The entropy of a complex molecule

Grme Faure,^{1,2} Rafael Delgado-Buscalioni,³ and Pep Espaol⁴

¹CEA, DAM, DIF, F-91297 Arpajon, France

²Universit Paris-Est, CERMICS (ENPC), F-77455 Marne-la-Valle, France

³Departamento de Fsica Terica de la Materia Condensada Universidad Autnoma de Madrid, Madrid 28049, Spain

⁴Departamento de Fsica Fundamental, Universidad Nacional de Educacin a Distancia, Apartado 60141, E-28080 Madrid, Spain

(Received 6 April 2017; accepted 22 May 2017; published online 12 June 2017)

Entropy is a central concept in the theory of coarse-graining. Through Einstein's formula, it provides the equilibrium probability distribution of the coarse-grained variables used to describe the system of interest. We study with molecular dynamics simulations the equilibrium probability distribution of *thermal blobs* representing at a coarse-grained level star polymer molecules in melt. Thermal blobs are characterized by the positions and momenta of the centers of mass, and internal energies of the molecules. We show that the entropy of the level of description of thermal blobs can be very well approximated as the sum of the thermodynamic entropy of each single molecule considered as isolated thermodynamic systems. The entropy of a single molecule depends on the intrinsic energy, involving only contributions from the atoms that make the molecule and not from the interactions with atoms of other molecules. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4984965>]

I. INTRODUCTION

There is much current interest in developing coarse-grained (CG) models for complex molecular systems that, while capturing some essential microscopic details, allow one to simulate large time and length scales that otherwise would be impossible to address with direct MD simulations. Typically, a complex molecule is represented by one or several blobs, in de Gennes terminology.¹ These blobs contain several atoms that are treated as a single entity.² The strategy based on the theory of coarse-graining^{3–6} has proved to be an interesting route providing a sounded foundation for these types of models.^{7–12} The resulting dynamics of the blobs is akin to the Dissipative Particle Dynamics (DPD)¹³ model, with a truly microscopic foundation. Unfortunately the DPD model is isothermal and does not allow one to study thermal processes in the system. Shock disruption of cell membrane for drug delivery,¹⁴ cavitation microsurgery,¹⁵ or multiscale simulation of explosives¹⁶ require the consideration of energy transport in the system. Although an energy conserving version, called Dissipative Particle Dynamics with Energy conservation (DPDE), of DPD exists,^{17,18} it was lacking a microscopic foundation.

Recently, we have considered in Ref. 19 (I, from now on) a first principles derivation of the blob description to non-isothermal situations. The resulting blobs are referred to as *thermal blobs* and are characterized by the position and momenta of the center of mass (CoM) of the atoms that make up the blob, as well as its internal energy. This additional degree of freedom allows one to describe temperature gradients in coarse-grained descriptions of complex molecules. Therefore, Ref. 19 provides a sounded foundation from first principles for the DPDE model. There is much interest in the microscopic description of heat conduction for solid materials at the nano-scale as well, see, for example, Ref. 20 and references therein.

One crucial ingredient of the theory of coarse-graining is the notion of entropy of the level of description. A level of description is characterized by the coarse-grained (CG) variables used to describe the system at that particular resolution. The entropy is a function of the CG variables of that level. We stress that there is no such thing as *the* entropy of a system. Rather, each level of description (of the same system) has its own entropy, which is a distinct function of the CG variables of that level. A particularly important level of description is that of macroscopic thermodynamics. This level is characterized by CG variables that are dynamic invariants of the system and describes the equilibrium state of the system. We may refer to the entropy of the level of description of macroscopic thermodynamics as the *thermodynamic entropy*.

The entropy of a particular level of description determines, through Einstein's formula, the probability distribution of the CG variables of that level.²¹ It also determines the irreversible "forces" acting on the CG variables. In I we made a key approximation (KA) in the calculation of the entropy of the level of description of thermal blobs. This KA consists in assuming that the microscopic intermolecular potential of interaction between complex molecules can be approximated (within phase space integrals and, hence, in a statistical sense) by a function of the coordinates of the centers of mass alone. All the details of the microscopic configurations of the complex molecules in the intermolecular potential energy are assumed to be sufficiently well represented by just the positions of the center of mass. Note that pairwise additivity was not assumed and, in principle, many body or density dependent potentials can be encompassed in the framework.²² The main consequence of this KA is that the entropy of the level of description of thermal blobs can be written as the sum of the thermodynamic entropies of each molecule, as if the molecule was an isolated thermodynamic system. This gives

meaning and usefulness to the notion of the *entropy of a single molecule*. We note in this respect that there is much current interest in the thermodynamics of small systems.²³ The KA simplifies enormously the CG description in terms of thermal blobs presented in I. If the KA is not satisfied, much more thinking is required. The purpose of the present work is to asses and validate the KA of Ref. 19 in a particular model system.

This article is organized as follows. First we establish in Sec. II the coarse-grained description for star polymers based on the formalism developed in Ref. 19. In Sec. III, we then check the KA by studying the independence of selected coarse-grained variables, such as positions, momenta, and energies of molecules, and comparing the entropy of star polymers in a melt and in vacuum.

II. THE COARSE-GRAINED DESCRIPTION OF STAR POLYMERS

We consider a model for a melt of identical star polymer molecules (as shown in Fig. 1) that has been already used in the isothermal CG version of star polymer molecules in Ref. 9. Each star polymer molecule is described at a coarse level as a single thermal blob. The CG variables used to describe the blob are the position, $\hat{\mathbf{R}}_\mu$, and momentum, $\hat{\mathbf{P}}_\mu$, of the centers of mass (CoMs) of the blobs, together with the internal energy $\hat{\mathcal{E}}_\mu$ of the blob. The Greek index μ runs from $\mu = 1, \dots, M$. The coarse variables are the following functions of the microstate z (coordinates $\mathbf{q}_i$ and momenta $\mathbf{p}_i = m_i \mathbf{v}_i$ of each atom in the system):

$$\begin{aligned}\hat{\mathbf{R}}_\mu(z) &= \frac{1}{M_\mu} \sum_i m_i \mathbf{q}_i \delta_\mu(i), \\ \hat{\mathbf{P}}_\mu(z) &= \sum_i \mathbf{p}_i \delta_\mu(i), \\ \hat{\mathcal{E}}_\mu(z) &= \sum_i \frac{m_i}{2} (\mathbf{v}_i - \hat{\mathbf{V}}_\mu(z))^2 \delta_\mu(i) + \hat{\phi}_\mu(z).\end{aligned}\quad (1)$$

FIG. 1. Star polymer molecules (each molecule with a distinct color) in a melt are coarse-grained at the level of thermal blobs described by the position and momentum of the center of mass, plus the internal energy of the molecule with respect to the center of mass.

The indicator symbol $\delta_\mu(i)$ takes the value 1 if the atom i belongs to the blob μ and zero otherwise. This indicator satisfies

$$\begin{aligned}\sum_\mu \delta_\mu(i) &= 1, \\ \delta_\mu(i) \delta_\nu(i) &= \delta_{\mu\nu} \delta_\mu(i).\end{aligned}\quad (2)$$

These expressions reflect that the atom i belongs to one and only one blob. The mass, the velocity, and the potential energy of the μ th blob in (1) are defined as

$$\begin{aligned}M_\mu &= \sum_i m_i \delta_\mu(i), \\ \hat{\mathbf{V}}_\mu(z) &= \frac{\hat{\mathbf{P}}_\mu(z)}{M_\mu}, \\ \hat{\phi}_\mu(z) &= \frac{1}{2} \sum_{ij} \phi_{ij} \delta_\mu(i),\end{aligned}\quad (3)$$

with ϕ_{ij} the pairwise potential energy between atoms i and j . The total energy and total momentum of the system can be expressed explicitly in terms of the CG variables (1) as follows:

$$\begin{aligned}\hat{H}(z) &= \sum_\mu \left[\frac{\hat{\mathbf{P}}_\mu^2(z)}{2M_\mu} + \hat{\mathcal{E}}_\mu(z) \right], \\ \hat{\mathbf{P}}(z) &= \sum_\mu \hat{\mathbf{P}}_\mu(z),\end{aligned}\quad (4)$$

as can be seen by using definitions (1) and properties (2).

Another microscopic quantity that turns to be crucial in the formulation of the model is the *intrinsic* energy, defined in microscopic terms as

$$\begin{aligned}\hat{U}_\mu(z) &= \sum_i \frac{m_i}{2} (\mathbf{v}_i - \hat{\mathbf{V}}_\mu(z))^2 \delta_\mu(i) \\ &\quad + \frac{1}{2} \sum_{ij} \phi_{ij}^{\text{intra}} \delta_\mu(i) \delta_\mu(j),\end{aligned}\quad (5)$$

where the last contribution contains the potential ϕ^{intra} of interaction of atoms within the molecule. The difference between the *internal* energy $\hat{\mathcal{E}}_\mu(z)$ and the *intrinsic* energy $\hat{U}_\mu(z)$ is the part of the potential energy of the molecule which is due to the interactions with other molecules. This *intermolecular* potential energy is assumed, under the KA, to be a function of the microstate of the system only through the position of the centers of mass of the blobs, at least in a statistical sense.

As shown in I, under the KA the joint probability density of finding the values $\mathbf{R}, \mathbf{P}, U = \{\mathbf{R}_1, \dots, \mathbf{R}_M, \mathbf{P}_1, \dots, \mathbf{P}_M, U_1, \dots, U_M\}$ has the following form:¹⁹

$$\begin{aligned}P^{\text{eq}}(\mathbf{R}, \mathbf{P}, U) &= \frac{1}{\Omega} \delta \left(\sum_\mu \left[\frac{\mathbf{P}_\mu^2}{2M_\mu} + U_\mu + \phi_\mu(\mathbf{R}) \right] - E_0 \right) \\ &\quad \times \delta \left(\sum_\mu \mathbf{P}_\mu - \mathbf{P}_0 \right) \exp\{S(\mathbf{R}, U)/k_B\}\end{aligned}\quad (6)$$

under the understanding that we know with absolute certainty that the total energy of the system is E_0 and the total momentum

is $\mathbf{P}_0$, which is the case in usual MD simulations. Equation (6) is Einstein's formula giving the probability of the CG variables in terms of the entropy of the level of description,²¹ adequately modified to include dynamic invariants.²⁴ Here $S(\mathbf{R}, U)$ is the entropy of the present level of description that depends, in principle, on all the CG variables. Due to Galilean invariance, though, in this level of description the entropy does not depend on momenta. As shown in I, under the KA the intermolecular potential is approximated with a function of the position of the centers of mass of the blobs, and the entropy becomes purely additive, and of the form

$$S(\mathbf{R}, U) = \sum_{\mu} S_{\mu}(U_{\mu}), \quad (7)$$

where the thermodynamic entropy of the blob $S_{\mu}(U)$ is the Boltzmann entropy of blob μ as a function of the intrinsic energy, defined through the logarithm of “the number of microstates compatible with the macrostate U ,” this is

$$S_{\mu}(U) = k_B \ln \int dz_{\mu} \delta(\hat{\mathbf{R}}_{\mu}) \delta(\hat{\mathbf{P}}_{\mu}) \times \delta \left(U - \sum_i^N \frac{m_i}{2} \mathbf{v}_i^2 \delta_{\mu}(i) - \hat{\phi}_{\mu}^{\text{intra}}(z_{\mu}) \right), \quad (8)$$

where the phase space integration is over the degrees of freedom z_{μ} (including positions and momenta) of a *single isolated* molecule μ . Entropy (8) is the usual definition of the entropy of macroscopic thermodynamics. In the present case, the “thermodynamic system” is a single molecule.

Probability (6) is microcanonical in the total energy and momenta at the coarse-grained level. If we assume equivalence of ensembles for a sufficiently large number of molecules, we can approximate the microcanonical distribution with a canonical one, leading to

$$P^{\text{eq}}(\mathbf{R}, \mathbf{P}, U) = \frac{1}{Z} \exp -\beta_0 \sum_{\mu} \left[\frac{\mathbf{P}_{\mu}^2}{2M_{\mu}} + \mathcal{F}_{\mu}(U_{\mu}) + \Phi_{\mu}(\mathbf{R}) \right] \quad (9)$$

where Z stands for the normalization, and where we have introduced the free energy of blob μ as

$$\mathcal{F}_{\mu}(U_{\mu}) = U_{\mu} - T_0 S_{\mu}(U_{\mu}). \quad (10)$$

The temperature T_0 and $\beta_0 = (k_B T_0)^{-1}$ govern the overall thermodynamic point of the system and are related in a standard way with the total energy E_0 of the system.

The most salient feature of the probability distribution (9) is that it factorizes and, therefore, *positions, momenta, and intrinsic energies are statistically independent*. The purpose of this paper is to check that the probability distribution (9) adequately describes the distribution of the coarse-grained variables for a system of star polymers in a melt, by focusing on the statistical independence displayed. Because (9) is a joint probability in a multidimensional space, it is not feasible to make histograms, and we need to select a number of observables out of this distribution function in order to validate the statistical independence of the variables. This will be done below.

III. RESULTS

A. Simulation details

We consider a system of star polymers similar to the ones described in Refs. 9 and 10. A number N_b of branches containing N_m identical monomers of mass m_0 are attached to a central monomer in order to construct a star polymer molecule. We will denote by a “ N_b - N_m star polymer” such a molecule (e.g., a 6-3 star polymer has six branches with three monomers each). All monomers interact with a purely repulsive Lennard-Jones potential whose expression is given by

$$\mathcal{U}_{LJ,\text{rep}}(r) = \begin{cases} 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 + \frac{1}{4} \right], & \text{if } r \leq 2^{\frac{1}{6}} \sigma, \\ 0, & \text{if } r > 2^{\frac{1}{6}} \sigma. \end{cases}$$

Results will be given in Lennard-Jones reduced units, where $\sigma = 1$, $\epsilon = 1$, and $m_0 = 1$. The bonds are represented with a quadratic potential with stiffness $k = 20$,

$$\mathcal{U}_{\text{bond}}(r) = \frac{1}{2} k (r - l_{\text{bond}})^2.$$

The equilibrium bond length is set to $l_b = 1.147$ for monomers inside the branches and to $l_c = 1.615$ for bonds involving the central monomer.

Typically, we consider 1000 star polymer molecules. The molecules are first set on a lattice and successive NVT simulations are carried for 25 000 steps using a time step $\Delta t = 0.001$ with a Langevin thermostat at a temperature $T = 1$ while the monomer density is progressively increased to reach the desired value. The Langevin friction is set to $\gamma = 1$. This preparation allows one to have the system in a controlled thermodynamic state with different monomer densities $\rho = 0.1, 0.4, 0.7$ at the same temperature $T = 1$ in Lennard-Jones units. We then perform our simulations in the NVE ensemble using the velocity-Verlet integration scheme with a time step $\Delta t = 0.001$ during a total time $t_{\text{sim}} = 5000$.

In Secs. III B–III F we focus on star polymers with 6 branches of 3 monomers each. We will extend the study to somewhat larger molecules in Sec. III G.

B. The distribution of $\mathbf{P}_{\mu}$ and U_{μ}

In Fig. 2 we plot the probability distribution of the momentum of the center of mass of a single molecule. According to probability (9), the momenta are distributed according to a Maxwellian. This is what is observed in our simulations. Note that the temperature is fixed to 1 on average during the thermostating NVT process. When switching to NVE, the system may have a slightly different value of the average kinetic energy. The measured inverse thermal energy in the NVE simulation is $\beta = 1.05$.

In Fig. 3 we plot the distribution of the intrinsic energy of a single star polymer molecule.

C. Correlations of U and $\mathbf{P}$

In order to validate the independence of momentum and intrinsic energy, we compute the covariance of the kinetic energy of the center of mass of a molecule with the intrinsic energy $\left\langle \Delta \left(\frac{\mathbf{P}_{\mu}^2}{2M_{\mu}} \right) \Delta U_{\mu} \right\rangle$. According to form (9), this should be zero. Here and in what follows, we refer to $\Delta A = A - \langle A \rangle_{\text{eq}}$

FIG. 2. The probability density $P(P_x)$ that the x component of the momentum of the center of mass of one given molecule is P_x . Also shown is a fit to a Maxwellian with temperature $\beta = 1.05$.

as the fluctuations with respect to the equilibrium average. In order to check how small this correlation is, we introduce the normalized covariance

$$C_{KU} \equiv \frac{\left\langle \Delta \left(\frac{\mathbf{p}_\mu^2}{2M_\mu} \right) \Delta U_\mu \right\rangle}{\sqrt{\left\langle \left(\Delta \left(\frac{\mathbf{p}_\mu^2}{2M_\mu} \right) \right)^2 \right\rangle \left\langle (\Delta U_\mu)^2 \right\rangle}}. \quad (11)$$

The average correlation C_{KU} computed over all the molecules at a monomer density $\rho = 0.3$ is $C_{KU} = 5.08 \times 10^{-2}$ that we deem very small. Perhaps a more illustrative test is obtained from the scatter plot in Fig. 4, where we plot the instantaneous kinetic energy and intrinsic energy for each blob at a given time. There is no particular trend, indicating that the probability factorizes $P(K, U) = P(K)P(U)$. Therefore, momentum and intrinsic energy are not correlated. A more quantitative test is by looking at the relative error $\left| \frac{P(K, U) - P(K)P(U)}{P(K, U)} \right|$. This error should be zero under statistical independence. In Fig. 5

FIG. 3. The probability density $P(U)$ that the intrinsic energy of a molecule is U .

FIG. 4. Scatter plot of the value of instantaneous kinetic energy and intrinsic energy for each blob at a given time. These energies have been normalized to have 0 mean and a variance of 1.

the relative error is plotted. In order to avoid spurious errors in regions with small values of $P(K, U)$, the error is not computed in those regions. We observe that the relative error is very small in the region where the probability is appreciable, indicating a good degree of statistical independence.

It may appear as a surprise that there is no appreciable correlation between the intrinsic energy (that includes kinetic energy in its definition) and the kinetic energy of the blob. However, note that the kinetic energy contribution in the intrinsic energy is with respect to the center of mass of the molecule. Such “thermal” kinetic energy need not be strongly correlated with the translational kinetic energy $\mathbf{P}^2/2M_\mu$.

D. The blob entropy

A very stringent test of the KA comes from the comparison of the two different and independent routes for computing the entropy of a molecule as a function of the intrinsic energy of the molecule. First, consider the microscopic definition (8) in which an isolated molecule is considered as a thermodynamic system in its own right. From the virial theorem,²⁵ we obtain

FIG. 5. Density plot colored with respect to the relative error $\left| \frac{P(K, U) - P(K)P(U)}{P(K, U)} \right|$. If the value of $P(K, U)$ is too small, the relative error is not computed and displayed in gray. Note that in the region where the probability is noticeable different from zero, the relative error is very small.

the equipartition of energy in the form

$$\left\langle \sum_{i_\mu}^{N_\mu} \frac{\mathbf{p}_{i_\mu}^2}{2m_{i_\mu}} \right\rangle^U = \frac{3}{2} N_\mu k_B T(U), \quad (12)$$

where N_μ is the number of atoms of molecule μ , and the average $\langle \dots \rangle^U$ is a microcanonical average of an isolated molecule at the energy U . By running an MD simulation of an isolated molecule for different energies U , we may plot the function $T(U)$ and, through integration, recover the entropy function

$$\frac{S_\mu(U)}{k_B} = \frac{S_\mu(U_0)}{k_B} + \int_{U_0}^U dU' \frac{1}{k_B T(U')}. \quad (13)$$

A second independent route to compute the entropy function of the blob is a consequence of form (9) that implicitly assumes that the microscopic potential energy of interaction between molecules may be approximated by a function of the centers of mass of the blobs. If we integrate (9) over all variables except the intrinsic energy of one blob, we obtain the probability density $P(U)$ that one blob has a particular intrinsic energy U . This function is plotted in Fig. 3. Using (10) we obtain easily the identity

$$\frac{S_\mu(U)}{k_B} = \ln P(U) + \beta U, \quad (14)$$

where β is obtained from fitting a Maxwellian to the momentum distribution of a single molecule. Note that while (13) is computed with a single isolated molecule in vacuum, (14) is computed for the molecules in the melt. Both functions must be identical if the KA on the microscopic intermolecular potential is valid. Alternatively, and in order to avoid the integral, we may take the derivative of (14) in order to obtain the temperature

$$\frac{1}{k_B T(U)} = \frac{P'(U)}{P(U)} + \beta \quad (15)$$

and compare it with the temperature obtained from the equipartition theorem (12) in an isolated molecule.

FIG. 6. The entropy of an isolated molecule from Eq. (13) (solid line) compared with the entropy of a molecule inside a melt, computed from Eq. (14) (dashed lines), for different monomer densities. The curves have been superimposed taking advantage of the fact that the entropy is defined up to an additive constant.

FIG. 7. Temperature of a single molecule in vacuum as a function of the intrinsic energy, computed from (12) (solid line). Temperature of a molecule in the melt as a function of the intrinsic energy, computed from (15) (dashed lines) for different monomer densities.

In Fig. 6 we plot the two entropy functions (13) and (14) as a function of the intrinsic energy, for different monomer densities. The agreement of the two curves for the entropy of a molecule in the melt and in vacuum is quite remarkable. The temperature function for an isolated molecule (12) and for a molecule in melt (15) is practically linear, as shown in Fig. 7. The slope of the temperature function is the heat capacity of a molecule. In Fig. 8 we plot the heat capacity of a star polymer molecule in the melt for different monomer densities, compared to the heat capacity of an isolated star polymer molecule. The heat capacity of a molecule in the melt is not strongly dependent on the monomer density and it is within an $\sim 5\%$ difference from the heat capacity of a molecule in vacuum. It should be clear that the heat capacity of the whole sample of star polymers *is not* the sum of the heat capacity of each molecule. Energy can be stored in both the intrinsic energy of the molecules and also in the intermolecular potential energy between molecules.

FIG. 8. The heat capacity of a molecule in the melt for different monomer densities (diamonds) and in vacuum (red line).

E. Correlations of $\mathbf{R}$ and $\mathbf{U}$

Consider the joint probability distribution $P(\mathbf{R}_\mu, \mathbf{R}_\nu, U_\mu, U_\nu)$ that molecule μ has position and intrinsic energy $\mathbf{R}_\mu, U_\mu$ and molecule ν has position and intrinsic energy $\mathbf{R}_\nu, U_\nu$. It is given by

$$P(\mathbf{R}_\mu, \mathbf{R}_\nu, U_\mu, U_\nu) = \int dz \rho_{\text{eq}}(z) \times \delta(\hat{\mathbf{R}}_\mu(z) - \mathbf{R}_\mu) \delta(\hat{\mathbf{R}}_\nu(z) - \mathbf{R}_\nu) \times \delta(\hat{U}_\mu(z) - U_\mu) \delta(\hat{U}_\nu(z) - U_\nu), \quad (16)$$

where $\rho_{\text{eq}}(z)$ is the equilibrium ensemble. Translation and rotation invariance of the Hamiltonian, and hence of the equilibrium ensemble, implies that the above probability is, in fact, just a function of three variables,

$$P(\mathbf{R}_\mu, \mathbf{R}_\nu, U_\mu, U_\nu) = P(R_{\mu\nu}, U_\mu, U_\nu), \quad (17)$$

where $R_{\mu\nu} = |\mathbf{R}_{\mu\nu}|$ is the distance between the centers of mass of molecule μ and ν . The above probability is defined on a three-dimensional space and is too costly to evaluate by histograms. We may, therefore, look at the second moments with respect to intrinsic energy variables and at the following conditional covariance:

$$\begin{aligned} \langle \hat{U}_\mu \hat{U}_\nu \rangle^{R_{\mu\nu}} &= \frac{\int dU_\mu dU_\nu U_\mu U_\nu P(R_{\mu\nu}, U_\mu, U_\nu)}{\int dU_\mu dU_\nu P(R_{\mu\nu}, U_\mu, U_\nu)} \\ &= \frac{\int dz \rho_{\text{eq}}(z) \hat{U}_\mu(z) \hat{U}_\nu(z) \delta(\hat{\mathbf{R}}_\mu) \delta(\hat{\mathbf{R}}_\mu - \mathbf{R}_{\mu\nu})}{\int dz \rho_{\text{eq}}(z) \delta(\hat{\mathbf{R}}_\mu) \delta(\hat{\mathbf{R}}_\mu - \mathbf{R}_{\mu\nu})}. \end{aligned} \quad (18)$$

The intuitive meaning of this conditional correlation is that it gives the correlation of the intrinsic energies of two molecules separated a distance $R_{\mu\nu}$. Clearly, if the position and intrinsic energies are statistical independent, then $P(R_{\mu\nu}, U_\mu, U_\nu) = P(R_{\mu\nu})P(U_\mu, U_\nu)$, and the above covariance should be independent of $R_{\mu\nu}$. In order to increase statistics, we will instead consider the following covariance:

$$\langle U_1 U_2 \rangle^r = \frac{\sum'_{\mu\nu} \int dz \rho_{\text{eq}}(z) \hat{U}_\mu(z) \hat{U}_\nu(z) \delta(r - \hat{R}_{\mu\nu})}{\sum'_{\mu\nu} \int dz \rho_{\text{eq}}(z) \delta(r - \hat{R}_{\mu\nu})}, \quad (19)$$

where the prime in the summation means that the terms $\mu = \nu$ are excluded. If the KA is valid, $P(R_{\mu\nu}, U_\mu, U_\nu) = P(R_{\mu\nu})P(U_\mu)P(U_\nu)$ and, from (18) we should have that the conditional expectation (19) becomes identical to the unconditional expectation

$$\langle U_1 U_2 \rangle^r = (\langle U \rangle)^2 \quad (20)$$

and, therefore, independent of the distance. If this is the case, we see that Eq. (19) implies the identity

$$u(r) = g(r), \quad (21)$$

where the radial distribution function is

$$g(r) = \frac{1}{N_0} \sum'_{\mu\nu} \int dz \rho_{\text{eq}}(z) \delta(r - \hat{R}_{\mu\nu}), \quad (22)$$

FIG. 9. Radial distribution function $g(r)$ of the centers of mass of the molecules for two different monomer densities. Also shown is the correlation $u(r)$ defined in (23). Identity (21) which is a consequence of the KA is satisfied very well.

where N_0 is the usual normalization of the radial distribution function, and we introduce the following quantity:

$$u(r) \equiv \frac{1}{N_0} \sum'_{\mu\nu} \int dz \rho_{\text{eq}}(z) \frac{\hat{U}_\mu(z) \hat{U}_\nu(z)}{(\langle U \rangle)^2} \delta(r - \hat{R}_{\mu\nu}), \quad (23)$$

which is a sort of normalized covariance similar to the radial distribution function. Therefore, verification of identity (21) is a proof of independence of intrinsic energy and positions of the blobs.

We show in Fig. 9 both functions $g(r)$ and $u(r)$ that turn out to be indistinguishable. It is possible, however, to detect some differences at very short distances, by using Eq. (20) directly. In Fig. 10 we plot Eq. (20) and the radial distribution function. We observe that when two molecules have their centers of mass “on top of each other,” then the intrinsic energy of each molecule which is entirely due to self-interaction is affected by the presence of the other interdigitated molecule. This small effect occurs rarely, as can be seen from the fact that it occurs

FIG. 10. The conditional covariance (20) (solid lines) plotted together with the radial distribution function (dotted lines), whose full curve is given in Fig. 9.

in the region where the radial distribution function is very small.

F. Correlations of U and ρ

Another illuminating criterion is the calculation of the correlations between the intrinsic energy $\hat{U}_\mu$ of the blob and its “volume.” Borrowing an idea from Smoothed Particle Hydrodynamics (SPH),²⁶ we may define the volume of a blob as a function of the positions of neighbour blobs through the inverse of the local density of molecules which in turn is defined as

$$\hat{\rho}_\mu(z) = \sum_\nu W(\hat{R}_{\mu\nu}(z)), \quad (24)$$

where $W(r)$ is a bell-shaped weight function normalized to unity,

$$\int d\mathbf{r} W(|\mathbf{r}|) = 1. \quad (25)$$

When other molecules are close to molecule μ , the contribution to the sum is large and, consequently, the density of molecule μ is larger. A usual selection in SPH is the Lucy function for $W(r)$,

$$W(r) = \frac{105}{16\pi h^3} \left(1 + 3\frac{r}{h}\right) \left(1 - \frac{r}{h}\right)^3 \theta(h - r), \quad (26)$$

where h is the support of the function, which should be larger than the typical distance between blobs, and θ is the Heaviside function. In the following, we choose h such that the second peak of the radial distribution function is included in the kernel support (i.e., $h = 10$ for $\rho = 0.3$).

We will compute the following correlation:

$$\begin{aligned} \langle \Delta \hat{U}_\mu \Delta \hat{\rho}_\mu \rangle^{\text{eq}} &= \int dR dP dU P(R, P, U) \\ &\times \Delta U_\mu \Delta \left(\sum_\nu W(R_{\mu\nu}) \right) \end{aligned} \quad (27)$$

FIG. 11. Scatter plot of the value of instantaneous density and intrinsic energy for each particle in a given time. The density and energy have been normalized to have 0 mean and a variance of 1. The isotropy of the plot indicates that the probability factorizes $P(\rho, U) = P(\rho)P(U)$.

FIG. 12. The conditional expectation $\langle U \rangle^\rho$ of the intrinsic energy of a molecule, conditional to the value ρ of the local density of that molecule, at a global density $\rho = 0.3$. Within the large error, there is no systematic trend.

and normalize it as

$$C_{UD} = \frac{\langle \Delta \hat{U}_\mu \Delta \hat{\rho}_\mu \rangle^{\text{eq}}}{\sqrt{\langle (\Delta \hat{U}_\mu)^2 \rangle \langle (\Delta \hat{\rho}_\mu)^2 \rangle}}.$$

This correlation should vanish if positions and intrinsic energies of a molecule are statistically independent. At a monomer density $\rho = 0.3$, the average correlation computed over all particles is indeed very small $C_{UD} = 8.59 \times 10^{-4}$. We also show in Fig. 11 the scatter plot of the instantaneous density and intrinsic energy of each particle in a given time. As the plot shows no trend, we infer that these variables are statistically independent. A more stringent test in order to see that independence between intrinsic energy and local density can be obtained by looking at the following average:

$$\langle U \rangle^\rho = \frac{\int dU U P(U, \rho)}{\int dU P(U, \rho)}. \quad (28)$$

This gives the average of the intrinsic energy of a molecule conditional to the value of its local density. We plot in Fig. 12

FIG. 13. Radial distribution function $g(r)$ of the centers of mass of the molecules for the 6-6 and 6-12 star polymers, compared to the correlation $u(r)$ defined in (23).

FIG. 14. The entropy of an isolated 6-6 star polymer from Eq. (13) (solid line) compared with the entropy of a molecule inside a melt, computed from Eq. (14) (dashed lines), for different monomer densities.

the conditional expectation $\langle U \rangle^p$ as a function of the number density. Within the larger error, we do not appreciate any systematic trend, indicating that the intrinsic energy and the local density of a blob are uncorrelated.

G. Similar results for larger molecules

In order to further check the KA, we consider a few larger molecules, namely, 6-6 star polymers and 6-12 star polymers. The results obtained for these molecules are very similar to what we got for the 6-3 star polymers presented above.

Figure 13 is the same as Fig. 9, but for these larger molecules. We observe that apart from the obvious fact that the peak of the radial distribution function is located at larger distances for larger molecules, the quantity defined by (23) agrees very well with the radial distribution function g (22). Thus, the covariance $\langle U_1 U_2 \rangle^r$ defined by (19) also appears to be independent of the distance r between these larger molecules. The entropy of a molecule in the melt for these larger molecules

FIG. 15. The entropy of an isolated 6-12 star polymer from Eq. (13) (solid line) compared with the entropy of a molecule inside a melt, computed from Eq. (14) (dashed lines), for different monomer densities.

agrees also quite well with the entropy of an isolated molecule as Figs. 14 and 15 show.

IV. CONCLUSION

We have studied through MD simulations the equilibrium state of a melt of star polymer molecules in order to validate a key assumption (KA) made in a CG model of complex molecules in terms of thermal blobs. The KA for the present system assumes that the microscopic intermolecular potential of interaction between star polymer molecules is well approximated (in a statistical sense) by a function of the positions of the centers of mass of the molecules alone. This seems *a priori* a very crude approximation. However, it allows one to express the entropy of this CG level of description as a sum of the thermodynamic entropies of each molecule, as if they were isolated thermodynamic subsystems of the whole system. In this respect, the KA simply says that “entropy is additive,” in quotes because what is really meant is that the entropy of the level of description of thermal blobs can be expressed as the sum of the entropy of isolated molecules at the level of description of macroscopic thermodynamics. In addition, the entropy of the blobs is shown to depend on the *intrinsic* energy of the molecule, capturing only those contributions that depend solely on the atoms of the molecule.

The KA gives sense and usefulness to the notion of *entropy of a molecule* and makes a model based on thermal blobs practical. It was therefore important to validate this assumption through explicit MD simulations. One consequence of the KA is that the entropy of a molecule in vacuum and the entropy of a molecule in the melt are the same function of the intrinsic energy of the molecule. Quite remarkably this is what is observed in the simulations. Another consequence of the KA is that the positions, momenta, and intrinsic energies are statistically independent of each other. From a mathematical point of view, the validation of statistical independence of a large number of statistical variables is not a trivial issue. We have constructed several reasonable measures of independence between the different pairs of variables. In all cases, we observe a high degree of statistical independence, giving a rather strong support to the validity of the KA. This is quite fortunate because if the KA is not valid, the entropy of the system becomes an extremely complex function of *all* the CG variables, instead of a very simple sum of functions of just one variable. In that case, the modeling of the entropy is a rather challenging issue. We have shown that the KA is valid for the present model system of star polymer molecules. This result is quite robust with respect to the length of the branches of the star polymer that has been varied from 3 (quite compact molecules) to 12 monomers. It is also robust with respect to variations in a broad range of the monomer number density. Excluded volume interactions are relevant and compressibility effects are mainly seen in the shrinkage of the peaks of the center of mass radial distribution, towards smaller values as the monomer number density is increased. However, for the star polymers in this work there are no dramatic entanglement effects that could, perhaps, diminish the validity of the KA. Whether the KA is valid in other molecular systems requires further investigation.

The present work validating the KA is a preliminary and necessary step towards the construction of CG models of complex molecules from first principles that are able to simulate thermal processes. Note that the CG models formulated up to date are isothermal by construction. The next necessary step is to produce faithful potentials of mean force (see Refs. 27 and 28 for recent pertinent examples). To complete the derivation of the CG model, the transport coefficients need to be computed explicitly. As the intrinsic energy is non-linear in the degrees of freedom, as opposed to the position and momentum of the centers of mass, we cannot use the constrained dynamics proposed in Ref. 9 to compute the transport coefficients. In the present case, the most promising strategy seems to be a relative entropy/Bayesian approach, first introduced in Ref. 29 and implemented with variations recently in Ref. 12.

ACKNOWLEDGMENTS

We thank Aleksandar Donev for clarifying the discussions on the issue of statistical independence and Gabriel Stolz for his useful comments on the manuscript. This research is supported by the Ministerio de Economía y Competitividad of Spain (MINECO) under Grant Nos. FIS2013-47350-C5-3-R (P.E.) and FIS2013-47350-C5-1-R (R.D.B.). G. Faure was funded by the International Mobility grant of Université Paris-Est and the doctoral school MSTIC, and thanks UNED for its hospitality.

¹P.-G. de Gennes, *Scaling Concepts in Physics* (Cornell University Press, 1979).

²W. G. Noid, *J. Chem. Phys.* **139**, 090901 (2013).

³M. Green, *J. Chem. Phys.* **20**, 1281 (1952).

⁴R. Zwanzig, *Phys. Rev.* **124**, 983 (1961).

⁵K. Kawasaki and J. Gunton, *Phys. Rev. A* **8**, 2048 (1973).

⁶H. Grabert, *Projection Operator Techniques in Non-Equilibrium Statistical Mechanics* (Springer, 1982).

⁷R. L. C. Akkermans and W. J. Briels, *J. Chem. Phys.* **115**, 6210 (2001).

⁸T. Kinjo and S.-a. Hyodo, *Phys. Rev. E* **75**, 051109 (2007).

⁹C. Hijón, P. Español, E. Vanden-Eijnden, and R. Delgado-Buscalioni, *Faraday Discuss.* **144**, 301 (2010).

¹⁰Z. Li, X. Bian, B. Caswell, and G. E. Karniadakis, *Soft Matter* **10**, 8659 (2014).

¹¹S. Trément, B. Schnell, L. Petitjean, M. Couty, and B. Rousseau, *J. Chem. Phys.* **140**, 134113 (2014).

¹²A. Dequidt and J. G. S. Canchaya, *J. Chem. Phys.* **143**, 084122 (2015).

¹³P. Español and P. B. Warren, *J. Chem. Phys.* **146**, 150901 (2017).

¹⁴G. C. Ganzenmüller, S. Hiermaier, and M. O. Steinhauser, *Soft Matter* **7**, 4307 (2011).

¹⁵G. A. Kaminski, *J. Chem. Theory Comput.* **10**, 4972 (2014).

¹⁶J. B. Maillet, E. Bourasseau, N. Desbiens, G. Vallverdu, and G. Stoltz, *Europhys. Lett.* **96**, 68007 (2011).

¹⁷P. Español, *Europhys. Lett.* **40**, 631 (1997).

¹⁸J. B. Avalos and A. D. Mackie, *Europhys. Lett.* **40**, 141 (1997).

¹⁹P. Español, M. Serrano, I. Pagonabarraga, and I. Zúñiga, *Soft Matter* **12**, 4821 (2016).

²⁰X. Li, *Phys. Rev. E* **90**, 032112 (2014).

²¹L. E. Reichl, *A Modern Course in Statistical Physics* (University of Texas Press, Austin, 1980).

²²The framework of I allows one to deal with many-body potentials by assigning to each atom a part of the many-body potential. Note though that the way to perform this partition is somewhat arbitrary, and the results may be affected by different partitioning.

²³C. Jarzynski, *Phys. Rev. X* **7**, 011008 (2017).

²⁴J. Español and F. J. de la Rubia, *Physica A* **187**, 589 (1992).

²⁵M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford Graduate Texts, 2010).

²⁶G. Liu and M. Liu, *Smoothed Particle Hydrodynamics* (World Scientific, 2003).

²⁷J. G. Solano Canchaya, A. Dequidt, F. Goujon, and P. Malfreyt, *J. Chem. Phys.* **145**, 054107 (2016).

²⁸J. K. Moore, J. D. Barnes, B. C. Izvekov, S. Lísál, M. Sellers, M. S. Taylor, and D. E. Brennan, *J. Chem. Phys.* **144**, 104501 (2016).

²⁹P. Español and I. Zúñiga, *Phys. Chem. Chem. Phys.* **13**, 10538 (2011).