Molecular Simulation of Chemical Reaction Equilibrium by Computationally Efficient Free Energy Minimization

William R. Smith*†‡§∥ and Weikai Qi†

†Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1, Canada
‡Department of Chemistry, University of Guelph, Guelph, Ontario N1G 2W1, Canada
§Department of Chemical Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
∥Faculty of Science, University of Ontario Institute of Technology, Oshawa, Ontario L1H 7K4, Canada

ABSTRACT: The molecular simulation of chemical reaction equilibrium (CRE) is a challenging and important problem of broad applicability in chemistry and chemical engineering. The primary molecular-based approach for solving this problem has been the reaction ensemble Monte Carlo (REMC) algorithm [Turner et al. Molec. Simulation 2008, 34, (2), 119–146], based on classical force-field methodology. In spite of the vast improvements in computer hardware and software since its original development almost 25 years ago, its more widespread application is impeded by its computational inefficiency. A fundamental problem is that its MC basis inhibits the implementation of significant parallelization, and its successful implementation often requires system-specific tailoring and the incorporation of special MC approaches such as replica exchange, expanded ensemble, umbrella sampling, configurational bias, and continuous fractional component methodologies. We describe herein a novel CRE algorithm (reaction ensemble molecular dynamics, ReMD) that exploits modern computer hardware and software capabilities, and which can be straightforwardly implemented for systems of arbitrary size and complexity by exploiting the parallel computing methodology incorporated within many MD software packages (herein, we use GROMACS for illustrative purposes). The ReMD algorithm utilizes these features in the context of a macroscopically inspired and generally applicable free energy minimization approach based on the iterative approximation of the system Gibbs free energy function by a mathematically simple convex ideal solution model using the composition at each iteration as a reference state. Finally, we additionally describe a simple and computationally efficient a posteriori method to estimate the equilibrium concentrations of species present in very small amounts relative to others in the primary calculation. To demonstrate the algorithm, we show its application to two classic example systems considered previously in the literature: the N₂–O₂–NO system and the ammonia synthesis system.

1. INTRODUCTION

The chemical reaction equilibrium (CRE) problem, most typically at a specified absolute temperature \( T \) and pressure \( P \), is of considerable importance in chemistry and chemical engineering. It can be posed as the solution of the nonlinear optimization problem of minimizing the system Gibbs function subject to element abundance conservation constraints, and macroscopically based thermodynamic models are widely used for this purpose (see, for example, Leal et al.†).

The CRE problem has been approached by means of molecular simulation in three different ways, each of which is based on a different molecular model of the system. When classical force fields (CFFs) are used to model the molecular environment, the reaction ensemble Monte Carlo (REMC) algorithm provides a numerically exact solution of the problem. The algorithm is essentially a direct translation of the macroscopic thermodynamic approach to the corresponding molecular environment. It requires a specified list of species (which may be expanded to include species present in minor amounts by means of an a posteriori approach described in Section 3), and implicitly incorporates the nonideal (residual) portions of the chemical potentials by molecular simulation involving the species force fields (FFs). Standard state ideal-gas quantities \( \mu_i^0(T, P) \) or ideal-gas reaction free energy changes, \( \Delta G^0_i(T; P) \) are also required input information, both of which may be separately obtained either from thermochemical tables, e.g., the JANAF Thermochemical Tables, or from partition functions calculated using quantum mechanical software. We emphasize that such quantities are not required in the simulation of phase equilibria, since they cancel when the total interphase species chemical potentials are equated.

Received: June 6, 2018
Published: August 23, 2018
REMC has been applied to many systems (see Turner et al. \textsuperscript{4} for a review).

A second methodology describes the system using a quantum mechanical (QM) model involving its atoms. The recently introduced “first-principles” RxPFMC algorithm of Fetisov et al. \textsuperscript{6} solves the CRE problem by means of an approximate QM simulation of the atoms of the system and an empirically based cluster analysis of its equilibrated atomic structure using a molecular identification criterion. Standard chemical potential data are not required. The original applications\textsuperscript{8} have thus far been applied to a relatively simple system involving N and O atoms. It is very computationally expensive, and its extension to larger molecules and their structural isomers will require more detailed identification criteria for particular species, somewhat analogously to the use of a “species list”.

A third methodology describes the molecular environment by means of a “reactive force field” (ReaxFF) approach,\textsuperscript{7–9} which uses force fields (whose parameters are typically obtained by matching to QM data), to model chemical reactions by means of the formation and dissociation of chemical bonds. van der Waals and Coulomb interaction terms are also included, and the ReaxFFs can model reaction kinetics in addition to reaction equilibria when implemented within a molecular dynamics (MD) simulation. This approach also does not require standard chemical potential data, and its underlying molecular description thus lies intermediate between the CFF and QM methods. FF parameter sets have thus far been developed primarily for atoms involved in combustion (C/B/N/H/O) systems and in aqueous (Ni/C/H/O) systems,\textsuperscript{8} with some extensions to systems involving other atomic species (see Senffle et al.\textsuperscript{4} for a review).

We note that, as for macroscopic CRE calculations using sets of chemical reactions to implement mass conservation and contrary to statements in the literature (e.g., Fetisov et al.\textsuperscript{6}), REMC does not require a prescribed set of chemical reactions. As described in Smith and Missen,\textsuperscript{10} an equilibrium algorithm based on the use of chemical reactions incorporates the atom conservation constraints by calculating an appropriate stoichiometric matrix $\nu$ from the formula matrix $A$, and recalculating it as necessary during the course of the algorithm. This is a nontrivial task for multireaction systems, but for simple systems described by a single reaction (such as the examples considered herein), the stoichiometry is usually obvious, and it does not change during the course of the algorithm. In such cases, the reaction may be provided to the equilibrium algorithm at the outset.

It has also been stated\textsuperscript{6} that “RxPFMC simulations are capable of generating new molecules, whereas the REMC (sic) simulations are limited to only those species appearing in the reaction set”. We show in this paper to the contrary that species not considered in the initial species list may be considered by means of a simple \textit{a posteriori} procedure, which may be viewed as the molecular counterpart of an approach implemented by macroscopic algorithms.

A common disadvantage of all currently available molecular-simulation-based CRE algorithms for their routine implementation is their computational complexity. Although REMC\textsuperscript{5} is arguably the currently most general and useful algorithm for the simulation of a broad range of CRE problems, its routine implementation for systems involving complex molecules, strong intermolecular interactions, high densities, and low temperatures is not easily formulated, since its computational efficiency typically depends on specific features of the molecules involved. REMC functions by implementing sequences of “reaction moves” interspersed with equilibration sequences; this process is carried out until numerical convergence is deemed to have been achieved. The major computational bottleneck lies in the efficient implementation of the reaction moves, which are equivalent to alchemical changes involving reactant and product species according to prescribed probabilities. For all but the simplest of systems, such as replica exchange\textsuperscript{11,12} and configurational bias,\textsuperscript{13,14} continuous fractional component,\textsuperscript{15,16} and related\textsuperscript{17,18} approaches.

The goal of this paper is to describe a novel classical FF-based CRE algorithm (ReMD), which is both computationally efficient and can be routinely applied to a wide range of systems, and whose implementation does not depend on the specific nature of the molecules involved. The algorithm is based on the core use of modern molecular dynamics (MD) simulation technology, which has undergone vast improvements in computational efficiency since the original development of the REMC algorithm, and which are currently implemented in widely available MD simulation packages such as GROMACS.\textsuperscript{19} The relevant improvements for our purposes are due to both the widespread availability of increased hardware computing speeds (roughly 2 orders of magnitude since 1994\textsuperscript{20}) and utilization of the software parallelization capabilities inherent in MD algorithms.\textsuperscript{21–23} The key feature of the proposed ReMD algorithm is the coupling of these features with a computationally efficient free energy minimization strategy based on classical thermodynamic considerations.

In the next section of the paper, we describe the ReMD algorithm. The subsequent section describes the example systems studied, and the following section describes the simulation details and protocols. For illustrative purposes, we consider two classic single-phase CRE simulation problems: the Ni\textsubscript{2}−O\textsubscript{2}−NO system\textsuperscript{2,6,24} and the ammonia synthesis system,\textsuperscript{25−27} leaving application to more complex problems to future work. A Results and Discussion section follows, with a final section containing our conclusions.

### 2. ReMD Algorithm

We describe the algorithm for a reacting system at specified overall thermodynamic variables ($T$, $P$), involving a single fluid phase using mole fraction concentration variables; extensions to other thermodynamic and composition variables, and multiple phases, are straightforward in principle.

The ReMD algorithm is based on the translation to the molecular-simulation context of physical chemistry concepts underlying the macroscopic thermodynamic formulation of the CRE problem,\textsuperscript{10} which we first briefly review. The CRE problem at specified ($T$, $P$) seeks to minimize the total system Gibbs free energy $G(T, P; N)$ in a closed system, given by

$$G(T, P; N) = \sum_{i=1}^{N} N_{i} \mu_{i}(T, P; x)$$

where $N$ is the species mole number vector with entries $N_{i}$ giving the molar amount of species $i$, $N_{i}$ is the number of species, $\mu_{i}$ is the chemical potential of species $i$, $x$ is the system mole fraction vector, with entries $x_{i} = N_{i}/N_{t}$, and $N_{t}$ is the total number of moles.

The atom conservation constraints may be incorporated by means of a stoichiometric matrix that can be interpreted as
describing a set of chemical reactions, and the mole number variables can be expressed as

\[ N_i = N_i^0 + \sum_{j=1}^{R} v_{ij} \xi_j \]  

(2)

where \( R \) is the maximum number of linearly independent chemical reactions, and \( v_{ij} \) is the stoichiometric coefficient of species \( i \) in reaction \( j \) (we use the convention \( v_{ij} < 0 \) for a reactant species and \( v_{ij} > 0 \) for a product species). We emphasize that any reaction set may be used, whose sole purpose is to incorporate the element conservation equations of the system (see Smith and Missen, Chapter 2). \( N_i^0 \) is an arbitrary set of molar amounts satisfying the constraints, and \( \xi_j \) is a reaction extent variable. The necessary condition for the minimum of \( G \) in terms of the \( \xi_j \) variables is

\[ \frac{\partial G}{\partial \xi_j} \equiv \Delta G(T, P; x) = \sum_{i=1}^{R} v_{ij} \mu_i(T, P; x) = 0, \quad j = 1, 2, ..., R \]  

(3)

The molar species chemical potentials calculated using molecular simulation can be expressed in the following form:

\[ \mu_i(T, P; x) = \mu_i^0(T; P^0) + RT \ln \left( \frac{P}{P^0} \right) + RT \ln \left( \frac{x_i}{\xi_i} \right) \]

(4)

where \( \mu_i^0(T; P^0) \) is the ideal-gas chemical potential of species \( i \) at \( T \) and the standard state pressure \( P^0 \) (typically 1 bar). \( \mu_i^{\text{res},\text{NPT}}(T, P; x) \) is the residual chemical potential with respect to the underlying \( (T, P) \) ideal-gas model in the NPT ensemble. We have explicitly incorporated the fact that the reference state composition is \( x^0 = 1 \), in view of the following development.

The basis of the ReMD algorithm is the iterative construction of ideal solution approximations to the system Gibbs function at a sequence of reference state compositions, which are obtained by the rapid solution of simple ideal solution CRE problems and which ultimately converge to the solution of eq 3.

We construct each ideal solution approximation by first using eq 4 to write an exact expression for each chemical potential at an arbitrary composition \( x \) in terms of its value at a composition \( \bar{x} \), which will be used in the algorithm as a reference state composition:

\[ \mu_i(T, P; x) = \mu_i^0(T; P^0) + RT \ln \left( \frac{P}{P^0} \right) + RT \ln \left( \frac{x_i}{\bar{x}_i} \right) \]

\[ + \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) + RT \ln \left( \frac{\bar{x}_i}{x_i} \right) \]

\[ + [\mu_i^{\text{res},\text{NPT}}(T, P; x) - \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x})] \]

\[ = \mu_i^*(T, P) + RT \ln \left( \frac{x_i}{\bar{x}_i} \right) \]

\[ + [\mu_i^{\text{res},\text{NPT}}(T, P; x) - \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x})] \]  

(5)

where

\[ \mu_i^*(T, P) = \mu_i^0(T; P^0) + RT \ln \left( \frac{P}{P^0} \right) + \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) \]

(6)

The ideal solution approximation is constructed by neglecting the bracketed term in eq 5, to give

\[ \mu_i(T, P; x) = \mu_i^*(T, P) + RT \ln \left( \frac{x_i}{\bar{x}_i} \right) \]  

(7)

At each iteration, whose composition is denoted by \( \bar{x} \), the ReMD algorithm solves the CRE problem for this ideal solution approximation to yield the composition at the next iteration. Solving this problem is a relatively simple task, even in the case of a multireaction system, for which many computationally efficient algorithms are available. We can examine the quality of the ideal solution approximation of eq 7 by considering deviations of the values of \( G \) and the components of its gradient \( \partial G/\partial \xi_j \) from the corresponding exact values at the reference state \( \bar{x} \), which are given by

\[ G(T, P; N) - G(T, P; \bar{N}) = \sum_{i=1}^{N} N_i[\mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) - \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x})] \]

(8)

\[ \frac{\partial G}{\partial \xi_j} (T, P; N) - \frac{\partial G}{\partial \xi_j} (T, P; \bar{N}) = \sum_{i=1}^{N} v_{ij} \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) - \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) \]

(9)

These equations indicate that the ReMD algorithm uses a convex approximation (the ideal solution model) that matches the exact values of both \( G(\bar{x}) \) and its gradient at each iteration \( \bar{x} \). In comparison, a Newton–Raphson strategy would also match the exact values of \( G \) and its gradient at each iteration, but would use a linear approximation to \( G(\bar{x}) \) beyond \( \bar{x} \). The general strategy of approximating a complex convex function by a sequence of simpler convex functions was first suggested by Folkman and Shapiro. Its implementation in the context of a macroscopic CRE algorithm using eq 7 was applied to an aqueous electrolyte system of 12 species and 5 elements (including the atomic charge) in Example 7.1 of Smith and Missen.

3. MINOR SPECIES CONCENTRATIONS

Provided that their concentrations are relatively small (“minor species”), the ReMD algorithm can accurately calculate the concentrations of species not included in the original list (thereby “creating new molecules”) using a relatively simple a posteriori procedure by means of a macroscopic thermodynamic device described in Section 9.5 of Smith and Missen. This can also be viewed as a generalization of a somewhat related approach recently used by several research groups for solubility simulations of sparingly soluble solids. The amounts and chemical potentials of the major species at the equilibrium composition calculated by omitting the minor species will be relatively unchanged from their values when the additional species are added to the system. In such cases, we calculate \( \mu_i^{\text{res},\text{NPT}}(T, P; \bar{x}) \) for a single molecule of the minor species \( m \) in the equilibrium mixture, and use eqs 6 and 7 to approximate its chemical potential. We then write the chemical reaction forming 1 mol of the minor species from the major species, and setting its \( \Delta G \) value to zero yields the result.
The JANAF thermochemical tables,\textsuperscript{5} yielding chemical potentials for all species considered are taken from the relevant force-fields. The species which we employ here (see the original papers for text at Several N:O Atomic Ratios Using the Indicated Total Number of Particles).

The values in parentheses in the last row are previous REMC algorithm results\textsuperscript{2,24} and the RxFPMC(BLYP) results of Fetisov et al.\textsuperscript{6} using 192 from eq 22 of the text on each iteration. The row indicated as \( \Delta \) denotes the higher precision calculations described in the text at the approximately converged equilibrium composition. The final row gives the result of the ideal solution extrapolation using these results. Subscripts on the numerical results denote the uncertainties in the indicated number of final digits. Subscripts denote the uncertainties of the indicated values. The values in parentheses in the last row are previous REMC algorithm results\textsuperscript{2,24} and the RxFPMC(BLYP) results of Fetisov et al.\textsuperscript{6} using 192 atoms (see the indicated reference for details).

\[
x_{\text{m}} = \exp \left[ \frac{\sum_{j \neq m} \mu_j (T, P; \bar{k}) - \mu_{\text{m}}^\infty (T, P)}{RT} \right]
\]  

(10)

For each minor species, eq 10 is used in conjunction with the reaction forming the species from the system’s major species to calculate its equilibrium composition. The assumption that the amounts of the hypothesized minor species are small may be tested following the implementation of the indicated procedure, and if a species amount is not small, it is added to the species list and the equilibrium composition quickly recalculated using the equilibrium composition of the major species as the initial estimate.

4. EXAMPLE SYSTEMS

We consider two classic example reacting systems previously considered by others.

(1) The following reacting system at 3000 K and 30 GPa

\[ \text{N}_2 + \text{O}_2 = 2\text{NO} \]  

(11)

was one of the first systems studied by the REMC algorithm.\textsuperscript{2} It was originally considered by Shaw,\textsuperscript{24} and Fetisov et al.\textsuperscript{6} recently considered it using their RxFPMC approach. We consider here calculations at several overall N:O ratios; Shaw,\textsuperscript{24} and Smith and Triska used EXP-6 FF models for the species which we employ here (see the original papers for the relevant force-field parameters). Ideal-gas standard chemical potentials for all species considered are taken from the JANAF thermochemical tables,\textsuperscript{5} yielding

\[
\Delta G_r^0 = 2\mu^0(\text{NO}) - \mu^0(\text{N}_2) - \mu^0(\text{O}_2) = 104,878 \text{ kJ mol}^{-1}
\]  

(12)

for reaction 11. In addition to simulating the equilibrium of the N\textsubscript{2}–O\textsubscript{2}–NO system, we also calculate the compositions of the minor species \{NO\textsubscript{2}, N\textsubscript{2}O, N, O, O\textsubscript{3}\}. The FFs for these species are taken from Fried et al.\textsuperscript{33}

(2) The ammonia synthesis reaction

\[ \text{N}_2 + 3\text{H}_2 = 2\text{NH}_3 \]  

(13)

has been considered by Turner et al.,\textsuperscript{25} Lisal et al.,\textsuperscript{27} and Poursaeidesfahani et al.\textsuperscript{16} All three studies used the same Lennard-Jones-based FFs, which we also use herein (see the papers for the parameters). We consider calculations at several isotherms as a function of pressure for the stoichiometric N:H ratio, using ideal-gas data from the JANAF thermochemical tables.\textsuperscript{5} Numerical values of \( \Delta G_r^0 \) for reaction 13 are given in the second row of Table 3.

5. SIMULATION DETAILS

We used GROMACS 2016.3 for all simulations. For the EXP-6 FF, we used the following \( \lambda \)-scaling scheme\textsuperscript{34} within GROMACS for implementation of the BAR method:

\[
a(\xi) = \lambda^a \frac{\epsilon}{\alpha - 6} \left[ 6 \exp \left( -1 \frac{\xi}{\xi_0} \right) - 4 \left( \frac{\xi}{\xi_0} \right)^6 \right] \left[ \frac{1}{\gamma(1 - \lambda) + \xi^2} \right] \]  

(14)
where \( r_g \) is the interparticle distance, and \( n = 5.5, m = 1.0, \) and \( \gamma = 0.9 \). \( \{ \epsilon, \alpha, r_g \} \) are the FF parameters, whose values are given in the original papers.\(^{24,33}\)

We used between 400 and 408 particles for the \( \text{N}_2-\text{O}_2-\text{NO} \) system, depending on the N:O ratio. For the ammonia synthesis system, the total number of particles used is equivalent to 250 particles of pure NH\(_3\).

The following general procedure was used to implement the ReMD algorithm:

1. The initial composition estimate was obtained by means of an ideal-gas equilibrium calculation (this is for convenience only; any other estimate can be used).
2. At each iteration, the residual chemical potentials are simultaneously calculated (in parallel) for all species using the BAR method,\(^{35}\) using parallel calculations over 20 equally spaced \( \lambda \) intervals.
3. The ideal solution approximation of eqs 6 and 7 is used in eq 3, which is solved for the extent of reaction; the particle numbers on the next iteration are then obtained from eq 2.
4. When the iterations are deemed to have converged, a final iteration is performed using 40 equally spaced \( \lambda \) intervals. The average of a set of 10 runs is used to calculate the residual chemical potentials and their uncertainties (one standard deviation), which are used to obtain a final equilibrium composition from the ideal solution approximation at the final state. The uncertainties (one standard deviation) in the mole fractions are calculated by the method described in the Supporting Information.

6. RESULTS AND DISCUSSION

A. \( \text{N}_2-\text{O}_2-\text{NO} \) System. For the \( \text{N}_2-\text{O}_2-\text{NO} \) system, Step 3 of the algorithm solves the following equation for \( \xi_{\text{eq}} \):

\[
\frac{(n_{\text{NO}}^{(k)} + 2\xi)^2}{(n_{\text{NO}}^{(k)} - \xi)(n_{\text{O}_2}^{(k)} - \xi)} = \exp \left( \frac{-\Delta G_{\text{eq}}^{(k)}}{RT} \right)
\]

(15)

The iterations of the algorithm for several N:O atomic ratios are shown numerically in Table 1. Convergence is seen to be very rapid, requiring at most two iterations, followed by a final iteration obtained from more precise values of the residual chemical potentials. The equilibrium compositions in the final rows for each ratio are in good agreement with previously obtained results for this system shown in parentheses\(^{22,24}\) and in slightly poorer agreement with the RxFPMC results of Fetisov et al.\(^6\)

The progress of the iterations for the representative 1:1 N:O atomic ratio is shown graphically in Figures 1 and 2. The solid curves are not needed for the algorithm implementation, but are shown to demonstrate the accuracy of the ideal solution extrapolation of eqs 8 and 9 that uses the initial estimate as the reference state for the ideal solution chemical potential approximation. This approximation gives results for \( G \) and \( \Delta G \) that are almost indistinguishable from the exact curves. A significant source of this accuracy is that the approximating functions incorporate the logarithmic form of compositional dependence of the correct functions. For example, the S-shaped \( \Delta G \) curves shown in Figure 2 are generic, in the sense that their logarithmic dependence on the species compositions results in limiting values of \( -\infty \) when product species approach zero amounts (\( \xi = 0 \)) and \( +\infty \) when reactant species approach zero amounts (\( \xi = 1 \)).

For the \( \text{N}_2-\text{O}_2-\text{NO} \) system, the RxFPMC algorithm of Fetisov et al.\(^5\) requires the selection of an approximate DFT method for implementing the calculation of the interatomic

Figure 1. Exact and ideal solution approximation for the dimensionless relative Gibbs energy for the reacting system \( \text{N}_2 + \text{O}_2 = 2\text{NO} \) at \( T = 3000 \text{ K} \) and \( P = 30 \text{ GPa} \) with an atomic ratio of \( \text{N}:\text{O} = 1:1 \) and a total of 400 particles on the scaled extent of reaction parameter, \( \eta = n(\text{NO}) / 400, \) where \( n(\text{NO}) \) is the number of particles of NO. \( G(\xi') \) is the Gibbs energy of the system at \( \xi' = 0 \), and \( G(0) \) is the Gibbs energy at \( \xi' = 0 \), corresponding to 200 particles of \( \text{N}_2 \) and 200 particles of \( \text{O}_2 \). The points are simulation results for directly simulated values of \( G(\xi') \) with their indicated uncertainties (one standard deviation), and the solid curve drawn through them is shown as an aid to the eye. The dashed curve shows \( G(\xi') \) for the ideal solution approximation to \( G(\xi') \) of eqs 1 and 7 using the initial estimate as the reference state, which is the equilibrium composition of an ideal-gas model of the system (\( \xi = 0.06 \), corresponding to 24 particles of NO and 188 particles of each of \( \text{N}_2 \) and \( \text{O}_2 \)). The vertical dashed lines show the progress of the ReMD algorithm from the initial estimate to the first iteration. The equilibrium solution for the system is indicated by the arrow.

Figure 2. Dependence of the dimensionless reaction free energy change \( \Delta G = \mu(\text{NO}) - \mu(\text{N}_2) + \mu(\text{O}_2) \) on the scaled extent of reaction parameter, \( \xi' \), for the system described in Figure 1. The solid and dashed curves show the gradients with respect to \( \xi' \) of the functions described by the corresponding curves shown in that figure. The simulation uncertainties lie within the symbol sizes.
forces, and a method for analyzing the equilibrium atomic distributions to infer that particular configurations indicate separate molecular aggregates. Using four different DFT methods and an interatomic distance criterion indicating molecular aggregates based on the location of the first minimum of the radial distribution function, they obtained equilibrium compositions reasonably compatible with our results. Their indicated uncertainties for each DFT method are calculated from 32 independent simulation runs. The spread of these results arises from the different DFT approximations and shows much larger overall uncertainties in the computed compositions; the corresponding spread of our results reflects the use of different force fields within an essentially exact calculation. Overall, our results using the Fried FF are more consistent with those obtained using the Cheetah equation of state (EOS) than with those obtained using the Shaw force field. This is consistent with the fact that the Fried Exp-6 FF results are from Fetisov et al., and the Cheetah results are obtained from the Cheetah 8.0 thermochemical software package as quoted in Fetisov et al., which employs an equation of state whose pure fluids are tailored to the EXP-6 FF and dipolar interactions between molecular gas products.

We calculate the equilibrium composition for each reaction using eq 10 for each minor species keeping the residual chemical potentials of the (major) reactant species fixed at their previously calculated equilibrium values. For example, for N_2O_3, we solve the following equation for \( \xi \), where \( \mathbf{x} \) denotes the previously calculated equilibrium composition of the major species:

\[
\begin{align*}
\mu^0_{\text{N}_2\text{O}_3}(T) + RT \ln \left( \frac{P}{P^0} \right) + \mu^\text{N}_2\text{O}_3^\text{N,P,T}(T, P; \mathbf{x}) + RT \ln \left( \frac{\xi}{n_i - \xi/2} \right) \\
= \mu^0_{\text{N}_2}(T) + RT \ln \left( \frac{P}{P^0} \right) + \mu^\text{N}_2^\text{N,P,T}(T, P; \mathbf{x}) + RT \ln \left( \frac{n_i - \xi}{n_i - \xi/2} \right) \\
+ \frac{1}{2} \left[ \mu^0_{\text{O}_2}(T) + RT \ln \left( \frac{P}{P^0} \right) + \mu^\text{O}_2^\text{N,P,T}(T, P; \mathbf{x}) \\
+ RT \ln \left( \frac{n_i - \xi/2}{n_i - \xi/2} \right) \right] \\
+ \frac{3}{2} \mu^0_{\text{O}_3}(T) + RT \ln \left( \frac{P}{P^0} \right) + \mu^\text{O}_3^\text{N,P,T}(T, P; \mathbf{x}) + RT \ln \left( \frac{n_i - \xi/2}{n_i - \xi/2} \right) \\
\end{align*}
\]
The resulting mole fractions of \( N_2, O, \) and \( O_3 \) were very small (less than \( 10^{-4} \) in all cases), but those of \( NO_2 \) and \( N_2O \) were both near 0.05; it was deemed that these species should be included in the equilibrium calculations along with \( N_2, O_2, \) and \( NO \). The equilibrium composition of the resulting system with 3 reactions was then determined, requiring one or two additional iterations in all cases. At convergence, the residual chemical potentials arising from adding one particle each of \( N, O, \) and \( O_3 \) were again calculated and their equilibrium mole fractions determined. The equilibrium compositions of all species are shown in Table 2, and compared with the results of Fetisov et al.,6 and with those of the Cheetah thermochemical software program, quoted in Fetisov et al.6

We remark in passing that the ability of the RxFPMC algorithm to detect the presence of minor species in a CRE calculation is limited by the length of the simulation run. The RxFPMC simulations of Fetisov et al.6 involved 96 or 192 atoms, and they averaged their results over 32 independent simulations. This only permitted concentrations down to 0.001 to be detected. The ReMD approximation method for minor species has no such limitations.

### B. Ammonia Synthesis Reaction System

For the ammonia synthesis reaction of eq 13, the ideal-gas values \( \Delta G_i^0(T) \) are given in the upper portion of Table 3. Our simulations converged in at most 3 iterations, similarly as in the case of the \( N_2-O_2-NO \) system. Numerical results are given in Table 3 at several isotherms as a function of pressure in the case of the stoichiometric N:H ratio, and are compared in Figure 3 with those of other research groups. All three sets of results are very similar. Our values are seen to lie between those of Turner et al.,25 and Poursaeidesfahani et al.,16 consistent with the relative values of the ideal-gas quantities \( \Delta G_i^0(T) \) shown in Table 3.

#### Table 3. Ideal-Gas Standard Free Energy Changes \( \Delta G_i^0(T; P^0) \) in kJ mol\(^{-1}\) and Equilibrium Concentrations of Species for the Ammonia Synthesis Reaction of Eq 13 at the Indicated Temperatures and Pressures Using the Methodology of This Work \[\Delta G_i^0(T) = 2\mu(NH_3) - \mu(N_2) - 3\mu(H_2)\]

| \( T \) (K) | \( P \) (bar) | \( x(NH_3) \) |
|-----------|-------------|--------------|
| 100       | 0.54620     | 0.26192      |
| 200       | 0.68375     | 0.39322      |
| 300       | 0.76397     | 0.48306      |
| 400       | 0.81818     | 0.55369      |
| 500       | 0.85661     | 0.61189      |
| 600       | 0.88080     | 0.65838      |
| 700       | 0.90677     | 0.69999      |
| 800       | 0.92277     | 0.73287      |
| 900       | 0.93485     | 0.76251      |
| 1000      | 0.94215     | 0.78645      |

### 7. CONCLUSIONS

We have introduced a new molecular-mechanism simulation-algorithm for calculating chemical reaction equilibrium, which we call the reaction ensemble molecular dynamics (ReMD) algorithm. We have illustrated its use in the case of specified \((T, P)\) problems and provided examples. It minimizes the system Gibbs free energy based on iterations involving calculations of species residual chemical potentials at a sequence of reference state composition, and their subsequent extrapolation using a macroscopic ideal solution chemical potential model, whose equilibrium composition may be rapidly calculated to yield the next iteration.

The species residual chemical potentials are calculated by means of the BAR method within a standard MD simulation package; we have used GROMACS 2016.3 for illustrative purposes, which implements multicore parallelized calculations for all particle trajectories. Each iteration can also be performed rapidly by parallelizing both the set of species chemical potentials and their increments over each \( \lambda \) window within the BAR method.

The algorithm is general and can be readily implemented for any reacting system independently of the complexity of the molecules composing it; it requires only a list of chemical species and their ideal-gas free energy data, in addition to force-field models for the molecules. The ReMD algorithm thus alleviates the difficulties encountered by the REMC algorithm in the calculation of chemical reaction equilibrium.
for challenging systems, which require the incorporation of special system-dependent Monte Carlo approaches\textsuperscript{11–18} to ensure convergence.

We have also described a general procedure for calculating the equilibrium compositions of species that have not been included in the species list and are present in very small amounts (called "minor species"), which entails an \textit{a posteriori} calculation of their chemical potentials due to the presence of a single molecule in the converged equilibrium system involving the major species.

Both main features of the algorithm (the iterative ideal solution extrapolation and the minor species calculation) are based on similar approaches previously used in the case of equilibrium algorithms employing macroscopic thermodynamic models.

We tested the algorithm for two classic benchmark reaction equilibrium problems at specified (\(T, P\)) considered by previous workers: the \(\text{N}_2–\text{O}_2–\text{NO}\) system and the ammonia synthesis system. Our results are in agreement with those of other workers for these systems, and convergence is achieved in most cases within two iterations.

\section*{ASSOCIATED CONTENT}

\subsection*{Supporting Information}

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscentsci.8b00361.

Derivation of the uncertainties in the equilibrium mole fractions calculated by the ReMD algorithm (PDF)

\section*{AUTHOR INFORMATION}

\subsection*{Corresponding Author}

*E-mail: bilsmith@uoguelph.ca.

\section*{ORCID}

William R. Smith: 0000-0002-1982-2050

\section*{Notes}

The authors declare no competing financial interest.

\section*{ACKNOWLEDGMENTS}

Support for this work was provided by the Natural Sciences and Engineering Research Council of Canada (Strategic Program Grant STP GP 479466-15), and by the SHARCNET (Shared Hierarchical Academic Research Computing Network) HPC consortium (www.sharcnet.ca) and Compute Canada (www.computecanada.ca).

\section*{REFERENCES}

(1) Leal, A. M. M.; Kulik, D. A.; Smith, W. R.; Saar, M. O. An overview of computational methods for chemical equilibrium and kinetic calculations for geochemical and reactive transport modeling. \textit{Pure Appl. Chem.} \textbf{2017}, \textit{89}, 597–643.

(2) Smith, W. R.; Triska, B. The reaction ensemble method for the computer simulation of chemical and phase equilibria. I. Theory and basic examples. \textit{J. Chem. Phys.} \textbf{1994}, \textit{100}, 3019–3027.

(3) Johnson, K. J.; Panagiotopoulos, Z. A.; Gubbins, E. K. A new simulation technique for reacting or associating fluids. \textit{Mol. Phys.} \textbf{1994}, \textit{81}, 717–733.

(4) Turner, C. H.; Brennan, J. K.; Lisal, M.; Smith, W. R.; Karl Johnson, J.; Gubbins, K. E. Simulation of chemical reaction equilibria by the Reaction Ensemble Monte Carlo method: A review. \textit{Mol. Simul.} \textbf{2008}, \textit{34}, 119–146.

(5) Chase, M. W., Jr. Thermochemical Tables, Fourth ed., Parts I and II. In \textit{J. Phys. Chem. Ref. Data, Monograph #9}, American Chemical Society and American Institute of Physics: Woodbury, NY, 1998.

(6) Fetisov, E. O.; Kuo, I. F.; Knight, C.; VandeVondele, J.; Van Voorhis, T.; Siepmann, J. I. First-Principles Monte Carlo simulations of reaction equilibria in compressed vapors. \textit{ACS Cent. Sci.} \textbf{2016}, \textit{2}, 409–415.

(7) Han, Y.; Jiang, D.; Zhang, J.; Li, W.; Gan, Z.; Gu, J. Development, applications and challenges of ReaxFF reactive force field in molecular simulations. \textit{Front. Chem. Sci. Eng.} \textbf{2016}, \textit{10}, 16–38.

(8) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grana, A.; van Duijn, A. C. T. The ReaxFF reactive force-field: development, applications and future directions. \textit{npj Computational Materials} \textbf{2016}, \textit{2}, 15011.

(9) Le, T. T. B.; Striolo, A.; Cole, D. R. Structural and dynamical properties predicted by reactive force fields simulations for four common pure fluids at liquid and gaseous non-reactive conditions. \textit{Mol. Simul.} \textbf{2018}, \textit{44}, 826.

(10) Smith, W. R.; Missen, R. W. Chemical Reaction Equilibrium Analysis: Theory and Algorithms; Krieger Publishing Co.: Reprint of same title; Willey-Interscience: Malabar, FL, 1991.

(11) Brennan, J. K. Cavity-bias sampling in reaction ensemble Monte Carlo simulations. \textit{Mol. Phys.} \textbf{2005}, \textit{103}, 2647–2654.

(12) Turner, C. H.; Brennan, J. K.; Lisal, M. Replica Exchange for Reactive Monte Carlo Simulations. \textit{J. Phys. Chem. C} \textbf{2007}, \textit{111}, 15706–15715.

(13) Hansen, N.; Jakobtorweihen, S.; Keil, F. J. Reactive Monte Carlo and grand-canonical Monte Carlo simulations of the propene metathesis reaction system. \textit{J. Chem. Phys.} \textbf{2005}, \textit{122}, 164705.

(14) Jakobtorweihen, S.; Hansen, N.; Keil, F. J. Combining reactive and configurational-bias Monte Carlo: confinement influence on the propene metathesis reaction system in various zeolites. \textit{J. Chem. Phys.} \textbf{2006}, \textit{125}, 224709.

(15) Rosch, T. W.; Maginn, E. J. Reaction Ensemble Monte Carlo simulation of complex molecular systems. \textit{J. Chem. Theory Comput.} \textbf{2011}, \textit{7}, 269–279.

(16) Poussaidesfahani, A.; Hens, R.; Rahbari, A.; Ramdin, M.; Dubbeldam, D.; Flugt, T. J. H. Efficient application of continuous fractional component Monte Carlo in the Reaction Ensemble. \textit{J. Chem. Theory Comput.} \textbf{2017}, \textit{13}, 4452–4466.

(17) Leiding, J.; Coe, J. D. Reactive Monte Carlo sampling with an ab initio potential. \textit{J. Chem. Phys.} \textbf{2016}, \textit{144}, 174109.

(18) Mullen, R. G.; Maginn, E. J. Reaction ensemble Monte Carlo simulation of xylene isomerization in bulk phases and under confinement. \textit{J. Chem. Theory Comput.} \textbf{2017}, \textit{13}, 4054–4062.

(19) Pronk, S.; Pall, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. \textit{Bioinformatics} \textbf{2013}, \textit{29}, 845–854.

(20) Edwards, S. \textit{History of Processor Performance}. http://www.cs. columbia.edu/~sedwards/classes/2012/3827-spring/advanced-arch-2011.pdf (accessed Aug 5, 2018).

(21) Stone, J. E.; Hardy, D. J.; Ulfimtsev, I. S.; Schultz, K. GPU-accelerated molecular modeling coming of age. \textit{J. Mol. Graphics Modell.} \textbf{2010}, \textit{29}, 116–125.

(22) Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E. Biomolecular simulation: a computational microscope for molecular biology. \textit{Annu. Rev. Biophys.} \textbf{2012}, \textit{41}, 429–452.

(23) Le Grand, S.; Götz, A. W.; Walker, R. C. SPPF: Speed without compromise-A mixed precision model for GPU accelerated molecular dynamics simulations. \textit{Comput. Phys. Commun.} \textbf{2013}, \textit{184}, 374–380.

(24) Shaw, M. S. Monte Carlo simulation of equilibrium chemical composition of molecular fluid mixtures in the N atoms PT ensemble. \textit{J. Chem. Phys.} \textbf{1991}, \textit{94}, 7550–7553.

(25) Turner, C. H.; Johnson, J. K.; Gubbins, K. E. Effect of confinement on chemical reaction equilibria: The reactions 2NO ⇔ 2N + O.
(NO)₂ and N₂+3H₂ ⇔ 2NH₃ in carbon micropores. J. Chem. Phys. 2001, 114, 1851.

(26) Brennan, J. K.; Lisal, M.; Gubbins, K. E.; Rice, B. M. Reaction ensemble molecular dynamics: direct simulation of the dynamic equilibrium properties of chemically reacting mixtures. Phys. Rev. E 2004, 70, 061103.

(27) Lisal, M.; Bendová, M.; Smith, W. R. Monte Carlo adiabatic simulation of equilibrium reacting systems: The ammonia synthesis reaction. Fluid Phase Equilib. 2005, 235, 50–57.

(28) Folkman, J.; Shapiro, N. Z. Approximating one convex function by another. SIAM J. Appl. Math. 1968, 16, 993–997.

(29) Paluch, A. S.; Parameswaran, S.; Liu, S.; Kolavennu, A.; Mobley, D. L. Predicting the excess solubility of acetonilide, acetaminophen, phenacetin, benzocaine, and caffeine in binary water/ethanol mixtures via molecular simulation. J. Chem. Phys. 2015, 142, 044508.

(30) Noroozi, J.; Ghobat, C.; Sardroodi, J. J.; Karimi-Sabet, J.; Robert, M. A. Solvation free energy and solubility of acetaminophen and ibuprofen in supercritical carbon dioxide: Impact of the solvent model. J. Supercrit. Fluids 2016, 109, 166–176.

(31) Noroozi, J.; Paluch, A. S. Microscopic structure and solubility predictions of multifunctional solids in supercritical carbon dioxide: A molecular simulation study. J. Phys. Chem. B 2017, 121, 1660–1674.

(32) Li, L.; Totton, T.; Frenkel, D. Computational methodology for solubility prediction: Application to the sparingly soluble solutes. J. Chem. Phys. 2017, 146, 214110.

(33) Fried, L. E.; Howard, W. M.; Souers, P. C. EXP6: A New Equation of State Library for High Pressure Thermochemistry. 12th International Detonation Symposium, San Diego, CA; 2002.

(34) Sellers, M. S.; Lisal, M.; Brennan, J. K. Exponential-six potential scaling for the calculation of free energies in molecular simulations. Mol. Phys. 2015, 113, 45–54.

(35) Bennett, C. H. Efficient estimation of free energy differences from Monte Carlo data. J. Comput. Phys. 1976, 22, 245–268.

(36) Gillespie, L. J.; Beattie, J. A. The thermodynamic treatment of chemical equilibria in systems composed of real gases. I. An approximate equation for the mass action function applied to the existing data on the Haber equilibrium. Phys. Rev. 1930, 36, 743–753.