

## Variational Approach to Correlations in Charged Polymers

Bo Jönsson\*

*Physical Chemistry 2, Chemical Center, Box 124, S-221 00 Lund, Sweden*

Carsten Peterson† and Bo Söderberg‡

*Department of Theoretical Physics, University of Lund Sölvegatan 14A, S-22362 Lund, Sweden*

(Received 9 November 1992)

A deterministic algorithm for calculating polymer properties is presented. It is based on a variational approach where the bond and Coulomb potentials are approximated by a quadratic trial energy. The parameters, which describe average atom positions and Gaussian fluctuations, are the solutions of matrix equations. By a judicious choice of parameter representations and the use of incremental matrix inversion, an efficient iterative algorithm is constructed that is several orders of magnitude faster than Monte Carlo methods. The method is explored numerically at realistic temperatures with results that agree within 5% with those of Monte Carlo calculations.

PACS numbers: 61.41.+e, 31.20.Wb, 36.20.-r

We consider polymers with nearest neighbor harmonic oscillator bonds and Coulombic self-repulsion as a model for, e.g., polyelectrolytes. These play an important role as modulators in colloidal interactions and in the interactions of large "particles" in general. Both synthetic and natural polyelectrolytes are widely used in industrial processes. From a more general point of view one could regard proteins as polyelectrolytes, since they normally carry a large number of charges at neutral pH.

The energy of such a polymer of  $N$  atoms is given by

$$\tilde{E} = \frac{k}{2} \sum_{i \neq N} \tilde{r}_{i,i+1}^2 + \frac{q^2}{4\pi\epsilon_r\epsilon_0} \sum_{i < j} \frac{1}{\tilde{r}_{ij}}, \quad (1)$$

where  $i, j = 1, \dots, N$  label the atoms and  $\tilde{r}_{ij}$  is the distance between atom  $i$  and  $j$ .

To avoid an unnecessary complication of expressions, we will work with dimensionless entities in what follows: all lengths are expressed in units of the  $N=2$  equilibrium distance  $r_0 = (q^2/4\pi\epsilon_r\epsilon_0k)^{1/3}$ , energies in  $kr_0^2$ , and temperatures in  $kr_0^2/k_B$ . With  $r_0 = 6 \text{ \AA}$ ,  $q = e$ , and  $\epsilon_r = 78.3$ , room temperature (298 K) corresponds to  $T = 0.838$ .

In dimensionless variables, the Boltzmann distribution at a finite temperature  $T$  reads

$$\frac{1}{Z} \exp(-E/T), \quad (2)$$

with

$$E = \frac{1}{2} \sum_i r_{i,i+1}^2 + \sum_{i < j} \frac{1}{r_{ij}}, \quad (3)$$

where it is understood that the global zero mode center of mass position is to be excluded from integration.

Such a polymer system is strongly fluctuating and hence standard Monte Carlo methods [1] are quite time consuming. Typically  $N = O(200)$  are within reach with good statistics for interesting quantities like atom-to-atom distances. Being able to compute such quanti-

ties reliably for larger polymers, and to extend the calculations to proteins which contain orders of magnitudes more degrees of freedom, is highly desirable. Hence a fast deterministic approach which well approximates the system is called for.

In Refs. [2,3] the variational method of Refs. [4,5] was revisited in this context. (Variational schemes have been used in the past for polymer studies [6] and in particular for polyelectrolytes [7].) The approach is based on an effective energy ansatz  $E_0$ , given by

$$E_0/T = \frac{1}{2} \sum_{ij} G^{-1}_{ij}(\mathbf{x}_i - \mathbf{a}_i) \cdot (\mathbf{x}_j - \mathbf{a}_j), \quad (4)$$

where  $\mathbf{a}_i$  define average atom positions, around which Gaussian fluctuations are given by the symmetric positive definite correlation matrix  $G_{ij}$ .

Using the effective energy, the exact free energy  $F = -T \ln Z$  of the polymer is approximated from above [4] by

$$\hat{F} = F_0 + \langle E - E_0 \rangle_0 \geq F, \quad (5)$$

where  $F_0 = -T \ln Z_0$ , and  $\langle \rangle_0$  refers to averages with respect to the trial Boltzmann distribution  $\exp(-E_0/T)$ . (For potentials more singular than  $1/r^2$ ,  $\langle E \rangle_0$  will be divergent. Such potentials are not physical and we do not consider this limitation of the approach a serious one.) The parameters  $G_{ij}$  and  $\mathbf{a}_i$  are to be determined such that the variational free energy  $\hat{F}$  is minimized. The resulting effective Boltzmann distribution is then used to approximate expectation values  $\langle f \rangle$  by effective ones  $\langle f \rangle_0$ .

In general the variational approach is expected to be most accurate at high dimensions [2,5]. Apparently, this has discouraged the community from pushing the approach for three-dimensional polymers into a numerical confrontation. Also, using the method in a naive way would give scaling behavior like  $N^4$ , which makes the method intractable for large sizes. We have also found

empirically that such a naive implementation is plagued with bad convergence properties.

In this Letter an algorithm is developed that lowers the computational costs to  $N^3$  with controlled and nice convergence properties. To this end we start by simplifying the formalism. The symmetry and positivity constraints on  $G_{ij}$  are automatic if  $G_{ij}$  is expressed as the product of a matrix with its transpose:

$$G_{ij} = \sum_{\mu=1}^{N-1} z_{i\mu} z_{j\mu} = \mathbf{z}_i \cdot \mathbf{z}_j. \quad (6)$$

The interpretation of the local parameter  $\mathbf{z}_i$  is simple: it gives the fluctuating part of  $\mathbf{x}_i = \mathbf{a}_i + \sum_{\mu} z_{i\mu} \boldsymbol{\eta}_{\mu}$ , where

$\boldsymbol{\eta}_{\mu}$  are a set of independent Gaussian noise variables of unit variance. (Note that  $\mathbf{z}_i$  are vectors not in  $R^3$  but in  $R^{N-1}$ .) In terms of  $\mathbf{a}_i$  and  $\mathbf{z}_i$ , the variational free energy becomes (ignoring the trivial constant  $\langle E_0 \rangle_0$ )

$$\hat{F} = -3T \ln \det z + \frac{1}{2} \sum_i (3z_{i,i+1}^2 + a_{i,i+1}^2) + \sum_{i < j} \frac{1}{a_{ij}} \operatorname{erf} \left( \frac{a_{ij}}{\sqrt{2} z_{ij}} \right). \quad (7)$$

Equations for a local extremum of  $\hat{F}$  are obtained by differentiation with respect to  $\mathbf{z}_i$  and  $\mathbf{a}_i$ ,

$$\frac{\partial \hat{F}}{\partial \mathbf{z}_i} = 0, \quad \frac{\partial \hat{F}}{\partial \mathbf{a}_i} = 0, \quad (8)$$

where

$$\frac{\partial \hat{F}}{\partial \mathbf{z}_i} = -3T \mathbf{w}_i + 3(2\mathbf{z}_i - \mathbf{z}_{i+1} - \mathbf{z}_{i-1}) - \sqrt{\frac{2}{\pi}} \sum_{j(\neq i)} \frac{\mathbf{z}_i \cdot \mathbf{z}_j}{z_{ij}^3} \exp \left( -\frac{\mathbf{a}_i \cdot \mathbf{z}_j^2}{2z_{ij}^2} \right), \quad (9)$$

$$\frac{\partial \hat{F}}{\partial \mathbf{a}_i} = 2\mathbf{a}_i - \mathbf{a}_{i+1} - \mathbf{a}_{i-1} - \sqrt{\frac{2}{\pi}} \sum_{j(\neq i)} \frac{\mathbf{a}_i \cdot \mathbf{z}_j}{a_{ij}^3} \int_0^{a_{ij}/z_{ij}} x^2 e^{-x^2/2} dx. \quad (10)$$

In Eqs. (9) and (10)  $w_{i\mu}$  is the (transposed) matrix inverse of  $z_{i\mu}$ .

Note that Eqs. (9) and (10) are consistent with  $\mathbf{a}_i = 0$ . Such a solution does indeed exist, and provides the global minimum for high enough  $T$ . For this solution the equation for  $\mathbf{z}_i$  simplifies to

$$0 = -3T \mathbf{w}_i + 3(2\mathbf{z}_i - \mathbf{z}_{i+1} - \mathbf{z}_{i-1}) - \sqrt{\frac{2}{\pi}} \sum_{j(\neq i)} \frac{\mathbf{z}_i \cdot \mathbf{z}_j}{z_{ij}^3}. \quad (11)$$

Empirically and theoretically it is advantageous to work with relative coordinates,  $\mathbf{x}_i^{(r)} = \mathbf{x}_{i+1} - \mathbf{x}_i$ ,  $\mathbf{a}_i^{(r)} = \mathbf{a}_{i+1} - \mathbf{a}_i$ , and  $\mathbf{z}_i^{(r)} = \mathbf{z}_{i+1} - \mathbf{z}_i$ . A simple gradient descent method with a large step size  $\epsilon$  can then be used,

$$\Delta \mathbf{z}_i^{(r)} = -\epsilon_z \frac{\partial \hat{F}}{\partial \mathbf{z}_i^{(r)}}, \quad \Delta \mathbf{a}_i^{(r)} = -\epsilon_a \frac{\partial \hat{F}}{\partial \mathbf{a}_i^{(r)}}, \quad (12)$$

that gives fast convergence to a solution of Eqs. (8). The reciprocal variables  $\mathbf{w}_j^{(r)}$  are updated using incremental matrix inversion [8]—the change due to  $\Delta \mathbf{z}_i^{(r)}$  is given by

$$\Delta \mathbf{w}_j^{(r)} = -\frac{\mathbf{w}_i^{(r)} (\mathbf{w}_j^{(r)} \cdot \Delta \mathbf{z}_i^{(r)})}{1 + \mathbf{w}_i^{(r)} \cdot \Delta \mathbf{z}_i^{(r)}} \quad (13)$$

to be applied in parallel for  $j$ . As a by-product, the denominator  $(1 + \mathbf{w}_i^{(r)} \cdot \Delta \mathbf{z}_i^{(r)})$  gives the multiplicative change in  $\det z^{(r)}$  needed to keep track of  $\hat{F}$ .

The number of computations in each iteration step for this procedure is proportional to  $N^3$ . The number of iterations required for convergence is a slowly growing function  $g(N)$ . In total the execution time  $\tau$  of the algorithm scales like

$$\tau = N^3 g(N). \quad (14)$$

For the exact thermal distribution, Eq. (2), the following virial theorem holds, due to the scaling properties of the energy [Eq. (3)]:

$$2\langle E_G \rangle - \langle E_C \rangle = 3(N-1)T, \quad (15)$$

where  $E_G$  and  $E_C$  are the Gaussian and Coulomb parts of the energy, respectively—the temperature factor  $3(N-1)$  reflects the number of degrees of freedom. The virial theorem provides a partial thermalization check in Monte Carlo simulations. A nontrivial result of the scaling properties of the effective energy  $E_0$  [8] is that the virial theorem is *respected* by the variational approach.

As mentioned above, at high enough temperatures the global minimum of  $\hat{F}$  is characterized by  $\mathbf{a}_i = 0$ , to be referred to as a purely fluctuating solution. This type of solution is a local minimum also at lower  $T$ , where, however, also a rigid ( $\mathbf{a}_i \neq 0$ ) solution exists. Below a certain critical temperature  $T_c$ , the latter gives the global minimum, indicating a first order phase transition. This is probably an artifact of the variational approach—the original system shows no sign of having a phase transition. It turns out that for realistic choices of  $T$  one is in the region where  $\mathbf{a}_i = 0$  gives rise to good results (see Fig. 1 below). Hence we will consistently use the  $\mathbf{a}_i = 0$  solutions in the numerical explorations. This implies faster performance since only the  $\mathbf{z}_i^{(r)}$  equations in Eq. (12) are needed.

In the high  $T$  limit, the variational results can be expanded in  $1/T$ . Thus, for the expectation value of the Gaussian part  $E_G$  of the energy, the first two terms of the expansion yield

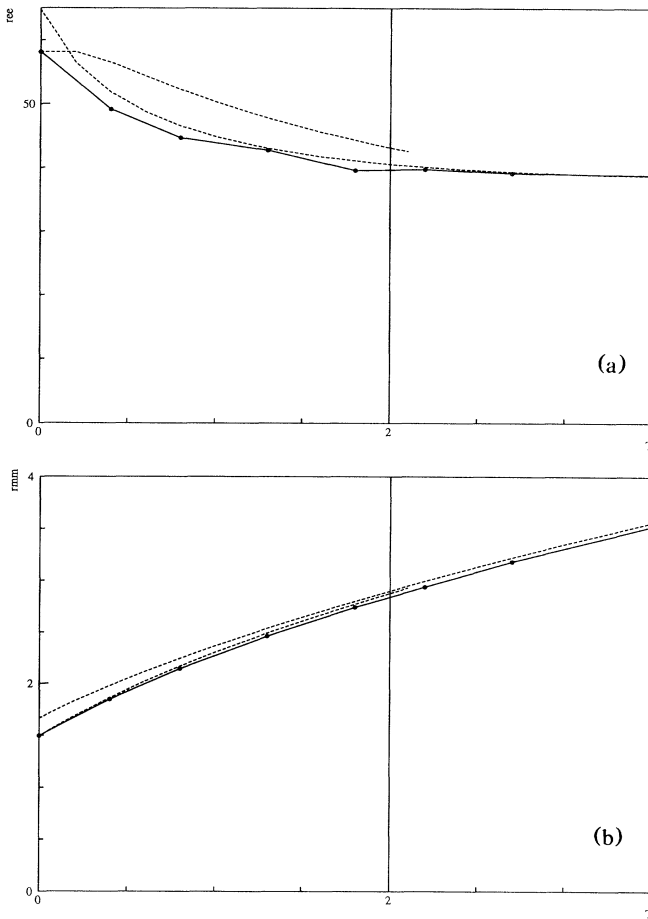


FIG. 1. (a)  $\langle r_{ee}^2 \rangle^{1/2}$  and (b)  $\langle r_{mm}^2 \rangle^{1/2}$  versus temperature from Monte Carlo (solid) and from the variational calculation (dashed), for  $N = 40$ . In this case,  $T_c \approx 2$ .

$$\langle E_G \rangle_0 = \frac{3(N-1)}{2}T + \frac{1}{\sqrt{2\pi T}} \sum_{i < j} \frac{1}{\sqrt{j-i}}, \quad (16)$$

which agrees with the exact result to the order shown—the first discrepancy occurs in the  $O(T^{-2})$  term. The same holds for any quadratic expectation value  $\langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle$ . By the virial theorem, this also holds for the Coulomb energy  $\langle E_C \rangle$ , and thus for the total energy as well.

In the low  $T$  limit, the exact result for the total internal energy, expanded in powers of  $T$ , is

$$\langle E \rangle = E_{cl} + \frac{(3N-5)T}{2} + O(T^2), \quad (17)$$

where  $E_{cl}$  is the classical ( $T = 0$ ) energy and  $3N-5$  is the number of degrees of freedom, modified for symmetry. In this limit, the variational free energy is minimized by the rigid solution, for which the corresponding expansion is

$$\langle E \rangle_0 = E_{cl} + \frac{(3N-3)T}{2} + O(T^2), \quad (18)$$

TABLE I.  $\langle r_{mm}^2 \rangle^{1/2}$  and  $\langle r_{ee}^2 \rangle^{1/2}$  in units of  $r_0 = 6 \text{ \AA}$  as computed with the variational method (V) and Monte Carlo (MC). The errors originating from the Monte Carlo runs are estimated to be  $O(0.1\%)$ .

	$N$	20	40	80	160
$r_{mm}$	V	2.173	2.266	2.351	2.428
	MC	2.093	2.172	2.238	2.305
	diff.	3.8%	4.3%	5.1%	5.3%
$r_{ee}$	V	20.33	46.17	105.3	237.5
	MC	19.83	46.67	100.7	230.1
	diff.	2.5%	1.0%	4.6%	3.2%

where the first term is the correct one and the second term is qualitatively correct for large  $N$ . For the purely fluctuating variational solution, the low  $T$  result would be

$$\langle E \rangle_0 = (6/\pi)^{1/3} E_{cl} + O(T), \quad (19)$$

which is off by 24% already in the first term. The same factor results for any quadratic expectation value (and for any single term in  $\langle E_C \rangle$ ) in this limit. Thus, for rms atom-to-atom distances,  $\langle r_{ij}^2 \rangle^{1/2}$ , this amounts to an error of 11%.

Thus the variational approach is exact in both  $T$  limits, and can be expected to give a reasonable approximation also at finite temperatures. In Fig. 1 we give a qualitative picture of the approximation as a function of  $T$  for both types of solution.

To evaluate the accuracy of our method, Monte Carlo calculations (Metropolis) utilizing translational and rotational symmetries (cf. the pivot method in Ref. [1]) were performed for sizes ranging from 20 to 160 (see Table I) at a fixed  $T = 0.838$  (corresponding to 298 K). The resulting comparisons with variational calculations for nearest neighbor ( $\langle r_{mm}^2 \rangle^{1/2}$ ) and end-to-end ( $\langle r_{ee}^2 \rangle^{1/2}$ ) distances are shown in Table I. The results from the variational approach are impressive—the deviation from the MC results is well below the 11% bound discussed above. Similar judgment can be made on the correlations between the first link and the others as shown in Fig. 2. The algorithm is extremely fast: only  $O(100)$  iterations are needed for convergence, which corresponds to approximately 30 CPU minutes on a DEC 3100 workstation for an  $N = 160$  problem.

We have presented a very cost effective and accurate deterministic approach for computing atom-to-atom correlations in polymer chains. This paper only concerns the unscreened Coulomb case. The case of screened potentials is presently being investigated. At low temperatures the performance of the approach is inferior to the pure Coulomb case. However, using an anisotropic low- $T$  variational ansatz should improve the situation [8]. The variational approach is also directly applicable to more general topologies. Proteins could also be treated in this

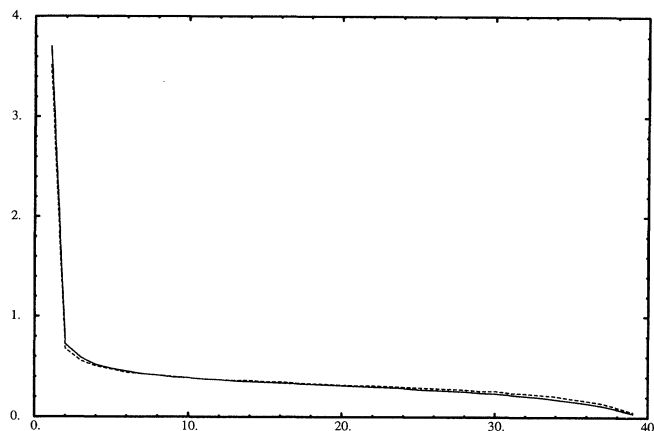


FIG. 2. Link correlations  $\langle \mathbf{v}_1 \cdot \mathbf{v}_i \rangle$ , where  $\mathbf{v}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ . Full line: variational result; dashed line: Monte Carlo.

way provided the traditional Lennard-Jones potentials are replaced by forms that are less singular at the origin.

One should point out that in this approach the free energy is naturally available in contrast to the case in Monte Carlo calculations.

A potential improvement strategy that should be ex-

plored is to perform Monte Carlo calculations using variables adapted to the variational effective energy [9].

\* Electronic address: fk2boj@grosz.fkem2.lth.se

† Electronic address: carsten@thep.lu.se

‡ Electronic address: bs@thep.lu.se

- [1] K. Kremer and K. Binder, *Comput. Phys. Rep.* **7**, 261 (1988).
- [2] J.-P. Bouchaud, M. Mezard, G. Parisi, and J.S. Yedida, *J. Phys. A* **24**, L1025 (1991).
- [3] J.-P. Bouchaud, M. Mezard, and J.S. Yedida, in *Neural Networks: from Biology to High Energy Physics*, edited by O. Benhar, C. Bosio, P. Del Giudice, and E. Tabet (ETS Editrice, Pisa, 1992).
- [4] R.P. Feynman, *Statistical Mechanics, Frontiers in Physics* (Benjamin, Reading, MA, 1972).
- [5] J. de Cloizeaux, *J. Phys. (Paris)* **31**, 715 (1970).
- [6] S.F. Edwards and M. Muthukumar, *J. Chem. Phys.* **89**, 2435 (1988).
- [7] Y. Chen and A.L. Kholodenko, *J. Chem. Phys.* **86**, 1540 (1987).
- [8] B. Jönsson, C. Peterson, and B. Söderberg (to be published).
- [9] This was suggested to us by R. Blankenbecler (Stanford Linear Accelerator Center).