Breaking the curse of dimensionality: Solving configurational integrals for crystalline solids by tensor networks

Duc P. Truong [1], Benjamin Nebgen [1], Derek DeSantis [1], Dimiter N. Petsev [2], Kim Ø. Rasmussen [1], and Boian S. Alexandrov [1]

[1] Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

[2] Department of Chemical and Biological Engineering, and Center for Microengineered Materials, University of New Mexico, Albuquerque, New Mexico 87131, USA

(Received 28 May 2025; accepted 13 August 2025; published 27 August 2025)

Accurately evaluating configurational integrals for dense solids remains a central and difficult challenge in the statistical mechanics of condensed systems. Here, we present a tensor network approach that reformulates the high-dimensional configurational integral for identical-particle crystals into a sequence of computationally efficient summations. We represent the integrand as a high-dimensional tensor and apply tensor-train (TT) decomposition together with a custom TT-cross interpolation. This approach circumvents the need to explicitly construct the full tensor. We introduce tailored rank-1 and rank-2 schemes optimized for sharply peaked Boltzmann probability densities, typical for identical-particle crystals. When applied to the calculation of internal energy and pressure-temperature curves for crystalline Cu and Ar at high (GPa) pressures, as well as the alpha-to-beta phase transition diagram of Sn, our method accurately reproduces molecular dynamics simulation results using tight-binding, machine learning, hierarchical interacting particle–neural network, and modified embedded atom method potentials,all within seconds of computation time.

DOI: 10.1103/xrbw-xr49

I. INTRODUCTION

For over a century, physicists have sought to rigorously elucidate the emergence of macroscopic properties of matter from the underlying microscopic dynamics of atoms and molecules. This foundational link, formalized by 19th-century statistical mechanics, continues to drive modern computational approaches that provide systematic frameworks for deriving thermodynamic behavior across scales from quantum mechanical to macroscopic systems. The foundations of statistical mechanics were established through the seminal contributions of Maxwell [1], Boltzmann [2], and Gibbs [3]. It was Gibbs who introduced the canonical partition function in statistical mechanics, a central construct that characterizes a system of N particles in three-dimensional space using a statistical ensemble formulation:

$$Q_N = \frac{1}{N!\Lambda^{3N}} Z_N, \quad Z_N = \int_{\mathbb{R}^{3N}} e^{-\beta U(\mathbf{q})} d\mathbf{q}, \tag{1}$$

where $\mathbf{q} = [x_1, y_1, z_1, \ldots, x_N, y_N, z_N]$ are the particle coordinates, and $U(\mathbf{q})$ is the multiparticle interaction potential. The parameter $\Lambda = h/(2\pi m k_B T)^{1/2}$ is the de Broglie wavelength, where h is Planck's constant, k_B is Boltzmann's constant, $\beta = 1/k_B T$ where T is the temperature, and m is the particle mass. The quantity Z_N is the configurational integral, which accounts for the interactions among particles (atoms and/or molecules) via their potential energy. The evaluation of the configurational integral in systems with strong interparticle interactions remains one of the central challenges in the statistical mechanics of condensed matter.

The physical interpretation of the partition function and the configurational integral in Eq. (1) is conceptually straightforward: it represents a summation over all configurations compatible with the system's intermolecular interactions and macroscopic thermodynamic constraints. However, the mathematical complexity escalates dramatically in condensed-phase systems, such as liquids and solids, where interactions between atoms or molecules dominate system behavior. In macroscopic systems, the number of constituent particles can exceed 10^{23} making the computation of total potential energy exceedingly challenging. In addition, the configuration integral Z_N [see Eq. (1)] is $3N$-fold, and traditional numerical techniques are not well-suited to deal with such large-scale problems. For example, a simple midpoint numerical integration approach to solve a 100-fold integral (dimensionality $d = 100$) on $n + 1$ nodes in each direction requires n^d function evaluations. This exponential increase in computational complexity is known as the curse of dimensionality [4]. Consider the example of $n = d = 100$ integration nodes. This choice requires 10^{200} function evaluations and similar number of floating-point operations. Such a computation would need at least 2×10^{188} seconds on a teraflop computer capable of 10^{12} floating point operations per second. This time is roughly equal to 4.58×10^{170} times the age of the Universe [5].

These computational bottlenecks have driven transformative methodological advances throughout the 20th and 21st

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

centuries. Kirkwood's coupling parameter approach [6] and later Bennett's acceptance ratio method [7] established rigorous foundations for free energy calculations, while pioneering Monte Carlo methods [8] opened new avenues for studying phase transitions. The development of enhanced sampling techniques—including replica exchange [9], metadynamics [10], and umbrella sampling [11]—has enabled exploration of rare events and complex energy landscapes. The integration of quantum mechanics through Car-Parrinello molecular dynamics [12] and modern density functional theory calculations [13] revolutionized our understanding of chemical processes, while recent advances in machine learning potentials [14] and theorems addressing nonequilibrium [15] have further expanded the accessible time and length scales. While these developments, comprehensively reviewed in the seminal works [16,17], have transformed computational statistical mechanics each method remains constrained by inherent limitations in accuracy, accessible time scales, or convergence properties, see for example, Refs. [18,19]. Recent advances in tensor networks methods and specifically of the Tensor Train (TT) network [20–23] have demonstrated their effectiveness in dramatically reducing the complexity of high-dimensional integrals. By reformulating previously intractable problems into sequences of low-dimensional contractions, the TT format offers a powerful numerical alternative to traditional quadrature or sampling-based methods. A substantial body of research has explored tensor network contraction and renormalization techniques for computing partition functions and thermodynamic observables in lattice-based models of spin-systems [24–33].

In this work, we build upon these developments to provide a scalable and accurate approach for computing configurational integrals in dense solids, with relevance to thermodynamic modeling in physics and chemistry.

We have previously demonstrated that applying the TT format to high-dimensional integrals yields exceptionally high accuracy, dramatic speed-ups, and extreme compression, outperforming traditional Monte Carlo methods by achieving up to a 1000× speed up, and several orders of magnitude better accuracy [5]. Similarly, when applied to partial differential equations, TT-based methods consistently outperform both standard numerical schemes and state-of-the-art solvers - achieving yottabyte compression in some cases [34]. Furthermore, we also have investigated how TT-rank selection impacts solution accuracy, which in turn affects both computational efficiency and compression rates [35].

This work applies the TT format to solve configurational integrals in three-dimensional (3D) crystals. Our methodology involves developing a TT-cross interpolation technique that represents the configurational integral's integrand as a high-dimensional tensor. By combining TT decomposition with custom TT-cross interpolation, the approach avoids explicitly constructing the full tensor—a computationally prohibitive task. The framework incorporates TT rank-1 and rank-2 schemes designed for sharply peaked Boltzmann probability densities characteristic of identical-particle crystals. Testing on internal energy calculations and pressure-temperature curves for crystalline Cu and Ar, plus alpha-to-beta phase transition analysis of Sn, demonstrates accurate reproduction of molecular dynamics results across tight-binding, machine learning, hierarchical interacting particle–neural network (HIP-NN), and modified embedded atom method (MEAM) potentials—achieving these results in few seconds. Our selection of TT topology stems from the custom cross interpolation scheme we introduced here.

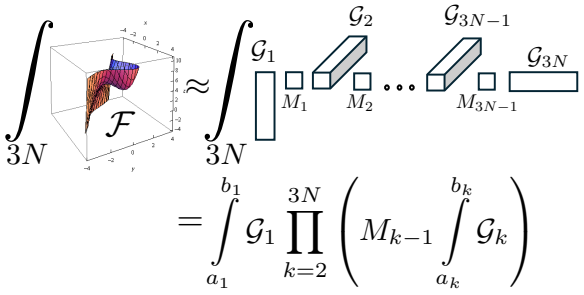


FIG. 1. Schematic illustration of the tensor-train (TT) decomposition of the Boltzmann tensor $\mathcal{F}$ and its application to computing the configurational integral. The TT format expresses $\mathcal{F}$ as a sequence of low-rank tensor cores $\mathcal{G}_k$ connected by intersection matrices $\mathbf{M}^k$. This structured representation enables the high-dimensional integral over $\mathcal{F}$ to be reduced to a product of one-dimensional on the interval $[a_k, b_k]$, significantly improving computational efficiency.

II. METHODS

A. Tensor formulation of Z_N

The integrand in Z_N is a $3N$-dimensional function $f(\mathbf{q}) := e^{-\beta U(\mathbf{q})}$. To solve the integral, this function must be numerically evaluated in a set of n^{3N} grid points, where n is the number of points in each dimension. This evaluation yields the Boltzmann factor, the array $\mathcal{F}(i_1, \ldots, i_{3N})$. The utilization of the cubature rule approximates Z_N by weighted summations over all dimensions give us

$$Z_N \approx \sum_{i_1=1}^{n} w_{i_1} \cdots \sum_{i_{3N}=1}^{n} w_{i_{3N}} \mathcal{F}(i_1, i_2, \ldots, i_{3N}). \tag{2}$$

It is important to note that the sheer size of the tensor $\mathcal{F}$ makes it impossible to construct explicitly.

B. Tensor train decomposition

TT decomposition provides an approximation that compresses high-dimensional tensors, formatting them as a product of small 3D tensor cores, ordered in a linear chain [36]. Below, we adopt a slightly modified TT format of the Boltzmann factor $\mathcal{F}$, as illustrated in Fig. 1:

$$\begin{aligned}\mathcal{F}(i_1, \ldots, i_{3N}) = \sum_{\alpha}^{\mathbf{r}} \mathcal{G}_1(1, i_1, \alpha_1)\mathbf{M^1}(\alpha_1, \alpha_1) \\ \mathcal{G}_2(\alpha_1, i_2, \alpha_2)\mathbf{M^2}(\alpha_2, \alpha_2) \\ \cdots \mathbf{M^{3N-1}}(\alpha_{3N-1}, \alpha_{3N-1})\mathcal{G}_{3N}(\alpha_{3N-1}, i_{3N}, 1) + \varepsilon.\end{aligned}$$

The components of the vector $\mathbf{r} = [r_1, \ldots, r_{3N-1}]$, which are the ranges of the indices α_i, are called TT ranks, and ε is the approximation error of the decomposition. Each of the 3D core tensors, $\mathcal{G}_k$, is uniquely associated with a single index of the original tensor $\mathcal{F}$, establishing the TT decomposition

as a discrete manifestation of separation of variables. The matrices $\mathbf{M}^k$ are referred to as *intersection matrices* and are used to couple adjacent TT cores, capturing key interactions across dimensions.

C. Computing Z_N from TT format

Once the tensor $\mathcal{F}$ is in the TT format, we can calculate Z_N by,

$$Z_N \approx \prod_{k=1}^{3N} \mathbf{R_k}, \tag{3}$$

where: $\mathbf{R_1}$ is a row vector of size r_1; $\mathbf{R_k}$, $k = 2, \ldots, 3N-1$ are matrices of size $r_{k-1} \times r_k$;

$$\mathbf{R_k} = \sum_{i_k=1}^{n_k} w_{i_k}^k \mathcal{G}_k(:, i_k, :)\mathbf{M^k}, \tag{4}$$

as illustrated schematically in Fig. 1. This formulation reduces the original $3N$-dimensional integral to a sequence of $3N$ one-dimensional weighted summations, exploiting the structure of the TT decomposition to enable efficient evaluation. The key computational challenge lies in constructing the TT format of $\mathcal{F}$ without accessing all of its entries.

D. TT-cross interpolation

TT cross is an algorithm for TT decomposition, needed when accessing all entries in a tensor $\mathcal{F}$ is computationally impossible [21]. Specifically, instead of performing consecutive Singular Value Decompositions (SVDs) and reshapings of the unfolding of the tensor $\mathcal{F}$ (as in the TT-SVD algorithm [36]), TT -cross adaptively samples a small subset (arguably the most important) of r columns and r rows of the unfolding, based on the ideas of pseudoskeleton [37] and **CUR** [38] decompositions. In this way TT cross selects a submatrix with maximum volume (i.e., maximum absolute determinant), which ensures an optimal accuracy [36,37] of the approximation. Since computing the exact submatrix with the maximum volume is NP-hard, practical implementations rely on the heuristic quasimaxvol algorithms [39].

When calculating the **CUR** decomposition, it is necessary to choose the target rank r, which depends on the structure of the Boltzmann tensor and directly affects the accuracy of the computed partition function Z_N. Specifically, constructing a submatrix with maximal volume—and thus achieving an accurate CUR approximation—requires a reasonable estimate of r.

The **CUR** is a matrix decomposition which approximate the matrix **A** by the product of subset of r of its columns **C** and rows **R**, and the pseudoinverse of their intersection **U**: $\mathbf{A} \approx \mathbf{CUR}$. In block-matrix form,

$$\begin{bmatrix} \mathbf{A}_{11}^{r\times r} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \approx \begin{bmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21} \end{bmatrix} \mathbf{A}_{11}^{-1} [\mathbf{A}_{11} \quad \mathbf{A}_{12}] = \mathbf{CUR}.$$

The r columns and rows can be selected by the principle of maximum volume (maxvol) [40], that is, through searching for a submatrix $\mathbf{A}_{11}$ with maximum absolute determinant.

The error of the approximation can be evaluated by comparison with SVD. The SVD which is finding the vectors (typically obtained after truncating the SVD to some small number r of terms) providing the best rank-r approximation to the matrix **A**. If $r \leqslant R = \mathrm{rank}(\mathbf{A})$ and we define $\mathbf{A}_r = \sum_{i=1}^{r} \sigma_i u_i v_i^T$, where σ_i, u, and v are the singular values and singular vectors, then $\|\mathbf{A}–\mathbf{A}_r\|_F^2 = \min\|\mathbf{A}–\mathbf{X}\|_F^2$, for all **X** with $\mathrm{rank}(\mathbf{X}) \leqslant r$. Drineas, Kannan, and Mahoney constructed an additive-error **CUR** matrix decomposition by randomly construct a matrix **C** of columns, a matrix **R** of rows, and a matrix **U** such as, $\|\mathbf{A}–\mathbf{CUR}\|_F^2 \leqslant \|\mathbf{A}–\mathbf{A}_r\|_F^2 + \varepsilon\|\mathbf{A}\|_F^2$, where ε is a small number. The error of **CUR** approximation depends on the rank r, which is related to the rank r of the truncated SVD of **A**. Thus, when the spectrum of **A** decays slowly, i.e., the singular values σ_k for $k > r$ are not negligible to achieve a better accuracy requires a higher rank r, and several rank-adaptive algorithms [41], such as density martrix renormalization group [42] and AMEn-cross [43], are used in TT cross.

E. TT cross of Z_N

TT-cross interpolation circumvents the need for full access to the tensor $\mathcal{F}$'s elements. However, the Boltzmann factor poses significant challenges to TT cross: (i) Since most of the tensor mass is concentrated near the minimal energy configuration $\bar{\mathbf{q}}$, TT cross may miss the associated sharp peak in $\mathcal{F}$, leading to large errors; (ii) TT cross uses algorithms for sampling of tensor fibers; they may not resolve the peak unless the grid aligns well with $\bar{\mathbf{q}}$. These problems, related to sharply concentrated probability densities, were highlighted in recent works on TT inverse Rosenblatt transport maps [44]. Consequently, applying standard TT cross to the Boltzmann factor frequently results in tensor train decompositions with undesirably high ranks or large errors. To resolve these issues, we sample the integrant based on the symmetry of the Boltzmann factor.

F. Discretization of the Boltzmann factor

Assuming that the exponential function, $f(\mathbf{q})$, favors a unique dominant peak at $\bar{\mathbf{q}} = [\bar{q}_1, \ldots, \bar{q}_{3N}]$, the integrand will be maximized at these coordinates. To calculate Z_N, we apply a cubature rule. For each dimension $k = 1, \ldots, 3N$, we choose the integration domain $[a_k, b_k]$ to be symmetric around $\bar{q}_k$. This choice generates a grid of size n^{3N}, with $\bar{\mathbf{q}}$, at its center. Importantly, due to the translational and permutational invariance of the potential, the resulting discretization of $\mathcal{F}$ on this grid displays a superdiagonal dominance. Specifically, for all indices i, $\mathcal{F}(i, i, ..., i) = f(\bar{q}_1, \bar{q}_2, ..., \bar{q}_{3N}) = \bar{f}$, where the coordinates along the superdiagonal correspond to configurations preserving the relative positions of particles in the system.

G. Rank-1 TT cross of the Z_N integrand

In Fig. 2(a), we illustrate a rank-1 TT cross for a 1D system, with two particles with coordinates $\mathbf{q} = [q_1, q_2]$. We assume that the system possesses a global minimum at $\bar{\mathbf{q}} = [\bar{q}_1, \bar{q}_2]$, maximizing the integrand to $\bar{f} = f(\bar{\mathbf{q}})$. Integration is performed over $[a_1, b_1] \times [a_2, b_2]$, chosen to be symmetric around $[\bar{q}_1, \bar{q}_2]$. For rank-1, the maximum volume is reduced to the task of finding the maximum value of the function $f(\mathbf{q})$. This value, $\bar{f}$ occurs at the cross section of two specific fibers

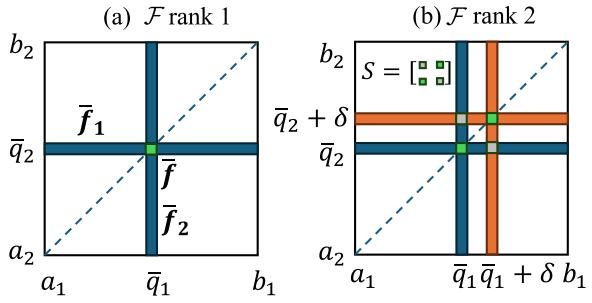


FIG. 2. TT-cross interpolation of the Boltzmann factor. (a) Rank-1 and (b) Rank-2.

of the tensor $\mathcal{F}$. The two blue fibers (one row and one column) in Fig. 2(a), represent these two fibers: $\bar{\mathbf{f}}_1$ and $\bar{\mathbf{f}}_2$ of the 2D tensor $\mathcal{F}$. We formulate the following result (see the proof in Appendix A):

Let the integrand $f(\mathbf{q}) = e^{-\beta U(\mathbf{q})}$ attain its global maximum at $\bar{\mathbf{q}}$, which corresponds to the multi-index $\mathbf{j_1} = [j_1^1, \ldots, j_1^{3N}]$ on the superdiagonal of the tensor $\mathcal{F}$, and $j_1^1 = \ldots = j_1^{3N}$. Let $\bar{f} := \mathcal{F}(\mathbf{j}_1)$ denote its value at that point. For each dimension k, we define the fiber $\bar{\mathbf{f}}_k \in \mathbb{R}^n$ as the k-mode fiber passing through the index $\mathbf{j}_1$. This set of fibers, $\bar{\mathbf{f}}_k(i_k) = \mathcal{F}(j_1^1, \ldots, j_1^{k-1}, i_k, j_1^{k+1}, \ldots, j_1^{3N})$ for $i_k = 1, \ldots, n$, intersects at $\bar{\mathbf{q}}$ with indices $\mathbf{j_1}$. Then, the rank-1 TT-cross interpolation of $\mathcal{F}$ is

$$\hat{\mathcal{F}} = \bar{\mathbf{f}}_\mathbf{1} \frac{1}{\bar{f}} \circ \bar{\mathbf{f}}_\mathbf{2} \frac{1}{\bar{f}} \circ \ldots \circ \bar{\mathbf{f}}_\mathbf{3N} = (\bar{f}^{-3N+1}) \circ_{k=1}^{3N} \bar{\mathbf{f}}_\mathbf{k}, \tag{5}$$

where $\circ$ denotes tensor product, and Z_N is

$$Z_N \approx (\bar{f}^{-3N+1}) \prod_{k=1}^{3N} \left(\mathbf{w}_k^T \bar{\mathbf{f}}_\mathbf{k}\right). \tag{6}$$

For crystal structures with identical particles, symmetry reduces computational complexity. In perfectly ordered solids, lattice periodicity makes all particles symmetrically equivalent, allowing us to calculate just one representative fiber, which yields

$$Z_N \approx (\bar{f}^{-3N+1}) \prod_{k=1}^{3} \left(\mathbf{w}_k^T \bar{\mathbf{f}}_\mathbf{k}\right)^N. \tag{7}$$

For crystal structures, such as Face-Centered Cubic (FCC), Body-Centered Cubic (BCC) or Diamond Cubic (DC), where the x, y, and z directions are symmetrically equivalent, only a single integral is needed:

$$Z_N \approx (\bar{f}^{-3N+1}) \left(\mathbf{w}_1^T \bar{\mathbf{f}}_\mathbf{1}\right)^{3N}. \tag{8}$$

H. Rank-2 TT cross of the Z_N integrand

In Fig. 2(b), we illustrate our rank-2 TT-cross approach. Again we choose the integration domain to be symmetric around the minimum of the potential- at $[\bar{q}_1, \bar{q}_2]$ points. To calculate the integral we will compute TT cross of $f(\mathbf{q})$ in this domain. We select two points on the superdiagonal at distance δ, where each point has the same value, $\bar{f}$. These two points, $\bar{\mathbf{q}} = \{\bar{q}_1, \bar{q}_2\}$, and the corresponding two points $\mathbf{q}_2^\delta = \{\bar{q}_1 + \delta, \bar{q}_2\}$ outside the superdiagonal determine four fibers whose intersection forms the matrix $\mathbf{S}$.

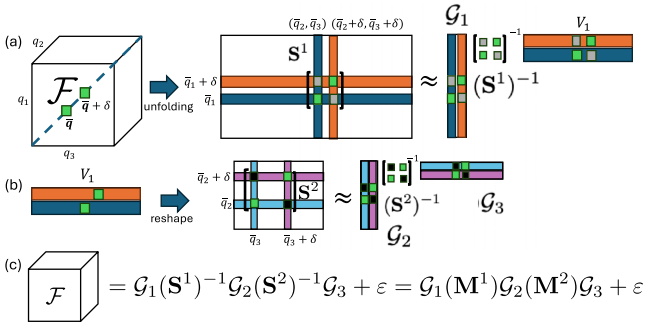


FIG. 3. Illustration of a rank-2 TT format of $\mathcal{F}$ for a three particle system in 1D. (a) The tensor is unfolded along its first dimension, and the rank-2 approximation is defined by selecting two points (green) on the superdiagonal. These two points determine four fibers (orange and blue), which intersect at four points defining: $\mathcal{G}_1$, $(\mathbf{S}^1)^{-1}$, and $\mathbf{V}_1$. (b) $\mathbf{V}_1$ is reshaped, and the next **CUR**'s components are extracted. (c) The final TT format is $\mathcal{F} = \mathcal{G}_1\mathbf{M}^1\mathcal{G}_2\mathbf{M}^2\mathcal{G}_3 + \varepsilon$ where $\mathbf{M}^1 := (\mathbf{S}^1)^{-1}$ and $\mathbf{M}^2 := (\mathbf{S}^2)^{-1}$.

Assuming that $f(\mathbf{q})$ is a very steep function, than out of the diagonal: $f(\mathbf{q}_1^\delta) \approx f(\mathbf{q}_2^\delta) \approx 0$, and we have $|\det(\mathbf{S})| \approx \bar{f}^2$. Hence, the choice of this submatrix is close to the theoretical maximum volume, we need to calculate for the rank-2 TT-cross, and we can formulate the following result:

For N identical particles, let the point $\bar{\mathbf{q}}$ and point $\bar{\mathbf{q}} + \delta$ are on the superdiagonal. Let $\mathbf{S}^k$, for $k = 1, \ldots, 3N-1$, denote the submatrix of the k$^{\text{th}}$ subsequent unfolding of $\mathcal{F}$. Then, for a very steep function $f(\mathbf{q})$, $|\det(\mathbf{S}^k)| \approx \bar{f}^2$, which in many cases is close to the maximum volume we need to calculate for the rank-2 TT-cross interpolation.

This rank-2 approach for TT cross of the Boltzmann factor is illustrated in detail for three particles in 1D, in Fig. 3. We now detail the TT-core construction for rank-2 interpolation. Assume the superdiagonal indices of two peak points are

$$\mathbf{j}_1 := \left[j_1^1, j_1^2, \ldots, j_1^{3N}\right], \quad \mathbf{j}_2 := \left[j_2^1, j_2^2, \ldots, j_2^{3N}\right].$$

Then, for each $k = 1, \ldots, 3N-1$, define the 2×2 interpolation matrix:

$$\begin{aligned}
\mathbf{S}_{1,1}^k &= \mathcal{F}\left(j_1^1, \ldots, j_1^k, j_1^{k+1}, \ldots, j_1^{3N}\right), \\
\mathbf{S}_{1,2}^k &= \mathcal{F}\left(j_1^1, \ldots, j_1^k, j_2^{k+1}, \ldots, j_2^{3N}\right), \\
\mathbf{S}_{2,1}^k &= \mathcal{F}\left(j_2^1, \ldots, j_2^k, j_1^{k+1}, \ldots, j_1^{3N}\right), \\
\mathbf{S}_{2,2}^k &= \mathcal{F}\left(j_2^1, \ldots, j_2^k, j_2^{k+1}, \ldots, j_2^{3N}\right),
\end{aligned}$$

with $\mathbf{M}^k = (\mathbf{S}^k)^{-1}$. These matrices encode coupling between consecutive TT cores.

The TT cores themselves are built from tensor fibers: For the first core $\mathcal{G}_1$:

$$\begin{aligned}
\mathcal{G}_1(1, :, 1) &= \left\{\mathcal{F}\left(i_1, j_1^2, \ldots, j_1^{3N}\right)\right\}_{i_1=1}^n, \\
\mathcal{G}_1(1, :, 2) &= \left\{\mathcal{F}\left(i_1, j_2^2, \ldots, j_2^{3N}\right)\right\}_{i_1=1}^n.
\end{aligned}$$

For the last core $\mathcal{G}_{3N}$:

$$\begin{aligned}
\mathcal{G}_{3N}(1, :, 1) &= \left\{\mathcal{F}\left(j_1^1, \ldots, j_1^{3N-1}, i_{3N}\right)\right\}_{i_{3N}=1}^n, \\
\mathcal{G}_{3N}(2, :, 1) &= \left\{\mathcal{F}\left(j_2^1, \ldots, j_2^{3N-1}, i_{3N}\right)\right\}_{i_{3N}=1}^n.
\end{aligned}$$

For intermediate cores $\mathcal{G}_k$, with $k \in \{2, \ldots, 3N-1\}$:

$$\mathcal{G}_k(1,:,1) = \left\{\mathcal{F}\left(j_1^1, \ldots, j_1^{k-1}, i_k, j_1^{k+1}, \ldots, j_1^{3N}\right)\right\}_{i_k=1}^{n},$$
$$\mathcal{G}_k(2,:,1) = \left\{\mathcal{F}\left(j_2^1, \ldots, j_2^{k-1}, i_k, j_1^{k+1}, \ldots, j_1^{3N}\right)\right\}_{i_k=1}^{n},$$
$$\mathcal{G}_k(1,:,2) = \left\{\mathcal{F}\left(j_1^1, \ldots, j_1^{k-1}, i_k, j_2^{k+1}, \ldots, j_2^{3N}\right)\right\}_{i_k=1}^{n},$$
$$\mathcal{G}_k(2,:,1) = \left\{\mathcal{F}\left(j_2^1, \ldots, j_2^{k-1}, i_k, j_1^{k+1}, \ldots, j_1^{3N}\right)\right\}_{i_k=1}^{n}.$$

After constructing the TT representation, Z_N is computed using Eqs. (3) and (4).

For crystalline systems with translational symmetry (e.g., FCC, BCC, DC), where all directions are equivalent, the matrices $\mathbf{M}^k$ and cores $\mathcal{G}_k$ for $k \in \{2, \ldots, 3N-1\}$ are identical, allowing for efficient reuse.

Finally, this construction generalizes to rank-r TT cross by selecting r interpolation points along the superdiagonal and constructing the corresponding $r \times r$ matrices and TT-cores.

III. NUMERICAL EXPERIMENTS

It is well known from statistical mechanics that for a system of N particles in a constant volume V, the mean internal energy of a system in the canonical ensemble, at temperature T, is given by

$$E = -\frac{\partial \ln Q_N}{\partial \beta} = \frac{3N}{2\beta} - \frac{\partial \ln Z_N}{\partial \beta}. \tag{9}$$

Numerically, we can approximate E as

$$E \approx \frac{3N}{2\beta} - \frac{\ln Z_N(\beta + \Delta\beta) - \ln Z_N(\beta)}{\Delta\beta}. \tag{10}$$

The thermodynamic pressure, of the considered phase, at volume V and temperature T, is approximated by

$$P = \frac{\partial \ln Q_N}{\beta \partial V} = \frac{\partial \ln Z_N}{\beta \partial V} \approx \frac{\ln Z_N(V + \Delta V) - \ln Z_N(V)}{\beta \Delta V}. \tag{11}$$

Both ΔV and $\Delta\beta$ are chosen empirically by gradually reducing their values until the computed internal energy and pressure become numerically stable. For all numerical experiments presented in this work, each dimension is discretized using a Gaussian quadrature scheme with eight points per cell across 100 cells, yielding a total of 800 quadrature points per dimension. This quadrature configuration was selected as a balance between accuracy and computational efficiency, and we have verified that it ensures convergence of the computed integrals. To validate and demonstrate our approach, we calculate in two different ways these two thermodynamical quantities: (a) we calculate them by Molecular Dynamics (MD) simulations, and (b) by our TT-cross method. The MD parameters are given in Appendix B. We consider two materials systems: (a) copper (Cu) and (b) argon (Ar). For Cu, see Fig. 4, we use tight-binding potential [45] that requires a rank-1 decomposition. For Ar, Fig. 5, we use a HIP-NN many-body machine learning (ML) potential [46], trained with an MP2 dataset generated in Appendix C. Interestingly, accurately reproducing the nonpairwise HIP-NN MD simulations required a rank-2 decomposition, in contrast to the other potentials

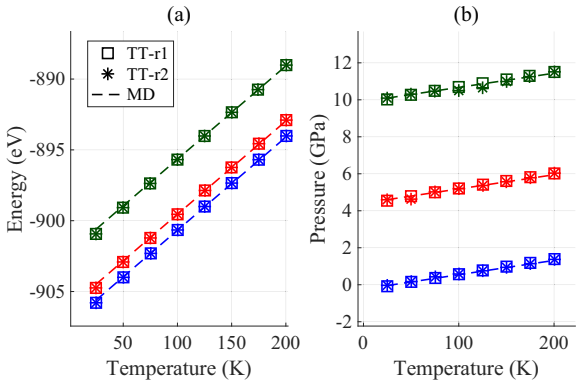


FIG. 4. TT-rank 1 (squares) and TT-rank 2 (stars) predictions compared with MD results for three densities 9.49 (green), 9.21 (red), and 8.94 (blue) g/cm^3, for Cu, calculated with tight-binding SMATB potential. Both TT-rank 1 and TT-rank 2 approximations of the internal energy (a) and pressure (b) are in excellent agreement with MD results.

examined in this paper which required only a rank-1 decomposition. For each structure with a given density, we fix the lattice constant and determine the coordinates of each particle using Atomic Simulation Environment (ASE) [47]. Next, for each structure with periodic boundaries conditions, we calculate the potential, using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [48], and evaluate the elements of the chosen fibers of $\mathcal{F}$.

Finally, we compare the computed phase transition temperature between $\alpha - \mathrm{Sn}$ (DC) and $\beta - \mathrm{Sn}$ (BCT) between a MD based thermodynamic integration and the TT-cross methodology. For this test a MEAM potential [49] was utilized due to the large number of simulations required to generate the themodynamic integration phase transition line. Specifically, the MEAM potential developed by Ko and colleagues [50] is known to accurately model the $\alpha - \mathrm{Sn}$ to $\beta - \mathrm{Sn}$ phase transition temperature. Figure 6 shows that our TT-cross method

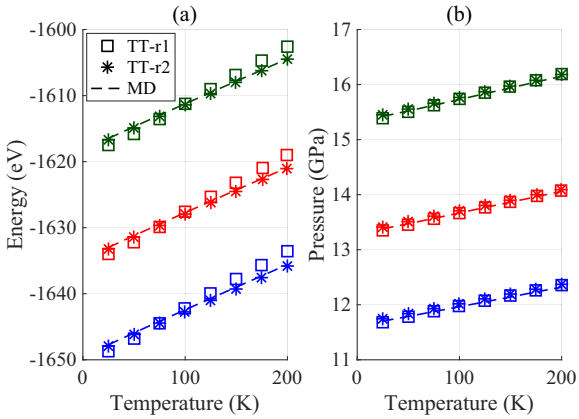


FIG. 5. TT-rank 1 (squares) and TT-rank 2 (stars) predictions compared with MD results for FCC Argon calculated with the HIP-NN potential across three densities 3.48 (green), 3.35 (red), and 3.23 (blue) g/cm^3. Rank-2 TT approximation of the internal energy (a) and pressure (b) consistently improves the approximation, yielding closer agreement with MD results.

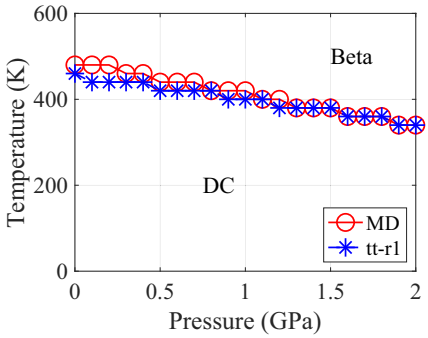


FIG. 6. Phase diagram of Tin (Sn) using a MEAM potential, showing the α-β-Sn transition. The rank-1 TT (blue stars) reconstructs the phase boundary in close agreement with MD results (red circles).

reproduces the solid-solid phase transition between the $\alpha-$ Sn(DC) and $\beta-$ Sn(β) crystal structures. To create the initial configurations of Sn(β) or Sn(DC), at a given pressure, we run MD for fixed: crystal structure, number of particles, and volume. The MD simulations for the phase diagram used Calphy [51], more details are given in Appendix B. As part of our method, for each phase and temperature, we compute the configurational integral Z_N and thermodynamic pressure P, and evaluate the Gibbs free energy,

$$G = \frac{1}{N}\left(-\frac{1}{\beta}\ln Z_N + PV\right). \tag{12}$$

The phase boundary is located at the temperature at which $G_{\mathrm{DC}} \approx G_{\beta}$.

For the phase diagram in Fig. 6, each data point—representing a specific pressure–temperature pair—requires computing the Gibbs free energy for both α-Sn and β-Sn structures. The TT method computes each free energy in approximately 20 seconds on our Linux system using a single CPU core in serial execution. Covering the full range from 0 to 600 K in 25 K increments (25 temperature points) and 0 to 2 GPa in 0.1 GPa increments (21 pressure points) for both phases results in a total of 1050 configurations. The total runtime is approximately 5.8 core hours. In comparison, thermodynamic integration using molecular dynamics, performed via the `Calphy` package, required 2560 core hours to generate the same phase diagram. This corresponds to a speed-up factor of over 400$\times$ for this task.

IV. CONCLUSIONS

The statistical mechanics of strongly interacting atomic and molecular systems exemplifies how innovative tensor networks techniques can yield substantial benefits. Central to these systems is the configurational integral, a quantity that encapsulates the essence of particle interactions, which historically been deemed analytically and numerically intractable due to the curse of dimensionality.

In this paper, we demonstrate that our TT approach successfully evaluates configurational integrals in crystalline solids from first principles, enabling an accurate, fast, and efficient computation of thermodynamic functions, such as Helmholtz free energies and equations of state. In addition, our approach supports the use of more realistic ML potentials, thereby capturing the underlying physics more accurately than the pairwise additivity approximation. While the presented results focus on crystalline solids with identical particles, we are optimistic that further analysis of Boltzmann integrands and its symmetries will enable extension of our method.

ACKNOWLEDGMENTS

We acknowledge the support of NNSA for the U.S. DOE at LANL under Contract No. DE-AC52-06NA25396 and Laboratory Directed Research and Development (LDRD) for support through 20230067DR. LANL is managed by Triad National Security, LLC, for the National Nuclear Security Administration of the U.S. DOE under Contract No. 89233218CNA000001.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: PROOF OF THE RESULT FOR RANK-1 TT-CROSS

To show this result we begin by unfolding the tensor $\mathcal{F}$ along the first mode to obtain the matrix $\mathbf{F}_1 \in \mathbb{R}^{n\times n^{3N-1}}$, where each entry is indexed by $[i_1, I_{i_2 i_3 \ldots i_{3N}}]$, and $I_{i_2 i_3 \ldots i_{3N}}$ is a flattened multi-index combining the remaining $3N-1$ dimensions.

According to the maximum volume principle, the best rank-1 approximation of $\mathbf{F}_1$ that passes through the pivot index $[j_1^1, I_{j_1^2 j_1^3 \ldots j_1^{3N}}]$ is given by

$$\mathbf{F}_1 \approx \mathbf{F}_1\big(:, I_{j_1^2 j_1^3 \ldots j_1^{3N}}\big)\cdot\frac{1}{\bar{f}}\cdot\mathbf{F}_1\big(j_1^1, :\big) = \bar{\mathbf{f}}_1\cdot\frac{1}{\bar{f}}\cdot\mathbf{F}_1\big(j_1^1, :\big).$$

We then reshape the row vector $\mathbf{F}_1(j_1^1, :) \in \mathbb{R}^{1\times n^{3N-1}}$ into a matrix $\mathbf{F}_2 \in \mathbb{R}^{n\times n^{3N-2}}$, corresponding to unfolding along the second mode. Applying the maximum volume rank-1 approximation again at index j_1^2, we obtain

$$\mathbf{F}_2 \approx \mathbf{F}_2\big(:, I_{j_1^3 j_1^4 \ldots j_1^{3N}}\big)\cdot\frac{1}{\bar{f}}\cdot\mathbf{F}_2\big(j_1^2, :\big) = \bar{\mathbf{f}}_2\cdot\frac{1}{\bar{f}}\cdot\mathbf{F}_2\big(j_1^2, :\big).$$

We continue this sequential procedure, reshaping and applying the max-volume rank-1 decomposition through the fixed multi-index $\mathbf{j}_1$ at each unfolding. After repeating this process up to the $(3N-1)^{\text{th}}$ mode, we arrive at the final rank-1 TT-cross interpolation of $\mathcal{F}$ as described in Eq. (5):

$$\hat{\mathcal{F}} = \bar{f}^{-3N+1}\circ_{k=1}^{3N}\bar{\mathbf{f}}_k.$$

Finally, applying the TT cubature rule, Eq. (4) to the rank-1 TT format yields the approximate integral Z_N in Eq. (6):

$$Z_N \approx \bar{f}^{-3N+1}\prod_{k=1}^{3N}\left(\mathbf{w}_k^T\bar{\mathbf{f}}_k\right).$$

APPENDIX B: MOLECULAR DYNAMIC PROTOCOLS

1. Argon and Copper Molecular dynamics (MD) simulations

The MD are performed using LAMMPS to study FCC argon and copper at various densities under fixed temperature

conditions. A $4 \times 4 \times 4$ FCC supercell is generated using the ASE package [47] and rescaled to match the target density. The Second Moment Approximation to Tight Binding (SMATB) potential [52] is used for copper, while argon is modeled using the HIP-NN machine learning potential [53]. A small custom training dataset for the argon HIP-NN potential was generated through active learning, as described in Appendix E. Simulations are conducted in the NVT ensemble using a Nosé-Hoover thermostat with a time step 0.2 fs, and a total run length of 20 000 steps.

2. Tin phase diagram

The phase diagram of tin Sn is computed using the calphy package, an open-source Python framework for automated free energy and phase boundary calculations [51]. calphy employs the thermodynamic integration method combined with molecular dynamics simulations to evaluate the Gibbs free energy of different crystalline phases. In this study, the DC, α and β-Sn phases are considered. For these simulations, CALPHY was used in "temperature sweep" mode to scan Gibbs free energy over a variety of different initial pressures. Specifically, a simulation was run every 0.05 GPa from 0 to 2 GPa from 100 to 600 K starting in both the $\alpha - \mathrm{Sn}$ and $\beta - \mathrm{Sn}$ configuration. 10 000 equilibration steps were performed before running the temperature sweep over 50 000 steps, with a 1 fs MD time step. For each initial pressure, the temperature at which the $\alpha - \mathrm{Sn}$ and $\beta - \mathrm{Sn}$ Gibbs free energies crossed is reported as the phase transition temperature. These temperatures were then fit to a 25-K grid, over which the TT method was performed, since this method required a grid in both temperature and pressure space.

TABLE I. Test set accuracies for the developed HIP-NN argon model. Model was primarily used to demonstrate that higher- order tensor integrations are required for modeling nonpairwise potential energy surfaces.

	MAE	RMSE
Energy (eV)	0.40	0.65
Force (eV/Å)	0.35	0.72

APPENDIX C: AN ACTIVE LEARNED ARGON MP2 DATASET

In order to train the Argon HIP-NN machine learning potential, a dataset of 425 MP2 calculations was generated via active learning. RI-MP2 implemented in Q-Chem [54] utilizing the CC-PVTZ [55] basis set was utilized as the underlying reference method. Preliminary MD simulations of up to 40 Ar atoms were performed held together with a linear spherical well potential with a radius of 7.0 Å and a force of 0.01 eV/Å. These MD simulations were performed with an ensemble of eight HIP-NN models, and stopped when the standard deviation of the force predictions exceeded 0.25 eV/Å. After 20 active learning generations, a HIP-NN model with 60 atomic features, two interaction layers, three atomic layers, a soft cutoff of 4.0 Å and a hard cutoff of 5.5 Å was trained to the gathered data. A random split of 80% training, 10% validation, and 10% testing was generated from the 425 structures. The resulting model had test set accuracies as reported in Table I.

[1] J. C. Maxwell, Philos. Trans. R. Soc. London **157**, 49 (1867).
[2] L. Boltzmann, *Lectures on Gas Theory* (Dover, New York, 2011).
[3] J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Cambridge University Press, New York, 2010).
[4] R. Bellman, Science **153**, 34 (1966).
[5] B. Alexandrov, G. Manzini, E. W. Skau, P. M. D. Truong, and R. G. Vuchkov, Mathematics **11**, 534 (2023).
[6] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[7] C. H. Bennett, J. Comput. Phys. **22**, 245 (1976).
[8] D. Frenkel and A. J. Ladd, J. Chem. Phys. **81**, 3188 (1984).
[9] Y. Sugita and Y. Okamoto, Chem. Phys. Lett. **314**, 141 (1999).
[10] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. USA **99**, 12562 (2002).
[11] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[12] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[13] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1999).
[14] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).
[15] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[16] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Elsevier, Amsterdam, 2002).
[17] J. Ihm, Rep. Prog. Phys. **51**, 105 (1988).
[18] H. J. Berendsen, in *Computational Molecular Dynamics: Challenges, Methods, Ideas: Proceedings of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin, May 21–24, 1997* (Springer, New York, 1999), pp. 3–36.
[19] M. K. Cowles and B. P. Carlin, J. Am. Stat. Assoc. **91**, 883 (1996).
[20] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, D. P. Mandic *et al.*, Found. Trends Mach. Learn. **9**, 249 (2016).
[21] I. Oseledets and E. Tyrtyshnikov, Linear Algebra Appl. **432**, 70 (2010).
[22] V. Khoromskaia, B. N. Khoromskij, and R. Schneider, SIAM J. Sci. Comput. **35**, A987 (2013).
[23] S. Dolgov and D. Savostyanov, Comput. Phys. Commun. **246**, 106869 (2020).
[24] G. Evenbly and G. Vidal, Phys. Rev. Lett. **115**, 180405 (2015).
[25] M. S. J. Tepaske and D. J. Luitz, Phys. Rev. Res. **3**, 023236 (2021).
[26] J. Colbois, Tensor network investigation of frustrated ising models, Master's thesis, École Polytechnique Fédérale de Lausanne, 2022.
[27] F. F. Song, T. Y. Lin, and G. M. Zhang, Phys. Rev. B **108**, 224404 (2023).
[28] B. Lanthier, J. Côté, and S. Kourtis, Front. Phys. **12**, 1431810 (2024).
[29] A. A. Klishin and G. van Anders, Soft Matter **16**, 6523 (2020).
[30] M. C. Bañuls, Annu. Rev. Condens. Matter Phys. **14**, 173 (2023).

[31] L. Ying, Multiscale Model. Simul. **15**, 1423 (2017).
[32] C.-Y. Huang, S.-H. Chan, Y.-J. Kao, and P. Chen, Phys. Rev. B **107**, 205123 (2023).
[33] B. Vanhecke, J. Colbois, L. Vanderstraeten, F. Verstraete, and F. Mila, Phys. Rev. Res. **3**, 013041 (2021).
[34] D. P. Truong, M. I. Ortega, I. Boureima, G. Manzini, K. Ø. Rasmussen, and B. S. Alexandrov, J. Comput. Phys. **507**, 112943 (2024).
[35] M. E. Danis, D. Truong, I. Boureima, O. Korobkin, K. Ø. Rasmussen, and B. S. Alexandrov, J. Comput. Phys. **529**, 113891 (2025).
[36] I. V. Oseledets, SIAM J. Sci. Comput. **33**, 2295 (2011).
[37] S. A. Goreinov, E. E. Tyrtyshnikov, and N. L. Zamarashkin, Linear Algebra Appl. **261**, 1 (1997).
[38] M. W. Mahoney and P. Drineas, Proc. Natl. Acad. Sci. USA **106**, 697 (2009).
[39] D. V. Savostyanov, Linear Algebra Appl. **458**, 217 (2014).
[40] S. A. Goreinov and E. E. Tyrtyshnikov, Contemp. Math. **280**, 47 (2001).
[41] https://github.com/oseledets/TT-Toolbox.
[42] I. Oseledets, Comput. Methods Appl. Math. **11**, 382 (2011).
[43] S. V. Dolgov and D. V. Savostyanov, SIAM J. Sci. Comput. **36**, A2248 (2014).
[44] T. Cui, S. Dolgov, and O. Zahm, J. Comput. Phys. **485**, 112103 (2023).
[45] F. Cleri and V. Rosato, Phys. Rev. B **48**, 22 (1993).
[46] N. Lubbers, J. S. Smith, and K. Barros, J. Chem. Phys. **148**, 241715 (2018).
[47] A. H. Larsen, J. J. Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus *et al.*, J. Phys.: Condens. Matter **29**, 273002 (2017).
[48] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, and S. J. Plimpton, Comput. Phys. Commun. **271**, 108171 (2022).
[49] B.-J. Lee, M. I. Baskes, H. Kim, and Y. K. Cho, Phys. Rev. B **64**, 184102 (2001).
[50] W.-S. Ko, D.-H. Kim, Y.-J. Kwon, and M. H. Lee, Metals **8**, 900 (2018).
[51] S. Menon, Y. Lysogorskiy, J. Rogal, and R. Drautz, Phys. Rev. Mater. **5**, 103801 (2021).
[52] V. Rosato, M. Guillope, and B. Legrand, Philos. Mag. A **59**, 321 (1989).
[53] M. Chigaev, J. S. Smith, S. Anaya, B. Nebgen, M. Bettencourt, K. Barros, and N. Lubbers, J. Chem. Phys. **158**, 184108 (2023).
[54] Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng *et al.*, Mol. Phys. **113**, 184 (2015).
[55] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).