Accelerating molecular dynamics simulations by linear prediction of time series

B. Brutovsky, T. Mülders, and G. R. Kneller

Citation: The Journal of Chemical Physics **118**, 6179 (2003); doi: 10.1063/1.1559033
View online: http://dx.doi.org/10.1063/1.1559033
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/118/14?ver=pdfcov
Published by the AIP Publishing

Accelerating molecular dynamics simulations by linear prediction of time series

B. Brutovsky,[a] T. Mülders, and G. R. Kneller[b]
Centre de Biophysique Moléculaire, CNRS UPR 4301, Rue Charles Sadron, F-45071 Orléans Cedex 2, France

(Received 12 March 2002; accepted 15 January 2003)

We present a molecular dynamics simulation scheme which allows to speed up molecular dynamics simulations by linear prediction of force time series. The explicit calculation of nonbonding forces is periodically replaced by linear prediction from past values. Applying our method to liquid oxygen consisting of flexible molecules we obtained real speedups between 5.4 and 6.5, compared to conventional molecular dynamics simulations. Here only the bond-stretching forces were calculated at each time step. We demonstrate that essential dynamical quantities, such as the mean-square displacement and the velocity autocorrelation function, are preserved. © *2003 American Institute of Physics.* [DOI: 10.1063/1.1559033]

I. INTRODUCTION

Molecular dynamics (MD) simulations of condensed matter systems have proved to provide valuable information for the interpretation of experimental data and for the development of theoretical models. Since the early days of MD simulations significant effort has been made to develop efficient simulation schemes which allow to extend the accessible time scale. Well-known examples are the multiple time step (MTS) methods developed by Streett *et al.*[1,2] and Tuckerman *et al.*,[3,4] but other approaches can be found as well.[5–9] The MTS simulation algorithm developed by Tuckerman and co-workers has the property of being symplectic, preserving thus the phase space volume of the simulated system. Since symplectic integrators have also the desirable property of ensuring long time stability of MD simulations, many articles have been devoted to the development of such integrators. Comparative studies of symplectic integrators can be found in the articles by Gray *et al.*[10] and Skeel *et al.*[11] Recent developments of efficient Verlet-type integrators have been published in Ref. 12.

A crucial point in developing efficient simulation methods is to recognize that, although the slower, large-amplitude motions are in most cases the interesting ones, the fast and more localized motions are nevertheless important and cannot simply be neglected. Because of the high atomic density in condensed matter systems slow motions due to soft forces and fast motions due to stiff forces cannot be disentangled. In the MTS method one exploits the fact that the ''soft,'' long-ranged forces, whose evaluation is the most time consuming part of an MD simulation, vary more slowly than ''hard,'' short-ranged forces. In the time-reversible and symplectic RESPA scheme developed by Tuckerman *et al.* the long ranged forces are updated less often than the short-ranged forces, and in the MTS method proposed by Streett *et al.* the long-ranged forces are in addition predicted on the basis of a Taylor expansion. The latter method is, however, not time-reversible.

Here we propose a method where soft, long-ranged forces are estimated on the basis of *linear prediction* of time series which is well established in the theory of signal processing.[13] This means that we do not use a Taylor expansion to predict forces, but a prediction on a *statistical* basis. Although nonlinear estimators have proved to be useful in practice, linear estimators are most often used since they can be realized by very efficient algorithms.[13–15] Moreover, they can be more directly related to stochastic models for time series, such as the autoregressive model.[13,16] This model has been used recently to compute memory functions from MD simulations.[17] Keeping in mind the limitation of long-time predictability of MD trajectories, which has been profoundly discussed in the context of Lyapunov instabilities,[18] we try to model the short time evolution of the underlying force trajectory as close as possible to reality. The essential point is to intertwine these predictions with real MD simulations, such that long-time stability is assured.

II. THEORY

A. Linear prediction (LP)

Linear prediction[13,19] is a method to estimate the value of a signal x at time t by a linear combination of values in the past. In the following it will always be assumed that $x(t)$ is a signal with zero mean. If the time parameter takes only discrete values, the linear predictor of order P for the true value $x(t)$ is written as

$$\hat{x}(t)=\sum_{n=1}^{P} a_n x(t-n\Delta t). \tag{1}$$

Here $\Delta t>0$ is the sampling interval. Defining the prediction

[a]On leave from Department of Biophysics, P. J. Šafárik University, Jesenná, Košice, Slovakia.
[b]Author to whom correspondence should be addressed. Electronic mail: kneller@cnrs-orleans.fr

error $\epsilon(t) := x(t) - \hat{x}(t)$, the predictor weights $\{a_n\}$ can be determined by the requiring that the time (or ensemble) averaged square of $\epsilon(t)$ is minimized,

$$\langle \epsilon^2(t) \rangle = \min(\{a_n\}). \tag{2}$$

The necessary condition for $\langle \epsilon^2(t) \rangle$ to have a minimum is that all partial derivatives with respect to the predictor coefficients vanish. This yields

$$\left\langle \left(x(t) - \sum_{n=1}^{P} a_n x(t - n\Delta t) \right) x(t - k\Delta t) \right\rangle = 0, \quad k = 1,...,P. \tag{3}$$

The above relation states that the prediction error is uncorrelated ("orthogonal") to the P preceding values of the signal in the past. Introducing the autocorrelation function,

$$r(t_1, t_2) \equiv \langle x(t_1) x(t_2) \rangle, \tag{4}$$

and assuming a stationary process, such that $r(t_1, t_2) = r(t_1 - t_2, 0) = r(t_2 - t_1, 0) \equiv r(|t_2 - t_1|)$, Eq. (3) can be written in the compact form,

$$\sum_{j=1}^{P} r(|i-j|\Delta t) a_j = r(i\Delta t), \quad i = 1,...,P. \tag{5}$$

One obtains thus a system of P linear equations for the P predictor coefficients $\{a_k\}$. Equation (5) is known as the Yule–Walker equation.[13,16] It should be noted that the matrix of coefficients has Toeplitz-form. This fact has been exploited to develop efficient methods for the calculation of the predictor coefficients from a given sequence $r(i\Delta t)(i = 0,...,P)$ of the autocorrelation function.[20,21]

Since stationarity of $x(t)$ has been assumed, the mean-square prediction error becomes time independent and is given by

$$\langle \epsilon^2 \rangle = r(0) - \sum_{n=1}^{P} a_n r(n\Delta t). \tag{6}$$

To obtain this expression one uses the definition of $\epsilon(t)$ together with the orthogonality relation (3).

B. Determination of the predictor weights

When predictor coefficients for linear prediction of forces are to be determined from MD trajectories one must realize that the predictor coefficients for physically identical atoms are the same. Each of the $3N$ individual force trajectories in a simulation, N being the number of physically equivalent atoms, can be viewed as a section of one generic force trajectory, $f(\tau)$. The goal is not to determine for each atom and each force component a set of predictor coefficients which is optimal for the respective trajectory, but a set which gives on average good predictions for *all* $3N$ trajectories. To determine the predictor weights we use linear regression[19] which allows us to capture as many features of the dynamical system as possible in a finite set of observations.

If k enumerates the observations $(k = 1,...,M)$, we define j_k to be a randomly chosen component of the $3N$ force trajectories under consideration and t_k a randomly chosen point on the time axis. We have $1 \leqslant j_k \leqslant 3N$ and $P\Delta t \leqslant t_k \leqslant (N_t - 1)\Delta t$, where N_t is the number of time steps in the input trajectory. For each choice of $t_k = (k-1)\Delta t$, one must dispose of P preceding values of the force. Therefore the lower limit for t_k is $P\Delta t$.

Now we define the M-dimensional vector of observations,

$$\mathbf{f} := [f_{j_1}(t_1),...,f_{j_M}(t_M)]^T, \tag{7}$$

and the $M \times P$ matrix $\mathbf{F}$,

$$\mathbf{F} := \begin{bmatrix} f_{j_1}(t_1 - \Delta t) & \cdots & f_{j_1}(t_1 - P\Delta t) \\ f_{j_2}(t_2 - \Delta t) & \cdots & f_{j_2}(t_2 - P\Delta t) \\ \vdots & & \vdots \\ f_{j_M}(t_M - \Delta t) & \cdots & f_{j_M}(t_M - P\Delta t) \end{bmatrix}. \tag{8}$$

With $\mathbf{a} \equiv [a_1,...,a_P]^T$ we try to find $\mathbf{a}$ such that

$$\mathbf{f} \approx \mathbf{F} \cdot \mathbf{a}, \tag{9}$$

where "$\approx$" stands for "as close as possible to" in a least-squares sense. The predictor coefficients are now determined from the condition

$$\|\mathbf{f} - \mathbf{F} \cdot \mathbf{a}\|^2 = \mathrm{Min}(\mathbf{a}). \tag{10}$$

The unique solution of minimum norm of Eq. (10) is given by

$$\mathbf{a}_{\mathrm{opt}} = \mathbf{F}^+ \cdot \mathbf{f}, \tag{11}$$

where $\mathbf{F}^+$ is the *generalized inverse* of $\mathbf{F}$.[22] The solution as given by Eq. (11) is also the unique solution if all columns in $\mathbf{F}$ are linearly independent. In this case we have

$$\mathbf{F}^+ = (\mathbf{F}^T \cdot \mathbf{F})^{-1} \cdot \mathbf{F}^T, \tag{12}$$

and $\mathbf{a}_{\mathrm{opt}}$ is the solution of the well-known Gaussian normal equations,

$$(\mathbf{F}^T \cdot \mathbf{F}) \cdot \mathbf{a} = \mathbf{F}^T \cdot \mathbf{f}. \tag{13}$$

Using the structure of the matrix $\mathbf{F}$ as given by Eq. (8) shows that the elements of the $P \times P$ matrix $\mathbf{F}^T \cdot \mathbf{F}$ are proportional to an estimator for the autocorrelation function $r(t)$ of the generic time series $f(\tau)$. Defining

$$\mathbf{R} := \frac{1}{M} \mathbf{F}^T \cdot \mathbf{F}, \tag{14}$$

and assuming a (wide-sense) stationary process for the generic time series $f(\tau)$, we have

$$R_{ij} = \frac{1}{M} \sum_{k=1}^{M} f_{j_k}(t_k - i\Delta t) f_{j_k}(t_k - j\Delta t) \approx r(|j-i|\Delta t). \tag{15}$$

The above considerations show that the Gaussian normal equations play the role of the Yule–Walker equations (5) in traditional linear prediction, where the predictor weights are optimized on the basis of a *single* time series.

It is well known that linear regression using the Gaussian normal equations leads to numerical problems if the columns of $\mathbf{F}$ become linearly dependent (or almost linearly dependent). In this case $\mathbf{F}^T \cdot \mathbf{F}$ is (almost) singular. A numeri-

cally stable approach is to use singular value decomposition (SVD).[22–24] The matrix **F** is then decomposed as

$$\mathbf{F}=\mathbf{U}\cdot\mathbf{D}(\eta_1,\ldots,\eta_Q,0,...,0)\cdot\mathbf{V}^T, \tag{16}$$

where **U** and **V** are orthogonal matrices of dimensions $M\times M$ and $P\times P$, respectively, and **D** is an $M\times P$ diagonal matrix containing $Q\leqslant P$ nonzero elements $\eta_k\,(k=1,...,Q)$ with $\eta_k>0$. Using SVD, the pseudoinverse of **F** reads

$$\mathbf{F}^+=\mathbf{V}\cdot\mathbf{D}^T(\eta_1^{-1},\ldots,\eta_Q^{-1},0,...,0)\cdot\mathbf{U}^T. \tag{17}$$

If the columns of **F** are linearly independent we have $Q=P$, and with Eq. (15) we can write

$$\mathbf{R}=\mathbf{V}\cdot\mathrm{diag}\left(\frac{\eta_1^2}{M},\ldots,\frac{\eta_P^2}{M}\right)\cdot\mathbf{V}^T. \tag{18}$$

The squares of the singular values of **F** are thus proportional to the positive eigenvalues (principal components) of the estimated autocorrelation matrix given in Eq. (15).

III. THE ALGORITHM

In the following we present a MD simulation scheme in which the explicit calculation of "soft" forces is periodically replaced by linear prediction from past values. "Hard," rapidly varying forces are computed explicitly at each time step. The corresponding notation is f^s and f^h, respectively. In the following we associate the hard forces with bond-stretching forces ("bonding forces") and the soft forces with Lennard-Jones forces ("nonbonding forces").

A. Preparation phase

The first step in the preparation phase consists in a conventional MD simulation. In the work presented here we use the leapfrog integration scheme,

$$v_i(t+\Delta t/2)\leftarrow v_i(t-\Delta t/2)+\frac{\Delta t}{m}(f_i^h(t)+f_i^s(t)), \tag{19}$$

$$r_i(t+\Delta t)\leftarrow r_i(t)+\Delta t v_i(t+\Delta t/2). \tag{20}$$

The index i runs over all $3N$ Cartesian coordinates, Δt is the simulation time step, and m is the mass of the atoms considered.

During the initial MD phase we record for each Cartesian component of the soft, nonbonding forces about hundred consecutive trajectory segments of length $P+1$ from which the system of equations (9) is constructed. An important technical point is that the calculation of predictor coefficients should be based on *smooth* force trajectories. In MD simulations one uses typically a cutoff for the calculation of nonbonding forces, in order to save computer time. Therefore, the corresponding force time series exhibit small nonphysical jumps when initially interacting particles move out of the cutoff sphere and initially not interacting particles move into the cutoff sphere. We found that the use of smooth force trajectories for the determination of predictor coefficients reduces the mean-square (one-step) prediction error for the nonbonding forces by two orders of magnitude, which has a clear impact on the stability and the efficiency of the simulation scheme. Therefore, the vector of observations **f** and the matrix **F**, defined in Eqs. (7) and (8), respectively, are not directly constructed from the MD trajectory, but from smooth trajectory segments which are constructed separately. For this purpose we compute for each atom α trajectory segments of length $P+1$ for the non-bonding force trajectory to which all atoms $\beta\neq\alpha$ contribute which are initially within the cut-off, irrespective if they leave the cut-off sphere during the following P steps. Figure 1 illustrates this point. It shows the trajectory of a nonbonding force from a MD simulation together with the corresponding segments used to construct the system of Eq. (9). The example is taken from the application described below, where $P=4$.

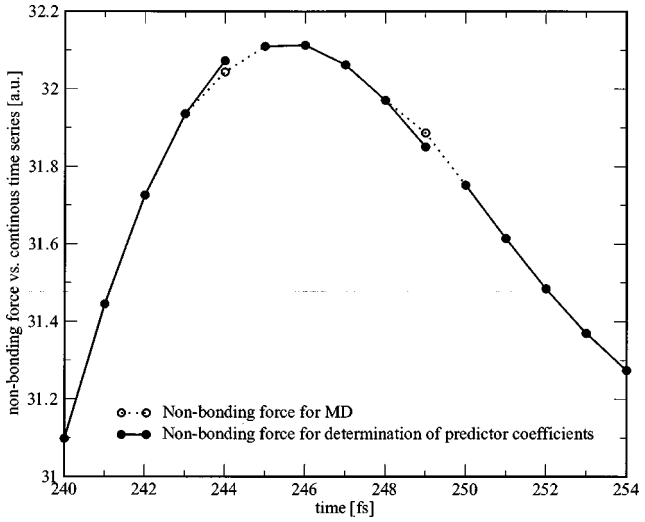

FIG. 1. Comparison of soft, nonbonding force trajectories used for MD simulation (dotted line, open circles) and for construction of the set of Eq. (9) (solid line and filled circles). The example is taken from the application of our hybrid dynamics MD scheme to liquid oxygen.

The choice of the dimensions M and P is motivated by the following considerations:

(a) M should be of the order of a few thousands in order to ensure good statistics, i.e., a representative description of the whole force trajectory. In practice, M is essentially limited by the available computer memory.

(b) To fix P we consider the speedup G of the method. Somewhat simplifying one can say that $G\approx(N_{\mathrm{pred}}+N_{\mathrm{MD}})/(N_{\mathrm{MD}})$, where N_{MD} is the number of MD steps and N_{pred} is the number of predicted steps. In doing so, one neglects any overhead during the prediction phase, such as the calculation of hard (bonding) forces and the construction of Verlet neighbor lists. We note that the CPU time needed for prediction of the forces itself can be safely neglected. Therefore the ratio $N_{\mathrm{pred}}/N_{\mathrm{MD}}$ should be as large as possible and N_{MD} should be as small as possible. It is well known that Lyapunov instabilities limit the long time predictability of the dynamics of many-body systems.[18] Therefore N_{pred} is limited effectively to a few ten steps. On the other hand, each new prediction phase must start from a set of at least P real MD steps to avoid an exponential accumulation of errors. Therefore we may write

$$G\approx 1+\frac{N_{\mathrm{pred}}}{P}. \tag{21}$$

We found that the upper limit for N_{pred} is about 30 and the lower limit of P is 4 for the system presented below.

B. Hybrid dynamics (HD) simulation

The predictor weights obtained in the preparation phase are now used in a hybrid MD simulation scheme, which consists of periodic switching between longer intervals of partially predicted dynamics and shorter intervals of conventional MD.

1. Prediction phase

To integrate the equations of motions within the leapfrog scheme, we write

$$v_i(t+\Delta t/2) \leftarrow v_i(t-\Delta t/2) + \frac{\Delta t}{m}(f_i^h(t) + \hat{f}_i^s(t)), \tag{22}$$

$$r_i(t+\Delta t) \leftarrow r_i(t) + \Delta t v_i(t+\Delta t/2), \tag{23}$$

where $f_i^h(t)$ are the hard, bonding forces, which are always explicitly computed, and $\hat{f}_i^s(t)$ are the corresponding *predicted* soft, nonbonding forces,

$$\hat{f}_i^s(t) = \sum_{n=1}^{P} a_n f_i^s(t-n\Delta t), \quad i=1,...,3N. \tag{24}$$

We recall that the predicted forces are always calculated from smooth force trajectory segments. Except for the first prediction phase, the latter are constructed during the respective preceding MD phase (see next step).

2. MD phase

During this phase P MD simulation steps are performed.

a. Scaling of the velocities: Although the errors accumulated during the prediction phase are small, they lead nevertheless to a small drift of the total energy. To restore the correct total energy we adjust the kinetic energy $E_{\text{kin}}(t)$ to the difference between the total (constant) energy E_{total} and the potential energy $E_{\text{pot}}(t)$ after the first step of this phase. All velocities $v_i(t)$ are multiplied by a corresponding scaling factor χ,

$$v_i'(t) \leftarrow \chi v_i(t), \tag{25}$$

where χ is given by

$$\chi = \left(\frac{E_{\text{total}} - E_{\text{pot}}(t)}{E_{\text{kin}}(t)} \right)^{1/2}. \tag{26}$$

χ is always very close to one (Fig. 6). Since the velocities for the computation of $E_{\text{kin}}(t)$ are not immediately available within the leapfrog scheme they are computed from

$$v(t) = \frac{v(t+\Delta t/2) + v(t-\Delta t/2)}{2}. \tag{27}$$

b. Partial randomization of the velocities: To ensure long-time stability of the simulation method, any accumulation of small systematic errors must be avoided. A simple way to achieve this is to partially randomize the particle velocities. Such a partial randomization can be realized by reducing all velocity components by a certain amount, and to add a corresponding stochastic component δv_i, such that the average kinetic energy equals $\frac{3}{2}Nk_BT$. Starting from the scaled velocities $v_i'(t)$ introduced in Eq. (25) one computes new velocities,

$$v_i''(t) \leftarrow (1-Q)^{1/2} v_i'(t) + \delta v_i, \quad i=1,...,3N, \tag{28}$$

where $0<Q<1$. To keep the desired average kinetic energy the random contributions δv_i are chosen from a Gaussian distribution with zero mean and variance $Q(k_BT/m)$. Once the corrected velocities $v_i''(t)$ have been calculated, the corrected velocities at time $t+\Delta/2$ are obtained via

$$v_i''(t+\Delta/2) = v_i''(t) + \frac{1}{2}\frac{\Delta t}{m}(f_i^h(t) + f_i^s(t)). \tag{29}$$

Note that the forces $f_i^s(t)$ are here explicitly calculated.

c. Construction of smooth force trajectory segments: During the MD phase smooth force trajectory segments are constructed to initiate the consecutive prediction phase. The approach is the same as in the preparation phase.

IV. APPLICATION

To demonstrate the efficiency of the simulation method described above we apply it to a simple model system consisting of 150 flexible oxygen molecules in the liquid phase at a temperature 73.0 K. Here the soft, nonbonding intermolecular forces are modeled by a Lennard-Jones potential ($\epsilon = 0.878\,64$ a.m.u.$\times$nm^2/ps^2 and $\sigma = 0.295\,99$ nm), and the O–O stretching motions are described by a hard, harmonic force with a force constant k corresponding to a vibration frequency of $\nu = 47.5$ THz. The predictor weights have been determined using the least-squares algorithm described in Sec. II B, with $M = 8000$ for the number of observations. Each line of the system (9) has been chosen from an ensemble of 100 consecutive smooth segments for the $3N$ Cartesian components of the nonbonding forces. As already mentioned, the choice of M is not very important as long as M is large enough to allow for good statistics. In all results presented here the prediction order (as well as the length of MD phase) was chosen to be $P=4$. We tested prediction

TABLE I. Five sets of predictor coefficients and the corresponding mean-square prediction error computed from different randomly chosen sets of Eq. (9). The magnitude of the forces is of the order of 10–100 in the internal units of our simulation program.

Set No.	a_1	a_2	a_3	a_4	$\langle \epsilon^2 \rangle$
1	3.908 291 293	−5.810 909 169	3.896 798 193	−0.994 206 850 4	$8.696\,98\times10^{-6}$
2	3.913 087 725	−5.825 192 814	3.911 026 860	−0.998 947 783 1	$9.558\,06\times10^{-6}$
3	3.912 751 015	−5.824 132 964	3.909 971 218	−0.998 617 145 6	$8.943\,10\times10^{-6}$
4	3.913 527 766	−5.826 594 080	3.912 450 849	−0.999 411 842 7	$9.342\,93\times10^{-6}$
5	3.908 286 315	−5.811 242 920	3.897 578 883	−0.994 649 829 1	$9.577\,80\times10^{-6}$

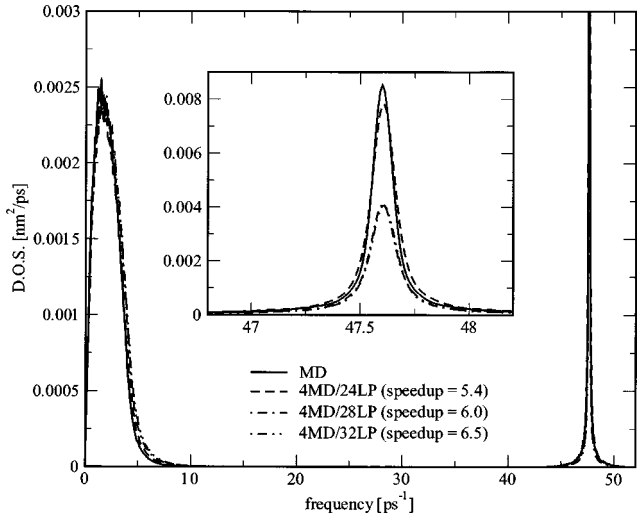

FIG. 2. Comparison of the density of states obtained by conventional MD simulation and from hybrid dynamics simulations. Results are shown for different lengths of the prediction phase, fixed length $P=4$ for the MD phase, and a randomization factor $Q=0.01$ for the velocities.

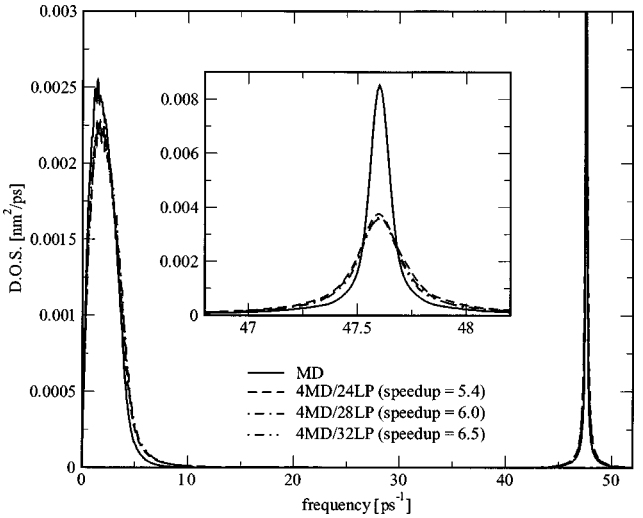

FIG. 3. As Fig. 1, but with a randomization factor $Q=0.05$ for the velocities.

intervals of $N_{\text{pred}}=24$, 28, 32, providing a measured speedup of 5.4, 6.0, and 6.5, respectively. The singular value decomposition of the matrix $\mathbf{F}$ has been performed with a routine from the LAPACK library.[25] For later analysis 50 ps trajectories were recorded, each of which was preceded by a stable simulation of 1 ns length. Table I shows the coefficients $\{a_1,a_2,a_3,a_4\}$ and the corresponding mean-square prediction error, $\langle \epsilon^2 \rangle$, for five different random sets of Eqs. (9). The mean-square prediction error was computed from the input trajectory segments which served to establish the ''training set'' (9).

As already mentioned, long time stability is a problem one has to cope with when intertwining linear prediction of force time series with normal MD simulation. This effect can be corrected for by the randomization described above. We found that a randomization of 1% ($Q=0.01$) is sufficient to guarantee long-time stability of the simulation, leaving nevertheless the dynamical properties of the system unchanged. To verify this point we considered the Fourier spectrum of the (average) velocity autocorrelation function of the oxygen atoms, sometimes also called density of states (DOS),

$$g(\omega) := \frac{1}{N}\sum_{i=1}^{N} \int_0^\infty dt \exp[-i\omega t]\langle \mathbf{v}_i(0)\cdot\mathbf{v}_i(t)\rangle, \tag{30}$$

and the corresponding mean-square displacement,

$$\mathrm{MSD}(t) := \frac{1}{N}\sum_{i=1}^{N} \langle [\mathbf{r}_i(t)-\mathbf{r}_i(0)]^2\rangle. \tag{31}$$

Figure 2 shows the DOS computed from conventional MD and from our hybrid dynamics simulation scheme, using a randomization factor $Q=0.01$. In the case of 4 MD and 24 prediction steps the results are very close to those obtained from conventional MD. For each length of the prediction phase the real speedups are given in the figure. The influence of increased randomization can be seen in Fig. 3. Here we have set $Q=0.05$, leaving all other parameters unchanged compared to Fig. 2. For all parameter combinations a broadening of the DOS can be observed.

Figures 4 and 5 show the mean-square displacements corresponding to Figs. 2 and 3. Here no clear difference can be seen between the results for $Q=0.01$ and $Q=0.05$, except for the case of $Q=0.05$ and the longest prediction interval, $N_{\text{pred}}=32$. In the latter case not only the amplitude, but also the slope of the mean-square displacement are reduced.

The relation between scaling, randomization, and prediction length is illustrated in Fig. 6. It shows the distribution of the velocity scaling factor χ introduced in Eq. (25) for different values of Q and different prediction lengths. It should be noted that the scaling factor is always close to one and approaches the ideal value of $\chi=1$ with decreasing randomization and decreasing length of the prediction phase.

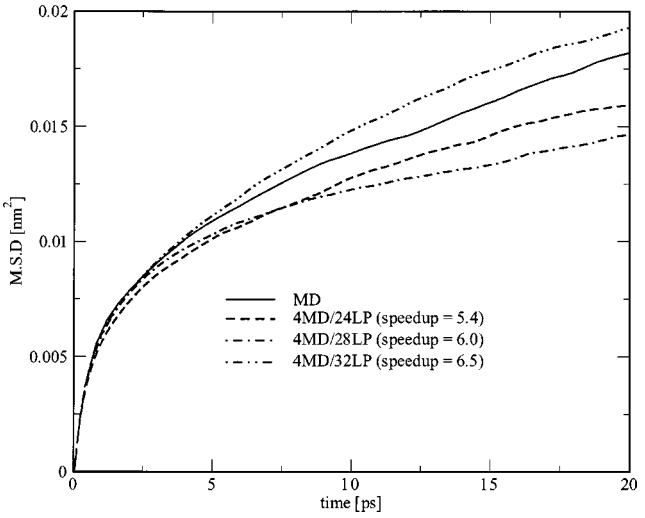

FIG. 4. Comparison of the mean-square displacement obtained by conventional MD simulation and from hybrid dynamics simulations. Results are shown for different lengths of the prediction phase, fixed length $P=4$ for the MD phase, and a randomization factor $Q=0.01$ for the velocities.

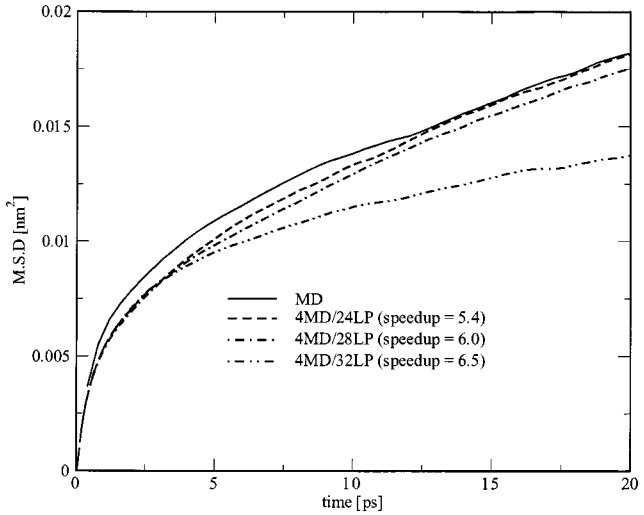

FIG. 5. As Fig. 3, but with a randomization factor $Q=0.05$ for the velocities.

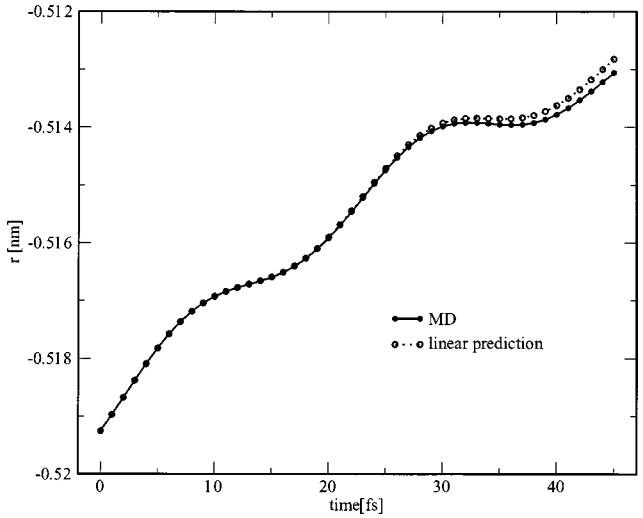

FIG. 7. Continuation of the MD trajectory of a randomly chosen particle by normal MD (solid line, filled circles) and linear prediction ($P=4$, $N_{\text{pred}}=24$, $Q=0.01$). The initial point is identical for both trajectories.

Figure 7 shows a comparison of the predicted and the corresponding ''real'' trajectory of a randomly chosen particle in which prediction has been replaced by normal MD simulation. It should be noted that the predicted trajectory not only reproduces the high frequency bond-stretching vibrations but also the slower drift of the particle.

To illustrate the impact of partial linear force prediction on collective quantities we show in Fig. 8 a zoom into the potential energy. The figure shows the evolution of the potential energy during the alternation of MD and prediction phases. For each simulation cycle we show also the corresponding values which are obtained if linear prediction is replaced by real MD simulation, i.e., if the initial common MD phase is continued.

Figure 9 gives finally an impression of the variation of the total energy compared to conventional MD. Due to the partial randomization of the velocities the total energy obtained from our hybrid MD method shows larger fluctuations, but the stability is exactly the same. In fact we found no stability problem for the longest simulations of about 1 ns.

V. DISCUSSION

We have demonstrated that linear force prediction combined with normal MD simulations and a minimal partial randomization of the velocities can be used as an efficient tool to simulate the dynamics of molecular liquids. For the system studied here we obtained speedups ranging between 5.4 and 6.5. Using an appropriate choice of the simulation parameters allows to preserve the characteristics of the short and the long time dynamics of the simulated system. The choice of the model system was motivated by the require-

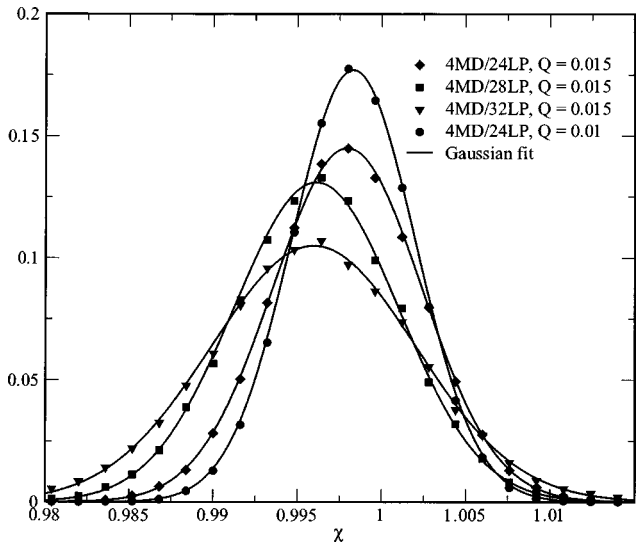

FIG. 6. Distribution of the scaling factors χ (26) during a hybrid dynamics simulation of 1 ns. The results are given for different randomization factors, different lengths of the prediction phase, and fixed length $P=4$ for the MD phase.

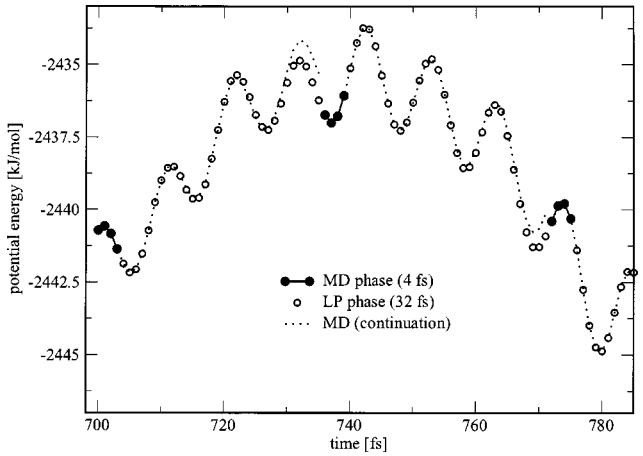

FIG. 8. Zoom on the potential energy during a hybrid dynamics simulation. Shown are the alternating phases of MD simulations (solid line and filled circles) and prediction (empty circles). The dotted lines show the piecewise continuous evolution of the potential energy for the case that each MD phase of $P=4$ steps would be continued by 32 normal MD instead of 32 prediction steps.

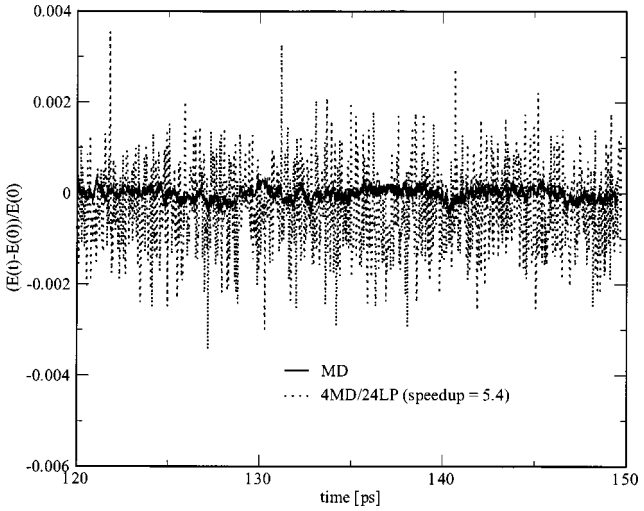

FIG. 9. Evolution of the total energy for normal MD and hybrid dynamics simulation with a randomization factor $Q=0.01$, 4 MD steps, and 24 prediction steps.

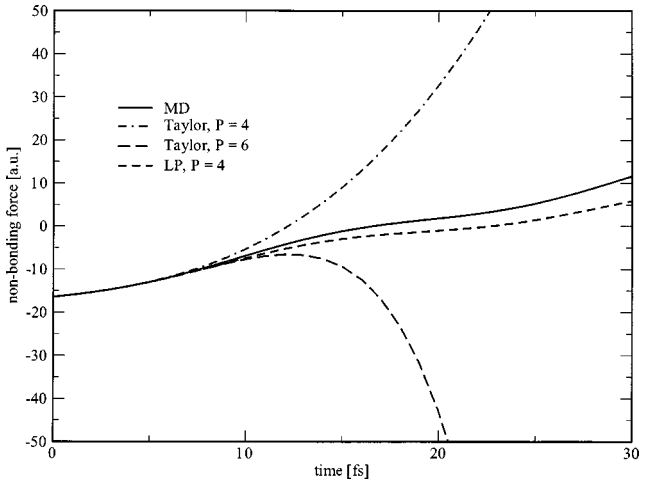

FIG. 10. Illustration of linear prediction and Taylor extrapolation of a soft, nonbonding force trajectory of randomly chosen atom. Shown is the MD force trajectory (solid line) compared to the trajectories resulting from a Taylor expansion of order 4 (dashed–dotted line), a Taylor expansion of order 6 (long dashes), and linear prediction of order 4 (short dashes). All curves include start from the same initial point, and the predictions/extrapolations start after $P=4$ or $P=6$ steps, depending on the order of the model.

ment that the atoms in the simulated system should exhibit at the same time diffusive motions known from simple liquids and fast vibrations due to chemical bonds. In MD simulations these motions are usually not considered by applying appropriate geometrical constraints. It is well known that such constraints have not only an impact on the dynamics but also on thermodynamic averages.[26] In addition, imposing certain thermodynamic ensembles in a proper way is very difficult in presence of geometrical constraints.[27] The method presented here offers the possibility to avoid constraints, yielding at the same time an even larger speedup.

It is worthwhile to compare the linear prediction of force time series as presented here to the extrapolation in terms of a discrete Taylor series (see, e.g., the MTS method by Streett *et al.*[1,2]). One knows from signal theory[13] that a function $f(t)$ whose Fourier spectrum is bandwidth limited, such that $\tilde{f}(\omega)\equiv 0$ for $|\omega|>\omega_c$, can be exactly predicted from an infinite set of past values if the sampling interval is sufficiently small,

$$f(t)=\lim_{P\to\infty}\sum_{n=1}^{P}(-1)^{n+1}\binom{P}{n}f(t-n\Delta t),\quad \text{if } \Delta t<\frac{\pi}{3\omega_c}. \tag{32}$$

The above expression corresponds to a Taylor series where all derivatives are replaced by corresponding differences constructed from the available time series for $f(t)$. Choosing a finite value for P, relation (32) can be interpreted as linear prediction with predictor coefficients,

$$a_n^{(P)}=(-1)^{n+1}\binom{P}{n}. \tag{33}$$

For $P=4$ one obtains in particular $\{a_n^{(4)}\}=\{4,-6,4,-1\}$. These values may be compared to the predictor coefficients listed in Table I. One notes that the latter are similar but not close to the Taylor coefficients obtained from Eq. (33). Here ''not close'' means that typical differences between different sets of coefficients in Table I are clearly much smaller than the differences between these coefficients and the Taylor coefficients. Figure 10 illustrates the difference between linear prediction and Taylor extrapolation. It is obvious that prediction coefficients optimized for $P=4$ give much better predictions than the Taylor coefficients of the same order. One might think that increasing the order of the Taylor extrapolation could improve the results. Figure 10 shows that using $P=6$ instead of $P=4$ improves the prediction for a few steps, but becomes then even more rapidly unstable. This behavior is well known from polynomial extrapolation of continuous functions. Moreover, the order of the model should not increased too much since this reduces the efficiency of the prediction method. This aspect has been discussed in Sec. III A. In this context it is interesting to apply the linear prediction method as described in this paper to a function which is exactly known. The results for a ''toy example'' are shown in the Appendix. They confirm that linear predictors constructed by the regression method lead to very precise predictions.

We compared our simulation method also to a conventional MTS method. For this purpose we used the simulation program DL_POLY.[28] The official release 2.12 offers the MTS algorithm by Streett *et al.*,[1] with an extension by to Coulombic systems.[29] Details are described in the user manual of DL_POLY. In conventional MTS algorithms the frequency for explicit calculation of forces is reduced when the distance between interacting particles exceeds a certain distance. One works with two update frequencies which are associated with corresponding minimal distances. To obtain a significant speedup the radius of the inner sphere, in which all interactions are computed explicitly, must be made as small as possible. Since our model system has a box size of 17.65 Å, and the cutoff distance for the Lennard-Jones forces was chosen to be 8.4 Å, we chose a radius of 3 Å for the innermost shell and an update interval of 5 MD steps for the outer shell. This gave a speedup of about two for the MTS algorithm imple-

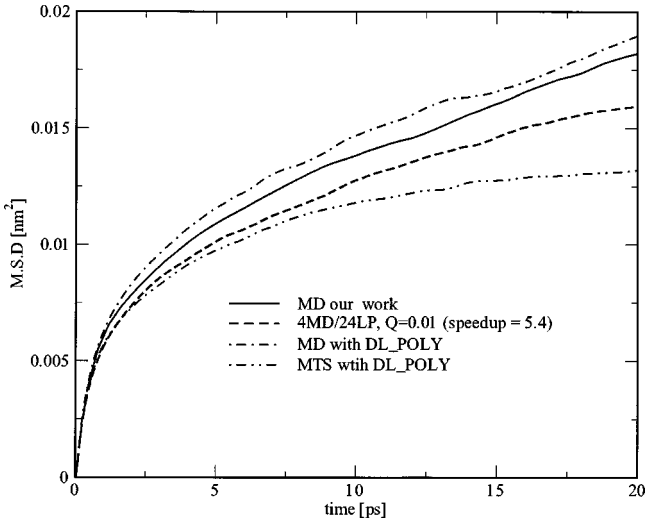

FIG. 11. Average mean-square displacements of the oxygen atoms from conventional MD simulation, hybrid MD with ($Q=0.01$, $P=4$, $N_{\text{pred}}=24$) and the MTS algorithm implemented in DL_POLY.

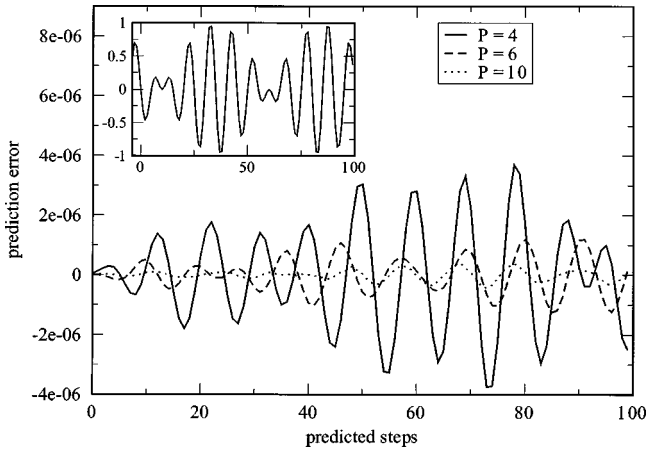

FIG. 12. Prediction errors for the discrete periodic function $f(n)$ given by Eq. (A1). The predictor weights have been obtained by linear regression and the predictor orders are $P=4$ (solid line), $P=6$ (dashed line), and $P=10$ (dotted line). The inset shows the true function $f(n)$ in the predicted time interval, augmented by the initial values for the predicted time series.

mented in DL_POLY, using a Verlet neighbor list. Figure 11 shows the comparison of the mean-square displacements obtained from conventional MD with our program, from MD with DL_POLY, from MTS simulation with DL_POLY, and from our hybrid dynamics simulation scheme with a speedup of 5.4. The figure makes clear that the diffusion constant in case of the MTS simulation is strongly reduced. During the 50 ps simulation we used to compute the mean-square displacement depicted in Fig. 11 we observed also a systematic drift of kinetic energy from the low frequency band into the high frequency band of the DOS. This second observation can be related to the reduction of the mean-square displacement, since the latter is essentially dominated by the low-frequency part of the DOS. One important point should mentioned here: Ideally one should compare our MTS method with the ''best'' known MTS method—the time-reversible MTS RESPA method—using the *same* force splitting scheme. Time-reversible integrators are known to be yield trajectories which exhibit a better long-time stability than nonreversible ones. It can be envisaged to use our splitting method with the MTS RESPA method and also to use a cutoff-based splitting with our method. For more complex systems, such as liquid water and proteins, which exhibit less clear separations of fast and slow motions, the cutoff splitting seems actually more suitable. In any case, the reader should not be left with the idea that the force splitting we use here is *a priori* essential for the method. Practical tests must show which force splitting method can or should be used.

Another point that should be addressed concerns the issues of time reversibility and symplecticity of our algorithm. All we can say is that our method is ''statistically time reversible,'' in the sense that the predictor coefficients for forward and backward motion are the same if the force trajectories can be considered as sections of a stationary stochastic process. In this case the coefficients for backward prediction follow the same Yule–Walker Eq. (5) as those for forward prediction, since the (force-) correlation function is invariant with respect to translation and reflection of the time axis. Nothing strict can be said with respect to symplecticity. However, regarding Fig. 7 suggests that at least on the short time scale of the prediction phases our method gives trajectories which are as symplectic as any MD trajectory since the predicted and real trajectories are very close.

To summarize, we can say that linear prediction of trajectories, which relies on correlations in the signal, is more adapted to molecular dynamics trajectories than conventional prediction schemes based on Taylor series expansions. This fact has been exploited to combine the conventional MTS approach with prediction of the forces which are considered to be ''slow.'' Nevertheless the MD integrator we proposed here is certainly not yet an ''all-purpose'' integrator and the next obvious step is to apply it to polar liquids like water, which exhibit also electrostatic interactions and fast molecular rotations.

ACKNOWLEDGMENTS

B.B. acknowledges financial support by a Marie Curie Fellowship of the European Community program ''Improving Human Research Potential and the Socio-Economic Knowledge Base'' under Contract No. HPMF-CT-2000-00470. T.M. acknowledges financial support by a postdoctoral fellowship (LESTUDIUM) of the Région Center and the Center National de la Recherche Scientifique.

APPENDIX: PREDICTION OF AN ANALYTICAL SIGNAL

To illustrate the application of linear prediction to a signal which is exactly known we consider the discrete function

$$f(n)=\cos\left(\frac{2\pi}{T_1}n\Delta t\right)\sin\left(\frac{2\pi}{T_2}n\Delta t\right), \tag{A1}$$

with periods $T_1=2\pi/10$, $T_2=2\pi/100$, and a sampling step of $\Delta t=0.001$. To find an optimal linear predictor we follow the regression method described in Sec. II B. The vector **f** defined by Eq. (7) is constructed from $M=100$ randomly chosen observations, and the $M\times P$ matrix **F** defined by Eq.

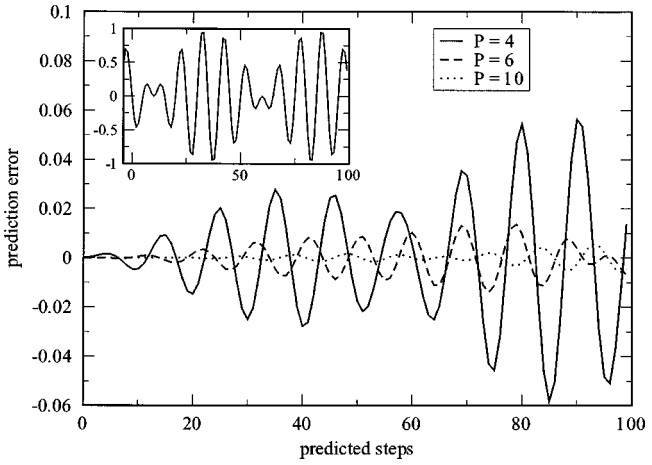

FIG. 13. The same as Fig. 12, except that the predictor weights have been obtained by the Burg method.

(8) contains the corresponding preceding values of $f(n)$. The predictor order is chosen to be $P=4,6,10$. The optimal predictor coefficients are then obtained from Eq. (11). Figure 12 shows the deviation of the predicted signal from the true signal for 100 predicted values from a randomly chosen time series of P initial values. The true function in that interval, augmented by the initial values for the predicted time series is shown in the inset of Fig. 12. The respective order of magnitude for each prediction error remains unchanged if different initial values are chosen for the predicted time series. Using the MATHEMATICA software[30] we tested also other methods to estimate the predictor coefficients for our ''toy example,'' namely the Yule–Walker method and the Burg algorithm.[13–15] These standard methods are mainly used for spectral estimation. The estimation of the parameters of an autoregressive (AR) stochastic process is actually equivalent to the determination of an optimal linear predictor. The differences between the various existing methods for spectral estimation of AR processes and linear prediction, respectively, are due to different explicit or implicit estimations of the autocorrelation function of the underlying signal. The interested reader should consult the textbooks of Papoulis and Haykin for more details.[13,19] The Yule–Walker as well as the Burg method start both from one contiguous time series. For our tests we used the same time series of $N_t=1000$ sampled points. We found that the Yule–Walker gives very bad results for our example, whereas the Burg method yields more satisfactory predictions, which are, however, not as precise as those obtained by the regression method. Figure 13 shows the prediction errors for the Burg method, using the same predictor orders as in Fig. 12 and the same initial values. It can be seen that the relative reduction of the prediction error with increasing predictor order is the same as for the linear regression method, but the *absolute* error is about four orders of magnitude bigger. In spite of the different prediction errors the predictor coefficients are very close. For $P=4$ we obtain, for example, $\mathbf{a}_{\mathrm{opt}}=(3.229\,68,\ -4.602\,26,\ 3.229\,68,\ -1.000\,00)^T$ from linear regression and $\mathbf{a}_{\mathrm{opt}}=(3.229\,66,\ -4.602\,20,\ 3.229\,66,\ -1.000\,00)^T$ from the Burg algorithm.

[1] W. B. Streett, D. J. Tildesley, and G. Saville, Mol. Phys. **35**, 639 (1978).
[2] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, Oxford, 1987).
[3] M. Tuckerman, B. J. Berne, and G. J. Martyna, J. Chem. Phys. **97**, 1990 (1992).
[4] G. J. Martyna, M. E. Tuckerman, D. J. Tobias, and M. L. Klein, Mol. Phys. **87**, 1117 (1996).
[5] R. Elber, J. Meller, and R. Olender, J. Phys. Chem. B **103**, 899 (1999).
[6] T. Huber and W. F. van Gunsteren, J. Phys. Chem. A **102**, 5937 (1998).
[7] A. Lejeune, J. Perdang, and J. Richert, Phys. Rev. E **60**, 2601 (1999).
[8] B. Mayer, G. Kohler, and S. Rasmussen, Phys. Rev. E **55**, 4489 (1997).
[9] L. S. Liebovitch, N. D. Arnold, and L. Y. Selector, J. Biol. Sys. **2**, 193 (1994).
[10] S. K. Gray, D. W. Noid, and B. G. Sumpter, J. Chem. Phys. **101**, 4062 (1994).
[11] R. D. Skeel, Guihua-Zhang, and T. Schlick, SIAM J. Sci. Comput. (USA) **18**, 203 (1997).
[12] I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. E **65**, 056706/1-5 (2002).
[13] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. (McGraw–Hill, New York, 1991).
[14] J. Burg, ''Maximum entropy spectral analysis,'' Ph.D. thesis, Stanford University, Stanford, CA, May, 1975.
[15] J. Makhoul, IEEE Trans. Acoust., Speech, Signal Process. **ASSP-25**, 423 (1977).
[16] J. Makhoul, Proc. IEEE **63**, 561 (1975).
[17] G. R. Kneller and K. Hinsen, J. Chem. Phys. **115**, 11097 (2001).
[18] W. G. Hoover, *Time Reversibility, Computer Simulation, and Chaos*, Vol. 13 in *Advanced Series in Nonlinear Dynamics* (World Scientific, Singapore, 1999).
[19] S. Haykin, *Adaptive Filter Theory* (Prentice-Hall, Englewood Cliffs, 1996).
[20] N. Levinson, J. Math. Phys. **25**, 261 (1947).
[21] J. Durbin, Rev. Inst. Int. Statist. **28**, 233 (1960).
[22] C. L. Lawson and R. J. Hansen, *Solving Least Square Problems* (Prentice–Hall, Englewood Cliffs, 1974).
[23] V. C. Klema and A. J. Laub, IEEE Trans. Autom. Control **29**, 164 (1980).
[24] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis* (Springer, Berlin, 1980).
[25] E. Anderson, Z. Bai, C. Bischof *et al.*, LAPACK *Users' Guide*, 3rd ed. (Society for Industrial and Applied Mathematics, Philadelphia, 1999).
[26] M. Fixman, Proc. Natl. Acad. Sci. U.S.A. **8**, 3050 (1974).
[27] G. R. Kneller and T. Mülders, Phys. Rev. E **54**, 6825 (1996).
[28] T. R. Forester and W. Smith, The DLPOLY user manual, version 2.0, December 1995. CCLRC, Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England.
[29] T. Forester and W. Smith, Mol. Simul. **13**, 195 (1994).
[30] MATHEMATICA 4.1.2.0, including the ''Time Series'' Package, Wolfram Research, Inc., 100 Trade Center Drive, Champaign, Illinois 61820–7237.