Kirkwood-Buff integrals from molecular simulation

Dawass, Noura; Krüger, Peter; Schnell, Sondre K.; Simon, Jean Marc; Vlugt, T. J.H.

DOI
10.1016/j.fluid.2018.12.027

Publication date
2019

Document Version
Accepted author manuscript

Published in
Fluid Phase Equilibria

Kirkwood-Buff Integrals from Molecular Simulation

Noura Dawass[a], Peter Krüger[b], Sondre K. Schnell[d], Jean-Marc Simon[c], T.J.H. Vlugt[a,*]

[a] *Process & Energy Laboratory, Delft University of Technology, Leeghwaterstraat 39, 2628CB Delft, The Netherlands*
[b] *Graduate School of Science and Engineering, Molecular Chirality Research Center, Chiba University, Chiba 263-8522, Japan*
[c] *ICB, UMR 6303 CNRS - Université de Bourgogne Franche-Comté, F-21078, Dijon, France*
[d] *Department of Materials Science and Engineering, NTNU, N-7491, Trondheim, Norway*

Abstract

The Kirkwood-Buff (KB) theory provides a rigorous framework to predict thermodynamic properties of isotropic liquids from the microscopic structure. Several thermodynamic quantities relate to KB integrals, such as partial molar volumes. KB integrals are expressed as integrals of RDFs over volume but can also be obtained from density fluctuations in the grand-canonical ensemble. Various methods have been proposed to estimate KB integrals from molecular simulation. In this work, we review the available methods to compute KB integrals from molecular simulations of finite systems, and particular attention is paid to finite-size effects. We also review various applications of KB integrals computed from simulations. These applications demonstrate the importance of computing KB integrals for relating findings of molecular simulation to macroscopic thermodynamic properties of isotropic liquids.

Keywords: Kirkwood-Buff Theory, Kirkwood-Buff Integrals, Small System Method, Solution Theory, Molecular Simulations, Density fluctuations

*Corresponding author
Email address: t.j.h.vlugt@tudelft.nl (T.J.H. Vlugt)

1. Introduction

The prediction of thermodynamic properties of multicomponent isotropic fluids from molecular information is of a great interest [1–6]. Molecular-based methods provide predictions for experimental thermodynamic and transport data and, contribute to developing predictive models, both needed for many industrial applications [7]. In that regard, the Kirkwood-Buff (KB) theory provides an important connection between the microscopic structure of isotropic liquid mixtures and the corresponding macroscopic properties [8]. Kirkwood and Buff [8] expressed thermodynamic quantities such as partial derivatives of the chemical potential with respect to composition, partial molar volumes, and isothermal compresibility in terms of integrals of radial distribution functions (RDFs) over infinite and open volumes. These integrals, which are considered the key quantity in the KB theory, are referred to as KB integrals. Alternatively, KB integrals can be obtained from density fluctuations in the grand-canonical ensemble [9, 10]. Rooted in statistical mechanics, the theory applies to any type of intermolecular interactions, making it one of the most general and important theories of isotropic fluids [8, 9, 11, 12].

The KB theory was derived in 1951 [8], however, it has not gained much interest until the late 70s after Ben-Naim [11] proposed the inversion of the KB theory. The inversion of the theory allows the calculation of KB integrals from experimental data [13–16]. KB integrals provide useful information about the local inhomogeneity and the affinity between components [11]. Later, other than using macroscopic experimental data, KB integrals were also obtained using small angle scattering techniques [17–22]. Perera et al. [23] concluded that KB integrals obtained from scattering experiments generally agree well with those calculated using the inversion method.

Thirty years following the inversion of the KB theory, molecular simulation emerged as a powerful tool for studying pure liquids and liquid mixtures [24]. There are two main types of molecular simulation techniques [25, 26]: Molecular Dynamics (MD), where trajectories of molecules are found by solving Newton's

equation of motion numerically; and Monte Carlo (MC) simulations where relevant states of the system are sampled according to their statistical weight [24–28]. In both simulation techniques, RDFs and local density fluctuations are easily computed, thus in principle enabling the calculation of KB integrals. Molecular simulations can be used to study closed systems with fixed number of molecules, or open systems in which the number of molecules fluctuate [25]. It is important to note that molecular simulations can only be performed for finite systems, while the KB theory requires KB integrals to be computed for infinite and open systems [8]. This has important consequences as will be shown below. When computing KB integrals from molecular simulations, it is common to simply truncate the KB integrals at a finite distance, corresponding to the size of the simulation box [22, 29–32]. This results in KB integrals that converge poorly to the thermodynamic limit, and we will discuss the underlying physical reasons [9, 33, 34]. Many studies have recognized the disparity between KB integrals computed from molecular simulation of finite systems and the integrals defined by Kirkwood and Buff [9, 35, 36]. A practical approach to deal with finite-size effects was proposed by Krüger and co-workers [35, 37] where an expression for KB integrals for finite volumes was derived. The method of Krüger and co-workers [34, 35, 37, 38] has been used in many studies [6, 39–55]. According to Hill's thermodynamics of small systems [56, 57] (also called nanothermodynamics), KB integrals for finite volumes scale with the inverse of the characteristic length scale of the small system and this scaling can be extrapolated to the thermodynamic limit (i.e to KB integrals as defined by Kirkwood and Buff). Several other methods were proposed to compute KB integrals from molecular simulation of finite systems [36, 58–60]. The basic idea is to extrapolate information provided from RDFs and density fluctuations from finite systems to the thermodynamic limit, and thus obtain estimates for KB integrals corresponding to the thermodynamic limit.

The objective of this paper is to review the available methods for computing KB integrals from molecular simulation. The most important applications of KB integrals will be discussed. KB integrals are being used to study different types

of systems from simple, like Lennard-Jones fluids to more complicated systems such as salt solutions [16, 41, 61, 62], alcohol solutions [39, 63–65], and biological systems [31]. For such systems, KB integrals can be used to study a number of thermodynamic properties. As mentioned earlier, the following properties are directly obtained from KB integrals: partial derivatives of the chemical potential with respect to composition, partial molar volumes, and isothermal compresibility. In addition, the KB theory is also useful for relating information obtained from molecular simulation with macroscopic properties. For example, KB integrals can be used to connect Maxwell-Stefan (MS) diffusivities computed by molecular simulations to Fick diffusivities found from experiments [39, 40, 66, 67].

The paper is organized as follows: In section 2, a summary of the most important parts of the KB theory is presented. In section 3, the method of Krüger and co-workers [34, 35, 37, 38] for computing KB integrals from molecular simulation of finite systems is explained. Other methods of calculating KB integrals from finite systems are reviewed in section 4. In section 5, the connection between KB integrals and nanothermodynamics is discussed. In section 6, several important applications of KB integrals computed using molecular simulation are presented. In section 7, the inversion of the KB theory is discussed. Section 8 provides a summary of the main findings of this review.

2. The Kirkwood-Buff theory

In this section, we review the most important relations derived by Kirkwood and Buff [8]. For the original formulation of the theory, the reader is referred to the paper by Kirkwood and Buff [8]. A very detailed derivation was presented by Newman [68], and an alternative derivation was provided by Hall [12].

In the grand-canonical (μTV) ensemble, thermodynamic quantities are related to KB integrals $G_{\alpha\beta}^{\infty}$ for an open and infinite system as [8]:

$$G_{\alpha\beta}^{\infty} = \int_0^{\infty} \mathrm{d}r 4\pi r^2 \left[g_{\alpha\beta}^{\infty}(r) - 1\right] \tag{1}$$

where r is the particle distance and $g^{\infty}_{\alpha\beta}(r)$ is the RDF of species α and β for an infinitely large system. In Eq. (1), species α and β can be the same. For a shell centered around a molecule of type β in an infinite system (see Fig. 1a), the number of molecules of type α is $4\pi r^2 \mathrm{d}r\rho_\alpha$ and $4\pi r^2 \mathrm{d}r\rho_\alpha g^{\infty}_{\alpha\beta}(r)$ for an ideal gas and real fluid, respectively. Here, $\rho_\alpha = \langle N_\alpha \rangle / V$ is the average number density of species α. Integrating to infinity over the excess number of molecules of type α, $(4\pi r^2 \mathrm{d}r\rho_\alpha[g^{\infty}_{\alpha\beta}(r) - 1])$, yields $\rho_\alpha G^{\infty}_{\alpha\beta}$. Hence, KB integrals $G^{\infty}_{\alpha\beta}$ provide the average excess (or depletion) per unit density of α molecules around a β molecule, and they reflect the affinity between components α and β. It is important to note that this interpretation of KB integrals only holds for infinite systems, as indicated by the upper bound of the integral in Eq. (1). Truncating the integral of Eq. (1) to a distance R yields the average excess of type α within a sphere of radius R. We will see later that the resulting integral does not represent the KB integral in the thermodynamic limit.

Kirkwood and Buff [8] formulated a relation between integrals over radial distribution functions and fluctuations in the number of molecules in the grand-canonical ensemble,

$$G^{\infty}_{\alpha\beta} = \int_0^{\infty} \mathrm{d}r 4\pi r^2 \left[g^{\infty}_{\alpha\beta}(r) - 1\right] = \lim_{V\to\infty} \left[V \frac{\langle N_\alpha N_\beta \rangle - \langle N_\alpha \rangle \langle N_\beta \rangle}{\langle N_\alpha \rangle \langle N_\beta \rangle} - \frac{V\delta_{\alpha\beta}}{\langle N_\alpha \rangle} \right] \tag{2}$$

where N_α, and N_β are the number of molecules of type α and β, inside the volume V, and the brackets $\langle \cdots \rangle$ denote the ensemble average in an open system. Hence, $\langle N_\alpha \rangle$ is the average number of molecules and $\langle N_\alpha N_\beta \rangle$ is the average product of the number of molecules of components α and β. $\delta_{\alpha\beta}$ is the Kronecker delta (equal to 1 when $\alpha = \beta$ and is zero otherwise). It is important to note that Eq. (2) holds for any isotropic fluid. Fluctuations in the number of molecules relate to several thermodynamic properties [69, 70]. For a binary system, the following relations can be derived that relate KB integrals to [8]:

1. partial derivatives of chemical potential with respect to the number of

molecules,

$$\left(\frac{\partial \mu_\alpha}{\partial N_\alpha}\right)_{T,P,N_\beta} = \frac{\rho_\beta kT}{\rho_\alpha V \eta}$$
$$\left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha} = \left(\frac{\partial \mu_\beta}{\partial N_\alpha}\right)_{T,P,N_\beta} = -\frac{kT}{V\eta} \tag{3}$$

2. partial molar volumes,

$$\bar{V}_\alpha = \left(\frac{\partial V}{\partial N_\alpha}\right)_{T,P,N_\beta} = \frac{1+\rho_\beta(G_{\beta\beta}-G_{\alpha\beta})}{\eta} \tag{4}$$

3. the isothermal compressibility,

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_{T,N_\alpha,N_\beta} = \frac{\zeta}{kT\eta} \tag{5}$$

where η and ϵ are auxiliary quantities that were defined for convenience [9],

$$\eta = \rho_\alpha + \rho_\beta + \rho_\alpha\rho_\beta(G_{\alpha\alpha}+G_{\beta\beta}-2G_{\alpha\beta}) \tag{6}$$
$$\zeta = 1 + \rho_\alpha G_{\alpha\alpha} + \rho_\beta G_{\beta\beta} + \rho_\alpha\rho_\beta(G_{\alpha\alpha}G_{\beta\beta}-G_{\alpha\beta}^2) \tag{7}$$

Expressions for ternary and multi-component mixtures of these thermodynamic quantities in terms of KB integrals are available in literature [9, 13, 71]. In Eq. (6), the term $G_{\alpha\alpha}+G_{\beta\beta}-2G_{\alpha\beta}$ can be used to indicate the thermodynamic ideality of a binary mixture. It has the value of zero for ideal solutions. The non-ideality of solutions is often quantified by the so-called thermodynamic correction factor Γ [10, 42, 72]. For a binary mixture, $\Gamma_{\alpha\beta}$ is defined as [9, 40, 66]:

$$\Gamma_{\alpha\beta} = 1 - \frac{x_\alpha\rho_\beta(G_{\alpha\alpha}+G_{\beta\beta}-2G_{\alpha\beta})}{1+\rho_\beta x_\alpha(G_{\alpha\alpha}+G_{\beta\beta}-2G_{\alpha\beta})} \tag{8}$$

Expressions for the thermodynamic factor for ternary mixtures can be found in Refs. [9, 39, 67, 71]. For a specific solution, the thermodynamic factor provides an indication of the phase stability, since Γ relates to the second derivative of the Gibbs energy with respect to composition [73]. For a binary system, Γ is

positive for a thermodynamically stable mixture and negative for an unstable one. As will be shown later, the thermodynamic factor can be used to connect Fick diffusion coefficients to MS diffusivities [72, 74]. Moreover, thermodynamic factors can be used to predict the finite-size effects of MS diffusion coefficients computed using molecular simulation [52].

To compute KB integrals, one can consider local density fluctuations in finite and closed systems rather than computing fluctuations in the grand-canonical ensemble of infinite systems (R.H.S of Eq. (2)). In practice, molecular simulation can only access finite systems and simulating open systems critically relies on insertion and deletion of molecules, as in the grand-canonical ensemble [25]. In the case of medium to high density fluids, the probabilities of accepting insertions and deletions of molecules are very low even with the use of advanced insertion schemes e.g. Continuous Fractional Component Monte Carlo (CFCMC) [75–84]. To avoid insertion of molecules, Schnell et al. [10] developed the so-called small system method (SSM), where macroscopic properties are computed using small and open subvolumes embedded in a larger reservoir (see Fig. 1b). Since molecules can enter and leave the subvolume, it is possible to compute density fluctuations in the grand-canonical (μVT) ensemble. In such subvolumes, it is possible to realize configurations that are not allowed in a periodic repetition of the subvolume (which is much closer to the thermodynamic limit) [25]. For example, the two green-colored molecules in Fig. 1b would overlap in a periodic repetition of the subvolume, resulting in a zero statistical weight of such a configuration. When the subvolume is not periodically repeated but embedded in a large simulation box, such a configuration is perfectly allowed [85]. This leads to effective surface effects, and properties of the subvolume should be studied by considering the thermodynamics of small systems [86]. In section 5, a brief explanation of thermodynamics of small systems is provided. To use the SSM for computing KB integrals in the thermodynamic limit, an expression for KB integrals of finite subvolumes was derived in Ref. [35]. In the following section, this derivation and the method proposed by Krüger and co-workers [34, 35, 37, 38] to compute KB integrals using subvolumes embedded in a finite reservoir is

discussed.

3. KB Integrals of Finite Volumes

In 2013, Krüger et al. [35] derived an expression for KB integrals of finite and open subvolumes embedded in a reservoir. An example of such subvolumes is provided in Fig. 1b. In this work, the volume of the subvolume will be denoted by V and L will be used for the characteristic length of the subvolume. The volume and length of the reservoir will be denoted by V_{box} and L_{box}, respectively. We consider an isotropic fluid, where translational and rotational effects have been integrated out and focus on a finite and open subvolume V. Krüger et al. [35] defined the finite-size KB integrals $G^V_{\alpha\beta}$ in terms of fluctuations in the number of molecules, as well as double integrals of particle positions over the RDF,

$$G^V_{\alpha\beta} \equiv \frac{1}{V}\int_V\int_V [g_{\alpha\beta}(r_{12}) - 1]\,\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 = V\frac{\langle N_\alpha N_\beta\rangle - \langle N_\alpha\rangle\langle N_\beta\rangle}{\langle N_\alpha\rangle\langle N_\beta\rangle} - \frac{V\delta_{\alpha\beta}}{\langle N_\beta\rangle} \quad (9)$$

where $r = |\mathbf{r}_1 - \mathbf{r}_2|$ is the pair distance. KB integrals computed from small subvolumes V scale as the surface area A to volume V ratio, $G^V_{\alpha\beta} = G^\infty_{\alpha\beta} + F\frac{A}{V}$, where F becomes constant in the limit $V \to \infty$. [35]. This scaling law can be explained by the concept of thermodynamics of small systems (see section 5). Alternatively, Krüger et al. [35] showed that as RDFs have a finite range, splitting the integral domain in Eq. (9) over the surrounding, $\int_V\int_{V_{\text{box}}}$ and $\int_V\int_{V_{\text{box}}-V}$, also results in the scaling of $G^V_{\alpha\beta}$ as mentioned above. From extrapolating to $A/V \to 0$ (or $1/L \to 0$), KB integrals in the thermodynamic limit, $G^\infty_{\alpha\beta}$, are obtained.

In Eq. (9), the R.H.S and L.H.S are equivalent and $G^V_{\alpha\beta}$ can be computed either from fluctuations in the number of molecules inside the subvolume V, or by integrating RDFs. Most molecular simulation packages readily compute RDFs and computing local density fluctuations is a bit more cumbersome. To use RDFs for computing KB integrals, a simpler expression is required instead

of the double integrals of the L.H.S of Eq. (9). In the limit $V \to \infty$, the double integrals on the L.H.S. of Eq. (9) reduce to a single integral over the distance between molecules 1 and 2. By applying the transformation $\mathbf{r_2} \to \mathbf{r} = \mathbf{r_1} - \mathbf{r_2}$, the original KB integral (Eq. (1)) is retrieved. However, it is not possible to apply the same transformation for finite volumes, as $\mathbf{r}$ depends on the position $\mathbf{r_1}$. Simply ignoring this dependency results in truncated KB integrals, here referred to as G_0,

$$G_0 = \int_0^{L_{\max}} \mathrm{d}r \left[g_{\alpha\beta}(r) - 1\right] 4\pi r^2 \tag{10}$$

where $L_{\max}$ is the maximum possible distance between two points inside in the subvolume V. Note that for finite $L_{\max}$, G_0 in Eq. (10) is different from the LHS of Eq. (9) and so it does not yield the number fluctuations in the finite volume (RHS of Eq. (10)). This important fact seems to have been overlooked prior to the work of Kruger and et al. [35]. As shown in Refs. [35] and [34], the truncation of the infinite KB integrals (Eq. (1)) to finite distances results in poor convergence. Later in this section, a physical argument for the poor convergence of Eq. (10) will be provided. Still, it is desired to deal with a single integral as opposed to the six-dimensional integration in Eq. (9). For isotropic fluids, it is possible to re-write the double integrals in Eq. (9) over the interparticle distance as [35]

$$G_{\alpha\beta}^V = \int_0^{\infty} \left[g_{\alpha\beta}(r) - 1\right] w(r, L_{\max}) \mathrm{d}r \tag{11}$$

where $w(r, L_{\max})$ is a geometric weight function that is proportional to the probability that two points inside V are at distance r,

$$w\left(r, L_{\max}\right) = \frac{1}{V} \int_V \mathrm{d}\mathbf{r}_1 \int_V \mathrm{d}\mathbf{r}_2 \delta(r - |\mathbf{r}_1 - \mathbf{r}_2|) \tag{12}$$

The function $w(r, L_{\max})$ depends on the dimensionality and shape of the subvolume V. From the definition, it follows that $w(r, L_{\max}) = 0$ for $r >= L_{\max}$.

As a result, Eq. (11) can be written as an integral over a finite range,

$$G_{\alpha\beta}^{V} = \int_{0}^{L_{\max}} \left[g_{\alpha\beta}(r) - 1\right] w(r, L_{\max}) \mathrm{d}r \tag{13}$$

Note that the integrand of Eq. (13) depends on the integration boundary. An analytic expression for the function $w(r, L_{\max})$ was derived for hyperspheres in $1D$, $2D$, and $3D$. The derivation can be found in Ref. [37]. For a $3D$ sphere, the expression is:

$$w(r, L_{\max}) = 4\pi r^2 \left(1 - \frac{3x}{2} + \frac{x^3}{2}\right) \tag{14}$$

where x is the dimensionless distance $x = r/L_{\max}$, and $L_{\max}$ is the diameter of the sphere. For a cubic subvolume with side L, an analytic expression for $w(r, L_{\max})$ was recently derived by Krüger and Vlugt [34],

$$w(r, L_{\max}) \begin{cases} = 4\pi r^2 (1 - \frac{3\sqrt{3}x}{2} + \frac{6x^2}{\pi} - \frac{3\sqrt{3}x^3}{4\pi}), & x < \frac{1}{\sqrt{3}}, \\ = r^2(-8\pi + 6\sqrt{3}x + 6\sqrt{3}x^3 + \frac{6\pi - 1}{\sqrt{3}x} + \\ 24\sqrt{3}x \arccos(1/\sqrt{3}x) - 8(6x^2+1)\sqrt{1 - 1/3x^2}), & \frac{1}{\sqrt{3}} < x < \sqrt{2/3} \\ \approx 0, & \sqrt{2/3} < x < 1 \end{cases} \tag{15}$$

with $x = r/L_{\max}$ and $L_{\max} = \sqrt{3}L$. For any other arbitrary concave shape of the subvolume, Dawass et al. [38] proposed a method to compute the function $w(r, L_{\max})$ numerically. These authors used umbrella sampling MC simulations to find the probability distribution of finding two points separated by a distance r inside a subvolume V. The normalization of this probability distribution function directly leads to the function $w(r, L_{\max})$. This approach can be used for any concave shape of the subvolume in any dimension. Fig. 2 displays the function $w(r, L_{\max})$ for the following shapes of the subvolume: a sphere, cube, spheroid, and cuboid. The functions $w(r, L_{\max})$ were found numerically using the approach of Dawass et al. [38]. For a sphere and cube, analytic functions $w(r, L_{\max})$ are plotted as well (Eqs. (14) and (15)). As shown in Fig. 2, numerical and analytic results are in excellent agreement.

When computing KB integrals from density fluctuations or integrating RDFs, effects related to the sizes of the system and the subvolume have to be considered [37]. In Fig. 3a, KB integrals of a binary Weeks-Chandler-Andersen (WCA) [87] mixture are shown for different sizes for the simulation box at the same temperature and density. The RDFs of the system are computed from MD simulations for different sizes of the subvolume, and KB integrals $G^V_{\alpha\beta}$ are computed from integrating the RDFs (Eqs. (13) and (14)). In Fig. 3a, G^V_{12} is plotted as a function of $1/L$ (L is the diameter of a spherical subvolume). To obtain KB integrals in the thermodynamic limit, G^∞_{12}, the linear part of the scaling of G^V_{12} vs. $1/L$ is extrapolated to the thermodynamic limit ($1/L \to 0$). In Fig. 3a, effects related to the size of the system (V_{box}) are shown. Unlike large systems, small systems do not provide a linear regime that is sufficient to extrapolate to $1/L \to 0$. For all system sizes, not the whole range of distances should be considered for this extrapolation. As outlined in Ref. [37], it is not advisable to use subvolumes that extend beyond half the size of the simulation box and L should always be smaller than $L_{\text{box}}/2$. The use of a larger reservoir or simulation box V_{box} allows larger subvolumes, and increases the linear regime when $G^V_{\alpha\beta}$ is plotted as a function of $1/L$.

Besides the finite-size effects of $G^V_{\alpha\beta}$ due to finite V_{box}, RDFs from finite and closed systems have to be corrected for finite-size effects. In literature, several methods have been proposed [35–37, 64] to estimate RDFs in the thermodynamic limit. These methods will be reviewed in section 3.1. The corrections are needed as the KB theory requires RDFs of open and infinite systems. Fig. 3 presents a comparison between KB integrals computed using corrected RDFs and integrals from RDFs that are not corrected. In Fig. 3b, RDFs are corrected using the correction of Ganguly and van der Vegt [64] (more details will be provided in section 3.1). As shown in Fig. 3b, applying a RDF correction for a small system results in a better convergence of KB integrals.

In Fig. 3, KB integrals are computed using spherical subvolumes. However, it is possible to use subvolumes of any shape, provided that the function $w(r, L_{\max})$ is known. In Refs. [34, 38, 47], the effect of the shape of the subvolume on

computing KB integrals was examined. It was demonstrated that KB integrals in the thermodynamic limit are independent of the shape of the subvolume, and only depend on its size. In section 3.2, shape effects will be discussed. We will show that understanding shape effects leads to an expression [34] for computing $G^{\infty}_{\alpha\beta}$ directly from RDFs of finite subvolumes (section 3.3). The new expression can be used as an alternative to Eq. (13) where KB integrals of finite subvolumes are computed and then extrapolated to the thermodynamic limit.

3.1. RDF Corrections

RDFs of open systems, $g^{\infty}_{\alpha\beta}(r)$, are required for computing KB integrals. However, RDFs of finite and closed systems are typically obtained from molecular simulation. As a result $g^{\infty}_{\alpha\beta}(r)$ has to be estimated from RDFs of closed systems before applying the KB integration. The following corrections have been compared in Ref. [37].

3.1.1. Ganguly and van der Vegt correction

Ganguly and van der Vegt [64] address the asymptotic behaviour of RDFs computed from simulation of finite systems. For large systems, RDFs should converge to the value of 1 [9]. For finite and closed systems, RDFs do not approach this value, e.g. for a single-component ideal gas with N molecules inside a fixed volume V, we have $g(r) = (N-1)/N$ [88]. The authors proposed that RDFs can be corrected to the thermodynamic limit by using the correct bulk density when normalizing $g_{\alpha\beta}(r)$. For a spherical shell with a central molecule of type α (see Fig. 1a), the excess or depletion in the number of molecules of type β can be computed. Since the number of molecules of type β in the system N_β is fixed, the bulk density needs to be compensated by the excess or depletion in the number of molecules inside a volume V with radius r,

$$\Delta N_{\alpha\beta}(r) = \int_0^r \mathrm{d}r' 4\pi r'^2 \rho_\beta \left[g_{\alpha\beta}(r') - 1\right] \tag{16}$$

The excess or depletion in the number of molecules is then used to correct the

RDF computed from finite systems,

$$g_{\alpha\beta}^{\mathrm{vdV}}(r) = g_{\alpha\beta}(r)\frac{N_\beta\left(1-\frac{V}{V_{\mathrm{box}}}\right)}{N_\beta\left(1-\frac{V}{V_{\mathrm{box}}}\right)-\Delta N_{\alpha\beta}(r)-\delta_{\alpha\beta}} \tag{17}$$

where $g_{\alpha\beta}^{\mathrm{vdV}}(r)$ is the corrected RDF, and $g_{\alpha\beta}(r)$ is the RDF obtained from a molecular simulation of a finite and closed system. As shown in Eq. (17), the Ganguly and van der Vegt correction provides a relatively simple method that can be applied to RDFs computed from finite and closed systems. For an ideal gas, it also reflects the correct physical behaviour of the RDF. For a single component ideal gas inside a closed volume V, the excess or depletion is

$$\Delta N(r) = \int_0^r \mathrm{d}r' 4\pi r'^2 \rho\left[g(r')-1\right] = V\frac{N}{V_{\mathrm{box}}}\left[\frac{N-1}{N}-1\right] = \frac{-V}{V_{\mathrm{box}}} \tag{18}$$

Substituting the result in Eq.(17) results in:

$$g^{\mathrm{vdV}}(r) = \left(\frac{N-1}{N}\right)\frac{N\left(1-\frac{V}{V_{\mathrm{box}}}\right)}{N\left(1-\frac{V}{V_{\mathrm{box}}}\right)-\frac{-V}{V_{\mathrm{box}}}-1} = 1 \tag{19}$$

which is the correct answer for an ideal gas in the thermodynamic limit.

3.1.2. $1/N$ correction

In the book of Ben Naim [9], the difference between the behaviour of RDFs of closed systems $g_{\alpha\beta}^{N}(r)$, with total number of molecules N, and RDFs of open systems $g_{\alpha\beta}^{\infty}(r)$ is explained. Specifically, it is shown that $g_{\alpha\beta}^{\infty}(r)$ converges to 1 for $r\to\infty$, while $g_{\alpha\beta}^{N}(r)$ converges to $1-1/N$ [9]. Therefore, one can consider the difference between $g_{\alpha\beta}^{N}(r)$ and $g_{\alpha\beta}^{\infty}(r)$ as a Taylor series in $1/N$ [89],

$$g_{\alpha\beta}^{N}(r) = g_{\alpha\beta}^{\infty}(r) + \frac{c(r)}{N} + \mathcal{O}\left(\frac{1}{N^2}\right) \tag{20}$$

where N is the total number of molecules of the system, and $c(r)$ is a function of the distance r. The function $c(r)$ can be estimated using two systems with

different N but the same thermodynamic state. This results in the following estimate for the RDF in the thermodynamic limit [35],

$$g^{\infty}_{\alpha\beta}(r) = \frac{N_1 g^{N_1}_{\alpha\beta}(r) - N_2 g^{N_2}_{\alpha\beta}(r)}{N_1 - N_2} \tag{21}$$

where $g^{N_1}_{\alpha\beta}(r)$ and $g^{N_2}_{\alpha\beta}(r)$ are the radial distribution functions for a closed system with N_1 and N_2 molecules, respectively. Hence, two simulations with different number of molecules, N_1 and N_2 need to be performed to find $g^{\infty}_{\alpha\beta}(r)$, in contrast to the method of Ganguly and van der Vegt [64] where only a single simulation is sufficient. Another disadvantage of this method is that it results in numerical inaccuracies in the computed RDFs and hence the values of KB integrals [37].

3.1.3. Cortes-Huerto et al. correction

In the work of Cortes-Huerto et al. [36], an expression is derived to compute KB integrals in the thermodynamic limit from KB integrals of finite subvolumes. In section 4, we provide more details on the approach of Cortes-Huerto et al. [36] for computing KB integrals from molecular simulation. To correct for RDF-related effects, the following relation between $g^{\infty}_{\alpha\beta}(r)$ and $g_{\alpha\beta}(r)$ is used from the book of Ben-Naim [9],

$$g_{\alpha\beta}(r) = g^{\infty}_{\alpha\beta}(r) - \frac{1}{V_{\text{box}}}\left(\frac{\delta_{\alpha\beta}}{\rho_\alpha} + G^{\infty}_{\alpha\beta}\right) \tag{22}$$

In Eq. (22), $g_{\alpha\beta}(r)$ is corrected by a constant value, independent of r. However, the difference between $g_{\alpha\beta}(r)$ and $g^{\infty}_{\alpha\beta}(r)$ depends on r, as shown in Ref. [37]. In principle, Eq. (22) only applies in the limit $r \to \infty$ and not for finite r. This approximation implied in Eq. (22) may affect the accuracy of the computed KB integrals, as will be discussed in the following section.

3.1.4. Comparison between correction methods

In Ref. [37], a quantitative comparison between the RDF correction methods considered above was carried out. The corrections were applied to RDFs of a binary WCA [87] fluid computed from MD simulations of closed and finite boxes of different sizes. KB integrals were computed using RDFs corrected using

the Ganguly and van der Vegt [64] correction (Eq. (17)), the $1/N$ correlation (Eq. (21)), and the method by Cortes-Huerto et al. [36] (Eq.(22)). The effects of the used RDF correction method on the accuracy of the computed KB integrals were investigated. Based on these comparisons, it was shown that the Ganguly and van der Vegt correction [64] provides the most accurate KB integrals. It is also the most simple correction in practice. The $1/N$ correction (Eq. (21)) requires simulating two systems and resulted in numerical inaccuracies, especially when the difference between N_1 and N_2 is not chosen carefully. The correction by Cortes-Huerto et al. [36] was found to reduce inaccuracies due to finite-size effects of the RDFs, however, not as much as the correction of Ganguly and van der Vegt. For instance, for the WCA fluid studied in Ref. [37], the correction of Cortes-Huerto et al. reduced the difference between KB integrals computed from very large systems and KB integrals from small systems to less than 5%, while the correction of Ganguly and van der Vegt reduced these differences to less than 1% [37].

3.2. Shape Effects of the Subvolume

The effect of the shape of the subvolume V on computing KB integrals was studied in Refs. [47], [38], and [34]. Strøm et al. [47] found that when approaching the thermodynamic limit, the KB integrals become independent of the shape of the subvolume. These authors used arguments from nanothermodynamics to illustrate that for large subvolumes, KB integrals should be a function of the size of the subvolume and the surface area to volume ratio, A/V. Dawass et al. [38] studied KB integrals for different shapes of the subvolumes and presented numerical results that agree with the findings of Ref. [47]. Here, we will summarize the main findings of investigating shape effects. This includes a universal first order expansion of the function $w(r, L_{\max})$, valid for all shapes of the subvolume.

To illustrate the shape effects related to computing KB integrals, we combine the use of spherical (Eq. (14)) and cubic (Eq. (15)) subvolumes with an analytic

RDF. We use the following analytic RDF [90],

$$g(r) - 1 = \begin{cases} \frac{3/2}{r/\sigma} \exp\left[\frac{1-r/\sigma}{\chi}\right] \cos\left[2\pi\left(\frac{r}{\sigma} - \frac{21}{20}\right)\right] & \frac{r}{\sigma} \geq \frac{19}{20}, \\ -1, & \frac{r}{\sigma} < \frac{19}{20} \end{cases} \tag{23}$$

where σ is the diameter of the molecule, and χ is the length scale at which the fluctuations of the RDF decay. We apply the function in Eq. (23) for a pure fluid, hence we drop the indices α and β for the remaining of this section. The use of an analytic RDF eliminates its finite-size effects, and hence we can focus on effects due to the size and shape of the subvolume.

In Fig 4, we show KB integrals computed for the liquid modeled by the RDF of Eq. (23), with $\sigma = 1$ and $\chi = 5$. KB integrals of finite subvolumes, G^V, are plotted as a function of the inverse size of the subvolume, $1/L$ (for a cube, L is the length of one side, and for a sphere, $L = 2R$). Fig. 4a shows that in the thermodynamic limit both shapes of the subvolume lead to the same estimate for G^∞. However, the shape of the subvolume affects the slope of the lines of G^V versus $1/L$. Fig. 4b shows that as $G^V_{\alpha\beta}$ approaches the thermodynamic limit ($L \to \infty$) shape effects can be corrected when plotting the integrals as a function of the area to volume ratio A/V. This is due to the fact that in the limit $L \to \infty$, KB integrals are a function of the ratio A/V, and not the shape of the subvolume. This was shown theoretically by the study of Strøm et al. [47]. The behavior of KB integrals in the thermodynamic limit can be explained using the function $w(r, L_{\max})$ at small distances. In the thermodynamic limit, the values of the function $w(r, L_{\max})$ at small distances have the largest contribution to KB integrals G^V (Eq. (13)). Dawass et al. [38] found that for all shapes studied, numerically computed values of the function $w(r, L_{\max})$ have the value of $4\pi r^2$ at the limit $r \to 0$. The function $w(r, L_{\max})$ can be expanded around $r = 0$ to find the following universal expression for any concave and continuous volume [34, 38],

$$w(r, L_{\max}) \approx 4\pi r^2 \left(1 - \frac{r}{4}\frac{A}{V} + \mathcal{O}(r^2)\right) \tag{24}$$

where A is the surface area of the subvolume. Eq. (24) shows that the function $w(r, L_{\max})$ depends on the size r and the ratio A/V. The shape contribution originates from the term $\mathcal{O}(r^2)$. Therefore, properties of large subvolumes are independent of shape. This is referred to as the so-called shape thermodynamics limit, where properties of the subvolume are dependent on the size but not the shape of the subvolume [85]. In the conventional thermodynamic limit, properties are independent of both size and shape of the subvolume. It is illustrative to see how the universal expression of the function $w(r, L_{\max})$ of Eq. (24) compares to analytic functions of $w(r, L_{\max})$ for a sphere and a cube, Eq. (14) and Eq. (15), respectively. To compare Eq. (24) with the leading terms of Eqs. (14) and (15), one can express all equations in terms of the distance r and the linear length of the subvolume L (for a sphere, L is the diameter and for a cube, L is the length of one size). For a sphere, using $x = r/L_{\max} = r/L$ in Eq. (14) and $A/V = 6/L$ in Eq. (24), one will arrive at the same weight function: $w(r, L_{\max}) = 4\pi r^2(1 - \frac{3}{2}\frac{r}{L})$. The same result is obtained for a cube, if one uses $x = r/L_{\max} = r/\sqrt{3}L$ and $A/V = 6/L$ in Eq. (15) and Eq. (24), respectively. Moreover, it would be interesting to see if Eq. (24) would provide a physical reasoning for the poor behavior of truncated KB integrals (Eq. (10)). If we consider a subvolume V with zero surface area $A = 0$, this will yield the weight function $w(r, L_{\max}) = 4\pi r^2$. Substituting the function $w(r, L_{\max})$ in the expression of KB integrals of KB integrals of finite subvolumes (Eq. (13)), one arrives at the expression of KB integrals truncated to finite distances of Eq. (10). Therefore, truncated KB integrals correspond to the nonphysical case of finite-size KB integrals (Eq. (9)) with subvolumes V and zero surface area [38].

3.3. Direct Extrapolation of $G^V_{\alpha\beta}$ to the thermodynamic limit

To estimate KB integrals for an infinite system, the scaling of G^V (Eq. (13)) with $1/L$ is extrapolated to the limit $L \to \infty$. However, using linear extrapolation to estimate the values of $G^\infty_{\alpha\beta}$ can result in numerical errors and, for some systems, a linear regime may be difficult to identify. Hence, it would be advantageous to have an expression that provides $G^\infty_{\alpha\beta}$ directly using RDFs computed

from simulations of finite and closed systems. Recently, Krüger and Vlugt [34] proposed the following general extrapolation formula

$$G^{\infty} \approx G_k(L_{\max}) = \int_0^{L_{\max}} \left[g_{\alpha\beta}(r) - 1\right] u_k(r) \mathrm{d}r \tag{25}$$

where $u_k(r)$ is a weight function that takes into account the finite-size effects of G^V. The subscript k, in G and $u(r)$, refers to the different forms of $u(r)$ and different levels of approximation. It will be shown that the function $u_k(r)$ may depend on the integration boundary $L_{\max}$. In Ref. [34], three functions are considered. The easiest estimation of G^{∞} is obtained from

$$u_0(r) = 4\pi r^2 \tag{26}$$

Substituting u_0 in Eq. (25) results in the truncated KB integrals (Eq. (10)), which is known to yield poor estimates of KB integrals in the thermodynamic limit. A better way of estimating KB integrals in the thermodynamic limit, is to take the derivative of Eq. (13) with respect to the integration boundary, and extrapolate it to the limit $1/L \to 0$ [35]. The final result is,

$$u_1(r) = 4\pi r^2 \left(1 - x^3\right) \tag{27}$$

where $x = r/L_{\max}$. The function $u_1(r)$ is zero when $x = 1$. The third estimate of $u_k(r)$ is based on the scaling of G^V with $1/L$ as well as the universal expression of the function $w(r, L_{\max})$ (Eq. (24)) [34]. As discussed earlier, KB integrals of finite volumes scale with A/V or $1/L$,

$$G^V \approx G^{\infty} + \frac{1}{L} F^{\infty} \tag{28}$$

The expansion above ignores terms of order $(1/L)^2$ or higher. Krüger and Vlugt [34] have demonstrated that one can obtain the surface term, F^{∞}, with the knowledge of the function $w(r, L_{\max})$ for all shapes (Eq. (24)). Using Eq. (1) and Eq. (13) for G^{∞} and G^V, respectively, in combination with Eq. (24) yields

an explicit expression for the surface term,

$$F^{\infty} = \int_0^{\infty} [g(r) - 1] \left(-\frac{3}{2} r \right) 4\pi r^2 \mathrm{d}r \tag{29}$$

The integral in the surface term is similar to the expression of KB integrals of infinite volumes (Eq. (1)), but now with $[g(r)-1]$ replaced by $(-3/2r)[g(r)-1]$. Hence, the methodology to estimate KB integrals of finite subvolumes can be used to find a suitable estimate of the surface term [34],

$$F^{\infty} \approx \int_0^{L_{\mathrm{box}}} [g(r) - 1] \left(-\frac{3}{2} r \right) \left(1 + \frac{3}{2} x \right) w(r, L_{\mathrm{max}}) \mathrm{d}r \tag{30}$$

From Eqs. (30), (28), and the function $w(r, L_{\mathrm{max}})$ of a sphere (Eq. (14)), a new weight function can be derived and used to find KB integrals in the thermodynamic limit from knowledge of RDFs of finite systems [34],

$$u_2(r) = 4\pi r^2 \left(1 - \frac{23}{8} x^3 + \frac{3}{4} x^4 + \frac{9}{8} x^5 \right) \tag{31}$$

Note that the function $u_2(r)$ and its derivative vanishes at $x = 1$. This is shown in Fig. 5a, where the weight functions $u_0(r)$, $u_1(r)$ and $u_2(r)$ are plotted. Krüger and Vlugt [34] derived a finite-range integral (Eq. (25)) to estimate KB integrals of infinite systems. The quality of the estimation depends on the weight function, $u_k(r)$. While this mathematical solution was derived for the problem of KB integrals, it is valid for the estimation of similar integrals of infinite distances. The proposed estimate discussed above, $u_2(r)$, was derived for any shape of the subvolume in $3D$. In Ref. [91], the approach of Krüger and Vlugt [34] was extended to higher dimensions.

In Fig. 6, a guide on how to compute KB integrals from molecular simulations using the method of Krüger and co-workers [34, 35, 37, 38] is provided. It is shown in this section that it is possible to compute G^{∞} in four different ways:

1. From truncated KB integrals, G_0 (Eqs. (25) and (26)).
2. From numerically extrapolating KB integrals of finite subvolumes G^V (Eq. (13)) to the thermodynamic limit.

3. Using the analytical extrapolation $G_1(L_{\max})$ (Eqs. (25) and (27)), derived form the slope of G^V vs $1/L$.

4. Using the analytical extrapolation $G_2(L_{\max})$ (Eqs. (25) and (31)), based on the scaling of G^V and Eq. (24).

In Fig. 5b, a comparison between the four methods to compute KB integrals is shown. We have created a python package to calculate those integrals [92]. RDFs used to compute KB integrals in Fig. 5b were computed from MD simulation of a binary WCA fluid [87] and we corrected the RDFs to the thermodynamic limit with the correction of Ganguly and van der Vegt (Eq. (17)). Fig. 5b shows that all estimates of G^∞ lead to the same value in the limit $1/L \to 0$. However, the truncated integrals, $G_0(L_{\max})$, results in very large oscillations when compared to the other KB integrals. This means that it is almost impossible to obtain G^∞ using Eq. (10) using simulations of small systems. The large oscillations of the integrals $G_0(L_{\max})$ can be attributed to oscillations of $g(r)$, which are amplified when performing the integration in Eq. (10). The integrals $G_1(L_{\max})$ and $G_2(L_{\max})$ result in smoother oscillations, while the function G^V results the smoothest lines as it scales with $1/L$. The integrals $G_1(L_{\max})$ and $G_2(L_{\max})$ scale with $1/L^3$ as indicated by the leading terms of the weight functions $u_1(r)$ and $u_2(2)$, Eqs. (27) and (31) respectively. As expected from Ref. [34], $G_2(L_{\max})$ provides a very smooth convergence when compared to $G_0(L_{\max})$ and $G_1(L_{\max})$. This can be explained by the weight functions $u_k(r)$ and their behaviour in the limit $x = r/L_{\max} = 1$. Fig. 5a shows that, for a subvolume with $L_{\max} = 1\sigma$, the functions u_1 and u_2 have the value of zero at $r = 1\sigma$ (i.e. $x = 1$), unlike the function u_0. As a result, we can conclude that the integrals $G_2(L_{\max})$ and the numerical extrapolation of G^V provide the most accurate estimate of KB integrals in the thermodynamic limit.

4. Other methods for estimating KB integrals in the thermodynamic limit

In the previous section, the method by Krüger and co-workers [34, 35, 37, 38] to formulate expressions of KB integrals of finite subvolumes was discussed. We showed that this approach provide integrals that converge smoothly to KB integrals of an infinite system. Besides the approach presented in section 3, a number of methods are available in literature to compute KB integrals from molecular simulations of finite systems. Here, some of these methods are discussed.

The approach of Cortes-Huerto et al. [36] is the very similar to the method of Krüger and co-workers [34, 35, 37, 38] (section 3). An expression was derived relating KB integrals of finite subvolumes $G^V_{\alpha\beta}$ to KB integrals of infinite systems, $G^\infty_{\alpha\beta}$. Cortes-Huerto et al. [36] computed $G^V_{\alpha\beta}$ from fluctuations in the number of molecules (R.H.S of Eq. (9)). In principle, KB integrals of finite subvolumes $G^V_{\alpha\beta}$ can be computed from integrals of RDFs as well. Cortes-Huerto et al. [36] apply corrections to compensate for two finite-size effects: (1) RDFs-related effects and, (2) boundary or surface effects. For the first effect, the correlation of Eq. (22) is used to estimate RDFs in the thermodynamic limit. The RDF correction used in the work of Cortes-Huerto et al. [36] (Eq. (22)) is discussed in section 3.1. For surface effects, resulting from computing fluctuations inside small subvolumes, Cortes-Huerto et al. [36] adopt the same scaling approach of Krüger et al. [35] where $G^V_{\alpha\beta}$ scales with A/V, which equals $1/V^{1/3}$ in $3D$. Including the two corrections in the definition of $G^V_{\alpha\beta}$ of Eq. (9) leads to a final working expression where $G^V_{\alpha\beta}$ is written as a function of $(V/V_{\text{box}})^{1/3}$,

$$G^V_{\alpha\beta} = G^\infty_{\alpha\beta}\left(1 - \frac{V}{V_{\text{box}}}\right) - \frac{V}{V_{\text{box}}}\frac{\delta_{\alpha\beta}}{\rho_\alpha} + \frac{C_{\alpha\beta}}{V^{1/3}} \tag{32}$$

$C_{\alpha\beta}$ is a constant originating from the scaling of $G^V_{\alpha\beta}$ with A/V, and it is specific to each thermodynamic state. From the slope of the line of $G^V_{\alpha\beta}$ vs. $(V/V_{\text{box}})^{1/3}$, KB integrals of infinite systems $G^\infty_{\alpha\beta}$ are found.

The methods of Krüger and co-workers [34, 35, 37, 38], and Cortes-Huerto et

al. [36] provide practical approaches to computing KB integrals for any isotropic fluid, while addressing system size effects and RDF-related effects. Other available methods for computing KB integrals are more complicated, and found to be difficult to extend to systems with internal degrees of freedom. Wedberg et al. [59, 93] presented a method for extending KB integrals to the thermodynamic limit using Verlet's extension method [90]. The Verlet extension method [90] can be applied to estimate RDFs beyond the size of the finite simulation box, which are then used to extrapolate to KB integrals to the thermodynamic limit, by truncating Eq. (1) to a value much larger than half the box size. The approach of Wedberg et al. [59] was verified using pure LJ and Stockmayer fluids. A drawback of this approach is the complexity of the numerical procedure. Moreover, it is not trivial to extend the method to systems of molecules with intramolecular degrees of freedom.

KB integrals can be computed from molecular simulations of finite number of molecules using static structure factors [58, 60]. The structure factor of a liquid, $S(q)$, is related to the Fourier transform of pair distribution functions, and q is the magnitude of change of a reciprocal lattice vector. Structure factors can be measured from scattering experiments, where q is a function of the wave length and the scattering angle. At the zero wavelength limit, $q = 0$, structure factors are directly related to KB integrals [94, 95]. However, the values of $S(q = 0)$ cannot be measured directly. Similarly, with molecular simulation, structure factors can be computed for a set of values of q, and then extrapolated to the limit $q = 0$ to find KB integrals. In the work of Nichols et al. [58], structure factors are computed from fluctuations in the number of molecules of finite systems. Each lattice vector q, corresponds to a set of different sampling volumes, or sub-cells inside the simulation box, from which the density fluctuations are computed. Rather than considering subvolumes formed by a central molecule, Nichols et al. [58] considered fluctuations in slab-like regions that resulted from dividing the simulation box. As a result, the whole volume is considered and all the information is used. From fluctuations, written as a $3D$ Fourier series, structure factors are computed and this was used to obtain

the thermodynamic properties that relate to KB integrals (partial deviates of chemical potential with respect to composition, molar volumes, and isothermal compressibility). For a LJ fluid, Nichols et al. [58] found it difficult to extrapolate structure factors to $q = 0$. Instead, thermodynamic properties computed from subcells (i.e specific range of q) were extrapolated to the limit $q = 0$. Extrapolation of thermodynamic properties is needed to remedy finite-size effects. While the method of Nichols et al. [58] provide accurate thermodynamic properties, compared to truncated KB integrals, it is computationally involved even for systems with no intramolecular interactions. Structure factors are also used in the work of Rogers [60] to compute KB integrals from simulations of closed and finite systems. As in the work of Nichols et al. [58], information from the entire volume of the simulation box was used. However, both methods were applied to compute KB integrals of systems of molecules with no intramolecular degrees of freedom such as LJ fluids.

5. Thermodynamic properties of small systems

In section 3, we discussed the method of Krüger and co-workers [34, 35, 37, 38] for computing KB integrals from molecular simulation. This approach applies a similar concept to that of the SSM discussed earlier (section 2). In both approaches, the desired property is computed in the thermodynamic limit from finite and small subvolumes embedded in a larger reservoir (i.e. simulation box). Since open and small subvolumes can be of the order of a few molecular diameters, thermodynamics of small systems applies. In this section, the basics of thermodynamics of small systems are presented. Classical, or bulk, thermodynamics will be referred to as standard thermodynamics.

5.1. A small system with constant μ, V and T

The starting point is a small and open subvolume embedded in a reservoir (for example: see Fig. 1b) [85, 96]. As discussed earlier, this set up can be treated as grand-canonical. For the sake of simplicity, we will only present the

case of pure component systems, as the extension to multi-component system is presented in the book of Hill [86]. Under these conditions, as we have seen from Eq. (9), KB integrals are directly related to the density fluctuations. Another important quantity that directly relates to these fluctuations and then to KB integrals is the thermodynamic factor, Γ (see Eq. (8) for a binary system). Following standard thermodynamics, for a single-component system Γ provides the evolution of the chemical potential with a density change. Using standard statistical thermodynamics one can easily show that in the thermodynamic limit, the density fluctuations in the μVT ensemble are proportional to Γ of a pure component,

$$\Gamma^{-1} = \frac{1}{\beta}\left(\frac{\partial \ln\langle N\rangle}{\partial \mu}\right)_{T,V} = \left(\frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle}\right)_{\mu,V,T} \tag{33}$$

where $\langle N\rangle$ is the average number of particles in the subvoume and $\beta = 1/(k_B T)$ with k_B being Boltzmann constant. For finite volumes, Γ^{-1} scales linearly with A/V beyond a few molecular diameters, as KB integrals [35, 37, 47]:

$$\Gamma^{-1}\left(\frac{A}{V}\right) = \Gamma_\infty^{-1} + C\frac{A}{V} \tag{34}$$

where C is a constant and Γ_∞^{-1} is the value of Γ^{-1} in the thermodynamic limit ($A/V \to 0$). The density fluctuations of Eq. (33) can then be considered as a sum of two contributions: a volume term ($\langle N\rangle\Gamma_\infty^{-1}$) and a surface term ($AC\rho$ with $\rho = \langle N\rangle/V$). Keeping in mind that the density, on average, is the same everywhere, $AC\rho$ can be understood as an excess fluctuation term that becomes negligible compared to $\langle N\rangle\Gamma_\infty^{-1}$ as the size of the subvolume increases.

As described in Refs. [85, 96], this implies that the grand-canonical partition function of this system also has an extra surface term compared to its standard expression in the thermodynamic limit. In that respect, thermodynamic properties like Γ^{-1} are no longer intensive, like in the thermodynamic limit, according to Gibbs' thermodynamics of heterogeneous systems. This result, which also applies to KB integrals as described above, clearly illustrates

590 that standard thermodynamics does not apply to small systems. Here, this result was obtained from simulations and statistical mechanics. Hill [86] arrived at a similar result from a thermodynamic derivation.

In the 1960's, Hill proposed a systematic extension of standard thermodynamics to small systems, also called nanothermodynamics [57, 97]. Considering an open and small system, Hill showed that the pressure is no longer intensive. To differentiate from classical pressure, the pressure of a small system will be referred to as $\hat{p}$. For a grand-canonical ensemble, $\hat{p}$ is a function of the grand-canonical partition function of the small system, Ξ:

$$\hat{p}V = k_{\mathrm{B}}T \ln \Xi \tag{35}$$

This expression is very similar to that provided by classical statistical mechanics except for the $\hat{p}$ instead of p. Like for Γ^{-1}, it follows directly that the change in
595 the grand partition function due to the surface contribution originates from the difference between p and $\hat{p}$. In the next section, we will see how Hill introduced $\hat{p}$ in thermodynamic relations and how it is used to extend the thermodynamics to small and open systems.

5.2. Basic relations for a small system with constant μ, V and T

Hill [57] considered $\mathcal{N}$ replicas of the small system, constructing thereby an ensemble (the total system), which is large enough to follow the laws of standard thermodynamics. The Gibbs' equation for this new ensemble is then:

$$E_t = TS_t - p\mathcal{N}V + \mu N_t + \left(\frac{\partial E_t}{\partial \mathcal{N}}\right)_{S_t,V,N_t} \mathcal{N} \tag{36}$$

subscript t refers to the total system, and the symbol E is used for internal energy, S for entropy. For convenience, we define

$$X \equiv \left(\frac{\partial E_t}{\partial \mathcal{N}}\right)_{S_t,V,N_t} \tag{37}$$

which can be interpreted as the reversible work needed to add one replica of the small system at constant S_t, V and N_t. The addition of a replica at constant

S_t, V and N_t, implies that S_t and N_t have to be redistributed over one more replica. By integrating Eq. (36) at constant T, V, μ and X, we obtain:

$$E_t = TS_t + \mu N_t + X\mathcal{N} \tag{38}$$

where T and μ are determined from the values of these quantities in the reservoir. The average variables of the small system are related to the variables of the total system by:

$$E_t \equiv \mathcal{N}\langle E\rangle, \quad S_t = \mathcal{N}S, \quad N_t \equiv \mathcal{N}\langle N\rangle \tag{39}$$

where the brackets $\langle\cdots\rangle$ are used to denote averages of a single replica. The entropy S is determined by the probability distribution over N and E, which is the same for each replica [86]. By introducing the variables for the small system into Eq. (38), we obtain:

$$\langle E\rangle = TS + \mu\langle N\rangle - \hat{p}V \tag{40}$$

in which we used the definition $X = -\hat{p}V$ [57]. The small system can be described by standard thermodynamics if $\hat{p} = p$. The system can be considered small when $\hat{p} \neq p$, as $\hat{p}$ may deviate significantly from p because of the effective surface energy of the system. The corresponding Gibbs-Duhem equation in this particular case is

$$S\mathrm{d}T + \langle N\rangle\mathrm{d}\mu - V\mathrm{d}\hat{p} = (\hat{p} - p)\,\mathrm{d}V \tag{41}$$

or

$$S\mathrm{d}T + \langle N\rangle\mathrm{d}\mu + p\mathrm{d}V = \mathrm{d}(\hat{p}V) \tag{42}$$

The relation between $\hat{p}V$ and the partition function proposed by Hill, Eq. (35), allows us to derive thermodynamic properties. We can then derive p, $\langle N\rangle$ and the thermodynamic correction factors, Γ, from

$$p = \left(\frac{\partial\hat{p}V}{\partial V}\right)_{T,\mu} \tag{43}$$

$$\langle N\rangle = \left(\frac{\partial\hat{p}V}{\partial\mu}\right)_{T,V} \tag{44}$$

$$\frac{1}{\Gamma} = \frac{1}{\beta}\left(\frac{\partial\ln\langle N\rangle}{\partial\mu}\right)_{T,V} \tag{45}$$

The first two quantities are special for small systems as they require partial derivatives of $\hat{p}$. It should be pointed out that the density $\langle N \rangle / V$ obviously does not exhibit any small volume effects as was shown from molecular simulations. Only the second derivative of $\hat{p}$ with the chemical potential shows a size dependence and also the fluctuation around $\langle N \rangle$. By this thermodynamic derivation, we have shown that results of Γ^{-1} obtained from finite subvolumes could be interpreted in the context of the thermodynamics of small systems (Γ^{-1} is no longer an intensive property). Bedeaux, Kjelstrup and co-workers [47, 98] have shown that this size effect can also be explained using Gibbs' thermodynamics of surfaces. This illustrates the equivalence between Hill's thermodynamic, Gibbs' surface thermodynamics and the Kirkwood-Buff approach.

6. Applications of KB Integrals from Molecular simulation

6.1. Partial molar enthalpies

In Ref. [6], Schnell et al. proposed a method to compute partial molar enthalpies from molecular simulation in the canonical ensemble. Following the SSM, enthalpies of small subvolumes $\hat{H}$ embedded in a larger reservoir are used. From nanothermodynamics, an expression for the change of $\hat{H}$ with respect to the average number of molecules $\langle N_\alpha \rangle$ was derived in terms of fluctuations in density and energy,

$$\left(\frac{\partial \hat{H}}{\partial \langle N_\alpha \rangle} \right)_{T,V,\mu_{\beta \neq \alpha}} = \frac{\langle E N_\alpha \rangle - \langle N_\alpha \rangle \langle E \rangle + \langle N_\alpha \rangle k_B T}{\langle N_\alpha^2 \rangle - \langle N_\alpha \rangle^2} \tag{46}$$

in which E is the energy of the subvolume. As shown in the previous section, properties of small subvolumes scale with the inverse size of the subvolume $(1/L)$. Extrapolating the derivatives of Eq. (46) to the thermodynamic limit yields partial enthalpies at constant volume $\left(\frac{\partial H}{\partial N_\alpha} \right)_{T,V,N_{\beta \neq \alpha}}$. To find partial molar enthalpies in the grand-canonical ensemble, $\left(\frac{\partial H}{\partial N_\alpha} \right)_{P,V,N_{\beta \neq \alpha}}$, a Legendre transform was performed. To convert from enthaplies in the canonical ensemble to partial molar enthalpies in the grand-canonical ensemble, KB integrals of the

studied system are needed. The method of Krüger and co-workers [34, 35, 37, 38] of KB integrals for finite subvolumes was used. Additionally, this approach was applied by Skorpa et al. [99] to compute the heat of reaction of a dissociation H_2 using a reactive force field.

6.2. Properties of Single-ions in Salt Solutions

Simulating closed and finite systems to compute KB integrals has the advantage of accessing single-ion properties [41]. Essentially, to apply the KB theory to a salt solution, the system has to be treated as a binary mixture where ions are indistinguishable [9], as shown in Ref [16]. In this case, relations between KB integrals and thermodynamic properties of binary mixtures, presented in section 2, can be applied. For a ternary mixture of a dissociating monovalent substance ($AB \rightarrow A + B$) and a solvent (e.g. water, W), KB integrals are subject to the following electroneutrality conditions,

$$\rho\, G_{WA} = \rho\, G_{WB} \tag{47}$$

$$1 + \rho\, G_{AA} = \rho\, G_{AB} \tag{48}$$

$$1 + \rho\, G_{BB} = \rho\, G_{AB} \tag{49}$$

where ρ is the number density of the salt ($\rho_A = \rho_B = \rho$). Eqs. (47), (48) and (49) imply that the number of molecules of species A and B cannot be varied independently. Ben-Naim [9] showed that the above constrains introduce a singularity to the equations relating KB integrals, $G^{\infty}_{\alpha\beta}$, to thermodynamic quantities. It is important to note that the KB theory is general for any type of interactions and the issue of singularity is not due to the strong electrostatic interactions present in salt solutions. Rather, it is a result of the closure constraints imposed by Eqs. (47), (48) and (49), and it does not apply to KB integrals defined in open systems [8]. Eqs. (47), (48) and (49) hold for any dissociating molecule AB where the number of molecules has to be conserved simultaneously in the system, i.e. $N_A = N_B$. The approach of using KB integrals of finite subvolumes

of Krüger and co-workers [34, 35, 37, 38] (section 3) allows KB integrals of single ions to be computed from simulations in the canonical ensemble with open subvolumes embedded in the simulation box. As a result, the charge neutrality of the reservoir is maintained ($N_A = N_B$), while the electroneutrality condition is not applied inside the subvolume, and therefore the grand-canonical ensemble is accessed. In the work of Schnell et al. [41], KB integrals of a sodium chloride (NaCl) solution were computed to find molar volumes of water, Na^+, and Cl^-. The partial molar volumes of one of the salt ions can have a negative value [41]. In Ref. [100], a similar observation was reported when computing molar volumes of Na^+ and Cl^-. The authors of Ref. [100] investigated the possibility of computing single-ion properties using molecular simulations by considering two methods. The first method is based on the changes in average potential energy and box volume when inserting an ion into a pure liquid, while the second method depends on evaluating the reversible work associated with inserting an ion into the a liquid.

6.3. Mass transfer in multicomponent liquids

KB integrals computed from molecular simulation can be applied to connect Fick diffusion coefficients to Maxwell- Stefan (MS) diffusivities. The generalized Fick's law relates the molar flux, J_i, to the Fick diffusivity, D_{ij} [72, 101],

$$J_i = -c_t \sum_{j=1}^{n-1} D_{ij} \nabla x_j \tag{50}$$

where c_t is the total molar concentration and ∇x_j is the mole fraction gradient, which is the driving force in Fick's law. MS diffusivities can be predicted from MD simulations and Fick diffusivities can be measured by experiments [40, 72, 101]. The MS diffusivity can be considered as an inverse friction term in an equation where the gradient in chemical potential is related to differences in the

average velocities between species:

$$-\frac{1}{RT}\nabla\mu_{ij} = \sum_{j=1(j\neq i)}^{n} \frac{x_j(u_i - u_j)}{Đ_{ij}} \tag{51}$$

where $(u_i - u_j)$ is the difference between the average velocities of the components. As chemical potentials cannot be measured directly, it is not possible to directly compare MS diffusivities to experiments. It is more convenient to compute MS diffusivites using molecular simulation. Fick diffusivities often depend more strongly on concentration than MS diffusivities [40, 101]. Moreover, it is possible to predict diffusion of multicomponent mixtures ($n > 2$) from the knowledge of MS diffusivites of a binary mixture [39, 40, 66]. For a mixture with more than two components, Fick diffusivities depend on the type of reference frame, unlike MS diffusivities [40, 101]. Fick diffusivities and the thermodynamic factor can be used to compare MS diffusivites with experimental data [72],

$$[D] = [B]^{-1}[\Gamma] \tag{52}$$

where $[D]$ is the Fick coefficients matrix. The elements of the matrix $[B]$ can be found using,

$$B_{ii} = \frac{x_i}{Đ_{ij}} + \sum_{j=1(j\neq i)}^{n} \frac{x_j}{Đ_{ij}} \qquad \text{with} \quad i = 1, 2, ...(n-1) \tag{53}$$

$$B_{ij} = -x_i\left(\frac{1}{Đ_{ij}} - \frac{1}{Đ_{ij}}\right) \qquad \text{with} \quad i = 1, 2, ...(n-1) \quad \text{and} \quad i \neq j \tag{54}$$

while the elements of the matrix $[\Gamma]$ can be expressed as a function of KB integrals. The relation between Γ_{ij} and KB integrals is provided by Eq. (8) for binary systems. The thermodynamic factors are provided for ternary mixtures can be found in Refs. [9, 39]. In Refs. [39] and [40], KB integrals from simulations of finite systems were computed using the approach of Krüger and co-workers [34, 35, 37, 38] (section 3). KB integrals of binary and ternary mixtures were used to compute thermodynamic factors and convert MS diffusivities

obtained from simulations to Fick diffusivities measured by experiments. The proposed method was applied to binary and ternary alcohol mixtures [39, 40].

In Ref. [52], it was shown that KB integrals can be used to correct finite-size effects of computed MS diffusion coefficients. MS diffusion coefficients are dependent on the size of the simulated system, and these finite-size effects were found to originate from hydrodynamic interactions [52, 102]. In the study of Jamali et al. [52], a correction based on viscosity, and the thermodynamic factor was used to compensate for this effect. For binary and LJ mixtures, KB integrals were obtained from molecular simulation and used to compute thermodynamic factors. The finite-size correction was applied to molecular systems such as organic fluids. Jamali et al. [52] found the finite-size effects of MS diffusivites to be significant, especially when thermodynamic factors approach zero (i.e when mixtures are close to demixing).

6.4. Other Applications

In section 1, we presented the relations that link KB integrals to partial derivatives of the chemical potential with respect to the number of molecules (Eq. (3)), partial molar volumes (Eq. (4)), and isothermal compressibilities (Eq. (5)) for binary systems. Based on these relations, other properties can be estimated from KB integrals. Galata et al. [103] used the KB theory to compute thermodynamic mixing properties and excess properties of liquid mixtures. In their work, the authors focus on computing partial derivatives of chemical potential with composition and the Gibbs energy of mixing, $\Delta_{\text{mix}}G$, which are important quantities for the prediction of phase equilibria of liquid mixtures. The prediction of $\Delta_{\text{mix}}G$ and other mixing properties from KB integrals was validated using binary ideal and real LJ mixtures [103]. The KB integrals were found using simulations of finite volumes, and finite-size effects were corrected using the approach of Cortes-Huerto et al. [36] (discussed in section 4).

KB integrals can be used to interpret findings from simulations of biological molecules. In Ref. [31], Pierce et al. presented a review of the applications of

the KB theory to biological systems. One of the valuable applications of the KB theory is to study the effects of co-solvents on biomolecules. Molecular simulation provide local information on the cosolvents surrounding biomolecules and how such an environment affects the structure of biomolecules [29, 31, 104–106]. In 2004, Smith [29] demonstrated how KB integrals can be used to relate simulation results which provide preferential interaction to macroscopic thermodynamic data [107]. Other than studying solvents surrounding biomolecules, the KB theory can be applied directly to systems with interacting biomolecules. However, this application can be hindered by difficulties associated with sampling the phase space of such systems.

KB integrals can be used for the development and parameterization of force fields [61, 108, 109]. Weerasinghe and Smith provide KB derived force fields (KBFF) for a number of mixtures such as, sodium chloride in water [108], urea and water [110], acetone and water [111], and methanol and water [112]. The force fields were parameterized so that KB integrals obtained from experimental data are reproduced (more on the use of experimentally obtained KB integrals are provided in section 7).

The authors of Refs. [61, 108, 109] found that macroscopic properties can be accurately computed using the KBFF models, while addressing solute-solute and solvent-solute molecular structure of the systems considered. For instance, in Ref. [61] the derived KBFF was able to reproduce microstructure of alkaline earth halide salts in water. Ion-ion and ion-water distances provided by the force field were found to agree with those measured by neutron scattering experiments. The same KB force field yielded satisfactory prediction of several macroscopic quantities including molar volumes, and partial derivatives of chemical potential with respect to density. Mijaković et al. [113] compared several force fields, including KBFF, for ethanol-water mixtures. The authors reported that the KB derived force field performed better than other force fields when computing KB integrals and several thermodynamic properties including: excess volumes, excess enthalpy, and self-diffusion coefficients.

7. Inversion of the KB theory

In previous sections, we showed how KB integrals are used to compute several thermodynamic properties of multicomponent liquids. Before molecular simulation were used to compute KB integrals, experimental data were used to obtain KB integrals. This approach is referred to as the inversion of the KB theory [9, 11]. In this section we will briefly discuss the inversion procedure, and some of its applications.

For a binary mixture with components α and β, molar volumes, the isothermal compressibility, and partial derivatives of chemical potential with respect to number of molecules are related to KB integrals $G^{\infty}_{\alpha\alpha}$, $G^{\infty}_{\beta\beta}$ and $G^{\infty}_{\alpha\beta}$ (Eqs. (4), (5), and (3)). Moreover, the Gibbs-Duhem relations apply to these thermodynamic quantities,

$$\begin{aligned}
\rho_\alpha \left(\frac{\partial \mu_\alpha}{\partial N_\alpha}\right)_{T,P,N_\beta} + \rho_\beta \left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha} &= 0 \\
\rho_\beta \left(\frac{\partial \mu_\beta}{\partial N_\beta}\right)_{T,P,N_\alpha} + \rho_\alpha \left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha} &= 0 \\
\rho_\alpha \bar{V}_\alpha + \rho_\beta \bar{V}_\beta &= 1
\end{aligned} \tag{55}$$

where $\bar{V}_\alpha$ and $\bar{V}_\beta$ are the partial molar volumes of components α and β, respectively. Using Eqs. (3), (4), (5), and (55), Ben-Naim [11] derived the following expression for KB integrals of binary mixtures,

$$G^{\infty}_{\alpha\beta} = kT\kappa_T - \frac{\delta_{\alpha\beta}}{\rho_\alpha} + \rho kT \frac{(1-\rho_\alpha \bar{V}_\alpha)(1-\rho_\beta \bar{V}_\beta)}{\rho_\alpha \rho_\beta \left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha}} \tag{56}$$

where the isothermal compressibility κ_T and molar volumes $\bar{V}_\alpha$ and $\bar{V}_\beta$ are obtained from experiments. The term $\left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha}$ can be obtained using second derivatives of the Gibbs excess energy, or experimental vapor pressure data [9]. In Refs. [71] and [114], equations for KB integrals in terms of thermodynamic properties were derived for ternary mixtures.

Ben-Naim [11] introduced the inversion procedure in 1977 and applied it to a mixture of water and ethanol. For water (W) and solute (S) systems, it was

shown that KB integrals obtained from experimental data are useful to study several local phenomena: (1) the quantity $G^{\infty}_{WS} = G^{\infty}_{SW}$ indicates the affinity between the solvent and the solute; (2) KB integrals of water, G^{∞}_{WW}, reflect the water-water affinity, which can be used to study the changes in the molecular structure of water when adding solutes; and (3) KB integrals of solutes, G^{∞}_{SS}, are of particular interest for studying hydrophobic interactions.

Following the paper of Ben-Naim [11], the inversion of the KB theory was applied to study various types of binary and ternary mixtures at the molecular level [15, 115–123]. For instance, Patil [119] computed KB integrals of water-butanol mixtures from experimental data of molar volumes, isothermal compresibility, and vapor pressures. KB integrals of the system considered were used to study local structure at various concentrations. Similarly, Matteoli et al. [118] used molar volumes and isothermal compresibility of mixtures of water and different organic co-solvents to find KB integrals. The KB integrals obtained from the inversion procedure were taken as a measure of the net attraction or repulsion, indicating the hydrophobicity of these mixtures. More recently, Kobayashi et al. [124] used KB integrals to study properties of residual water in ionic liquids. The authors found that the values of KB integrals computed using molecular simulation agree with integrals obtained from experimental data. However, the inversion of the KB theory requires the partial derivatives, $\left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha}$, which are difficult to obtain accurately from experimental data [125]. Matteoli et al. [118] demonstrated how the accuracy of KB integrals obtained from experimental data is very sensitive to uncertainties in partial derivatives of the chemical potential. Alternatively, KB integrals can be obtained from fluctuations in number of molecules measured by angle scattering [19] such as SANS and SAXS [17–19, 21, 22, 95]. Perera et al. [23] examined a number of water-alcohol mixtures using KB integrals and demonstrated that both methods are reliable and should provide similar values of KB integrals. In their study, Perera et al. [23] pointed out possible sources of errors leading to inaccurate KB integrals when using experimental data. For instance, the largest differences between the two methods were observed at the range where the val-

ues of the term $\left(\frac{\partial \mu_\alpha}{\partial N_\beta}\right)_{T,P,N_\alpha}$ in Eq. (56) is close to zero. Furthermore, Almásy et al. [126] obtained KB integrals from SANS as well as from vapor pressure data for an ionic liquid. The authors found that scattering experiments and thermodynamic data provided similar KB integrals.

8. Conclusions

The KB theory provides a solid connection between the microscopic structure of isotropic liquids and their thermodynamic properties, such as partial derivatives of chemical potential with respect to the number of molecules, isothermal compressibility, and partial molar volumes. The key quantities in the KB theory are the KB integrals which are expressed as volume integrals over the radial distribution function. Although developed more than 60 years ago, the theory did not gain mainstream interest until Ben-Naim [11] derived the inversion of the KB theory, where experimental data are used to compute KB integrals. Only in the past few decades the KB theory has been applied to predict thermodynamic properties, mostly by using molecular simulation to compute KB integrals. However, molecular simulation cannot be applied directly to compute KB integrals of infinite and open systems. To connect RDFs and fluctuations of small systems to KB integrals, several methods have been proposed. One of the most practical approaches was derived by Krüger and co-workers [34, 35, 37, 38], where KB integrals of finite subvolumes, that scale with the inverse size of the subvolume, were derived and then extrapolated to the thermodynamic limit. A python package was created to automatically compute the finite-size KB integrals from RDFs (see supplemental information). This and the other methods that we have reviewed here, provide a tool to compute KB integrals of infinite and open subvolumes from simulations of finite systems. From computed KB integrals, various thermodynamics properties can be predicted and used in different applications, which include: predicting properties of single ions in salt solutions, connecting Maxwell-Stefan diffusivities obtained from simulations to Fick diffusivities measured by experiments, and deriving force fields. The KB

theory is a general theory and in principle applies to all isotropic fluids, regardless of the molecular interactions involved. KB integrals were computed for different types of binary and ternary fluids including aqueous alcohol mixtures, salt solutions, and solvents present in biological systems. Difficulties may arise when applying the KB theory for biomolecules and macromolecules as sampling of the phase space can be challenging. Many molecular systems were studied using KB integrals computed simply by truncating the infinite volume integrals introduced by Kirkwood and Buff to finite distances. As shown in this work, truncated integrals do not converge easily and such truncation does not provide the same information as KB integrals in the thermodynamic limit. In fact, truncated integrals correspond to the nonphysical case of subvolumes with zero surface area. Currently, several methods are available to compute KB integrals accurately, and this opens the way to expand the applications of KB theory to real fluids, especially to systems where complex interactions are present and computing thermodynamic properties is challenging. Additionally, it would be interesting to explore the possibilities of computer simulations on KB integrals of non-isotropic fluids.

9. Acknowledgments

This work was sponsored by NWO Exacte Wetenschappen (Physical Sciences) for the use of supercomputer facilities with financial support from the Nederlandse Organisatie voor weten schappelijk Onderzoek (Netherlands Organization for Scientific research, NWO). TJHV acknowledges NWO-CW for a VICI grant. SKS acknowledges financial support from the Research Council of Norway through project number 275754.

References

[1] D. Patterson, Effects of molecular size and shape in solution thermodynamics, Pure and Applied Chemistry 47 (4) (1976) 305–314.

[2] Y. Koga, Solution thermodynamics and its application to aqueous solutions: a differential approach, 2nd Edition, Elsevier, 2017.

[3] L. L. Lee, Molecular thermodynamics of non-ideal fluids, Butterworth-Heinemann, Oxford, UK, 2016.

[4] A. Z. Panagiotopoulos, Molecular simulation of phase equilibria: simple, ionic and polymeric fluids, Fluid Phase Equilibria 76 (1992) 97–112.

[5] F. L. Ning, K. Glavatskiy, Z. Ji, S. Kjelstrup, T. J. H. Vlugt, Compressibility, thermal expansion coefficient and heat capacity of CH_4 and CO_2 hydrate mixtures using molecular dynamics simulations, Physical Chemistry Chemical Physics 17 (4) (2015) 2869–2883.

[6] S. K. Schnell, R. Skorpa, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, J.-M. Simon, Partial molar enthalpies and reaction enthalpies from equilibrium molecular dynamics simulation, The Journal of Chemical Physics 141 (14) (2014) 144501.

[7] E. Hendriks, G. M. Kontogeorgis, R. Dohrn, J.-C. de Hemptinne, I. G. Economou, L. F. Žilnik, V. Vesovic, Industrial requirements for thermodynamics and transport properties, Industrial & Engineering Chemistry Research 49 (22) (2010) 11131–11141.

[8] J. G. Kirkwood, F. P. Buff, The statistical mechanical theory of solutions. I, The Journal of Chemical Physics 19 (6) (1951) 774–777.

[9] A. Ben-Naim, Molecular theory of solutions, Oxford University Press, Oxford, UK, 2006.

[10] S. K. Schnell, X. Liu, J.-M. Simon, A. Bardow, D. Bedeaux, T. J. H. Vlugt, S. Kjelstrup, Calculating thermodynamic properties from fluctuations at

small scales, The Journal of Physical Chemistry B 115 (37) (2011) 10911–10918.

[11] A. Ben-Naim, Inversion of the Kirkwood–Buff theory of solutions: application to the water–ethanol system, The Journal of Chemical Physics 67 (11) (1977) 4884–4890.

[12] D. G. Hall, Kirkwood–Buff theory of solutions. An alternative derivation of part of it and some applications, Transactions of the Faraday Society 67 (1971) 2516–2524.

[13] P. E. Smith, On the Kirkwood–Buff inversion procedure, The Journal of Chemical Physics 129 (12) (2008) 124509.

[14] J. D. Pandey, R. Verma, Inversion of the Kirkwood–Buff theory of solutions: application to binary systems, Chemical Physics 270 (3) (2001) 429–438.

[15] E. Matteoli, A study on Kirkwood–Buff integrals and preferential solvation in mixtures with small deviations from ideality and/or with size mismatch of components. Importance of a proper reference system, The Journal of Physical Chemistry B 101 (47) (1997) 9800–9810.

[16] R. Chitra, P. E. Smith, Molecular association in solution: a Kirkwood–Buff analysis of sodium chloride, ammonium sulfate, guanidinium chloride, urea, and 2,2,2–trifluoroethanol in water, The Journal of Physical Chemistry B 106 (6) (2002) 1491–1500.

[17] I. Shulgin, E. Ruckenstein, Kirkwood–Buff integrals in aqueous alcohol systems: comparison between thermodynamic calculations and X-ray scattering experiments, The Journal of Physical Chemistry B 103 (13) (1999) 2496–2503.

[18] K. Nishikawa, H. Hayashi, T. Iijima, Temperature dependence of the concentration fluctuation, the Kirkwood–Buff parameters, and the correlation

length of tert–butyl alcohol and water mixtures studied by small–angle X–ray scattering, The Journal of Physical Chemistry 93 (17) (1989) 6559–6565.

[19] K. Nishikawa, Simple relationship between the Kirkwood–Buff parameters and the fluctuations in the particle number and concentration obtained by small–angle X–ray scattering: application to tert–butyl alcohol and water mixtures, Chemical Physics Letters 132 (1) (1986) 50–54.

[20] M. Matsumoto, N. Nishi, T. Furusawa, M. Saita, T. Takamuku, M. Yamagami, T. Yamaguchi, Structure of clusters in ethanol–water binary solutions studied by mass spectrometry and X-ray diffraction, Bulletin of the Chemical Society of Japan 68 (7) (1995) 1775–1783.

[21] L. Almásy, G. Jancsó, L. Cser, Application of SANS to the determination of Kirkwood–Buff integrals in liquid mixtures, Applied Physics A 74 (1) (2002) s1376–s1378.

[22] M. A. Blanco, E. Sahin, Y. Li, C. J. Roberts, Reexamining protein–protein and protein–solvent interactions from Kirkwood–Buff analysis of light scattering in multi–component solutions, The Journal of Chemical Physics 134 (22) (2011) 06B606.

[23] A. Perera, F. Sokolić, L. Almásy, Y. Koga, Kirkwood–Buff integrals of aqueous alcohol binary mixtures, The Journal of Chemical Physics 124 (12) (2006) 124515.

[24] J. M. Haile, I. Johnston, A. J. Mallinckrodt, S. McKay, Molecular dynamics simulation: elementary methods, Computers in Physics 7 (6) (1993) 625–625.

[25] D. Frenkel, B. Smit, Understanding molecular simulation: from algorithms to applications, 2nd Edition, Vol. 1, Academic press, London, UK, 2002.

[26] M. P. Allen, D. J. Tildesley, Computer simulation of liquids, 2nd Edition, Oxford University Press, 2017.

[27] D. C. Rapaport, The art of molecular dynamics simulation, Cambridge University Press, 2004.

[28] C. P. Robert, Monte Carlo methods, Wiley Online Library, New Jersey, US, 2004.

[29] P. E. Smith, Cosolvent interactions with biomolecules: relating computer simulation data to experimental thermodynamic data, The Journal of Physical Chemistry B 108 (48) (2004) 18716–18724.

[30] D. Trzesniak, N. F. A. van der Vegt, W. F. van Gunsteren, Computer simulation studies on the solvation of aliphatic hydrocarbons in 6.9 M aqueous urea solution, Physical Chemistry Chemical Physics 6 (4) (2004) 697–702.

[31] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, P. E. Smith, Recent applications of Kirkwood–Buff theory to biological systems, Cell Biochemistry and Biophysics 50 (1) (2008) 1–22.

[32] M. B. Gee, N. R. Cox, Y. Jiao, N. Bentenitis, S. Weerasinghe, P. E. Smith, A Kirkwood–Buff derived force field for aqueous alkali halides, Journal of Chemical Theory and Computation 7 (5) (2011) 1369–1380.

[33] P. Ganguly, N. F. A. van der Vegt, Convergence of sampling Kirkwood–Buff integrals of aqueous solutions with molecular dynamics simulations, Journal of Chemical Theory and Computation 9 (3) (2013) 1347–1355.

[34] P. Krüger, T. J. H. Vlugt, Size and shape dependence of finite–volume Kirkwood-Buff integrals, Physical Review E 97 (5) (2018) 051301.

[35] P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, T. J. H. Vlugt, J.-M. Simon, Kirkwood–Buff integrals for finite volumes, Journal of Physical Chemistry Letters 4 (2) (2013) 235–238.

[36] R. Cortes-Huerto, K. Kremer, R. Potestio, Communication: Kirkwood-buff integrals in the thermodynamic limit from small-sized molecular dy-

namics simulations, The Journal of Chemical Physics 145 (14) (2016) 141103.

[37] N. Dawass, P. Krüger, S. K. Schnell, D. Bedeaux, S. Kjelstrup, J.-M. Simon, T. J. H. Vlugt, Finite-size effects of Kirkwood–Buff integrals from molecular simulations, Molecular Simulation (2017) 1–14.

[38] N. Dawass, P. Krüger, J.-M. Simon, T. J. H. Vlugt, Kirkwood–Buff integrals of finite systems: shape effects, Molecular Physics (2018) 1–8.

[39] X. Liu, A. Martín-Calvo, E. McGarrity, S. K. Schnell, S. Calero, J.-M. Simon, D. Bedeaux, S. Kjelstrup, A. Bardow, T. J. H. Vlugt, Fick diffusion coefficients in ternary liquid systems from equilibrium molecular dynamics simulations, Industrial & Engineering Chemistry Research 51 (30) (2012) 10247–10258.

[40] X. Liu, S. K. Schnell, J.-M. Simon, P. Krüger, D. Bedeaux, S. Kjelstrup, A. Bardow, T. J. H. Vlugt, Diffusion coefficients from molecular dynamics simulations in binary and ternary mixtures, International Journal of Thermophysics 34 (7) (2013) 1169–1196.

[41] S. K. Schnell, P. Englebienne, J.-M. Simon, P. Krüger, S. P. Balaji, S. Kjelstrup, D. Bedeaux, A. Bardow, T. J. H. Vlugt, How to apply the Kirkwood–Buff theory to individual species in salt solutions, Chemical Physics Letters 582 (2013) 154–157.

[42] S. Kjelstrup, S. K. Schnell, T. J. H. Vlugt, J.-M. Simon, A. Bardow, D. Bedeaux, T. Trinh, Bridging scales with thermodynamics: from nano to macro, Advances in Natural Sciences: Nanoscience and Nanotechnology 5 (2) (2014) 023002.

[43] C.-F. Fu, S. X. Tian, Different aggregation dynamics of benzene–water mixtures, Physical Chemistry Chemical Physics 16 (2014) 21957–21963.

[44] M. I. Chaudhari, S. B. Rempe, D. Asthagiri, L. Tan, L. R. Pratt, Molecular theory and the effects of solute attractive forces on hydrophobic interactions, The Journal of Physical Chemistry B 120 (8) (2016) 1864–1870.

[45] D. J. Bonthuis, S. I. Mamatkulov, R. R. Netz, Optimization of classical nonpolarizable force fields for OH– and H3O+, The Journal of Chemical Physics 144 (10) (2016) 104503.

[46] V. Satarifard, S. Kashefolgheta, A. Vila Verde, A. Grafmüller, Is the solution activity derivative sufficient to parametrize ion–ion interactions? ions for TIP5P water, Journal of Chemical Theory and Computation 13 (5) (2017) 2112–2122.

[47] B. A. Strøm, J.-M. Simon, S. K. Schnell, S. Kjelstrup, J. He, D. Bedeaux, Size and shape effects on the thermodynamic properties of nanoscale volumes of water, Physical Chemistry Chemical Physics 19 (2017) 9016–9027.

[48] S. Kashefolgheta, A. Vila Verde, Developing force fields when experimental data is sparse: AMBER/GAFF–compatible parameters for inorganic and alkyl oxoanions, Physical Chemistry Chemical Physics 19 (2017) 20593–20607.

[49] J. Milzetti, D. Nayar, N. F. A. van der Vegt, Convergence of Kirkwood–Buff integrals of ideal and non-ideal aqueous solutions using molecular dynamics simulations, The Journal of Physical Chemistry B 122 (21) (2018) 5515–5526.

[50] G. Tesei, V. Aspelin, M. Lund, Specific cation effects on scn– in bulk solution and at the air–water interface, The Journal of Physical Chemistry B 122 (19) (2018) 5094–5105.

[51] S. H. Jamali, T. van Westen, O. A. Moultos, T. J. H. Vlugt, Optimizing nonbonded interactions of the OPLS force field for aqueous solutions of carbohydrates: How to capture both thermodynamics and dynamics, Journal of Chemical Theory and Computation 14 (12) (2018) 6690–6700.

[52] S. H. Jamali, L. Wolff, T. M. Becker, A. Bardow, T. J. H. Vlugt, O. A. Moultos, Finite–size effects of binary mutual diffusion coefficients from molecular dynamics, Journal of Chemical Theory and Computation 14 (5) (2018) 2667–2677.

[53] S. I. Mamatkulov, K. F. Rinne, R. Buchner, R. R. Netz, D. J. Bonthuis, Water–separated ion pairs cause the slow dielectric mode of magnesium sulfate solutions, The Journal of Chemical Physics 148 (22) (2018) 222812.

[54] R. Fingerhut, J. Vrabec, Kirkwood–Buff integration: A promising route to entropic properties?, Fluid Phase Equilibria, In press. doi:10.1016/j.fluid.2018.12.015.

[55] L. Wolff, S. H. Jamali, T. M. Becker, O. A. Moultos, T. J. H. Vlugt, A. Bardow, Prediction of composition–dependent self–diffusion coefficients in binary liquid mixtures: The missing link for Darken–based models, Industrial & Engineering Chemistry Research 57 (43) (2018) 14784–14794.

[56] T. L. Hill, A different approach to nanothermodynamics, Nano Letters 1 (5) (2001) 273–275.

[57] T. L. Hill, Thermodynamics of small systems, Dover, New York, US, 1994.

[58] J. W. Nichols, S. G. Moore, D. R. Wheeler, Improved implementation of Kirkwood–Buff solution theory in periodic molecular simulations, Physical Review E 80 (5) (2009) 051203.

[59] R. Wedberg, J. P. O'Connell, G. H. Peters, J. Abildskov, Accurate Kirkwood–Buff integrals from molecular simulations, Molecular Simulation 36 (15) (2010) 1243–1252.

[60] D. M. Rogers, Extension of Kirkwood–Buff theory to the canonical ensemble, The Journal of Chemical Physics 148 (5) (2018) 054102.

[61] N. Naleem, N. Bentenitis, P. E. Smith, A Kirkwood–Buff derived force field for alkaline earth halide salts, The Journal of Chemical Physics 148 (22) (2018) 222828.

[62] A. Narayanan Krishnamoorthy, C. Holm, J. Smiatek, Influence of cosolutes on chemical equilibrium: a Kirkwood–Buff theory for ion pair association–dissociation processes in ternary electrolyte solutions, The Journal of Physical Chemistry C 122 (19) (2018) 10293–10302.

[63] A. Perera, L. Zoranić, F. Sokolić, R. Mazighi, A comparative molecular dynamics study of water–methanol and acetone–methanol mixtures, Journal of Molecular Liquids 159 (1) (2011) 52–59.

[64] P. Ganguly, N. F. A. van der Vegt, Convergence of sampling Kirkwood-Buff integrals of aqueous solutions with molecular dynamics simulations, Journal of Chemical Theory and Computation 9 (3) (2013) 1347–1355.

[65] M. Ogawa, Y. Ishii, N. Ohtori, Dynamic behavior of mesoscopic concentration fluctuations in an aqueous solution of 1-propanol by MD simulation, Chemistry Letters 45 (1) (2016) 98–100.

[66] X. Liu, S. K. Schnell, J.-M. Simon, D. Bedeaux, S. Kjelstrup, A. Bardow, T. J. H. Vlugt, Fick diffusion coefficients of liquid mixtures directly obtained from equilibrium molecular dynamics, The Journal of Physical Chemistry B 115 (44) (2011) 12921–12929.

[67] X. Liu, Diffusion in liquids: equilibrium molecular simulations and predictive engineering models, Ph.D. thesis, Delft University of Technology (1 2013).

[68] K. E. Newman, Kirkwood–Buff solution theory: derivation and applications, Chemical Society Reviews 23 (1) (1994) 31–40.

[69] W. H. Stockmayer, Light scattering in multi-component systems, The Journal of Chemical Physics 18 (1) (1950) 58–61.

[70] J. G. Kirkwood, R. J. Goldberg, Light scattering arising from composition fluctuations in multi-component systems, The Journal of Chemical Physics 18 (1) (1950) 54–57.

[71] E. Ruckenstein, I. Shulgin, Entrainer effect in supercritical mixtures, Fluid Phase Equilibria 180 (1-2) (2001) 345–359.

[72] R. Krishna, J. Wesselingh, The Maxwell–Stefan approach to mass transfer, Chemical Engineering Science 52 (6) (1997) 861–911.

[73] R. Taylor, H. A. Kooijman, Composition derivatives of activity coefficient models (for the estimation of thermodynamic factors in diffusion), Chemical Engineering Communications 102 (1) (1991) 87–106.

[74] R. Krishna, Describing the diffusion of guest molecules inside porous structures, The Journal of Physical Chemistry C 113 (46) (2009) 19756–19781.

[75] T. J. H. Vlugt, M. G. Martin, B. Smit, J. I. Siepmann, R. Krishna, Improving the efficiency of the configurational-bias Monte Carlo algorithm, Molecular Physics 94 (4) (1998) 727–733.

[76] W. Shi, E. J. Maginn, Continuous fractional component Monte Carlo: an adaptive biasing method for open system atomistic simulations, Journal of Chemical Theory and Computation 3 (4) (2007) 1451–1463.

[77] W. Shi, E. J. Maginn, Improvement in molecule exchange efficiency in Gibbs ensemble Monte Carlo: development and implementation of the continuous fractional component move, Journal of Computational Chemistry 29 (15) (2008) 2520–2530.

[78] D. Dubbeldam, A. Torres-Knoop, K. S. Walton, On the inner workings of Monte Carlo codes, Molecular Simulation 39 (14-15) (2013) 1253–1292.

[79] A. Torres-Knoop, S. P. Balaji, T. J. H. Vlugt, D. Dubbeldam, A comparison of advanced Monte Carlo methods for open systems: CFCMC vs CBMC, Journal of Chemical Theory and Computation 10 (3) (2014) 942–952.

[80] A. Torres-Knoop, N. C. Burtch, A. Poursaeidesfahani, S. P. Balaji, R. Kools, F. X. Smit, K. S. Walton, T. J. H. Vlugt, D. Dubbeldam,

Optimization of particle transfers in the Gibbs ensemble for systems with strong and directional interactions using CBMC, CFCMC, and CB/CFCMC, The Journal of Physical Chemistry C 120 (17) (2016) 9148–9159.

[81] A. Poursaeidesfahani, A. Torres-Knoop, M. Rigutto, N. Nair, D. Dubbeldam, T. J. H. Vlugt, Computation of the heat and entropy of adsorption in proximity of inflection points, The Journal of Physical Chemistry C 120 (3) (2016) 1727–1738.

[82] A. Poursaeidesfahani, A. Rahbari, A. Torres-Knoop, D. Dubbeldam, T. J. H. Vlugt, Computation of thermodynamic properties in the continuous fractional component Monte Carlo Gibbs ensemble, Molecular Simulation 43 (3) (2017) 189–195.

[83] A. Torres-Knoop, A. Poursaeidesfahani, T. J. H. Vlugt, D. Dubbeldam, Behavior of the enthalpy of adsorption in nanoporous materials close to saturation conditions, Journal of Chemical Theory and Computation 13 (7) (2017) 3326–3339.

[84] A. Rahbari, R. Hens, I. K. Nikolaidis, A. Poursaeidesfahani, M. Ramdin, I. G. Economou, O. A. Moultos, D. Dubbeldam, T. J. H. Vlugt, Computation of partial molar properties using continuous fractional component Monte Carlo, Molecular Physics 116 (21-22) (2018) 3331–3344.

[85] S. K. Schnell, T. J. H. Vlugt, J.-M. Simon, D. Bedeaux, S. Kjelstrup, Thermodynamics of small systems embedded in a reservoir: a detailed analysis of finite size effects, Molecular Physics 110 (11-12) (2012) 1069–1079.

[86] T. L. Hill, Thermodynamics of small systems, The Journal of Chemical Physics 36 (12) (1962) 3182–3197.

[87] J. D. Weeks, D. Chandler, H. C. Andersen, Role of repulsive forces in

determining the equilibrium structure of simple liquids, The Journal of Chemical Physics 54 (12) (1971) 5237–5247.

[88] M. Tuckerman, Statistical mechanics: theory and molecular simulation, Oxford university press, 2010.

[89] J. J. Salacuse, A. R. Denton, P. A. Egelstaff, Finite-size effects in molecular dynamics simulations: static structure factor and compressibility. I. Theoretical method, Physical Review E 53 (3) (1996) 2382.

[90] L. Verlet, Computer "experiments" on classical fluids. II. Equilibrium correlation functions, Physical Review 165 (1) (1968) 201.

[91] A. Santos, Finite-size estimates of Kirkwood–Buff and similar integrals, arXiv preprint arXiv:1806.00821.

[92] https://github.com/sondresc/pykbi.

[93] R. Wedberg, J. P. O'Connell, G. H. Peters, J. Abildskov, Total and direct correlation function integrals from molecular simulation of binary systems, Fluid Phase Equilibria 302 (1-2) (2011) 32–42.

[94] N. H. March, M. P. Tosi, Atomic dynamics in liquids, Dover, New York, US, 1991.

[95] A. Perera, F. Sokolić, L. Almásy, P. Westh, Y. Koga, On the evaluation of the Kirkwood–Buff integrals of aqueous acetone mixtures, The Journal of Chemical Physics 123 (2) (2005) 024503.

[96] S. K. Schnell, T. J. H. Vlugt, J.-M. Simon, D. Bedeaux, S. Kjelstrup, Thermodynamics of a small system in a μT reservoir, Chemical Physics Letters 504 (2011) 199–201.

[97] T. L. Hill, Perspective: Nanothermodynamics, Nano Letters 1 (3) (2001) 111–112.

[98] D. Bedeaux, S. Kjelstrup, Hill's nano-thermodynamics is equivalent with Gibbs' thermodynamics for surfaces of constant curvatures, Chemical Physics Letters 707 (2018) 40 – 43.

[99] R. Skorpa, J.-M. Simon, D. Bedeaux, S. Kjelstrup, The reaction enthalpy of hydrogen dissociation calculated with the small system method from simulation of molecular fluctuations, Physical Chemistry Chemical Physics 16 (2014) 19681–19693.

[100] B. Dahlgren, M. M. Reif, P. H. Hünenberger, N. Hansen, Calculation of derivative thermodynamic hydration and aqueous partial molar properties of ions based on atomistic simulations, Journal of Chemical Theory and Computation 8 (10) (2012) 3542–3564.

[101] R. Taylor, R. Krishna, Multicomponent mass transfer, Vol. 2, John Wiley & Sons, 1993.

[102] B. Dünweg, K. Kremer, Molecular dynamics simulation of a polymer chain in solution, The Journal of Chemical Physics 99 (9) (1993) 6983–6997.

[103] A. A. Galata, S. D. Anogiannakis, D. N. Theodorou, Thermodynamic analysis of Lennard-Jones binary mixtures using Kirkwood–Buff theory, Fluid Phase Equilibria 470 (2018) 25–37.

[104] E. A. Oprzeska-Zingrebe, J. Smiatek, Preferential binding of urea to single–stranded DNA structures: a molecular dynamics study, Biophysical Journal 114 (7) (2018) 1551–1562.

[105] A. Ben-Naim, Theoretical aspects of self-assembly of proteins: A Kirkwood–Buff theory approach, The Journal of Chemical Physics 138 (22) (2013) 06B609–1.

[106] S. Chiba, T. Furuta, S. Shimizu, Kirkwood–buff integrals for aqueous urea solutions based upon the quantum chemical electrostatic potential and interaction energies, The Journal of Physical Chemistry B 120 (31) (2016) 7714–7723.

[107] M. Aburi, P. E. Smith, A combined simulation and Kirkwood–Buff approach to quantify cosolvent effects on the conformational preferences of peptides in solution, The Journal of Physical Chemistry B 108 (22) (2004) 7382–7388.

[108] S. Weerasinghe, P. E. Smith, A Kirkwood–Buff derived force field for sodium chloride in water, The Journal of Chemical Physics 119 (21) (2003) 11342–11349.

[109] M. Kang, P. E. Smith, A Kirkwood–Buff derived force field for amides, Journal of Computational Chemistry 27 (13) (2006) 1477–1485.

[110] S. Weerasinghe, P. E. Smith, A Kirkwood–Buff derived force field for mixtures of urea and water, The Journal of Physical Chemistry B 107 (16) (2003) 3891–3898.

[111] S. Weerasinghe, P. E. Smith, Kirkwood–Buff derived force field for mixtures of acetone and water, The Journal of Chemical Physics 118 (23) (2003) 10663–10670.

[112] S. Weerasinghe, P. E. Smith, A Kirkwood–Buff derived force field for methanol and aqueous methanol solutions, The Journal of Physical Chemistry B 109 (31) (2005) 15080–15086.

[113] M. Mijaković, K. D. Polok, B. Kežić, F. Sokolić, A. Perera, L. Zoranić, A comparison of force fields for ethanol-water mixtures, Molecular Simulation 41 (9) (2015) 699–712.

[114] E. Ruckenstein, I. Shulgin, Effect of a third component on the interactions in a binary mixture determined from the fluctuation theory of solutions, Fluid Phase Equilibria 180 (1-2) (2001) 281–297.

[115] L. Lepori, E. Matteoli, Excess Gibbs energies of the ternary system ehtanol+ tetrahydrofuran+ cyclohexane at 298.15 K, Fluid Phase Equilibria 134 (1-2) (1997) 113–131.

[116] D. M. Pfund, L. L. Lee, H. D. Cochran, Application of the Kirkwood–Buff theory of solutions to dilute supercritical mixtures. II. The excluded volume and local composition models, Fluid Phase Equilibria 39 (2) (1988) 161–192.

[117] A. Ben-Naim, Preferential solvation in two– and in three–component systems, Pure and Applied Chemistry 62 (1) (1990) 25–34.

[118] E. Matteoli, L. Lepori, Solute–solute interactions in water. II. An analysis through the Kirkwood–Buff integrals for 14 organic solutes, The Journal of Chemical Physics 80 (6) (1984) 2856–2863.

[119] K. J. Patil, Application of Kirkwood–Buff theory of liquid mixtures to water–butanol system, Journal of Solution Chemistry 10 (5) (1981) 315–320.

[120] K. J. Patil, G. R. Mehta, S. S. Dhondge, Application of Kirkwood–Buff theory of liquid mixtures to binary aqueous solutions of alcohols, India Journal of Chemistry 33 (12) (1994) 1069–1074.

[121] A. K. Nain, Inversion of the Kirkwood–Buff theory of solutions: application to tetrahydrofuran+ aromatic hydrocarbon binary liquid mixtures, Journal of Solution Chemistry 37 (11) (2008) 1541–1559.

[122] A. K. Nain, Application of the Kirkwood–Buff theory of solutions to acetonitrile+ amide binary mixtures by using inversion procedure and regular solution theory, Journal of Chemical Sciences 121 (3) (2009) 361–367.

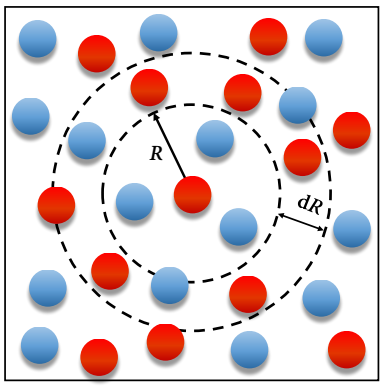
[123] J. J. Shephard, S. K. Callear, S. Imberti, J. S. O. Evans, C. G. Salzmann, Microstructures of negative and positive azeotropes, Physical Chemistry Chemical Physics 18 (28) (2016) 19227–19235.

[124] T. Kobayashi, J. E. S. J. Reid, S. Shimizu, M. Fyta, J. Smiatek, The properties of residual water molecules in ionic liquids: a comparison between direct and inverse Kirkwood–Buff approaches, Physical Chemistry Chemical Physics 19 (29) (2017) 18924–18937.

[125] T. Kato, T. Fujiyama, H. Nomura, Estimation of parameters, G11, G22, and G12 in the Kirkwood–Buff solution theory on the basis of the concentration fluctuation data obtained from Rayleigh scattering, Bulletin of the Chemical Society of Japan 55 (11) (1982) 3368–3372.

[126] L. Almásy, M. Turmine, A. Perera, Structure of aqueous solutions of ionic liquid 1-butyl-3-methylimidazolium tetrafluoroborate by small-angle neutron scattering, The Journal of Physical Chemistry B 112 (8).

(a)



(b)

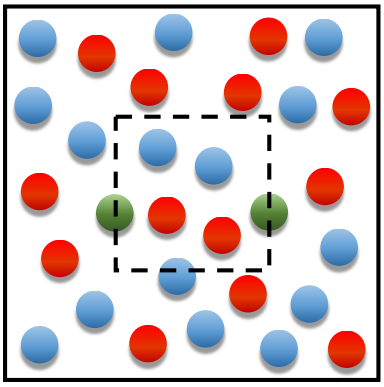


Figure 1 (a) Correcting RDFs using the method of Ganguly and van der Vegt by adjusting the bulk density using the access or depletion of molecules of type β (blue) in a volume V with radius R. (b) Computing KB integrals in the canonical ensemble from local density fluctuations [35, 96]. By using small and open subvolumes embedded in a larger reservoir (simulation box) one can mimic the grand-canonical ensemble.

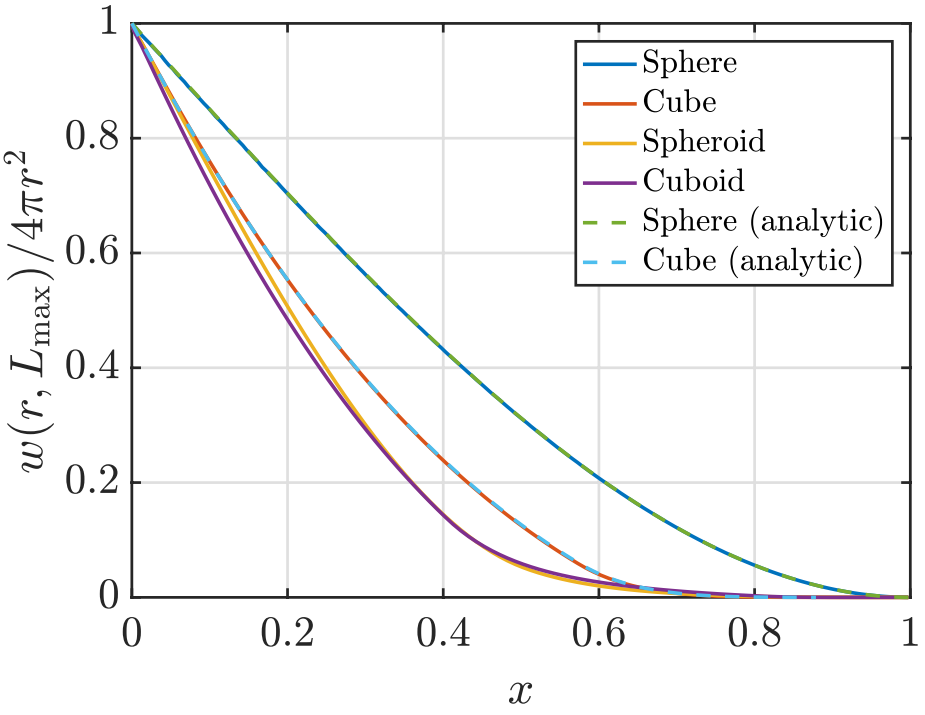


Figure 2: The function $w(r, L_{\max})$ computed numerically for different shapes of the subvolume (numerical procedure is outlined in Ref. [38]). Dashed lines are analytic functions $w(r, L_{\max})$ of Eqs. (14) and (15) for a sphere and cube, respectively.

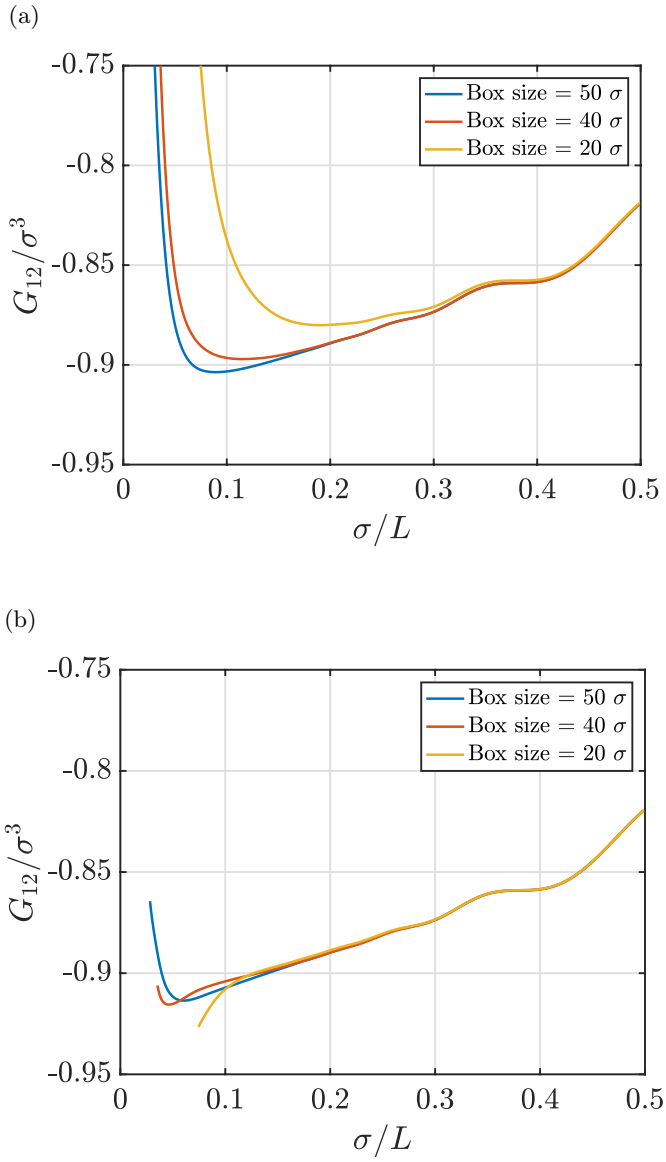


Figure 3 KB integrals of finite (spherical) subvolumes, G_{12}^V, versus $1/L$ (L is the diameter of the sphere) from RDFs computed from MD simulations, in the NVT ensemble, of an equimolar binary WCA fluid. The parameters of the WCA potential are: $\sigma_{11} = \sigma_{22} = 1.0$, $\epsilon_{11} = \epsilon_{22} = 1.0$ and $\epsilon_{12} = 0.5$. The dimensionless temperature and density are fixed at $T = 1.8$ and $\rho = 0.7$, respectively. Eq. (13) is used to integrate $g_{12}(r)$ and find G_{12}^V. The radial distribution function $g_{12}(r)$ is (a) not corrected and (b) corrected using the method by Ganguly and van der Vegt (Eq. (17)).

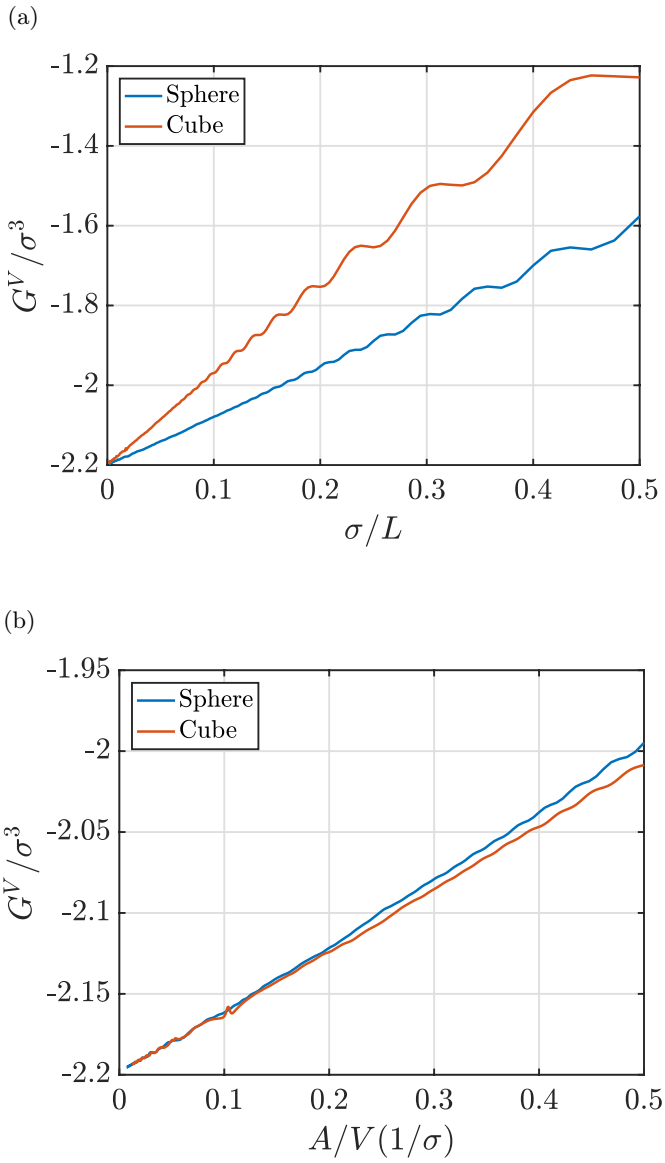


Figure 4 KB integrals of the fluid described by Eq. (23) for a sphereical and cubic subvolumes, versus (a) the inverse of the size of the subvolume, $1/L$ and (b) the ratio of the surface area to the volume (A/V) of the subvolume.

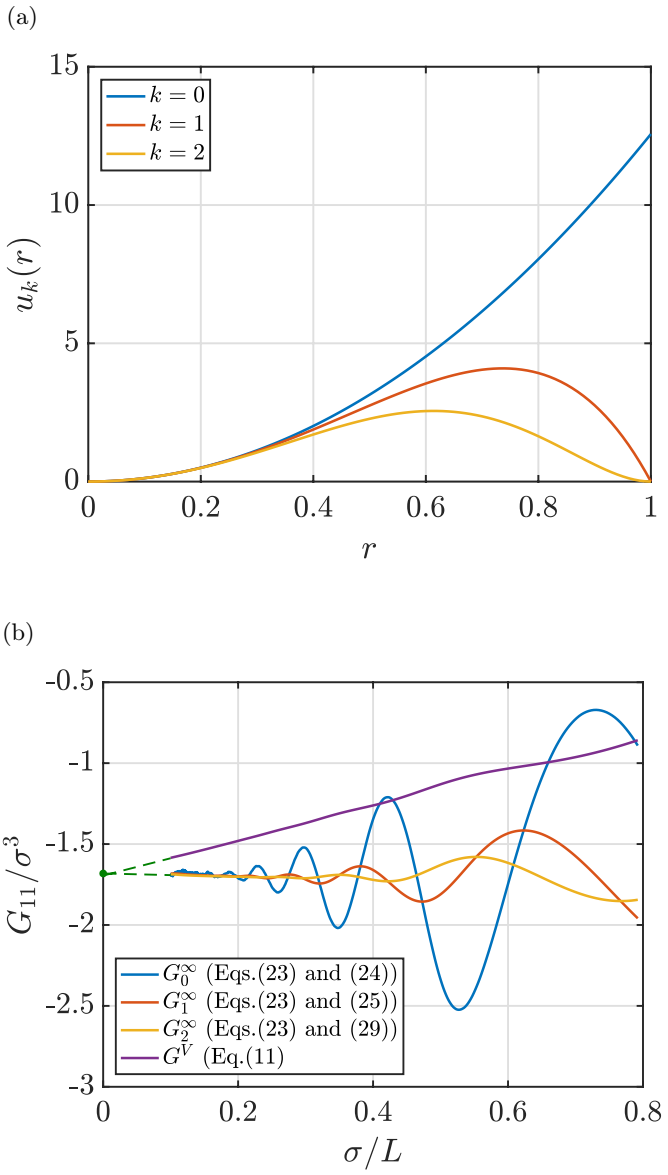


Figure 5 (a) The weight functions $u_k(r)$ discussed in section 3.3 to estimate KB integrals in the thermodynamic limit. (b) KB integrals, G_{11}, versus $1/L$ from direct extrapolation (Eq. (25)) as well as from the expression of KB integrals of finite subvolumes (Eq. (13)). Green dashed lines indicate extrapolation to the thermodynamic limit.

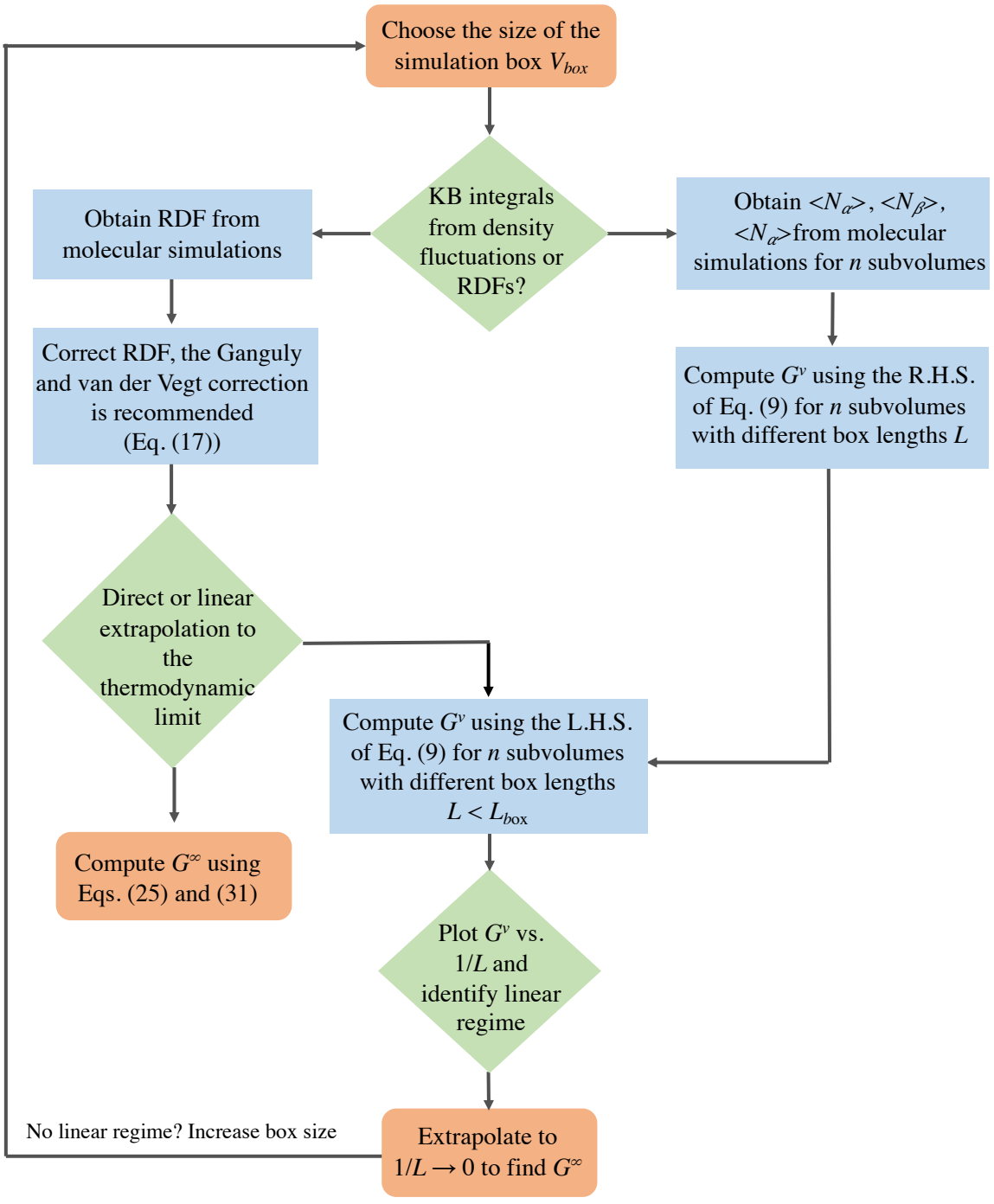


Figure 6: A guideline for computing KB integrals using the method of Krüger and co-workers [34, 35, 37, 38].