

1 **ms2: Ein Werkzeug zur Berechnung thermodynamischer**
2 **Stoffeigenschaften mittels molekularer Simulation**

3
4 Stephan Deublein¹, Bernhard Eckl², Jürgen Stoll², Sergey V. Lishchuk³,
5 Gabriela Guevara-Carrion¹, Colin W. Glass⁴, Thorsten Merker¹, Martin Bernreuther⁴,
6 Hans Hasse¹, Jadran Vrabec^{5,*}

7
8 ¹ *Dipl.-Ing. Stephan Deublein, Dipl.-Ing. Gabriela Guevara-Carrion, Dipl.-Ing. Thorsten Merker,*
9 *Prof. Dr. Ing. Hans Hasse, Lehrstuhl für Thermodynamik, Universität Kaiserslautern, 67653 Kaiserslautern*

10 ² *Dr. Ing. Bernhard Eckl, Dr. Ing Jürgen Stoll, Institut für Technische Thermodynamik und Thermische*
11 *Verfahrenstechnik, Universität Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart*

12 ³ *Dr. Ing. Sergey V. Lishchuk, Department of Mathematics, University of Leicester, University Road,*
13 *Leicester LE1 7RH, Großbritannien*

14 ⁴ *Dr. Colin W. Glass, Dr. Ing. Martin Bernreuther, Höchstleistungsrechenzentrum Stuttgart (HLRS), 70550*
15 *Stuttgart*

16 ⁵ *Prof. Dr. Ing. Jadran Vrabec, Lehrstuhl für Thermodynamik und Energietechnik, Universität Paderborn,*
17 *Warburger Str. 100, 33098 Paderborn*

19
20
21 In diesem Beitrag wird *ms2*, ein Programm zur Ermittlung thermodynamischer
22 Stoffdaten von Fluiden im Gleichgewicht mit molekularer Simulation, vorgestellt. Mit
23 *ms2* können thermodynamische Stoffdaten von Reinstoffen und Mischungen
24 bestehend aus einer Vielzahl von Komponenten ermittelt werden. Es lassen sich
25 Molekulardynamik und Monte-Carlo Simulationen in allen gängigen Ensembles
26 durchführen. Damit zugängliche Zielgrößen sind unter anderem thermische und
27 kalorische Eigenschaften, Dampf-Flüssigkeits-Gleichgewichte und Transportgrößen.
28 *ms2* ist optimiert auf kurze Antwortzeiten auf unterschiedlicher Hardware und
29 einfache Bedienbarkeit. Hilfsprogramme erleichtern die Verwendung von *ms2*.

30 **Schlagwörter:** Molekulare Simulation, Molekulardynamik, Monte-Carlo,
31 Phasengleichgewicht, Transporteigenschaften

ms2: A Molecular Simulation Tool for Thermodynamic Properties

This work presents the molecular simulation program *ms2* that is designed for the calculation of thermodynamic properties of bulk fluids in equilibrium consisting of small electro-neutral molecules. *ms2* features molecular dynamics and Monte-Carlo simulations. It supports the calculation of vapor-liquid equilibria of pure fluids and multi-component mixtures on the basis of the Grand Equilibrium method. It yields numerous thermodynamic properties by sampling various classical ensembles. *ms2* is designed to achieve short response times on different hardware and straightforward handling. Auxiliary feature tools facilitate the interpretation of input and output files.

Keywords: Molecular simulation, molecular dynamics, Monte-Carlo, vapor-liquid equilibria, transport properties

1 1 Einleitung

2 Durch die ständigen Fortschritte in der Entwicklung von Computerhardware,
3 Simulationsmethoden und molekularen Modellen (Kraftfeldern) können
4 thermodynamische Eigenschaften heute auch für komplexe Fluide mit molekularen
5 Simulationen zuverlässig ermittelt werden. So geht eine Gruppe überwiegend US-
6 amerikanischer Unternehmen sowie das National Institute of Standards and
7 Technology (NIST) davon aus, dass „die molekulare Modellierung und Simulation
8 eine Schlüsseltechnologie wird, die von der chemischen Industrie weithin anerkannt
9 und zukünftig in einen Zusammenspiel mit anderen prädiktiven Methoden den
10 zunehmenden Bedarf an Stoffdaten decken wird“ [1]. Die molekulare Modellierung
11 und Simulation basiert auf einem sicheren physikalischen Fundament und erlaubt
12 den Zugriff auf strukturelle, energetische und dynamische Eigenschaften auf der
13 molekularen Ebene - Größen, die das makroskopische Stoffverhalten festlegen. Aus
14 einem Kraftfeld lassen sich grundsätzlich alle thermodynamischen Stoffdaten
15 (statisch und dynamisch) an allen Zustandspunkten ermitteln. Die Methode ist unter
16 anderem dann interessant, wenn Experimente nicht oder nur mit hohem Aufwand
17 eingesetzt werden können, z.B. bei toxischen, explosiven oder anderweitig
18 sicherheitsrelevanten Fluiden [2]. Generell ist sie auch als Ergänzung etablierter
19 Methoden zur Berechnung thermodynamischer Stoffdaten attraktiv, da sie einen
20 unabhängigen Zugang bietet.

21 Die Modellierung von thermodynamischen Stoffdaten hat sich in den letzten Jahren
22 gewandelt. Mit modernen Methoden werden die thermodynamischen Stoffdaten nach
23 dem Bottom-Up Prinzip abgebildet, d.h. ausgehend von molekularen Eigenschaften
24 hin zu makroskopischen Größen. Beispiele sind Zustandsgleichungen, die auf der
25 „Statistical Associated Fluid Theory“ beruhen wie PC-SAFT [3], oder Ansätze für die
26 Gibbssche Exzessenergie, die auf dem „Continuum Solvation Model“ beruhen wie
27 COSMO-RS [4].

28 Die molekulare Modellierung und Simulation mit Kraftfeldmethoden folgt ebenfalls
29 dem Bottom-Up Prinzip. Es können mit dieser Methode alle thermodynamischen
30 Stoffdaten [5] von homogenen Phasen sowie Phasengleichgewichten ermittelt
31 werden. Darüber hinaus werden detaillierte Einblicke in die Vorgänge auf der
32 molekularen Ebene möglich. Basis sind dabei Kraftfelder, die die molekularen
33 Wechselwirkungen zwischen den Molekülen beschreiben. Deren Parameter werden
34 heute an quantenmechanische Rechnungen und einige wenige experimentelle
35 Stoffdaten angepasst. Alle weiteren Größen ergeben sich prädiktiv. Hinsichtlich der
36 behandelbaren Stoffsysteme bestehen kaum Einschränkungen.

37 In Deutschland wird die molekulare Modellierung und Simulation zur Ermittlung von
38 Stoffdaten nach wie vor fast nur an Universitäten und Forschungseinrichtungen
39 betrieben, obwohl der Ansatz mittlerweile fast 60 Jahre alt ist [6]. Im industriellen
40 Umfeld wird sein Potenzial zur Reduzierung von Zeit und Kosten bei der Prozess-
41 und Produktentwicklung [7] bisher noch kaum genutzt [8]. Die Gründe sind hierfür
42 meist die verhältnismäßig langen Rechenzeiten und fehlende Erfahrung. Die
43 Möglichkeiten der molekularen Modellierung und Simulation werden seit einigen
44 Jahren von der Industrial Fluid Properties Simulation Collective (IFPSC) [1]
45 ausgelotet. Die IFPSC schreibt dazu regelmäßig einen internationalen Wettbewerb
46 aus [5, 9-12], bei dem die wissenschaftliche Community industriell relevante
47 Fragestellungen mit molekularen Methoden bearbeitet.

48 In der vorliegenden Veröffentlichung wird das Programm *ms2* [13] vorgestellt, das zur
49 molekularen Simulation thermodynamischer Stoffdaten fluider Systeme mit

1 Kraftfeldmethoden entwickelt wurde. Der Fokus liegt dabei auf thermischen und
2 kalorischen Eigenschaften, Dampf-Flüssigkeits-Gleichgewichten und
3 Transportgrößen von Reinstoffen und Mischungen niedermolekularer elektroneutraler
4 Moleküle. Die zu simulierenden Systeme sind im thermodynamischen Gleichgewicht
5 und bestehen typischerweise aus etwa 10^3 Molekülen. *ms2* wurde bereits in einer
6 Vielzahl an Anwendungen erfolgreich eingesetzt, unter anderem bei zwei
7 Wettbewerben der IFPSC. Im Jahr 2006 wurde damit der zweite Platz erzielt [11], im
8 Folgejahr wurde der Wettbewerb gewonnen [5].

9 *ms2* minimiert die Rechenzeiten durch eine effiziente Nutzung der zur Verfügung
10 stehenden Computer-Hardware, vom Einzel-PC bis zum Höchstleistungsrechner. Die
11 Bedienung des Programms ist einfach, die zugehörigen Hilfsprogramme erlauben
12 einen raschen Einstieg und unterstützen bei der Auswertung von
13 Simulationsergebnissen. Die Distribution des Programms erfolgt im Fortran90
14 Quellcode, der zur Ausführung lediglich kompiliert werden muss. Es wird keine
15 zusätzliche Software zur Ausführung benötigt. Die Hilfsprogramme basieren auf Java
16 bzw. OpenGL. Weiterhin werden validierte Kraftfelder für über einhundert
17 niedermolekulare elektroneutrale Moleküle zur Verfügung gestellt, die für die
18 Beschreibung von Mischungen miteinander kompatibel sind. Für die akademische
19 Nutzung wird das Programm kostenlos zur Verfügung gestellt.

20

21 **2 Simulationsprogramm *ms2***

22 2.1 Simulationstechniken

23 *ms2* unterstützt die zwei grundlegenden Simulationstechniken Molekulardynamik
24 (MD) und Monte-Carlo (MC).

25 MD Simulationen beruhen auf der zeitdiskretisierten numerischen Lösung der
26 Newtonschen Bewegungsgleichungen. Für einen gegebenen Zeitpunkt t_0 werden die
27 molekularen Wechselwirkungen berechnet, wobei die Kräfte und Momente eine
28 Schlüsselrolle einnehmen. Diese Größen treiben die molekularen Bewegungen an
29 und führen zu Änderungen der räumlichen Molekülanordnung bei $t_0 + \Delta t$. Eine
30 wiederholte Durchführung von Wechselwirkungsberechnung und Integration führen
31 zu einer Approximation der Molekülbewegungen und -anordnungen.
32 Thermodynamische Stoffdaten lassen sich daraus durch zeitliche Mittelungen
33 berechnen. Für die Lösung der Bewegungsgleichungen in der MD stellt *ms2* zwei
34 unterschiedliche Integratoren zur Verfügung, Leapfrog und Gear Prädiktor-
35 Korrektor [14].

36 MC Simulationen erzeugen molekulare Ensembles stochastisch. Die per
37 Zufallszahlen generierten Konfigurationen werden hinsichtlich ihrer Energie evaluiert
38 und akzeptiert oder verworfen. Im Ergebnis liefert MC eine Markov Kette
39 repräsentativer molekularer Anordnungen. Die thermodynamischen Stoffdaten
40 werden durch Ensemblemittelungen berechnet. Es ist zu beachten, dass die
41 physikalische Größe Zeit in MC Simulationen nicht auftritt, so dass damit der Zugang
42 zu dynamischen Zustandsgrößen wie Diffusionskoeffizienten, anders als in der MD,
43 nicht möglich ist.

44 Unabhängig von der Simulationstechnik gelten für alle Simulationen mit *ms2* die
45 periodischen Randbedingungen, die Minimum-Image Konvention [14] und die
46 Annahme der paarweisen Additivität der molekularen Wechselwirkungen. Die
47 zwischenmolekularen Wechselwirkungen werden bis zu einem wählbaren

1 Abschneideradius explizit ausgewertet, die verbleibenden langreichweitigen Beiträge
2 werden über Korrekturterme implizit berücksichtigt. Zahlreiche weitere in *ms2*
3 implementierte Algorithmen und Methoden, wie beispielsweise die
4 Geschwindigkeitsskalierung bzw. der Andersen Barostat zur Vorgabe von
5 Temperatur bzw. Druck, sind bereits seit langem bekannt und in der Literatur
6 beschrieben [14, 15].

7 Die Genauigkeit der aus molekularen Simulationen berechneten Stoffdaten wird
8 durch zwei Faktoren bestimmt: 1. die Qualität des Kraftfelds, das prinzipiell alle
9 thermodynamischen Größen eindeutig festlegt, und 2. das Sampling der molekularen
10 Konfigurationen, das mit statistischen Unsicherheiten behaftet ist.

11 Generell können mit *ms2* alle starren Kraftfelder simuliert werden, die auf dem
12 Lennard-Jones (LJ) Potenzial basieren und mit elektrostatischen Potenzialen
13 überlagert sind. Auch an der Entwicklung genauer Kraftfelder wurde in der
14 Vergangenheit intensiv mit *ms2* gearbeitet, so dass Kraftfelder für mehr als
15 einhundert molekulare Spezies zur Verfügung stehen [16]. Diese Kraftfelder
16 beschreiben die thermodynamischen Stoffgrößen mit Genauigkeiten, die industriellen
17 Anforderungen genügen. Das Sampling kann durch die Größe des simulierten
18 Systems und die Anzahl an berechneten Konfigurationen verbessert werden. Je
19 mehr Daten erhoben werden, desto geringer werden die statistischen
20 Unsicherheiten.

21

22 2.2. Ensembles

23 Die Berechnung von unterschiedlichen thermodynamischen Stoffdaten erfordert
24 unterschiedliche Simulationsrandbedingungen, d.h. sogenannte Ensembles.
25 Folgende Ensembles werden derzeit unterstützt:

- 26 • Kanonisches Ensemble (NVT) - MD und MC
- 27 • Mikrokanonisches Ensemble (NVE) - MD
- 28 • Isobar-isothermes Ensemble (NpT) - MD und MC
- 29 • Grand Equilibrium Methode (pseudo- μVT) - MC

30 Die Ensembles sind bereits seit langem bekannt und werden breit
31 angewandt [14, 15]. Die Grand Equilibrium Methode dient der effizienten Berechnung
32 von Dampf-Flüssigkeits-Gleichgewichten [17-19].

33

34 2.3. Thermodynamische Stoffdaten

35 *ms2* berechnet die thermodynamischen Stoffgrößen während der Simulation on-the-
36 fly aus Trajektorien (MD) oder Markov Ketten (MC) und die Ergebnisse werden in
37 Intervallen in Dateien geschrieben. Die statistischen Unsicherheiten aller Daten
38 werden durch eine Blockmittelung [20] und das Fehlerfortpflanzungsgesetz
39 abgeschätzt.

40

41 *Thermische und kalorische Eigenschaften.* Die zugänglichen Stoffdaten hängen vom
42 gewählten Ensemble ab. Bei vorgegebener Dichte, Temperatur und
43 Zusammensetzung (NVT Ensemble) werden Druck, residuelle innere Energie,
44 residuelle Enthalpie, residuelle isochore Wärmekapazität und die partielle Ableitung

1 der residuellen potenziellen Energie nach dem Volumen bei konstanter Temperatur
2 ermittelt. Bei vorgegebenem Druck, Temperatur und Zusammensetzung
3 (NpT Ensemble) werden Dichte, residuelle innere Energie, residuelle Enthalpie,
4 residuelle isobare Wärmekapazität, isotherme Kompressibilität, Wärmeausdehnung,
5 Schallgeschwindigkeit und die partielle Ableitung der residuellen Enthalpie nach dem
6 Druck bei konstanter Temperatur ermittelt. Die residuellen Größen lassen sich durch
7 Addition der Idealanteile in absolute Größen umrechnen [21].

8
9 *Chemisches Potenzial.* Das chemische Potenzial einer Komponente i kann
10 aufgegliedert werden in einen rein temperaturabhängigen Anteil $\mu_i^{\text{id}}(T)$ des idealen
11 Gases, der für die Berechnung des Phasengleichgewichts irrelevant ist, und den
12 verbleibenden Anteil $\mu_i(T, \mathbf{x}, p) - \mu_i^{\text{id}}(T)$, der vom Realverhalten abhängt und Beiträge
13 des idealen Gases beinhaltet. Letzterer kann in *ms2* mittels zweier Methoden
14 bestimmt werden: Widomsche Testteilchenmethode [22] und graduelle
15 Einsetzung [18, 19]. Beide Methoden sind sowohl für Reinstoffe als auch für
16 Mischungen anwendbar. Die Widomsche Methode zeichnet sich durch einen
17 geringen Rechenaufwand aus und ist sowohl in MD als auch in MC Simulationen
18 einsetzbar. Die graduelle Einsetzung ist ausschließlich als MC Methode
19 implementiert. Sie wird für sehr dichte Fluide verwendet, bei denen die Berechnung
20 nach Widom mit hohen statistischen Unsicherheiten behaftet ist.

21
22 *Dampf-Flüssigkeits-Gleichgewicht.* Die Berechnung des Dampf-Flüssigkeits-
23 Gleichgewichts erfolgt nach der Grand Equilibrium Methode [17-19] in zwei
24 aufeinander folgenden Simulationen. In der ersten Simulation wird für die
25 Flüssigphase bei vorgegebener Temperatur, Druck und Zusammensetzung das
26 chemische Potenzial aller Komponenten bestimmt. Hierfür stehen zwei
27 unterschiedliche Techniken zur Verfügung, s.o. In der zweiten, darauf folgenden
28 Simulation der Gasphase im pseudo- μVT Ensemble stellt sich das thermische,
29 mechanische und stoffliche Gleichgewicht zu den vorgegebenen Werten von
30 Temperatur und Flüssigkeitszusammensetzung ein. Es ergeben sich Stoffdaten des
31 Phasengleichgewichts, unter anderem der Dampfdruck, die
32 Dampfzusammensetzung und die Sättigungsdichten beider Phasen.

33
34 *Henry Konstante.* Die Löslichkeit eines Stoffs in einem Lösungsmittel wird durch die
35 Henry Konstante charakterisiert und ist bei gegebener
36 Lösungsmittelzusammensetzung eine rein temperaturabhängige Größe. Die Henry
37 Konstante H_i lässt sich aus dem chemischen Potenzial des gelösten Stoffs i bei
38 unendlicher Verdünnung berechnen [23]. Dies ist in *ms2* über die Widomsche
39 Testteilchenmethode bzw. die graduelle Einsetzung möglich. Hierbei wird das
40 Lösungsmittel im NpT Ensemble beim Dampfdruck simuliert. Die gelöste
41 Komponente wird nur in Form von Testmolekülen (Widom) bzw. eines fluktuierenden
42 Moleküls (graduelle Einsetzung) in die Simulation eingeführt.

43
44 *Transportgrößen.* Transportgrößen werden in *ms2* durch MD Simulationen im
45 Gleichgewicht bestimmt. Mikroskopische Fluktuationen des Systems um den
46 Gleichgewichtszustand bei konstantem Volumen und Temperatur bzw. Energie (NVT
47 bzw. NVE Ensemble) werden als Funktionen der Zeit aufgenommen und nach dem

1 Green-Kubo Formalismus ausgewertet [24]. Zugängliche Größen sind dabei der
2 Selbstdiffusionskoeffizient, der Maxwell-Stefan Diffusionskoeffizient in binären und
3 ternären Mischungen [25], sowie die Scherviskosität und die Volumenviskosität.

4

5 **3 Detaillierte Beschreibung**

6 *ms2* ist für eine genaue und schnelle Berechnung von thermodynamischen
7 Stoffdaten bei niedriger Arbeitsspeicherbelastung ausgelegt. Die hier vorgestellten
8 Konzepte zeigen die Realisierung dieser Anforderungen auf.

9

10 *Reduzierte Größen.* Die Einführung von reduzierten Größen verringert
11 Rundungsfehler. Alle dimensionsbehafteten Werte werden durch eine Referenzlänge
12 σ_R , eine Referenzenergie ε_R und eine Referenzmasse m_R reduziert, die als
13 Eingabeparameter frei gewählt werden können. In reduzierten Größen haben alle
14 Daten, die zur Ermittlung der thermodynamischen Stoffgrößen kumuliert werden,
15 etwa die gleiche Größenordnung, was zu einer erhöhten Genauigkeit der
16 Berechnungen führt.

17

18 *Zwischenmolekulare Wechselwirkungen.* Um die Anisotropie der Wechselwirkungen
19 zwischen den Molekülen zu modellieren, werden die Wechselwirkungszentren auf
20 Positionen (Sites) außerhalb des Schwerpunkts des Moleküls verteilt. Als
21 zwischenmolekulare Wechselwirkungen werden Dispersion und Repulsion sowie
22 Elektrostatik berücksichtigt.

23 Die dispersiven und repulsiven Wechselwirkungen werden auf paarweise
24 Interaktionen u_{ij} zwischen zwei Sites i und j von unterschiedlichen Molekülen
25 reduziert. Diese Interaktionen werden durch das weithin verbreitete Lennard-Jones
26 (LJ) Potenzial [14] modelliert

$$27 \quad u_{ij}(r_{ij}) = 4\varepsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right). \quad (1)$$

28 Hierbei beschreibt r_{ij} den Abstand zwischen den Sites und σ_{ij} bzw. ε_{ij} bezeichnen den
29 LJ Längen- bzw. Energieparameter. Für unterschiedliche LJ Sites sind die
30 Wechselwirkungen standardmäßig durch die Lorentz-Berthelot
31 Kombinationsregel [26] definiert. Für Mischungen können diese Parameter mittels
32 der modifizierten Lorentz-Berthelot Regel [26] frei skaliert werden.

33 Die elektrostatischen Wechselwirkungen werden durch Punktladungen, Punktdipole
34 und Punktquadrupole beschrieben. Punktdipole und Punktquadrupole fassen zwei
35 bzw. drei Punktladungen zu einer elektrostatischen Site zusammen und reduzieren
36 daher die Anzahl an Wechselwirkungsberechnungen. Dies führt zu einer erheblichen
37 Beschleunigung der Simulation und vereinfacht die Kraftfeldentwicklung.

38

39 *Speicherung von Simulationsdaten.* Die Speicherung der Ortskoordinaten erfolgt auf
40 der Basis von molekularen Massenschwerpunkten und Orientierungen. Daraus
41 werden die Positionen aller Sites während der Simulation fortlaufend berechnet. Die
42 Schwerpunktpositionen werden in kartesischen Koordinaten gespeichert, die

1 Orientierungen in Form von Quaternionen [14]. Dieser Ansatz erlaubt eine robuste
2 und effiziente Berechnung aller Positionen mit wenig Arbeitsspeicher.

3

4 *Implementierung.* *ms2* wurde in Fortran90 geschrieben, eine Sprache, die aufgrund
5 ihrer einfachen Struktur von Ingenieuren gern eingesetzt wird, und mit der eine hohe
6 Ausführungsgeschwindigkeit erreicht werden kann. Die Implementierung weist eine
7 modulare Struktur und einen objektorientierten Aufbau auf, so dass in der Simulation
8 nur die Module zusammengeschaltet werden, die für die Berechnung notwendig sind.
9 Der modulare Aufbau ermöglicht zudem eine gute Übersicht über das Programm und
10 vereinfacht Erweiterungen. Der Aufbau von *ms2* sowie einige Module sind in
11 Abbildung 1 skizziert.

12

13 *Parallelisierung.* *ms2* erlaubt eine parallele Ausführung auf einer wählbaren Anzahl
14 von Prozessoren N_p . Hierfür wird der Message Passing Interface (MPI) Standard
15 verwendet. Die Parallelisierungsstrategie variiert je nach Simulationstechnik. In MD
16 Simulationen wird die Berechnung der zwischenmolekularen Wechselwirkungen auf
17 die Prozessoren gemäß der Kraftzerlegung nach Plimpton [27] verteilt. Jeder
18 Prozessor bearbeitet so die Wechselwirkungen für N/N_p Moleküle, wobei N die
19 Anzahl der Moleküle in der Simulation ist. Die Ergebnisse der Berechnungen werden
20 am Ende jedes Zeitschritts an den Root-Prozessor kommuniziert, der die neuen
21 Koordinaten der Moleküle ermittelt.

22 In MC werden die M Simulationsschritte auf N_p Prozessoren verteilt. Jeder Prozessor
23 bearbeitet so M/N_p Konfigurationen, die hinsichtlich der Stoffdaten ausgewertet
24 werden. Die Ergebnisse werden lediglich zum Schluss der Simulation an den Root-
25 Prozessor kommuniziert, der sie zum Endergebnis mittelt, was zu einer optimalen
26 Ausnutzung aller zur Verfügung stehender Rechenressourcen führt.

27

28 **4 Benchmarking**

29 *Parallelisierung.* Die Performance der beiden Parallelisierungsstrategien wurde auf
30 einem Intel „Nehalem“ Dual-Quadcore PC-Cluster mit Infiniband-Netzwerk getestet.
31 MD Simulationen skalieren darauf gut bis zu 16-32 Prozessoren, während MC
32 Simulationen bis hin zu hohen Prozessorenzahlen optimal skalieren, vgl.
33 Abbildung 2.

34

35 *Vergleiche mit anderen Simulationsprogrammen*

36 *Molekulardynamik.* Die Performance von *ms2* hinsichtlich MD wurde mit dem
37 Programm *Gromacs V4.0.3* [30] verglichen. In diesem Vergleich wurde eine
38 äquimolare flüssige Mischung von Methanol und Ethanol aus 1372 Molekülen im
39 *NVT* Ensemble simuliert. Die Simulationen wurden auf einem Intel „Harpertown“ PC-
40 Cluster mit Infiniband Netzwerk ausgeführt. Im Vergleich ist *Gromacs* um einen
41 Faktor 2 schneller als *ms2* und weist auch eine bessere Skalierung bei parallelen
42 Rechnungen auf. Der Testfall eines industriell relevanten Phasengleichgewichts
43 wurde mit *Gromacs* nicht berechnet, da die hierfür notwendigen Algorithmen nicht
44 implementiert sind.

45

1 *Monte-Carlo*. Die Performance von *ms2* hinsichtlich MC Simulationen im *NVT*
2 Ensemble wurde mit dem Programm *MCCCS Towhee V6.2.7* [28] verglichen, wobei
3 erneut das System Methanol und Ethanol im *NVT* Ensemble simuliert wurde. *ms2*
4 und *Towhee* wurden jeweils mit dem Intel ifortran Compiler übersetzt und auf einem
5 einzelnen Prozessor die gleiche Anzahl an Simulationsschritten ausgeführt. Hierbei
6 war *ms2* um einen Faktor 20 schneller als *Towhee*.

7 In einem zweiten Vergleich wurden *ms2* und *Towhee* in Bezug auf die Berechnung
8 von Dampf-Flüssigkeits-Gleichgewichten untersucht. Betrachtet wurde die
9 äquimolare flüssige Mischung von Methanol und Ethanol bei 393 K. In den Vergleich
10 wurde auch der *Errington* Gibbs Ensemble Monte Carlo Code [29] einbezogen.
11 *Towhee* und *Errington* verwenden zur Phasengleichgewichtsberechnung das Gibbs-
12 Ensemble, während *ms2* die Grand Equilibrium Methode einsetzt. Die Simulationen
13 der Flüssigphase wurden mit 1372 Molekülen durchgeführt, die der Gasphase mit
14 500 Molekülen, wobei diese Molekülanzahl für die Gibbs Ensemble Berechnungen
15 mit *Towhee* bzw. *Errington* als Startwert diente. Der Vergleich wurde über 96 h
16 Simulationszeit auf einem Prozessor durchgeführt. Innerhalb dieser Zeit durchlief
17 *ms2* 52000 Loops für die Flüssigphase und 100000 Loops für die Gasphase, wobei
18 jeder Loop aus 4116 unterschiedlichen Molekülanordnungen besteht. Der *Errington*
19 Code durchlief in der vorgegebenen Zeit insgesamt 22000 Loops, während *Towhee*
20 6000 Loops berechnete. Die Geschwindigkeit der Programme spiegelt sich deutlich
21 in den Ergebnissen wider. *ms2* reproduzierte die experimentellen Ergebnisse mit
22 hoher Genauigkeit, während bei *Errington* bzw. *Towhee* die statistischen
23 Ungenauigkeiten um einen Faktor 10 bzw. 50 höher lagen.

24

25 *Berechnung von Transportgrößen*. Der Rechenaufwand zur Berechnung von
26 Transportgrößen mit *ms2* wurde anhand der Simulation einer äquimolaren flüssigen
27 Mischung aus Methanol und Ethanol bei 298 K und 0.1 MPa bestimmt. Die
28 Transportgrößen wurden mit einer Autokorrelationslänge von 13.82 ps ermittelt,
29 wobei jeder Zeitschritt ausgewertet wurde. Die Zeitspanne zwischen den unabhängig
30 ausgewerteten Autokorrelationsfunktionen lag bei 197.4 fs. Insgesamt wurden
31 100 Autokorrelationsfunktionen ausgegeben. Der Mehraufwand an Rechenzeit für
32 die Transportgrößen lag bei 78 % im Vergleich zu einer Rechnung, in der die
33 Transportgrößen nicht berechnet wurden.

34

35 **5 Zusatzprogramme**

36 Drei Zusatzprogramme erleichtern den Umgang mit dem Simulationsprogramm *ms2*.
37 Die Spezifikation der Simulationsszenarien sowie die Auswertung der Ergebnisse
38 werden damit vereinfacht und die grafische Darstellung der Simulationen ermöglicht.

39

40 *ms2par*. Das Java-Programm *ms2par* unterstützt mit einem graphischen Interface die
41 Erstellung der Eingabedatei, die für die Ausführung einer molekularen Simulation mit
42 *ms2* notwendig ist.

43

44 *ms2chart*. Das Java-Programm *ms2chart* dient der grafischen Analyse der
45 Simulationsergebnisse von *ms2*. Es stellt die thermodynamischen Stoffdaten über
46 dem Verlauf der Simulation dar, sowohl als Blockmittel als auch als laufende Mittel.

1 Die Grafiken lassen sich in Form von Bildern speichern. Die Anwendung von
2 *ms2chart* ist jederzeit möglich, auch während des Simulationsablaufs.

3

4 *ms2molecules*. Das in C geschriebene und auf OpenGL basierende *ms2molecules*
5 ist ein Visualisierungsprogramm für Konfigurationen, die von *ms2* erzeugt wurden.
6 Die LJ Sites werden als farbige Kugeln abgebildet, deren Durchmesser
7 standardmäßig dem LJ Längenparameter σ entspricht. Die Visualisierung bietet
8 Funktionen wie Rotation oder Zoom und kann in Form von Bildern oder Videos
9 gespeichert werden. Die Bedienung des Programms erfolgt durch Konsolenbefehle.
10 Abbildung 3 zeigt einen entsprechenden Schnappschuss.

11

12 **6 Zusammenfassung**

13 Das molekulare Simulationsprogramm *ms2* ermöglicht die Berechnung von
14 thermodynamischen Stoffdaten *in silico* auf der Basis von Kraftfeldern (molekularen
15 Modellen), die die zwischenmolekularen Wechselwirkungen beschreiben. Der Fokus
16 der Berechnungen liegt derzeit auf Fluiden, die durch starre, elektroneutrale
17 Kraftfelder dargestellt werden können. Die Simulationen liefern genaue Stoffdaten
18 bei niedrigen Ausführungszeiten. Die zugänglichen Stoffdaten reichen von statischen
19 Eigenschaften wie thermischen und kalorischen Größen und Dampf-Flüssigkeits-
20 Gleichgewichten zu Transportgrößen wie Diffusionskoeffizienten und Viskositäten. Es
21 können sowohl Reinstoffe als auch Mischungen simuliert werden.

22 Die Genauigkeit und Effizienz der Simulationen mit *ms2* wurde vielfach in
23 Veröffentlichungen und bei Simulationswettbewerben nachgewiesen. Die eigens
24 entwickelten Kraftfelder für mehr als einhundert industrierelevante Komponenten
25 werden gemeinsam mit dem Programm zur Verfügung gestellt.

26 Trotz seiner breiten Anwendbarkeit ist die Implementierung von *ms2* aufgrund seines
27 modularen und objektorientierten Aufbaus übersichtlich und leicht erweiterbar. Die
28 Anwendung ist auch für Neueinsteiger in die molekulare Simulation möglich.
29 Hilfsprogramme unterstützen den Anwender bei der Generierung von
30 Eingabedateien und bei der grafischen Analyse der Simulationsergebnisse. Das
31 Programm wurde für die derzeitige Hardwaretechnologie optimiert und erreicht eine
32 hohe Effizienz.

33 Derzeitig wird *ms2* weiterentwickelt, um die Anwendbarkeit des Programms zu
34 erweitern. Dies bezieht sich insbesondere auf innere molekulare Freiheitsgrade für
35 die Simulation von größeren Molekülen und auf die Ewald-Summation für die
36 Simulation von elektrisch geladenen Molekülen.

37 Der Quellcode ist für den akademischen, nichtkommerziellen Gebrauch kostenlos
38 verfügbar unter <http://www.ms-2.de>.

39

40

41 **7. Danksagung**

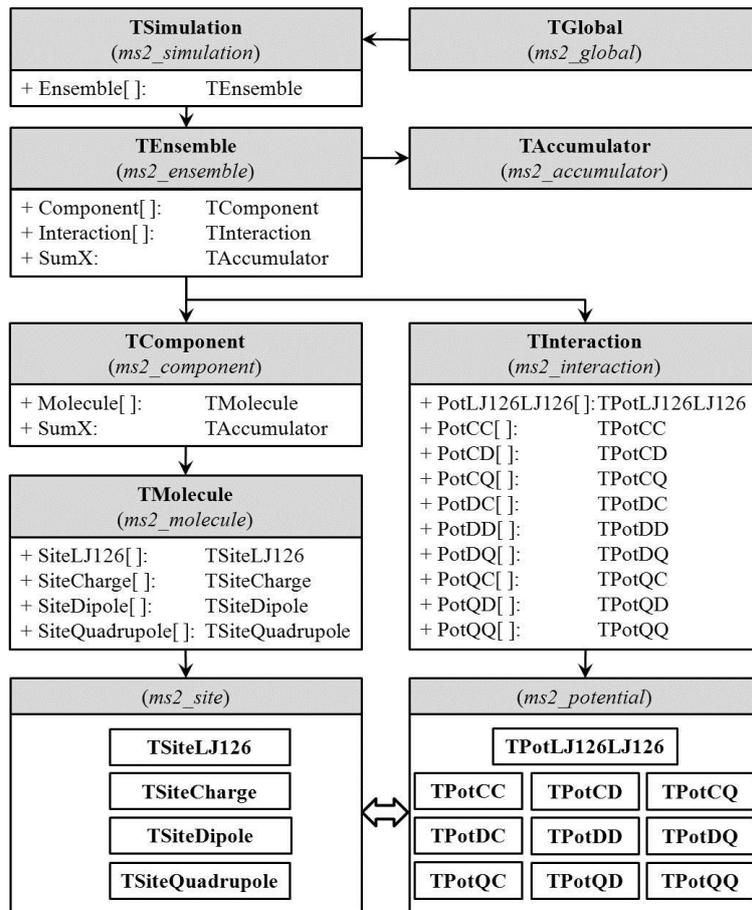
42 Die Autoren bedanken sich beim BMBF für die finanzielle Unterstützung des Projekts
43 „Innovative HPC-Methoden und Einsatz für hochskalierbare Molekulare Simulation
44 (01H08013A)“. Die Arbeiten wurden durch die Boltzmann-Zuse Society of
45 Computational Molecular Engineering (BZS) koordiniert.

Literatur

- [1] Industrial Fluid Properties Simulation Collective, <http://www.fluidproperties.org>, **2011**.
- [2] K. E. Gubbins, J. D. Moore, *J. Chem. Phys.* **2010**, 49, 3026.
DOI: 10.1021/je901909c.
- [3] M. Kleiner, F. Turnakaka, G. Sadowski, *Molecular Thermodynamics of Complex Systems* **2009**, 131, 75. DOI: 10.1007/430_2008_2.
- [4] A. Klamt, F. Eckert, *Fluid Phase Equilib.* **2000**, 172, 43.
DOI: 10.1016/S0378-3812(00)00357-5.
- [5] F. H. Case, J. Brennan, A. Chaka, K. D. Dobbs, D. G. Friend, P. A. Gordon, J. D. Moore, R. D. Mountain, J. D. Olson, R. B. Ross, M. Schiller, V. K. Shen, E. A. Stahlberg, *Fluid Phase Equilib.* **2008**, 274, 2. DOI: 10.1016/j.fluid.2008.05.009.
- [6] N. Metropolis, A. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, *J. Chem. Phys.* **1953**, 21, 1087. DOI: 10.1063/1.1699114.
- [7] The Economist, **2010**, March 13th, 68.
- [8] S. Gupta, J. Olson, *Ind. Eng. Chem. Res.* **2003**, 42, 6359.
DOI: 10.1021/ie030170v.
- [9] F. H. Case, A. Chaka, D. G. Friend, D. Frurip, J. Golab, R. Johnson, J. D. Moore, R. D. Mountain, J. Olson, M. Schiller, J. Storer, *Fluid Phase Equilib.* **2004**, 217, 1.
DOI: 10.1016/S0378-3812(03)00208-5.
- [10] F. H. Case, A. Chaka, D. G. Friend, D. Frurip, J. Golab, P. Gordon, R. Johnson, P. Kolar, J. Moore, R. D. Mountain, J. Olson, R. Ross, M. Schiller, *Fluid Phase Equilib.* **2005**, 236, 1. DOI: 10.1016/j.fluid.2005.06.015.
- [11] F. H. Case, J. Brennan, A. Chaka, K. D. Dobbs, D. G. Friend, D. Frurip, P. A. Gordon, J. D. Moore, R. D. Mountain, J. D. Olson, R. B. Ross, M. Schiller, V. K. Shen, *Fluid Phase Equilib.* **2007**, 260, 153. DOI: 10.1016/j.fluid.2007.08.021.
- [12] F. H. Case, A. Chaka, J. Moore, R. D. Mountain, J. D. Olson, R. B. Ross, M. Schiller, V. K. Shen, E. A. Stahlberg, *Fluid Phase Equilib.* **2009**, 285, 1.
DOI: 10.1016/j.fluid.2009.08.005.
- [13] S. Deublein, B. Eckl, J. Stoll, S. V. Lishchuk, G. Guevara-Carrion, C. W. Glass, T. Merker, M. Bernreuther, H. Hasse, J. Vrabec, *Comput. Phys. Commun.* **2011**, accepted.
- [14] M. Allen, D. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford **1987**.
- [15] D. Frenkel, B. Smith, *Understanding Molecular Simulation*, Academic Press, Elsevier, San Diego **1993**.

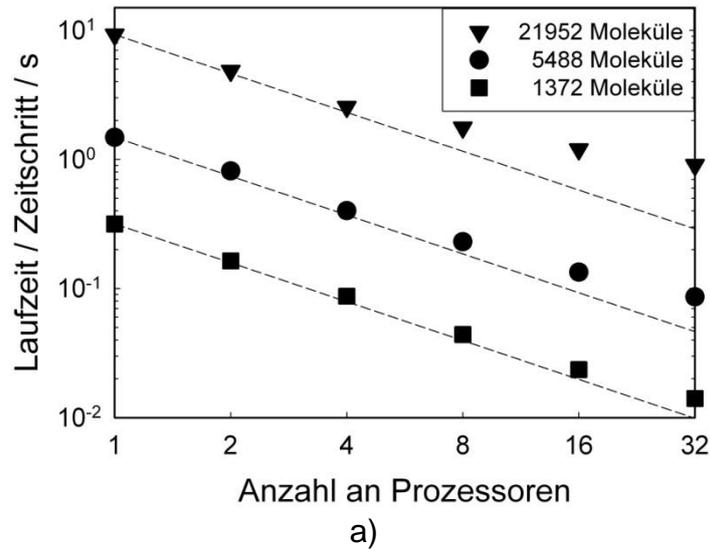
1
2 [16] B. Eckl, J. Vrabec, H. Hasse, *Chem. Ing. Tech.* **2008**, 80, 25.
3 DOI: 10.1002/cite.200700102.
4
5 [17] J. Vrabec, H. Hasse, *Mol. Phys.* **2002**, 100, 3375.
6 DOI: 10.1080/00268970210153772.
7
8 [18] A. P. Lyubartsev, A. A. Martsinovskii, S. V. Shevkunov, P. N. Vorontsov-
9 Velyaminov, *J. Chem. Phys.* **1992**, 96, 1776. DOI: 10.1063/1.462133.
10
11 [19] I. Nezbeda, J. Kolafa, *Mol. Simul.* **1991**, 5, 391.
12 DOI: 10.1080/08927029108022424
13
14 [20] H. Flyvbjerg, H. Petersen, *J. Chem. Phys.* **1989**, 91, 461.
15 DOI: 10.1063/1.457480.
16
17 [21] R. Rowley, W. V. Wilding, J. L. Oscarson, Y. Yang, N. A. Zundel, T. E. Daubert,
18 R. P. Danner, *DIPPR Data Compilation of Pure Compound Properties*, Design
19 Institute for Physical Properties, **2003**.
20
21 [22] B. Widom, *J. Chem. Phys.* **1963**, 39, 2808. DOI: 10.1063/1.1734110.
22
23 [23] K. Shing, K. Gubbins, K. Lucas, *Mol. Phys.* **1988**, 65, 1235.
24 DOI: 10.1080/00268978800101731.
25
26 [24] K. Gubbins, *Statistical Mechanics: Volume 1*, The Chemical Society Burlington
27 House, London **1972**.
28
29 [25] R. Krishna, J. M. van Baten, *Ind. Eng. Chem. Res.* 2005, 44, 6939.
30 DOI: 10.1021/ie050146c.
31
32 [26] T. Schnabel, J. Vrabec, H. Hasse, *J. Mol. Liq.* **2007**, 135, 170.
33 DOI: 10.1016/j.molliq.2006.12.024.
34
35 [27] S. Plimpton, *J. Comput. Phys.* **1995**, 117, 1. DOI: 10.1006/jcph.1995.1039.
36
37 [28] MCCC'S Towhee, <http://www.towhee.sourceforge.org>, **2008**.
38
39 [29] J. Errington, A. Z. Panagiotopoulos,
40 <http://kea.princeton.edu/jerring/gcmc/index.html>.
41
42 [30] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory*
43 *Comput.* **2008**, 4, 435. DOI: 10.1021/ct700301q.
44

1

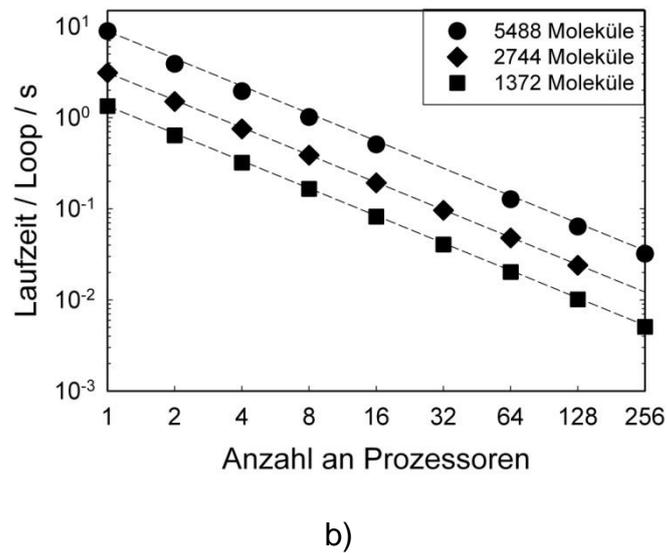


2

3 **Abbildung 1: UML Diagramm des Aufbaus von *ms2*. Die Beschriftung hervorgehoben**
 4 **durch fetten Zeichensatz kennzeichnet die Modulklassen, die kursiven Namen**
 5 **bezeichnen die Datei von *ms2*, in der die Klassen implementiert sind. Die Pfeile**
 6 **geben den hierarchischen Aufbau des Programms wieder, während die großen Pfeile**
 7 **den Datenfluss zwischen einzelnen Ebenen kennzeichnen. Zur einfacheren**
 8 **Überblick über die Struktur wurden folgende Abkürzungen eingeführt:**
 9 **CC: ChargeCharge; CD: ChargeDipol; CQ: ChargeQuadrupol; DC: DipolCharge;**
 10 **DD: DipolDipol; DQ: DipolQuadrupol; QC: QuadrupolCharge; QD QuadrupolDipol;**
 11 **QQ QuadrupolQuadrupol.**



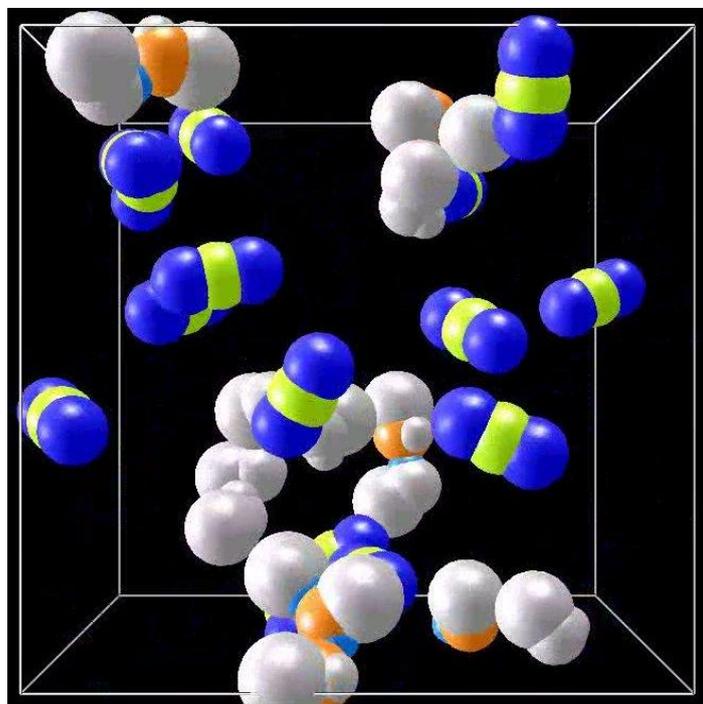
1
2



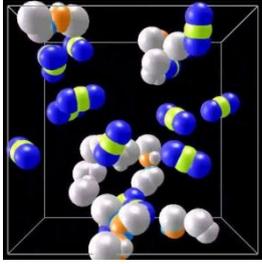
3
4
5
6
7
8
9
10
11
12
13

Abbildung 2: Laufzeiten für MD (a) und MC (b) Simulationen mit *ms2* auf einem Intel „Nehalem“ PC-Cluster mit „Infiniband“ Netzwerk unter Verwendung des Intel ifortran Compilers und Intel MPI V4.0. Es wurde eine äquimolare flüssige Mischung aus Methanol und Ethanol bei 298 K und 0.1 MPa simuliert. Die gestrichelten Linien stellen die optimale Skalierung dar. Die Symbole kennzeichnen Simulationen mit unterschiedlicher Systemgröße.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16



17
18 **Abbildung 3:** Momentaufnahme aus der Simulation einer binären Mischung,
19 dargestellt mit *ms2molecules*. Der weiße Rahmen stellt das Simulationsvolumen dar.
20



Das Programm *ms2* zur Ermittlung von thermodynamischen Stoffdaten mittels molekularer Simulation wird vorgestellt. Hauptanwendungsgebiete sind Eigenschaften homogener Phasen, Dampf-Flüssigkeits-Gleichgewichte und Transportgrößen. *ms2* wurde ausgelegt, dass die Stoffeigenschaften zeiteffizient berechnet werden können und dass die Bedienbarkeit des Programms einfach ist.