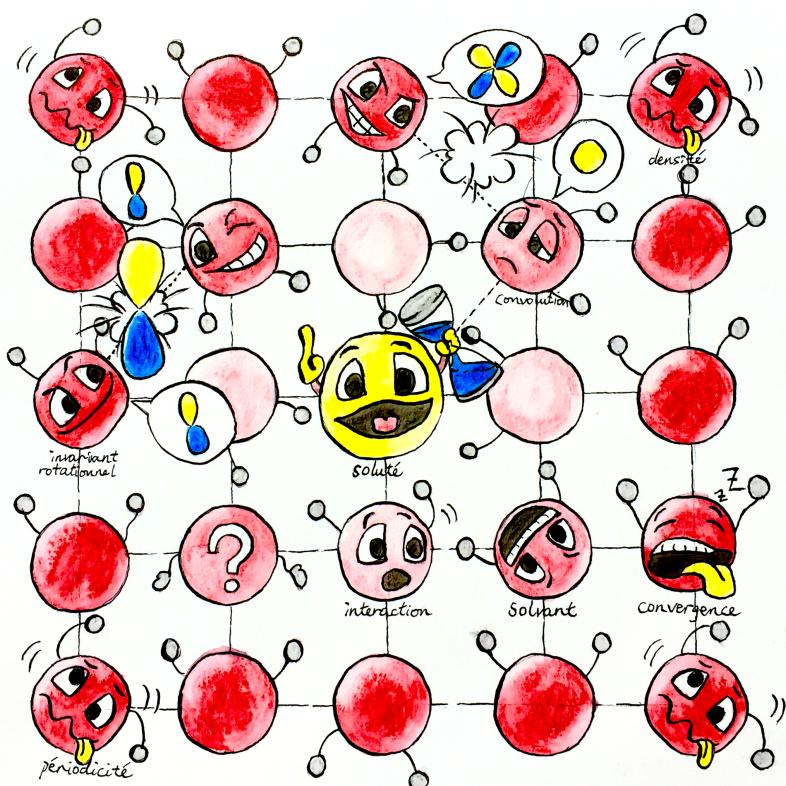


MOLECULAR DENSITY FUNCTIONAL THEORY UNDER HOMOGENEOUS REFERENCE FLUID APPROXIMATION

LU DING

Under the direction of
DANIEL BORGIS, LUC BELLONI & MAXIMILIEN LEVESQUE



SYSTEMATIC PREDICTION OF SOLVATION PROPERTIES
WITH MOLECULAR-SCALE LIQUID THEORY

*When you are studying any matter, or considering any philosophy, ask yourself only,
what are the facts and what is the truth that the facts bear out.
Never let yourself be diverted either by what you wish to believe, or by what you think
would have beneficent social effects if it were believed.
But look only, and solely, at what are the facts.*

— Bertrand Russell

ACKNOWLEDGEMENTS

I would love to express my most respectful gratitude to my thesis advisors Daniel Borgis and Luc Belloni, who developed the main theory used in this thesis, and whose great knowledge of liquid theory as well as the genius way of thinking and explaining have given me a solid guide on doing this research. I would also love to thank them for their being engaged in the endless work of correcting this manuscript.

I'm also grateful to Maximilien Levesque, who also belong to this project and joined the supervision of my thesis. Thanks for sharing some results and data that is useful to this thesis.

I wish to acknowledge all the great heads having accepted to evaluate and give ideas about my work: (**noms des jurés**), thank you to be part of the jury of this thesis.

For a master student issue from pure chemistry speciality, a lack of knowledge about informatics brought me a lot of difficulties. I would like to express my sincere gratitude to my colleagues Pierre Kestener, Matthieu Haefele, Yacine Ould-Rouis, and Martial Mancip for huge aide in informatics and very useful advices during this thesis.

This thesis undergoes at *Maison de la Simulation, CEA Saclay*, financially supported by Bourse Idex-CEA. I'm grateful to all the organizations and staffs that gave me the chance to have this three-year experience.

I am also grateful to ... for the help in correcting huge amount of grammar faults in the manuscript.

I'm deeply indebted to my tutors of license and master, Monsieur Hongwei Tan and Madame Michelle Gupta, whose clear logic and warm encouragement gave me all the necessities to be in loved with theoretical chemistry.

Finally I would like to thank my father, who made the right decision to send me here in France and support me in every aspect.

CONTENTS

1	INTRODUCTION	1
1.1	Simulation of solvent effects,	1
1.2	Scope of this thesis,	3
I	STATE OF THE ART: SOLVATION, MODELS AND METHODS	5
II	APPENDIX	7
A	BASIC CONCEPTS ABOUT COMPUTING PERFORMANCE	9
A.1	Algorithm complexity,	9
A.2	Roofline model and memory delay,	10
A.3	Scalability of parallelized code,	11
A.4	Profiling and tracing toolkits,	12
	BIBLIOGRAPHY	13

LIST OF FIGURES

Figure 1.1	The solvation process,	2
Figure A.1	Function growth,	9
Figure A.2	The roofline model and performance pattern,	10
Figure A.3	Memory usage in hardware level,	11

LIST OF TABLES

Table 1.1	Theories of solvation simulation,	3
-----------	-----------------------------------	---

LISTINGS

NOTATIONS

$\mathcal{F}[\rho]$	solvation free energy functional
\mathcal{F}_{exc}	excess free energy functional
$\rho(\mathbf{r}, \Omega)$	density of solvent
$\gamma(\mathbf{r}, \Omega)$	gradient of excess free energy functional,

ACRONYMS

DCF direct correlation function

DFT discret Fourier transform, also refer to density functional theory

HRF homogeneous reference fluid (approximation)

- IET integral equation theory
MC Monte Carlo
MD molecular dynamics
MDFT molecular density functional theory
OZ Ornstein-Zernike (equation)
QM quantum mechanics
RISM reference interaction site model

INTRODUCTION

This thesis is about to develop an original numerical toolkit for physical chemists and structural biologists, based on the molecular density functional theory (MDFT), which makes it possible to predict efficiently, and with a microscopic accuracy, the solvation properties of arbitrary molecular objects in arbitrary molecular solvents (mainly water). This introduction will help to understand the objective of this thesis, it explains why people are interested in the nature of solvation, and where we are in the computing trends in solvation simulation.

1.1 SIMULATION OF SOLVENT EFFECTS

Solvation is a fundamental phenomenon in chemistry. The chemical behavior of lots of systems has a strong dependence on the nature of solvent. For example, for some popular issue as metal-organic reacting center [7, 21], or pharmaceutical etudes [16, 22, 23]. The solvation properties required by etudes are very variable, such as the Gibbs free energy of solvation, solubility, partition coefficient, saturated vapor pressure, pH value, as well as the 3D solvation structure, etc. Overall, the interest of these solvation properties comes from many domains, such as chemistry, biochemistry, pharmaceutics, medicine, environmental and agrochemical industries. Unlike the well-studied quantum mechanics (QM) for chemical interaction and macroscopic finite element model for physical process, the theories of solvation are very variable and still under developing, owing to the ambiguous compromise between the accuracy and the computing cost. In a word, the studies in this domain are quite important and vibrant.

To change a phenomenon to a model, we must understand its process. Solvation is defined as the process of moving a molecule from the gas phase (or vacuum) to a condensed phase (figure 1.1), which builds a stabilizing interaction with the solute (or solute moiety like protein residues, interfaces, etc.) [19]. Such interactions are mostly classical, involving electrostatic forces and van der Waals forces, with also chemically more specific effects such as hydrogen bond formation, and quantic effects for some small solvents whose vibrational or rotational energy states is at the same magnitude as $k_B T$, etc. [11].

As not all kinds of interactions is important in applications, according to the usage, different models and methods are developed.

In the most of the 20th century [4], the study of solvation effects has been dominated by continuum (implicit) models, which is simply depending on the dielectric constants and not costly on computation resource. They provide an accurate way to treat the strong, long-range electrostatic interactions which dominate many solvation phenomena, but lack of detail informations in the first solvation shell. The later, which mainly includes the cavity formation energy and solute-solvent van der Waals interactions, is often rudely treat by introducing an artificial form of cavity, that links to the form of solute. And the methods for electrostatic interactions involves like generalized Born model, or through better estimates via Poisson-Boltzmann calculations. They are widely integrated within QM simulations of the solvent, by add extra solvation terms onto the Fock or Kohn-Sham operator [10, 17, 25]. However, the improper treatment of the first-shell, where the microscopic interactions are primarily located, often introduce sometimes huge error

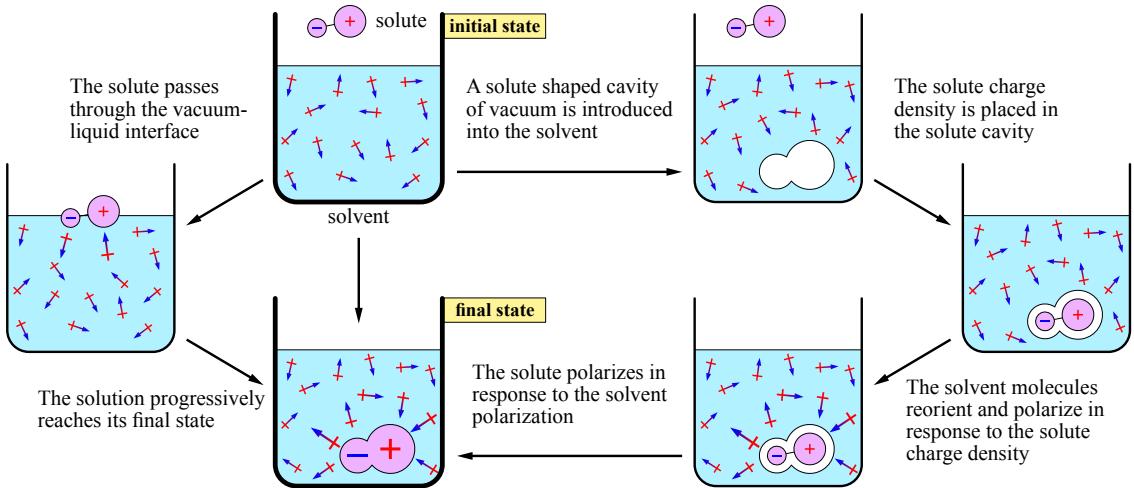


Figure 1.1: The solvation process. To a thermodynamic system, whose properties only depends on the initial and final states, it can go through different paths. The physical process of solvation (left path) takes the solute from vacuum into bulk solvent, progressively passing through the vacuum-liquid interface. Theoretically, the solvation energy is defined as the energy consumed in such a progress. In theoretical studies, the process can be decomposed to some artificial unphysical process (right path), involving the growth of an uncharged solute-sized cavity within the bulk solvent, the transfer of the solute charge distribution from vacuum into the cavity, and the interaction between the solute and solvent, etc.

in free energy evaluation, especially for polar solvents (like water), despite the accuracy that the QM calculation alone can achieve. Therefore, classical molecular simulations, which describing the individual solvent molecules (explicit), particularly the molecular dynamics (MD) and Monte Carlo method (MC), became the alternative solution during the last few decades. They generate trajectories and configurations, then estimate free energy changes by statistical mechanic technics, such as free energy perturbation (FEP) theory or thermodynamic integration (TI) [18]. These calculation is very demanding on computing cost, due to the requirement of many (hundreds or thousands of) solvent molecules to form a realistic model.

Recently, a third domain of theory to describe solvent, based on the statistical mechanics of fluid, is growing rapidly. It is generally called liquid theory, involves mainly the integral equation theory (IET), and the classical density functional theory for liquids. These approaches are cope to give the molecular nature of the first-shell, but without calculate all the instantaneous micro-states with respect to time, which can be integrated over positions and momentums theoretically. Therefore, they are of magnitudes faster than those simulations by micro-states.

The integral equation theory (IET) is about solving the Ornstein-Zernike (OZ) equation with a specific closure equation [11, 14]. It was firstly limited to so called “simple liquid” - a system of spherical particles. A part, Chandler and Andersen in 1971 [3] developed the reference interaction site model (RISM), which discretizes the distribution and correlation functions into a site-site set of functions, and solve the OZ equation in matrix [15]. Another part, Blum [1, 2], Fries and Patey [8] extend the OZ equation to molecular case, where the distribution and correlation functions depend on both position and orientation. In their theory, the orientation part of OZ equation is simplified by expending the distribution and correlation functions on Wigner generalized spherical harmonics.

The classical density functional theory approach deal with inhomogeneous liquids, which uses the same variation principle and minimization strategy [6, 13, 20] as electronic

density functional theory DFT that treats electric interactions and has a great success in computational chemistry. It gives the Helmholtz free energy and the equilibrium solvent density, by minimizing the free energy functional of the solvent density in the presence of a given external potential. Borgis and collaborators [too many ref] have recently generalized it into molecular case, named molecular density functional theory (MDFT), where the solvent density depends on both position and orientation, $\rho(\mathbf{r}, \Omega)$. The main theoretical difficulty lies in the definition of well-founded and reliable functionals of the excess free energy $\mathcal{F}_{\text{exc}}[\rho]$, according to the geometric complexity of the solvent molecule. Some recent researches have shown that it is cope with linear solvents like acetonitrile, but still have little non-satisfaction with the most complex solvent, i. e. water. MDFT can be proved to be mathematically equivalent to the two-component molecular IET.

The majority of work of all these theories have been focused on water, since it is one of the most difficult systems to model due to its molecular geometry, ineligible multi-body interaction, quantum effect, hydrogen bond, etc. The importance of including instantaneous polarization in potential functions is also an issue [5, 24]. However, since polarizable force fields are not yet in common use, the simulations by micro-states and the liquid theory which feed on force field also have their own limit, compared to the continuum model which can be polarizable. The advantages and disadvantages of each branch of theory are listed in table 1.1.

THEORY	SPEED	LONG-RANGE	FIRST-SHELL	POLARIZABLE SOLVENT
Continuum model	fast	yes	no	fully
Simulation by time	costly	yes	yes	partially, very costly
Liquid theory	fast	yes	yes	partially

Table 1.1: Theories of solvation simulation

This thesis consists in the development of the MDFT, focusing on the generalization and algorithmic acceleration of the excess free energy functional \mathcal{F}_{exc} evaluation under homogenous reference fluid (HRF) approximation, which will be discussed in detail in later chapters.

1.2 SCOPE OF THIS THESIS

Chapter I reviews a selection of models and methods to the solvent effect. It includes the mainly used continuum model, the basic of liquid theory, as well as its two frontier research domains, IET and MDFT. The code structure of MDFT, which all the development in this thesis is based on, is also presented. There is also a brief introduction to MD and MC, as well as the generation of direct correlation function (DCF) used in this thesis by such methods.

Chapter II presents all the theory developed and newly used in this thesis. In this thesis, two algorithms of excess energy functional evaluation are proposed, one is extension of the previous algorithm, other is a new algorithm, that combines the molecular OZ equation treatment of angular part with MDFT. The output solvation properties is mainly the two: free energy, and solvent structure.

Chapter III takes note of all the implementation result, that divided into two aspects, the “accuracy”, which involves comparisons between algorithms, and with IET and MD results; and the “efficiency”, which evaluate the computing cost of the code, both in sequential and parallelized version.

Chapter VI gives some application to ions and molecules.

Chapter I

STATE OF THE ART: SOLVATION, MODELS AND METHODS

(Chapter introduction is to clear the motivation and relation between sections.)

This chapter is a summery of all the previous work that this thesis is based on.

In section ??, we begin by introducing the models used in our study, as well as some others of different description scale, by way of comparison to illustrate why we chose our models.

Once the model is chosen, all the theories become mathematical problems. Section ?? reviews some basic concepts of statistical mechanics for liquids (liquid theory), which present the deduction of formalisms from the model of the system, without introducing any artificial terms. The following two sections, section ?? and ?? give two frontier domains of the liquid theory: the integral equation theory (IET), and the molecular density functional theory (MDFT) that this thesis works upon. The clear mathematical equivalence between these two theories is presented in section ??, which gave us the idea to use the expansion technics in IET to serve MDFT.

Finally, a brief introduction of the simulations we used is made in section ???. Depending on micro configuration, this includes molecular dynamics (MD) and Monte Carlo (MC) simulation methods, This is done to explain how the direct correlation functions (DCF) portrayed in this thesis are obtained.

Chapter II

APPENDIX

A

BASIC CONCEPTS ABOUT COMPUTING PERFORMANCE

In addition to the theory work, the performance of the code being developed is also an important aspect of this thesis. It is essential to have a fast and accurate method. To evaluate a code in a strict and systematic way, some basic concepts of computing performance are listed here.

a.1 ALGORITHM COMPLEXITY

Algorithm complexity is a crucial criteria for sequential code. A definition is given below.

Let f and g be two real (or even complex) functions defined over the natural numbers \mathbb{N} . We write

$$f = O(g) \quad (\text{A.1})$$

if there is a constant $c > 0$ such that from certain number $n > n_0$ we always have $|f(n)| \leq c|g(n)|$. The O is also named as the big-O notation [9], or order of growth. Figure A.1 shows the growth tendency of some frequent functions; from this we can conclude the following:

$$O(1) > O(\log_2 n) > O(n) > O(n \log_2 n) > O(n^2) > O(2^n) > O(n!) \quad (\text{A.2})$$

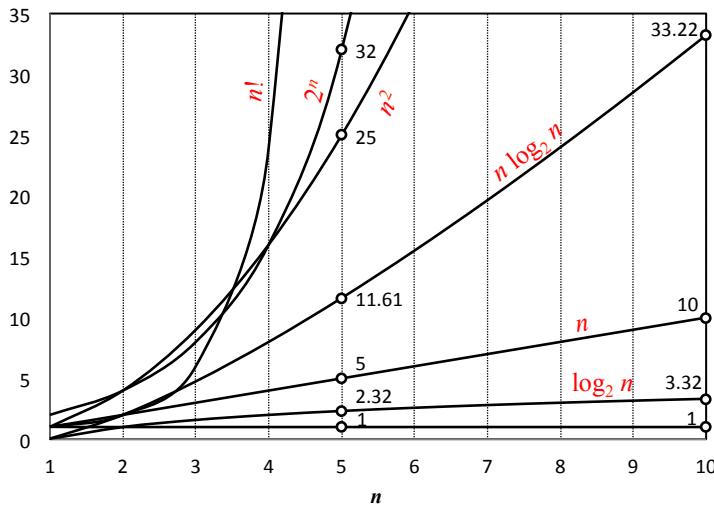


Figure A.1: Function growth

In this thesis, the big-O notation is used to measure algorithm complexity. Other notations can also be used for the same purpose, such as:

- $f = o(g)$ if $f(n)/g(n) \rightarrow 0$, $n \rightarrow \infty$
- The inverse of big-O notation $f = \Omega(g)$ if $g = O(f)$

- The notation $f = \Theta(g)$ means that both $f = O(g)$ and $g = O(f)$ hold, and we can also say they are of the same order.

In a code we always search algorithms with a lower algorithm complexity. Ideally, the implementation of code matches the model and have the same growth tendency as its complexity, but in the practical case, overheads and memory delay can also limit the performance. (part to be modified to adapt implementation results)

a.2 ROOFLINE MODEL AND MEMORY DELAY

The simplest model aiming to distinguish whether a piece of code is limited by the computing power (CPU) or the memory bandwidth (RAM to Caches) is the roofline model [26] for single loop:

$$P = \min(P_{\max}, I \cdot b_S) \quad (\text{A.3})$$

where

- P is the applicable peak performance of a loop, assuming that data comes from the level 1 cache, of unity GFlop/s.
- I is the computational intensity (“work” per byte transferred) over the slowest data path utilized, of unity Flop/Byte.
- b_S is the applicable peak bandwidth of the slowest data path utilized, of unity GByte/s.

As shown in figure A.2, the overall performance is limited by both the peak performance and the memory bandwidth. The computational intensity I depends on the code, while the other two terms in eq. (A.3) depend on hardware. The optimal use of resources occurs at the intersection point.

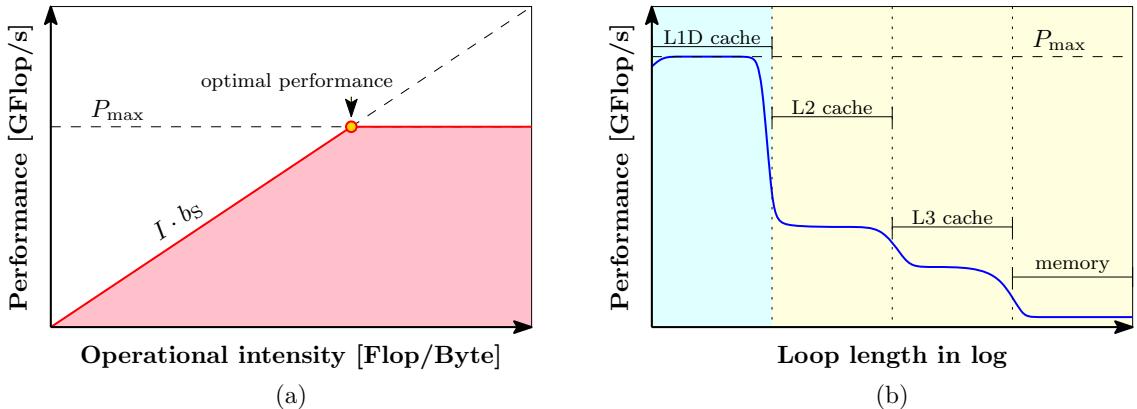


Figure A.2: The roofline model and performance pattern. (a) The roofline model. (b) Performance pattern of a simple loop with respect to the loop length in logarithm. The blue part is limited by operation execution, and the yellow part is limited by memory bottleneck.

The roofline model can give an idea of whether the diminuition of algorithm complexity is the most important optimization strategy, because it only counts the number of operations. In most of cases, avoiding slow data paths is the key to performance optimization.

As shown in figure A.3, the memory hardware has hierarchical architectures. The fastest ones are the registers included in the microprocessor, which are used for temporary storage of data, instructions and addresses required by the arithmetic logic unit (ALU) and the

control unit (CU) in CPU during execution of a program. The lowest is normally the input/output (I/O) process. The reading strategy of data (contiguous or not), as well as the size and initialized location of arrays, both play pivotal roles in the overall computing performance.

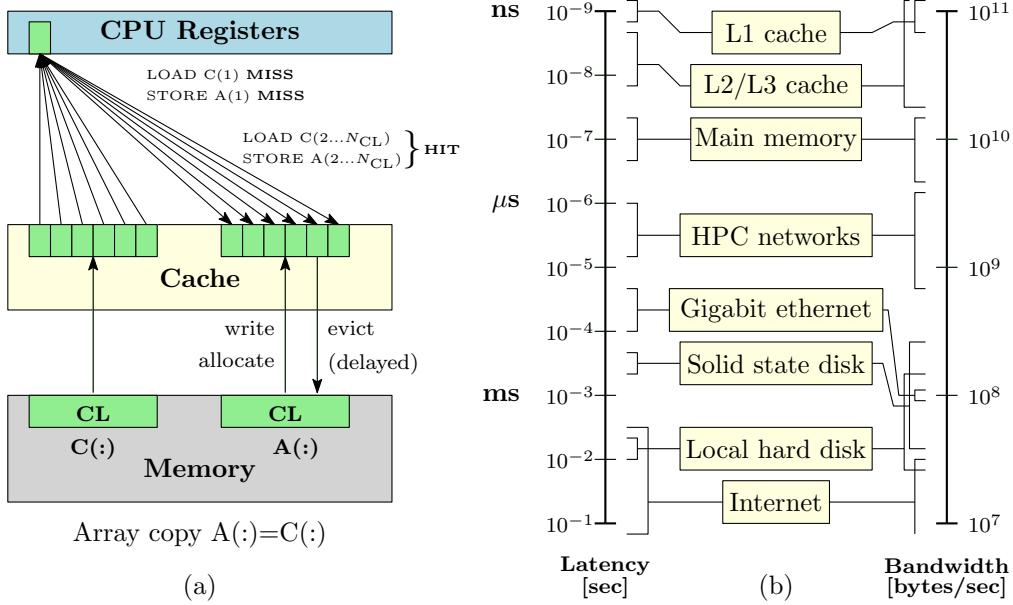


Figure A.3: Memory usage in hardware level [12]. (a) An example of array copy $A(:)=C(:)$. Caches are organized in cache lines (CL), only complete cache lines are transferred between memory hierarchy levels (except registers). HIT/MISS: Load or store instruction does/doesn't find the data in a cache level. (b) Computing latency and memory bandwidth vary by magnitude, from the fastest cache transfers to the lowest processes.

a.3 SCALABILITY OF PARALLELIZED CODE

For parallelized code, scalability is the key issue. Highly scalable codes can take advantage of numerous nodes of HPC centers, so that single core performance no longer matters.

The speed-up is defined as:

$$S(N) = \frac{t(1)}{t(N)} \quad (\text{A.4})$$

And the relative efficiency

$$E(N) = \frac{S(N)}{N} = \frac{t(1)}{Nt(N)} \quad (\text{A.5})$$

$S(N) \sim N$ or $E(N) \sim 100\%$ means the application scales. By contrast, $S(N) < N/2$ or $E(N) < 50\%$ means the application does not scale.

Amdahl's Law gives the theoretical speedup in latency of the execution of a task at fixed workload:

$$S(N) = \frac{1}{\alpha_s + \alpha_p/N} \quad (\text{A.6})$$

where α_s is the serial fraction and α_p the parallel fraction of the source code. Therefore the overall computing speed is limited by the unscalable part:

$$\lim_{N \rightarrow \infty} S(N) = \frac{1}{\alpha_s} \quad (\text{A.7})$$

making it the focus we wish to reduce.

a.4 PROFILING AND TRACING TOOLKITS

There are several types of software and toolkits for performance evaluation. They are of two categories: profiling and tracing. A trace is a collection of events or timestamps. A profile is a collection of timings. Profiling tools are usually more simple and rapid, but for subroutines that are called by a large number of times, the overhead in time measurement is not negligible.

The tool used in this thesis is mainly VTune, where application execution is interrupted every $\sim 100 \text{ ns}$ and information is stored (call stack, hardware counters, etc.). The execution time overhead is small. (To be detailed.)

BIBLIOGRAPHY

- [1] L. Blum. “Invariant Expansion. II. The Ornstein Zernike-Equation for Nonspherical Molecules and an Extended Solution to the Mean Spherical Model.” In: *The Journal of Chemical Physics* 57.5 (1972), pp. 1862–1869. DOI: <http://dx.doi.org/10.1063/1.1678503>. URL: <http://scitation.aip.org/content/aip/journal/jcp/57/5/10.1063/1.1678503> (cit. on p. 2).
- [2] L. Blum and A. J. Torruella. “Invariant Expansion for Two-Body Correlations: Thermodynamic Functions, Scattering, and the Ornstein-Zernike Equation.” In: *The Journal of Chemical Physics* 56.1 (1972), pp. 303–310. DOI: <http://dx.doi.org/10.1063/1.1676864>. URL: <http://scitation.aip.org/content/aip/journal/jcp/56/1/10.1063/1.1676864> (cit. on p. 2).
- [3] David Chandler and Hans C. Andersen. “Optimized Cluster Expansions for Classical Fluids. II. Theory of Molecular Liquids.” In: *The Journal of Chemical Physics* 57.5 (1972), pp. 1930–1937. DOI: <http://dx.doi.org/10.1063/1.1678513>. URL: <http://scitation.aip.org/content/aip/journal/jcp/57/5/10.1063/1.1678513> (cit. on p. 2).
- [4] Christopher J. Cramer and Donald G. Truhlar. “Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics.” In: *Chemical Reviews* 99.8 (1999). PMID: 11849023, pp. 2161–2200. DOI: [10.1021/cr960149m](https://doi.org/10.1021/cr960149m). eprint: <http://dx.doi.org/10.1021/cr960149m>. URL: <http://dx.doi.org/10.1021/cr960149m> (cit. on p. 1).
- [5] Liem X. Dang, Julia E. Rice, and Peter A. Kollman. “The effect of water models on the interaction of the sodium–chloride ion pair in water: Molecular dynamics simulations.” In: *The Journal of Chemical Physics* 93.10 (1990), pp. 7528–7529. DOI: <http://dx.doi.org/10.1063/1.459714>. URL: <http://scitation.aip.org/content/aip/journal/jcp/93/10/10.1063/1.459714> (cit. on p. 3).
- [6] R. Evans. “The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids.” In: *Advances in Physics* 28.2 (1979), pp. 143–200. DOI: [10.1080/00018737900101365](https://doi.org/10.1080/00018737900101365). eprint: <http://dx.doi.org/10.1080/00018737900101365>. URL: <http://dx.doi.org/10.1080/00018737900101365> (cit. on p. 2).
- [7] Sebastiao Formosinho and Monica Barroso, eds. *Proton-Coupled Electron Transfer. A Carrefour of Chemical Reactivity Traditions*. RSC Catalysis Series. The Royal Society of Chemistry, 2012, P001–157. DOI: [10.1039/9781849733168](https://doi.org/10.1039/9781849733168). URL: <http://dx.doi.org/10.1039/9781849733168> (cit. on p. 1).
- [8] P. H. Fries and G. N. Patey. “The solution of the hypernetted-chain approximation for fluids of nonspherical particles. A general method with application to dipolar hard spheres.” In: *The Journal of Chemical Physics* 82.1 (Jan. 1985), pp. 429–440. DOI: [doi:10.1063/1.448764](https://doi.org/10.1063/1.448764). URL: http://jcp.aip.org/resource/1/jcpa6/v82/i1/p429_s1 (visited on 01/15/2013) (cit. on p. 2).
- [9] Peter Gács and László Lovász. *Complexity of Algorithms*. Lecture Notes. 1999 (cit. on p. 9).
- [10] Gaussian 09 keyword: SCRF. 2014. URL: http://www.gaussian.com/g_tech/g_ur/k_scrf.htm (cit. on p. 1).

- [11] C.G. Gray and K.E. Gubbins. *Theory of Molecular Fluids: I: Fundamentals*. International Series of Monographs on Chemistry. OUP Oxford, 1984. URL: <http://books.google.de/books?id=3mz2RcnnMGwC> (cit. on pp. 1, 2).
- [12] Georg Hager and Gerhard Wellein. *Formation lecture: Node-Level Performance Engineering*. Dec. 2015. URL: <http://moodle.rrze.uni-erlangen.de/course/view.php?id=274> (cit. on p. 11).
- [13] J. P. Hansen. “Basic Statistical Theory of Liquids.” In: *The Physics and Chemistry of Aqueous Ionic Solutions*. Ed. by M.-C Bellissent-Funel and G. W. Neilson. Dordrecht: Springer Netherlands, 1987, pp. 1–59. DOI: [10.1007/978-94-009-3911-0_1](https://doi.org/10.1007/978-94-009-3911-0_1). URL: http://dx.doi.org/10.1007/978-94-009-3911-0_1 (cit. on p. 2).
- [14] Jean-Pierre Hansen and Ian R. McDonald. *Theory of Simple Liquids*. Ed. by Jean-Pierre Hansen and Ian R. McDonald. Fourth Edition. Oxford: Academic Press, 2013 (cit. on p. 2).
- [15] Fumio Hirata, ed. *Molecular Theory of Solvation*. en. Vol. 24. Understanding Chemical Reactivity. Dordrecht: Kluwer Academic Publishers, 2004. URL: <http://link.springer.com/10.1007/1-4020-2590-4> (visited on 08/10/2016) (cit. on p. 2).
- [16] Laura D. Hughes, David S. Palmer, Florian Nigsch, and John B. O. Mitchell. “Why are some properties more difficult to predict than others? A study of QSPR models of solubility, melting point, and Log P.” In: *Journal of Chemical Information and Modeling* 48.1 (Jan. 2008), pp. 220–232. DOI: [10.1021/ci700307p](https://doi.org/10.1021/ci700307p) (cit. on p. 1).
- [17] Frank Jensen. *Introduction to Computational Chemistry*. John Wiley & Sons, 2006 (cit. on p. 1).
- [18] William L. Jorgensen and Julian Tirado-Rives. “Free energies of hydration for organic molecules from Monte Carlo simulations.” In: *Perspectives in Drug Discovery and Design* 3.1 (1995), pp. 123–138. DOI: [10.1007/BF02174470](https://doi.org/10.1007/BF02174470). URL: <http://dx.doi.org/10.1007/BF02174470> (cit. on p. 2).
- [19] A. D. Mcnaught and A. Wilkinson. *IUPAC Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). WileyBlackwell; 2nd Revised edition edition (cit. on p. 1).
- [20] N. David Mermin. “Thermal Properties of the Inhomogeneous Electron Gas.” In: *Physical Review* 137.5A (Mar. 1965), A1441–A1443. DOI: [10.1103/PhysRev.137.A1441](https://doi.org/10.1103/PhysRev.137.A1441). URL: <http://link.aps.org/doi/10.1103/PhysRev.137.A1441> (visited on 02/10/2014) (cit. on p. 2).
- [21] Trenton H. Parsell, Meng-Yin Yang, and A. S. Borovik. “C-H Bond Cleavage with Reductants: Re-Investigating the Reactivity of Monomeric MnIII/IV-Oxo Complexes and the Role of Oxo Ligand Basicity.” In: *Journal of the American Chemical Society* 131.8 (2009). PMID: 19196005, pp. 2762–2763. DOI: [10.1021/ja8100825](https://doi.org/10.1021/ja8100825). eprint: <http://dx.doi.org/10.1021/ja8100825>. URL: <http://dx.doi.org/10.1021/ja8100825> (cit. on p. 1).
- [22] German L. Perlovich, Tatyana V. Volkova, and Annette Bauer-Brandl. “Towards an understanding of the molecular mechanism of solvation of drug molecules: A thermodynamic approach by crystal lattice energy, sublimation, and solubility exemplified by paracetamol, acetanilide, and phenacetin.” In: *Journal of Pharmaceutical Sciences* 95.10 (), pp. 2158–2169. DOI: [10.1002/jps.20674](https://doi.org/10.1002/jps.20674). URL: <http://dx.doi.org/10.1002/jps.20674> (cit. on p. 1).

- [23] GermanL Perlovich and Annette Bauer-Brandl. “Solvation of Drugs as a Key for Understanding Partitioning and Passive Transport Exemplified by NSAIDs.” In: *Chemical Processes with Participation of Biological and Related Compounds*. CRC Press, 2008, pp. 291–325. DOI: doi:10.1201/b12241-9. URL: <http://dx.doi.org/10.1201/b12241-9> (cit. on p. 1).
- [24] Michiel Sprik and Michael L. Klein. “A polarizable model for water using distributed charge sites.” In: *The Journal of Chemical Physics* 89.12 (1988), pp. 7556–7560. DOI: <http://dx.doi.org/10.1063/1.455722>. URL: <http://scitation.aip.org/content/aip/journal/jcp/89/12/10.1063/1.455722> (cit. on p. 3).
- [25] Jacopo Tomasi and Maurizio Persico. “Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent.” In: *Chemical Reviews* 94.7 (1994), pp. 2027–2094. DOI: 10.1021/cr00031a013. eprint: <http://dx.doi.org/10.1021/cr00031a013>. URL: <http://dx.doi.org/10.1021/cr00031a013> (cit. on p. 1).
- [26] Samuel Williams, Andrew Waterman, and David Patterson. “Roofline: An Insightful Visual Performance Model for Multicore Architectures.” In: *Commun. ACM* 52.4 (Apr. 2009), pp. 65–76. DOI: 10.1145/1498765.1498785. URL: <http://doi.acm.org/10.1145/1498765.1498785> (cit. on p. 10).