

Corrections to the quasiharmonic approximation for evaluating molecular entropies

Olivia L. Rojas^{a)} and Ronald M. Levy

Department of Chemistry, Rutgers University, New Brunswick, New Jersey 08903

Attila Szabo

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

(Received 29 January 1986; accepted 10 March 1986)

Molecular dynamics and Monte Carlo simulations are now being used to analyze thermodynamic problems in chemical and biochemical systems. The evaluation of entropies requires, in principle, the calculation of partition functions which are difficult to compute accurately from simulations. One approach involves the approximation of the partition function by a multivariate Gaussian with coefficients evaluated from the simulation—the so-called quasiharmonic approximation. In this paper we show that the quasiharmonic approximation is the leading term in a series expansion of the configurational distribution function in its moments. We derive the formula for the cubic correction to the quasiharmonic entropy and apply it to two model systems. The exact entropies are calculated for comparison with the approximate results. We also show how the quasiharmonic model can be used as a reference system for the evaluation of molecular entropies using importance sampling methods. Importance sampling with the quasiharmonic reference appears to be a very promising approach for evaluating the entropy of macromolecules from computer simulations.

I. INTRODUCTION

Molecular dynamics and Monte Carlo simulations are now being used to analyze thermodynamic problems in chemical and biochemical systems. Examples of this kind of analysis include the evaluation of equilibrium constants for conformational changes of peptides in solution,^{1,2} simple substitution reactions in organic chemistry,³ and the binding of ions to clathrates.⁴ While enthalpies can be calculated directly from molecular dynamics or Monte Carlo simulations, the evaluation of molecular entropies and free energies requires, in principle, the calculation of partition functions which are very difficult to compute accurately from simulations. Two solutions to this problem have been developed. One approach makes use of the fact that, although free energies are difficult to extract from simulations, free energy differences can be more readily calculated using importance sampling^{5,6} or perturbation-coupling parameter methods.^{7,8} A second approach involves the approximation of the partition function by one that is analytically tractable. Using this second approach, a method has recently been proposed for estimating the configurational entropy of macromolecules.^{9–11} The method is applicable to nondiffusive systems. The exact probability distribution is replaced by one with a Gaussian form with the coefficients chosen so that the correlations between fluctuations of coordinates evaluated from an ensemble average within the harmonic approximation to the potential have the same values as the results obtained from molecular dynamics trajectories on the complete anharmonic potential surface. At low temperatures this re-

duces to the standard harmonic approximation for which the potential is expanded in a Taylor series and truncated after the quadratic term. At higher temperatures it is possible to define a quasiharmonic Hamiltonian from which effective vibrational frequencies and normal modes can be obtained.¹² The quasiharmonic oscillator model has been shown to be particularly useful for the analysis of vibrational spectroscopic properties of molecules^{13–15} and in the development of quantum Monte Carlo simulation methods.^{16,17}

With respect to the evaluation of molecular entropies from computer simulations using the quasiharmonic approximation to the potential, two questions arise: (i) how accurate is the method, and (ii) how can more accurate approximations be developed. In this paper we show that the quasiharmonic approximation to the potential is the leading term in a certain series expansion of the configurational distribution function in its moments. We evaluate the cubic correction to the quasiharmonic entropy for two model systems; the exact entropy is then calculated for comparison with the approximate results. We also show how the quasiharmonic model can be used as a reference system for the more accurate calculation of molecular entropies using importance sampling methods. The theory is considered in the following section. Illustrative applications are presented in Sec. III. The results are summarized in Sec. IV. In Appendix A we develop an explicit expression for the cubic correction to the quasiharmonic entropy and in Appendix B we discuss the use of normal modes in the computation of the molecular entropy from computer simulations.

II. THEORY

The classical partition function for a polyatomic molecule with N atoms can be written as

^{a)} Present address: Ecole Polytechnique—ENSTA, INSERM U 275, Batterie de L'Yvette, 91120 Palaiseau, France.

$$Z = (1/h)^{3N} \prod_{i=1}^N (2\pi m_i k_B T)^{3/2} Z_r^c, \quad (2.1a)$$

$$Z_r^c = \int e^{-\beta V(\mathbf{r})} d\mathbf{r}, \quad (2.1b)$$

where $\beta = 1/k_B T$, the term containing the masses comes from the evaluation of the momentum integrals, and Z_r^c is the configurational partition function in Cartesian coordinates. For the development of the configurational partition function, Z_r^c in terms of moments of the configurational distribution it is useful for both computational and physical reasons to express the configurational partition function as an integral over internal coordinates. The computational reason for this transformation is the need to separate out translations and rotations from the fluctuations of coordinates about mean values. This is accomplished naturally by transforming to internal coordinates, although the calculations could also be performed in Cartesian coordinates using a series of Eckart transformations¹⁸ to eliminate the translations and rotations. The use of internal coordinates also provides better physical insight into the molecular origins of the thermodynamic effects since it is possible to determine the contributions of particular molecular degrees of freedom to the entropy. Transforming to internal coordinates, the configurational partition function can be expressed as

$$Z_r^c = 8\pi^2 V \int e^{-\beta V(\mathbf{q})} g^{1/2}(\mathbf{q}) d\mathbf{q}, \quad (2.2a)$$

$$\simeq 8\pi^2 V g^{1/2}(\mathbf{q}_0) Z_q^c, \quad (2.2b)$$

$$Z_q^c = \int e^{-\beta V(\mathbf{q})} d\mathbf{q}. \quad (2.2c)$$

The factor $g^{1/2}(\mathbf{q})$ is the Jacobian for transforming from Cartesian to internal coordinates. Since $g(\mathbf{q})$ depends only on bond lengths and bond angles,^{19,20} the change in $g(\mathbf{q})$ with conformation is expected to be small and removing it from the integral [Eq. (2.2b)] is a good approximation. Although it is difficult to evaluate Z_q^c exactly, a very useful simplification results when Z_q^c is approximated by a multidimensional Gaussian:

$$Z_q^c \simeq Z_q^{\text{QH}} = \int e^{-\beta(\mathbf{q}^\dagger \cdot \mathbf{F} \cdot \mathbf{q})/2} d\mathbf{q}. \quad (2.3)$$

The approximate partition function Z_q^{QH} can be evaluated analytically. In the quasiharmonic approximation the elements of \mathbf{F} are constructed from the correlation matrix of internal coordinate fluctuations^{9,12}

$$\sigma_{ij} = \langle \Delta q_i \Delta q_j \rangle, \quad (2.4a)$$

$$\Delta q_i = q_i - \langle q_i \rangle, \quad (2.4b)$$

where the averages $\langle \Delta q_i \Delta q_j \rangle$ are readily calculated from a molecular dynamics or Monte Carlo simulation. The elements of \mathbf{F} are given by^{9,12}

$$\mathbf{F} = k_B T \boldsymbol{\sigma}^{-1}. \quad (2.5)$$

The configurational probability distribution is

$$P^{\text{(QH)}}(\mathbf{q}) = (1/2\pi)^{n/2} (\det \boldsymbol{\sigma})^{-1/2} \times \exp\left\{-\frac{1}{2} \Delta \mathbf{q}^\dagger \cdot \boldsymbol{\sigma}^{-1} \cdot \Delta \mathbf{q}\right\}, \quad (2.6)$$

where n is the number of degrees of freedom. The quasiharmonic approximation is self-consistent in the sense that the

ensemble averages of internal coordinate fluctuations obtained from Z_q^{QH} are equal to the values obtained with the complete potential and the molecular dynamics or Monte Carlo trajectory. Using the relation

$$S = -k_B \int P(\mathbf{q}) \ln P(\mathbf{q}) d\mathbf{q}, \quad (2.7)$$

the configurational entropy in the quasiharmonic approximation is given by

$$S_{\text{QH}}^c = \frac{1}{2} n k_B + \frac{1}{2} k_B \ln[(2\pi)^n \det \boldsymbol{\sigma}]. \quad (2.8)$$

We wish to develop a procedure for incorporating self-consistently, higher order correlations (moments) of the coordinate fluctuations in the approximate expressions for the configurational probability distribution and entropy. Our approach, which is in the spirit of the quasiharmonic model, is to express the configurational distribution function as a Gaussian times a power series correction in the positional moments. Using $P^{\text{QH}}(\mathbf{q})$ as the reference Gaussian distribution

$$P(\mathbf{q}) = P^{\text{QH}}(\mathbf{q}) \left[1 + \sum_k a_k \Delta q_k + \sum_{kl} b_{kl} \Delta q_k \Delta q_l + \sum_{klm} c_{klm} \Delta q_k \Delta q_l \Delta q_m + \dots \right] \quad (2.9)$$

in analogy to the well known Gram-Charlier series expansion,²¹ the coefficients $\{a_k\}$, $\{b_{kl}\}$, etc. are defined so that when the series is truncated at n th order, the moments of the truncated series up to n th order are equal to the corresponding moments of the exact probability distribution. Clearly, the quasiharmonic probability distribution is equivalent to a second-order series expansion with coefficients $\{a_k\}$ and $\{b_{kl}\}$ zero. In Appendix A we derive the third-order coefficients $\{c_{klm}\}$. The coefficients are shown to be functions of the first three moments $\langle q_i \rangle$, $\langle \Delta q_i \Delta q_j \rangle$, and $\langle \Delta q_i \Delta q_j \Delta q_k \rangle$, of the exact configurational distribution which are evaluated from the computer simulation. We also show in Appendix A that for the configurational probability distribution of Eq. (2.9) including terms up to third order, the configurational entropy is given by the following expression:

$$S_3^c = S_{\text{QH}}^c - \frac{k_B}{12} \sum_{\alpha\beta\gamma} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle \times (\boldsymbol{\sigma}^{-1})_{\alpha\alpha} (\boldsymbol{\sigma}^{-1})_{\beta\beta} (\boldsymbol{\sigma}^{-1})_{\gamma\gamma} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle. \quad (2.10)$$

We note that the cubic correction always lowers the quasiharmonic entropy as is to be expected since the quasiharmonic entropy is an upper bound to the exact entropy. This follows from the well known theorem (see, e.g., Ref. 22) that of all distributions with the same second moment, the Gaussian has the maximum entropy. In the following section we use Eq. (2.10) to evaluate the entropy for two model potentials and compare the results with the quasiharmonic and harmonic approximations to the entropy.

The series expansion for the configurational entropy developed in Appendix A is tractable for calculating low-order corrections to the quasiharmonic entropy. An alternative and very promising approach to the evaluation of the molecular entropy for nondiffusive systems is to use the quasiharmonic model as a reference system to calculate the exact partition function using importance sampling. We note that

the quasiharmonic model has been demonstrated to be an excellent reference system for the Monte Carlo evaluation of path integrals via importance sampling.¹⁶ Consider the ratio of the exact to the quasiharmonic partition functions

$$\frac{Z^c}{Z_{\text{QH}}^c} = \frac{\int e^{-\beta V} d\mathbf{q}}{\int e^{-\beta V^{\text{QH}}} d\mathbf{q}} = \frac{\int e^{-\beta V} e^{+\beta V^{\text{QH}}} e^{-\beta V^{\text{QH}}} d\mathbf{q}}{\int e^{-\beta V^{\text{QH}}} d\mathbf{q}} = \langle e^{-\beta(V-V^{\text{QH}})} \rangle_{\text{QH}}, \quad (2.11)$$

where the brackets imply a Monte Carlo average with respect to $e^{-\beta V^{\text{QH}}}$ where $\beta V^{\text{QH}} = \frac{1}{2} \Delta \mathbf{q}^\dagger \boldsymbol{\sigma}^{-1} \Delta \mathbf{q}$. Thus, the ratio of the exact to the quasiharmonic partition functions can be evaluated by calculating the average of $e^{-\beta(V-V^{\text{QH}})}$ over the quasiharmonic surface. This is just the standard method of evaluating the partition function via importance sampling and is numerically stable because the fluctuations in $e^{-\beta(V-V^{\text{QH}})}$ are much smaller than those in $e^{-\beta V}$ in the important regions of conformation space. An even more efficient algorithm can be obtained by inverting Eq. (2.11):

$$\frac{Z_{\text{QH}}^c}{Z^c} = \frac{\int e^{-\beta V^{\text{QH}}} d\mathbf{q}}{\int e^{-\beta V} d\mathbf{q}} = \frac{\int e^{-\beta V^{\text{QH}}} e^{+\beta V} e^{-\beta V} d\mathbf{q}}{\int e^{-\beta V} d\mathbf{q}} = \langle e^{+\beta(V-V^{\text{QH}})} \rangle, \quad (2.12)$$

where now the average is evaluated using simulations performed on the exact potential surface. This second approach has the advantage that only a single simulation need be performed. The entropy difference between the exact and quasiharmonic values is given by

$$S^c - S_{\text{QH}}^c = -k_B \ln \langle e^{\beta(V-V^{\text{QH}})} \rangle + \frac{\langle V \rangle}{T} - \frac{nk_B}{2}. \quad (2.13)$$

We now consider corrections to the quasiharmonic entropy for two model potentials.

III. ILLUSTRATIVE APPLICATIONS

As a first application of the expression derived in Appendix A for the cubic correction to the quasiharmonic entropy we consider the one-dimensional system described by the potential

$$V(x) = e^{-x} + x. \quad (3.1)$$

This potential provides one of the simplest nontrivial examples for which the configurational partition function as well as all moments can be calculated analytically. The exact configurational partition function and second and third moments are given by²³

$$Z = \int_{-\infty}^{\infty} e^{-\beta V(x)} dx = \Gamma(\beta) / \beta^\beta, \quad (3.2a)$$

$$\langle (\Delta x)^2 \rangle = d\psi/d\beta = \psi'(\beta), \quad (3.2b)$$

$$\langle (\Delta x)^3 \rangle = -d^2\psi/d\beta^2 = -\psi''(\beta), \quad (3.2c)$$

where $\Gamma(x)$ and $\psi(x)$ are the gamma and digamma func-

tions, respectively. Using Eqs. (2.7), (2.8), and (2.10), the exact, harmonic, and quasiharmonic entropies and the cubic correction to the quasiharmonic entropy for this potential are given by the following expressions:

$$S^c/k_B = \beta [1 - \psi(\beta)] + \ln \Gamma(\beta), \quad (3.3a)$$

$$S_H^c/k_B = -\frac{1}{2} \ln \beta + \frac{1}{2} \ln 2\pi + \frac{1}{2}, \quad (3.3b)$$

$$S_{\text{QH}}^c/k_B = \frac{1}{2} \ln \psi'(\beta) + \frac{1}{2} \ln 2\pi + \frac{1}{2}, \quad (3.3c)$$

$$S_3^c/k_B = S_{\text{QH}}^c/k_B - \psi''(\beta)^2 / [12\psi'(\beta)^3]. \quad (3.3d)$$

In Table I, we compare the harmonic and quasiharmonic (with and without the cubic correction) entropies with the exact configurational entropy for three temperatures. It can be seen that the approximations improve as the temperature is lowered as expected. At each temperature the harmonic approximation is less accurate than the quasiharmonic estimate of the entropy, while the cubic correction further improves the result. For the highest temperature considered ($\beta = \frac{1}{2}$), the harmonic entropy underestimates the exact result by 14% while the quasiharmonic entropy overestimates it by 8%. The inclusion of the cubic correction reduces the error to 1.6%.

The second example we consider is a small test molecule, butane. The empirical potential for butane contains all the terms present in the molecular mechanics potentials for macromolecules, viz. bond stretching, bending, and torsional terms as well as nonbonded interactions. The butane model has been used extensively for computer simulation studies of both structural and dynamic processes.²⁴⁻²⁶ A harmonic analysis of entropy differences for this butane model was reported by Karplus and Kushick⁹ and a quasiharmonic analysis of the vibrational spectrum has also been carried out.¹³

The form of the empirical potential used for the Monte Carlo simulations of butane is

$$V = \frac{1}{2} \sum_{\text{bonds}} K_{b_i} (b_i - b_{i0})^2 + \frac{1}{2} \sum_{\text{angles}} K_{\theta_i} (\theta_i - \theta_{i0})^2 + \frac{1}{2} K_\phi [1 + \cos(3\phi)] + \frac{1}{2} \left[\frac{A}{r^{12}} - \frac{C}{r^6} \right], \quad (3.4)$$

where the b_{i0} and θ_{i0} are the equilibrium values for the isolated bond and bond angle coordinates and the K_{b_i} , K_{θ_i} , and K_ϕ are force constants. The last term of Eq. (3.4) corresponds to the van der Waals interaction between the end atoms. The parameters used in the present study are the same as those used in an earlier harmonic analysis for butane.¹³ The meth-

TABLE I. Configurational entropies for the potential $V(x) = e^{-x} + x$.

β S/k_B	0.5	1.0	2.0
Harmonic	1.766	1.419	1.072
Quasiharmonic	2.217	1.668	1.200
QH + cubic correction	2.021	1.560	1.149
Exact	2.054	1.577	1.154

yl and methylene groups of the butane molecule were treated as extended atoms so that the model contained four interaction centers. The conformation of the molecule is therefore determined by the 12 Cartesian coordinates or 6 internal coordinates (three bonds, two angles, and one torsional coordinate).

For comparison with the quasiharmonic and importance sampling results, the harmonic approximations to the entropy of *trans* and *gauche* butane were first calculated. The harmonic force constants in internal coordinates were used to calculate the correlation matrix [Eqs. (2.4) and (2.5)] and entropy in the harmonic approximation. Monte Carlo simulations were performed using Metropolis sampling with normal mode eigenvectors as the independent coordinates for sampling as described previously.¹³ *Trans* conformations were those for which the dihedral angle ϕ was restricted to vary between 120° and 240° while for *gauche* conformations the dihedral angle was restricted to vary between 0° and 120°. These restrictions were incorporated into the Monte Carlo calculations by simply increasing the energy by 100 kcal/mol whenever a trial step yielded a dihedral angle outside the restricted ranges. The Monte Carlo trajectory that was used for further analysis consisted of 75 000 points for each of the conformers at an average temperature of 301 K.

The trajectories were used to calculate the second and third moments of the configurational probability distributions in internal coordinates. The quasiharmonic configurational entropies were calculated using Eq. (2.8) and the cubic correction using Eq. (2.10). The results are summarized in Table II. We note that the quasiharmonic entropy is larger than the harmonic result by 0.529 cal/mol K for *trans* butane and by 0.773 cal/mol K for *gauche* butane. The entropy difference between *trans* and *gauche* conformers is 0.377 cal/mol K within the harmonic approximation while it is only 0.133 cal/mol K in the quasiharmonic approximation. The cubic correction is essentially zero for the *trans* conformer while it is 0.105 cal/mol K for the *gauche* conformer. The *trans-gauche* entropy difference including the third moment corrections is 0.236 cal/mol K. Thus, the entropy dif-

ference for this model potential is quite sensitive to the level of approximation used. One can readily understand why the cubic correction is so small for the *trans* conformer as follows. All the anharmonicity in the potential [Eq. (3.4)] arises from the torsional and Van der Waals contributions. For the *trans* conformer the potential energy is symmetric about $\phi = \pi$, i.e., there are no cubic terms proportional to $(\phi - \pi)^3$. As a result of this, the cubic moments are expected to be very small. For the *gauche* conformer, although there are still no cubic terms arising from the torsional part of the potential, the van der Waals contribution is large enough to make the potential asymmetric about $\phi = \pi/3$.

We now consider the use of importance sampling to estimate the exact entropy for the model butane potential. We first use the Monte Carlo trajectory run on the complete butane potential to calculate the quasiharmonic potential and entropy and then use the quasiharmonic potential as a reference to calculate the difference between the "exact" and quasiharmonic entropies [Eq. (2.13)]. An attractive feature of this approach is that only one simulation is required. Once the quasiharmonic potential is found, ΔS can be obtained in a computationally simple way. Thus it is possible to obtain an estimate of the accuracy of the quasiharmonic entropies and entropy differences between different states of the system with minimal computational effort.

The configurational *trans-gauche* entropy difference calculated by importance sampling is 0.220 cal/mol K. The results, Table II, indicate that the exact configurational entropy difference between the *trans* and *gauche* conformers is overestimated by the harmonic and underestimated by the quasiharmonic approximations with the exact entropy difference being close to the average value. For this model potential the addition of a cubic correction of the quasiharmonic entropy significantly improves the accuracy of the estimate of ΔS ; ΔS^c is within 7% of the exact value. It should be noted that the entropy difference between *trans* and *gauche* butane is only 1.5% of the absolute value of the entropy for the individual conformers. This model problem therefore constitutes a rather severe test of the quasiharmonic approximation in that small differences between large numbers need to be calculated with high precision. If the estimates of the entropy for the individual conformers are compared in Table II it can be seen that all the approximations are within a few percent of the absolute values.

TABLE II. Internal coordinate configurational entropies^a of *trans* and *gauche* butane at 301 K.

	<i>Trans</i>	<i>Gauche</i>	<i>Trans-Gauche</i>
S_H^c ^b	-17.748	-18.125	0.377
S_{QH}^c ^c	-17.219	-17.352	0.133
S_3^c ^d	-17.221	-17.457	0.236
S^c ^e	-17.320	-17.527	0.207
S^{cf} ^f	-17.290	-17.510	0.220

^aThe values of S^c in this table depend on the units of the internal coordinates; however, all ΔS^c 's are in units of cal/mol K.

^bHarmonic entropy.

^cQuasiharmonic entropy.

^dQuasiharmonic entropy + cubic correction.

^eEntropy obtained from numerical integration of Eq. (B8).

^fEntropy obtained using importance sampling with the quasiharmonic reference [Eq. (2.12)].

IV. CONCLUSIONS

There is considerable interest in using molecular mechanics to model the binding of ligands to proteins and nucleic acids and to study problems associated with drug design. Central to this effort is the development of approximate methods for calculating thermodynamic properties. It has been demonstrated that at room temperature, anharmonicity can have a substantial effect on the thermodynamics of biopolymers.¹² The quasiharmonic approximation utilizes molecular dynamics or Monte Carlo simulations for partially incorporating anharmonic effects in the statistical thermodynamic properties. We have derived an expression for incorporating corrections to the quasiharmonic entropy of molecules in terms of higher-order moments of the configu-

rational distribution function which may be calculated from simulations. The cubic correction term has been evaluated explicitly for two model systems and has been shown to significantly improve the accuracy of the quasiharmonic approximation. One important goal is to develop computational methods suitable for reliably estimating the entropy differences associated with conformational changes of macromolecules. The moment method provides a way to systematically improve the quasiharmonic approximation to the entropy. The weakness of the method is that a large number of higher moments must be calculated and it may be difficult to evaluate these with great accuracy from limited macromolecular simulations. On physical grounds it should be possible to ignore certain higher-order correlations, e.g., between internal coordinates which are separated by large distances. It appears that a very promising approach for estimating the entropy of macromolecules is to use importance sampling with the quasiharmonic reference.

ACKNOWLEDGMENTS

This work has been supported by a grant from the National Institutes of Health (GM-30580). RML is an Alfred P. Sloan Fellow and the recipient of a N. I. H. Research Career Development Award and thanks the members of the Laboratoire du Cristallographie Biologique, I. B. M. C., C. N. R. S., Strasbourg, France for their hospitality during the writing of this manuscript.

APPENDIX A: DERIVATION OF THE CUBIC CORRECTION TO THE QUASIHARMONIC ENTROPY

In this appendix we derive an expression [Eq. (2.10)] for the leading correction to the quasiharmonic entropy in terms of the cubic moments $\langle \Delta q_i \Delta q_j \Delta q_k \rangle$ obtained from a simulation. The approach we use can be readily generalized to yield higher order corrections involving higher-order moments. In analogy to the Gram-Charlier series, we consider an expansion of the exact probability distribution of the form

$$P(\mathbf{q}) = P^{\text{QH}}(\mathbf{q}) \left[1 + \sum_k a_k \Delta q_k + \sum_{kl} b_{kl} \Delta q_k \Delta q_l + \sum_{klm} c_{klm} \Delta q_k \Delta q_l \Delta q_m + \dots \right], \quad (\text{A1})$$

where $\Delta q_k = q_k - \langle q_k \rangle$ and where $P^{\text{QH}}(\mathbf{q})$ is the quasiharmonic configurational distribution given in Eq. (2.6). To obtain the leading correction to the quasiharmonic entropy we truncate the expansion [Eq. (A1)] incorporating only the terms up to third order:

$$P^3(\mathbf{q}) = P^{\text{QH}}(\mathbf{q}) \left[1 + \sum_k a_k \Delta q_k + \sum_{kl} b_{kl} \Delta q_k \Delta q_l + \sum_{klm} c_{klm} \Delta q_k \Delta q_l \Delta q_m \right]. \quad (\text{A2})$$

The conditions used to determine the coefficients a_k , b_{kl} , and c_{klm} in the expansion (A2) are

(1) The first moment values $\langle \Delta q_i \rangle$ computed with the truncated probability function [Eq. (A2)] must be zero for all the internal coordinates q :

$$\langle \Delta q_\alpha \rangle = \langle q_\alpha - \langle q_\alpha \rangle \rangle = 0, \quad \alpha = 1, 2, \dots, 3N - 6. \quad (\text{A3})$$

(2) The computed third moments $\langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle$ must be equal to the values obtained from the numerical simulation.

(3) Finally, the second moments $\langle \Delta q_\alpha \Delta q_\beta \rangle$ computed with the quasiharmonic probability function [Eq. (2.6)] and the probability function including third-order corrections [Eq. (A2)] must be equal to each other and to the values obtained from the numerical simulation

$$\begin{aligned} \langle \Delta q_\alpha \Delta q_\beta \rangle &= \langle \Delta q_\alpha \Delta q_\beta \rangle_{\text{QH}} \\ &= \langle \Delta q_\alpha \Delta q_\beta \rangle_{P^3}, \quad \alpha, \beta = 1, 2, \dots, 3N - 6. \end{aligned} \quad (\text{A4})$$

This third condition simplifies Eq. (A2), since the coefficients b_{kl} must be zero. We have then, to deal with a simpler expression for $P(\mathbf{q})$ with only two types of coefficients:

$$P^3(\mathbf{q}) = P^{\text{QH}}(\mathbf{q}) \left[1 + \sum_k a_k \Delta q_k + \sum_{klm} c_{klm} \Delta q_k \Delta q_l \Delta q_m \right]. \quad (\text{A5})$$

In order to determine the coefficients a_k and c_{klm} we need two sets of equations which can be obtained using conditions (1) and (2). These equations are

$$\begin{aligned} \langle \Delta q_\alpha \rangle &= 0 = \sum_k a_k \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \\ &+ \sum_{klm} c_{klm} \langle \Delta q_\alpha \Delta q_k \Delta q_l \Delta q_m \rangle_{\text{QH}} \end{aligned} \quad (\text{A6})$$

and

$$\begin{aligned} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle &= \sum_k a_k \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \Delta q_k \rangle_{\text{QH}} \\ &+ \sum_{klm} c_{klm} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \Delta q_k \Delta q_l \Delta q_m \rangle_{\text{QH}}, \end{aligned} \quad (\text{A7})$$

where the subscript on the moment values refer to the moments obtained with the quasiharmonic probability function. The fourth and sixth moments obtained with the quasiharmonic probability function can be expressed in terms of the second moments. For example,

$$\begin{aligned} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \Delta q_l \rangle_{\text{QH}} &= \langle \Delta q_\alpha \Delta q_\beta \rangle_{\text{QH}} \langle \Delta q_\gamma \Delta q_l \rangle_{\text{QH}} \\ &+ \langle \Delta q_\alpha \Delta q_\gamma \rangle_{\text{QH}} \langle \Delta q_\beta \Delta q_l \rangle_{\text{QH}} \\ &+ \langle \Delta q_\alpha \Delta q_l \rangle_{\text{QH}} \langle \Delta q_\beta \Delta q_\gamma \rangle_{\text{QH}}. \end{aligned} \quad (\text{A8})$$

Using Eq. (A8) and the analogous relation for the sixth moments and the symmetry properties of the c_{klm} and the second moments, Eqs. (A6) and (A7) become

$$\begin{aligned} \sum_k a_k \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \\ + 3 \sum_{klm} c_{klm} \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \langle \Delta q_l \Delta q_m \rangle_{\text{QH}} = 0, \end{aligned} \quad (\text{A9})$$

$$\begin{aligned}
\langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle &= 3 \sum_k a_k \langle \Delta q_\beta \Delta q_\gamma \rangle_{\text{QH}} \cdot \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \\
&+ 6 \sum_{klm} c_{klm} \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \langle \Delta q_\beta \Delta q_l \rangle_{\text{QH}} \langle \Delta q_\gamma \Delta q_m \rangle_{\text{QH}} \\
&+ 9 \sum_{klm} c_{klm} \langle \Delta q_\beta \Delta q_\gamma \rangle_{\text{QH}} \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \langle \Delta q_l \Delta q_m \rangle_{\text{QH}}.
\end{aligned} \tag{A10}$$

Multiplying Eq. (A9) by $3 \langle \Delta q_\beta \Delta q_\gamma \rangle_{\text{QH}}$ and subtracting the result from Eq. (A10), we have

$$\begin{aligned}
\langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle &= 6 \sum_{klm} c_{klm} \langle \Delta q_\alpha \Delta q_k \rangle_{\text{QH}} \langle \Delta q_\beta \Delta q_l \rangle_{\text{QH}} \\
&\times \langle \Delta q_\gamma \Delta q_m \rangle_{\text{QH}}.
\end{aligned} \tag{A11}$$

By substituting the probability function given by Eq. (A1) into the equation which defines the configurational entropy Eq. (2.7), and using the approximation $\ln(1+y) \simeq y - (y^2/2)$ for small values of y , we obtain the expression for the configurational entropy including third-order corrections:

$$\begin{aligned}
S_3^c &= S_{\text{QH}}^c - \frac{k_B}{2} \left[\sum_{\alpha} a_\alpha a_\alpha \langle \Delta q_\alpha \Delta q_l \rangle_{\text{QH}} \right. \\
&+ 2 \sum_{\alpha klm} a_\alpha c_{klm} \langle \Delta q_\alpha \Delta q_k \Delta q_l \Delta q_m \rangle_{\text{QH}} \\
&\left. + \sum_{\substack{klm \\ \alpha \beta \gamma}} c_{klm} c_{\alpha \beta \gamma} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \Delta q_k \Delta q_l \Delta q_m \rangle_{\text{QH}} \right],
\end{aligned} \tag{A12}$$

where S_{QH}^c refers to the configurational entropy computed with the quasiharmonic probability function. Using Eqs. (A6) and (A7) to simplify Eq. (A12), we find

$$S_3^c = S_{\text{QH}}^c - \frac{k_B}{2} \sum_{klm} c_{klm} \langle \Delta q_k \Delta q_l \Delta q_m \rangle. \tag{A13}$$

Finally solving Eq. (A11) for c_{klm} and substituting the result into Eq. (A13) we have

$$\begin{aligned}
S_3^c &= S_{\text{QH}}^c - \frac{k_B}{12} \sum_{\substack{klm \\ \alpha \beta \gamma}} \langle \Delta q_k \Delta q_l \Delta q_m \rangle \\
&\times (\sigma^{-1})_{k\alpha} (\sigma^{-1})_{l\beta} (\sigma^{-1})_{m\gamma} \langle \Delta q_\alpha \Delta q_\beta \Delta q_\gamma \rangle.
\end{aligned} \tag{A14}$$

APPENDIX B: QUASIHARMONIC NORMAL MODES

In this Appendix we discuss the use of quasiharmonic normal modes in the evaluation of the quasiharmonic entropy and in developing corrections to this approximation. The quasiharmonic approximation to the potential which we have used to evaluate the entropy [Eqs. (2.5) and (2.8)] can also be used to define a quasiharmonic Hamiltonian from which it is possible to obtain normal mode eigenvectors and frequencies.^{12,13} The frequencies are temperature dependent and the normal modes are rotated and shifted with respect to their harmonic values. In fact, the configurational entropy in the quasiharmonic approximation can be simply expressed in terms of the quasiharmonic frequencies. Using

Eqs. (2.2), (2.3), and (2.5) the configurational partition function in Cartesian coordinates is given by

$$Z_r^c = 8\pi^2 V g^{1/2}(\mathbf{q}_0) (2\pi)^{(3N-6)/2} (\det \boldsymbol{\sigma})^{1/2}. \tag{B1}$$

The Jacobian $g^{1/2}$ is given by^{19,20}

$$g^{1/2}(\mathbf{q}_0) = \det[\mathbf{I}(q_0) \cdot \mathbf{G}^{-1}(q_0)]^{1/2} \left(\sum_i^N m_i / \prod_i^N m_i \right)^{3/2}, \tag{B2}$$

where $\mathbf{I}(q_0)$ and $\mathbf{G}(q_0)$ are the moment of inertia tensor and kinetic energy matrix for the molecule evaluated at the equilibrium conformation q_0 . Substituting Eq. (B2) into Eq. (B1) and using the fact that $\boldsymbol{\sigma}$ is proportional to \mathbf{F}^{-1} we have

$$\begin{aligned}
Z_r^c &= 8\pi^2 V (2\pi k_B T)^{(3N-6)/2} \left(\sum_i^N m_i / \prod_i^N m_i \right)^{3/2} \\
&\times \det \mathbf{I}^{1/2}(q_0) \cdot \det[\mathbf{G}^{-1}(q_0) \cdot \mathbf{F}^{-1}]^{1/2} \\
&= 8\pi^2 V (2\pi k_B T)^{(3N-6)/2} \left(\sum_i^N m_i / \prod_i^N m_i \right)^{3/2} \\
&\times \det \mathbf{I}^{1/2}(q_0) \cdot \det \boldsymbol{\Lambda}^{-1},
\end{aligned} \tag{B3}$$

where $\boldsymbol{\Lambda}^{-1}$ is the diagonal matrix of quasiharmonic normal mode frequencies. Thus the configurational partition function is proportional to the product of a rotational partition function $\det \mathbf{I}^{1/2}(q)$ times a vibrational partition function $\det \boldsymbol{\Lambda}^{-1}$. The quasiharmonic normal modes \mathcal{Q}^{QH} have the property that the ensemble average of the product of any two different quasiharmonic modes is zero:

$$\langle \mathcal{Q}_i^{\text{QH}} \mathcal{Q}_j^{\text{QH}} \rangle = \delta_{ij} \frac{k_B T}{\lambda_i^2}. \tag{B4}$$

Note this ensemble average orthogonal property does not hold for the harmonic normal modes. The quasiharmonic eigenvectors represent a best set of linearly independent basis coordinates which are uncoupled up to second order in an ensemble average sense. The orthogonal property (B4) can be exploited in developing corrections to the quasiharmonic entropy. For example, the sum over six indices required for the third order correction to the quasiharmonic entropy can be reduced to a sum over three indices, greatly simplifying the computational effort, since Eq. (A14) can be rewritten as

$$S_3^c = S_{\text{QH}}^c - \frac{k_B}{12} \sum_{ijk} \langle \mathcal{Q}_i^{\text{QH}} \mathcal{Q}_j^{\text{QH}} \mathcal{Q}_k^{\text{QH}} \rangle^2 \frac{\lambda_i^2 \lambda_j^2 \lambda_k^2}{(k_B T)^3}, \tag{B5}$$

where we have used the fact that the matrix of quasiharmonic normal modes diagonalizes $\boldsymbol{\sigma}^{-1}$.

An alternative approach to improving the quasiharmonic approximation to the entropy involves estimating the total configurational entropy as a sum of uncorrelated or diagonal non-Gaussian contributions plus the entropy due to correlations between coordinates which is estimated in the

Gaussian approximation. Edholm and Berendsen¹⁰ have used this correction procedure with the diagonal non-Gaussian contribution to the entropy calculated from histograms of the one-dimensional spatial probability distribution generated by molecular dynamics trajectories. We have used a similar separation to estimate the configurational entropy of *trans* and *gauche* butane for comparison with our quasiharmonic importance sampling results. This approximation for estimating the entropy should be accurate for the butane model because almost all of the anharmonicity for this potential is associated with the torsional coordinate. A quasiharmonic normal mode analysis for this model¹³ has demonstrated that at temperatures up to at least 300 K, the torsional coordinate couples only harmonically to the other coordinates and that one of the six quasiharmonic normal modes at 300 K is a pure torsion. Therefore a good approximation to the entropy can be obtained by constructing and numerically integrating over the one dimensional torsional probability distribution and integrating over all the other degrees of freedom analytically in the quadratic approximation. Expanding V in a Taylor series involving q_i , $i = 1, \dots, 5$ ($\phi = q_6$) we have

$$V \simeq V(\phi, \{q_i^0\}) + \sum_{i=1}^5 \left(\frac{\partial V}{\partial q_i} \right)_0 (q_i - q_i^0) + \frac{1}{2} \sum_{ij=1}^5 \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 (q_i - q_i^0)(q_j - q_j^0), \quad (\text{B6})$$

where $V(\phi, \{q_i^0\})$ is the potential energy calculated using Eq. (3.4) with all coordinates except ϕ set equal to their equilibrium values. We now further expand $(\partial V / \partial q_i)$ and $(\partial^2 V / \partial q_i \partial q_j)$ about $\phi = \phi_0$ and hence we find

$$V_A \simeq V(\phi, \{q_i^0\}) + \sum_{i=1}^5 F_{i6}^H (\phi - \phi_0)(q_i - q_i^0) + \frac{1}{2} \sum_{ij=1}^5 F_{ij}^H (q_i - q_i^0)(q_j - q_j^0), \quad (\text{B7})$$

since $(\partial V / \partial q_i)_0 = 0$ and the second derivatives are simply the harmonic force constants. The configurational partition function corresponding to the approximate potential in Eq. (B7) is

$$Z_A^c = \int \{dq_i\} \int d\phi e^{-\beta V_A} = (2\pi/\beta)^{5/2} (\det \mathbf{F}_5^H)^{-1/2} \int d\phi e^{-\beta V_{\text{eff}}(\phi)}, \quad (\text{B8})$$

$$V_{\text{eff}}(\phi) = V(\phi, \{q_i^0\}) - \frac{1}{2} F_{66}^H (\phi - \phi_0)^2 + \frac{1}{2} \frac{\det \mathbf{F}^H}{\det \mathbf{F}_5^H} (\phi - \phi_0)^2, \quad (\text{B9})$$

where \mathbf{F}^H is the full 6×6 harmonic force constant matrix, \mathbf{F}_5^H is the 5×5 harmonic force constant matrix involving all internal coordinates except ϕ , and F_{66}^H is the diagonal harmonic force constant matrix element corresponding to ϕ . The integral over ϕ was performed numerically for both the *trans* and *gauche* conformers. The resulting entropy is listed in Table II. The entropy difference estimated by this separation procedure is 0.207 cal/mol K which is close to the value estimated by importance sampling (0.220 cal/mol K).

- ¹J. Brady and M. Karplus, *J. Am. Chem. Soc.* **107**, 6103 (1985).
- ²M. Mezei, P. K. Mehrota, and D. L. Beveridge, *J. Am. Chem. Soc.* **107**, 2239 (1985).
- ³J. Chandrasekhar, F. S. Smith, and W. L. Jorgensen *J. Am. Chem. Soc.* **106**, 3049 (1984).
- ⁴T. P. Lybrand, J. A. McCammon, and G. Wipff, *Proc. Natl. Acad. Sci. U. S. A.* **83**, 833 (1986).
- ⁵J. P. Valleau and G. M. Torrie, in *Statistical Mechanics Part A*, edited by B. J. Berne (Plenum, New York, 1977), p. 169.
- ⁶C. Pangali, M. Rao, and B. J. Berne, *J. Chem. Phys.* **71**, 2975 (1979).
- ⁷J. M. Torrie and J. P. Valleau, *Chem. Phys. Lett.* **28**, 578 (1974).
- ⁸J. A. Barker and D. Henderson, *Rev. Mod. Phys.* **48**, 578 (1974).
- ⁹M. Karplus and J. N. Kushick, *Macromolecules* **14**, 325 (1981).
- ¹⁰O. Edholm and H. J. C. Berendsen, *Mol. Phys.* **51**, 1011 (1984).
- ¹¹A. DiNola, H. J. C. Berendsen, and O. Edholm, *Macromolecules* **17**, 2044 (1984).
- ¹²R. M. Levy, M. Karplus, J. N. Kushick, and D. Perahia, *Macromolecules* **17**, 1370 (1984).
- ¹³R. M. Levy, O. L. Rojas, and R. A. Friesner, *J. Phys. Chem.* **88**, 4233 (1984).
- ¹⁴R. M. Levy, A. R. Srinivasan, W. K. Olson, and J. A. McCammon, *Biopolymers* **23**, 1099 (1984).
- ¹⁵S. Nakai, F. Hirata, and R. M. Levy (to be published).
- ¹⁶R. A. Friesner and R. M. Levy, *J. Chem. Phys.* **80**, 4488 (1984).
- ¹⁷R. M. Levy, P. Zhang, and R. A. Friesner (to be published).
- ¹⁸C. Eckart, *Phys. Rev.* **47**, 552 (1935).
- ¹⁹N. Go and H. A. Scheraga, *J. Chem. Phys.* **51**, 4751 (1969).
- ²⁰M. Fixman, *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3050 (1974).
- ²¹D. McQuarrie, *Statistical Mechanics* (Random, New York, 1976).
- ²²A. A. Sveshnikov, *Problems in Probability Theory, Mathematical Statistics and the Theory of Random Functions* (Dover, New York, 1970), p. 159.
- ²³The substitution $y = \exp(-x)$ converts the integrals into forms that are tabulated.
- ²⁴D. W. Rebertus, B. Berne, and D. Chandler, *J. Chem. Phys.* **70**, 3395 (1979).
- ²⁵R. M. Levy, M. Karplus, and J. A. McCammon, *Chem. Phys. Lett.* **65**, 4 (1979).
- ²⁶B. Bigot and W. Jorgensen, *J. Chem. Phys.* **75**, 1977 (1981).