approx A_1\, e^{-t'/\tau_1} + A_2\, e^{-t'/\tau_2}, \tag{5.4}$$

where $t' = t/\tau$ is the time expressed in units of time steps, while $A_{1,2}$ and $\tau_{1,2}$ are fitting parameters. As can be seen in Figure 5.14c, a simple exponential decay is inappropriate to fit the simulation data.

The relaxation of the density autocorrelation is reported at low temperature ($T = 300$ K) in Figure 5.14a and at very high temperature ($T \to \infty$) in Figure 5.14b. If the temperature is low, the relaxation is different for different loadings; in particular, at high $\langle n \rangle$ the function $C(t)/C(0)$ goes to zero more rapidly than at low $\langle n \rangle$. This feature is also exhibited by the autocorrelation of density fluctuations computed in MD simulations (see Ref.,[41] Figure 8). When the temperature is increased, the relaxations at different loadings become similar. This can be seen from Figure 5.15, where the quantity $1/\tau_{\mathrm{rel}}$ is shown for various loadings and temperatures. The parameter τ_{rel} is the instantaneous relaxation time, defined as the time employed by the fit of $C(t)/C(0)$ to reach $1/e$ of its initial value. As the temperature becomes very large, the relaxation becomes independent of temperature, as well as the chemical diffusion coefficient. This behavior supports the observations about the separation of the mean life times: the local density relaxes in time in a way dependent on the degree of inhomogeneity of the lattice, which becomes important at low temperature.

Finally, as can be seen from the proportionality between $1/\tau_{\mathrm{rel}}$ and D_{chem} shown in Figure 5.15, the decay of the density fluctuations controls the transport of density in the lattice.

Chapter 6

Interacting case: Numerical Simulations

6.1 Cellular Automata from static models

In this Section, ThPCA simulations will be discussed where the energy parameters vary as functions of the cell occupancy. The modeling of that dependence further enlarges the configuration space the model can access. On the other hand, a dependence of the local energy on local observables is a reasonable choice since in real systems it is not uncommon for the energy of adsorption sites to be dependent on whether the neighboring sites are occupied or not,[32] and in principle such a situation can be reproduced by adjusting the site free-energy parameters, or by estimating them through coarse-graining of the interactions in a zeolite cage described at the atomistic level.

In this Section, a local thermodynamic model developed by Ayappa *et al.*[22] to study adsorption of xenon in NaA zeolite at 360 K will be implemented on the ThPCA, and the resulting dynamical behaviors shall be analyzed. In their approach, Ayappa *et al.* made use of an energy function of the form

$$E(n) = n\varepsilon + \frac{zn^2\varepsilon^{\mathrm{pp}}(n)}{2K} \tag{6.1}$$

where n is the cell occupancy, ε is the adsorbate-lattice interaction energy, z is the assumed number of nearest-neighbors per lattice site, and $\varepsilon^{\mathrm{pp}}(n)$ is the adsorbate-adsorbate interaction parameter given by the following Lennard-Jones-like functional form:

$$\varepsilon^{\mathrm{pp}}(n) = 4\epsilon^{\mathrm{lj}}\left[\left(\frac{\sigma}{r(n)}\right)^{12} - \left(\frac{\sigma}{r(n)}\right)^{6}\right] \tag{6.2}$$

with

$$r(n) = (r_2 - r_1)\frac{n}{K} + r_2. \tag{6.3}$$

In Eqs. (6.2) and (6.3), K is the number of adsorption sites per cell, ϵ^{lj} and σ are respectively the Lennard-Jones energy parameter and molecular diameter, $r_2 = 2^{1/6}\sigma$, and r_1 is a free parameter.

The effective site volume $v(n)$ changes with the occupancy according to

$$v(n) = v_0 \left[\frac{1}{2}\left(\frac{l-\sigma}{2R} + 1\right)^2\right]^{zn/2K}, \tag{6.4}$$

where v_0 is the volume of the adsorption site associated with a single adsorbate in the absence of neighboring adsorbates, R is the site radius obtained by assuming a cubic site of volume v_0, and l is the lattice parameter. The resulting partition function of an n-occupied cell can be written as

$$Q^{\mathrm{Ay}}(n) = e^{-\beta n \phi^{\mathrm{Ay}}(n)} \binom{K}{n} e^{-\beta n \varepsilon}, \tag{6.5}$$

where $\phi^{\mathrm{Ay}}(n)$ is the interaction free-energy term given by

$$\phi^{\mathrm{Ay}}(n) = -\frac{1}{\beta} \ln \frac{v(n)}{\Lambda^3} + \frac{zn\varepsilon^{\mathrm{pp}}(n)}{2K}. \tag{6.6}$$

To simulate the adsorption of xenon in NaA zeolite at $T = 360$ K, Ayappa *et al.*[22] assumed $K = 12$, $z = 4$, $\sigma = 4.10$ Å, $\varepsilon = -25.6$ kJ mol^{-1}, $\epsilon^{\mathrm{lj}} = 1.837$ kJ mol^{-1}, $r_2 = 3.906$ Å, $v_0 = 7.465$ Å^3, $l = 3.58$ Å.

Eqs. (6.1) to (6.5) assume equivalence of inner and exit sites. However, due to the heterogeneity of the zeolite cage structure, multiple types of adsorption site mutually differing in adsorption energy are present,[31,42–44] causing the probability of a guest molecule to occupy locations close to the windows to differ from the probability to occupy any other location in the cell. In the ThPCA paradigm such a differentiation is mimicked through the modeling of the difference in statistical weights between exit and inner sites. Such a modeling induces a dynamics characterized by a hierarchy criterion where guests must be promoted to the exit sites in order to leave their host cell, so that the attitude of the cells to transmit their contents outside becomes strictly dependent on local conditions.

The ThPCA paradigm can be used to extend a static thermodynamic model as the one of Ayappa *et al.* to the simulation of transport properties by

assuming two fixed site free-energies, $f^o_{\rm ex}$ and $f^o_{\rm in}$, entering a cellular partition function of the form

$$Q(n) = e^{-\beta n \phi(n)} Q^o(n), \tag{6.7}$$

where

$$Q^o(n) = \sum_{n_{\rm ex}, n_{\rm in}}^{(n)} \prod_{\alpha={\rm ex,in}} \binom{K_\alpha}{n_\alpha} e^{-\beta n_\alpha f^o_\alpha} \tag{6.8}$$

and $\phi(n)$, with $n = 1, \ldots, K$, defined as

$$\phi(n) = \phi^{\rm Ay}(n) + \varepsilon + \frac{1}{n\beta}\left[\ln Q^o(n) - \ln \binom{K}{n}\right], \tag{6.9}$$

in order Eqs. (6.7) and (6.5) to be equivalent.

Obviously the use of Eqs. (6.7) to (6.9) implies that the equivalence $\phi_{\rm ex}(n) = \phi_{\rm in}(n) \equiv \phi(n)$ is assumed for each n in Eq. (2.21). Although as a result the trend of $\langle n_{\rm ex}(n)\rangle$ *vs.* n will depend only on the difference

$$\Delta f^o = f^o_{\rm ex} - f^o_{\rm in} \tag{6.10}$$

and not on the interaction parameter ϕ, that assumption will be enough to show the most general diffusion profiles that are available in the ThPCA parameter space. Anyway, a finer tuning of the transport properties can be acquired through differentiation of the $\phi_{\rm ex}(n)$ and $\phi_{\rm in}(n)$ trends while preserving (numerically) the total value of $Q(n)$, thus further enlarging the range of possible dynamical behaviors of the model.

From pore properties, to cell properties, to macroscopic properties. In the left side of Figure 6.1 it is shown how local interactions and effective volume are embedded together in the effective interaction free-energy $\phi^{\rm Ay}$. Then, in the right side it is shown how different values of Δf^o generate different diffusivity trends. In details, in the right side of Figure 6.1:

(i) In the first line it is shown how the local free-energy parameters, $f^o_{\rm ex}(n)$ and $f^o_{\rm in}(n)$, are modified by adapting the cell termodynamics to different values of Δf^o by means of Eq. (6.9).

(ii) In the second line a mesoscopic quantity, the *average exit sites accessibility*, given by the ratio $\langle n_{\rm ex}(n)\rangle/n$ (where $\langle n_{\rm ex}(n)\rangle$ is the average number of occupied exit sites in a n-occupied cell), is shown for each value of Δf^o. Such a quantity is of remarkable importance, since it definitely determines the qualitative trend of the memoryless diffusivity.

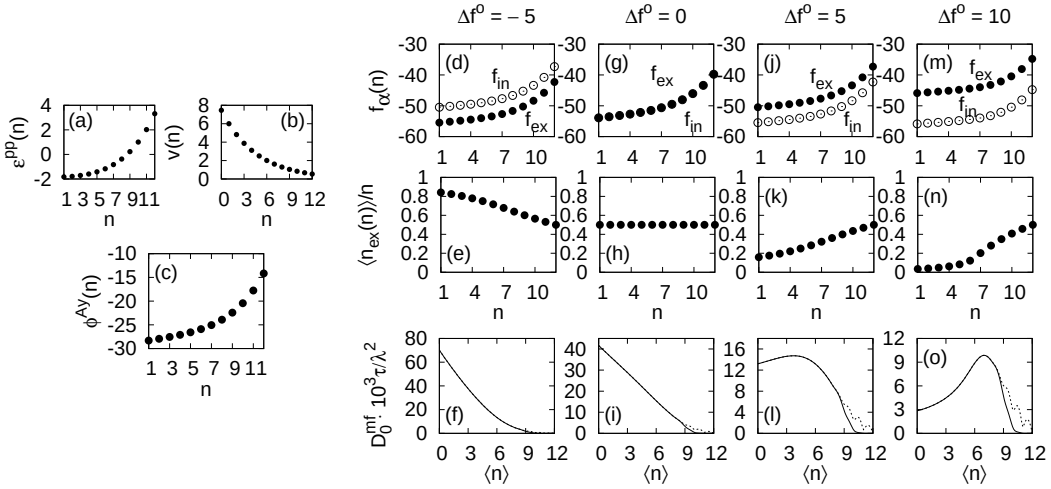

Figure 6.1: Possible diffusive behaviors modeled on the thermodynamic model of Ayappa *et al.*[22] for adsorption of xenon in NaA zeolite at 360 K. On the left, an effective interaction free-energy function $\phi(n)$ is constructed starting from the given effective Lennard-Jones interaction potential $\varepsilon^{\mathrm{lj}}(\mathrm{n})$ and volume $v(n)$. On the right, four different balances between average occupancies of exit and inner sites are assumed by varying the difference $\Delta f^o = f^o_{\mathrm{ex}} - f^o_{\mathrm{in}} = -5$, 0, 10, and 10 kJ mol^{-1}. For each column relative to each value of Δf^o, the following properties are plotted: (top) The effective site energy profiles $f_{\mathrm{ex}}(n)$ and $f_{\mathrm{in}}(n)$ in units of kJ mol^{-1} *vs.* the occupancy n; (middle) The average exit sites occupancy $\langle n_{\mathrm{ex}}(n)\rangle$; (bottom) The mean-field self-diffusivity D_s^{mf} *vs.* the loading $\langle n\rangle$. Solid and dashed lines represent D_s^{mf} respectively with and without the kinetic barrier modeling introduced in Section 3.3.1. See text for further details.

(iii) In the third line the resulting memoryless diffusivity (computed through the mean-field equation, Eq. (A.10), which will be introduced in Appendix A) is shown. Solid lines are obtained through a test modeling of the harmonic force constant $k(n)$ introduced in Eq. (3.31), where the jump barrier for jumps between highly occupied cells has been lowered by assuming

$$k(n) = C_1 + \theta(n' - 1)\left(C_2 n' + C_3 n'^2\right), \tag{6.11}$$

where $n' = n - \xi + 1$ and the Heaviside function $\theta(x)$ (which has value 1 for $x \geq 0$ and 0 otherwise) allows $k(n)$ to remain constant (and equal to C_1) for occupancies $0 \leq n \leq \xi - 1$, and rises quadratically for $\xi \leq n \leq K - 1$. In the example of Figure 6.1 the parameters have been assigned the values $C_1 = 0$, $C_2 = 2$, $C_3 = 20$, and $\xi = 9$. Dotted lines are instead diffusivities obtained with no modeling of the kinetic barrier, i.e. $k = 0$.

It should be noted that the choice of Δf^o determines the qualitative shape of the exit sites accessibiliy (a local equilibrium property), which in turn

induces a specific qualitative trend of the self-diffusivity (a global equilibrium property). In general, the diffusivity trend embed the accessibility trend together with the effect of repulsion at high loadings. Since according to Eqs. (A.8), (A.9), and (A.10) the quantity D_s^{mf} is $\lambda^2/2d\tau$ times the escape probability p^{esc}, such trends can be easily explained. The discussion of the diffusivity trend follows the same lines of Section 5.

If $\Delta f^o < 0$ (so that the exit sites are deeper than the inner sites, see Figs. 6.1d-f) then the exit sites accessibility exhibits a *decreasing* trend from high values at $n = 0$ to K_{ex}/K at $n = K$. Due to the high accessibility, the exit sites are very frequently visited so that during propagation there is a high probability of two adjacent exit sites to be simultaneously occupied. This gives no transfer between the two respective cells. As a consequence, the D_0^{mf} trend will be *decreasing*.

If exit and inner sites are equivalent (i.e. $\Delta f^o = 0$, see see Figs. 6.1g-i) then the value $\langle n_{\mathrm{ex}}(n)\rangle/n$ is held constant at K_{ex}/K. Therefore, since due to the constant accessibility the probability of an exit site to be occupied increases linearly with the occupancy, then the probability of two adjacent exit sites to be both occupied during propagation is also linearly increasing, therefore the D_0^{mf} is expected to (approximately) linearly decrease with increasing loading $\langle n\rangle$.

If the inner sites are set as slightly deeper than the exit sites (e.g. the case of $\Delta f^o = 5$ kJ mol^{-1} in Figure 6.1j-l), then the exit sites accessibility exhibits a slightly increasing trend from values between 0.1 and 0.2 at $n = 0$ to K_{ex}/K at $n = K$. This produces a slight inflection of D_0^{mf} at the lowest loadings, since a promotion mechanism is induced which causes the guests to occupy preferentially the inner sites: this lowers the probability of two adjacent exit sites to be both occupied.

At last, when the inner sites are much deeper than the exit sites (e.g. the case of $\Delta f^o = 10$ kJ mol^{-1} in Figure 6.1j-l), then (i) low values of D_s^{mf} are expected for low loadings due to the fact that at the lowest occupancies most of the guests are located in the inner sites so that the exit sites are poorly occupied and the intercell transfers are rare, (ii) an increase in D_s^{mf} is expected at intermediate-high loadings where almost all the inner sites are occupied and the guests start filling the exit sites, and (iii) a maximum and then a rapid decrease are expected at higher loadings, where exit sites start being saturated so that events at the propagation steps in which both communicating exit sites of a pair of neighboring cells are occupied become frequent.

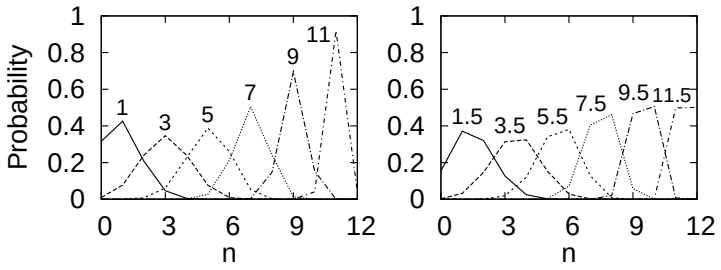

Figure 6.2: Occupancy probability distributions for the same system of Figure 6.1. Each curve refers to a particular value of the loading $\langle n \rangle$ specified on its top. Distributions for integer and half-integer loadings are shown respectively on the left and on the right part of the figure.

Modeling the self-diffusivity. As mentioned above, in each D_s^{mf} plot of Figure 6.1, two self-diffusivity trends are shown for each reported value of Δf^o. Solid and dashed lines represent D_s^{mf} trends obtained respectively with and without a modeling of the kinetic barrier ϵ_{ki} introduced in Section 3.3.1. The non-modeled case exhibits local D_s^{mf} maxima at the highest loadings. This is an effect of the repulsive part of the interaction function $\varepsilon^{\mathrm{pp}}(n)$ at high occupancies. The reason of such a behavior can be found in the propagation probability defined in Eq. (3.21), according to which migration events followed by a decrease in free-energy are favoured. First of all, the positions of the D_s^{mf} peaks at high loadings are remarkable, since they are located at half-integer values of $\langle n \rangle$. More generally, they can exist only if in the occupancy proability distribution $p(n)$ there are two (or more) occupancies much more probable than all the others. In the case of the current parametrization the shape of $p(n)$ is shown in Figure 6.2 for integer (left) and half-integer loadings (right).

At integer loadings the peak in $p(n)$ corresponds to the occupancy $n_{\max} = \langle n \rangle$ and is much more pronounced than all other occupancies. At half-integer loadings instead *two* occupancies (i.e. the ones located around $n_1 = \langle n \rangle - \frac{1}{2}$ and $n_2 = \langle n \rangle + \frac{1}{2}$) are much more probable than all the others. Taking account of the shape of the free-energy interaction term $\phi^{\mathrm{Ay}}(n)$ shown in Figure 6.1, jumps from a more to a less occupied cell are favoured. Therefore at high integer loadings, migrations from a $n_{\max}$-occupied to less occupied cells will be favoured, while the reverse jumps (and all jump to more occupied cells) will happen with a very low probability. This will cause the number of migrations per time step to be much lower than the case of high half-integer loadings, where instead not only migrations from n_1- and n_2- to less-occupied cells are favoured, but also exchanges between n_1- and n_2-occupied cells become very frequent thus producing diffusivity peaks near the saturation limit. Of

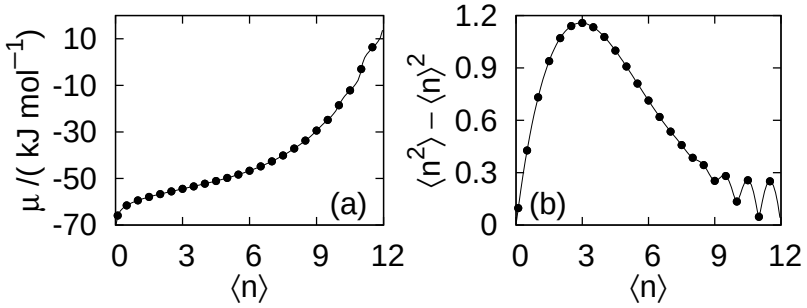

Figure 6.3: (a) Chemical potential and (b) variance of the occupancy probability distribution **p** for the same system of Figure 6.1.

course such a behavior is enhanced by irregularities in the chemical potential and occupancy variance trends, which in the parametrization by Ayappa *et al.* are present as shown in Figure 6.3.

p

6.2 Numerical simulations to fit reference data

In the present Section, the ThPCA rule (with memoryless randomization) shall be used to produce thermodynamic and transport properties in agreement with literature data about real reference systems. In particular, the following systems will be investigated:

- Xenon atoms in Na*A* zeolite, where the interactions have been modeled in order to fit equilibrium properties available from both experimental measurements and grand-canonical Monte Carlo (GCMC) simulations of Jameson *et al.*,[76,77] without an explicit modeling of the kinetics.

- Methane in ZK4 zeolite, where guest-host interactions and the kinetics have been modeled in order to fit MD results relevant to equilibrium properties from Fritzsche *et al.*[60] and self-diffusivities from Dubbeldam *et al.*[32] while neglecting the guest-guest interactions.

- Ethylene in Na*A* zeolite, where both interactions and kinetics have been modeled to fit adsorption and self-diffusion data from experimental results of Ruthven and Derrah[88] and from a lattice MC model developed by Gladden *et al.*[89]

- Xenon in zeolite Ca*A* at 300 K, where equilibrium properties are modeled so as to fit experimental adsorption data from Jameson *et al.*[57]

	Xe in NaA	CH_4 in ZK4	C_2H_4 in NaA		Xe in CaA
n	$\phi(n)$	$k(n)$	$\phi(n)$	$k(n)$	$\phi(n)$
0		18		3.5	
1	0	18	0	11	−0.3467
2	−0.3309	18	−0.2672	11.5	−0.3768
3	−0.9057	18	−0.5494	9	−0.3975
4	−1.3248	21	−0.7848	1.3	−0.3998
5	−1.5085	21	−0.9763	0.5	−0.3690
6	−1.2631	21	−0.1333	0	−0.2814
7	−0.7891	0	$+\infty$		−0.0999
8	−0.1603	2	$+\infty$		0
9		2.5			
10		12.8			
11		28			
12		30			
13		26			
14		19			
	$f^o_{\text{in}} = -44.920$ $f^o_{\text{ex}} = -41.920$ $K = 8$	$f^o_{\text{ex}} - f^o_{\text{in}} = 10$ $\phi = 0$ $K = 15$	$f^o_{\text{in}} = -79.126$ $f^o_{\text{ex}} = -73.126$ $K = 8$		$f^o_{\text{in}} = -50.0$ $f^o_{\text{ex}} = -44.0$ $K = 8$
	$T = 300$ K		$T = 358$ K		$T = 300$ K

Table 6.1: All ThPCA parameters used in this work to fit data from literature (see text for details). The free energies f^o_{in}, f^o_{ex}, and ϕ are expressed in units of kJ mol^{-1}. All reported values of the force constant $k(n)$ refer to $\alpha = 1$.

All the numerical simulations in this Section have been performed on a lattice of $32 \times 32 \times 32$ cells, for a total observation time of 10^6 time steps after an equilibration time between 10^3 and $2 \cdot 10^4$ time steps, depending on the average intercell transfer rate of the guests. All input parameters of the simulations are listed in Table 6.1.

6.2.1 Application: xenon in zeolite Na*A*

The ThPCA has been modeled in order to reproduce the adsorption isotherm of xenon in zeolite Na*A* at 300 K reported by Jameson *et al.*[76] (experimental and GCMC simulation data). The obtained parameters produced equilibrium probability distributions of occupancies in quantitative agreement with experimental NMR results.[77]

In this application the presence of cations has been taken into account in an implicit way by fixing the maximum occupancy at $K = 8$[76,77] (instead of 15 as it would be in a cation-free zeolite like ZK4[15]). According to the cell

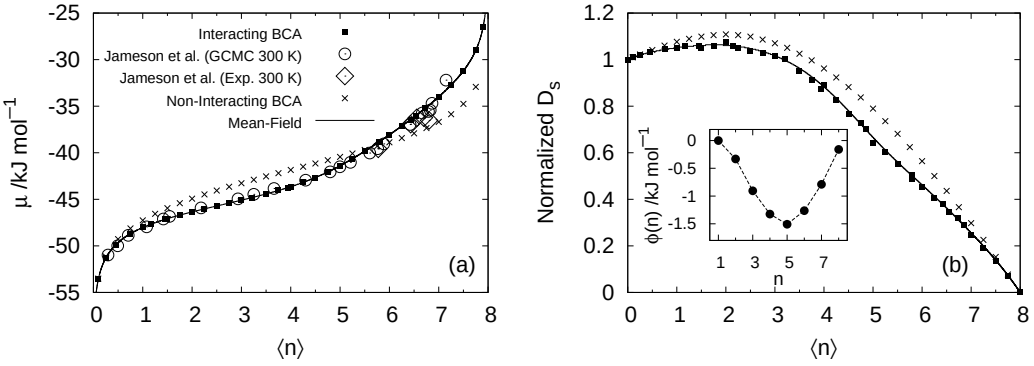

Figure 6.4: (a) Adsorption isotherm of Xe in NaA zeolite at 300 K obtained by introducing the effective parameters shown in Table 6.1 in the ThPCA, in direct comparison with GCMC and experimental results from Jameson *et al.*,[76] taken as the reference data to fit. The adsorption isotherm of a non-interacting lattice-gas[54] characterized by the same values of K, $f^o_{\rm ex}$, and $f^o_{\rm in}$ is also shown. In (b) the resulting ThPCA diffusivity trend (normalized w.r.t. the diffusivity in the limit of infinite dilution) is shown together with a picture of the trend of the fitting parameter ϕ (inset).

structure defined in Section 2.2, only the topology of the $K_{\rm ex} = 6$ exit sites is defined in an octahedral arrangement as confirmed by measurements of Lim *et al.*,[90] which justifies the cubic topology of the entire grid of cells. The remaining $K - K_{\rm ex} = 2$ sites are taken as structureless inner sites, referring to locations of a real zeolite cell which are not necessarily close to the center of an α-cage (where no Xe atom was found in any of the measurements of Jameson *et al.*[10,76,77,91]), but simply they represent locations which do not give access to any of the 6 windows. In order to mimic the effect of window blocking[92] all exit sites are assumed to be available but (i) their accessibility is less than the inner sites (this is achieved by setting $f^o_{\rm in} < f^o_{\rm ex}$), and (ii) they are affected by a mean field diffusion barrier $\epsilon_{\rm ki}$ (see Section (3.3.1)) that scales by a factor $e^{-\beta\epsilon_{\rm ki}} = 0.1$ the escape probability of a guest close to a window, therefore inducing a homogeneous slowdown of the diffusion process.[55] Since the present interest is to obtain a qualitative diffusivity trend, $\epsilon_{\rm ki}$ is treated as a fixed parameter. Therefore, a strategy to find the best values of f^o_α and ϕ would be to use mean field equations which will be presented in Appendix A to (i) extract the values of f^o_α providing a good fit of the isotherm for low loadings under the approximation of non-interacting LGCA, and to (ii) introduce the interaction ϕ as a correction function, in order to mimic repulsive/attractive effects which improve the fit for higher loadings.

Both the reference isotherm and an excellent fitting curve can be found in Figure 6.4a. In the same figure the adsorption isotherm for a non-interacting

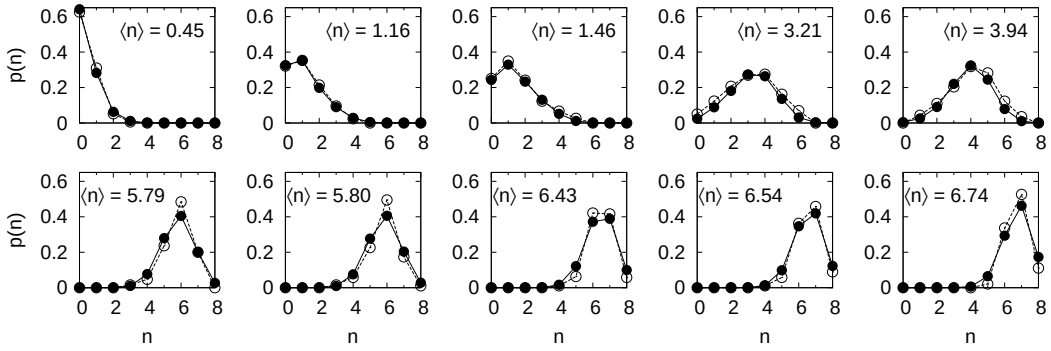

Figure 6.5: The occupancy probability distributions of guests in the cells from ThPCA simulations (black dots) for Xe in NaA zeolite are shown for various loadings in direct comparison with experimental results of Jameson *et al.*[77] (white dots).

lattice gas is reported, which is characterized by the same structure and by the values of the free energy parameters f^o_{ex} and f^o_{in} providing the best fit to the reference data. In Figure 6.4b the obtained trends of self-diffusivity *vs.* loading and of the interaction function ϕ *vs.* occupancy n (inset) are reported. All the obtained values of parameters are listed in Table 6.1.

It should be noted (inset of Figure 6.4b) that ϕ reaches a minimum at intermediate values of n and then increases due to repulsion. This finding is in qualitative agreement with (i) the trend of the average energy of clusters of n Xe atoms inside NaA zeolite cavities as obtained by Li and Berry[93] through atomistic simulations, and with (ii) the interaction potentials of previous thermodynamic models by Ayappa[22] and Cheung.[21]

As can be seen in Figure 6.4b, the obtained self-diffusivity profile exhibits small variations up to loading $\langle n \rangle = 3$, then decreases until it goes to zero at the saturation limit $\langle n \rangle = 8$. Such a trend corresponds to the type-II diffusivity observed by Kärger and Ruthven in the pulsed field gradient-NMR (PFG-NMR) measurements of intracrystalline self-diffusion coefficient depending on sorbate concentration.[40] Therefore, although the model is simple in its structure and in this application the kinetic parameter ϵ is homogeneous, the (qualitative) diffusive behavior it produces is meaningful. This represents an improvement over the lattice model of Jameson *et al.*[77] where the trend is monotonically increasing *near saturation* (which is unphysical for the problem of diffusion in zeolites), and over the models of Ayappa[22] and Cheung[21] which provide no diffusivity trend.

In Figure 6.5 it is shown that the obtained free energy parameters give the same probability distributions of guests in the α-cages at different loadings as obtained experimentally.[10] This is a remarkable result. The ThPCA after a proper setting of its internal parameters is able to describe the nature

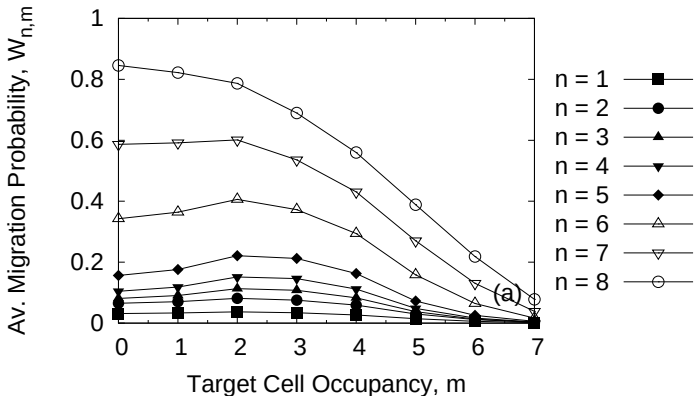

Figure 6.6: Average probability of a n-occupied cell to release a guest to a neighboring m-occupied cell. The line serves only as a guide for the eye. Each curve corresponds to a particular value of the occupancy of the departure cell.

of the equilibrium distributions of xenon atoms trapped inside the α-cages. Therefore, it represents an alternative to lattice MC calculations[77] to get a quantitative explanation of ^{129}Xe chemical shifts obtained by NMR measurements.[10] Consequently it is reasonable to expect that the very good agreement obtained in the description of the spatial distribution could be extended to the cage-to-cage migration process of the adsorbed fluid in the zeolite by using the same parameters.

In order to confirm such expectation the *average migration probability*, $W_{n,m}$, of a guest to migrate from a n- to a m-occupied neighboring cell during propagation, has been computed. In the ThPCA paradigm such a quantity is independent of loading, since it is determined only by the current occupancies, n and m, of the two adjacent cells. It can be obtained by simply coupling two cells with the given occupancies and then averaging over all events that cause one cell to change occupancy from n to $n-1$, and the other one to change from m to $m+1$. The trend of the probability $W_{n,m}$ with varying n and m (according to $1 \leq n \leq K$, and $0 \leq m \leq K-1$) arises directly from the mathematical structure of both randomization and propagation probabilities, and from the values of the energy parameters. In Figure 6.6 curves for the probability of $W_{n,m}$ relative to different values of the occupancy n of the departure cell are plotted *vs.* the occupancy m of the target cell. After small variations at low occupancies, the migration probability decreases rapidly due to a decrease in the availability of the target exit site and to the fact that high occupancies are unfavourable due to an increased interaction potential (see inset of Figure 6.4b).

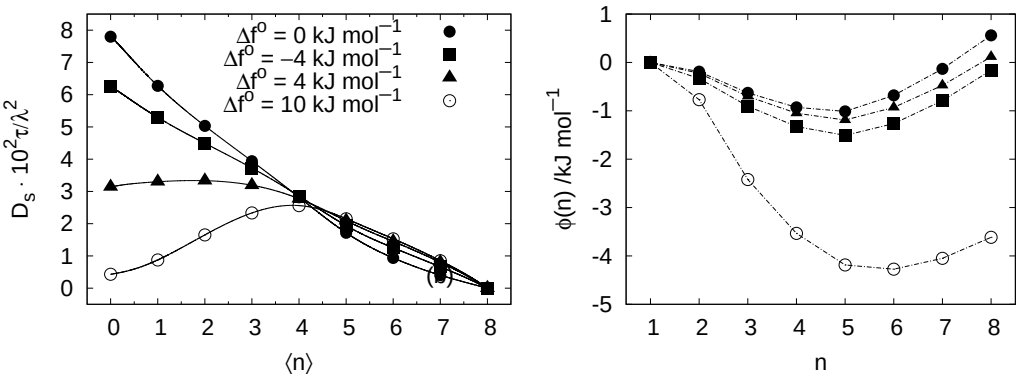

Figure 6.7: ThPCA diffusivity trends connected with the same set of partition functions fitting the isotherm of Xe in NaA, but arising from different values of Δf^o, i.e., changing the relative accessibility of inner and exit sites: (○) $\Delta f^o = 10$, (▲) $\Delta f^o = 4$, (●) $\Delta f^o = 0$, (■) $\Delta f^o = -4$. (All values are in units of kJ mol^{-1}) In the inset the trends of ϕ preserving the partition function $Q(n)$ are shown.

Modeling the diffusion profile. The local ThPCA parametric structure makes it a widely flexible environment where it is easy to modify some properties while leaving untouched the other ones, in order to refine the modeling of a given reference system. For example it is possible to work on the free energy parameters to vary the dependence of diffusivity upon the loading without producing any change in the adsorption isotherm. Such a procedure constitutes a first modeling of the ThPCA kinetic behavior, and can be easily realized through (i) modification of the difference $\Delta f^o = f^o_{\rm ex} - f^o_{\rm in}$ (see Section 6.1) and (ii) determination of the new function $\phi(n)$ allowing the partition functions $Q(0), Q(1), \ldots, Q(K)$ to keep the same values they covered before the modification. As for the above discussed determination of fitting parameters for Xe in NaA, once a new value of Δf^o is assumed, only one choice of the new interaction parameter $\phi(n)$ will result to be consistent with the fixed values of $Q(n)$. This procedure will produce different diffusion profiles as Δf^o changes (this can be seen in Figure 6.7), without affecting the global thermodynamic properties of the system. The physical meaning of Δf^o is based upon the knowledge of the probability of a guest to accede to a window; when such an input information is not available, Δf^o can be deduced by adapting the model to the kinetic behavior of a particular system. After such a procedure, agreement with reference data can be further improved by adjusting the kinetic barrier $\epsilon_{\rm ki}$.

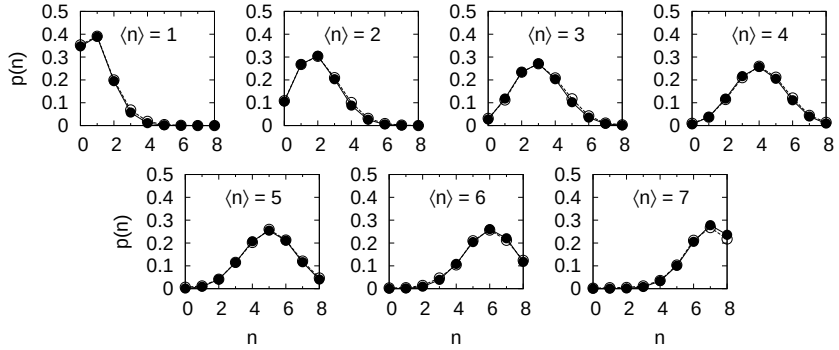

Figure 6.8: ThPCA occupancy probability distributions for methane in ZK4 zeolite shown in direct comparison with Molecular Dynamics simulation data of Fritzsche *et al.*[60] (white dots).

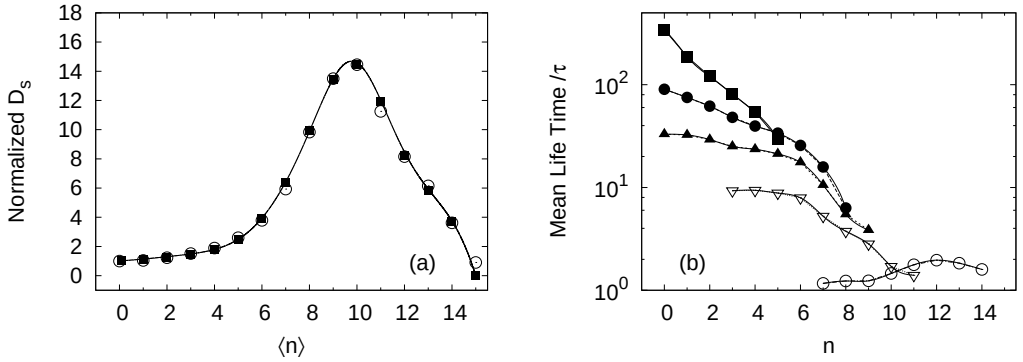

Figure 6.9: (a) A fit of the MD simulation data of self-diffusion coefficient of methane in ZK4 at 300 K from Dubbeldam *et al.*[32] with the ThPCA using the parameters of Table 6.1. Data are normalized w.r.t. the zero-loading diffusivity. (b) Mean life times of the cells having the most frequent occupancies at various loadings: (■) $\overline{n} = 1$, (●) $\overline{n} = 3$, (▲) $\overline{n} = 5$, (▽) $\overline{n} = 7$, (○) $\overline{n} = 11$

6.2.2 Application: methane in zeolite ZK4

The occupancy distributions of Fritzsche *et al.*[60] and the self-diffusion coefficient values reported by Dubbeldam *et al.*[32] have been taken as reference data.

As can be seen in Figure 6.8, in order to obtain occupancy distribution trends in agreement with MD data[60] it is enough to set $K = 15$, $\Delta f^o = 10$ kJ mol^{-1} and $\phi = 0$ independent of n. In this application it is shown how the kinetic barrier ϵ can be modeled as a function of the cell occupancy in order to obtain results in good agreement with reference data of diffusivity. In Table 6.1 the values of k giving the best agreement with MD data[32] shown in Figure 6.9a are reported.

In Figure 6.9b the mean life time (MLT) of a cell is plotted. This quantity

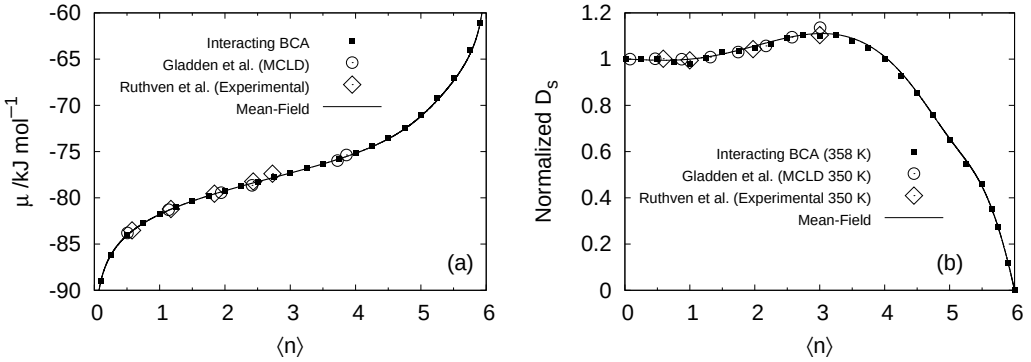

Figure 6.10: (a) Adsorption isotherm and (b) self-diffusivity profile of Ethylene in NaA zeolite obtained by introducing the effective parameters shown in Table 6.1 into the ThPCA to fit experimental data of Ruthven and Derrah[88] and Monte Carlo data from Gladden *et al.*[89]

is sensitive to the loading (see Demontis *et al.*[41,55] for a full discussion about this topic) and as one can see by direct comparison between Figure 6.9a and Figure 5.13c, the ThPCA model captures the qualitative trend found in MD simulations of methane in ZK4 at 360 K.[41] Again, the ThPCA is able to effectively reproduce the essential features of the physical problem under study, from both the thermodynamic and the kinetic point of view.

6.2.3 Application: ethylene in zeolite Na*A*

In this application the chosen reference data are the adsorption isotherm at 358 K and the self-diffusivity trend at 350 K of Ethylene in zeolite Na*A* from the experimental data of Ruthven and Derrah[88] and from a Monte Carlo Lattice Dynamics (MCLD) model developed by Gladden *et al.*[89] The best set of parameters giving the fits reported in Figure 6.10 is reported in Table 6.1. Besides the fact that the ThPCA simulates equilibrium and transport properties all at once in the canonical ensemble, another important difference between this approach and the aforementioned MCLD should be remarked concerning the structure of the single cell: while in a MCLD cell the saturation limit (6 molecules per cell) coincides with the number of exit sites,[89] in a CA cell a *nominal maximum occupancy* has been set as $K = 8$ in order to make accessible also the inner part of the cell (since $K_{\mathrm{ex}} = 6$ due to the LTA topology, the inner space is made up of $K_{\mathrm{in}} = 2$ inner sites). Then, an *effective maximum occupancy* $K^{\mathrm{eff}} = 6$ is imposed via the introduction of a strongly repulsive n-particle potential for occupancies $n > 6$ (i.e. setting $\phi(n > K^{\mathrm{eff}}) = \infty$).

This procedure together with $f^o_{\mathrm{in}} < f^o_{\mathrm{ex}}$ enables the CA cell to mimic

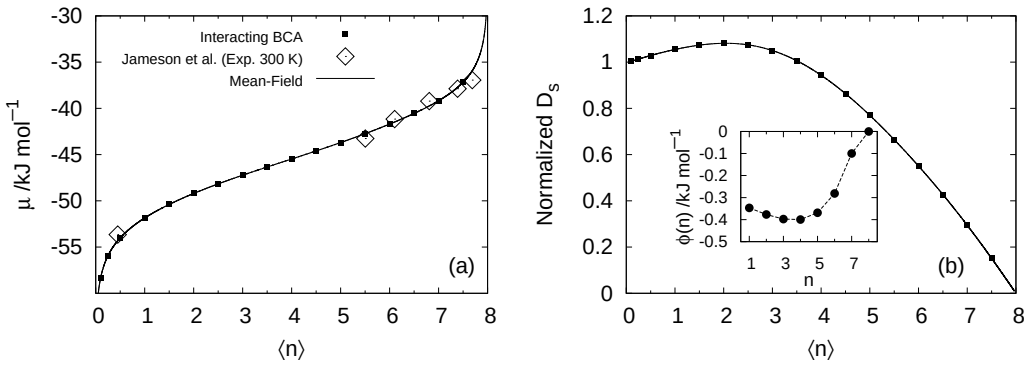

Figure 6.11: (a) Adsorption isotherm of Xe in CaA zeolite, and (b) resulting diffusion profile and interaction potential (inset)

the reduced accessibility of exit sites due to the presence of cations blocking the windows. With this example it can be seen clearly that the cellular, occupancy-dependent structure of the ThPCA energetics and the block synchronous propagation scheme,[35,72] allow a straighforward implementation of a constraint like the effective maximum occupancy.

6.2.4 Application: xenon in zeolite Ca*A*

In Figure 6.11a the results of the fitting of a set of experimental data from Jameson *et al.*[57] about the adsorption of Xe in Ca*A* zeolite are reported. In Figure 6.11b the expected diffusion trend together with the interaction potential $\phi(n)$ is shown. The maximum occupancy has been set as $K = 8$, while the framework parameters have been set as $f_{\text{ex}} = -44$ kJ mol^{-1} and $f_{\text{ex}} = -50$ kJ mol^{-1}. The numerical values of the parameters adopted are reported in Table 6.1.

Chapter 7

Exploring correlation effects

7.1 Correlated motion

The motion of real guests confined in a zeolite framework is always affected by correlations. The backscattering effect, according to which a molecule which has just moved from one location to another one has a large probability to return back to its former position, is the major source of correlations and becomes increasingly important when adding guests to the system, since that causes the number of vacancies (locations available to the diffusing molecule) to be reduced. In the previous Section, such a correlation effect has been not explicited, but has been incorporated in the effective intercell barrier $\epsilon_{\rm ki}(n,m)$, which then confer on the factor $e^{-\beta\epsilon_{\rm ki}(n,m)}$ the role of a *transmission coefficient*[32] of a guest from an n- to an m-occupied cell.

In the present Section the correlations will be set up in the model through the jump randomization strategy introduced in Section 3.2.2, which mimics the memory effects in the self-motion of the guests by means of correlated jumps inside the cell. Such memory effects will be here studied through numerical simulations.

To get accurate statistical averages, all simulations were performed on a grid of $16 \times 16 \times 16$ cells for an observation time of 10^5 time steps. Anyway, equilibrium properties of systems of various sizes above $L^3 = 4^3$ have been found to converge to the same values, therefore size effects can be considered unimportant in the cases presented here. Several (all three-dimensional) systems will be investigated characterized by different parameter sets:

- *System # 1*— This is a system of non-interacting guests ($\Phi = 0$ in Eq. (2.24)). Each cell has $K_{\rm ex} = 6$ exit sites with fixed energy $f^o_{\rm ex} = -10$ kJ mol^{-1}, and $K_{\rm in} = 9$ inner sites with energy $f^o_{\rm in} = -20$ kJ mol^{-1}.

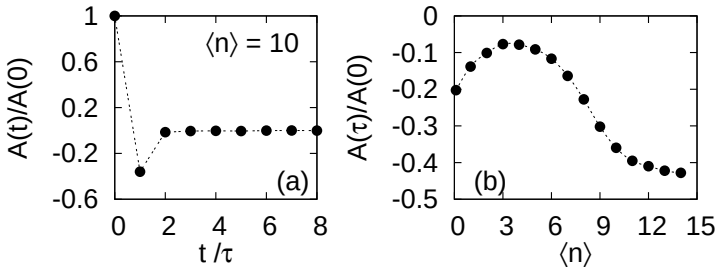

Figure 7.1: (a) Displacement autocorrelation function *vs.* time step for some loadings, and (b) backscattering effect after one time step τ represented by the ratio $C(\tau)/C(0)$ plotted *vs.* the loading for the System # 1.

The jump probability (3.9) reads

$$p^R_{\text{jump}}(\boldsymbol{\eta} \to \boldsymbol{\eta}') = A_{jk} e^{\beta f^o_j}$$

where exit-exit jumps are forbidden by means of the same A_{jk} defined in Eq. (3.12), and the assumed value of γ is $\exp(-\beta f^o_{\text{ex}})$.

- *System # 2*— Same of System # 1 apart from the parameter $\gamma = 1$.
- *System # 3*— Same of System # 1 apart from

$$p^R_{\text{jump}}(\boldsymbol{\eta} \to \boldsymbol{\eta}') = A'_{jk} e^{\beta f^o_j}.$$

where the A'_{jk} (defined in Eq. (3.12)) allows each jumping guest to choose any of the K cell site during randomization, with $\gamma = \exp(-\beta f^o_{\text{ex}})$.

- *System # 4*— Same of System # 3 apart from $\gamma = 1$.
- *System # 5*— Same of System # 1 apart from $f^o_{\text{ex}} = -20$ and $f^o_{\text{in}} = -10$ kJ mol^{-1}, i.e. the depth of the adsorption sites is reversed.
- *System # 6*— System with the same parametrizations as for Figure 6.1.

The entity of correlations is well represented by the displacement autocorrelation function (DACF) of a single guest,

$$A(t) = \langle \delta\mathbf{r}(t) \cdot \delta\mathbf{r}(0) \rangle \,, \tag{7.1}$$

which has been introduced in Eq. (4.21). From Figure 7.1a, where the DACF is pictured for some loadings ($\langle n \rangle = 1, 5, 10, 14$), it is clear the predominant role played in the entire guest's motion history by the correlation after *one* time step. The latter quantity $A(\tau)/A(0)$ plotted in Figure 7.1b expresses the ratio of self-diffusivity lost due to the backscattering effect after one time

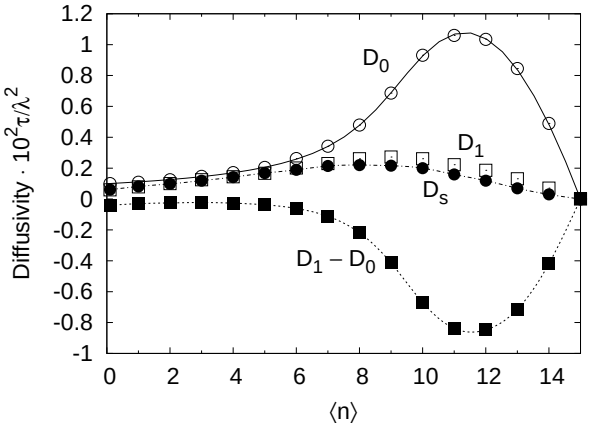

Figure 7.2: Self diffusivity profile D_s (black circles) for the System # 1 plotted *vs.* the loading together with the zero-time diffusivity D_0 (white circles), the diffusivity after one time step D_1 (white squares), and the backscattering contribution after one time step $D_1 - D_0$ (black squares). The continuous line is the mean-field memoryless diffusivity D_0^{mf}, Eq. (A.10). Dotted lines are to guide the eye.

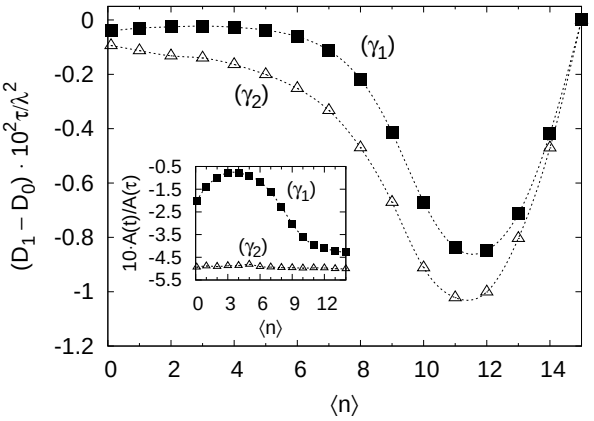

Figure 7.3: Backscattering contribution $D_1 - D_0$ *vs.* loading for two different values of γ. Here, $\gamma_2 < \gamma_1$ (where γ_1 refers to System # 1 and γ_2 to System # 2.) Dotted lines are to guide the eye.

step. As can be seen, after a value of D_s remarkably lower than D_0 at the limit of zero loading, increasing the number of guests in the system causes the backscattering to have not much effect on the intercell migration process until some loading ($3 < \langle n \rangle < 4$ in the case of Figure 7.1b) above which such a tendency is reversed due to a decrease in the probability of the guests to find empty sites to jump in during randomization. For loadings higher than $\langle n \rangle = K_{\mathrm{in}}$ most of the inner sites are occupied so that many successful events of migration at time t are canceled by the poor availability of empty sites to jump in during randomization at time $t + \tau$ which causes the guest to have

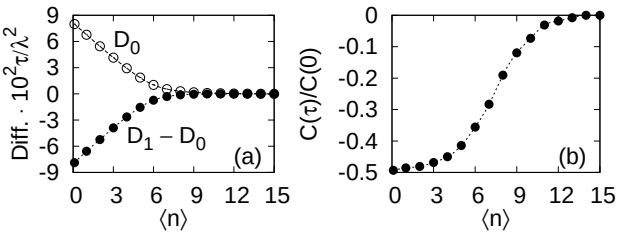

Figure 7.4: (a) Zero-time diffusivity, D_0, compared with the backscattering contribution, $D_1 - D_0$, and (b) Normalized DACF $A(t)/A(0)$ for System # 5.

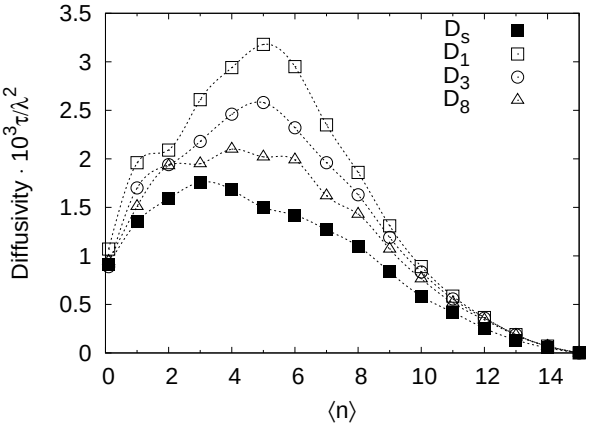

Figure 7.5: Self diffusivity profile D_s *vs.* loading for the System # 5, plotted together with the diffusivity after 1, 3, and 8 time steps. Dotted lines are to guide the eye.

a high probability to return back to the departure cell, so that $A(\tau)/A(0)$ in Figure 7.1b becomes more negative.

It shall be useful to define the diffusivity after j time steps as:

$$D_j = \sum_{z=0}^{j} \frac{\langle \delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0) \rangle}{d\tau} - \frac{\langle \delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0) \rangle}{2d\tau}. \tag{7.2}$$

In Figure 7.2 the zero-time diffusivity D_0, the one-time diffusivity D_1, the self diffusivity D_s given in Eq. (4.23), and the backscattering contribution to the diffusivity after one time step $D_1 - D_0 = \langle \delta\mathbf{r}(\tau) \cdot \delta\mathbf{r}(0) \rangle / d\tau$ are plotted *vs.* the loading to get a precise idea of the variation in diffusivity caused by the backscattering effect. The mean-field diffusivity computed through Eq. (A.10) has exactly the same definition of D_0. The one-time correlation due to the backscattering effect is most of the total entity of correlations in the motion, (i.e. in this case) so that $D_1 \approx D_s$.

The reducing of γ to the value of 1 gives an increase in correlation effect. This is shown in Figure 7.3 where the backscattering contribution after one time step, $D_1 - D_0$, is more pronounced (i.e. more negative) for the case of

$\gamma_2 = 1$ rather than $\gamma_1 = \exp(-\beta f^o_{\text{ex}})$. Moreover, the use of low values of γ (i.e. high jump barrier during randomization) causes correlations to level off, as can be seen in the inset of Figure 7.3 where it is shown that the ratio $A(\tau)/A(0)$ becomes less loading-dependent as the value of γ is lowered.

The change in shape of the D_s *vs.* $\langle n \rangle$ plot becomes dramatic when $f^o_{\text{ex}} < f^o_{\text{in}}$. This can be seen in Figure 7.4a where the tendency of a guest to *escape* a cell, as given by D_0, is counter-balanced by a backscattering effect, $D_1 - D_0$, almost equal in modulus and with opposite sign. This is confirmed in the $A(\tau)/A(0)$ plot in Figure 7.4b where the importance of the backscattering with respect to the escape ability is very high at low loadings (large negative values of $A(\tau)/A(0)$) and decreases with increasing loading until for a highly dense system it becomes null ($A(\tau)/A(0) \approx 0$). Moreover, the present case remarkably differs from the case of $f^o_{\text{ex}} > f^o_{\text{in}}$ because the diffusing guest finds itself very often unable to escape an exit site during randomization at time t. Then it migrates during propagation and at time $t+\tau$ is again trapped into the destination exit site during randomization, and so on cyclically until it reaches sufficient energy to perform a displacement during randomization. Such a 'rebound process' causes the approximation $D_1 \approx D_s$ to be no longer valid. Higher moments of D_j become important, so that more than one time step are necessary to define precisely the trend of D_s, as can be seen in Figure 7.5. Therefore, the resulting trend of D_s appears very different from D_0 since it exhibits a maximum at low-intermediate loading. This is due to the fact that the rebound process becomes weaker while increasing the loading, since more frequently the guests in the exit sites stop rebounding because of one of the involved exit sites being filled by another guest. This causes the correlations to weaken producing an initial diffusivity increase, which for higher loadings is canceled by the repulsion effect during propagation.

Self *vs.* collective diffusivity. One of the most realistic phenomena induced by the jump randomization is the differentiation between the *self* and the *collective* diffusivity. If the randomization is memoryless, then the migration process is governed only by the ability of guests to escape from their host cell. Therefore in that case, neither correlation in the self-motion nor in the collective motion are expected,[55] since the mixed displacement correlations for different guests cancel out.

In the case of jump randomization instead, some correlations between different guests will cancel out with some correlations in the self-motion, and this will result in a D_c higher than D_s. This is shown in Figs. 7.6 and 7.7 where D_s and D_c trends are compared directly respectively for the non-

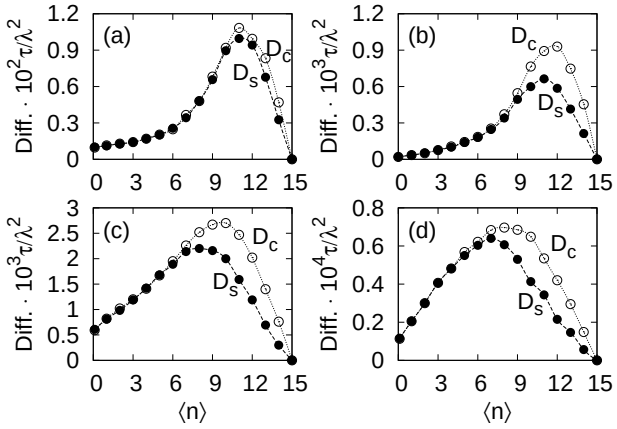

Figure 7.6: Self and collective diffusivity for (a) System # 3, (b) System # 4, (c) System # 1, and (d) System #2.

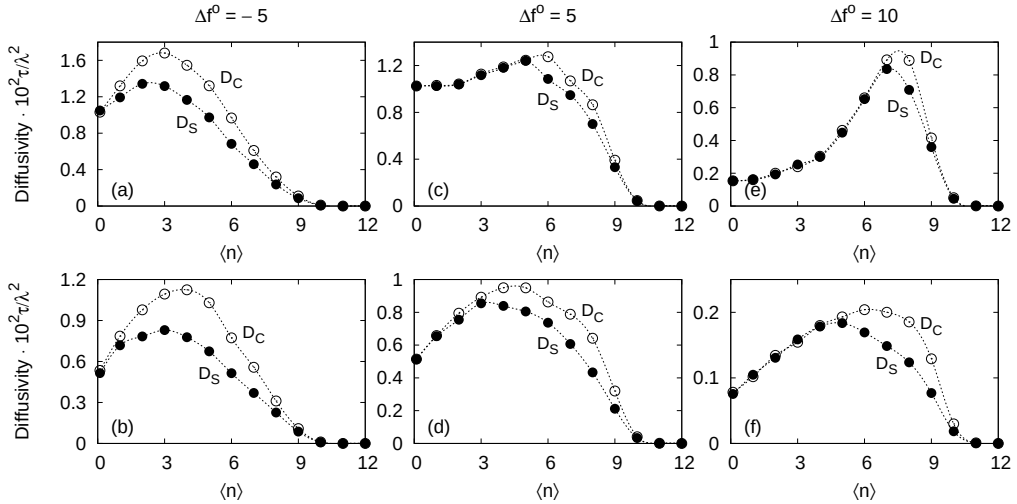

Figure 7.7: Self and collective diffusivities for the same system of Figure 6.1 while using the jump randomization. In figures (a), (c), and (e) jumps among different exit sites are allowed during randomization. In figures (b), (d), and (f) instead, such jumps are not allowed. Dotted lines are to guide the eye.

interacting and the interacting case. It should be noticed that in all these cases the difference between D_c and D_s increases in those regions where the migration sites are more involved in the migration process than the inner sites (i.e., from low to intermediate loadings for $f^o_{\rm ex} < f^o_{\rm in}$, and from intermediate to high loadings for $f^o_{\rm ex} > f^o_{\rm in}$). By comparison of the trends in Figs. 7.6 and 7.7 with the plots of collective- and self-diffusivity obtained from molecular dynamics simulation (see e.g. Dubbeldam *et al.*[32] for the case of methane in ZK4), it can be argued that the local, discrete ThPCA rule effectively

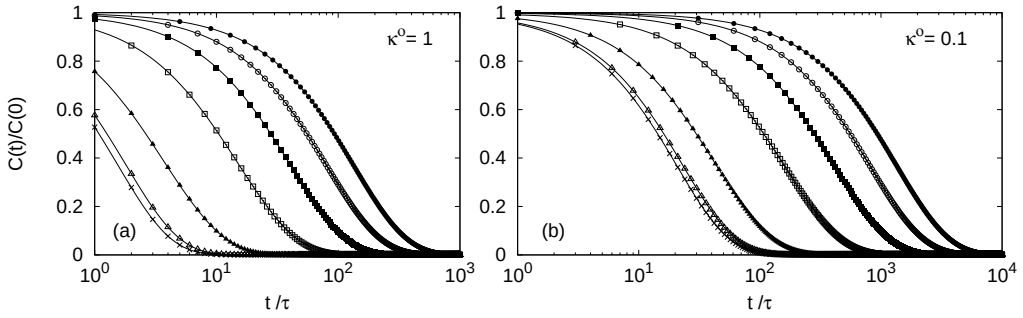

Figure 7.8: Biexponential fits of the normalized time-autocorrelation function $C(t)/C(0)$ for two different values of the scaling factor $\kappa^o = e^{-\beta\epsilon_{\rm ki}}$ defined in Eq. (A.53): (a) $\kappa^o = 0.1$ and (b) $\kappa^o = 1.0$. For both figures $K_{\rm ex} = 6$, $K_{\rm in} = 9$, $f_{\rm ex} - f_{\rm in} = 10$, and $T = 300$ K. In each figure, fits for increasing integer loadings are plotted as solid lines, while numerical data are reported as symbols: (●): $\langle n\rangle = 1$, (○): $\langle n\rangle = 3$, (■): $\langle n\rangle = 5$, (□): $\langle n\rangle = 7$, (▲): $\langle n\rangle = 9$, (△): $\langle n\rangle = 11$, (×): $\langle n\rangle = 13$.)

captures the most relevant aspects of correlated diffusion of simple chemical species in zeolites.

7.2 Correlations in local density fluctuations

In Section 5.2.2 the behavior of the LDFACF function has been discussed for the case of simulations of non-interacting guests. Here the same function will be analyzed both in the mean-field approach and in numerical simulation in the presence of correlations induced by the jump randomization.

The mean-field theory of the LDFACF will be derived in Appendix A. Here only the main results will be discussed.

Mean-field LDFACF decay. For $f^o_{\rm ex} - f^o_{\rm in} \neq 0$ and $T < \infty$, as it has been found in the simulations illustrated in Section 5.2.2 also the mean-field LDFACF relaxes to equilibrium following a bi-exponential decay:

$$C(t)/C(0) \sim b_1 e^{-t/\tau_1} + b_2 e^{-t/\tau_2}. \tag{7.3}$$

This can be seen in Figure 7.8 where biexponential fits are shown together with mean-field values of the normalized function $C(t)/C(0)$ for a reference three-dimensional system (i.e. $K_{\rm ex} = 6$) with $K_{\rm in} = 9$ inner sites and effective site energies differing of an amount $f^o_{\rm ex} - f^o_{\rm in} = 10$. Increases in temperature (or decreases in $|f^o_{\rm ex} - f^o_{\rm in}|$) cause the separation between the two timescales τ_1 and τ_2 to decrease, and in the limit of $f^o_{\rm ex} - f^o_{\rm in} = 0$ and/or $T \to \infty$ the decay is single-exponential, i.e. $C(t)/C(0) \sim e^{-t/\tau}$. The biexponential prefactors b_1 and b_2 in Eq. (7.3) give the balance between the timescales τ_1 and τ_2. As

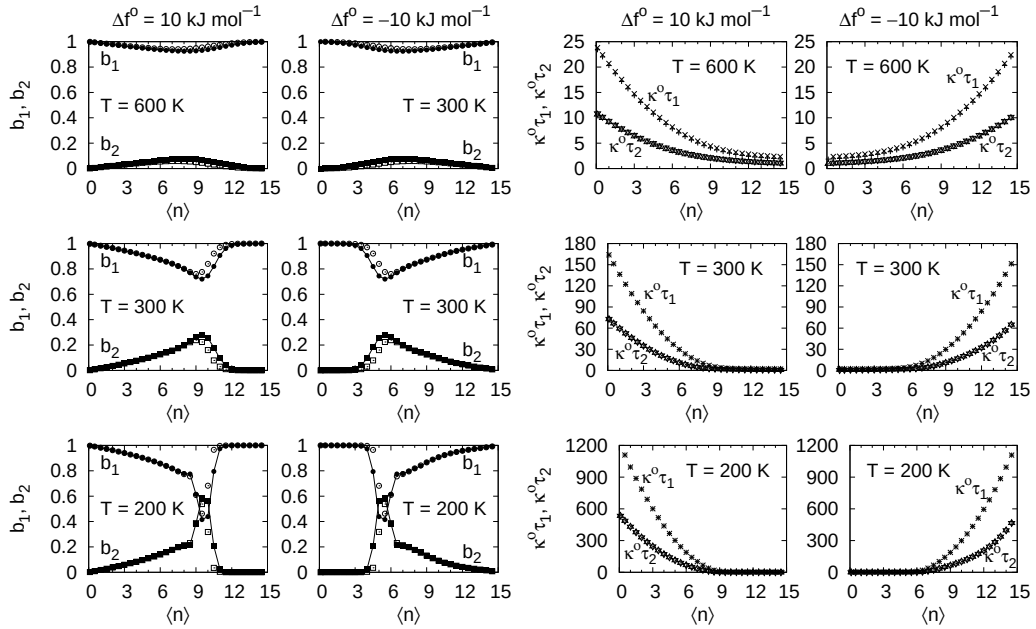

Figure 7.9: (left) The biexponential prefactors b_1 and b_2 (see Eq. (7.3)) of the function $A(t)/A(0)$ for a non-interacting system characterized by $K_{\rm ex} = 6$, $K_{\rm in} = 9$ is shown for different temperatures (600, 300, and 200 K). Figures (a), (b), and (c) refer to $f_{\rm ex} - f_{\rm in} = 10$, while (d), (e), and (f) refer to $f_{\rm ex} - f_{\rm in} = -10$. In all figures, filled and empty symbols refer respectively to $\kappa^o = 1.0$ and $\kappa^o = 0.1$ (right) The time parameters τ_1 and τ_2 (insets) of the biexponential fits (see Eq. (7.3)) for the same system of Figure 7.9 are shown in units of time steps. Different symbols are used for different temperatures: squares (200 K), circles (300 K), and triangles (600 K). Filled and empty symbols refer respectively to $\kappa^o = 1.0$ and $\kappa^o = 0.1$. .

can be seen in Figure 7.9, they are strongly sensitive to different conditions of loading and temperature.

One of the most remarkable features of the mean field LDFACF is that deviations become much higher while decreasing the temperature. The most interesting aspect is that minima of b_1 (or maxima of b_2) are located around the loading corresponding to the diffuse phase transition occurring in such a kind of systems (see Section 5), i.e. $\langle n \rangle = K_{\rm in}$ for $f^o_{\rm ex} - f^o_{\rm in} > 0$ and $\langle n \rangle = K_{\rm ex}$ for $f^o_{\rm ex} - f^o_{\rm in} < 0$.

The connection between minima (or maxima) in the prefactor b_1 (or b_2) and the diffuse phase transition becomes even clearer when we analyze the case of multiple kinds of inner sites. In Figure 7.10 the trend of b_1 is plotted for the case of a (mean-field) cell having one, two and three non-equivalent inner sites, together with the variance of the distribution function $\langle n^2 \rangle - \langle n \rangle^2$. In the first case (i.e. Figs. 7.10a-b) the cell is constituted by two kinds of non-equivalent sites (exit and inner) so that at low temperature *one* diffuse phase

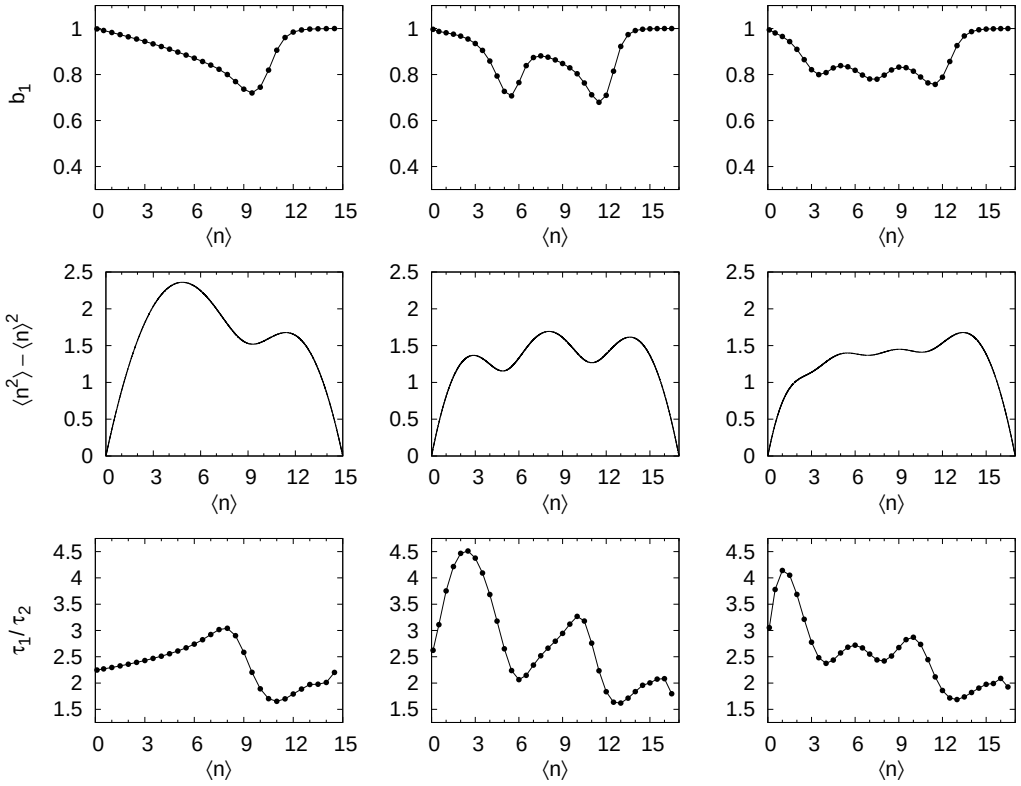

Figure 7.10: The biexponential prefactor b_1 (top) together with the occupancy variance (middle) and the ratio τ_1/τ_2 between the two biexponential timescales (down) for three kinds of cell: (a-c) 9 equivalent inner sites of energy -20 kJ mol^{-1}. (d-f) 6 inner sites at -20 kJ mol^{-1} and 5 at -30 kJ mol^{-1}. (g-i) 4 inner sites at -18 kJ mol^{-1}, 4 at -25 kJ mol^{-1}, and 3 at -32 kJ mol^{-1}. In all three cases there are 6 exit sites of energy -10 kJ mol^{-1}, and the temperature is 300 K.

transition is expected corresponding to the condition

$$\langle n \rangle \approx K_{\text{in}} \text{ if } f^o_{\text{in}} < f^o_{\text{ex}}, \text{ or}$$

$$\langle n \rangle \approx K_{\text{ex}} \text{ if } f^o_{\text{in}} > f^o_{\text{ex}}.$$

In the second and the third case (i.e. Figs. 7.10c-d and 7.10e-f) different energies have been associated respectively to two and three kinds of inner sites. Therefore, in general, if there are z different site energies ordered as $f^o_1 < f^o_2 < \cdots < f^o_z$, at low temperatures $z-1$ diffuse phase transitions are expected to take place at the loadings

$$\langle n \rangle \approx K_1,\ K_1 + K_2,\ \ldots,\ \sum_{i=1}^{z-1} K_i.$$

Diffuse phase transitions are caused by heterogeneities in the lattice sites causing the curvature of the variance to deviate from the hypergeometric trend given by the formula $\langle n\rangle(1-\langle n\rangle/K)$, which is the expected one in case of all-equivalent sites. If the temperature is not very low (as it is the case of Figure 7.10 where $T = 300$ K) the diffuse phase transition may be not very much pronounced so that deviations from the hypergeometric model are small or even hardly pronounced, while minima (or maxima) in the bi-exponential prefactors will be much more marked. This suggests that the analysis of the LDFACF decay could be used to locate more efficiently the critical loading associated to the diffuse phase transition. Of course, the mean-field representation provides only an ideal decay where all the collective phenomena are neglected. Nevertheless the connection with the diffuse phase transition is a remarkable result, since even in such an essential representation a time-correlation function turns out to be a good indicator of a phase transition.

The scaling of the propagation probability by a homogeneous factor $\kappa^o = e^{-\beta\epsilon_{\mathrm{ki}}}$ in Eq. (A.53) such that $0 < \kappa^o \leq 1$ causes not simply the occupancy transition matrix $\mathbf{\Omega}^{(z)}$ (defined in Section A.3) to be scaled by the same factor, but instead the entire relaxation of differently occupied cells to undergo slight modifications. This can be seen clearly in the first two columns of Figure 7.9, while in the last two is shown that the time parameters τ_1 and τ_2 are unaffected by changes in κ^o.

Decay of local density fluctuations in numerical simulations. The same function has been analyzed from the output of numerical simulation. In some aspects, the resulting behavior is somewhat similar to what found in the mean-field case: for instance, decreases in temperature enhance the difference between the two timescales and the intensity of the peaks correspondingly to the diffuse phase transition. Nevertheless, some remarkable differences deserve some comments.

First of all, numerical simulations with all equivalent sites do not show a single-exponential relaxation as in the mean-field case. The presence of two timescales in the decay of the occupancy fluctuation seems to be an intrinsic feature of the collective correlations which are of course totally lost in the mean-field representation: this can be seen clearly in the trends of the parameters b_1 and b_2 in Figure 7.11 (for the case of the memoryless randomization), which do not approach 0 nor 1 for very low or very high loadings, as instead was observed for the mean-field case. Enlarging the system will not change that feature.

The most interesting aspect is the change in the relaxation dynamics when

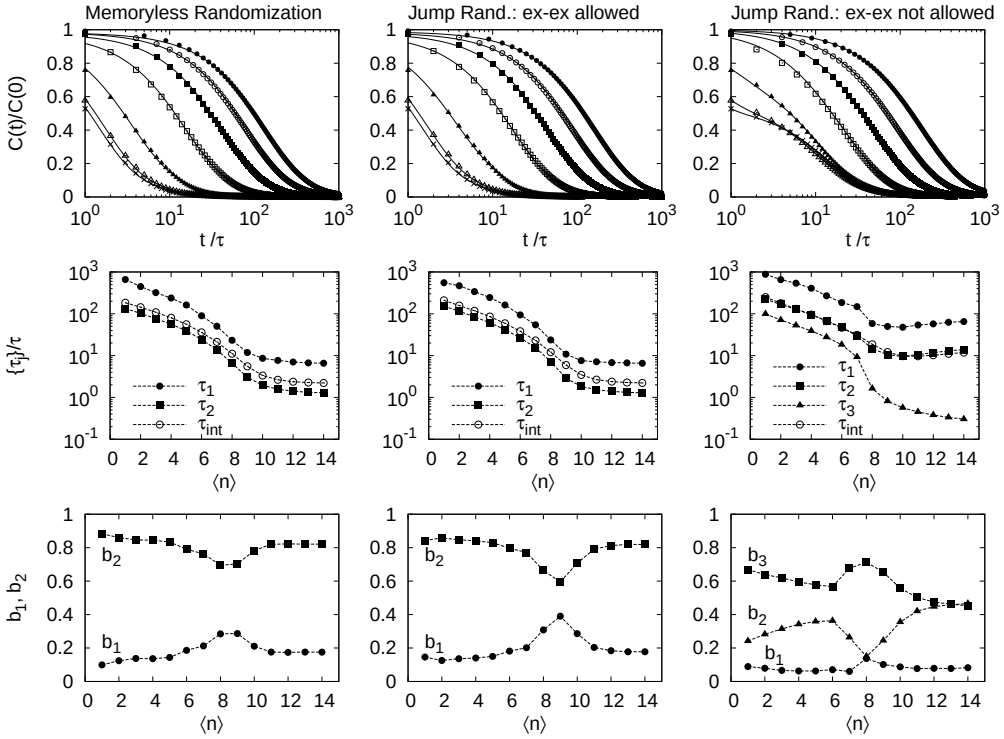

Figure 7.11: Decay of the LDFACF and related parameters for numerical simulations with $f^o_{\rm ex} - f^o_{\rm in} = 10$ kJ mol^{-1} at $T = 300$ K, when $K_{\rm ex} = 6$ and $K_{\rm in} = 9$. In each figure, fits for increasing integer loadings are plotted with solid lines, while symbols represent data from simulations: (●): $\langle n \rangle = 1$, (○): $\langle n \rangle = 3$, (■): $\langle n \rangle = 5$, (□): $\langle n \rangle = 7$, (▲): $\langle n \rangle = 9$, (△): $\langle n \rangle = 11$, (×): $\langle n \rangle = 13$.)

passing from a simulation without correlations to a simulation where instead the correlations play a significant role. If such correlations are present but weak, as in the case of the 'allowed ex-ex jumps' illustrated in Section 3.2.2, then the relaxation process appears to be similar to the uncorrelated case and the best fit is provided by a double exponential trend. If instead an inner topology is imposed, as in the case of 'not allowed ex-ex jumps', then the relaxation turns out to be much more complicated, and two main facts will emerge: (i) First of all, a double exponential is no longer appropriate to fit the relaxation curve, while a triple exponential becomes much more suitable:

$$C(t)/C(0) \sim \sum_{j=1}^{3} b_j e^{-t/\tau_j}, \tag{7.4}$$

where $\sum_j b_j = 1$. (ii) Secondly, the behavior of the relaxation time changes.

The integrated relaxation time, defined as

$$\tau_{\text{int}} = \sum_{j=1}^{z} b_j \tau_j, \tag{7.5}$$

where $z = 2$ for the bi-exponential and $z = 3$ for the tri-exponential fit, provides an essential information regarding the extent in time of the memory each cell has of its previous occupancies. In the less correlated cases the observed τ_{int} was monotonically decreasing with increasing the loading (first two columns of Figure 7.11): this was exactly the expected trend, since the chemical diffusion coefficient, D_{chem}, is inversely proportional to the relaxation time[85] (as seen in Section 5), and given that for $f^o_{\text{ex}} - f^o_{\text{in}} > 0$ the chemical diffusivity has an 's-shaped' increasing trend, the parameter τ_{int} is expected to decrease. In the more correlated case instead (where the ex-ex jumps are forbidden), adding adsorbate to the system above some (high) loading does not cause τ_{int} to decrease anymore and the parameter reaches a plateau. This is a dramatic sign of how an increase in self-correlations can induce high memory effects in the collective properties also. The phenomenon can be explained as follows: at high loadings the fact that the guests occupying the exit sites cannot access to other exit sites in the same cell becomes a critical condition in which the mobility is highly reduced. In this situation almost all the inner sites are filled, and direct jumps to other exit sites of the cell, that were energetically favoured in the less correlated cases, here are impossible. As a consequence, each exit sites can remain blocked for a very long time, not allowing the cell to vary its occupancy at all.

The situation becomes even more complicated if the site free-energies are reversed. In such a case the chemical diffusivity is decreasing; consequently, the relaxation parameter τ_{int} is expected to increase. In fact, as shown in Figure 7.12 this is exactly what happens for the memoryless randomization, but the trend changes dramatically for the more complicated randomization scheme where the ex-ex jumps are not allowed. When compared with the case of the memoryless randomization, at the lowest loadings the local fluctuations appear to be much more correlated due to the backscattering effect, that under these circumstances becomes very important. Increasing the loading causes correlations to decrease since more guests are present and the number of migrations increases, then above some threshold the correlations increase again, since the exit sites are almost all occupied and the migrations are very few.

Many of the aspects of the relaxation processes illustrated here were impossible to expect while designing the model. They emerged spontaneously in spite of the highly reductionistic character of the model. This makes the

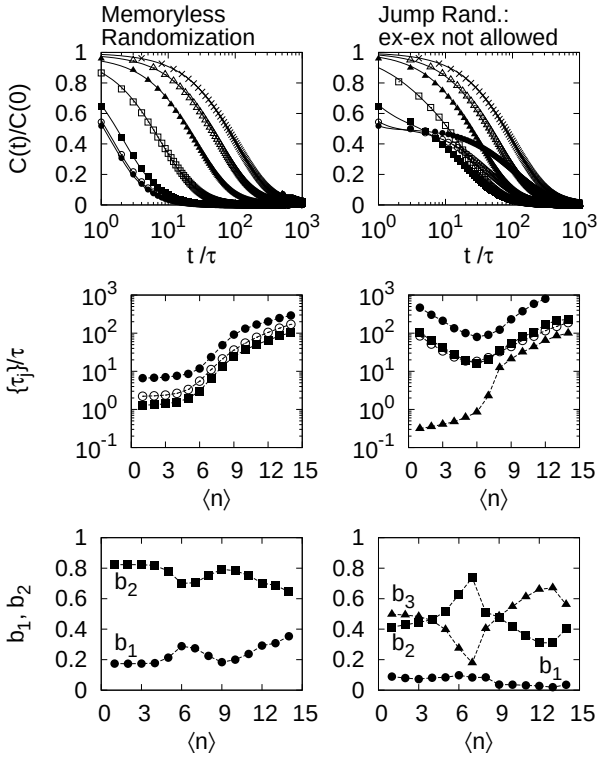

Figure 7.12: Decay of the LDFACF and related parameters for numerical simulations with $f^o_{\rm ex} - f^o_{\rm in} = -10$ kJ mol^{-1} at $T = 300$ K, when $K_{\rm ex} = 6$ and $K_{\rm in} = 9$. Fits are plotted with solid lines, while symbols represent data from simulations: (•): $\langle n \rangle = 1$, (○): $\langle n \rangle = 3$, (■): $\langle n \rangle = 5$, (□): $\langle n \rangle = 7$, (▲): $\langle n \rangle = 9$, (△): $\langle n \rangle = 11$, (×): $\langle n \rangle = 13$.)

model not only a cheap route to design coarse-graining reductions of real systems, but also a useful tool to investigate the deep reasons of confinement along with a purely statistical mechanical approach. Anyway, although the ThPCA with its simplified structure allows to sketch the essential experimental facts (competition mechanism for adsorption, loading dependence of diffusivity, backscattering and correlation), explaining all the complexities arising from confinement seems not to be straightforward even in this lattice model.

Chapter 8

Conclusions

The construction of an efficient model able to reproduce the main experimental facts required to identify the basic physics from a detailed analysis of the sorption properties of molecules confined in a zeolite crystal. Two main questions have been taken into account namely: (i) how to describe in a meaningful way the microstructure and the pore network connectivity of the zeolite crystal, and (ii) how to model the external fields due to the confining geometries. Answer to the first question has been found in the mapping of the zeolitic framework on a regular geometry of connected and structured pores (cells). The answer to the second question required the model to include the proper local interactions in order to reproduce adsorption isotherms and diffusivity. In these systems accurate determination of intermolecular forces to be used in simulations remains a largely unsolved problem and adsorption isotherms have been used to extract the data needed to obtain a mean molecular interaction potential within the cells. The model developed in the present thesis makes use of an automatic partitioning of the space in order to efficiently couple the locality of cellular automata together with the flexibility of static models of self-interacting cell, which in literature have been proven to be very efficient in reproducing the macroscopic properties of simple interacting systems by means of simple statistical-mechanical equations. Therefore it has been called *Thermodynamic Partitioning Cellular Automaton* (ThPCA). In such a reductionistic model, the molecules can hop from one cell to another, according to well defined rules satisfying detailed balance, the diffusion coefficient being deduced from their mean square displacement *vs.* time. It has been demonstrated that such a coarse-grained description representing the microscopic system at the mesoscopic length scale (10-1000 nm) can capture the proper density distribution while retaining microscopic information on particle fluctuations and dynamics. In the pore network modeled in this thesis, essentially an LTA type, a molecule can jump from the

parent cell to one of the six neighboring cell, and the diffusion behavior will be Einsteinian over time scales that are sufficiently large to allow molecules to move between several coarse-grained cells. On shorter time-scales in a real zeolite, the motion may be more complex, since molecules feel the interactions with the confining walls and possibly with other sorbed molecules within the same cage. The molecules wait in the cage until they gain sufficient energy to move to a neighboring cage. In the ThPCA this experimental fact has been reproduced through the action of the inner sites where molecules could be trapped until they are stochastically allowed to gain an exit site. In this way the necessary computer time can be reduced by orders of magnitude for zeolitic systems where rare event dynamics driven by strong guest-guest and guest-host interactions make full-scale molecular dynamics simulations still prohibitively expensive. We conclude that the developed ThPCA can be classified as a mesoscopic dynamical system able to give the correct behavior of various experimental systems that have been investigated in recent years, bringing out very clearly the essential elements involved in the adsorption and diffusion processes. Of course it cannot replace in any way MD simulations or any other atomistic method in the detailed description of the properties of a complex system like a zeolite since microscopic information about e.g. the framework composition and the structure of the sorbed molecules is lost. On the contrary, the ThPCA model can be interpreted as a complementary tool able to retain the essential physical and chemical properties of molecules sorbed in zeolites by averaging over less important detailed information at the atomistic level, which can be acquired through more sophisticated microscopic theoretical approaches and/or experimental measurements. The implementation of a systematic procedure to derive the ThPCA interaction parameters and kinetic barriers directly from atomistic calculations over a small zeolite's portion would add self-consistency to the model. In this thesis it has been shown that when constructed from a fitting procedure of experimental data the model is flexible enough to allow a coherent representation of the essential (thermodynamic and dynamic) properties of a host-guest system. Therefore the ThPCA can be proposed as a proper environment for a coarse-grained representation of such systems. Moreover, it results that a well-designed coarse-grained description of a microporous material is not simply a way to save computer time, but a route to focus on the main characteristics of sorption and diffusion processes, where details may be unimportant and nonessential. The advantage of the present model compared with other methods is that it provides a simplified description of the microscopic dynamics of guests adsorbed in micropore, yet most of the interesting details are intrinsically present and sometimes it is easier to see important correlations that are obscured in other methods. In this context

the ThPCA can be considered as the computationally cheapest model that has the essential properties of guests adsorbed in a micropore.

Many are the possible developments of the model. First of all, the modeling strategy developed in the present thesis could be exploited for designing coarse-grained representations of diffusion of guests in various zeolite topologies (also including defects), in and out od equilibrium. Secondly, the application of the model to the study of mixtures of guests adsorbed in zeolites will be of remarkable interest since it could find applications as test model in problems of separation and catalysis. Finally, extending the applicability to guest species of more complex molecular structure (like *n*-alkanes) would represent another important step towards the formulation of a complete coarse-grained theory able to embed in a single model a very wide variety of guest-host systems under a reductionistic approach.

8.1 Acknowledgements

All the work presented in this thesis has been carried out with financial support provided by Italian Ministero dell'Istruzione, dell'Università e della Ricerca, by Università degli Studi di Sassari, and by Istituto Nazionale per la Scienza e Tecnologia dei Materiali (INSTM), which are acknowledged. This work makes use also of results produced by the Cybersar Project managed by the Consorzio COSMOLAB, a project co-founded by the Italian Ministry of University and Research (MIUR) within the Programma Operativo Nazionale 2000-2006 "Ricerca Scientifica, Sviluppo Tecnologico, Alta Formazione" per le Regioni Italiane dell'Obiettivo 1 (Campania, Calabria, Pugla, Basilicata, Sicilia, Sardegna)—Asse II, Misura II.2 "Società dell'Informazione", Azione a "Sistemi di calcolo e simulazione ad alte prestazioni". More information is available at http://www.cybersar.it.

Appendix A

Mean field theory

A.1 Single cell at equilibrium: grand-canonical formulation

The properties of the model can be studied through the analysis of some probability distribution related to the lattice partition function $\mathcal{Q}^{\mathscr{L}}$ defined in Eq. (2.5). The probability distribution of some global observable, like the energy of the whole lattice, could in principle be formulated as a function of the terms contributig to the summatory in $\mathcal{Q}^{\mathscr{L}}$, and the systems behavior could be investigated through the analysis of such a distribution. Anyway, the direct use of $\mathcal{Q}^{\mathscr{L}}$ would make such a computation an impossible task for large systems. Partition functions restricted to a single cell are to be preferred to define the statistical distributions the model obeys. If correlations among different cells are negligible (and this is the case for the model considered here), and if the lattice contains a large number of cells, then such partition functions will provide information about the equilibrium properties of the entire system.

The thermodynamic approach allows a mean-field theory to be developed for the ThPCA under the assumption that occupancies are distributed in the neighborhood of an n-occupied cell according with the same occupancy probability distribution of the entire system. Such a distribution, denoted $\mathbf{p} = \{p(n)\}_{n=0,\dots,K}$, where $p(n)$ denotes the probability of a cell to be occupied by n guests, depends strongly on the loading (average occupancy) $\langle n \rangle = N/M$. Each loading is related to a particular value of the chemical potential μ according to the *adsorption isotherm* of the system. While in numerical simulations of ThPCA $\langle n \rangle$ is constant, in the mean-field approach it is more suitable to work with constant μ (where the subscript GC stands for 'grand-

canonical ensemble'), in order to compute the occupancy distribution as

$$p(n) = \frac{Q(n)e^{\beta\mu n}}{\Xi(\mu)} \tag{A.1}$$

where $Q(n)$ is the canonical partition function for a closed cell of occupancy n and the denominator,

$$\Xi(\mu) = \sum_{n=0}^{K} Q(n)e^{\beta\mu n}, \tag{A.2}$$

is the grand-canonical partition function of the open cell. The loading, $\langle n\rangle$, and the occupancy variance, $\langle n^2\rangle - \langle n\rangle^2$, can then be recovered through the first and second moment of $p(n)$:

$$\langle n\rangle = \sum_{n=0}^{K} n\, p(n) = \frac{1}{\beta}\frac{\partial \ln \Xi(\mu)}{\partial \mu}, \tag{A.3}$$

$$\langle n^2\rangle - \langle n\rangle^2 = \sum_{n=0}^{K} (n - \langle n\rangle)^2\, p(n) = \frac{1}{\beta}\frac{\partial \langle n\rangle}{\partial \mu}. \tag{A.4}$$

The *intercell transfer factor*, $W(n,m)$, (introduced in Section 6.2.1) can be defined as the conditional probability to observe migration of one guest from an n-occupied cell to an m-occupied neighboring cell:

$$W(n,m) = \frac{1}{{K_{\text{ex}}}^2} \sum_{n_{\text{ex}}=0}^{K_{\text{ex}}} \sum_{m_{\text{ex}}=0}^{K_{\text{ex}}} n_{\text{ex}}(K_{\text{ex}} - m_{\text{ex}}) \times P(n_{\text{ex}}|n)P(m_{\text{ex}}|m)p^P(n_{\text{ex}}, n, m_{\text{ex}}, m) \tag{A.5}$$

defined for $n = 1, \ldots, K$ and $m = 0, \ldots, K-1$ with $P(n_{\text{ex}}|n)$ given in Eq. (2.29), and

$$p^P(n_{\text{ex}}, n, m_{\text{ex}}, m) = \frac{e^{-\beta\epsilon_{\text{ki}}(n,m)}}{1 + e^{\beta\Delta F(n_{\text{ex}},n,m_{\text{ex}},m)}} \tag{A.6}$$

as the propagation probability defined in Eq. (3.22). In Eq. (A.6) the term

$$\Delta F(n_{\text{ex}}, n, m_{\text{ex}}, m) = \Phi(n_{\text{ex}} - 1, n - n_{\text{ex}}) + \Phi(m_{\text{ex}} + 1, m - m_{\text{ex}}) - \Phi(n_{\text{ex}}, n - n_{\text{ex}}) - \Phi(m_{\text{ex}}, m - m_{\text{ex}}) \tag{A.7}$$

is the free-energy difference between the pre- and the post-propagation state. From the occupancy distribution and the intercell transfer factor one can derive the equilibrium probability of a guest to escape from its current cell as

$$p^{\mathrm{esc}} = \frac{K_{\mathrm{ex}}}{\langle n \rangle} \sum_{n=1}^{K} \sum_{m=0}^{K-1} p(n)p(m)W(n,m) \tag{A.8}$$

$$= \frac{\tau}{\tau_{\mathrm{mrt}}} \tag{A.9}$$

where the equivalence (A.9) gives the relation between the escape probability and the is the mean residence time, τ_{mrt}, of a guest inside of a cell (expressed in time units).

Let us consider now the discrete self-diffusivity as a function of the DACF reported in Eq. (4.23). Its value at $t = 0$ i.e. $\langle \delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0) \rangle$ is λ^2 times the escape probability given in Eq. (A.8). Under the assumption of uncorrelated motion then the summation in Eq. (4.23) is null and the mean-field self-diffusivity results:

$$D_0^{\mathrm{mf}} = \frac{\lambda^2}{\tau} \frac{K_{\mathrm{ex}}}{2d\langle n \rangle} \sum_{n=1}^{K} \sum_{m=0}^{K-1} p(n)p(m)W(n,m), \tag{A.10}$$

which turns out to be related with the mean residence time introduced in Eq. (A.9) according to $D_0^{\mathrm{mf}} = \lambda^2 / 2d\tau_{\mathrm{mrt}}$. The quantity D_0^{mf} is exactly the *memoryless diffusivity* (see Section 3.2.1) produced in numerical simulations where the memoryless randomization is used.

A.1.1 Non-interacting case

Several simplifications are possible in the absence of guest-guest interactions.

In Appendix B.1 it is shown that in such a case Ξ reads

$$\Xi(\mu) = \left(1 + e^{\beta(\mu - f_{\mathrm{ex}}^o)}\right)^{K_{\mathrm{ex}}} \left(1 + e^{\beta(\mu - f_{\mathrm{in}}^o)}\right)^{K_{\mathrm{in}}}, \tag{A.11}$$

and that the thermodynamic equilibrium of the model is represented by a dual-site Langmuir isotherm:[43]

$$\langle n \rangle = K_{\mathrm{ex}} \frac{e^{-\beta f_{\mathrm{ex}}^o} \lambda_a}{1 + e^{-\beta f_{\mathrm{ex}}^o} \lambda_a} + K_{\mathrm{in}} \frac{e^{-\beta f_{\mathrm{in}}^o} \lambda_a}{1 + e^{-\beta f_{\mathrm{in}}^o} \lambda_a}, \tag{A.12}$$

where K_{ex} is the capacity for adsorption in the exit sites, K_{in} is the capacity for adsorption in the inner sites, $\lambda_a = e^{\beta\mu}$ is an absolute activity, and the

weights $e^{-\beta f^o_{\rm ex}}$ and $e^{-\beta f^o_{\rm in}}$ play the role of the equilibrium constants for adsorption in the exit sites and in the inner sites, respectively. In Eq. (A.12) the probability of an exit site to be occupied, $\rho_{\rm ex}$, and the probability of an inner site to be occupied, $\rho_{\rm in}$, are defined as:

$$\rho_{\rm ex} = \frac{e^{-\beta f^o_{\rm ex}}\lambda_a}{1+e^{-\beta f^o_{\rm ex}}\lambda_a},\ \rho_{\rm in} = \frac{e^{-\beta f^o_{\rm in}}\lambda_a}{1+e^{-\beta f^o_{\rm in}}\lambda_a}. \tag{A.13}$$

Inserting Eq. (A.11) in Eq. (A.1) and using the relations in Eq. (A.13) one obtains the following probability distribution of occupancies:

$$\begin{aligned} p(n) = &\sum_{n_{\rm ex}=0}^{K_{\rm ex}} \binom{K_{\rm ex}}{n_{\rm ex}} (\rho_{\rm ex})^{n_{\rm ex}} (1-\rho_{\rm ex})^{K_{\rm ex}-n_{\rm ex}} \\ &+ \binom{K_{\rm in}}{n-n_{\rm ex}} (\rho_{\rm in})^{n-n_{\rm ex}} (1-\rho_{\rm in})^{K_{\rm in}-n+n_{\rm ex}}, \end{aligned} \tag{A.14}$$

which arises from the composition of the probability distributions in two subsystems (of exit and inner sites, respectively) exchanging guests.

When the temperature is infinite (an ideal situation equivalent to assigning the same statistical weight to both kinds of adsorption site), then $\Xi(\mu) = (1+\lambda_a)^K$ and Eq. (A.12) reduces to the single-site Langmuir isotherm $\langle n\rangle = K\lambda_a/(1+\lambda_a)$. For the equilibrium distribution $\mathbf{p}$, this is the same as to consider $f^o_{\rm in} = f^o_{\rm ex}$, so that the system is trivially microcanonical and the mutual exclusion is the only restraint to the guest distribution in the lattice. Therefore the distribution $\mathbf{p}$ converges to an hypergeometric distribution,[41] denoted $\mathbf{p}^{\rm hyp}$ (see Appendix B.1 for details).

On the other hand, the entire system is canonical when $f^o_{\rm in} \neq f^o_{\rm ex}$. In this case energy effects add to the mutual exclusion, therefore in general $\mathbf{p}^{\rm eq} \neq \mathbf{p}^{\rm hyp}$ for $T < \infty$.

In the absence of guest-guest interactions also the uncorrelated self-diffusivity achieves a simpler formulation, and Eq. (A.10) becomes:

$$D_0^{\rm mf} = \frac{\lambda^2}{\tau}\frac{1}{2d}\frac{K_{\rm ex}\rho_{\rm ex}}{\langle n\rangle}(1-\rho_{\rm ex})\frac{e^{\beta\epsilon_{\rm ki}}}{2}, \tag{A.15}$$

where $\epsilon_{\rm ki}$ is the fixed kinetic barrier.

A.2 Mean-field displacement autocorrelation function

The present Appendix is devoted to the developement of a mean-field representation of the correlations arising as a consequence of the jump randomization scheme introduced in Section 3.2.2. In order to facilitate the reading, some of the concepts already described in the main body of the present thesis will be repeated.

Cell configurations of distinguished guests. Numerical simulations showed that the correlation effects arising from the application of the jump randomization are well represented by the displacement autocorrelation function (DACF), which in a discrete model can be interpreted as the discrete-counterpart of the velocity autocorrelation function (VACF) of a diffusing atom or molecule in a continuous system. In the present Appendix a mean-field analysis of the DACF is presented which is aimed to trace back the origin of correlations in the self-motion and to obtain a qualitative mean-field expression for the self-diffusivity emphasizing the main contributions of memory effects to the effective mobility of each diffusing guest.

While in the memoryless randomization all the cell configurations preserving the cell occupancy are available as output configurations, during the jump randomization the set of accessible outputs is much smaller. As a result, while the memoryless randomization pushes each cell towards a condition of local equilibrium, by construction the configuration changes occurring by means of the jump randomization are much less pronounced. This allows memory effects to arise spontaneously during the time evolution of the system.

In Chapter 2 the configuration of a generic cell $\mathbf{r}$ (implicitly referred to a generic time t) has been defined as

$$\boldsymbol{\eta}(\mathbf{r}) = \{\eta_i(\mathbf{r})\}, \qquad i = 1, \ldots, K \tag{A.16}$$

where η_i is a boolean variable which has value 1 if according to configuration $\boldsymbol{\eta}$ the i-th site of the cell is occupied, and 0 if unoccupied. With such a definition, Eq. (A.16) is a *configuration of occupied sites* and contains no information regarding the guests' identities. In words, the identity-less configuration $\boldsymbol{\eta}$ will be referred to as 'η-configuration'.

Identities are taken into account by the following σ-configuration:

$$\boldsymbol{\sigma}(\mathbf{r}) = \{\sigma_{iI}\}, \qquad i = 1, \ldots, K \vee I = 1, \ldots, N \tag{A.17}$$

where N is the number of guests, and σ_{iI} has value 1 if the I-th guest of the system is located at the i-th site of cell $\mathbf{r}$, and 0 otherwise.

Let us consider a single, closed cell with configuration $\boldsymbol{\sigma}$ *before* a randomization operation. Let us assume $\boldsymbol{\sigma}$ to be a particular realization of the identity-less configuration $\boldsymbol{\eta}$. Indicating with K the number of sites in the cell, and with n the numer of occupied sites, there are exactly $\Omega_\sigma = K!/(K-n)!$ possible output (i.e., post-randomization) σ-configurations, which can be grouped into $\Omega_\eta = K!/n!(K-n)!$ η-configurations. Now, the memoryless randomization selects an output η-configuration, $\boldsymbol{\eta}^R$, among all the Ω_η with a probability proportional to $\exp[-\beta F(\boldsymbol{\eta}^R)]$, where F is the cell free-energy function. After that, the guest names are assigned randomily, so all the possible $n!$ identity configurations have the same probability to be selected. As a result, once a value for the numer of sites per cell K has been set, the number of accessible output configurations Ω_η depends only on the cell-occupancy n and not on the particular configuration of occupied sites.

Everything changes under the jump randomization approach, where the number of possible outputs for a given input configuration is limited by the fact that each guest must overcome a barrier to reach a different site in the cell (so that the probability to change from $\boldsymbol{\eta}$ to $\boldsymbol{\eta}^R$ is no longer proportional to $\exp[-\beta F(\boldsymbol{\eta}^R)]$ but varies with the configuration path performed during the randomization operation), and that two guests' identities cannot be exchanged while keeping their position unvaried. Since at each time step the jump randomization moves sequentially the guests from site to site while preserving their identities the entire migration process contains memory effects. In Chapter 7 it has been shown via numerical simulations that such an approach affects the mobility of the guest so that the self-diffusivity results to be less than what expected in the absence of correlations (as it would be in the case of the memoryless randomization), and in some cases it causes relevant changes in the trend of self-diffusivity *vs* loading as observed when exit sites are more binding than inner sites.

A deeper understanding of the correlations introduced by the jump randomization required an analysis of the migration mechanism, in terms of the probability connected with every possible move a tagged guest can perform during its entire diffusion process. From a probabilistic point of view, the process of migration of a tagged guest is a sequence of events causing the guest to move from cell to cell and from site to site inside of each cell. Each event has a certain probability to occur which is in principle *dependent* on the previous moves of the guest itself. Moreover, a dependency exists also on the moves of all the surrounding guests, so that it would be very difficult (and somewhat useless) to calculate exactly the probability of every event. A mean-field approach must therefore be used to derive readable equations

linking correlations in motion to macroscopic observables (e.g. densities, rate of transfers, etc.).

A.2.1 Basic probability distributions and averages

Let us briefly revise the main quantities involved in the diffusion process.

The *instantaneous cell-to-cell displacement* of a single guest is given by

$$\delta\mathbf{r}_I(t) = \mathbf{r}_I(t+\tau) - \mathbf{r}_I(t) \tag{A.18}$$

where I denote the guest's identity and $\mathbf{r}_I(t)$ is its position (or, equivalently, the position of its host cell taking account of periodic boundary conditions) at time t. A hypercubic lattice shall be assumed so that Eq. (A.18) is equivalent to Eq. (4.19). From now on in the present Appendix, the guest identity, I, will be kept as implicit to maintain the notation as simple as possible.

The self-diffusivity is related to the displacement autocorrelation function (DACF), given by $\langle\delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle$ (where $z \geq 0$ is an integer and τ is the duration of a time step) via the Eq. (4.23). Such a quantity can be used in a very similar way as the velocity autocorrelation function of a guest in a continuous system. Its zero-time value is given by

$$\begin{aligned}\langle\delta\mathbf{r}(0)\cdot\delta\mathbf{r}(0)\rangle =& \lambda^2\frac{1}{\langle n\rangle}\sum_{n_{\mathrm{ex}}=1}^{K_{\mathrm{ex}}}\sum_{n_{\mathrm{in}}=0}^{K_{\mathrm{in}}}\sum_{m_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}-1}\sum_{m_{\mathrm{in}}=0}^{K_{\mathrm{in}}} n_{\mathrm{ex}}\left(1-\frac{m_{\mathrm{ex}}}{K_{\mathrm{ex}}}\right)\\ &\times p(n_{\mathrm{ex}},n_{\mathrm{in}})\,p(m_{\mathrm{ex}},m_{\mathrm{in}})\,\kappa(n_{\mathrm{ex}},n_{\mathrm{in}},m_{\mathrm{ex}},m_{\mathrm{in}})\end{aligned} \tag{A.19}$$

where $\langle n\rangle$ is the loading (average number of occupied sites in a cell), $p(n_{\mathrm{ex}},n_{\mathrm{in}})$ is the *total probability* of a cell to be occupied by n_{ex} guests in the exit sites and n_{in} guests in the inner sites:

$$p(n_{\mathrm{ex}},n_{\mathrm{in}}) = [\Xi(\mu)]^{-1}\binom{K_{\mathrm{ex}}}{n_{\mathrm{ex}}}\binom{K_{\mathrm{in}}}{n_{\mathrm{in}}}e^{\beta(n_{\mathrm{ex}}+n_{\mathrm{in}})\mu}e^{-\beta F(n_{\mathrm{ex}},n_{\mathrm{in}})} \tag{A.20}$$

with $\Xi(\mu)$ as the grand-canonical partition function of a single cell when the chemical potential of the system is μ. In Eq. (A.20) the quantity

$$\kappa(n_{\mathrm{ex}},n_{\mathrm{in}},m_{\mathrm{ex}},m_{\mathrm{in}})$$

is the *conditional migration rate*, that is the conditional probability of a tagged guest to migrate (during propagation) to the $(j+d)$-th exit site of the neighboring cell $\mathbf{r}^j$, given that (i) the tagged guest is located at the j-th site of its current guest cell $\mathbf{r}$, (ii) the cell $\mathbf{r}$ has respectively n_{ex} and n_{in} filled exit and inner sites including the one occupied by the tagged guest, and (iii)

the cell $\mathbf{r}^j$ has respectively $m_{\rm ex}$ and $m_{\rm in}$ filled exit and inner sites but (iv) its $(j+d)$-th exit site (the target site) is empty, i.e. $\eta_{j+d}(\mathbf{r}^j)=0$:

$$\begin{aligned}\kappa(n_{\rm ex},n_{\rm in},m_{\rm ex},m_{\rm in}) &= \exp\left[-\beta\epsilon_{\rm ki}(n,m)\right] \\ &\times\Big(1+\exp\big\{\beta\big[\Phi(n_{\rm ex}-1,n_{\rm in})+\Phi(m_{\rm ex}+1,m_{\rm in}) \\ &\quad -\Phi(n_{\rm ex},n_{\rm in})-\Phi(m_{\rm ex},m_{\rm in})\big]\big\}\Big)^{-1}, \quad \text{(A.21)}\end{aligned}$$

with $\epsilon_{\rm ki}(n,m)$ as the kinetic barrier to intercell migration, $n=n_{\rm ex}+n_{\rm in}$ and $m=m_{\rm ex}+m_{\rm in}$.

Another probability distribution function which will play a central role in our mean-field analysis is $g(n_\alpha-1,n_\nu|\{\alpha\})$ (where both α and ν can take values 'ex' and 'in'), that is the conditional probability of a cell with an α-type site already occupied to have n_ν filled sites of type ν and $n_\alpha-1$ of the remaining $K_\alpha-1$ filled sites of type α:

$$g(n_\alpha-1,n_\nu|\{\alpha\})=\frac{n_\alpha p(n_{\rm ex},n_{\rm in})}{\sum_{n'_{\rm ex}=1}^{K_{\rm ex}}\sum_{n'_{\rm in}=0}^{K_{\rm in}} n'_\alpha p(n'_{\rm ex},n'_{\rm in})}, \quad \text{(A.22)}$$

where $\{\alpha\}$ denotes the condition that there is a guest (i.e. the tagged particle) located in an α-type site, and the quantity

$$\frac{n_\alpha}{K_\alpha}p(n_{\rm ex},n_{\rm in})=[\Xi(\mu)]^{-1}\binom{K_\alpha-1}{n_\alpha-1}\binom{K_\nu}{n_\nu}e^{\beta(n_{\rm ex}+n_{\rm in})\mu}e^{-\beta F(n_{\rm ex},n_{\rm in})} \quad \text{(A.23)}$$

is the total probability of a cell to have occupied one particular α-type site, $n_\alpha-1$ of the remaining α-type sites, and n_ν sites of type ν.

The probability distribution in Eqs. (A.20) and (A.22) are very useful to compute averages of the observables involved in the intercell migration process. Such averages will represent the macroscopic observables the DACF will be related to. In fact, each of them represent a *probability* for a single guest, but since it is obtained by averaging observables over the probability distributions p and g then the notation will be used $\langle\,\cdot\,\rangle$ for them.

The probability of an α-site to be occupied is defined as

$$\langle\rho_\alpha\rangle=\sum_{n_{\rm ex}=0}^{K_{\rm ex}}\sum_{n_{\rm in}=0}^{K_{\rm in}}\frac{n_\alpha}{K_\alpha}\,p(n_{\rm ex},n_{\rm in}). \quad \text{(A.24)}$$

The conditional probability of an α-site in a cell to be empty given that there is one exit site occupied in the same cell is given instead by

$$\langle h_{\rm ex}\rangle = \sum_{n_{\rm ex}=1}^{K_{\rm ex}} \sum_{n_{\rm in}=0}^{K_{\rm in}} \left(1 - \frac{n_{\rm ex}-1}{K_{\rm ex}-1}\right) g(n_{\rm ex}-1, n_{\rm in}|\{\rm ex\}) \tag{A.25}$$

if $\alpha =$ 'ex', and

$$\langle h_{\rm in}\rangle = \sum_{n_{\rm ex}=1}^{K_{\rm ex}} \sum_{n_{\rm in}=0}^{K_{\rm in}} \left(1 - \frac{n_{\rm in}}{K_{\rm ex}}\right) g(n_{\rm ex}-1, n_{\rm in}|\{\rm ex\}) \tag{A.26}$$

if $\alpha =$ 'in'.

The probability of a guest to jump *from* an exit site during the jump randomization to any other site (independently of its type) of the cell shall be indicated as $\langle q_{\rm ex}\rangle$. The calculation of $\langle q_{\rm ex}\rangle$ will be carried out by supposing the tagged guest (located in a generic $\alpha =$ site, with $\alpha =$'ex', 'in') to point towards another exit or inner site during the jump randomization. In order to jump into it, it must: (i) Find it empty, (ii) Overcome the adsorption barrier of its host site, f_α^o, and (iii) Overcome an interaction free-energy barrier, $\Phi(\boldsymbol{\eta}') - \Phi(\boldsymbol{\eta})$ (where $\Phi(\boldsymbol{\eta})$ is the interaction free-energy of configuration $\boldsymbol{\eta}$, and $\boldsymbol{\eta}'$ is the configuration that the cell would reach after the jump move) to preserve the detailed balance. So in case of empty target site the total free-energy barrier to overcome will be $f_\alpha^o + \Phi(\boldsymbol{\eta}) - \max\left[\Phi(\boldsymbol{\eta}'), \Phi(\boldsymbol{\eta})\right]$. Since K_α/K is the probability to select an α site as target and since as it follows from the definition of the free-energy function $\Phi(n_{\rm ex}, n_{\rm in})$ if the guest is jumping between two sites of the same type then the cell free-energy will not change causing the interaction free-energy barrier to be zero, the result is

$$\langle q_\alpha\rangle = \sum_{\nu={\rm ex,in}} \frac{K_\nu}{K} \langle q_\alpha h_\nu\rangle \tag{A.27}$$

$$\tag{A.28}$$

where $\langle q_\alpha h_\nu\rangle$ are defined by the following relations, for α, ν taking values in the set $\{\rm ex, in\}$:

$$\langle q_\alpha h_\alpha\rangle = \gamma e^{\beta f_\alpha^o} \sum_{n_\alpha=1}^{K_\alpha} \sum_{n_\nu=0}^{K_\nu} \left(1 - \frac{n_\alpha - 1}{K_\alpha - 1}\right) g(n_\alpha - 1, n_\nu|\{\alpha\}), \tag{A.29}$$

and

$$\begin{aligned}\langle q_\alpha h_\nu\rangle =& \gamma e^{\beta f_\alpha^o} \sum_{n_\alpha=1}^{K_\alpha} \sum_{n_\nu=0}^{K_\nu} \left(1 - \frac{n_\nu}{K_\nu}\right) e^{\beta\Phi(n_\alpha,n_\nu)} e^{-\beta\max[\Phi(n_\alpha-1,n_\nu+1),\Phi(n_\alpha,n_\nu)]} \\ & \times g(n_\alpha - 1, n_\nu|\{\alpha\}), \qquad \alpha \neq \nu\end{aligned} \tag{A.30}$$

where γ is a scaling factor for randomization introduced in Eq. (3.12). The macroscopic quantity $\langle q_\alpha h_\nu \rangle$ can be interpreted as the average acceptance probability of a jump from a site of type α to a site of type ν).

Let us compute now the *transmission coefficient*, that is the probability that during propagation a guest located in an exit site effectively migrates into the neighboring cell. Before to do that, let us introduce a formalism which will simplify the final equations. If p indicates a probability, or a composition of probabilities (i.e. sum, product, etc.) then $\overline{p} = 1 - p$, while the notation $\langle p_1 p_2 \cdots p_u \rangle$ will be used in case the events connected with the u probabilities are treated as non-independent. Under this assumption the *transmission coefficient*, denoted $\langle \kappa \overline{\rho_{\mathrm{ex}}} \rangle$ is

$$\langle \kappa \overline{\rho_{\mathrm{ex}}} \rangle = \sum_{n_{\mathrm{ex}}=1}^{K_{\mathrm{ex}}} \sum_{n_{\mathrm{in}}=0}^{K_{\mathrm{in}}} \sum_{m_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}-1} \sum_{m_{\mathrm{in}}=0}^{K_{\mathrm{in}}} \left(1 - \frac{m_{\mathrm{ex}}}{K_{\mathrm{ex}}}\right) p(m_{\mathrm{ex}}, m_{\mathrm{in}}) \kappa(n_{\mathrm{ex}}, n_{\mathrm{in}}, m_{\mathrm{ex}}, m_{\mathrm{in}}) \times g(n_{\mathrm{ex}} - 1, n_{\mathrm{in}} | \{\mathrm{ex}\}). \tag{A.31}$$

If guests are non-interacting and the kinetic barrier ϵ_{ki} is homogeneous, then $\langle \kappa \overline{\rho_{\mathrm{ex}}} \rangle = \langle \kappa \rangle \overline{\rho_{\mathrm{ex}}}$, where $\langle \kappa \rangle = \frac{1}{2} \exp[-\beta \epsilon_{\mathrm{ki}}]$.

A.2.2 Mean-field evaluation of the displacement autocorrelation function

Let us suppose that at time zero (i.e. $z = 0$) the tagged guest has moved from its host cell, say $\mathbf{r}$, to the neighboring cell along the direction $\mathbf{e}_j$, i.e. $\mathbf{r}^j$, so that $\delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0) = 1$ and the guest find itself in the exit site $(\mathbf{r}^j, j + d)$. That represents the starting point for the evaluation of probabilities of all subsequent events. In this approach the choice of the cubic topology turns out to be the most suitable since a non-zero value of $\delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0)$ is obtained only if at the z-th time step the guest migrates to a cell along the direction of $\mathbf{e}_j$ (so that $\delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0) = 1$) or $\mathbf{e}_{j+d} = -\mathbf{e}_j$ (so that $\delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0) = -1$). Therefore, $\langle \delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0) \rangle / \langle \delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0) \rangle$ represents the conditional probability of a guest to migrate at time $z\tau$ in the same direction of displacement at time 0, given that at time 0 the displacement was not null, minus the conditional probability of a migration in the opposite direction. The calculation will be carried out for the first two time steps after the initial migration event and then generalized to an arbitrary time step.

Let us introduce a notation to denote the various events involved in the self-motion. At the initial time, $z = 0$:

$\Diamond$ The guest reaches any of the K_{ex} exit sites of the current cell during randomization, and subsequently migrates to the corresponding neighboring cell during propagation.

By assuming $\mathbf{e}_j$ to be the migration direction at time 0, the following events at any other time step $z \geq 1$ will be specified:

$\rightarrow$ The guest reaches the j-th exit site of the current cell during randomization (i.e. it jumps in if it occupies a different site, or it stays in if it already occupies it) and subsequently *cannot* migrate to the neighboring cell in the $\mathbf{e}_j$ direction during propagation.

$\Rightarrow$ The guest reaches the j-th exit site of the current cell during randomization, and subsequently *migrates* to the neighboring cell in the $\mathbf{e}_j$ direction during propagation.

$\leftarrow$ The guest reaches the $(j + d)$-th exit site of the current cell during randomization and subsequently *cannot* migrate to the neighboring cell in the $\mathbf{e}_j$ direction during propagation.

$\Leftarrow$ The guest reaches the $(j + d)$-th exit site of the current cell during randomization, and subsequently *migrates* to the neighboring cell in the direction $-\mathbf{e}_j$ during propagation.

$\updownarrow$ The guest reaches an exit site different from the j-th and the $(j+d)$-th during randomization and subsequently *cannot* migrate to a neighboring cell during propagation.

$\Updownarrow$ The guest reaches an exit site different from the j-th and the $(j+d)$-th during randomization and subsequently *migrates* to a neighboring cell during propagation.

$\bigcirc$ The guest reaches an inner site of the current cell during randomization.

Case one: allowed ex-ex jumps

The calculation of the DACF will be extensively illustrated for the case in which during the jump randomization the jumping guest can select any of the K sites of the cell as a target site.

	Ex-ex jumps allowed
$p(\Rightarrow \vert\Diamond)$	$= (1/K)\langle q_{\rm ex} h_{\rm ex}\rangle \langle \kappa\overline{\rho_{\rm ex}}\rangle$
$p(\rightarrow \vert\Diamond)$	$= (1/K)\langle q_{\rm ex} h_{\rm ex}\rangle \left(1 - \langle \kappa\overline{\rho_{\rm ex}}\rangle\right)$
$p(\Leftarrow \vert\Diamond)$	$= (1 - \gamma_{\rm ex} - \gamma_{\rm in}) \langle \kappa\overline{\rho_{\rm ex}}\rangle$
$p(\leftarrow \vert\Diamond)$	$= (1 - \gamma_{\rm ex} - \gamma_{\rm in}) \left(1 - \langle \kappa\overline{\rho_{\rm ex}}\rangle\right)$
$p(\Updownarrow \vert\Diamond)$	$= [(K_{\rm ex} - 2)/K]\langle q_{\rm ex} h_{\rm ex}\rangle \langle \kappa\overline{\rho_{\rm ex}}\rangle$
$p(\updownarrow \vert\Diamond)$	$= [(K_{\rm ex} - 2)/K]\langle q_{\rm ex} h_{\rm ex}\rangle \left(1 - \langle \kappa\overline{\rho_{\rm ex}}\rangle\right)$
$p(\bigcirc\vert\Diamond)$	$= (K_{\rm in}/K)\langle q_{\rm ex} h_{\rm in}\rangle$

Table A.1: Probability values for events of jump starting from initial condition $\Diamond$ at time 0 for the case of jump randomization with allowed ex-ex jumps, where $\gamma_{\rm ex} = [(K_{\rm ex} - 1)/K]\langle q_{\rm ex} h_{\rm ex}\rangle$ is the probability of the guest to jump into an exit site different from the departure one, and $\gamma_{\rm in} = (K_{\rm in}/K)\langle q_{\rm ex} h_{\rm in}\rangle$ is the probability to jump to an inner site.

Time $t = 0$. Since at the initial time $\delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0) = 1$ only if the guest migrates to a neighboring cell, then $p(\Diamond) = 2d\lambda^{-2}\tau D_0^{\rm mf}$ where

$$D_0^{\rm mf} = \frac{1}{2d\tau} \langle \delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0)\rangle \tag{A.32}$$

is the mean-field diffusivity in the absence of memory effects in the guest motion, with $\langle \delta\mathbf{r}(0) \cdot \delta\mathbf{r}(0)\rangle$ as given in Eq. (A.19).

Time $t = \tau$. After one time step (i.e. $z = 1$), the guest has just propagated from cell $\mathbf{r}$ to $\mathbf{r} + \mathbf{e}_j$, and find itself in the $(j + d)$-th exit site of the latter cell, indicated $(\mathbf{r} + \mathbf{e}_j, j + d)$. The events which are possible to happen from this position will be now listed. Mathematical formulas for the related (non-null) probabilities can be found in Table A.1 in terms of the averages defined in Section A.2.1.

$(\Diamond, \Rightarrow)$: The guest jumps from $(\mathbf{r}+\mathbf{e}_j, j+d)$ to $(\mathbf{r}+\mathbf{e}_j, j)$ with subsequent propagation to $(\mathbf{r} + 2\mathbf{e}_j, j + d)$, with probability $p(\Rightarrow \vert\Diamond)$ (indicating the conditional probability of an event $\Rightarrow$ given an initial event $\Diamond$).

$(\Diamond, \rightarrow)$: The guest jumps from $(\mathbf{r} + \mathbf{e}_j, j + d)$ to $(\mathbf{r} + \mathbf{e}_j, j)$ with *no* subsequent propagation, with probability $p(\rightarrow \vert\Diamond)$.

$(\Diamond, \Leftarrow)$: The guest stays in the site $(\mathbf{r} + \mathbf{e}_j, j + d)$ and subsequently propagates to the site $(\mathbf{r} - \mathbf{e}_j, j)$, with probability $p(\Leftarrow \vert\Diamond)$.

$(\Diamond, \leftarrow)$: The guest stays in the site $(\mathbf{r} + \mathbf{e}_j, j + d)$ with *no* subsequent propagation, with probability $p(\leftarrow \vert\Diamond)$.

$(\Diamond, \Updownarrow)$: The guest jumps from $(\mathbf{r}+\mathbf{e}_j, j+d)$ to an exit site $(\mathbf{r}+\mathbf{e}_j, i)$ with $i \neq j, j+d$, and subsequently propagate to the connected exit site $(\mathbf{r}+\mathbf{e}_j+\mathbf{e}_i, i+d)$, with probability $p(\Updownarrow|\Diamond)$.

$(\Diamond, \updownarrow)$: The guest jumps from $(\mathbf{r}+\mathbf{e}_j, j+d)$ to an exit site $(\mathbf{r}+\mathbf{e}_j, i)$ with $i \neq j, j+d$, and undergoes *no* subsequent propagation, with probability $p(\updownarrow|\Diamond)$.

$(\Diamond, \bigcirc)$: The guest jumps from $(\mathbf{r}+\mathbf{e}_j, j+d)$ to an inner site of the same cell, with probability $p(\bigcirc|\Diamond)$.

Let us introduce the quantity $\chi(\varsigma|\Diamond)$, returning a value 1 if the cell-to-cell displacement at time t (represented by the symbol ς) has equal sign of the displacement at time 0, a value -1 if the sign is opposite, and 0 in all other cases:

$$\chi(\varsigma|\Diamond) = \begin{cases} 1, & \text{if } \varsigma = \Rightarrow \\ -1, & \text{if } \varsigma = \Leftarrow \\ 0, & \text{otherwise} \end{cases} \tag{A.33}$$

Therefore, since the process is Markovian one can define

$$\langle \delta\mathbf{r}(\tau)\delta\mathbf{r}(0)\rangle = \lambda^2 p(\Diamond) \sum_{\varsigma\in S} \chi(\varsigma|\Diamond) p(\varsigma|\Diamond) \tag{A.34}$$

where

$$S = \{\Rightarrow, \rightarrow, \Leftarrow, \leftarrow, \Updownarrow, \updownarrow, \bigcirc\} \tag{A.35}$$

is the set of all possible displacements, which returns

$$\langle \delta\mathbf{r}(\tau)\cdot\delta\mathbf{r}(0)\rangle = \lambda^2 p(\Diamond)\left[p(\Rightarrow|\Diamond) - p(\Leftarrow|\Diamond)\right]. \tag{A.36}$$

Time $t = z\tau$. After two time steps ($z = 2$), there are 7 kinds of starting position according to the list in Table A.1. A guest starting from an inner site or an exit site (of any cell) *different* from j or $j+d$ will have equal probability to move towards the direction $\mathbf{e}_j$ or the direction $-\mathbf{e}_j$, therefore for the starting positions $\bigcirc$, $\updownarrow$, and $\Updownarrow$ the net average displacement is null. This fact extends over all times after $z > 1$. In other words, only the moves $\Rightarrow, \rightarrow, \Leftarrow, \leftarrow$ can generate net non-null contributions to the DACF so that the general expression

$$\langle \delta\mathbf{r}(z\tau)\delta\mathbf{r}(0)\rangle = \lambda^2 p(\Diamond) \sum_{\varsigma_1\in S} \cdots \sum_{\varsigma_z\in S} \chi(\varsigma_z|\Diamond) p(\varsigma_1|\Diamond) \prod_{j=1}^{z} p(\varsigma_{j+1}|\varsigma_1), \qquad z \geq 1 \tag{A.37}$$

$p(\varsigma \mid \Rightarrow)$	
$p(\Rightarrow \mid \Rightarrow) = p(\Rightarrow \mid \Diamond)$	$p(\rightarrow \mid \Rightarrow) = p(\rightarrow \mid \Diamond)$
$p(\Leftarrow \mid \Rightarrow) = p(\Leftarrow \mid \Diamond)$	$p(\leftarrow \mid \Rightarrow) = p(\leftarrow \mid \Diamond)$
$p(\Updownarrow \mid \Rightarrow) = p(\Updownarrow \mid \Diamond)$	$p(\updownarrow \mid \Rightarrow) = p(\updownarrow \mid \Diamond)$
$p(\bigcirc \mid \Rightarrow) = p(\bigcirc \mid \Diamond)$	
$p(\varsigma \mid \rightarrow)$	
$p(\Rightarrow \mid \rightarrow) = p(\Leftarrow \mid \Diamond)$	$p(\rightarrow \mid \rightarrow) = p(\leftarrow \mid \Diamond)$
$p(\Leftarrow \mid \rightarrow) = p(\Rightarrow \mid \Diamond)$	$p(\leftarrow \mid \rightarrow) = p(\rightarrow \mid \Diamond)$
$p(\Updownarrow \mid \rightarrow) = p(\Updownarrow \mid \Diamond)$	$p(\updownarrow \mid \rightarrow) = p(\updownarrow \mid \Diamond)$
$p(\bigcirc \mid \rightarrow) = p(\bigcirc \mid \Diamond)$	
$p(\varsigma \mid \Leftarrow)$	
$p(\Rightarrow \mid \Leftarrow) = p(\Leftarrow \mid \Diamond)$	$p(\rightarrow \mid \Leftarrow) = p(\leftarrow \mid \Diamond)$
$p(\Leftarrow \mid \Leftarrow) = p(\Rightarrow \mid \Diamond)$	$p(\leftarrow \mid \Leftarrow) = p(\rightarrow \mid \Diamond)$
$p(\Updownarrow \mid \Leftarrow) = p(\Updownarrow \mid \Diamond)$	$p(\updownarrow \mid \Leftarrow) = p(\updownarrow \mid \Diamond)$
$p(\bigcirc \mid \Leftarrow) = p(\bigcirc \mid \Diamond)$	
$p(\varsigma \mid \leftarrow)$	
$p(\Rightarrow \mid \leftarrow) = p(\Rightarrow \mid \Diamond)$	$p(\rightarrow \mid \leftarrow) = p(\rightarrow \mid \Diamond)$
$p(\Leftarrow \mid \leftarrow) = p(\Leftarrow \mid \Diamond)$	$p(\leftarrow \mid \leftarrow) = p(\leftarrow \mid \Diamond)$
$p(\Updownarrow \mid \leftarrow) = p(\Updownarrow \mid \Diamond)$	$p(\updownarrow \mid \leftarrow) = p(\updownarrow \mid \Diamond)$
$p(\bigcirc \mid \leftarrow) = p(\bigcirc \mid \Diamond)$	
$p(\varsigma \mid \Updownarrow)$	
$p(\Rightarrow \mid \Updownarrow) = p(\Rightarrow \mid \updownarrow)$	$p(\rightarrow \mid \Updownarrow) = p(\rightarrow \mid \updownarrow)$
$p(\Leftarrow \mid \Updownarrow) = p(\Leftarrow \mid \updownarrow)$	$p(\leftarrow \mid \Updownarrow) = p(\leftarrow \mid \updownarrow)$
$p(\Updownarrow \mid \Updownarrow) = p(\Updownarrow \mid \updownarrow)$	$p(\updownarrow \mid \Updownarrow) = p(\updownarrow \mid \updownarrow)$
$p(\bigcirc \mid \Updownarrow) = p(\bigcirc \mid \updownarrow)$	
$p(\varsigma \mid \updownarrow)$	
$p(\Rightarrow \mid \updownarrow) = p(\Leftarrow \mid \updownarrow)$	$p(\rightarrow \mid \updownarrow) = p(\leftarrow \mid \updownarrow)$
$p(\varsigma \mid \bigcirc)$	
$p(\Rightarrow \mid \bigcirc) = p(\Leftarrow \mid \bigcirc)$	$p(\rightarrow \mid \bigcirc) = p(\leftarrow \mid \bigcirc)$

Table A.2: Possible guest jumps after two time steps for the case where during the jump randomization each guest can select any of the K sites in the cell as target sites.

reduces simply to

$$\begin{aligned}\langle \delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0)\rangle = \lambda^2 p(\Diamond)\,[p(\Rightarrow \mid \Diamond) - p(\Leftarrow \mid \Diamond)] \\ \times\,[p(\Rightarrow \mid \Diamond) - p(\Leftarrow \mid \Diamond) + p(\leftarrow \mid \Diamond) - p(\rightarrow \mid \Diamond)]^{z-1}\end{aligned} \quad \text{(A.38)}$$

which gives

$$\begin{aligned}\langle \delta\mathbf{r}(z\tau) \cdot \delta\mathbf{r}(0)\rangle = -2d\tau D_0^{\mathrm{mf}} \langle \kappa\overline{\rho_{\mathrm{ex}}}\rangle\,(1 - 2\langle \kappa\overline{\rho_{\mathrm{ex}}}\rangle)^{z-1} \\ \times \left[1 - \frac{K_{\mathrm{ex}}}{K}\langle q_{\mathrm{ex}} h_{\mathrm{ex}}\rangle - \frac{K_{\mathrm{in}}}{K}\langle q_{\mathrm{ex}} h_{\mathrm{in}}\rangle\right]^z .\end{aligned} \quad \text{(A.39)}$$

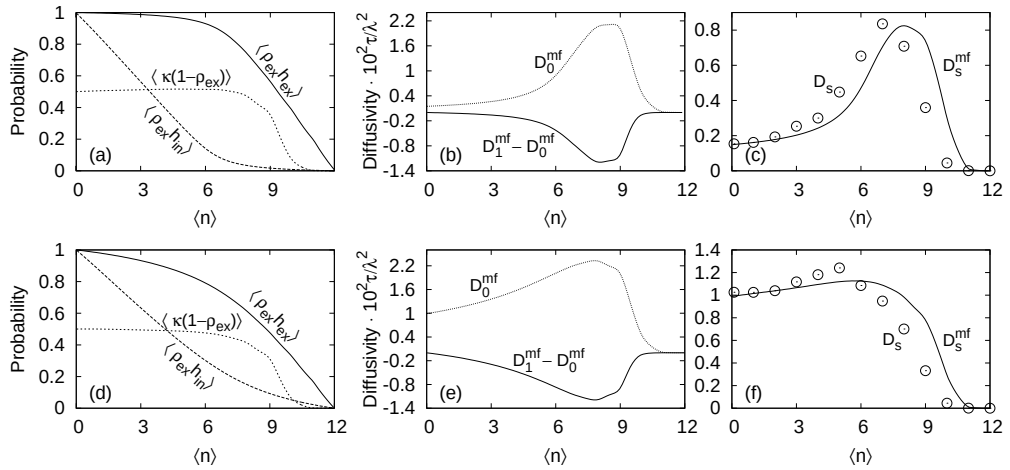

Figure A.1: Mean-field self diffusivity and macroscopic quantities involved for the same interacting system investigated in Section 6.1. Here, ex-ex jumps are allowed. Figs. (a)-(c) refer to $\Delta f^o = 10$ kJ mol^{-1}. Figs. (d)-(f) refer to $\Delta f^o = 5$ kJ mol^{-1}. In the first column, i.e. Figs. (a) and (c), the macroscopic quantities $\langle q_{\rm ex} h_{\rm ex}\rangle$, $\langle q_{\rm ex} h_{\rm in}\rangle$, and $\langle \kappa\overline{\rho_{\rm ex}}\rangle$ are plotted. In the second column, i.e. Figs. (b) and (d), the memoryless mean-field self-diffusivity $D_0^{\rm mf}$ is plotted together with the first-time correlation term, $D_1^{\rm mf} - D_0^{\rm mf}$. In the third column, i.e. Figs. (c) and (e), the resulting mean-field self-diffusivity $D^{\rm mf}$ (continuous line) is plotted together with diffusivities found in numerical ThPCA simulations.

By indicating as $D_s^{\rm mf}$ the mean-field value of the self-diffusivity of Eq. (4.23), and by using the following series

$$\sum_{z=1}^{\infty} A^z B^{z-1} = \frac{A}{1-AB} \tag{A.40}$$

to perform the summation of the correlated part, one obtains

$$D_s^{\rm mf} = D_0^{\rm mf}\left\{1 - 2\langle\kappa\overline{\rho_{\rm ex}}\rangle \frac{1 - \sum_\alpha \frac{K_\alpha}{K}\langle q_{\rm ex}\overline{\rho_\alpha}\rangle}{1-(1-2\langle\kappa\overline{\rho_{\rm ex}}\rangle)\left[1-\sum_\alpha \frac{K_\alpha}{K}\langle q_{\rm ex}\overline{\rho_\alpha}\rangle\right]}\right\}, \tag{A.41}$$

where the sums extend to $\alpha =$'ex', 'in'.

Case two: not allowed ex-ex jumps

The same procedure can be applied for the case where guests are not allowed to jump from an exit site to another exit site during the jump randomization. In such an evolution scheme, after a guest has propagated from one cell to the exit site of another cell it is forced to pass through the inner sites in order to migrate into a neighboring cell which is different from the one they have

Ex-ex jumps not allowed	
$p(\Leftarrow \vert \diamondsuit)$	$= [1 - (K_{\rm in}/K)\langle q_{\rm ex} h_{\rm in}\rangle]\, \langle \kappa\overline{\rho_{\rm ex}}\rangle$
$p(\leftarrow \vert \diamondsuit)$	$= [1 - (K_{\rm in}/K)\langle q_{\rm ex} h_{\rm in}\rangle]\, (1 - \langle \kappa\overline{\rho_{\rm ex}}\rangle)$

Table A.3: Probability values for events of jump starting from initial condition $\diamondsuit$ at time 0 for the jump randomization with not allowed ex-ex jumps.

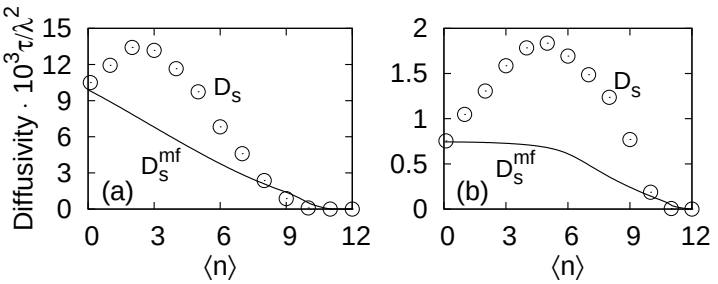

Figure A.2: Mean-field diffusivity compared with numerical values found in numerical ThPCA simulations for the same system of Figure A.1 but with (a) $\Delta f^o = -5$ kJ mol^{-1}, and (b) $\Delta f^o = 10$ kJ mol^{-1} but forbidden ex-ex jumps.

just left. The resulting DACF is

$$\langle \delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle = \lambda^2 p(\diamondsuit)(-1)^z p(\Leftarrow\vert\diamondsuit)\,[p(\Leftarrow\vert\diamondsuit) - p(\leftarrow\vert\diamondsuit)]^{z-1}\,, \qquad \text{(A.42)}$$

that by inserting the probabilities listed in Table A.3 becomes

$$\begin{aligned}\langle \delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle = &-2d\tau D_0^{\rm mf}\langle\kappa\overline{\rho_{\rm ex}}\rangle\,(1 - 2\langle\kappa\overline{\rho_{\rm ex}}\rangle)^{z-1} \\ &\times\left[1 - \frac{K_{\rm in}}{K}\langle q_{\rm ex}h_{\rm in}\rangle\right]^z, \qquad z \geq 1 \end{aligned} \qquad \text{(A.43)}$$

Therefore the mean-field equation for the diffusivity results

$$D_s^{\rm mf} = D_0^{\rm mf}\left\{1 - 2\langle\kappa\overline{\rho_{\rm ex}}\rangle\frac{1 - \frac{K_{\rm in}}{K}\langle q_{\rm ex}\overline{\rho_{\rm in}}\rangle}{1 - (1 - 2\langle\kappa\overline{\rho_{\rm ex}}\rangle)\left[1 - \frac{K_{\rm in}}{K}\langle q_{\rm ex}\overline{\rho_{\rm in}}\rangle\right]}\right\}, \qquad \text{(A.44)}$$

where the sums extend to $\alpha =$'ex', 'in'.

A.2.3 Discussion of the mean-field results.

Two of the possible mean-field diffusion profiles obtainable by introducing the simulation of dynamic events in a static thermodynamic model of adsorption of xenon in zeolite NaA (i.e. the same interacting system illustrated in Section 6.1 through numerical simulations) are plotted in Figure A.1 together

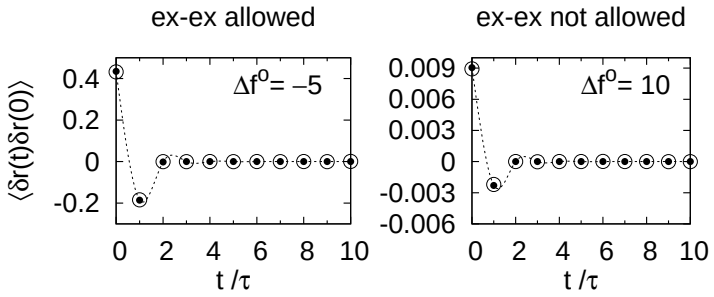

Figure A.3: Mean-field correlation function (black dots) compared with the values found by numerical simulation (white circles) for the same system of Figure A.2, in the limit of zero loading.

with the macroscopic observables involved, $\langle q_{\rm ex} h_{\rm in}\rangle$, $\langle q_{\rm ex} h_{\rm ex}\rangle$ and $\langle \kappa\overline{\rho_{\rm ex}}\rangle$. By defining the mean-field diffusivity after z time steps as

$$D_j^{\rm mf} = \sum_{z=0}^{j} \frac{\langle \delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle}{d\tau} - \frac{\langle \delta\mathbf{r}(0)\cdot\delta\mathbf{r}(0)\rangle}{2d\tau}, \tag{A.45}$$

the quantity $D_1^{\rm mf} - D_0^{\rm mf}$ represents the memory effect after one step. The more negative it is, the stronger the correlation is.

In particular, Figure A.1 refers to $\Delta f^o = f^o_{\rm ex} - f^o_{\rm in} > 0$, i.e. a condition where the inner sites are preferred w.r.t. the exit sites. Now the obtained trends will be analyzed by considering the system while increasing the loading, $\langle n\rangle$.

(i) From low to intermediate loading, the trend of $\langle q_{\rm ex} h_{\rm in}\rangle$ is rapidly decreasing, while $\langle q_{\rm ex} h_{\rm ex}\rangle$ decreases much more slowly. Together with an approximately constant value of the transmission coefficient ($\langle \kappa\overline{\rho_{\rm ex}}\rangle \approx \frac{1}{2}$) the consequence of that is that $D_0^{\rm mf}$ increases since almost all guests promoted to an exit sites during randomization have a probability of about one half to migrate to a neighboring cell, but the correlation effect also increases (because the last term in Eq. (A.39)).

(ii) At intermediate loadings many inner sites are already filled so the exit sites are being visited much more frequently. Therefore $\langle q_{\rm ex} h_{\rm ex}\rangle$ starts decreasing more rapidly whereas the decreasing trend of $\langle q_{\rm ex} h_{\rm in}\rangle$ becomes less pronounced. Anyway, the term $1 - \sum_\alpha \frac{K_\alpha}{K}\langle q_{\rm ex} h_\alpha\rangle$ of Eq. (A.39) keeps on increasing (even if less rapidly).

(iii) At intermediate-high loadings the exit sites are so frequently visited that the number of successful intercell migration per time step starts decreasing. Moreover, the kinetic barrier starts increasing. That situation causes $\langle \kappa\overline{\rho_{\rm ex}}\rangle$ to rapidly decrease, so that the correlation effect decreases together with the frequency of migrations.

Independence of events and memory effects. While if ex-ex jumps are allowed and $\Delta f^o > 0$ then a good agreement is obtained between D^{mf} and the real self-diffusivity D_s calculated through numerical ThPCA, the mean-field approach becomes instead unsuitable when $\Delta f^o < 0$ (i.e. when the exit sites have a higher binding ability) or in case of forbidden ex-ex jumps. The non-agreement shown in Figure A.2 is due to an overestimation of correlations. The reason of such an overestimation stands in the larger *memory ability* of the exit sites w.r.t. the inner sites: Since propagations occurr via the exit sites, and the exit sites' structure of connections determines the topology of the whole grid of cells, all events involving them are more correlated than the events occurring in the inner sites, which instead are structureless so that they can be considered as the less memory-preserving part of the cell. The mean-field approach required a strong approximation: independence of events. Although the closest events have been not assumed to be independent, (i.e. the probabilities $\langle q_\alpha h_\nu \rangle$, $\langle \kappa \overline{\rho_{\mathrm{ex}}} \rangle$ have *not* been separated respectively into $\langle q_\alpha \rangle \langle h_\nu \rangle$ and $\langle \kappa \rangle \langle \overline{\rho_{\mathrm{ex}}} \rangle$), there are other sources of correlations which are not included in the mean-field approach:

(1) When a tagged guest migrates from cell $\mathbf{r}_1$ to cell $\mathbf{r}_2$ during propagation, the probabilities related to every next move *do* depend on the configuration of both cells before propagation occurred. In other words, in principle the path probabilities $p(\diamondsuit, \varsigma)$, with $\varsigma \in S$ defined in Eq. (A.35) cannot be factorized as $p(\diamondsuit)p(\varsigma|\diamondsuit)$.

(2) After randomization in cell $\mathbf{r}_2$ has occurred, its K_{ex} exit sites undergo propagation in a random order, so that not only the outcome of the propagation of the tagged guests depends on the outcome of the propagations that previously occurred in other exit sites of the cell, but even the input configuration for the propagation of the tagged guest cannot be any of the possible cell configurations having one filled exit site since the number of inner sites cannot change after randomization has been completed. For example, if the σ-configuration of the cell after randomization is denoted as $\boldsymbol{\sigma}^R = \boldsymbol{\sigma}^R_{\mathrm{ex}} \cup \boldsymbol{\sigma}^R_{\mathrm{in}}$ where $\boldsymbol{\sigma}^R_{\mathrm{ex}}$ and $\boldsymbol{\sigma}^R_{\mathrm{in}}$ indicate respectively the exit and inner sites σ-configurations, then at the moment at which the tagged guest undergoes propagation the input configuration can differ only in the exit-site part while the inner configuration must remain equal to $\boldsymbol{\sigma}^R_{\mathrm{in}}$.

(3) The implicit assumption in both Eqs. (A.29) and (A.30) is that when the tagged guest is called to jump during the randomization process, the other guests in the cell are distributed according to the equilibrium distribution. This is equivalent to approximate the jump randomization

scheme to a slightly different local operation where, just before the tagged guest is invoked to perform its own jump, all the remaining guests in the cell perform a memoryless randomization. In fact, the more binding are the sites, the slower will be the cell in reaching local equilibrium. This will affect the acceptance probabilities, which will deviate from $\langle q_\alpha h_\nu \rangle$.

The macroscopic quantities $\langle q_\alpha h_\nu \rangle$ and $\langle \kappa \overline{\rho_{\text{ex}}} \rangle$ have been derived by assuming local equilibria and by forcing some connected events to be treated as they are independent. These assumptions are valid if the system evolve in such a way that the ability of each cell to preserve memory of previous configurations is weak, and the guests are easy to leave the exit sites. In turn, these requirements cannot be satisfied if $f^o_{\text{ex}} < f^o_{\text{in}}$ since guests are not easy to reach the (memoryless) inner sites, nor if the ex-ex jumps are not allowed because of the reduced guests ability to leave an exit site. To confirm these observation, in Figure A.1 agreement between mean-field approach and numerical simulation becomes poor at intermediate-high loadings, that is when inner sites are less available so that the tagged guest is forced to occupy exit sites.

All observations made above are embedded into the following one: in the ThPCA, if correlations in the single-guest motion are present, then the collective motion affects the motion of the single guests. This is emphasized by the difference discussed in Chapter 7 between collective- and self-diffusivity as found in numerical simulations. For all the investgated parametrizations, such differences reached their maximum exactly when guests were forced to visit frequently (or they were frequently trapped in) the exit sites, that is: (i) low-intermediate loadings if $\Delta f^o < 0$, and (ii) intermediate-high loadings if $\Delta f^o > 0$.

Exact theoretical predictions in the limit of infinite dilution. Despite their limitations, formally Eqs. (A.39) and (A.43) furnish the essential source of correlations in the self-motion of the guest. In particular, as shown in both Figs. A.2 and A.3, in the zero-loading limit (i.e. when the system contains only one guest) they are *quantitatively exact*:

$$\lim_{\langle n \rangle \to 0} \langle \delta \mathbf{r}(z\tau) \cdot \delta \mathbf{r}(0) \rangle = -\lambda^2 K_{\text{ex}} p^{(0)}_{\text{ex}} \langle \kappa \rangle^2 \left(1 - 2\langle \kappa \rangle\right)^{z-1} \left(1 - q_{\text{ex}}\right)^z , \qquad \text{(A.46)}$$

for the case where during the jump randomization all jumps are allowed, with $\langle \kappa \rangle = \frac{1}{2} e^{-\beta \epsilon_{\text{ki}}(1,0)}$ as the propagation probability, $q_{\text{ex}} = \gamma e^{\beta f^o_{\text{ex}}}$ as the jump randomization probability, and

$$p^{(0)}_{\text{ex}} = \frac{e^{-\beta f^o_{\text{ex}}}}{K_{\text{ex}} e^{-\beta f^o_{\text{ex}}} + K_{\text{in}} e^{-\beta f^o_{\text{in}}}} \qquad \text{(A.47)}$$

as the equilibrium probability of the guest to occupy one exit site pointing towards a specific direction of motion. Eq. (A.46) is independent of the number of exit/inner sites in the cell, while if jumps between different exit sites are forbidden then

$$\lim_{\langle n\rangle\to 0}\langle\delta\mathbf{r}(z\tau)\cdot\delta\mathbf{r}(0)\rangle = -\lambda^2 K_{\mathrm{ex}} p_{\mathrm{ex}}^{(0)}\kappa^2\,(1-2\kappa)^{z-1}\left(1-\frac{K_{\mathrm{in}}}{K}q_{\mathrm{ex}}\right)^{z}, \quad \text{(A.48)}$$

showing an explicit dependence on the number of sites constituting the cell. Therefore the accessibility of the sites plays a fundamental role in determining the entity of correlations.

A.3 Local density fluctuations

The crucial point in evaluating the LDFACF defined in Eq. (4.32) under the mean-field approximation is the calculation of the *patch correlation function*[85] $\langle n(t)n(0)\rangle$, given by

$$\langle n(z\tau)n(0)\rangle = \sum_{m=0}^{K}\sum_{n=0}^{K} nm\Omega_{n,m}^{(z)}p(m), \quad \text{(A.49)}$$

where $\mathbf{p}$ is the equilibrium probability distribution of occupancies given in Eq. (A.1), and the matrix

$$\mathbf{\Omega}^{(z)} = \{\Omega_{n,m}^{(z)}\}, \qquad n,m = 0,\ldots,K, \quad \text{(A.50)}$$

is the *occupancy transition matrix*. Each matrix element $\Omega_{n,m}^{(z)}$ gives the equilibrium probability of a cell of occupancy n to change to occupancy m after $z = t/\tau$ time steps. For $z = 0$ Eq. (A.49) corresponds to the second moment of the occupancy distribution $\mathbf{p}$, therefore $\mathbf{\Omega}^{(0)} = \mathbf{I}$ (identity matrix). Since the evolution rule is markovian, under the mean-field approximation it results

$$\mathbf{\Omega}^{(z)} = \mathbf{\Omega}^{z}, \quad \text{(A.51)}$$

where $\mathbf{\Omega}$ is the occupancy transition matrix after one time step. Its elements are given by

$$\Omega_{n,n+\Delta n} = \sum_{n_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}} \omega(\Delta n|n_{\mathrm{ex}},n)P(n_{\mathrm{ex}}|n), \quad \text{(A.52)}$$

where $P(n_{\mathrm{ex}}|n)$ has been given in Eq. (2.29), and $\omega(\Delta n|n_{\mathrm{ex}},n)$ is the probability of a n-occupied cell with n_{ex} occupied exit sites to change its occupancy from n to $m = n+\Delta n$. The matrix $\mathbf{\Omega}$ can be computed in two different ways depending on whether the guests are interacting or not. This will be the subject of the next two subsections.

Non-interacting case

In the non-interacting case the interaction free-energy term $\phi_\alpha(n)$ in Eq. (2.21) is zero, so that the evolution can be performed in a fully synchronous way[45,54,55] and a mean-field value of the the transition probability $\omega(\Delta n|n_{\rm ex}, n)$ can be computed exactly.

Let us focus on a single n-occupied cell (labeled $\mathbf{r}$) and suppose it to have $n_{\rm ex}$ guests located in its exit sites. Each exit site $(\mathbf{r}, j)$ is connected to one exit site of the adjacent cell, denoted as $(\mathbf{r}^j, j+d)$. These two adjacent exit sites constitute one *link* between $\mathbf{r}$ and $\mathbf{r}^j$. The link occupancy, given by the quantity $\delta(\mathbf{r}, \mathbf{r}^j)$ defined in Eq. (3.25), can have one of the following three possible states:

$\delta(\mathbf{r}, \mathbf{r}^j) = 0$: empty link (no transfer can occur);

$\delta(\mathbf{r}, \mathbf{r}^j) = 1$: open link (a transfer occurs from the occupied to the unoccupied site with probability

$$\widetilde{\kappa} = \frac{\kappa^o}{2} = \frac{1}{2} e^{-\beta \epsilon_{\rm ki}}, \tag{A.53}$$

where $\epsilon_{\rm ki}$ is a *homogeneous* kinetic barrier);

$\chi(\mathbf{r}, \mathbf{r}^j) = 2$: blocked link (no transfer can occur).

Let us call $z_{\rm ex}$ the sum of occupancies of the adjacent exit sites, i.e.

$$z_{\rm ex} = \sum_{j=1}^{K_{\rm ex}} \chi(\mathbf{r}, \mathbf{r}^j) - n_{\rm ex}.$$

If the values of the observables $n_{\rm ex}, z_{\rm ex}$ are known but no information is available about the value of each single link occupancy, then calculating the probability of having x blocked links will be very useful. Thus, if there are x blocked links, then $n_{\rm ex} - x$ guests can try to exit from the cell $\mathbf{r}$ and $z_{\rm ex} - x$ guests can try to enter in $\mathbf{r}$. It will be useful to define with $q(x|n_{\rm ex}, z_{\rm ex})$ the conditional probability that x links are blocked, given that $n_{\rm ex}$ guests occupy the exit sites of the cell and $z_{\rm ex}$ guests occupy the adjacent exit sites of the neighboring cells. This is given by:

$$q(x|n_{\rm ex}, z_{\rm ex}) = \frac{\binom{K_{\rm ex}}{x}\binom{K_{\rm ex}-x}{n_{\rm ex}+z_{\rm ex}-2x}\binom{n_{\rm ex}+z_{\rm ex}-2x}{\min(n_{\rm ex},z_{\rm ex})-x}}{\binom{K_{\rm ex}}{n_{\rm ex}}\binom{K_{\rm ex}}{z_{\rm ex}}}. \tag{A.54}$$

The probability of the adjacent exit site $(\mathbf{r}^j, j+d)$ of each (j-th) neighboring cell to be occupied is equal to the probability of one exit site to be occupied,

which is given by

$$\rho_{\mathrm{ex}} = \frac{\langle n_{\mathrm{ex}} \rangle}{K_{\mathrm{ex}}} = \frac{1}{K_{\mathrm{ex}}} \sum_{n=1}^{K} p(n) \sum_{i=1}^{K_{\mathrm{ex}}} n_{\mathrm{ex}} P(n_{\mathrm{ex}}|n) \tag{A.55}$$

and the probability of having z_{ex} guests in the first-neighbors' adjacent exit sites is

$$A(z_{\mathrm{ex}}) = \binom{K_{\mathrm{ex}}}{z_{\mathrm{ex}}} \rho_{\mathrm{ex}}^{z_{\mathrm{ex}}} \left(1 - \rho_{\mathrm{ex}}\right)^{K_{\mathrm{ex}} - z_{\mathrm{ex}}} . \tag{A.56}$$

Now the exiting guests shall be considered. The conditional probability of j_{ex} guests to leave the cell given that $n_{\mathrm{ex}} - x$ is the number of guests in open links attempting to leave the cell $\mathbf{r}$, is

$$B(j_{\mathrm{ex}}|n_{\mathrm{ex}} - x) = \binom{n_{\mathrm{ex}} - x}{j_{\mathrm{ex}}} \widetilde{\kappa}^{j} (1 - \widetilde{\kappa})^{n_{\mathrm{ex}} - j_{\mathrm{ex}} - x}, \tag{A.57}$$

By analogy with Eq. (A.57), h_{ex} guests will effectively penetrate into $\mathbf{r}$ with probability

$$C(h_{\mathrm{ex}}|z_{\mathrm{ex}} - x) = \binom{z_{\mathrm{ex}} - x}{h_{\mathrm{ex}}} \widetilde{\kappa}^{j} (1 - \widetilde{\kappa})^{z_{\mathrm{ex}} - h_{\mathrm{ex}} - x}, \tag{A.58}$$

therefore the conditional probability of the cell $\mathbf{r}$ to change its occupancy of an amount Δn given that its exit sites are occupied by n_{ex} guests, the adjacent exit sites of the neighboring cells are occupied by n_{ex} guests and x links are blocked, is given by

$$G(\Delta n|n_{\mathrm{ex}}, z_{\mathrm{ex}}, x) = \sum_{j_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}} B(j_{\mathrm{ex}}|n_{\mathrm{ex}} - x) C(j_{\mathrm{ex}} + \Delta n|z_{\mathrm{ex}} - x) \tag{A.59}$$

Now the quantity $\omega(\Delta n|n_{\mathrm{ex}}, n)$ can be computed by combining all obtained probabilities:

$$\omega(\Delta n|n_{\mathrm{ex}}, n) = \sum_{z_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}} \sum_{x=0}^{K_{\mathrm{ex}}} A(z_{\mathrm{ex}}) q(x|n_{\mathrm{ex}}, z_{\mathrm{ex}}) G(\Delta n|n_{\mathrm{ex}}, z_{\mathrm{ex}}, x), \tag{A.60}$$

which can be inserted into Eq. (A.52) to get the occupancy transition matrix.

Interacting case

If the guests are interacting then the partitioning scheme described in Sections 3.4.2 and E.3 must be used for propagation. As a consequence it would be rather tedious to compute $\omega(\Delta n|n_{\mathrm{ex}}, n)$ since every possible variation Δn (ranging from $-K_{\mathrm{ex}}$ to K_{ex}) can be caused by any of the $(2d)!$ paths associated with the different orders in which the exit sites of the cell can be updated. To simplify the problem, one shall instead compute the probability of the events *associated to the propagation along one direction*, of losing/gaining one guest or keeping the cell occupancy unchanged given respectively by the formulas:

$$\widetilde{\Omega}_{n,n-1} = \sum_{m=0}^{K-1} W(n,m)p(m) \tag{A.61}$$

$$\widetilde{\Omega}_{n,n+1} = \sum_{m=1}^{K} W(m,n)p(m) \tag{A.62}$$

$$\widetilde{\Omega}_{n,n} = 1 - \widetilde{\Omega}_{n,n-1} - \widetilde{\Omega}_{n,n+1} \tag{A.63}$$

where $W(n,m)$ is the intercell transfer factor given by Eq. (A.5). The one-step occupancy transition matrix can then be recovered under the following assumption:

$$\mathbf{\Omega} \approx \widetilde{\mathbf{\Omega}}^{2d}, \tag{A.64}$$

and the obtained approximate value of $\mathbf{\Omega}$ can be inserted into Eq. (A.51) to get the occupancy transition matrix. This is equivalent to approximate the propagation process described in Appendix E.3 to a sequence of operations in which one randomization is performed after propagation along each direction. Such an approximation becomes strictly valid only for low values of the kinetic prefactor $e^{-\beta\epsilon_{\mathrm{ki}}}$ (see Eq. (3.21)), when the probability of a cell to release/accept a high number of guests during the same propagation step is low. Indeed by testing the approximation with the non-interacting system investigated in Section A.3, it emerged that for $e^{-\beta\epsilon_{\mathrm{ki}}} = 0.1$ the $C(t)$ curves obtained with and without the approximation (A.64) coincide. The fitting parameters b_1, b_2, τ_1, and τ_2 are therefore found to coincide exactly with the empty symbols in Figure 7.9. Since their deviation from the corresponding values obtained for $e^{-\beta\epsilon_{\mathrm{ki}}} = 1.0$ is not much, the approximation (A.64) will be kept as valid to get qualitative information about the mean-field relaxation process in interacting systems.

Appendix B

Theory of the non-interacting case

B.1 Computation of the partition functions

First of all the possible configurations of the N guests in the KM sites of the lattice $\mathscr{L}$ will be considered. For a given configuration, let N_{ex} and N_{in} be the total number of occupied exit and inner sites, respectively. Obviously their sum is always the total number of guests $N = N_{\text{ex}} + N_{\text{in}}$, which is constant in time.

Since there is no interaction between the guests, the energy of the entire system depends only on the values of N_{ex} and N_{in}. As the observables n_{ex} and n specify a particular energy level of the single cell, in the same way the global quantities N_{ex} and N define a particular energy level of the system with degeneracy $\Omega(N_{\text{ex}}, N_{\text{in}}) = \binom{MK_{\text{ex}}}{N_{\text{ex}}}\binom{MK_{\text{in}}}{N_{\text{in}}}$, where M is the total number of cells. The energy of the level $(N_{\text{ex}}, N_{\text{in}})$ is

$$E^{\text{sys}}(N_{\text{ex}}, N_{\text{in}}) = N_{\text{ex}} f^o_{\text{ex}} + N_{\text{in}} f^o_{\text{in}} \tag{B.1}$$

Therefore the partition function of the system can be rewritten as

$$Q^{\text{sys}} = \sum_{N_{\text{ex}}=0}^{MK_{\text{ex}}} \Omega(N_{\text{ex}}, N_{\text{in}})\, e^{-\beta E^{\text{sys}}(N_{\text{ex}}, N_{\text{in}})}. \tag{B.2}$$

The probability of the level $(N_{\text{ex}}, N_{\text{in}})$ is

$$\frac{\Omega(N_{\text{ex}}, N_{\text{in}})\, e^{-\beta E^{\text{sys}}(N_{\text{ex}}, N_{\text{in}})}}{Q^{\text{sys}}}, \tag{B.3}$$

and the occupancy probability distribution of the single cell, indicated as

$$\mathbf{p} = \{p(0), \ldots, p(K)\},$$

reads

$$p(n) = \frac{1}{Q^{\mathrm{sys}}} \sum_{N_{\mathrm{ex}}=0}^{MK_{\mathrm{ex}}} \sum_{n_{\mathrm{ex}}=0}^{K_{\mathrm{ex}}} \binom{K_{\mathrm{ex}}}{n_{\mathrm{ex}}} \binom{(M-1)K_{\mathrm{ex}}}{N_{\mathrm{ex}}-n_{\mathrm{ex}}} \binom{K_{\mathrm{in}}}{n-n_{\mathrm{ex}}} \times \binom{(M-1)(K_{\mathrm{in}})}{N-N_{\mathrm{ex}}-n+n_{\mathrm{ex}}} e^{-\beta E^{\mathrm{sys}}(N_{\mathrm{ex}},N_{\mathrm{in}})}. \tag{B.4}$$

Computing probability distributions of states through Eqs. (B.2) and (B.4) is feasible only for a system containing a very small number of cells. Instead, it is much easier to compute the probability distribution $\mathbf{p}$ in terms of the *cellular partition function* in the grand canonical ensemble, $\Xi(\mu)$. Let us denote[70] $\mathbf{a} = \{a_0, \ldots, a_K\}$ a particular distribution *of occupancies* of the system of M cells, where a_n is the number of cells having occupancy n.

- Given the particular distribution of occupancies $\mathbf{a}$, the sum over states of an ordered sequence of the M cells preserving the occupancy of each cell is $\prod_{n=0}^{K} \left[Q^{\mathrm{cell}}(n)\right]^{a_n}$.
- Differently ordered sequences of the M cells according to the distribution $\mathbf{a}$ are equally probable. The number of ways that any particular distribution of the a_j's can be realized is $M!/\prod_{n=0}^{K} a_n!$.

Therefore the sum over states for a system described by the particular distribution of occupancies $\mathbf{a}$ is:

$$W(\mathbf{a}) = M! \prod_{n=0}^{K} \frac{\left[Q^{\mathrm{cell}}(n)\right]^{a_n}}{a_n!}. \tag{B.5}$$

Indicating as W^{TOT} the sum over states over all the possible configurations, for a very large system ($M \to \infty$) one obtains

$$\lim_{M\to\infty} W^{\mathrm{TOT}} = \lim_{M\to\infty} \sum_{\mathbf{a}} W(\mathbf{a}) \approx W(\mathbf{a}^{\mathrm{eq}}), \tag{B.6}$$

where $\mathbf{a}^{\mathrm{eq}} = M\mathbf{p}$ is the equilibrium occupancy distribution. Using the Stirling's approximation the Eq. (B.5) becomes

$$\ln W(\mathbf{a}^{\mathrm{eq}}) = M \ln M - \sum_{n=0}^{K} a_n^{\mathrm{eq}} \ln a_n^{\mathrm{eq}} + \sum_{n=0}^{K} a_n^{\mathrm{eq}} \ln Q^{\mathrm{cell}}(n). \tag{B.7}$$

This function can be maximized using the method of Lagrange multipliers with the constraints $\sum_{n=0}^{K} a_n^{\mathrm{eq}} = M$ and $\sum_{n=0}^{K} n a_n^{\mathrm{eq}} = N$ thus obtaining Eq. (2.29).

The cellular partition function is written as $\Xi(\mu) = \sum_{n=0}^{K} Q^{\text{cell}}(n)\, e^{\beta\mu n}$. This partition function is analogous to the grand-canonical one, so the following relations are satisfied:

$$\langle n \rangle = \frac{1}{\beta} \frac{\partial \ln \Xi}{\partial \mu}, \tag{B.8}$$

$$\sigma^2(\langle n \rangle) = \frac{1}{\beta} \frac{\partial \langle n \rangle}{\partial \mu}, \tag{B.9}$$

where $\sigma^2(\langle n \rangle)$ is the variance of the occupancy distribution $\mathbf{p}$ at the loading $\langle n \rangle$.

Since the energy parameters f^o_{ex} and f^o_{in} are fixed in space and time, the expression for Ξ can be manipulated[70] to obtain the partition function of the Fermi-Dirac statistics:

$$\Xi(\mu) = \prod_{j=1}^{K} \left(1 + e^{\beta(\mu - f^o_j)}\right), \tag{B.10}$$

which reduces to Eq. (A.11). Now, using the relation (B.8) and introducing the absolute activity $\lambda_a = e^{\beta\mu}$, one obtains the dual-site Langmuir isotherm reported in Eq. (A.12).

B.1.1 Limiting Distributions for $T \to \infty$

In this case, $\lim_{T\to\infty} f^{\text{eq}} = f^{\text{hyp}}$, i.e. the guests are distributed in the cells according to the hypergeometric distribution:

$$p^{\text{hyp}}(n) = \binom{K}{n} \binom{K(M-1)}{N-n} \Big/ \binom{MK}{N}. \tag{B.11}$$

For systems with a very large number of cells ($M \to \infty$), this expression reduces to:

$$p^{\text{hyp}}(n) = \binom{K}{n} (\rho)^n (1-\rho)^{K-n}, \tag{B.12}$$

with $\rho = \langle n \rangle / K$. This is because in the limit of $T \to \infty$, Eq. (4.11) has the solution $\rho_{\text{ex}} = \rho_{\text{in}} = \rho$. The partition function of a closed cell reduces to $Q^{\text{cell}}(n) = \binom{K}{n}$, and Eq. (A.14) reduces to

$$p(n) = \binom{K}{n} \frac{{\lambda_a}^n}{(1+\lambda_a)^K}. \tag{B.13}$$

Using the fact that $\rho = \lambda_a/(1+\lambda_a)$, it is easy to recognize in Eq. (B.13) the hypergeometric distribution in the limit of $M \to \infty$ reported in Eq. (B.12).

Using Eq. (B.9) the variance results

$$\sigma^2_{\text{hyp}} = \frac{1}{\beta}\frac{\partial\langle n\rangle}{\partial\lambda_a}\frac{\partial\lambda_a}{\partial\mu} = -\frac{\langle n\rangle^2}{K} + \langle n\rangle. \tag{B.14}$$

B.1.2 Limiting Distributions for $T \to 0$

In such a case the occupancy distribution splits up in two distinct hypergeometrics (referred to two distinct regions of loading) separated by a delta function at the critical loading:

$$\lim_{T\to 0} p(n) = \begin{cases} p^{\text{hyp}}_{\text{in}}(n), \text{ for } 0 \le \langle n\rangle < K_{\text{in}} \\ \\ \delta\left(n - \langle n\rangle\right), \text{ for } \langle n\rangle = K_{\text{in}} \\ \\ p^{\text{hyp}}_{\text{ex}}(n), \text{ for } K_{\text{in}} < \langle n\rangle \le K \end{cases} \tag{B.15}$$

where

$$p^{\text{hyp}}_{\text{in}}(n) = \binom{K_{\text{in}}}{n}(\rho_{\text{in}})^n(1-\rho_{\text{in}})^{K_{\text{in}}-n} \qquad \text{with } \rho_{\text{in}} = \frac{\langle n\rangle}{K_{\text{in}}}, \tag{B.16}$$

and

$$p^{\text{hyp}}_{\text{ex}}(n) = \binom{K_{\text{ex}}}{n-K_{\text{in}}}(\rho_{\text{ex}})^{n-K_{\text{in}}}(1-\rho_{\text{ex}})^{K-n} \qquad \text{with } \rho_{\text{ex}} = 1 - \frac{K-\langle n\rangle}{K_{\text{ex}}}. \tag{B.17}$$

This is because in the limit of $T \to 0$, Eq. (4.11) has the solutions (i) $\rho_{\text{ex}} = 0$ and (ii) $\rho_{\text{in}} = 1$. Using the relation

$$\langle n\rangle = K_{\text{ex}}\rho_{\text{ex}} + K_{\text{in}}\rho_{\text{in}}, \tag{B.18}$$

Since $0 \le \rho_{\text{ex}}, \rho_{\text{in}} \le 1$:

i) If $\rho_{\text{ex}} = 0$, then $\langle n\rangle = K_{\text{in}}\rho_{\text{in}}$, which is valid for $0 \le \langle n\rangle < K_{\text{in}}$.

   Since n_{ex} cannot be negative, then in order to satisfy the relation $\rho_{\text{ex}} = 0$ it must be $n_{\text{ex}} = 0$ for each occupancy $n = n_{\text{ex}} + n_{\text{in}}$. Therefore $n \in [0, K_{\text{in}}] \in \mathbb{N}$ and $Q^{\text{cell}}(n)$ reduces to

$$Q^{\text{cell}}_{\text{in}}(n_{\text{in}}) = \binom{K_{\text{in}}}{n_{\text{in}}} e^{-\beta f^o_{\text{in}} n_{\text{in}}},$$

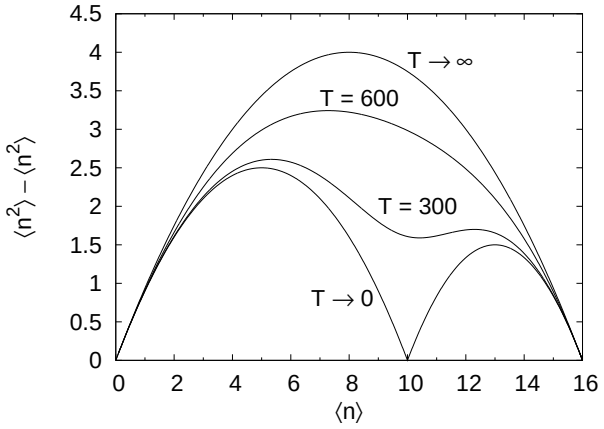

Figure B.1: The variance of the occupancy distributions at the two limiting temperatures $T \to \infty$ and $T \to 0$, and at the intermediate temperatures $T = 300, 600$ K.

with $n_{\text{in}} = n$. The sum over states in a cell can be represented through the inner partition function

$$\Xi_{\text{in}}(\mu) = (1 + e^{-\beta f^o_{\text{in}}} \lambda_{\text{in}})^{K_{\text{in}}}$$

(with $\lambda_{\text{in}} = e^{\beta \mu_{\text{in}}}$, with μ_{in} given by Eq. (4.9)). Using Eq. (A.14) and the relation $n = n_{\text{in}}$ one obtains

$$p_{\text{in}}(n) = \binom{K_{\text{in}}}{n} \frac{\left(e^{-\beta f^o_{\text{in}}} \lambda_{\text{in}}\right)^n}{\left(1 + e^{-\beta f^o_{\text{in}}} \lambda_{\text{in}}\right)^{K_{\text{in}}}}. \tag{B.19}$$

Using the relation $\rho_{\text{in}} = e^{-\beta f^o_{\text{in}}} \lambda_{\text{in}} / (1 + e^{\beta f^o_{\text{in}}} \lambda_{\text{in}})$ one obtains Eq. (B.16). This is a hypergeometric distribution of guests in K_{in} sites inside of a cell in the limit of $M \to \infty$, when the remaining K_{in} sites are all empty. Its variance is

$$\sigma^2_{\text{in}} = \frac{1}{\beta} \frac{\partial \langle n \rangle}{\partial \lambda_{\text{in}}} \frac{\partial \lambda_{\text{in}}}{\partial \mu_{\text{in}}} = -\frac{\langle n \rangle^2}{K_{\text{in}}} + \langle n \rangle. \tag{B.20}$$

ii) If $\rho_{\text{in}} = 1$, since then $\langle n \rangle = K_{\text{ex}} \rho_{\text{ex}} + (K_{\text{in}})$, which is valid for $K_{\text{in}} \leq \langle n \rangle < K$. Since n_{in} cannot be greater than K_{in}, then in order to satisfy the relation $\rho_{\text{in}} = 1$ it must be $n_{\text{in}} = K_{\text{in}}$ for each occupancy n. Therefore $n \in [K_{\text{in}}, K] \in \mathbb{N}$. Since n_{in} is fixed, for each value of n also the observable n_{ex} turns out to be fixed. Therefore $Q^{\text{cell}}(n)$ reduces to

$$Q^{\text{cell}}_{\text{ex}}(n_{\text{ex}}) = e^{-\beta f^o_{\text{in}}(K_{\text{in}})} \binom{K_{\text{ex}}}{n_{\text{ex}}} e^{-\beta f^o_{\text{ex}} n_{\text{ex}}},$$

with $n_{\mathrm{ex}} = n - K_{\mathrm{in}}$. Introducing $\lambda_{\mathrm{ex}} = e^{\beta\mu_{\mathrm{ex}}}$ exit partition function can be calculated as:

$$\Xi_{\mathrm{ex}} = e^{-\beta f^o_{\mathrm{in}}(K_{\mathrm{in}})}\left(1 + e^{-\beta f^o_{\mathrm{ex}}}\lambda_{\mathrm{ex}}\right)^{K_{\mathrm{ex}}}$$

Using Eq. (A.14) and the relation $n = K_{\mathrm{in}} + n_{\mathrm{ex}}$ one obtains

$$p_{\mathrm{ex}}(n) = \binom{K_{\mathrm{ex}}}{n - K_{\mathrm{in}}} \frac{\left(e^{-\beta f^o_{\mathrm{ex}}}\lambda_{\mathrm{ex}}\right)^{n-K_{\mathrm{in}}}}{\left(1 + e^{-\beta f^o_{\mathrm{ex}}}\lambda_{\mathrm{ex}}\right)^{K_{\mathrm{ex}}}}, \tag{B.21}$$

which, using $\rho_{\mathrm{ex}} = e^{-\beta f^o_{\mathrm{ex}}}\lambda_{\mathrm{ex}}/(1+e^{-\beta f^o_{\mathrm{ex}}}\lambda_{\mathrm{ex}})$, can be reduced to Eq. (B.17). This is a hypergeometric distribution of guests in K_{ex} sites inside of a cell in the limit of $M \to \infty$, when the other K_{ex} sites are all filled. Its variance is

$$\sigma^2_{\mathrm{ex}} = \frac{1}{\beta}\frac{\partial\langle n\rangle}{\partial\lambda_{\mathrm{ex}}}\frac{\partial\lambda_{\mathrm{ex}}}{\partial\mu_{\mathrm{ex}}} = -\frac{1}{K_{\mathrm{ex}}}\left[\langle n\rangle^2 - (2K - K_{\mathrm{ex}})\langle n\rangle + KK_{\mathrm{in}}\right]. \tag{B.22}$$

iii) At the transition point the relations $\rho_{\mathrm{ex}} = 0$ and $\rho_{\mathrm{in}} = 1$ hold simultaneously, therefore $\mathbf{p} = \mathbf{p}^{\mathrm{hyp}}_{\mathrm{in}} \cap \mathbf{p}^{\mathrm{hyp}}_{\mathrm{ex}} = \delta(n - \langle n\rangle)$ with $\langle n\rangle = K_{\mathrm{in}}$.

In Figure B.1 the variance of $\mathbf{p}$ at the two limiting temperatures (from Eqs. (B.14), (B.20), and (B.22)) and at intermediate temperature (obtained differentiating Eq. (A.12) w.r.t. μ at $T = 300$ and 600 K) are shown. As can be seen, the properties of the distributions at $0 < T < \infty$ are intermediate between the properties of the two limiting distributions, and the temperature will determine which one of the two limiting distributions the $\mathbf{p}$ will be more similar to.

p

$\sqrt{3}\ell$	2ℓ	$2\sqrt{2}\ell$	$\sqrt{11}\ell$	$2\sqrt{3}\ell$	4ℓ
$\ell_{1,7}$, $\ell_{7,15}$	$\ell_{7,8}$, $\ell_{1,15}$	$\ell_{1,2}$, $\ell_{7,10}$	$\ell_{1,11}$	$\ell_{7,13}$	$\ell_{1,4}$

Table C.1: The site-to-site distances in the structured cell of Figure 2.1. ℓ is an arbitrary factor. Not reported distances can be deduced from observation of the figure.

Appendix C

From a structured to a less-structured cell

In the present Appendix a systematic procedure will be shown which is aimed to transfer the most relevant thermodynamic properties of a detailed cell of structured interacting adsorption sites into a less detailed cell characterized by a very essential structure and two occupancy-dependent effective site energies. Such a less-detailed cell is constructed *ad-hoc* to find direct application in the ThPCA paradigm. In the following procedure, the cell partition function (which determines the shape of both the occupancy probability distribution and the adsorption isotherm[54]), and the average number of filled exit sites (which plays a major role in determining the loading dependence of the diffusivity[55]) will be reproduced *exactly* in the less detailed cell, at expense of the accuracy in the average cell energy.

In what follows such structured local interactions will be reduced to *two* effective occupancy-dependent site potentials by means of a coarse-graining procedure aimed to preserve the essential thermodynamic properties of the structured cell, which for the purposes of the present model are the free energy, the average energy, and the average fraction of guests near the windows of an n-occupied cell.

Thermodynamics of the structured cell. Numerical examples will refer to the structured cell of Figure 2.1.

The *partial configurations* of exit and inner sites shall be denoted as

$$\begin{aligned} \boldsymbol{\eta}_{\text{ex}} &= \{\eta_1, \ldots, \eta_{K_{\text{ex}}}\}, \\ \boldsymbol{\eta}_{\text{in}} &= \{\eta_{K_{\text{ex}}+1}, \ldots, \eta_K\}, \end{aligned} \tag{C.1}$$

and their relative *partial occupancies* as

$$n(\boldsymbol{\eta}) = \boldsymbol{\eta} \cdot \boldsymbol{\eta}, \qquad \alpha = \text{ex, in}, \tag{C.2}$$

where η_j is the *occupancy* of the j-th site of the cell and has value 0 if empty or 1 if occupied by one guest. The *cell configuration* and *cell occupancy* will be given respectively by $\boldsymbol{\eta} = \boldsymbol{\eta}_{\text{ex}} \cup \boldsymbol{\eta}_{\text{in}}$ and $n = n_{\text{ex}} + n_{\text{in}}$.

In the structured cell, the quantity ℓ_{ij} shall denote the distance between two sites of the same cell. The pair distances of the cell in Figure 2.1 are listed in Table C.1. If $\{\ell_{hk}\}$ indicates the set of all pair distances then $\ell^{\text{sd}} = \min\{\ell_{hk}\}$ is the 'shortest distance' ('sd') between two sites in the cell. The interaction potential between two sites at the distance ℓ^{sd} shall be indicated as ε^{sd}. The lattice interaction potential of the structured cell will be expressed in the following form:

$$\varepsilon_{ij} = [1 - \delta_{\text{kr}}(i,j)] \left(\frac{\ell^{\text{sd}}}{\ell_{ij}}\right)^a \theta\left(\frac{\ell^{\text{tr}} - \ell_{ij}}{\ell_{\text{ex}}/2}\right) \varepsilon^{\text{sd}}, \tag{C.3}$$

where

- $\delta_{\text{kr}}(i,j)$ is a Krœneker delta,
- the second factor (with $a > 0$) lowers the interaction energy depending on the distance between the two interacting sites,
- $\theta(x)$ is a Heaviside function (returning 1 for $x \geq 0$ and 0 otherwise) which 'turns off' all the interactions above a truncation distance ℓ^{tr},
- ℓ_{ex} is the (scalar) distance between the center of the cell and an exit site.

Attractive and repulsive components can be taken into account at the same time by introducing a dependence of ε^{sd} on n.

From now on in this Appendix, the superscript $[j]$ (where $j = 0, 1, 2$) will denote the *level of reduction*.

- The cell of level [0] (i.e. the structured cell) is characterized by a well-defined structure of the inner sites and an energy function

$$E^{[0]}(\boldsymbol{\eta}) = \sum_{i=1}^{K} \eta_i \varepsilon_i + \sum_{i=1}^{K-1} \sum_{j=i+1}^{K} \eta_i \eta_j \varepsilon_{ij}. \tag{C.4}$$

  which depends on the particular configuration $\boldsymbol{\eta}$ of the guest particles in the sites. The term ε_i in Eq. (C.4) indicates the energy of the i-th site.

- At constant temperature the energy function $E^{[0]}(\boldsymbol{\eta})$ can be reduced to the simpler function

$$E^{[1]}(n_{\mathrm{ex}}, n_{\mathrm{in}}; T) = \sum_{\alpha=\mathrm{ex,in}} n_\alpha \varepsilon_\alpha^{[1]}(n_{\mathrm{ex}}, n_{\mathrm{in}}; T), \tag{C.5}$$

  where the energy parameters ε_α depend on both partial occupancies n_{ex} and n_{in}.

- The energy function $E^{[1]}(\boldsymbol{\eta})$ can be further simplified to the energy function of level [2]:

$$E^{[2]}(n_{\mathrm{ex}}, n_{\mathrm{in}}; T) = \sum_{\alpha=\mathrm{ex,in}} n_\alpha \varepsilon_\alpha^{[2]}(n; T), \tag{C.6}$$

  where both energy parameters depend on the total occupancy of the cell, given by $n = n_{\mathrm{ex}} + n_{\mathrm{in}}$.

Since the reduced energy parameters are derived from averages over the accessible configurations of the structured cell (and each configuration has its own temperature-dependent weight in the cell partition function), all coarse-grained energies functions are necessarily temperature-dependent.

The derivation of the reduced energy parameters can be derived as follows. First of all Eq. (C.4) should be rewritten as

$$E^{[0]}(\boldsymbol{\eta}) = E^{[0]}_{\mathrm{ex-ex}} + E^{[0]}_{\mathrm{in-in}} + E^{[0]}_{\mathrm{ex-in}} \tag{C.7}$$

Guest-host and guest-guest interactions are contained respectively in the first and second term. It will be useful to express the interaction term as a sum of three contributions: ex-ex interactions (among exit sites only)

$$E^{[0]}_{\mathrm{ex-ex}}(\boldsymbol{\eta}_{\mathrm{ex}}) = \sum_{i=1}^{K_{\mathrm{ex}}} \eta_i \varepsilon_i + \sum_{i=1}^{K_{\mathrm{ex}}-1} \sum_{j=i+1}^{K_{\mathrm{ex}}} \eta_i \eta_j \varepsilon_{ij}, \tag{C.8}$$

in-in interactions (among inner sites only)

$$E^{[0]}_{\mathrm{in-in}}(\boldsymbol{\eta}_{\mathrm{in}}) = \sum_{i=K_{\mathrm{ex}}+1}^{K} \eta_i \varepsilon_i + \sum_{i=K_{\mathrm{ex}}+1}^{K-1} \sum_{j=i+1}^{K} \eta_i \eta_j \varepsilon_{ij}, \tag{C.9}$$

and ex-in interactions

$$E^{[0]}_{\mathrm{ex-in}}(\boldsymbol{\eta}) = \sum_{i=1}^{K_{\mathrm{ex}}} \sum_{j=K_{\mathrm{ex}}+1}^{K} \eta_i \eta_j \varepsilon_{ij}. \tag{C.10}$$

In order the less-structured cell to be set up as the coarse-grained version of the structured cell, the two kinds of cells must be similar in several basic properties. In particular they should satisfy the following *similarity conditions*:

(i) Similar (or equal) values of the *cell partition function*. That will produce similar (or equal) thermodynamic properties such as the adsorption isotherm.

(ii) Similar values of the *average number of filled exit sites in the cell*. That will ensure that at each time step the number of guests in the lattice attempting to migrate from one site to the other will be approximately the same in both the structured and the less-structured cell.

(iii) Similar values of the *average energy of the cell*, producing similarity of the entropy of the lattice.

Since the temperature will be kept constant, from now on in this Appendix the T-dependence will be implicit.

The cellular partition function can be expressed as

$$Q^{[j]}(n) = \sum_{n_{\mathrm{ex}}, n_{\mathrm{in}}}^{(n)} Q^{[j]}_c(n_{\mathrm{ex}}, n_{\mathrm{in}}), \tag{C.11}$$

where the structured cell's partition function for given values of partial occupancies is

$$Q^{[0]}_c(n_{\mathrm{ex}}, n_{\mathrm{in}}) = \sum_{\boldsymbol{\eta}_{\mathrm{ex}}}^{(n_{\mathrm{ex}})} \sum_{\boldsymbol{\eta}_{\mathrm{in}}}^{(n_{\mathrm{in}})} e^{-\beta E^{[0]}(\boldsymbol{\eta})}, \tag{C.12}$$

In Eq. (C.11) the compact notation introduced in Eq. (2.2) has been used.

For a less-structured cell:

$$Q^{[j]}_c(n_{\mathrm{ex}}, n_{\mathrm{in}}) = \binom{K_{\mathrm{ex}}}{n_{\mathrm{ex}}} \binom{K_{\mathrm{in}}}{n_{\mathrm{in}}} e^{-\beta E^{[j]}(n_{\mathrm{ex}}, n_{\mathrm{in}})}, \qquad j = 1, 2. \tag{C.13}$$

For all levels of reduction, the probability of the cell to have the partial occupancies $n_{\rm ex}, n_{\rm ex}$ for a given occupancy n is

$$P^{[j]}(n_{\rm ex}|n) = \frac{Q_c^{[j]}(n_{\rm ex}, n - n_{\rm ex})}{Q^{[j]}(n)}. \tag{C.14}$$

The average number of filled exit sites of a n-occupied cell is

$$\left\langle n_{\rm ex}^{[j]}(n) \right\rangle = \sum_{n_{\rm ex}=0}^{K_{\rm ex}} n_{\rm ex} P^{[j]}(n_{\rm ex}|n). \tag{C.15}$$

The average energy of a structured cell of occupancy n is

$$\left\langle E^{[0]}(n) \right\rangle = \sum_{\boldsymbol{\eta}}^{(n)} E(\boldsymbol{\eta}) \frac{e^{-\beta E(\boldsymbol{\eta})}}{Q(n)}, \tag{C.16}$$

while for a less-structured cell

$$\left\langle E^{[j]}(n) \right\rangle = \sum_{n_{\rm ex}, n_{\rm in}}^{(n)} E^{[j]}(n_{\rm ex}, n_{\rm in}) P^{[j]}(n_{\rm ex}|n), \qquad j = 1, 2. \tag{C.17}$$

The above defined similarity conditions can therefore be expressed as:

$$Q^{[j]}(n) \approx Q^{[0]}(n), \tag{C.18}$$
$$\left\langle n_{\rm ex}^{[j]}(n) \right\rangle \approx \left\langle n_{\rm ex}^{[0]}(n) \right\rangle, \tag{C.19}$$
$$\left\langle E^{[j]}(n) \right\rangle \sim \left\langle E^{[0]}(n) \right\rangle, \tag{C.20}$$

with $j = 1, 2$. It should be noted that the similarity (C.20) is the weakest one. Indeed, partition function and average occupied exit sites have the highest priority due to their central role in the determination of global equilibrium and transport properties.

The similarity (C.18) produces similarity of occupancy distribution and adsorption isotherms due to the grand-canonical expression

$$p^{[j]}(n) = \frac{Q^{[j]}(n) \exp(\beta\mu n)}{\sum_m Q(m) \exp(\beta\mu m)}, \tag{C.21}$$

where μ is the chemical potential. The *loading* (average occupancy) is related to the chemical potential through $\langle n^{[j]} \rangle = \sum_n n\, p^{[j]}(n)$.

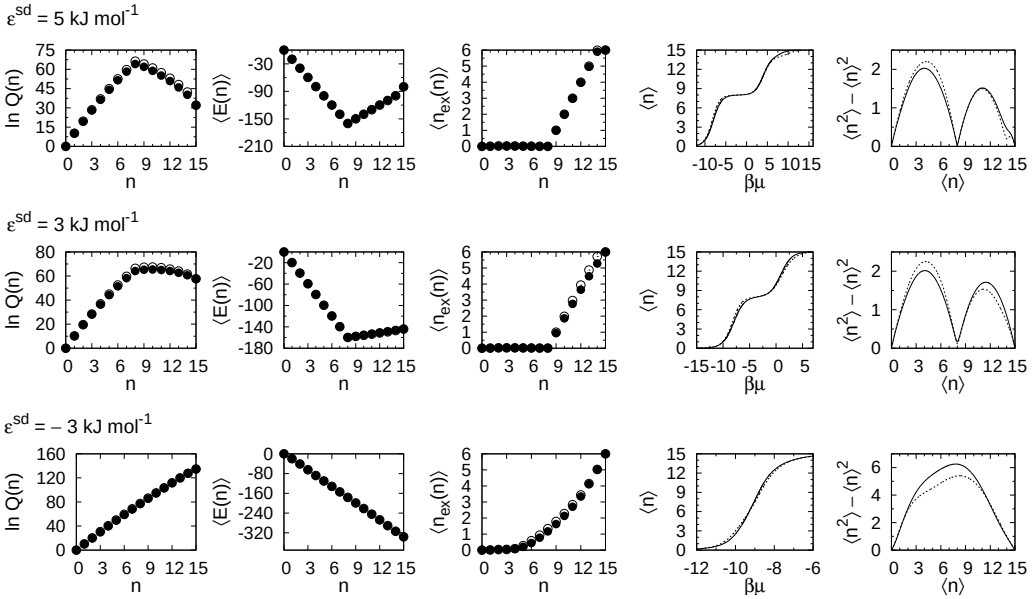

Figure C.1: Properties of the Test Cell # 1 before reduction (black dots, solid lines) and after reduction of level 1 (white circles, dashed lines). Each row refer to a different values of the energy parameter $\varepsilon^{\mathrm{sd}}$.

C.1 Less-structured cell of level $j = 1$

The parameters $\varepsilon_\alpha^{[1]}$ in Eq. (C.5) can be defined starting from the following conditional averages:

$$\left\langle E_\alpha^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle = \sum_{\boldsymbol{\eta}_{\mathrm{ex}}}^{(n_{\mathrm{ex}})} \sum_{\boldsymbol{\eta}_{\mathrm{in}}}^{(n_{\mathrm{in}})} \frac{E_{\alpha-\alpha}^{[0]}(\boldsymbol{\eta}) e^{-\beta E^{[0]}(\boldsymbol{\eta})}}{Q_{\mathrm{c}}^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}})}, \tag{C.22}$$

with $\alpha = \mathrm{ex}, \mathrm{in}$, representing the average energy of the α-type sites for a cell with partial occupancies $n_{\mathrm{ex}}, n_{\mathrm{in}}$, and

$$\left\langle E_{\mathrm{ex-in}}^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle = \sum_{\boldsymbol{\eta}_{\mathrm{ex}}}^{(n_{\mathrm{ex}})} \sum_{\boldsymbol{\eta}_{\mathrm{in}}}^{(n_{\mathrm{in}})} \frac{E_{\mathrm{ex-in}}^{[0]}(\boldsymbol{\eta}) e^{-\beta E^{[0]}(\boldsymbol{\eta})}}{Q_{\mathrm{c}}^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}})}, \tag{C.23}$$

representing the average interactions between occupied exit and inner sites for a cell with partial occupancies $n_{\mathrm{ex}}, n_{\mathrm{in}}$. Of course, the total average energy for given values of partial occupancies reads

$$\left\langle E^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle = \left\langle \left(E_{\mathrm{ex}}^{[0]} + E_{\mathrm{in}}^{[0]} + E_{\mathrm{ex-in}}^{[0]} \right) (n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle .$$

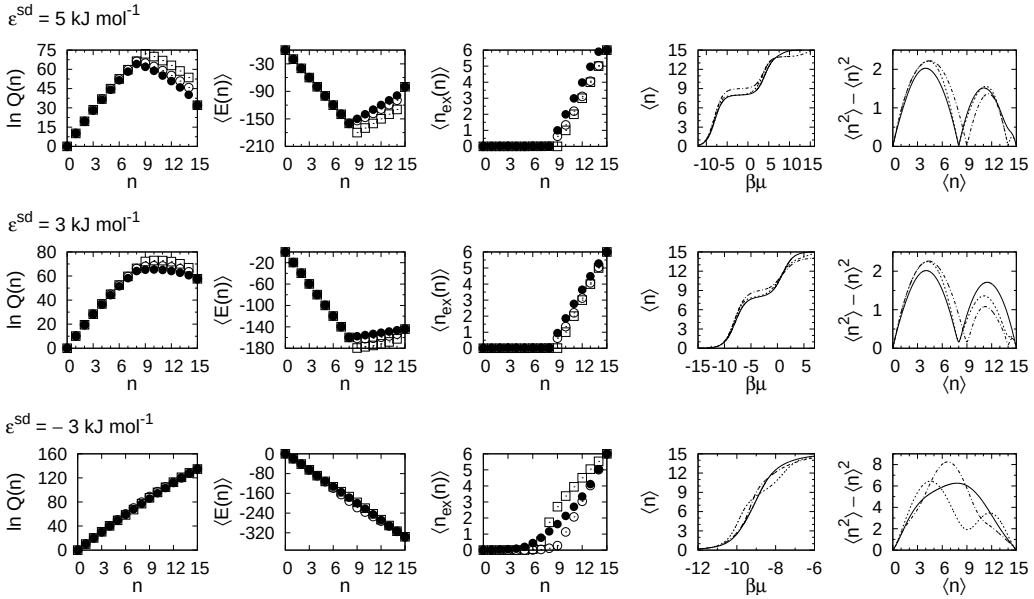

Figure C.2: Properties of the Test Cell # 1 before reduction (black dots, solid lines) and after reduction of level 2 (white circles and dashed lines refer to MODE 1; white squares and dash-dotted lines refer to mode 2). Each row refer to a different values of the energy parameter $\varepsilon^{\mathrm{sd}}$.

Then the first-level effective potentials can be defined as

$$\varepsilon_{\alpha}^{[1]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) = [1 - \delta_{\mathrm{kr}}(n_{\alpha}, 0)] \times \frac{1}{n_{\alpha}} \left[\left\langle E_{\alpha}^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle + \zeta_{\alpha} \left\langle E_{\mathrm{ex-in}}^{[0]}(n_{\mathrm{ex}}, n_{\mathrm{in}}) \right\rangle \right], \quad \text{(C.24)}$$

where ζ_{α} (satisfying $\zeta_{\mathrm{ex}} + \zeta_{\mathrm{in}} = 1$) is an arbitrary distribution function of the exit-inner average interaction energy between exit and inner sites. In this Appendix two examples of such distribution will be illustrated. At the end of the reduction procedure, the choice of the ζ_{α} distribution has no relevant effects on the cell thermodynamics.

i) MODE 1 shall indicate the case of $\zeta_{\alpha} = n_{\alpha}/n$, i.e. the exit-inner interaction distributed over both exit and inner site proportionally to their respective occupancy.

ii) MODE 2 shall indicate the case of $\zeta_{\mathrm{ex}} = 0$ and $\zeta_{\mathrm{in}} = 1$, i.e. the exit-inner interaction energy distributed over the inner sites only.

As a consequence of Eq. (C.24), it results $E^{[1]} = \langle E^{[0]} \rangle$ independently of the distribution ζ_{α}, therefore due to Eq. (C.13) the partition functions $Q_{\mathrm{c}}^{[1]}$ and

$Q^{[1]}$ are also independent of ζ_α. This means that the cell thermodynamics of the cell does not depend on the way the exit-inner contribution is distributed among exit and inner sites.

Using the procedure that is going to be described in Appendix C.2 one can determine the corrections $\Delta\varepsilon^\dagger_\alpha$ to the effective potentials in order to impose the strict equivalence of partition function and average occupied exit sites

$$Q^{[1]}(n) = Q^{[0]}(n), \tag{C.25}$$
$$\left\langle n^{[1]}_{\rm ex}(n)\right\rangle = \left\langle n^{[0]}_{\rm ex}(n)\right\rangle, \tag{C.26}$$

obtaining

$$\varepsilon^\dagger_\alpha(n_{\rm ex}, n_{\rm in}) = \varepsilon^{[1]}_\alpha(n_{\rm ex}, n_{\rm in}) + \Delta\varepsilon^\dagger_\alpha(n_{\rm ex} + n_{\rm in}),$$

with $\alpha = $ ex, in. The corrected first-level energy function results

$$E^\dagger(n_{\rm ex}, n_{\rm in}) = \sum_{\alpha={\rm ex,in}} n_\alpha \varepsilon^\dagger_\alpha(n_{\rm ex}, n_{\rm in}).$$

It should be noted that for different distributions ζ_α the correction leads to different effective site potentials, but, again, the resulting Hamiltonian $E^\dagger$ will be independent of ζ_α.

The first-level corrected energy function in Eq. (C.27) could replace the energy function $E^{[1]}$ in Eq. (C.13) to obtain a less-structured cell with the same cellular partition function and average number of occupied exit sites of the structured cell.

C.2 Corrected energy function

Starting from the coarse-grained energy function of level $j = 1, 2$, one can improve the similarity conditions (C.18) and (C.19) by making use of the following corrected energy function (temperature-dependence is implicit):

$$E^*(n_{\rm ex}, n_{\rm in}, \Delta\varepsilon_{\rm ex}(n), \Delta\varepsilon_{\rm in}(n)) = E^{[j=1,2]}(n_{\rm ex}, n_{\rm in}) + \Delta E^{[j=1,2]}(n_{\rm ex}, n_{\rm in}) \tag{C.27}$$

where $n = n_{\rm ex} + n_{\rm in}$, and the coarse-grained energy functions are given respectively by Eq. (C.5) for $[j = 1]$ and by Eq. (C.6) $[j = 2]$. The correction term is defined as

$$\begin{aligned}\Delta E^{[j=1,2]}(n_{\rm ex}, n_{\rm in}) &= n_{\rm ex}\Delta\varepsilon_{\rm ex}(n) + n_{\rm in}\Delta\varepsilon_{\rm in}(n) \\ &= n_{\rm ex}\delta\Delta\varepsilon(n) + n\Delta\varepsilon_{\rm in}(n),\end{aligned} \tag{C.28}$$

where $\Delta\varepsilon_{\rm ex}(n)$ and $\Delta\varepsilon_{\rm in}(n)$ are occupancy-dependent correction energies properly chosen in order to obtain the best approximations for the conditions (C.18) and (C.19). In this Appendix a way to determine their optimal value is proposed. In Eq. (C.28) the variable

$$\delta\Delta\varepsilon(n) = \Delta\varepsilon_{\rm ex}(n) - \Delta\varepsilon_{\rm in}(n)$$

has been introduced as the difference between the energy corrections to exit and inner sites, in order to evidence that, for a given occupancy n, w.r.t. the partition functions $Q^{[j]}$, $Q_c^{[j]}$, in the reduced partition functions

$$q_c^*(n_{\rm ex}, n_{\rm in}, \delta\Delta\varepsilon(n)) = \binom{K_{\rm ex}}{n_{\rm ex}}\binom{K_{\rm in}}{n_{\rm in}} \times \exp\left\{-\beta\left[E^{[j=1,2]}(n_{\rm ex}, n_{\rm in}) + n_{\rm ex}\delta\Delta\varepsilon(n)\right]\right\}, \quad \text{(C.29)}$$

$$q^*(n, \delta\Delta\varepsilon(n)) = \sum_{n_{\rm ex}, n_{\rm in}}^{(n)} q_c^*(n_{\rm ex}, n_{\rm in}, \delta\Delta\varepsilon(n)) \quad \text{(C.30)}$$

only one new independent variable $\delta\Delta\varepsilon(n)$ is appeared. The reduced partition functions in Eq. (C.30) are independent of $\Delta\varepsilon_{\rm in}(n)$ and proportional to the partition functions Q^*, Q_c^*:

$$\begin{aligned} Q_c^*(n_{\rm ex}, n_{\rm in}, \delta\Delta\varepsilon(n)) &= e^{-\beta n\Delta\varepsilon_{\rm in}(n)} q_c^*(n_{\rm ex}, n_{\rm in}, \delta\Delta\varepsilon(n)) \\ Q^*(n, \delta\Delta\varepsilon(n)) &= e^{-\beta n\Delta\varepsilon_{\rm in}(n)} q^*(n, \delta\Delta\varepsilon(n)) \end{aligned} \quad \text{(C.31)}$$

It turns out that the probability

$$\begin{aligned} P^*(n_{\rm ex}|n, \delta\Delta\varepsilon(n)) &= \frac{Q_c^*(n_{\rm ex}, n - n_{\rm ex}, \delta\Delta\varepsilon(n))}{Q^*(n, \delta\Delta\varepsilon(n))} \\ &= \frac{q_c^*(n_{\rm ex}, n - n_{\rm ex}, \delta\Delta\varepsilon(n))}{q^*(n, \delta\Delta\varepsilon(n))} \end{aligned} \quad \text{(C.32)}$$

is a function of the new variable $\delta\Delta\varepsilon(n)$ and it is independent of $\Delta\varepsilon_{\rm in}(n)$. Of course, this is valid also for the variable

$$\left\langle n_{\rm ex}^*(n, \delta\Delta\varepsilon(n))\right\rangle = \sum_{n_{\rm ex}=0}^{K_{\rm ex}} n_{\rm ex} P^*(n_{\rm ex}|n, \delta\Delta\varepsilon(n)). \quad \text{(C.33)}$$

Then, in order to find the corrections $\Delta\varepsilon_{\rm ex}(n)$ and $\Delta\varepsilon_{\rm in}(n)$ satisfying the requirements (C.20), (C.18), and (C.19), one can apply the following procedure: For each occupancy $n > 0$, we:

- find the value of $\delta\Delta\varepsilon(n)$ solving the equation

$$\left\langle n^{*}_{\mathrm{ex}}(n, \delta\Delta\varepsilon(n)) \right\rangle = \left\langle n^{[0]}_{\mathrm{ex}}(n) \right\rangle \tag{C.34}$$

  Let $\delta\Delta\varepsilon^{\dagger}(n)$ be the zero of Eq. (C.34), and $q^{\dagger}_{c}(n_{\mathrm{ex}}, n_{\mathrm{in}})$, $q^{\dagger}(n)$ the corresponding reduced partition functions;

- determine the correction to the effective inner-site potential

$$\Delta\varepsilon^{\dagger}_{\mathrm{in}}(n) = (\beta n)^{-1}[\ln q^{\dagger}(n) - \ln Q^{[0]}(n)]$$

  ensuring the partition function $Q^{\dagger}(n)$ of the new coarse-grained cell to be equivalent to $Q^{[0]}(n)$;

- determine the correction to the effective exit-site potential $\Delta\varepsilon^{\dagger}_{\mathrm{ex}}(n) = \delta\Delta\varepsilon(n) + \Delta\varepsilon^{\dagger}_{\mathrm{in}}(n)$.

C.3 Less-structured cell of level $j = 2$

A further reduction can be obtained by using the first level of reduction as the starting point to construct the energy function $E^{[2]}$ of Eq. (C.6), in which the energy parameters $\varepsilon^{[2]}_{\alpha}$ depend on the total occupancy of the cell instead of partial occupancies. They are defined as

$$\varepsilon^{[2]}_{\alpha}(n) = \sum_{n_{\mathrm{ex}}, n_{\mathrm{in}}}^{(n)} \varepsilon^{[1]}_{\alpha}(n_{\mathrm{ex}}, n_{\mathrm{in}}) P^{[1]}(n_{\mathrm{ex}}|n), \tag{C.35}$$

where $\alpha = \mathrm{ex}, \mathrm{in}$. The agreement of thermodynamic properties can be improved by determining the corrections to the potentials with the procedure described in Appendix C.2. The resulting corrections shall be indicated as $\Delta\varepsilon^{\ddagger}_{\alpha}$ where the superscript ‡ will be used instead of † to distinguish such a quantity from $\Delta\varepsilon^{\dagger}_{\alpha}$, since the same correction procedure returns different corrections when applied to $E^{[1]}$ and $E^{[2]}$. The corrected second-level potentials will be obtained as

$$\varepsilon^{\ddagger}_{\alpha}(n) = \varepsilon^{(2)}_{\alpha}(n) + \Delta\varepsilon^{\ddagger}_{\alpha}(n),$$

(with $\alpha = \mathrm{ex}, \mathrm{in}$) to form the corrected second-level energy function

$$E^{\ddagger}(n_{\mathrm{ex}}, n_{\mathrm{in}}) = \sum_{\alpha=\mathrm{ex,in}} n_{\alpha} \varepsilon^{\ddagger}_{\alpha}(n_{\mathrm{ex}} + n_{\mathrm{in}}),$$

satisfying the equivalences

$$Q^{[2]}(n) = Q^{[1]}(n), \tag{C.36}$$

$$\left\langle n_{\mathrm{ex}}^{[2]}(n) \right\rangle = \left\langle n_{\mathrm{ex}}^{[1]}(n) \right\rangle . \tag{C.37}$$

It should be noted that, in general the particular choice of the exit-inner distribution ζ_α does influence the functions $E^{[1]}$ and $E^{[2]}$. Nevertheless, the cell free energy is preserved by means of the equivalences (C.25), (C.26), (C.36), and (C.37), while the average cell energy is mimicked according to the similarity (C.20) with a very good approximation as can be seen from the example in Figure C.4 where a cell with a relatively complex (although local) energetics has been taken as example to prove the effectiveness of the reduction procedure.

C.4 Test of the reduction procedure

Several parametrizations of the structured cell of Figure 2.1 will be investigated which have been intentionally chosen as extreme (i.e. characterized by large repulsive or attractive interaction) in order the test the procedure of reduction to a less-structured cell to be more significant.

C.5 Test cell # 1

The fixed energy parameters ε_i are set as $\varepsilon_1 = \cdots = \varepsilon_6 = -10$ kJ mol^{-1} for the exit sites, and $\varepsilon_7 = \cdots = \varepsilon_{15} = -20$ kJ mol^{-1} for the inner sites. The values of the parameters in Eq. (C.3) are $\ell^{\mathrm{tr}} = \sqrt{3}\ell$ and $a = 6$. Three values of the interaction parameter $\varepsilon^{\mathrm{sd}}$ will be investigated, namely -3, 3, and 5 kJ mol^{-1}. Temperature is fixed at 300 K.

Test Cell #1 before reduction. The local properties of the cell before and after reduction of level 1 are shown in Figure C.1. As can be seen, in the case of purely attractive interactions the average cell energy decreases with occupancy. In the presence of purely repulsive interactions at some intermediate occupancy, say n_0, the interaction energy becomes too positive to stabilize a higher number of occupied inner sites so that the exit sites start being filled, causing a sudden increase in the average cell energy. This produces a singular point (a first-order phase transition[81]) in the chemical potential around the loading where the occupancy n_0 is highly the most probable one, say $\langle n \rangle_0$. It should be noted that in the case of non-interacting guests such a singular point would be observed around loading K_{in} (which in the present

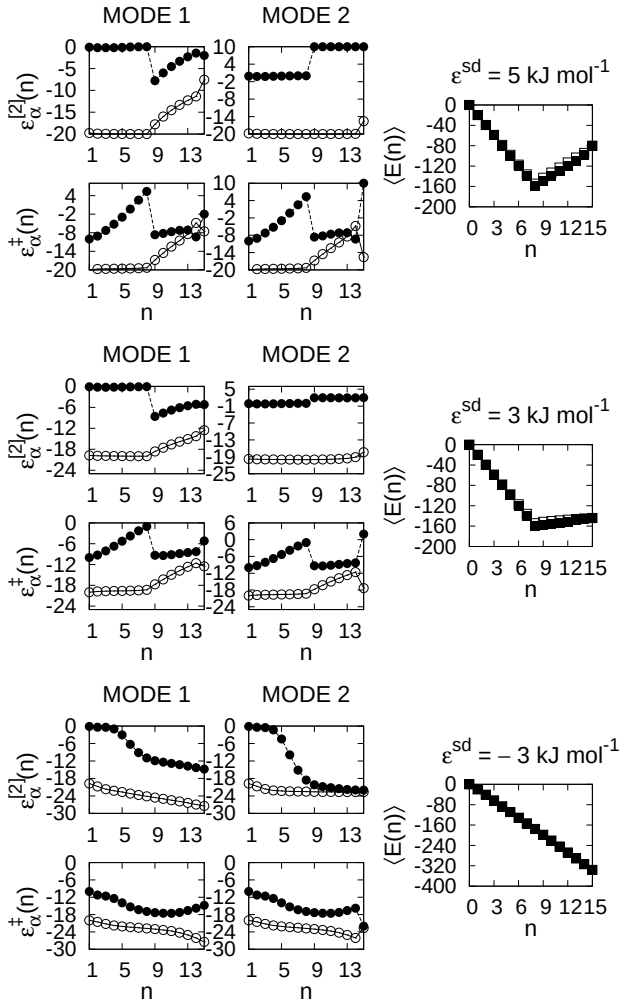

Figure C.3: Properties of the closed cell after reduction of level 2 of the Test Cell # 2. In the two columns at the left, filled and empty circles refer respectively to exit and inner site potentials. At the right, the average energies $\langle E^{[0]}(n)\rangle$ and $\langle E^{\ddagger}(n)\rangle$ are indicated respectively with filled and open squares.

case is 9) and only at low temperatures, due to the high energy difference between exit and inner sites which causes a K_{in}-occupied cell to preferably host all guests in the inner sites.[54] In the present parametrization instead $n_0 = 8$, which is less than K_{in}. This is because in a cell with occupancy $> n_0$ the repulsive interactions balance the attractive contribution due to the largely negative potential energy of the inner sites, thus starting to assist the occupation of exit sites; since in the present case each guest in an exit site interacts with four inner guests, and the exit sites potential is much less negative than the inner ones, the filling of the exit sites causes the repulsive

contribution to sharply increase.

First and second levels of reduction. The first-level reduction preserve exactly the value of the average cell energy, therefore at this stage the choice of MODE 1 or 2 is unimportant. Partition function and average filled exit sites are reproduced with good approximation; nevertheless, small deviations in the properties of the reduced cell w.r.t. the structured cell cause larger disagreement in the macroscopic properties μ and $\langle n^2\rangle - \langle n\rangle^2$. This becomes also more evident in the second-level reduction (Figure C.2), where in case of repulsive interactions the phase transition is produced at loading 8 for MODE 1, and at loading 9 for MODE 2. This is because in MODE 2 guest-guest interactions are entirely charged to the inner sites, while in MODE 1 they are equally distributed on all guest particles. The situation is reversed in case of purely attractive interactions, where instead the effect of high stabilization of configurations with higly-occupied inner sites is better reproduced (at least qualitatively) by MODE 2.

Correction. Anyway, the use of second-level energy parameters $\varepsilon_\alpha^{[2]}$ as obtained by simple averaging of interactions produces large disagreements between thermodynamic properties of structured and less-structured cell, due to the effect of topology which can be embedded into the less-structured cell representation only by means of the correction procedure. In Figure C.3 the effect of such correction on the effective potentials is shown (corrected chemical potential and variance are not shown since they have exactly the same values as those obtained for the structured cell). First of all, the choice of the MODE has little or no effects on the obtained trend of $\varepsilon_\alpha^\ddagger(n)$. In the case of attractive interactions, for MODE 1 the topology effect is reproduced by more negative site potentials at low occupancies (where guests are more likely to be found in the inner sites), while for MODE 2 in addition less negative potentials are required at higher occupancies (where exit sites start to be filled). For repulsive interactions, at low occupancies the corrected potentials largely differ in the exit sites, since a large contribution to the repulsion effect in the cell energy comes from the interaction between each guest in the exit sites and the respective 4 nearest inner sites (if occupied). Repulsive effects add to the effective inner sites potential also at the higher occupancies. The effect on the average energy is relatively small, therefore the corrected less-structured cell can be considered as a good approximation of the structured cell.

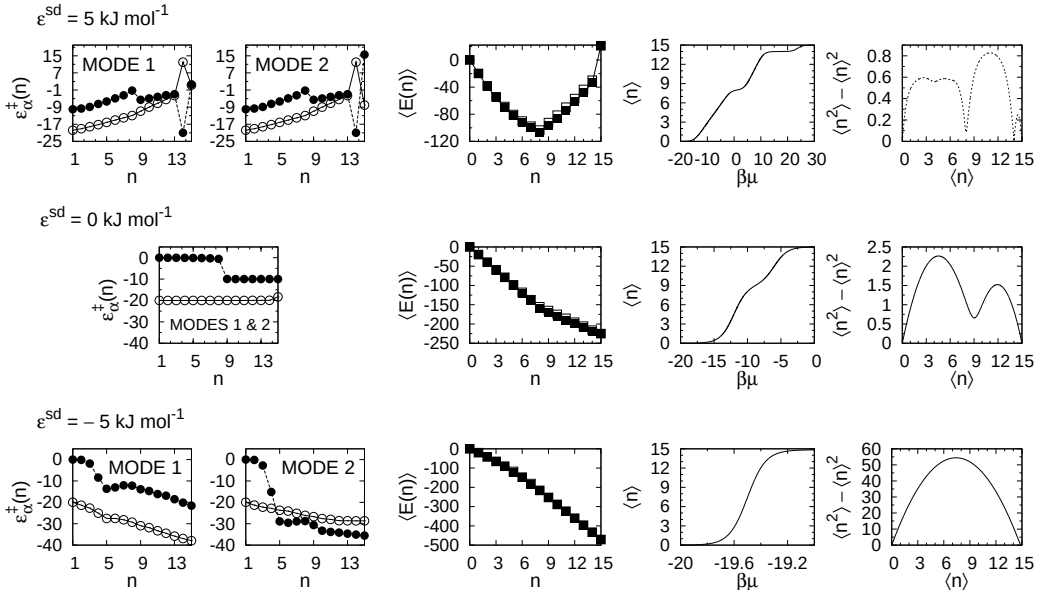

Figure C.4: Results of the reduction-correction procedure for the Test Cell # 2.

C.6 Test cell # 2

The structured cell's energy parameters are set as: $\varepsilon_1 = \cdots = \varepsilon_6 = -10$ kJ mol^{-1} for the exit sites, $\varepsilon_7 = \cdots = \varepsilon_{14} = -20$ kJ mol^{-1} and $\varepsilon_{15} = -5$ kJ mol^{-1} for the inner sites . The values of parameters in Eq. (C.3) are $\ell^{\mathrm{tr}} = 2\sqrt{2}\ell$ and $a = 3$. Three value of the interaction parameters $\varepsilon^{\mathrm{sd}}$ have been investigated, namely -5, 0, and 5 kJ mol^{-1}.
Temperature is fixed at a relatively low value, i.e. $T = 200$ K in order to enhance the presence of multiple phase transitions at various loadings causing adsorption isotherms to show several singularities.

In Figure C.4 results of the reduction procedure are shown. As can be seen, although the equilibrium properties are complicated by singularities and irregularities (shown in the last two columns) due to the higher heterogeneity in the energy structure the reduction procedure produces satisfactory results.

Appendix D

Collective diffusion coefficient

D.1 Collective diffusion coefficient

Since the coordinates of the center-of-mass at time t are given by

$$\mathbf{r}_{\mathrm{CM}}(t) = \frac{1}{N}\sum_{I=1}^{N}\mathbf{r}_I(t), \tag{D.1}$$

the center-of-mass displacement at time t is

$$\Delta\mathbf{r}_{\mathrm{CM}}(t) = \mathbf{r}_{\mathrm{CM}}(t+\tau) - \mathbf{r}_{\mathrm{CM}}(t) = \frac{1}{N}\sum_{I=1}^{N}\sum_{z=0}^{Z-1}\delta\mathbf{r}_I(z\tau), \tag{D.2}$$

where $Z = t/\tau$ is the number of time steps needed for the system to evolve from time 0 to time t, and $\delta\mathbf{r}_I(z\tau)$ is the instantaneous cell-to-cell displacement function, Eq. (4.19), for the I-th guest. The mean-squared center-of-mass displacement results therefore

$$\begin{aligned}\left\langle\left[\Delta\mathbf{r}_{\mathrm{CM}}(t)\right]^2\right\rangle &= \frac{1}{N^2}\left\langle\sum_{I=1}^{N}\sum_{J=1}^{N}\sum_{z=1}^{Z-1}\sum_{w=1}^{Z-1}\delta\mathbf{r}_I(z\tau)\cdot\delta\mathbf{r}_J(w\tau)\right\rangle \\ &= \frac{1}{N^2}\sum_{I=1}^{N}\sum_{J=1}^{N}\left\{Z\left\langle\delta\mathbf{r}_I(0)\cdot\delta\mathbf{r}_J(0)\right\rangle + 2\sum_{z=1}^{Z-1}(Z-z)\left\langle\delta\mathbf{r}_I(z\tau)\cdot\delta\mathbf{r}_J(0)\right\rangle\right\}\end{aligned} \tag{D.3}$$

The instantaneous variation of the average center-of-mass displacement results

$$\begin{aligned}
&\left\langle [\Delta \mathbf{r}_{\mathrm{CM}}(t+\tau)]^2 \right\rangle - \left\langle [\Delta \mathbf{r}_{\mathrm{CM}}(t)]^2 \right\rangle \\
=&\frac{1}{N^2} \sum_{I=1}^{N} \sum_{J=1}^{N} \left\{ \langle \delta \mathbf{r}_I(0) \cdot \delta \mathbf{r}_J(0) \rangle + 2 \sum_{z=1}^{Z} \langle \delta \mathbf{r}_I(z\tau) \cdot \delta \mathbf{r}_J(0) \rangle \right\} \qquad \text{(D.4)} \\
=&\frac{1}{N} \left[\left\langle [\Delta \mathbf{r}(t+\tau)]^2 \right\rangle - \left\langle [\Delta \mathbf{r}(t)]^2 \right\rangle \right] \\
&+ \frac{1}{N^2} \sum_{\substack{1 \le I, I \le N \\ I \ne J}} \left\{ \langle \delta \mathbf{r}_I(0) \cdot \delta \mathbf{r}_J(0) \rangle + 2 \sum_{z=1}^{Z} \langle \delta \mathbf{r}_I(z\tau) \cdot \delta \mathbf{r}_J(0) \rangle \right\}, \qquad \text{(D.5)}
\end{aligned}$$

where in the last equality, Eq. (D.5), the self-term $\left\langle [\Delta \mathbf{r}(t+\tau)]^2 \right\rangle$ (given in Eq. (4.21)) has been separated from the mixed terms. Application of Eq. (4.26) gives the expression for the collective diffusivity in terms of the DMCF, Eq. (4.27).

Appendix E

Sampling schemes and detailed balance

E.1 Mixed Arrhenius-Metropolis acceptance probability for jump randomization

In the present Appendix a jump acceptance probability $p^R_{\text{jump}}(\boldsymbol{\eta} \to \boldsymbol{\eta}')$ shall be derived.

During each event of jump, the selected guest changes its position in the cell. As a result, the relation between the configurations $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$ is that there are exactly two sites, i.e. the departure site j and the target site k, satisfying

$$\begin{aligned} \eta_i = \eta'_i \quad & \text{for each } i \in [1, K] \text{ with } i \neq j, k \\ \eta_j = 1, \quad \eta_k = 0 \quad & \text{(before the jump)} \\ \eta'_j = 0, \quad \eta'_k = 1 \quad & \text{(after the jump)} \end{aligned} \tag{E.1}$$

The detailed balance condition reads v

$$P(\boldsymbol{\eta}|n)\, p^R_{\text{jump}}(\boldsymbol{\eta} \to \boldsymbol{\eta}') = P(\boldsymbol{\eta}'|n)\, p^R_{\text{jump}}(\boldsymbol{\eta}' \to \boldsymbol{\eta}). \tag{E.2}$$

where $P(\boldsymbol{\eta}|n)$ is given by Eq. (2.30) and it is implicit that the two configurations, $\boldsymbol{\eta}$ and $\boldsymbol{\eta}'$, have the same occupancy. Inserting Eq. (2.30) into Eq. (E.2) one obtains

$$e^{-\beta\Phi(\boldsymbol{\eta})} \prod_{i=1}^{K} e^{-\beta\eta_i f_i^o}\, p^R_{\text{jump}}(\boldsymbol{\eta} \to \boldsymbol{\eta}') = e^{-\beta\Phi(\boldsymbol{\eta}')} \prod_{i=1}^{K} e^{-\beta\eta'_i f_i^o}\, p^R_{\text{jump}}(\boldsymbol{\eta}' \to \boldsymbol{\eta}). \tag{E.3}$$

Using the conditions (E.1) the detailed balance equation can be rewritten as

$$e^{-\beta\Phi(\boldsymbol{\eta})}e^{-\beta f_j^o}\prod_{\substack{i=1,\dots,K\\ i\neq j,k}} e^{-\beta\eta_i f_i^o}\, p^R_{\text{jump}}(\boldsymbol{\eta}\to\boldsymbol{\eta}') = e^{-\beta\Phi(\boldsymbol{\eta}')}e^{-\beta f_k^o}\prod_{\substack{i=1,\dots,K\\ i\neq j,k}} e^{-\beta\eta_i f_i^o}\, p^R_{\text{jump}}(\boldsymbol{\eta}'\to\boldsymbol{\eta}). \quad \text{(E.4)}$$

The product factor may be eliminated from both sides of the equation, thus obtaining

$$e^{-\beta\Phi(\boldsymbol{\eta})}e^{-\beta f_j^o}\, p^R_{\text{jump}}(\boldsymbol{\eta}\to\boldsymbol{\eta}') = e^{-\beta\Phi(\boldsymbol{\eta}')}e^{-\beta f_k^o}\, p^R_{\text{jump}}(\boldsymbol{\eta}'\to\boldsymbol{\eta}). \quad \text{(E.5)}$$

The jump probabilities can be defined as

$$\begin{aligned} p^R_{\text{jump}}(\boldsymbol{\eta}\to\boldsymbol{\eta}') &= e^{\beta\Phi(\boldsymbol{\eta})}e^{\beta f_j^o}\, u(\boldsymbol{\eta},\boldsymbol{\eta}'),\\ p^R_{\text{jump}}(\boldsymbol{\eta}'\to\boldsymbol{\eta}) &= e^{\beta\Phi(\boldsymbol{\eta}')}e^{\beta f_k^o} u(\boldsymbol{\eta}',\boldsymbol{\eta}) \end{aligned} \quad \text{(E.6)}$$

where to preserve the detailed balance it is necessary that the function u satisfies $u(\boldsymbol{\eta},\boldsymbol{\eta}') = u(\boldsymbol{\eta}',\boldsymbol{\eta})$. The jump probability can be optimized by defining the function u as

$$u(\boldsymbol{\eta},\boldsymbol{\eta}') = C\exp\left\{-\beta\max\left[\Phi(\boldsymbol{\eta}),\Phi(\boldsymbol{\eta}')\right]\right\}, \quad \text{(E.7)}$$

where $C\in(0,1]\in\mathbb{R}$ can be set as a constant or as a function $C(\boldsymbol{\eta},\boldsymbol{\eta}')$, provided that $C(\boldsymbol{\eta},\boldsymbol{\eta}') = C(\boldsymbol{\eta}',\boldsymbol{\eta})$. Therefore the jump will be accepted with probability $p^R_{\text{jump}}(\boldsymbol{\eta}\to\boldsymbol{\eta}')$ as given in Eq. (3.9).

Since for each jump p^R_{jump} satisfies the detailed balance, the entire randomization scheme satisfies the detailed balance.

Let us consider an n-occupied cell undergoing one randomization with $\mathcal{R}^{\text{jump}}$. In such a cell, n jump attempts (one for each guest) will generate the following sequence of transformations: $\boldsymbol{\eta}\to\boldsymbol{\eta}_1\to\boldsymbol{\eta}_2\to\cdots\to\boldsymbol{\eta}_n\equiv\boldsymbol{\eta}^R$ where the configurations path

$$\{\boldsymbol{\eta}_J\}_{J=1,\dots,n-1} = \boldsymbol{\eta}_1,\boldsymbol{\eta}_2,\dots,\boldsymbol{\eta}_{n-1}$$

connects the starting configuration $\boldsymbol{\eta}$ with the post-randomization one $\boldsymbol{\eta}^R$. The probability of a cell of configuration $\boldsymbol{\eta}'$ to pass to a configuration $\boldsymbol{\eta}''$ is given by:

$$\widetilde{p}^R(\boldsymbol{\eta}'\to\boldsymbol{\eta}'') = \omega(\boldsymbol{\eta}'\to\boldsymbol{\eta}'')\, p^R_{\text{jump}}(\boldsymbol{\eta}'\to\boldsymbol{\eta}'') \quad \text{(E.8)}$$

where $\omega(\boldsymbol{\eta}' \to \boldsymbol{\eta}'')$ is the probability of a trial configuration $\boldsymbol{\eta}''$ to be chosen given that the input one is $\boldsymbol{\eta}'$, and $p^R_{\text{jump}}(\boldsymbol{\eta}' \to \boldsymbol{\eta}'')$ has been defined in Eq. (E.2). According to the procedure illustrated in Section 3.2.2 for the jump randomization, $\omega(\boldsymbol{\eta}' \to \boldsymbol{\eta}'') = \omega(\boldsymbol{\eta}'' \to \boldsymbol{\eta}')$ thus Eq. (E.8) is satisfied. Now, for each passage of the path a detailed balance equation like Eq. (E.2) can be written, that is

$$P(\boldsymbol{\eta}|n)\, \widetilde{p}^R(\boldsymbol{\eta} \to \boldsymbol{\eta}_1) = P(\boldsymbol{\eta}_1|n)\, \widetilde{p}^R(\boldsymbol{\eta}_1 \to \boldsymbol{\eta}) \tag{E.9}$$

$$P(\boldsymbol{\eta}_1|n)\, \widetilde{p}^R(\boldsymbol{\eta}_1 \to \boldsymbol{\eta}_2) = P(\boldsymbol{\eta}_2|n)\, \widetilde{p}^R(\boldsymbol{\eta}_2 \to \boldsymbol{\eta}_1) \tag{E.10}$$

$$\dots$$

$$P(\boldsymbol{\eta}_{n-1}|n)\, \widetilde{p}^R(\boldsymbol{\eta}_{n-1} \to \boldsymbol{\eta}_n) = P(\boldsymbol{\eta}_n|n)\, \widetilde{p}^R(\boldsymbol{\eta}_n \to \boldsymbol{\eta}_{n-1}). \tag{E.11}$$

By multiplying side by side Eqs. (E.9) to (E.11) (thus obtaining the detailed balance equation for the path $\{\boldsymbol{\eta}_J\}$) and sum over all the possible paths one obtains the detailed balance equation for the jump randomization:

$$P(\boldsymbol{\eta}|n) \sum_{\{\boldsymbol{\eta}_J\}} \widetilde{p}^R\left(\boldsymbol{\eta} \to \boldsymbol{\eta}^R \middle| \{\boldsymbol{\eta}_J\}\right) = P(\boldsymbol{\eta}^R|n) \sum_{\{\boldsymbol{\eta}_J\}} \widetilde{p}^R\left(\boldsymbol{\eta}^R \to \boldsymbol{\eta} \middle| \{\boldsymbol{\eta}_J\}\right) \tag{E.12}$$

where the probabilities

$$\begin{aligned}
\widetilde{p}^R\left(\boldsymbol{\eta} \to \boldsymbol{\eta}^R \middle| \{\boldsymbol{\eta}_J\}\right) &= \widetilde{p}^R(\boldsymbol{\eta} \to \boldsymbol{\eta}_1) \prod_{J=1}^{n-1} \widetilde{p}^R(\boldsymbol{\eta}_J \to \boldsymbol{\eta}_{J+1}) \\
\widetilde{p}^R\left(\boldsymbol{\eta}^R \to \boldsymbol{\eta} \middle| \{\boldsymbol{\eta}_J\}\right) &= \widetilde{p}^R(\boldsymbol{\eta}_1 \to \boldsymbol{\eta}) \prod_{J=1}^{n-1} \widetilde{p}^R(\boldsymbol{\eta}_{J+1} \to \boldsymbol{\eta}_J),
\end{aligned} \tag{E.13}$$

represent the transition probability for the path $\{\boldsymbol{\eta}_J\}$ respectively in the forward and reverse direction.

E.2 Global detailed balance for randomization

At each instant of time, the *randomization partition function* can be written as:

$$Z^R(t) = \prod_{\mathbf{r} \in \mathscr{L}} Q\left(n(\mathbf{r}, t)\right), \tag{E.14}$$

that is, the post-randomization lattice configuration $\boldsymbol{\eta}^R(\mathscr{L}, t)$ will be chosen among the configurations contained (and weighted) in Eq. (E.14).

As shown in Section 3.2, the local operator $\mathcal{R}$ acts independently from cell to cell. Then if one indicates as

- $p\left(\boldsymbol{\eta}(\mathbf{r})\right)$ the probability of the cell $\mathbf{r}$ configuration to be $\boldsymbol{\eta}(\mathbf{r})$, and
- $p\left(n(\mathbf{r})\right)$ the probability of the cell $\mathbf{r}$ configuration to have occupancy n, and
- $p\left(\boldsymbol{\eta}(\mathbf{r})|n(\mathbf{r})\right)$ the conditional probability of the cell $\mathbf{r}$ to have configuration $\boldsymbol{\eta}(\mathbf{r})$ given that its occupancy is $n(\mathbf{r})$,

the detailed balance for the single cell reads:

$$p\left(\boldsymbol{\eta}(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}(\mathbf{r}) \to \boldsymbol{\eta}^R(\mathbf{r})\big) = p\left(\boldsymbol{\eta}^R(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}^R(\mathbf{r}) \to \boldsymbol{\eta}(\mathbf{r})\big) \tag{E.15}$$

which, using the relation $p(\boldsymbol{\eta}) = p(n(\boldsymbol{\eta}))p(\boldsymbol{\eta}|n(\boldsymbol{\eta}))$ and the fact that $n^R(\boldsymbol{\eta}) = n(\boldsymbol{\eta})$, can be reduced to

$$p\left(\boldsymbol{\eta}(\mathbf{r})|n(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}(\mathbf{r}) \to \boldsymbol{\eta}^R(\mathbf{r})\big) = p\left(\boldsymbol{\eta}^R(\mathbf{r})|n(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}^R(\mathbf{r}) \to \boldsymbol{\eta}(\mathbf{r})\big)\,. \tag{E.16}$$

The advantage of Eq. (E.16) w.r.t. Eq. (E.15) is that it involves only probabilities which are *locally* defined. Eq. (E.16) is satisfied both for the memoryless randomization and the jump randomization. First of all, the term $p\left(\boldsymbol{\eta}(\mathbf{r})|n(\mathbf{r})\right)$ is given by $P(\boldsymbol{\eta}, n)$ defined in Eq. (2.30). Secondly, for randomization with $\mathcal{R}^{\text{pf}}$ Eq. (E.16) is trivially satisfied by use of Eq. (3.7), while for $\mathcal{R}^{\text{jump}}$ the transition probability satisfies detailed balance as discussed in Appendix E.1.

Extending to the entire lattice, the probability of the lattice configuration $\boldsymbol{\eta}(\mathscr{L})$ is

$$\begin{aligned} p\left(\boldsymbol{\eta}(\mathscr{L})\right) &= p\left(n(\mathscr{L})\right)\, p\left(\boldsymbol{\eta}(\mathscr{L})|n(\mathscr{L})\right) \\ &= p\left(n(\mathscr{L})\right) \prod_{\mathbf{r}\in\mathscr{L}} p\left(\boldsymbol{\eta}(\mathbf{r})|n(\mathbf{r})\right). \end{aligned} \tag{E.17}$$

Using the independence of cells during randomization, the detailed balance for the lattice randomization, expressed by

$$p\left(\boldsymbol{\eta}(\mathscr{L})\right)\, p^R\big(\boldsymbol{\eta}(\mathscr{L}) \to \boldsymbol{\eta}^R(\mathscr{L})\big) = p\left(\boldsymbol{\eta}^R(\mathscr{L})\right)\, p^R\big(\boldsymbol{\eta}^R(\mathscr{L}) \to \boldsymbol{\eta}(\mathscr{L})\big) \tag{E.18}$$

can be written as

$$\begin{aligned} &\prod_{\mathbf{r}\in\mathscr{L}} p\left(\boldsymbol{\eta}(\mathbf{r})|n(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}(\mathbf{r}) \to \boldsymbol{\eta}^R(\mathbf{r})\big) \\ &\qquad = \prod_{\mathbf{r}\in\mathscr{L}} p\left(\boldsymbol{\eta}^R(\mathbf{r})|n(\mathbf{r})\right)\, p^R\big(\boldsymbol{\eta}^R(\mathbf{r}) \to \boldsymbol{\eta}(\mathbf{r})\big)\,. \end{aligned} \tag{E.19}$$

which, since Eq. (E.16) is satisfied, is also satisfied.

E.3 Block propagation: partitioning scheme in details

In the present Appendix a detailed description of the partitioning scheme of the ThPCA shall be given.

Cell lines. First of all, a *cell line* shall be defined. Let us consider a particular cell $\mathbf{R} = \{R_j\} \in \mathscr{L}$ (with $j = 1, \ldots, d$), where the set of coordinates $\{R_j\}$ (with $R_j = \mathbf{R} \cdot \mathbf{e}_j$ according to Eq. (2.9)) specifies its position in $\mathscr{L}$. If the specification of the k-th coordinate of $\mathbf{R}$ is omitted, then we are indicating the cell line at the position $\{R_j\}_{j \neq k}$ spanning the lattice perpendicularly from one side to the opposite one along the k-th direction. As example, the cell line $\{R_j\}_{j \neq 1}$ is the set of the following L cells:

$$\begin{pmatrix} 0 \\ R_2 \\ \vdots \\ R_d \end{pmatrix}, \begin{pmatrix} \lambda \\ R_2 \\ \vdots \\ R_d \end{pmatrix}, \ldots, \begin{pmatrix} (L-1)\lambda \\ R_2 \\ \vdots \\ R_d \end{pmatrix}. \tag{E.20}$$

An alternative notation for the cell coordinates will be useful for a compact definition of cell line's partition. Let us consider the particular cell $\mathbf{R}$. Of course it belongs to the cell line $\{R_j\}_{j \neq k}$, where k is a particular direction. Together with the cell line position $\{R_j\}_{j \neq k}$, the additional specification of R_k returns the exact coordinates of $\mathbf{R}$:

$$\mathbf{R} \equiv \left(r_k = R_k \,\middle|\, \{R_j\}_{j \neq k} \right). \tag{E.21}$$

As example,

$$\begin{pmatrix} \lambda \\ R_2 \\ \vdots \\ R_d \end{pmatrix} \equiv \left(r_1 = \lambda \,\middle|\, \{R_j\}_{j \neq 1} \right). \tag{E.22}$$

A pair of neighboring cells along the k-th direction shall be indicated as

$$\left\{ \mathbf{R}, \mathbf{R}^k \right\} = \bigcup_{z=0}^{1} \left(r_k = R_k + z\lambda \,\middle|\, \{R_j\}_{j \neq k} \right). \tag{E.23}$$

As example, if we take the cell of Eq. (E.22) and the next cell along the 1-st direction,

$$\left\{\begin{pmatrix}\lambda\\R_2\\\vdots\\R_d\end{pmatrix},\begin{pmatrix}2\lambda\\R_2\\\vdots\\R_d\end{pmatrix}\right\}\equiv\bigcup_{z=0}^{1}\begin{pmatrix}(1+z)\lambda\\R_2\\\vdots\\R_d\end{pmatrix}$$
$$\equiv\bigcup_{z=0}^{1}\left(r_1=(1+z)\lambda\,\middle|\,\{R_j\}_{j\neq1}\right). \quad \text{(E.24)}$$

Alternating pairs of neighboring cells on the same line form a partition. Two partitions of this kind (denoted as $g^{[0]}$ and $g^{[1]}$) are possible: Let us take as example the first partition of the cell line $\{R_j\}_{j\neq1}$:

$$g_1^{[0]}\left(\{R_j\}_{j\neq1}\right)=\bigcup_{i=0}^{L/2-1}\bigcup_{z=0}^{1}\left(r_1=(2i+z)\lambda\,\middle|\,\{R_j\}_{j\neq1}\right). \quad \text{(E.25)}$$

Taking into account the periodicity of the lattice, the second partition will be

$$g_1^{[1]}\left(\{R_j\}_{j\neq1}\right)=\bigcup_{i=0}^{L/2-1}\bigcup_{z=0}^{1}\left(r_1=(2i+1+z)\lambda\,\middle|\,\{R_j\}_{j\neq1}\right), \quad \text{(E.26)}$$

where the sum is a sum modulo L. So the general formulation of the two cell line partitions is obtained:

$$g_k^{[J]}\left(\{R_j\}_{j\neq k}\right)=\bigcup_{i=0}^{L/2-1}\bigcup_{z=0}^{1}\left(r_k=(2i+J+z)\lambda\,\middle|\,\{R_j\}_{j\neq k}\right), \quad \text{(E.27)}$$

$$\text{with } J=0,1. \quad \text{(E.28)}$$

With this formulation, the cell pairs belonging to each partition are independent and can undergo propagation simultaneously.

It should be noted that, since a sequence of the two line partitions $g_k^{[0]}$ and $g_k^{[1]}$ allows each of the L belonging cells to communicate with its two own neighbors on the same line, the two partitions are complementary. From now on, an overline over a partition will denote its complementary: (This notation applies only on partitions. It must not be confused with the notation of average, which applies instead on observables)

$$\overline{g_k^{[0]}}\left(\{R_j\}_{j\neq k}\right)=g_k^{[1]}\left(\{R_j\}_{j\neq k}\right),$$
$$\overline{g_k^{[1]}}\left(\{R_j\}_{j\neq k}\right)=g_k^{[0]}\left(\{R_j\}_{j\neq k}\right). \quad \text{(E.29)}$$

From cell lines to lattice partitions. Partitions of several cell lines can be combined together to construct the *lattice partitions*. From now on, the symbol $(\,\cdot_k)$ will contract the extended notation $(\{r_j\}_{j\neq k})$ of the position of a cell line along the k-th direction.

The cell line partitions defined above can be used to construct many different lattice partitions. First of all it will be shown how the ThPCA creates a lattice partition along the k-th direction at each time step t. For each cell line spanning the system along the k-th direction a random boolean $\zeta_k(\,\cdot_k, t)$ is picked. Then for the k-th direction the first lattice partition will be

$$\mathcal{G}_k(t) = \bigcup_{\{\,\cdot_k\}} g_k^{[\zeta_k(\,\cdot_k,t)]}(\,\cdot_k), \tag{E.30}$$

where $\{\,\cdot_k\}$ denotes all the cell lines along the k-th direction. The complementary cell lines will constitute the second lattice partition:

$$\overline{\mathcal{G}_k}(t) = \bigcup_{\{\,\cdot_k\}} \overline{g_k^{[\zeta_k(\,\cdot_k,t)]}}(\,\cdot_k). \tag{E.31}$$

The same thing is done for all the d lattice directions.

Next the sequence of directions is determined: at time t a permutation is randomly chosen, which is denoted $\mathbf{I}(t)$, among the $d!$ possible permutations of the elements of the array $(1,\ldots,d)$. The j-th element of $\mathbf{I}(t)$ is denoted as $I_j(t)$.

The following notation shall be used:

- $\mathcal{P}_k^t$ (and $\overline{\mathcal{P}_k^t}$) will indicate the propagation operator applied to the partition $\mathcal{G}_k(t)$ (and $\overline{\mathcal{G}_k}(t)$) of the lattice configuration which is the output of the previous lattice operation.

- the time associated to each propagation substep will be denoted as t_j (and $\overline{t_j}$), given by:

$$\begin{aligned} t_j &= t + \tau^R + (2j-2)\tau_*^P, \\ \overline{t_j} &= t + \tau^R + (2j-1)\tau_*^P. \end{aligned} \tag{E.32}$$

Then, the propagation at time t among the cells (assumed in the post-randomization state) of the entire lattice $\mathscr{L}$ is given by the following sequence:

1) Independent propagation on the pairs of cells of the partition $\mathcal{G}_{I_1}(t)$ along the direction I_1, where the input state of the cells is their respective post-randomization state. That is,

$$\mathcal{P}_{I_1}^t : \boldsymbol{\eta}(\mathscr{L}, t_1) \to \boldsymbol{\eta}(\mathscr{L}, \overline{t_1}) \tag{E.33}$$

2) Independent propagation on the pairs of cells of the complementary partition $\overline{\mathcal{G}_{I_1}}(t)$ along the direction I_1, which will use the output state of the previous step as input state. That is,

$$\overline{\mathcal{P}^t_{I_1}} : \boldsymbol{\eta}(\mathscr{L}, \overline{t_1}) \rightarrow \boldsymbol{\eta}(\mathscr{L}, t_2) \tag{E.34}$$

...

$2d-1$) Independent propagation on the pairs of cells of the partition $\mathcal{G}_{I_d}(t)$ along the direction I_d, which will use the output state of the previous step as input state. That is,

$$\mathcal{P}^t_{I_d} : \boldsymbol{\eta}(\mathscr{L}, t_d) \rightarrow \boldsymbol{\eta}(\mathscr{L}, \overline{t_d}) \tag{E.35}$$

$2d$) Independent propagation on the pairs of cells of the complementary partition $\overline{\mathcal{G}_{I_d}}(t)$ along the direction I_d, which will use the output state of the previous step as input state. That is,

$$\overline{\mathcal{P}^t_{I_d}} : \boldsymbol{\eta}(\mathscr{L}, \overline{t_d}) \rightarrow \boldsymbol{\eta}(\mathscr{L}, t+\tau) \tag{E.36}$$

Each substep of the propagation can be expressed through the following general form:

$$\begin{aligned} &\mathcal{P}^t_{I_j} : \boldsymbol{\eta}(\mathscr{L}, t_j) \rightarrow \boldsymbol{\eta}(\mathscr{L}, \overline{t_j}), \\ &\overline{\mathcal{P}^t_{I_j}} : \boldsymbol{\eta}(\mathscr{L}, \overline{t_j}) \rightarrow \boldsymbol{\eta}(\mathscr{L}, t_{j+1}), \end{aligned} \tag{E.37}$$

where $j = 1, \ldots, d$ and it is assumed that $t_{d+1} \equiv t + \tau$.

The sequence of operations can be summarized by the expression

$$\begin{aligned} \mathcal{P}^t_{\mathcal{L}} \left(\boldsymbol{\eta}^R(\mathscr{L}, t)\right) &= \overline{\mathcal{P}^t_{I_d}} \circ \mathcal{P}^t_{I_d} \circ \overline{\mathcal{P}^t_{I_{d-1}}} \circ \mathcal{P}^t_{I_{d-1}} \circ \cdots \circ \overline{\mathcal{P}^t_{I_1}} \circ \mathcal{P}^t_{I_1} \left(\boldsymbol{\eta}^R(\mathscr{L}, t)\right) \\ &= \overset{d-1}{\underset{j=0}{\circ}} \left(\overline{\mathcal{P}^t_{I_{d-j}}} \circ \mathcal{P}^t_{I_{d-j}}\right) \left(\boldsymbol{\eta}^R(\mathscr{L}, t)\right), \end{aligned} \tag{E.38}$$

where

$$\mathcal{P}^t_{I_j} \left(\boldsymbol{\eta}(\mathscr{L}, t_j)\right) = \bigcup_{\left\{\mathbf{r}, \mathbf{r}^{I_j(t)}\right\} \in \mathcal{G}_{I_j}(t)} \mathcal{P}\left(\left\{\boldsymbol{\eta}(\mathbf{r}, t_j), \boldsymbol{\eta}(\mathbf{r}^{I_j(t)}, t_j)\right\}\right), \tag{E.39}$$

for the cells belonging to the partition $\mathcal{G}_{I_j}(t)$, and

$$\overline{\mathcal{P}^t_{I_j}} \left(\boldsymbol{\eta}(\mathscr{L}, \overline{t_j})\right) = \bigcup_{\left\{\mathbf{r}, \mathbf{r}^{I_j(t)}\right\} \in \overline{\mathcal{G}_{I_j}}(t)} \mathcal{P}\left(\left\{\boldsymbol{\eta}(\mathbf{r}, \overline{t_j}), \boldsymbol{\eta}(\mathbf{r}^{I_j(t)}, \overline{t_j})\right\}\right), \tag{E.40}$$

for the cells belonging to the partition $\overline{\mathcal{G}_{I_j}}(t)$. Eq. (E.38) can be inserted into Eq. (3.1) to obtain the complete evolution equation of the model.

In this Appendix the time-dependent nature of the lattice propagation operator $\mathcal{P}^t_{\mathcal{L}}$ has been shown. It is due to the fact that it is composed of partial propagation operators working on different partitions of the lattice *which are time-dependent*. Therefore, the definition of *partitioning cellular automaton* is the most appropriate for this model since the lattice propagation is performed through an automatic partitioning of the lattice space from time to time.

E.4 Global detailed balance for propagation

At each instant of time, the *propagation partition function* of the lattice partitions $\mathcal{G}_{I_j}(t)$ and $\overline{\mathcal{G}_{I_j}}(t)$ can be written respectively as:

$$Z^P_{I_j}(t) = \prod_{\left\{\mathbf{r},\mathbf{r}^{I_j(t)}\right\}\in\mathcal{G}_{I_j}(t)} \prod_{\boldsymbol{\rho}=\mathbf{r},\mathbf{r}^{I_j(t)}} \prod_{\alpha=\mathrm{ex,in}} \left\{ \left[q_\alpha\big(n(\boldsymbol{\rho},t_j)\big)\right]^{n_\alpha(\boldsymbol{\rho},t_j)} + \left[q_\alpha\big(n^P(\boldsymbol{\rho},t_j)\big)\right]^{n^P_\alpha(\boldsymbol{\rho},t_j)} \right\}, \quad \text{(E.41)}$$

$$\overline{Z^P_{I_j}}(t) = \prod_{\left\{\mathbf{r},\mathbf{r}^{I_j(t)}\right\}\in\overline{\mathcal{G}_{I_j}}(t)} \prod_{\boldsymbol{\rho}=\mathbf{r},\mathbf{r}^{I_j(t)}} \prod_{\alpha=\mathrm{ex,in}} \left\{ \left[q_\alpha\big(n(\boldsymbol{\rho},\overline{t_j})\big)\right]^{n_\alpha\left(\boldsymbol{\rho},\overline{t_j}\right)} + \left[q_\alpha\big(n^P(\boldsymbol{\rho},\overline{t_j})\big)\right]^{n^P_\alpha\left(\boldsymbol{\rho},\overline{t_j}\right)} \right\}. \quad \text{(E.42)}$$

Similarly to the case of lattice randomization, (see Eq. (E.14)) the propagation partition function along each lattice partition is the product of independent factors.

First of all, proof of the detailed balance during propagation on a pair of neighboring cells (denoted $\{\mathbf{r},\mathbf{r}^j\}$) shall be given. The proability of the input configuration of the pair to be $\boldsymbol{\eta}(\mathbf{r},\mathbf{r}^j)$ can be expressed as the product between the probability to have input occupancy $n(\mathbf{r},\mathbf{r}^j)$ times the conditional probability to be in the input configuration given the value of the input occupancy:

$$p\left(\boldsymbol{\eta}(\mathbf{r},\mathbf{r}^j)\right) = p\left(n(\mathbf{r},\mathbf{r}^j)\right)\, p\left(\boldsymbol{\eta}(\mathbf{r},\mathbf{r}^j)|n(\mathbf{r},\mathbf{r}^j)\right). \quad \text{(E.43)}$$

Since the sum of the two occupancies, $n(\mathbf{r}, \mathbf{r}^j)$ is conserved during propagation, the detailed balance equation

$$p\left(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)\right)\, p^P\big(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j) \to \boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j)\big) = p\left(\boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j)\right)\, p^P\big(\boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j) \to \boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)\big) \quad \text{(E.44)}$$

with p^P given in Eq. (3.22), reduces to

$$p\left(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)|n(\mathbf{r}, \mathbf{r}^j)\right)\, p^P\big(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j) \to \boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j)\big) = p\left(\boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j)|n(\mathbf{r}, \mathbf{r}^j)\right)\, p^P\big(\boldsymbol{\eta}^P(\mathbf{r}, \mathbf{r}^j) \to \boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)\big) \quad \text{(E.45)}$$

As for the randomization, such reduction allows to work with an equation whose terms are easily computable. The probability of a cell pair configuration is given by:

$$p\left(\boldsymbol{\eta}(\mathbf{r}, \mathbf{r}^j)|n(\mathbf{r}, \mathbf{r}^j)\right) = \frac{q(\mathbf{r}, \mathbf{r}^j)}{Q(\mathbf{r}, \mathbf{r}^j)}, \quad \text{(E.46)}$$

where $Q(\mathbf{r}, \mathbf{r}^j)$ is the partition function of the pair, given by

$$Q\big(\mathbf{r}, \mathbf{r}^j\big) = \sum_{n=0}^{n(\mathbf{r}, \mathbf{r}^j)} Q(n)\, Q\big(n(\mathbf{r}, \mathbf{r}^j) - n\big)\,, \quad \text{(E.47)}$$

while the propagation probability is given by Eq. (3.22). Inserting Eqs. (3.22) and (E.46) into Eq. (E.45) one obtains that the detailed balance is satisfied during the propagation operation on each pair of cells. Following the same procedure as in Appendix E.2 for the randomization, it is straightforward to obtain that the detailed balance is satisfied during propagation on each lattice partition. Therefore, at each time step the entire propagation process satisfies detailed balance.

Bibliography

[1] E. van Steen, L. H. Callanan, and M. Claeys, editors, *Proceedings of the 14th International Zeolite Conference*, 2004.

[2] N. C. Jeong, H. S. Kim, and K. B. Yoon, *J. Phys. Chem. C* **111**, 10298 (2007).

[3] J. Klafter and J. M. D. eds., *Molecular Dynamics in Restricted Geometries*, John Wiley and Sons, New York, first edition, 1989.

[4] A. Huwe, F. Kremer, P. Beherens, and W. Schwieger, *Phys. Rev. Lett.* **82**, 2338 (1999).

[5] Fluid Transport in Nanoporous Materials: Proceedings of the NATO Advanced Study Institute, held in La Colle sur Loup, France, 16-28 June 2003, in *NATO Science Series II: Mathematics, Physics and Chemistry, vol. 219*, edited by W. Conner and J. Fraissard, pp. 125–238, Springer: Berlin, Germany, 2006, New York, NY, USA, 1993.

[6] C. Tunca and D. Ford, *J. Chem. Phys.* **111**, 2751 (1999).

[7] C. Saravanan, F. Jousse, and S. M. Auerbach, *Phys. Rev. Lett.* **80**, 5754 (1998).

[8] M.-O. Coppens, A. T. Bell, and A. K. Chakraborty, *Chem. Eng. Sci.* **53**, 2053 (1998).

[9] M.-O. Coppens and V. Iyengar, *Nanotechnology* **16**, S442 (2005).

[10] C. J. Jameson, A. K. Jameson, R. E. Gerald II, and A. C. de Dios, *J. Chem. Phys.* **96**, 1676 (1992).

[11] F. Trudu, G. Tabacchi, A. Gamba, and E. Fois, *J. Phys. Chem. A* **111**, 11626 (2007).

[12] C. Baerlocher, L. B. McCusker, and D. H. Olson, *Atlas of Zeolite Framework Types*, Elsevier, Amsterdam, sixth edition, 2007.

[13] M. M. J. Treacy and J. B. Higgins, *Collection of Simulated XRD Powder Patterns for Zeolites*, Elsevier, Amsterdam, fifth edition, 2007.

[14] B. Smit and R. Krishna, *Current Opinion in Solid State and Materials Science* **5**, 455 (2001).

[15] P. Demontis and G. B. Suffritti, *Chem. Rev.* **97**, 2485 (1997).

[16] B. Smit and J. I. Siepmann, *Science* **264**, 1118 (1994).

[17] S. M. Auerbach, F. Jousse, and D. P. Vercauteren, Dynamics of Sorbed Molecules in Zeolites, in *Computer Modelling of Microporous and Mesoporous Materials*, edited by C. R. A. Catlow, R. A. van Santen, and B. Smit, pp. 49–108, Elsevier, Amsterdam, 2004.

[18] E. Fois, A. Gamba, C. Medici, and G. Tabacchi, *ChemPhysChem* **6**, 1927 (2005).

[19] P. Demontis, J. Gulín-González, and G. B. Suffritti, *J. Phys. Chem. B* **110**, 7513 (2006).

[20] C. Alba-Simionesco, B. Coasne, G. Dosseh, G. Dudziak, K. E. Gubbins, R. Radhakrishnan, and M. Sliwinska-Bartkowiak, *J. Phys.: Condens. Matter* **18**, R15 (2006).

[21] T. T. P. Cheung, *J. Phys. Chem.* **97**, 8993 (1993).

[22] K. G. Ayappa, *J. Chem. Phys.* **111**, 4736 (1999).

[23] K. G. Ayappa, C. R. Kamala, and T. A. Abinandanan, *J. Chem. Phys.* **110**, 8714 (1999).

[24] J. Güémez, S. Velasco, and A. Calvo Hernández, *Physica A* **152**, 226 (1988).

[25] J. Güémez, S. Velasco, and A. Calvo Hernández, *Physica A* **152**, 243 (1988).

[26] M. W. Deem, *AICHE J.* **44**, 2569 (1998).

[27] M.-O. Coppens, A. T. Bell, and A. K. Chakraborty, *Chem. Eng. Sci.* **54**, 3455 (1999).

[28] S. Y. Bhide and S. Yashonath, *J. Chem. Phys* **111**, 1658 (1999).

[29] S. Y. Bhide and S. Yashonath, *J. Phys. Chem. B* **104**, 2607 (2000).

[30] S. M. Auerbach, *Int. Rev. Phys. Chem.* **19**, 155 (2000).

[31] R. Krishna, D. Paschek, and R. Baur, *Microporous and Mesoporous Materials* **76**, 233 (2004).

[32] D. Dubbeldam, E. Beerdsen, T. J. H. Vlugt, and B. Smit, *J. Chem. Phys.* **122**, 224712 (2005).

[33] C. Tunca and D. Ford, *Chem. Eng. Sci.* **58**, 3373 (2003).

[34] J. P. Boon, D. Dab, R. Kapral, and A. T. Lawniczak, *Phys. Rep.* **273**, 55 (1996).

[35] T. Toffoli and N. Margolus, *Cellular Automata Machines: A New Environment for Modeling*, MIT Press, Cambridge, Massachusetts, first edition, 1997.

[36] B. Chopard and M. Droz, *Cellular Automata Modeling of Physical Systems*, Cambridge University Press, Cambridge, England, first edition, 1998.

[37] Chapter III. Lattice Gases, in *Proceedings of the NATO ASI Summer School on: Microscopic simulations of complex hydrodynamic phenomena, held July 15-27, 1991 in Alghero, Sardinia*, edited by M. Marechal and B. L. Holian, pp. 125–238, Plenum Press, New York, NY, USA, 1993.

[38] J.-P. Rivet and J. P. Boon, *Lattice Gas Hydrodynamics*, Cambridge University Press, Cambridge, England, first edition, 2001.

[39] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models*, Cambridge University Press, Cambridge, England, first edition, 1999.

[40] J. Kärger and D. M. Ruthven, *Diffusion in Zeolites and Other Microporous Materials*, John Wiley and Sons, New York, first edition, 1992.

[41] P. Demontis, L. Fenu, and G. B. Suffritti, *J. Phys. Chem. B* **109**, 18081 (2005).

[42] R. S. Drago, C. E. Webster, and J. M. McGilvray, *J. Am. Chem. Soc.* **120**, 538 (1998).

[43] W. Zhu, F. Kapteijn, and J. A. Moulijn, *Phys. Chem. Chem. Phys.* **2**, 1989 (2000).

[44] A. J. Skoulidas and D. S. Sholl, *J. Phys. Chem. A* **107**, 10132 (2003).

[45] P. Demontis, F. G. Pazzona, and G. B. Suffritti, *J. Phys. Chem. B* **110**, 13554 (2006).

[46] Z. Chvoj, H. Conrad, and V. Cháb, *Surf. Sci.* **376**, 205 (1997).

[47] A. A. Tarasenko and L. Jastrabík, *Surf. Sci.* **507-510**, 108 (2002).

[48] A. N. Fitch, H. Jobic, and A. J. Renouprez, *J. Phys. Chem.* **90**, 1311 (1986).

[49] P. Demontis, S. Yashonath, and M. L. Klein, *J. Chem. Phys.* **93**, 5016 (1989).

[50] M. Gardner, *Scientific American* **220**, 120 (1970).

[51] S. Wolfram, *Rev. Modern Physics* **55**, 601 (1983).

[52] S. Wolfram, *A New Kind of Science*, Wolfram Media, Inc., 2002.

[53] U. Frisch, B. Hasslacher, and Y. Pomeau, *Phys. Rev. Lett* **56**, 1505 (1986).

[54] P. Demontis, F. G. Pazzona, and G. B. Suffritti, *J. Chem. Phys.* **126**, 194709 (2007).

[55] P. Demontis, F. G. Pazzona, and G. B. Suffritti, *J. Chem. Phys.* **126**, 194710 (2007).

[56] P. Demontis, F. G. Pazzona, and G. B. Suffritti, *J. Phys. Chem. B* **112**, 12444 (2008).

[57] C. J. Jameson, A. K. Jameson, R. E. Gerald II, and A. C. de Dios, *J. Chem. Phys.* **96**, 1690 (1992).

[58] J. Güémez and S. Velasco, *Am. J. Phys* **55**, 154 (1987).

[59] S. Fritzsche, R. Haberlandt, J. Kärger, H. Pfeifer, and M. Wolfsberg, *Chem .Phys. Lett.* **171**, 109 (1990).

[60] S. Fritzsche, R. Haberlandt, J. Kärger, H. Pfeifer, and K. Heinzinger, *Chem .Phys. Lett.* **198**, 283 (1992).

[61] S. Fritzsche, R. Haberlandt, J. Kärger, H. Pfeifer, and K. Heinzinger, *Chem. Phys.* **174**, 229 (1993).

[62] S. Fritzsche, R. Haberlandt, J. Kärger, H. Pfeifer, K. Heinzinger, and M. Wolfsberg, *Chem. Phys. Lett.* **242**, 361 (1995).

[63] P. Demontis and G. B. Suffritti, *J. Phys. Chem. B* **101**, 5789 (1997).

[64] S. Fritzsche, R. Haberlandt, G. Hofmann, J. Kärger, H. Pfeifer, K. Heinzinger, and M. Wolfsberg, *Chem. Phys. Lett.* **265**, 253 (1997).

[65] S. Fritzsche, M. Wolfsberg, R. Haberlandt, P. Demontis, G. B. Suffritti, and A. Tilocca, *Chem. Phys. Lett.* **296**, 253 (1998).

[66] S. Fritzsche, R. Haberlandt, and M. Wolfsberg, *Chem. Phys.* **253**, 283 (2000).

[67] R. Krishna and R. Baur, *Separation and Purification Technology* **33**, 213 (2003).

[68] E. Beerdsen, D. Dubbeldam, and B. Smit, *Phys. Rev. Lett.* **93**, 248301 (2004).

[69] E. Beerdsen, D. Dubbeldam, and B. Smit, *Phys. Rev. Lett.* **96**, 044501 (2006).

[70] D. A. McQuarrie, *Statistical Mechanics*, Harper and Row, New York, first edition, 1976.

[71] D. M. Ruthven, *Nature, Phys. ScL* **232**, 70 (1971).

[72] T. Toffoli and N. Margolus, *Physica D* **45**, 229 (1990).

[73] J. S. Rowlinson and G. B. Woods, *Physica A* **164**, 117 (1990).

[74] C. J. Jameson, *Ann. Rev. Phys. Chem.* **47**, 135 (1996).

[75] M. Hunger and E. Brunner, Characterization I-NMR Spectroscopy, in *Molecular Sieves - Science and Technology, vol. 4*, edited by D. Stauffer, pp. 201–293, Springer-Verlag, Heidelberg, 2004.

[76] C. J. Jameson, A. K. Jameson, B. I. Baello, and H.-M. Lim, *J. Chem. Phys.* **100**, 5965 (1994).

[77] A. K. Jameson, C. J. Jameson, and R. E. Gerald II, *J. Chem. Phys.* **101**, 1775 (1994).

[78] C. Uebing and R. Gomer, *J. Chem. Phys.* **95**, 7626 (1991).

[79] D. Chandler, *Introduction to Modern Statistical Mechanics*, Oxford University Press, New York, first edition, 1987.

[80] Z. Chvoj, H. Conrad, V. Cháb, M. Ondrejcek, and A. M. Bradshaw, *Surf. Sci.* **329**, 121 (1995).

[81] A. A. Tarasenko, Z. Chvoj, L. Jastrabík, F. Nieto, and C. Uebing, *Phys. Rev. B* **63**, 165423 (2001).

[82] N. A. Tarasenko, A. A. Tarasenko, Z. Bryknar, and L. Jastrabík, *Surf. Sci.* **562**, 22 (2004).

[83] D. M. Ruthven, *Principles of Adsorption and Adsorption Processes*, John Wiley and Sons, New York, first edition, 1984.

[84] D. M. Ruthven, *Fundamentals of Adsorption Equilibrium and Kinetics in Microporous Solids*, volume 7, Springer-Verlag Berlin, Heidelberg, 2008.

[85] R. Gomer, *Rep. Prog. Phys.* **53**, 917 (1990).

[86] C. J. Jameson, *Molec. Phys.* **102**, 723 (2004).

[87] E. Beerdsen, D. Dubbeldam, and B. Smit, *J. Phys. Chem. B* **110**, 22754 (2006).

[88] D. M. Ruthven and R. I. Derrah, *J. Colloid Interface Sci.* **52**, 397 (1975).

[89] L. F. Gladden, J. A. Sousa-Goncalves, and P. Alexander, *J. Phys. Chem. B* **101**, 10121 (1997).

[90] W. T. Lim, M. Park, and N. H. Heo, *Bull. Korean. Chem. Soc.* **21**, 75 (2000).

[91] C. J. Jameson, A. K. Jameson, H.-M. Lim, and B. I. Baello, *J. Chem. Phys.* **100**, 5977 (1994).

[92] D. Keffer, A. V. McCormick, and H. T. Davis, *J. Phys. Chem.* **100**, 967 (1995).

[93] F.-Y. Li and S. Berry, *J. Phys. Chem.* **99**, 2459 (1995).