Kirkwood-Buff integrals

From fluctuations in finite volumes to the thermodynamic limit

Simon, J. M.; Krüger, P.; Schnell, S. K.; Vlugt, T. J.H.; Kjelstrup, S.; Bedeaux, D.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Kirkwood–Buff integrals: From fluctuations in finite volumes to the thermodynamic limit F

J.-M. Simon, P. Krüger, S. K. Schnell, et al.

COLLECTIONS

F   This paper was selected as Featured

View Online   Export Citation   CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Computing chemical potentials of solutions from structure factors
The Journal of Chemical Physics **157**, 121101 (2022); https://doi.org/10.1063/5.0107059

Virial equation of state as a new frontier for computational chemistry
The Journal of Chemical Physics **157**, 190901 (2022); https://doi.org/10.1063/5.0113730

War and peace between electrostatic and van der Waals forces regulate translational and rotational diffusion
The Journal of Chemical Physics **157**, 080901 (2022); https://doi.org/10.1063/5.0098506

Kirkwood–Buff integrals: From fluctuations in finite volumes to the thermodynamic limit F

View Online   Export Citation   CrossMark

J.-M. Simon,[1,a)] ID P. Krüger,[2] ID S. K. Schnell,[3] ID T. J. H. Vlugt,[4] ID S. Kjelstrup,[5] ID and D. Bedeaux[5] ID

AFFILIATIONS

[1] Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR-6303 CNRS - Université de Bourgogne Franche-Comté,
F-21078 Dijon, France
[2] Graduate School of Science and Engineering, Molecular Chirality Research Center, Chiba University, Chiba 263-8522, Japan
[3] Department of Materials Science and Engineering, NTNU - Norwegian University of Science and Technology,
N-7491 Trondheim, Norway
[4] Process & Energy Department, Delft University of Technology, Leeghwaterstraat 39, 2628CB Delft, The Netherlands
[5] Center of Excellence PoreLab, Department of Chemistry, Faculty of Natural Sciences, NTNU - Norwegian University of Science
and Technology, Trondheim, Norway

[a)] Author to whom correspondence should be addressed: jmsimon@u-bourgogne.fr

ABSTRACT

The Kirkwood–Buff theory is a cornerstone of the statistical mechanics of liquids and solutions. It relates volume integrals over the radial distribution function, so-called Kirkwood–Buff integrals (KBIs), to particle number fluctuations and thereby to various macroscopic thermodynamic quantities such as the isothermal compressibility and partial molar volumes. Recently, the field has seen a strong revival with breakthroughs in the numerical computation of KBIs and applications to complex systems such as bio-molecules. One of the main emergent results is the possibility to use the finite volume KBIs as a tool to access finite volume thermodynamic quantities. The purpose of this Perspective is to shed new light on the latest developments and discuss future avenues.

Published under an exclusive license by AIP Publishing. https://doi.org/10.1063/5.0106162

I. INTRODUCTION

In the context of the statistical description of liquid solutions, Kirkwood and Buff (KB) proposed in the 1950s[1] a rigorous statistical approach that makes the link between structure, thermodynamics, and density fluctuations.[2,3] In this context, we can say that the structure provides access to thermodynamic properties and the inverse is also true (thermodynamics provides insight in structure), the key function to operate being the so-called Kirkwood–Buff integral (KBI). KB theory is an exact theory, valid for any isotropic solution in thermodynamic equilibrium, and it is widely used in physical chemistry and biochemistry.[3] Until recently, however, for systems of great structural complexity and/or long–range correlations, KBIs could not be obtained accurately from molecular simulation because of severe finite-size effects. Efficient solutions for this have been proposed in the past few years,[4–11] and KB theories have been applied to biological molecules,[12] force field development,[13] multicomponent fluids[14] and oligomeric blends,[15] and ionic solutions.[16] There are numerous other recent applications of KB theory, including hydration shell,[17] self-aggregation,[18] and protein stability.[19–21] Our list of applications is not exhaustive, and we have focused on applications related to the concept of finite volume KB theory, exposed in the first part of this article. Since this concept has considerably enlarged the scope of the KB theory, it is expected to stimulate more interesting applications in the future.

As below the thermodynamic limit, the KB theory is exact, it opens the question of the link with thermodynamics and beyond, with the type of thermodynamics that applies to small systems. Evidence has now accumulated to support the idea of a scaling method, called the Small System Method (SSM), first proposed by Schnell *et al.*,[7] and extended by Bedeaux *et al.*,[22] based on the thermodynamics of small systems by Hill.[23,24]

The present work is motivated by our interest in optimizing the computation of KBIs and improving the physical understanding at different length scales. The perspective is divided into three parts: the first one presents KBIs and the methods to calculate them,

the second is dedicated to the thermodynamic interpretation of the size-dependence of the KBIs, and the last one is focused on new applications and perspectives.

II. FINITE VOLUME KIRKWOOD–BUFF INTEGRALS

After some general considerations on KBIs, the focus is put on the methodology to compute them from molecular simulation while knowing the radial distribution function (RDF). In the last part, recent developments in the calculation of the RDF are added.

A. Basics

As first realized by Kirkwood and Buff,[1] particle number fluctuations can be computed from volume integrals over the pair-distribution function (PDF), this defines the so-called Kirkwood–Buff integral,

$$
\begin{aligned}
G_{ab}^V &= \frac{V}{\langle N_a \rangle \langle N_b \rangle} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2 \rho_a(\mathbf{r}_1) \rho_b(\mathbf{r}_2) (g_{ab}(\mathbf{r}_1, \mathbf{r}_2) - 1) \\
&= V \frac{\langle N_a N_b \rangle - \langle N_a \rangle \langle N_b \rangle}{\langle N_a \rangle \langle N_b \rangle} - \frac{\delta_{ab}}{\rho_a},
\end{aligned}
\tag{1}
$$

where $a, b \dots$ denote the particle species, N_i is the number of i-particles in volume V, $\rho_i(\mathbf{r})$ is the ensemble averaged local density of i-particles at position $\mathbf{r}$, and $\rho_i = \langle N_i \rangle / V$ is the average density in V. The $g_{ab}(\mathbf{r}_1, \mathbf{r}_2)$ is the PDF and $\langle \cdots \rangle$ denotes the ensemble average. By definition of g_{ab}, Eq. (1) holds in any system.

We consider an open subvolume V of an infinite isotropic fluid. Then, $\rho_i(\mathbf{r})$ is constant and g_{ab} depends only on the particle distance $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. In this situation, the PDF has no privileged orientation and it reduces to the radial distribution function. We can rewrite Eq. (1) as

$$
G_{ab}^V \equiv \frac{1}{V} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2 (g_{ab}(r_{12}) - 1).
\tag{2}
$$

In the limit $V \to \infty$, we have

$$
G_{ab}^\infty = \int_0^\infty (g_{ab}(r) - 1) 4\pi r^2 dr,
\tag{3}
$$

which is the expression introduced by Kirkwood and Buff.[1] In practice, the integral of Eq. (3) must be truncated at some upper bound L,

$$
\tilde{G}_{ab}^L = \int_0^L (g_{ab}(r) - 1) 4\pi r^2 dr,
\tag{4}
$$

which is called the "running KBI" (RKBI). In the past, the RKBI was used in most calculations,[3] but convergence to the thermodynamic limit (TL) is difficult for two reasons. First, the RDF is often obtained from molecular simulations in closed systems. This canonical RDF does not tend exactly to 1 for large r, but to $1 - c/V_0$, where c is a constant and V_0 is the simulation box size.[3] This small, but systematic difference makes the KBIs diverge as L^3. Second, even with an exact, grand-canonical RDF, the RKBI strongly oscillates as a function of L, and so extrapolation to $L \to \infty$ can be very difficult.

B. Practical computation of KBIs

1. RDF calculation

In molecular simulation software, pair distribution functions are nearly always computed by "counting" the number of molecules at distances between r and $r + \Delta r$ from a central molecule.[25] Although this is correct, it is not efficient as the number of molecules per bin can be small, resulting in poor statistics and noisy radial distribution functions, especially for small bin sizes Δr. Recently, Borgis *et al.* have derived a much more efficient way of obtaining RDFs from molecular dynamics simulations[26] using the forces between particles and molecules,

$$
g_{ab}(r) = \frac{V}{4\pi N_a N_b k_B T} \left\langle \sum_{i \in a < j \in b} \frac{H(r - r_{ij})(\mathbf{F}_i - \mathbf{F}_j)\mathbf{r}_{ij}}{r_{ij}^3} \right\rangle,
\tag{5}
$$

where T and k_B are, respectively, the temperature and the Boltzmann constant. In this equation, $\mathbf{F}_i$ is the net force of particle i and $H(r)$ is the step function [i.e., $H(r) = 1$ if $r > 0$ and $H(r) = 0$ otherwise].

In Refs. 26 and 27, it is shown that RDFs obtained using Eq. (5) are much smoother than those by the traditional "counting" approach. For the computation of Kirkwood–Buff integrals, having a smooth RDF is less relevant as integrals are usually much smoother than the integrand, as fluctuations in the integrand are integrated out.[28]

2. From closed to open ensembles

The RDF is usually obtained from molecular simulations in the canonical ensemble, i.e., with a constant number of particles and constant volume. As mentioned above, the canonical RDF has a systematic finite-size error. As a consequence, RDFs obtained from canonical ensemble simulations must be corrected before being used in KBI calculations. Without correction, KBIs diverge as $L^3 = V$. Several correction schemes have been proposed in the literature. A simple method consists in calculating $g_{ab}(r)$ for two different box sizes V_0 and extrapolating to $V_0 \to \infty$ according to the $1/V_0$ dependence of the error.[3,9] The drawback of this method is that two independent simulations must be performed and the choice of the two system sizes may affect the accuracy of the result. A more elegant method was proposed by Ganguly and van der Vegt,[29] which renormalizes the RDF by correcting the particle densities in the finite reservoir. The corrected RDF is

$$
g_{ab}(r) = \tilde{g}_{ab}(r) \frac{N_b(1 - V/V_0)}{N_b(1 - V/V_0) - \Delta N_{ab}(r) - \delta_{ab}},
\tag{6}
$$

where $\tilde{g}_{ab}(r)$ is the canonical RDF obtained with a simulation box of volume V_0 containing N_i particles of type i and $V = 4\pi r^3/3$. $\Delta N_{ab}(L)$ is the excess of particles b in a sphere of radius L around a particle a, given by

$$
\Delta N_{ab}(L) = \int_0^L dr \, 4\pi r^2 [\tilde{g}_{ab}(r) - 1].
\tag{7}
$$

Ben-Naim[3] has proposed a simpler, r-independent correction given by

$$
g_{ab}^\infty(r) = \tilde{g}_{ab}(r) + \frac{1}{V_0} \left(\frac{\delta_{ab}}{\rho_a} + G_{ab}^\infty \right).
\tag{8}
$$

Dawass $et\ al.$,[30] have compared the various schemes and found that the correction of Eq. (6) by Ganguly and van der Vegt[29] generally leads to the best results for KBIs. Before ending this section, it is worth to mention the extension method of the RDFs to large scales that Verlet proposed in 1968.[31] Despite the complexity of its numerical application, it was successfully used to obtain accurate KBIs; see Ref. 32.

3. Convergence of the KBI to the thermodynamic limit

The second finite-size problem of KBIs is that, even with the exact RDF, the RKBI, Eq. (4), converges badly with size L, because the oscillations of $g(r)$ become strongly amplified by the r^2 factor in the integrand. This problem can be avoided by using finite volume KBI, developed by the present authors.[9,30,33,34] First, we note that for $V < \infty$, G_{ab}^V in Eq. (2) cannot be written in terms of a RKBI (4). Indeed, the RKBI does not give the particle number fluctuations, but the excess particle number $\Delta N_{ab}(L)$, see Eq. (7), which strongly oscillates with L because of the shell structure in the fluid. In sharp contrast, the finite volume KBI in Eq. (2) oscillates only weakly and converges smoothly for $L \to \infty$. While Eq. (2) is exact, it is numerically cumbersome, since it involves a double volume integral. Krüger and co-workers[9,30,33–35] have shown that this six-dimensional integral can be reduced to the simple radial integral as

$$G_{ab}^V = \int_0^L (g_{ab}(r) - 1) w(r) \mathrm{d}r. \tag{9}$$

Here,

$$w(r) = \frac{1}{V} \int_V \mathrm{d}\mathbf{r}_1 \int_V \mathrm{d}\mathbf{r}_2 \delta(r - |\mathbf{r}_1 - \mathbf{r}_2|), \tag{10}$$

is a purely geometrical weight-function, which is zero for $r > L_{max}$, where L_{max} is the largest distance in V. Analytic expressions of $w(r)$ have been derived for hyperspheres[9,36,37] and cuboids.[34] For volumes of arbitrary shape, $w(r)$ can easily be obtained by numerical integration.[33] For a sphere of diameter L, we have[9]

$$w(r, L) = 4\pi r^2 (1 - 3x/2 + x^3/2), \quad x \equiv r/L. \tag{11}$$

The finite volume KBI, G_{ab}^V in Eq. (9), converges to G_{ab}^∞ as $1/L$, where L is the linear dimension of the volume V, conveniently defined as $L = V/(6A)$, where A is the surface area of V. For large L, we can expand in $1/L$ as

$$G_{ab}^V \equiv G_{ab}(L) = G_{ab}^\infty + F_{ab}^\infty/L + \mathcal{O}(L^{-2}), \tag{12}$$

where

$$F_{ab}^\infty = -\frac{3}{2} \int_0^\infty r(g_{ab}(r) - 1) 4\pi r^2 \mathrm{d}r \tag{13}$$

is an exact expression of the surface term, which holds for volumes of any shape.[34]

By plotting $G_{ab}(L)$ as a function of $1/L$, Eq. (12), one can obtain the thermodynamic limit G_{ab}^∞ by linear regression. From the knowledge of the size-dependence of thermodynamic quantities, the

thermodynamic limit can be found by extrapolation, which is linear here. This is what we called Small System Method[7,8] (SSM). We will come back to this approach in more details in Sec. III. An alternative scaling approach was proposed by Cortes-Huerto $et\ al.$[10] It is based on an approximate expression of the finite size KBI, which depends both on the subdivision volume V and the simulation box volume V_0. From this, the infinite volume KBI G_{ab}^∞ is obtained by non-linear regression. The finite size error of the RDF is corrected using Eq. (8).

The SSM is illustrated for a Lennard-Jones (LJ) liquid system in Fig. 1(a) where the quantity $G^V(L)/\sigma^3$ is plotted as a function of σ/L using different expressions for KBI, the RKBI Eq. (4), the exact finite volume KBI, Eq. (9), and the fluctuation expression from Eq. (1) (for spherical volumes). σ is the diameter of the LJ particle. The results show excellent agreement between the fluctuation approach and Eq. (9), despite the small noise for fluctuations. The trends are linear in the range 0.25 and 0.45 (for σ/L) where G^∞ can be linearly extrapolated using Eq. (13) to $\sim$−1.2. Below $\sigma/L = 0.25$, the trend is not linear due to the small size effects of the simulation box, this point has been discussed previously, see Ref. 35. The RKBI shows large oscillation that attenuates with length but not enough to obtain reliable results, i.e., a constant plateau. The statistical efficiency can be improved using larger simulations boxes but the results clearly show that the SSM is accurate enough to obtain reliable results while the usual approaches based on RKBI, find here their limit.

An equivalent method consists of plotting $L \times G_{ab}(L)$ as a function of L and extracting the linear slope at large L, which tends
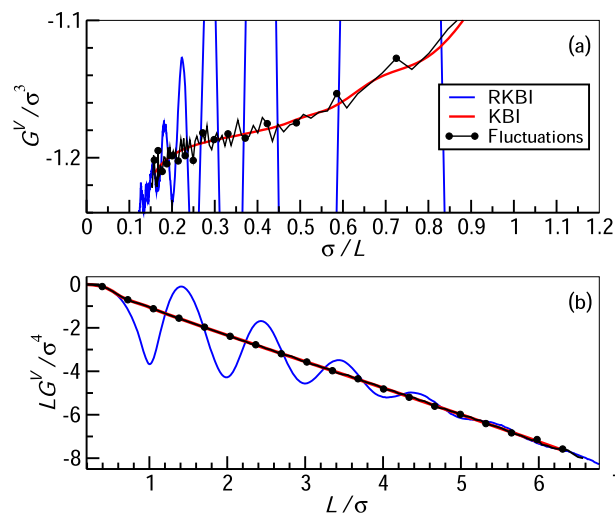
FIG. 1. For a Lennard-Jones (LJ) system, the functions (a) G^V (upper panel) and (b) $L \times G^V$ (lower panel), using Eq. (4), the RKBI, and Eq. (9), the exact KBI, respectively, as a function of the reduced values of L. These values are compared with the ones calculated from the fluctuation expression, the rhs of Eq. (1). The sets of data are the same in the two figures. The system was composed of 10 000 argon particles, the reduced density, i.e., the number of molecules per σ^3 where σ is the LJ parameter, was 0.551 and the reduced temperature, i.e., temperature divided by ϵ/k_B where ϵ and k_B are the other LJ parameter and the Boltzmann constant, respectively, was 1.40. A cut-off radius of 2.5σ was used for the LJ potential.

to G_{ab}^{∞}. This is presented in Fig. 1(b) using the same set of data as in (a). Here, we find $G_{ab}^{\infty} \approx -1.2$ in agreement with the result from Fig. 1(a). This type of representation reduces the statistical noise and enlarges the useful linear range for the regressions both for the KBI and RKBI. The two linear procedures are simple and safe.[8,30] However, it should be pointed that the choice of the linear regime for the interval of regression may affect the accuracy of the extrapolation.[30]

Alternatively, one can use one of the extrapolation expressions that have been proposed for G_{ab}^{∞}.[9,34] These are of the form,

$$G_{ab}^{\infty} \approx \int_0^{\infty} [g(r) - 1] u_k(r) \mathrm{d}r, \quad (14)$$

with some finite-range weight-function $u_k(r)$, in which k is the level of approximation. The most simple form is $u_0(r) = 4\pi r^2$, which corresponds to the "running" KBI used in most of the literature before 2013. It is obtained by simply truncating the infinite KBI at $r = L$, and it is identical to $\tilde{G}_{ab}^V$ in Eq. (4) for a sphere. However, as explained above, $\tilde{G}_{ab}^V$ is not related to number fluctuations in V. In practice, the running KBI with $u_0(r)$ converges poorly with L, because it hugely amplifies the oscillations of $g_{ab}(r)$. A better choice is $u_1(r) = 4\pi r^2 [1 - (r/L)^3]$, which was obtained by a Taylor expansion in $1/L$ of the exact KBI for a finite sphere, Eqs. (9) and (11).[9] Using Eq. (14) with $u_1(r)$ affords fast and stable convergence of the KBI.[9] Another extrapolation was developed based on Eq. (12) and a similar expansion applied to F_{ab}^{∞}.[34] The corresponding weight-function is $u_2(r, L) = 4\pi r^2 [1 + (-23x^3 + 6x^4 + 9x^5)/8]$, where $x = r/L$. This extrapolation was found to converge even faster than $u_1(r)$ in tests with a model RDF.[34] In applications with realistic, material specific RDFs,[38] both $u_1(r)$ and $u_2(r)$ have been found to be equally useful and far superior to $u_0(r)$. These new approximations for the infinite KBI, obtained by extrapolating the exact finite volume KBI—either numerically or using the functions u_1 and u_2—have led to a considerable improvement of the reliability and accuracy of KBI calculations in recent years.[14,30,39]

III. THERMODYNAMIC FORMULATION OF THE SIZE-DEPENDENCE OF KBI AND FLUCTUATIONS

One of the main interest in KBIs and fluctuations is to relate them to thermodynamic quantities.[1–3] This has been exploited at the thermodynamic limit where G_{ab}^{∞} is known. The KBI, like fluctuations, has a statistical basis, which is the same regardless of the size and shape of the volume, area, or the three-phase contact line of the system. This provides a direct path between any small system property and its KBI. This procedure has not yet been exploited in any systematic manner. We explain this possibility to find thermodynamic properties for small systems and point at some perspectives. The material in this section is a modified version of Ref. 22.

The finite volume KBI is provided by Eq. (1). It is first obtained from the density fluctuations as a function of T, V, the chemical potentials of all the species $\boldsymbol{\mu}$, and the size and the shape of the volume V. In the grand-canonical ensemble, we can rewrite Eq. (1),

$$G_{ab}^V(T, V, \boldsymbol{\mu}, \text{shape}) = V \frac{\langle N_a N_b \rangle - \langle N_a \rangle \langle N_b \rangle}{\langle N_a \rangle \langle N_b \rangle} - \frac{\delta_{ab}}{\rho_a^{\mathrm{GC}}}, \quad (15)$$

where $\rho_a^{\mathrm{GC}}(T, V, \boldsymbol{\mu}, \text{shape}) = N_a^{\mathrm{GC}}(T, V, \boldsymbol{\mu}, \text{shape})/V \equiv \langle N_a \rangle/V$ is the average number density of component a. The brackets $\langle \ldots \rangle$ denote the ensemble average in the grand-canonical ensemble. Equation (15), as well as its formulation for the grand-canonical ensemble, is the same for small and large volumes. Equation (15) and further equations in this section are completely general. In particular, they do not include any assumptions about the pair correlation function. In this section, we restrict most of the discussion to systems controlled in the grand-canonical ensemble.

Following Kirkwood and Buff, we introduce an auxiliary matrix, which is equal to the particle number correlation matrix B_{ab}^{GC}. In the grand-canonical ensemble, the matrix elements are defined by

$$B_{ab}^{\mathrm{GC}}(T, V, \boldsymbol{\mu}, \text{shape}) \equiv \frac{1}{V} [\langle N_a N_b \rangle - \langle N_a \rangle \langle N_b \rangle]$$
$$= \rho_a^{\mathrm{GC}} \rho_b^{\mathrm{GC}} G_{ab}^V(T, V, \boldsymbol{\mu}, \text{shape}) + \rho_a^{\mathrm{GC}} \delta_{ab}. \quad (16)$$

This results in

$$B_{ab}^{\mathrm{GC}}(T, V, \boldsymbol{\mu}, \text{shape}) = \frac{k_{\mathrm{B}} T}{V} \frac{\partial \langle N_a \rangle}{\partial \mu_b} = k_{\mathrm{B}} T \frac{\partial \rho_a^{\mathrm{GC}}}{\partial \mu_b}. \quad (17)$$

All elements B_{ab}^{GC} are functions of $(T, V, \boldsymbol{\mu})$ and system shape. Both matrices G_{ab}^V and B_{ab}^{GC} are symmetric. The equation applies to small systems, also those that are not isotropic.

Kirkwood and Buff[1] next defined the auxiliary matrix A_{ab}^{GC}. For environmental variables $(T, V, \mathbf{N}^{\mathrm{GC}})$ and given shape, we obtain

$$A_{ab}^{\mathrm{GC}}(T, V, \mathbf{N}^{\mathrm{GC}}, \text{shape}) = \beta V \left(\frac{\partial \mu_b}{\partial N_a^{\mathrm{GC}}} \right)_{T, V, \mathbf{N}_a^{\mathrm{GC}}}$$
$$= \beta \left(\frac{\partial \mu_b}{\partial \rho_a^{\mathrm{GC}}} \right)_{T, V, \rho_a^{\mathrm{GC}}}, \quad (18)$$

where

$$\mathbf{N}_a^{\mathrm{GC}} \equiv (N_1, \ldots, N_{a-1}, N_{a+1}, \ldots, N_k)^{\mathrm{GC}}$$
$$\equiv V(\rho_1, \ldots, \rho_{a-1}, \rho_{a+1}, \ldots, \rho_k)^{\mathrm{GC}} \equiv V\rho_a^{\mathrm{GC}}. \quad (19)$$

The **A** and **B** matrices are each other's inverse, meaning that

$$\sum_b B_{ab}^{\mathrm{GC}} A_{bk}^{\mathrm{GC}} = \delta_{ak}. \quad (20)$$

From the symmetry of $\mathbf{B}^{\mathrm{GC}}$, we conclude that also $\mathbf{A}^{\mathrm{GC}}$ is symmetric. The so-called matrix of thermodynamic factors $\boldsymbol{\Gamma}^{\mathrm{GC}}$ can now be obtained from $\mathbf{A}^{\mathrm{GC}}$ by

$$\Gamma_{ab}^{\mathrm{GC}}(T, V, \mathbf{N}^{\mathrm{GC}}, \text{shape}) \equiv \rho_a^{\mathrm{GC}} A_{ab}^{\mathrm{GC}}(T, V, \mathbf{N}^{\mathrm{GC}}, \text{shape})$$
$$= \beta N_a^{\mathrm{GC}} \left(\frac{\partial \mu_b}{\partial N_a^{\mathrm{GC}}} \right)_{T, V, \mathbf{N}_a^{\mathrm{GC}}}$$
$$= \beta \rho_a^{\mathrm{GC}} \left(\frac{\partial \mu_b}{\partial \rho_a^{\mathrm{GC}}} \right)_{T, V, \rho_a^{\mathrm{GC}}}. \quad (21)$$

We have thus obtained the thermodynamic factors as functions of $(T, V, \mathbf{N}^{\mathrm{GC}}, \text{shape})$ in a few steps directly from particle fluctuations. Other thermodynamic properties can be found likewise.[3,22] The formulas above reduce for a pure component to Ref. 3,

$$B_{11} = \rho_1^2 G_{11} + \rho_1, \tag{22}$$

$$\kappa_T = \frac{1 + \rho_1 G_{11}}{k_B T \rho_1}, \tag{23}$$

$$\left(\frac{\partial \mu_1}{\partial N_1}\right)_{T,V} = \frac{k_B T N_1}{1 + \rho_1 G_{11}}. \tag{24}$$

Here, κ_T is the isothermal compressibility and ρ_1 is the density of the component. Ben-Naim[3] provides these relations for the thermodynamic limit only, but it is important to note that these relations are also valid for any finite volume.

For a binary mixture, the corresponding information is obtained from the following equations:

$$\eta = \rho_1 + \rho_2 + \rho_1 \rho_2 (G_{11} + G_{22} - 2G_{12}), \tag{25}$$

$$\zeta = 1 + \rho_1 G_{11} + \rho_2 G_{22} + \rho_1 \rho_2 (G_{11} G_{22} - G_{12}^2). \tag{26}$$

In terms of these quantities, we have

$$B_{ab} = \rho_a \rho_b G_{ab} - \delta_{ab} \rho_a, \tag{27}$$

$$\kappa_T = \frac{\zeta}{k_B T \eta}, \tag{28}$$

$$V_1 = \frac{1 + \rho_2 (G_{22} - G_{12})}{\eta}, \tag{29}$$

$$V_2 = \frac{1 + \rho_1 (G_{11} - G_{12})}{\eta}, \tag{30}$$

where $V_i = (\partial V / \partial N_i)_{T,p,\mathbf{N}_i}$ is the partial molar volume of component i at constant pressure, p. These relations also apply to a small system with a finite volume. Ben-Naim also provides the inverse relations.[3] Thermodynamic properties can be, thus, computed from KBIs, but KBIs can also be computed from thermodynamic properties both for large and small systems. In summary, we have explained how thermodynamic properties, also for small systems, can be computed from KBIs, and how KBIs can be computed from thermodynamic properties.

From the symmetry of the $\mathbf{A}^{GC}$-matrix, it follows that $\Gamma_{ab}^{GC} \rho_b^{GC} = \Gamma_{ba}^{GC} \rho_a^{GC}$. The thermodynamic factors can be used to understand solute and solvent properties of small volumes of arbitrary shape, i.e., fluid mixtures in confinement.

For completeness, we also provide the corresponding results for the microcanonical (MC), and the canonical ensembles (C), to indicate that they are different,

$$\Gamma_{ab}^{MC}(U, V, \mathbf{N}, \text{shape}) \equiv \rho_a A_{ab}^{MC}(U, V, \mathbf{N}, \text{shape})$$

$$= \frac{N_a}{k_B T^{MC}} \left(\frac{\partial \mu_b^{MC}}{\partial N_a}\right)_{U,V,\mathbf{N}_a} = \frac{\rho_a}{k_B T^{MC}} \left(\frac{\partial \mu_b^{MC}}{\partial \rho_a}\right)_{U,V,\rho_a} \tag{31}$$

and

$$\Gamma_{ab}^{C}(T, V, \mathbf{N}, \text{shape}) \equiv \rho_a A_{ab}^{C}(T, V, \mathbf{N}, \text{shape})$$

$$= \beta N_a \left(\frac{\partial \mu_b^{C}}{\partial N_a}\right)_{T,V,\mathbf{N}_a} = \beta \rho_a \left(\frac{\partial \mu_b^{C}}{\partial \rho_a}\right)_{T,V,\rho_a}. \tag{32}$$

Inversions of the $\mathbf{A}^{MC}$ and the $\mathbf{A}^{C}$ matrices provide the $\mathbf{B}^{MC}$ $(U, V, \boldsymbol{\mu}^{MC}, \text{shape})$ and the $\mathbf{B}^{C}(T, V, \boldsymbol{\mu}^{C}, \text{shape})$ matrices. The above procedure is required, as these matrices are not directly obtainable from the fluctuations in particle numbers, unlike in the grand-canonical ensemble, cf. Eq. (16).

We are finally in a position to explain how the Small System Method[7,8] can further be applied, in practice, to find system properties at any scale, away from the thermodynamic limit. For this, we expand the property in question in the inverse characteristic length $L \equiv \sqrt[3]{V}$ of the system, as explained by Schnell *et al.*[7,8] From the particle number fluctuations, cf. Eq. (16), we obtain the matrix $\mathbf{B}^{GC}$, which is expanded as follows:

$$\mathbf{B}^{GC}(T, V, \boldsymbol{\mu}, \text{shape}) = \mathbf{B}^{\infty} + \mathbf{B}^{s} \frac{1}{L} + \mathbf{B}^{se} \frac{1}{L^2} + \mathbf{B}^{sc} \frac{1}{L^3} + \cdots. \tag{33}$$

All expansion coefficients depend on $(T, V, \boldsymbol{\mu})$ and the shape of the volume as well. The number of terms needed in the expansion depends on the system considered,[22] the $\mathbf{B}$ exponents s, se, sc stand for surface, edge, and corner contributions, respectively. The reason to use the expansion of $\mathbf{B}^{GC}$ in $1/L$ is that particle number fluctuations are additive. This improves the convergence in the thermodynamic limit. The thermodynamic limit values of the KBIs, G_{ab}^{∞} and B_{ab}^{∞}, are obtained from this expansion in $1/L$. They depend on the environmental variables $(T, V, \boldsymbol{\mu})$, but no longer on the shape. The $\mathbf{A}^{\infty}$- matrix and the corresponding thermodynamic factors $\boldsymbol{\Gamma}^{\infty}$ are found by inversion of the $\mathbf{B}^{\infty}$ matrix. Kirkwood and Buff[1] explained how we can find all thermodynamic properties of a mixture in the large system limit. In this limit (only), we can change from one to another set of environmental variables using Legendre transforms. Away from this limit, Legendre-Fenchel transforms have shown to be useful.[40]

In this section, we have shown how the formulas central to KB theory also can be used to obtain thermodynamic data for small systems. The procedure presented has to a large extent not yet been applied. Several systematic investigations may be worth while pursuing. For pure systems, one may consider the effect of shape variation on various properties, while for mixtures, the inverse size of the volume is expected to influence thermodynamic properties.

IV. APPLICATIONS AND PERSPECTIVES

The analysis of molecule fluctuations in a system in terms of thermodynamic properties and its link with the structural properties through KBIs, and RDF, is one of the first and major results of statistical mechanics. The possibility to simulate, explore, validate, and use these properties has been proposed since emergence of classical molecular simulations. A lot has been done to obtain access to the RDFs in an infinite range with good accuracy, as summarized in previous sections. Since ten years, it has been shown that the approach of KBIs to thermodynamic limit follows a well-defined size-dependence. The use of this dependence offers an efficient method, the SSM, to extrapolate small system thermodynamic

properties to the thermodynamic limit. Furthermore, it gives the possibility to investigate the physical interpretation of the KBI and related properties whatever the size range. Beyond the number of particle fluctuations (KBI), energy fluctuations also show this dependence and were used to get access to partial molar energies and enthalpies.[41] It offers new statistical tools that can be used to investigate challenging systems and phenomena.

KBI was built by Kirkwood and Buff as a solution theory that can be applied to pure systems and to mixtures. It is well adapted to analyze liquid systems and, in particular, to get access to the properties of both the whole system and individual species. In the context of solution theory, the applications are very broad and we will restrict ourselves to underline its importance in the context of transport properties for the determination of the diffusion coefficients. A review of other applications of KBI is provided in Ref. 35.

As was shown in Sec. III, the KB theory has a strong statistical basis and the question of applying KBIs to other type of systems, such as heterogeneous ones, is well opened. For example, in the case of hydrophobic drugs, Shimizu and Nagai Kanasaki in 2019[18] investigated the solute-hydrotrope affinity and its consequences on the balance between self-aggregation and solubilization. Different authors[19–21] discussed the effect of the cosolvent on the protein stability using KBIs as efficient tools. Tripathy *et al.* in 2020[17] studied the water density fluctuations around a generic hydrophobic polymer chain hydration shell and were able to compute the hydration shell compressibility using the SSM approach. Each type of systems has specificities that need to be taken into account in order to apply properly KB theory. In the following, we will restrict to give some insights and perspectives on the concept of finite volume KB theory, we will illustrate and discuss the use of KBIs on confined and adsorbed systems and on solids.

A. Self and collective diffusion coefficients in fluids and their finite-size effects

An important application of Kirkwood–Buff integrals is their use in computing transport diffusion coefficients from Molecular Dynamics (MD) simulations.[39,42–44] Transport diffusion coefficients are used to describe the transport of species due to a gradient in concentration or chemical potential. The thermodynamic driving force for such mass transfer is the gradient in chemical potential, but in experiments and MD simulations one only has direct access to concentration gradients. Converting gradients in concentration to gradients in chemical potential requires partial derivatives of activity coefficients, which can be computed directly from Kirkwood–Buff integrals.

1. Finite-size effect for self-diffusion

In the limit of very low concentrations, the process of transport diffusion reduces to self-diffusion.[45] Self-diffusion describes the motion of individual molecules in a system as a function of time.[25] It is well known that computed self-diffusivities have finite-size effects, i.e., for the same thermodynamic state the value of the computed self-diffusivity changes with system size.[46–49] In 1993, Dünweg and Kremer[47] and later in 2004, Ye and Hummer derived the following equation for the finite-size-dependence of self-diffusivities in three-dimensional cubic systems with periodic boundary conditions based on hydrodynamic arguments,[48]

$$D_{\text{self},a}^{\infty} = D_{\text{self},a}^{L} + \frac{k_B T \xi}{6\pi\eta L} = D_{\text{self},a}^{L} + D^{\text{YH}}, \tag{34}$$

in which $D_{\text{self},a}^{\infty}$ is the self-diffusivity of species a in the thermodynamic limit, $D_{\text{self},a}^{L}$ is the computed self-diffusivity of a system with simulation box length L (and volume L^3), and η is the viscosity of the system (which does not depend on L). The constant ξ has the dimensionless value of 2.837 298 for periodic cubic simulation boxes. Expressions for finite-size effects of self-diffusivities in confinement and rotational diffusion are also available.[50–52]

2. Collective diffusion

Collective diffusion (or transport diffusion) describes the transport of a large collection of molecules due to a driving force, e.g., a gradient in concentration or chemical potential.[45,53,54] The most well-known approach for transport diffusion is generalized Fick's law. For an n-component system in a molar reference frame, the diffusion flux J_a for component a equals[45]

$$J_a = -c_t \sum_{b=1}^{n-1} D_{ab} \nabla x_b, \tag{35}$$

in which c_t is the total molar concentration, x_b is the mole fraction of component b, and D_{ab} are the Fick diffusion coefficients. In a binary system ($n = 2$), there is a single Fick diffusion coefficient D that does not depend on the choice of the reference frame. We refer to Refs. 53 and 55 for other reference frames.

An alternative formulation for multicomponent diffusion is the Maxwell–Stefan (MS) approach, in which gradients in chemical potentials are considered as driving forces (instead of concentration gradients),[45]

$$-\frac{1}{RT}\nabla_{T,P}\mu_a = \sum_{b=1,b\neq a}^{n} \frac{x_b(u_a - u_b)}{Đ_{ab}}, \tag{36}$$

in which R is the universal gas constant, $\nabla_{T,P}\mu_a$ is the gradient in chemical potential of component a at constant temperature and pressure, and u_a is the average molar velocity of component a. For an n-component system, there are $n(n-1)/2$ Maxwell–Stefan diffusion coefficients, which (unlike the Fick diffusivities D_{ab}) are symmetric, i.e., $Đ_{ab} = Đ_{ba}$. MS diffusion coefficients can be computed from both equilibrium and non-equilibrium Molecular Dynamics simulations.[42] The Fick and Maxwell–Stefan approaches describe the same physical phenomena, so their diffusion coefficients D_{ab} and $Đ_{ab}$ are related. The Fick diffusivities in a molar reference frame are obtained from[45]

$$[D] = [\Delta][\Gamma], \tag{37}$$

in which the matrix $[\Gamma]$ is the so-called matrix of thermodynamic factors for diffusion[56] and $[\Delta]$ is a matrix function of MS diffusion coefficients.

The matrix $[B]$ is defined by the inverse of matrix $[\Delta]$, so $[B] = [\Delta]^{-1}$. MS diffusion coefficients are related by the elements of the matrix $[B]$ according to

$$B_{aa} = \frac{x_a}{\mathcal{D}_{an}} + \sum_{j=1, b\neq a}^{n} \frac{x_b}{\mathcal{D}_{ab}}, \tag{38}$$

for $a = 1, 2, \ldots, (n-1)$ and

$$B_{ab} = -x_a\left(\frac{1}{\mathcal{D}_{ab}} - \frac{1}{\mathcal{D}_{an}}\right), \tag{39}$$

for $a, b = 1, 2, \ldots, (n-1)$ and $b \neq a$. The resulting expressions for MS diffusion coefficients of binary, ternary, and quaternary systems can be found in Refs. 42, 43, and 57.

The elements of matrix of thermodynamic factor for diffusion, $[\Gamma]$, are[53]

$$\Gamma_{ab} = \delta_{ab} + x_a\left(\frac{\partial \ln \gamma_a}{\partial x_b}\right)_{T,p,\Sigma}, \tag{40}$$

in which γ_a is the activity coefficient of component a. The symbol Σ is used to point out that the partial derivative of the logarithm of the activity coefficient is performed at constant mole fractions of all other component in the system, except for the n-th component. It is important to note that the thermodynamic factor for diffusion, Eq. (40), is defined differently than the thermodynamic factor as defined in Eq. (21). The main difference is that the quantities that are kept constant in the differentiation are different. Both definitions are widely used in the literature, so it is important to make a clear distinction. It is important to note that for binary and multicomponent systems, the derivatives of Eq. (40) follow directly from Kirkwood–Buff coefficients.[3] For various activity coefficient models (i.e., $\ln \gamma_a$ as a function of composition), expressions for Γ_{ab} are available from Ref. 56. For a binary system, the thermodynamic factor for diffusion equals

$$\Gamma = 1 + x_1\left(\frac{\partial \ln \gamma_1}{\partial x_1}\right)_{T,p} = 1 + x_2\left(\frac{\partial \ln \gamma_2}{\partial x_2}\right)_{T,p}. \tag{41}$$

The thermodynamic factor for the diffusion of binary systems follows directly from the Kirkwood–Buff integrals according to Ref. 3,

$$\Gamma = 1 - \frac{\rho_1\rho_2\Omega_{12}}{\rho_1 + \rho_2 + \rho_1\rho_2\Omega_{12}}, \tag{42}$$

in which $\rho_a = N_a/L^3$ is the number density of species a and the auxiliary quantity Ω_{12} is defined as

$$\Omega_{12} = G_{11} + G_{22} - 2G_{12}, \tag{43}$$

and G_{ab} is the Kirkwood–Buff integral in the thermodynamic limit. Explicit expressions for Γ_{ab} for ternary and quaternary systems can be found in Refs. 14 and 43. The value of the thermodynamic factor for diffusion describes how much molecules of species 1 and 2 like each-other, compared with the $1-1$ and $2-2$ interactions. A condition for thermodynamic stability of binary systems is $\Gamma > 0$.[45] Close to the value $\Gamma = 0$, the system will demix in the individual components.

3. Finite-size effects for collective diffusion in binary systems

In binary systems, there is only a single Fick diffusion coefficients D and a single MS diffusion coefficient $\mathcal{D}_{12}$. The Fick diffusivity D equals

$$D = \Gamma \mathcal{D}_{12}. \tag{44}$$

Jamali *et al.* have developed a phenomenological finite-size correction for MS diffusivities in binary systems,[58]

$$\mathcal{D}_{12}^{\infty} = \mathcal{D}_{12}^L + \frac{1}{\Gamma}\frac{k_B T\xi}{6\pi\eta L} = \mathcal{D}_{12}^L + \frac{D^{YH}}{\Gamma}, \tag{45}$$

in which D^{YH} is the YH correction [Eq. (34)]. From Eq. (44), the finite-size correction for binary Fick diffusivities equals

$$D^{\infty} = D^L + D^{YH}. \tag{46}$$

This means that the YH correction should be directly applied to Fick diffusivities (as is done for self-diffusivities), and the corresponding finite-size effect for the MS diffusivity differs by a factor $1/\Gamma$. Therefore, finite-size effects of computed MS diffusivities can be very large if Γ is small, i.e., close to demixing conditions. For binary systems, the thermodynamic factor for diffusion can be calculated by computing finite-size effects of transport properties: the viscosity of the system follows from the finite-size-dependence of the self-diffusivities, and the thermodynamic factor for diffusion follows from the finite-size-dependence of the computed MS diffusivities. This could serve as an independent check for Γ computed from KBIs. The finite-size-dependence of diffusion coefficients for systems with three or more components is discussed in Refs. 49 and 59.

B. Application of nanothermodynamics to confined fluids

The expansion expressed by Eq. (33) originates in the nanothermodynamic theory of Hill.[23] It derives from the way Hill includes shape and size as a variables in the thermodynamic description of small systems.[24] Such systems are not extensive in Gibbs sense, but extensivity in the description can be restored by taking an ensemble of small systems into account. By doing this, Bedeaux *et al.*[22] were able to find scaling laws, characteristic for the ensemble in question. For instance, in the description of porous media, a scaling law was derived from the grand potential, giving the difference of the so-called integral pressure and the normal (differential) pressure times the volume in terms of the subdivision potential of the ensemble.[60–62] The difference between the integral and differential pressures is related to the disjoining pressure.[61] The law was studied for slit pores and cylindrical pores and a first example of a simple porous medium. While Legendre transforms apply to ensembles of large systems only, it was found that Legendre-Fenchel transforms were useful in two cases.[40] Whether or not they are generally valid remains to be investigated.

Clearly, there are new tools becoming available and "new thermodynamics" at the nanoscale waiting to be explored. This applies to ensembles of varying sets of control variables, to particular shapes and sizes, as well as to other small scale phenomena, where classical approaches find their limits.

C. Application to surfaces

An important class of phenomena that relates directly to confinement is adsorption on a surface. For molecules adsorbed on a flat surface, a two dimensional (2D) KBI can be easily determined using the 2D number fluctuations of a surface or simply knowing the 2D surface RDF. In such a case, the weight-function w has a different expression[9] than Eq. (10), for a circular surface we have

$$w(r, L) = 4r\left(\arccos(x) - x(1 - x^2)^{1/2}\right), \quad x \equiv r/L. \tag{47}$$

It should also be pointed out that for the specific case of an adsorbed layer on a surface, which is in chemical equilibrium with its environment, grand-canonical conditions apply, and canonical condition corrections should not be used. As was illustrated in Ref. 63, such an approach provides a direct way to estimate the chemical activity of adsorbed molecules on flat surfaces. Using a 3D approach, it has also been possible to get access to the adsorption properties of molecules adsorbed in zeolites.[8,64] The case of curved surfaces should, however, be dealt with using finite surface KBI.[65]

D. Extension of KBI theory to solids

The KB theory is a major theory of solutions and has been used extensively in liquid solutions. As the theory relies only on general statistical mechanical relations, it should also be valid and useful for the study of solid solutions. However, the KB theory was not applied to solids until very recently, probably because RKBI diverges in solids. This is related to the fact that in a crystal, the oscillations of the RDF decay very slowly with interatomic distance, because the crystal is periodic and so the atomic positions are correlated with arbitrary distance. The divergence of the RKBI is unphysical as can be seen by considering the compressibility equation,

$$1 + \rho G^\infty = \rho k_B T \kappa_T, \tag{48}$$

which relates the KBI G^∞ to the isothermal compressibility κ_T, where ρ is the number density. Since compressibility of solids is finite, G^∞ is also finite, so KBIs should converge in the limit $V \to \infty$. This is indeed the case, if the finite-volume KBI is used instead of the RKBI, as was recently shown by Miyaji et al.[38] and Krüger.[66] Figure 2 shows the convergence properties of the running KBI and the finite-volume KBI for a perfect fcc crystal at zero temperature. The RDF is shown in (a). The running KBI (b) has huge oscillations whose amplitude increases linearly and the integral diverges. In sharp contrast, the finite volume KBI (c) converges to $-1/\rho$, which is the correct theoretical value at $T = 0$, as can be seen from Eq. (48).

The KB theory exploits the fact that a fluid is homogeneous and isotropic, which implies that the PDF depends only on the particle distance and reduces to the radial distribution function (RDF). A crystalline solid is neither homogeneous nor isotropic, so the PDF and RDF are, in principle, different. An obvious solution to this problem consists in going back to Eq. (1), which is valid in any system. However, this would make the KBI method numerically much more demanding. Under certain assumptions, KB theory can be applied to solids in the same way as in liquids.[38,66,67] This is the case of polycrystal, which has the same symmetry as a fluid, i.e., it is
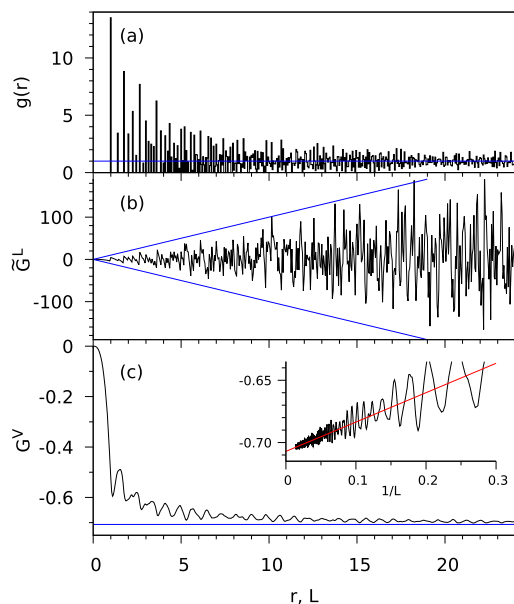

FIG. 2. (a) RDF of a perfect fcc lattice with nearest neighbor distance 1, as a histogram plot with bin size $\Delta r = 0.05$. The blue line indicates the uncorrelated limit, $g = 1$. (b) Running KBI $\tilde{G}^L$. The blue dotted lines are a guide to the eye ($y = \pm 10L$). (c) Finite volume KBI G^V for a sphere of diameter L. The blue line is the exact, infinite volume value $G^\infty = -1/\rho$ with $\rho = \sqrt{2}$. In the inset, G^V is plotted over $1/L$. The red line is a linear fit, $y = G^\infty + \alpha/L$ with $\alpha = \rho/3$.

homogeneous and isotropic. Krüger[66] has applied the KB theory to a monoatomic, harmonic crystal. The obtained isothermal compressibility agrees perfectly with continuum theory (Newton–Laplace equation). This proves that the compressibility equation [Eq. (48)] is exactly valid in harmonic solids with isotropic, linear phonon dispersion, provided that G^∞ is calculated using finite-volume KBI.

Miyaji et al.[38] have presented the first numerical application of KBIs to solids with a realistic model, namely, fcc argon with a Lennard-Jones potential. While RKBI strongly diverged, the finite-volume KBI converged for all temperatures. Compared with the liquid state, however, the convergence is very slow, which makes extrapolation to the thermodynamic limit difficult. To solve this problem, Miyaji et al.[38] introduced a convolution of the RDF, which leaves the infinite volume integral (G^∞) unchanged. The point-like particles in the usual definition of the RDF are replaced by spheres of finite diameter σ and constant density $6/(\pi\sigma^3)$. The parameter σ can be chosen arbitrarily, without changing the value of G^∞. The corresponding RDF can be written as a convolution,

$$\tilde{g}(r) = \int_{\max(0, r-\sigma)}^{r+\sigma} g(r')\chi(r', r)dr', \tag{49}$$

where $g(r)$ is the original RDF and the convolution function $\chi(r', r)$ has a simple analytic form.[38] When choosing the value of σ in the order of the average particle distance, the convoluted RDF $\tilde{g}(r)$ is extremely smooth, and the convergence of the KBI is dramatically improved. This method was applied to solid argon for temperatures

between 15 and 75 K. KBIs could be computed with good accuracy and G^∞ was converged to better than 1%. The computed isothermal compressibility from KBIs was somewhat underestimated but its temperature dependence agreed very well with the experiment. Recently, Miyaji et al.[68] presented the first application of the KB theory to a solid solution, namely, $Ar_x Xe_{1-x}$ for $x < 0.1$ at temperatures around 80 K. The isothermal compressibility of the mixture, the partial molar volumes of Ar and Xe, and the thermodynamic factor Γ were obtained from KBIs using Monte Carlo simulations. Additionally, the activity coefficients of each species were computed by integrating Γ. The analysis of the thermodynamic results evidenced the emergence of a liquid state around $x \approx 0.1$.

In summary, it was shown that the KB theory can be applied to solids, but only if the finite-volume KBIs are used instead of the RKBI. Both single atom crystals and solid solutions were studied successfully, but a few problems remain to be solved. The systematic deviation in the isothermal compressibility needs to be understood and corrected in a non-empirical way. The difference might be induced by a small shift of the simulated RDF, which converges very slowly in crystals. In addition, the assumption that solids can be described by a homogeneous, isotropic statistical ensemble should be critically reexamined. It is hoped that the possibility to compute KBIs in solids will open up many opportunities for thermodynamic studies of solid condensed materials.

V. CONCLUSIONS

This paper presents a perspective of our contribution to the KB theory. It emphasizes the finite volume dependence of KBIs and sheds light on its physical meaning. The linear $1/L$ size-dependence of KBIs is an important property that has been used to extrapolate KBIs to the thermodynamic limit without knowing the RDF to infinite size, the so-called small system method. In this limit, the KB theory provides a direct access to thermodynamic quantities such as the thermodynamic factor, Γ, which can be understood as a measure of non-ideality. This quantity is crucial to compute accurate mutual diffusion coefficients. We presented new applications and perspectives for surfaces, confined systems, and solids. This last is particularly promising since it offers an efficient route to compute partial molar quantities of solid mixtures that are difficult to access otherwise. Below the thermodynamic limit, the thermodynamic relations have to be rewritten by introducing new concepts such as those originated from nanothermodynamics. KBIs can be considered as a key function for this purpose because they relate to clear thermodynamic quantities whatever the size and shape of finite volume.

To conclude, it is important to underline that KB theories provide a rigorous thermodynamic background that can be applied to complex systems both at the small and large scales. The recent developments enlarge its domain of application to more complex ones and give new useful statistical tools to better understand dynamic phenomena such as nucleation or confinement.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

J.-M. Simon: Writing – original draft (equal). **P. Krüger**: Writing – original draft (equal). **S. K. Schnell**: Writing – original draft (equal). **T. J. H. Vlugt**: Writing – original draft (equal). **S. Kjelstrup**: Writing – original draft (equal). **D. Bedeaux**: Writing – original draft (equal).

DATA AVAILABILITY

The data that support the findings of this study are available within the article.

REFERENCES

[1] J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).
[2] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1975).
[3] A. Ben-Naim, *Molecular Theory of Solutions*, 1st ed. (Oxford University Press, Oxford, 2006).
[4] M. Heidari, K. Kremer, R. Potestio, and R. Cortes-Huerto, Mol. Phys. **116**, 3301 (2018).
[5] L. Belloni, J. Chem. Phys. **149**, 094111 (2018).
[6] B. Stenqvist, V. Aspelin, and M. Lund, J. Chem. Theory Comput. **16**, 3737 (2020).
[7] S. K. Schnell, T. J. H. Vlugt, J.-M. Simon, D. Bedeaux, and S. Kjelstrup, Chem. Phys. Lett. **504**, 199 (2011).
[8] S. K. Schnell, X. Liu, J.-M. Simon, A. Bardow, D. Bedeaux, T. J. H. Vlugt, and S. Kjelstrup, J. Phys. Chem. B **115**, 10911 (2011).
[9] P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, and J.-M. Simon, J. Phys. Chem. Lett. **4**, 235 (2013).
[10] R. Cortes-Huerto, K. Kremer, and R. Potestio, J. Chem. Phys. **145**, 141103 (2016).
[11] M. Sevilla and R. Cortes-Huerto, J. Chem. Phys. **156**, 044502 (2022).
[12] E. A. Oprzeska-Zingrebe and J. Smiatek, Biophys. J. **114**, 1551 (2018).
[13] P. S. Schmalhorst, F. Deluweit, R. Scherrers, C.-P. Heisenberg, and M. Sikora, J. Chem. Theory Comput. **13**, 5039 (2017).
[14] R. Fingerhut, G. Herres, and J. Vrabec, Mol. Phys. **118**, e1643046 (2020).
[15] F. Venetsanos, S. D. Anogiannakis, and D. N. Theodorou, Macromolecules **55**, 4852 (2022).
[16] S. K. Schnell, P. Englebienne, J.-M. Simon, P. Krüger, S. P. Balaji, S. Kjelstrup, D. Bedeaux, A. Bardow, and T. J. H. Vlugt, Chem. Phys. Lett. **582**, 154 (2013).
[17] M. Tripathy, S. Bharadwaj, S. J. B., and N. F. A. van der Vegt, Nanomaterials **10**, 1460 (2020).
[18] S. Shimizu and Y. Nagai Kanasaki, J. Mol. Liq. **274**, 209 (2019).
[19] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith, Cell Biochem. Biophys. **50**, 1 (2008).
[20] M. B. Gee and P. E. Smith, J. Chem. Phys. **131**, 165101 (2009).
[21] S. Shimizu, Chem. Phys. Lett. **514**, 156 (2011).
[22] D. Bedeaux, S. Kjelstrup, and S. K. Schnell, *Nanothermodynamics General Theory* (PoreLab: Trondheim, Norway, 2020).
[23] T. L. Hill, Nano Lett. **1**, 273 (2001).

[24] T. L. Hill, *Thermodynamics of Small Systems* (Courier Corporation, 1994).

[25] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed. (Academic Press, Cambridge, MA, 2002), Vol. 1.

[26] D. Borgis, R. Assaraf, B. Rotenberg, and R. Vuilleumier, Mol. Phys. **111**, 3486 (2013).

[27] D. de las Heras and M. Schmidt, Phys. Rev. Lett. **120**, 218001 (2018).

[28] S. H. Jamali, "Transport properties of fluids: Methodology and force field improvement using molecular dynamics simulations," Ph.D. thesis, Delft University of Technology, Delft, 2020.

[29] P. Ganguly and N. F. A. van der Vegt, J. Chem. Theory Comput. **9**, 1347 (2013).

[30] N. Dawass, P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, J. M. Simon, and T. J. H. Vlugt, Mol. Simul. **44**, 599 (2018).

[31] L. Verlet, Phys. Rev. **165**, 201 (1968).

[32] R. Wedberg, J. P. O'Connell, G. H. Peters, and J. Abildskov, Mol. Simul. **36**, 1243 (2010).

[33] N. Dawass, P. Krüger, J.-M. Simon, and T. J. H. Vlugt, Mol. Phys. **116**, 1573 (2018).

[34] P. Krüger and T. J. H. Vlugt, Phys. Rev. E **97**, 051301(R) (2018).

[35] N. Dawass, P. Krüger, S. K. Schnell, J.-M. Simon, and T. J. H. Vlugt, Fluid Phase Equilib. **486**, 21 (2019).

[36] S. Giorgini, L. P. Pitaevskii, and S. Stringari, Phys. Rev. Lett. **80**, 5040 (1998).

[37] A. Santos, Phys. Rev. E **98**, 063302 (2018).

[38] M. Miyaji, B. Radola, J.-M. Simon, and P. Krüger, J. Chem. Phys. **154**, 164506 (2021).

[39] G. Guevara-Carrion, R. Fingerhut, and J. Vrabec, J. Phys. Chem. B **124**, 4527–4535 (2020).

[40] O. Galteland, E. Bering, K. Kristiansen, D. Bedeaux, and S. Kjelstrup, Nanoscale Adv. **4**, 2660 (2022).

[41] S. K. Schnell, R. Skorpa, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, and J.-M. Simon, J. Chem. Phys. **141**, 144501 (2014).

[42] X. Liu, S. K. Schnell, J.-M. Simon, P. Krüger, D. Bedeaux, S. Kjelstrup, A. Bardow, and T. J. H. Vlugt, Int. J. Thermophys. **34**, 1169 (2013).

[43] X. Liu, A. Martín-Calvo, E. McGarrity, S. K. Schnell, S. Calero, J.-M. Simon, D. Bedeaux, S. Kjelstrup, A. Bardow, and T. J. H. Vlugt, Ind. Eng. Chem. Res. **51**, 10247 (2012).

[44] S. Pařez, G. Guevara-Carrion, H. Hasse, and J. Vrabec, Phys. Chem. Chem. Phys. **15**, 3985 (2013).

[45] R. Krishna and J. A. Wesselingh, Chem. Eng. Sci. **52**, 861 (1997).

[46] D. M. Heyes, M. J. Cass, J. G. Powles, and W. A. B. Evans, J. Phys. Chem. B **111**, 1455 (2007).

[47] B. Dünweg and K. Kremer, J. Chem. Phys. **99**, 6983 (1993).

[48] I.-C. Yeh and G. Hummer, J. Phys. Chem. B **108**, 15873 (2004).

[49] S. H. Jamali, A. Bardow, T. J. H. Vlugt, and O. A. Moultos, J. Chem. Theory Comput. **16**, 3799 (2020).

[50] P. Simonnin, B. Noetinger, C. Nieto-Draghi, V. Marry, and B. Rotenberg, J. Chem. Theory Comput. **13**, 2881 (2017).

[51] M. Linke, J. Köfinger, and G. Hummer, J. Phys. Chem. B **122**, 5630 (2018).

[52] M. Vögele, J. Köfinger, and G. Hummer, J. Phys. Chem. B **123**, 5099 (2019).

[53] R. Taylor and R. Krishna, *Multicomponent Mass Transfer*, 1st ed. (John Wiley & Sons, 1993), Vol. 2.

[54] S. Kjelstrup and D. Bedeaux, *Non-equilibrium Thermodynamics of Heterogeneous Systems*, 1st ed. (World Scientific, 2008), Vol. 16.

[55] J. M. Ortiz de Zárate and J. V. Sengers, Phys. Chem. Chem. Phys. **22**, 17597 (2020).

[56] R. Taylor and H. A. Kooijman, Chem. Eng. Commun. **102**, 87 (1991).

[57] R. Krishna and J. M. Van Baten, Ind. Eng. Chem. Res. **44**, 6939 (2005).

[58] S. H. Jamali, L. Wolff, T. M. Becker, A. Bardow, T. J. H. Vlugt, and O. A. Moultos, J. Chem. Theory Comput. **14**, 2667 (2018).

[59] A. T. Celebi, S. H. Jamali, A. Bardow, T. J. H. Vlugt, and O. A. Moultos, Mol. Simul. **47**, 831 (2021).

[60] M. T. Rauter, O. Galteland, M. Erdős, O. A. Moultos, T. J. H. Vlugt, S. K. Schnell, D. Bedeaux, and S. Kjelstrup, Nanomaterials **10**, 608 (2020).

[61] O. Galteland, D. Bedeaux, and S. Kjelstrup, Nanomaterials **11**, 165 (2021).

[62] O. Galteland, D. Bedeaux, B. Hafskjold, and S. Kjelstrup, Front. Phys. **7**, 60 (2019).

[63] T. T. Trinh, D. Bedeaux, J.-M. Simon, and S. Kjelstrup, Phys. Chem. Chem. Phys. **17**, 1226 (2015).

[64] N. Floquet, J. M. Simon, J. P. Coulomb, J. P. Bellat, G. Weber, and G. Andre, Micropourous Mesoporous Mater. **122**, 61 (2009).

[65] B. A. Strøm, D. Bedeaux, and S. K. Schnell, Nanomaterials **11**, 431 (2021).

[66] P. Krüger, Phys. Rev. E **103**, L061301 (2021).

[67] P. Krüger, Phys. Rev. B **101**, 205423 (2020).

[68] M. Miyaji, J.-M. Simon, and P. Krüger, Physchem **2**, 191–206 (2022).