Stability of molecular dynamics simulations of classical systems

Søren Toxvaerd
DNRF Centre "Glass and Time," IMFUFA, Department of Sciences, Roskilde University, Postbox 260, DK-4000 Roskilde, Denmark

(Received 13 September 2012; accepted 12 November 2012; published online 4 December 2012)

The existence of a shadow Hamiltonian $\tilde{H}$ for discrete classical dynamics, obtained by an asymptotic expansion for a discrete symplectic algorithm, is employed to determine the limit of stability for molecular dynamics (MD) simulations with respect to the time-increment h of the discrete dynamics. The investigation is based on the stability of the shadow energy, obtained by including the first term in the asymptotic expansion, and on the exact solution of discrete dynamics for a single harmonic mode. The exact solution of discrete dynamics for a harmonic potential with frequency ω gives a criterion for the limit of stability $h \leq 2/\omega$. Simulations of the Lennard-Jones system and the viscous Kob-Andersen system show that one can use the limit of stability of the shadow energy or the stability criterion for a harmonic mode on the spectrum of instantaneous frequencies to determine the limit of stability of MD. The method is also used to investigate higher-order central difference algorithms, which are symplectic and also have shadow Hamiltonians, and for which one can also determine the exact criteria for the limit of stability of a single harmonic mode. A fourth-order central difference algorithm gives an improved stability with a factor of $\sqrt{3}$, but the overhead of computer time is a factor of at least two. The conclusion is that the second-order "Verlet"-algorithm, most commonly used in MD, is superior. It gives the exact dynamics within the limit of the asymptotic expansion and this limit can be estimated either from the conserved shadow energy or from the instantaneous spectrum of harmonic modes. © *2012 American Institute of Physics*. [http://dx.doi.org/10.1063/1.4768891]

I. INTRODUCTION

In molecular dynamics (MD) simulations of classical dynamics[1] one uses algorithms which propagate the dynamics with a (constant) small time-increment h. This article deals with the problem of how large this time-increment can be chosen.

Mathematical investigations have established[2–5] that if a discrete algorithm is symplectic, then there exists a *shadow Hamiltonian* $\tilde{H}(\mathbf{q}, \mathbf{p})$[6] for sufficient small h such that the discrete points $\mathbf{q}(nh)$, for a particle, generated by the symplectic algorithm, lie on the analytic trajectory for the particle obtained by $\tilde{H}(\mathbf{q}(t), \mathbf{p}(t))$. Not all algorithms for solving Newton's classical mechanical differential equation are symplectic. But the most commonly used algorithm in MD, the "Verlet" algorithm,[7] is not only symplectic, it is also time-reversible and this quality ensures the existence of $\tilde{H}$ with an exact conserved energy, but for the unknown shadow Hamiltonian. In Ref. 8 we include the first term in the asymptotic expansion in the energy estimate of $H(\mathbf{q}, \mathbf{p})$ and, for traditional values of h used in MD, there is only a marginal difference between the energy estimate of $\tilde{H}(\mathbf{q}, \mathbf{p})$ and the energy of the Hamiltonian $H(\mathbf{q}, \mathbf{p})$ for the analytic dynamics. The article deals with determining the limit of energy conservation with respect to the time-increment. Inclusion of the first term in the energy estimate of $\tilde{H}$ reduces the energy fluctuations with a factor of hundred for a traditional choice of h[8] and makes it possible to determine the limit of energy conservation in MD accurately.

The Verlet algorithm is the central difference expression for the acceleration and it is the first expression in an infinite series in even powers of the time-increment.[9] All the higher-order algorithms in this expansion are in principle time-reversible and symplectic. Thus, all these algorithms have shadow Hamiltonians which rapidly converge to the Hamiltonian for the analytic solution. All the higher-order central difference algorithms depend, however, on positions as well as momenta at the discrete time where the positions are updated and this fact complicates their implementation.[9] So, although one can choose a larger h for these algorithms, there is a computational "overhead." In addition to the determination of the maximum time-increment for energy conservation this article also deals with the stability and effectiveness of these higher-order algorithms.

The discrete dynamics of a single normal mode can be derived exactly[6,10–12] and the shadow Hamiltonian can be obtained by an expansion in even powers of the frequency of the analytic dynamics ω for $h \leq 2/\omega$.[6] The expansion was used in Ref. 6 to determine the energy difference to first order between $\tilde{H}(\mathbf{q}, \mathbf{p})$ and $H(\mathbf{q}, \mathbf{p})$ for a complex system. The first-order term, derived in Ref. 6, is the harmonic approximation of the first term in the asymptotic expansion,[2–5] which ensures the existence of $\tilde{H}(\mathbf{q}, \mathbf{p})$ for sufficient small values of h and a conserved energy of the discrete dynamics for symplectic algorithms. Here it is shown that the stability of MD with the central difference algorithms can be estimated either from the conserved energies or from the instantaneous spectrum of normal frequencies in the complex systems. The

investigation makes use of the instantaneous normal-mode analysis (INM)[13] to determine the limit of stability of the algorithms by using the exact criterion for the stability of a discrete harmonic mode. Normal mode analysis has been very successful in investigations of viscous systems,[14–17] which is extremely time demanding to simulate and for which it is very important to optimize the MD simulations.

The article is organized as follows. Section II presents a new and simpler way to obtain the INM spectrum and the first energy term in the asymptotic expansion, and in Sec. III we use the INM spectrum and the exact relations obtained in Appendix B to determine the limit of stability for two representative systems, a Lennard-Jones system (LJ) and the Kob-Andersen system[18] for a viscous fluid. A summary and a discussion of the results are given in Sec. IV. The derivation of the fourth-order algorithm in the central difference expansion[9] is given in Appendix A and the exact stability limits for a harmonic mode[6] is derived in Appendix B.

II. THE ASYMPTOTIC EXPANSION OF THE SHADOW HAMILTONIAN

The stability of discrete dynamics for a harmonic mode given in Appendix B is used in the investigation of the stability of the discrete symplectic dynamics for complex systems. The Hessian $\mathbf{J}(\mathbf{r}_n^N)$ (the matrix of the second derivative of the potential energy $U(\mathbf{r}_n^N)$ of the N interacting particles at the nth discrete move) enters into the harmonic approximation of the first energy term in the asymptotic expansion[6,19] and the squares of the normal-mode frequencies, which are used in the INM,[13] are the eigenvalues of $\mathbf{J}(\mathbf{r}_n^N)$.

The harmonic approximation of the first term in the asymptotic expansion gives (Eq. (20) in Ref. 6)

$$E_n \simeq U(\mathbf{r}_i^N) + \frac{1}{2}\sum_i^N \frac{\mathbf{v}_i^2}{1-h^2\omega_i^2/4} - \sum_i^N h^2 A_i^2 \omega_i^4/24, \tag{1}$$

(masses included in the time unit), where ω_i is the frequency and A_i is the amplitude of the harmonic approximation of the ith particle's motion at time $t = nh$ at the nth discrete step. The "velocity" of particle no. i at time t in Eq. (1) is traditionally obtained from the positions at $t + h$ and $t - h$ as $\mathbf{v}_i \equiv (\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h))/2h$.

The square of the harmonic mode frequencies can consistently with the velocities be determined from the discrete values of forces and positions at $t + h$ and $t - h$ by

$$\omega_i^2(t) = -\frac{(\mathbf{f}_i(t+h) - \mathbf{f}_i(t-h)) \cdot \mathbf{e}_i}{|\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h)|}, \tag{2}$$

where $\mathbf{e}_i$ is the unit vector $\mathbf{e}_i(t) = (\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h))/|\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h)|$ along the ith particle's (discrete) trajectory at $\mathbf{r}_i(t)$. Correspondingly, the amplitude for the ith mass point is obtained as

$$A_i^2 = \left(\frac{\mathbf{f}_i}{\omega_i^2}\right)^2 + \left(\frac{\mathbf{v}_i}{\omega_i}\right)^2. \tag{3}$$

By inserting Eqs. (2) and (3) in Eq. (1) (and using $(1 - h^2\omega_i^2/4)^{-1} = 1 + h^2\omega_i^2/4 + \mathcal{O}(\omega^4)$), Eq. (1) transforms to

$$E_n \simeq U(\mathbf{r}_n^N) + \frac{1}{2}(\mathbf{v}_n^N)^2 + \frac{h^2}{12}(\mathbf{v}_n^N)^T \mathbf{J}(\mathbf{r}_n^N)(\mathbf{v}_n^N) - \frac{h^2}{24}\mathbf{f}_n^N(\mathbf{r}_n^N)^2, \tag{4}$$

(with $\mathbf{v}_n^N \equiv \mathbf{v}_1(t), \ldots, \mathbf{v}_N(t)$, $\mathbf{f}_n^N \equiv \mathbf{f}_1(t), \ldots, \mathbf{f}_N(t)$, and $\mathbf{r}_n^N \equiv \mathbf{r}_1(t), \ldots, \mathbf{r}_N(t)$) for the nth determination of the shadow energy E_n.[8,19] The first two terms in Eq. (4) are the traditional expression for the discrete value

$$E_{disc,n} = U(\mathbf{r}_n^N) + \frac{1}{2}(\mathbf{v}_n^N)^2 \tag{5}$$

of the total energy in the system. $E_{disc,\,n}$ fluctuates with n but it is conserved in mean with a relatively small standard deviation (SD)(h^2). Inclusion of the first (harmonic approximated) term in the asymptotic expansion in the energy reduces, however, the SD by a factor of hundred for a traditional choice of the time-increment.[8]

Figure 1 shows the record of the discrete obtained energies in a LJ system for a traditional choice of time-increment $h = 0.005$. Inclusion of the first term in the asymptotic expansion reduces the variation in energy by a factor of hundred and makes it possible to determine when the energy is conserved. The imaginary eigenvalues, corresponding to convex (saddle-point) transitions were ignored in Ref. 6. But, by using Eq. (2) these modes are included in the first term in the asymptotic expansion (Eq. (1)) and the standard deviation of E_n is reduced significantly and leads to a SD similar to the SD obtained from $\mathbf{J}(\mathbf{r}_n^N)$ in Eq. (4). Since the determination of the frequency spectrum by using Eq. (2) is much simpler than the determination of the eigenvalues of the Hessian and lead to almost the same value of the shadow energy (Figure 1) and conservation of energy per time step (Figure 2), it is used here in the INM analysis.

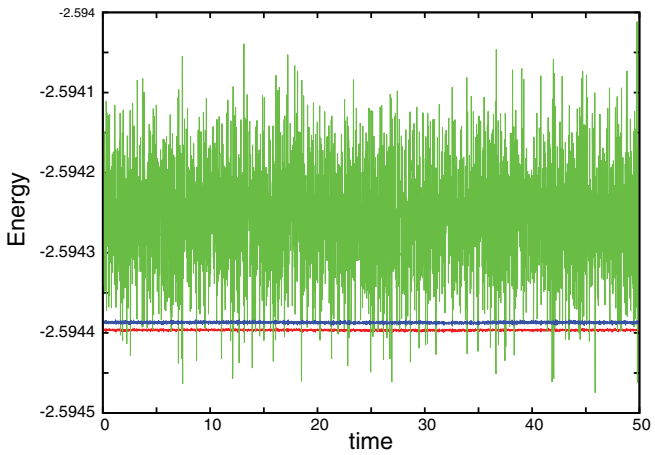

FIG. 1. Discrete determined energies of a LJ system with the Verlet algorithm at T, $\rho = 1.0$, 0.80 and for $h = 0.005$. Green: Traditional energy estimate obtained by Eq. (5); red: "Shadow" energy obtained by Eq. (4); blue "Shadow" energy obtained by Eqs. (1)–(3).

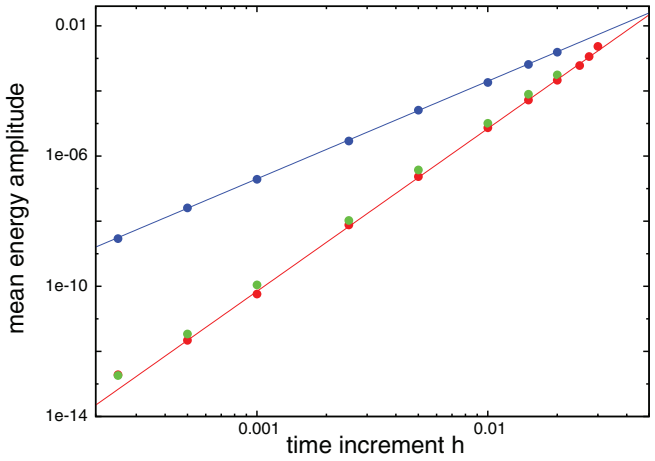

FIG. 2. The mean energy amplitudes $\langle|E_n - E_{n-1}|\rangle$ as a function of the time-increment h in a log − log plot. Red points are for the fourth-order algorithm (Eq. (A4)) (Appendix A) and green points are for the shadow energy of the Verlet algorithm (Eq. (4)). Blue points are for the Verlet algorithm with the traditional energy estimate (Eq. (5)). The straight lines through the red and blue points have the slopes of five and three, respectively.

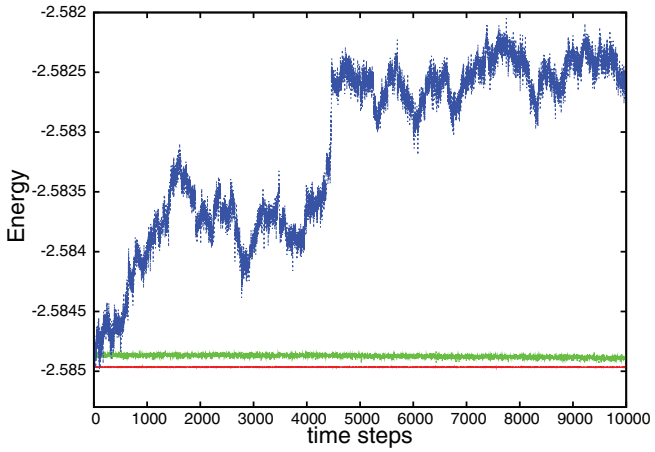

FIG. 3. The shadow energy per particle (Verlet algorithm) over 10 000 time steps for a LJ fluid at T, $\rho = 1.00, 0.80$, with shifted forces (SF) cutoff at $r_c = 2.5$ and for different time-increments h. Red: $h = 0.005$; green: $h = 0.01$; and blue: $h = 0.015$.

III. STABILITY OF DISCRETE DYNAMICS FOR COMPLEX SYSTEMS

The stability of discrete time-reversible dynamics is investigated for two systems: A system of particles interacting with the Lennard-Jones pair potential

$$u_{\mathrm{LJ}}(r) = 4\varepsilon[(\sigma/r)^{12} - (\sigma/r)^6], \ (r < r_c) \tag{6}$$

and a "Kob-Andersen binary mixture" (KABLJ). The KABLJ liquid is a mixture of 80% large (A) and 20% small (B) LJ particles ($\sigma_{\mathrm{BB}} = 0.8\sigma_{\mathrm{AA}}$) with a very strong AB attraction.[18] The A particles dominate the overall dynamics and the B particles are to a large extent slaves of the structure set by the A particles. Cut-offs of the interactions are given in units referring to σ_{AA},[20] computational MD details are given in Ref. 21. The system is very resistant against crystallization due to the strong exothermic mixture.[22] For this reason it is the standard system for simulations of highly viscous systems. Simulations of viscous systems are very time demanding and efficient MD is important.

The stability of the discrete dynamics for the LJ system is investigated at a typical condensed liquid state point with $T = 1.00$ at the density $\rho = N/V = 0.80$, with a cut-off $r_c = 2.5$. This system was the very first to be simulated[7] with a time-increment $h = 0.005$. The energy of the system is conserved for $h = 0.01$, but MD collapses sooner or later for $h = 0.02$. When the fourth-order algorithm, Eq. (A4), is used the system remains stable for $h \simeq 0.02$, but collapses for larger time-increments.

The discrete symplectic dynamics must be stable within the radius of convergence of the asymptotic expansion. But without knowledge of the higher-order coefficients in the asymptotic expansion we can only estimate the convergence limit. The dependence of the energy on the time-increment for the second-order Verlet algorithm, its first-order corrected shadow energy, and the fourth-order algorithm give, however, some indication of the limit of the expansion. Figure 2 shows the mean change in the energy per time step $\langle|E_n - E_{n-1}|\rangle$ as a function of h. The mean energy changes per step follow the predicted dependence (straight lines) and demonstrate that the first order correction to the shadow energy of the Verlet dynamics and the energy of the fourth-order algorithm Eq. (A4) (Appendix A) only improve the energy variation per time step for $h < 0.05$.

Although the behavior of the energy variation per time step indicates that the discrete dynamics might be stable for a significantly larger time-increment $h \approx 0.05$, the shadow energy of the Verlet dynamics is not conserved for $h > 0.01$. Figure 3 shows the (short time) energy evolution. The small growth of the energy appears instantaneously for $h = 0.015$ and is caused by ill-integrated modes (see below and Appendix B) and not by the truncated and shifted force field. (A truncated and shifted potential or a truncated and shifted force field only results in a very small drift in the energy which shows up for very long simulations[8,24]). The small drift was removed by a Nosé-Hoover thermostat, which maintains the time-reversibility and is "mean-symplectic."[23] This stabilized the MD without any change in energy, pressure, and the distribution of the particles. But the MD with the Verlet algorithm breaks down sooner or later for $h = 0.02$ even by employing a thermostat. For the fourth-order algorithm (Appendix A) the energy is conserved for $h \leq 0.02$ in accordance with the theory, which predicts an improvement in stability with respect to h with a factor of $\sqrt{3}$. These limits of stability are significantly less than the estimate $h < 0.05$, obtained from the energy conservation per time step (Figure 2). The reason for this behavior is that the spectrum of normal modes at the nth time step contains modes that cannot be integrated with $h > 0.01$ for the Verlet algorithm and correspondingly $h > 0.02$ for the fourth-order algorithm.

The distributions of harmonic frequencies for the LJ system at T, $\rho = 1.00, 0.80$ are shown in Figure 4. The first term in the asymptotic expansion depends only on the potential energy, explored by generated discrete positions. Thus, the instant distribution of the modes could depend on the time-increment, but the distribution is, however, not sensitive to h. The distribution of eigenvalues ω_i^2 is sorted in negative

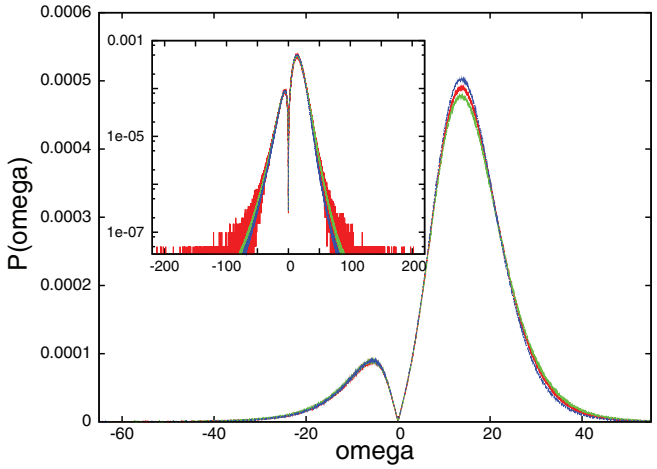

FIG. 4. Distribution $P(\omega)$ of normal modes with frequencies ω in a LJ-fluid at $T = 1.00$ and $\rho = 0.80$. Red lines show the distribution for a small time-increment $h = 0.0025$ and with green points the distributions for a time-increment $h = 0.015$ both obtained for the Verlet algorithm. The blue curve is the distribution obtained by the fourth-order algorithm (Eq. (A4)) (Appendix A) and for a time-increment $h = 0.020$. The inset shows the "instantaneous distribution" log $P(\omega)$ for $h = 0.0025$ (red) obtained from the first 10 000 steps together with the distributions for $h = 0.015$ (green) and $h = 0.02$ (blue), also shown in the main figure.

(imaginary "saddle-point transitions") and positive values, and the distribution of ω_i in Figure 4 shows both distributions. The distributions are obtained from one million time steps at T, $\rho = 1.00, 0.80$ and with a cut-off $r_c = 2.5$. The figure shows the distributions for $h = 0.0025$ (red) and $h = 0.015$ (green) for the Verlet algorithm and $h = 0.02$ (blue) for the fourth-order algorithm (Eq. (A4)) (Appendix A). (The two distributions for $h = 0.015$ and $h = 0.02$ are obtained by removing a small energy drift by a thermostat.[23]) The inset with a logarithmic ordinate shows the distribution of the modes for the first 10 000 steps with $h = 0.0025$ (red) together with the two mean distributions for $h = 0.015$ and $h = 0.02$ from the main figure. The main figure demonstrates that the distributions are not sensitive to h, even outside the stability region of the shadow energy. But the inset reveals that the instantaneous spectrum contains a few modes that cannot be integrated by the discrete dynamics. The instantaneous distribution in the inset (red) is obtained for $h = 0.0025$ and for 10 000 time steps, and it shows that there are modes within the intervals $[-200, -100]$ and $[100, 200]$. They appear with a very small probability, but there are a few of the $2000 \times 10\,000$ eigenvalues for the N = 2000 particles within the 10 000 time steps. According to the stability criterion (Eq. (B8)) (Appendix B), these modes require a maximum time-increment of $h = 0.01$ for the Verlet algorithm and this explains the instability of the energy, shown in Figure 3. The energy drift for $h = 0.015$ can be removed by a thermostat and the few unstable modes do not result in that the MD breaks down, nor in a shift in the distribution of normal modes. The few ill-integrated modes have no effect on the equilibrium behavior. In fact, the radial distribution function for $h = 0.015$ is equal to the distribution function for $h = 0.0025$, so the thermostat ensures that the thermodynamics is not affected by the mistreatment of these few unstable modes.

The general observation from simulations of the two systems (LJ and KABLJ) is that the distributions of frequencies are not sensitive to h as long as the shadow energy is conserved. At the LJ-state point T, $\rho = 1.00, 0.80$ frequencies above 50 are rare, but there are many frequencies $\omega \approx 100$ (inset in Figure 4). For a time-increment $h = 0.02$ the maximum frequency which can be integrated is 100 and the Verlet algorithm breaks down sooner or later, whereas the fourth-order algorithm remains stable in accordance with the stability criterion derived in Appendix B.

The KABLJ system is the MD standard system for simulation of a supercooled fluid system since it it very resistant against crystallization.[22] It was first simulated by Kob and Andersen[18] in the supercooled regime $T \approx 0.45$. The system was integrated (Verlet) with a time-increment $h = 0.0025$ and for shifted potentials[24] at a pair distance $r_c = 2.5$.[20] The energy of the KABLJ system at this state point remains, however, stable for $h \approx 0.01$ and for larger time-increments one can remove the small energy drift by a thermostat.

Figure 5 shows the distribution of normal modes obtained by the Verlet dynamics at T, $\rho = 0.45$, 1.2 and for different values of h. The distributions show some interesting features. First we notice that the imaginary modes are more important for the small solute B particles and second that the high frequency spectrum goes exponentially to zero with $|\omega_{max}| \approx 100$; but there appear a few frequencies outside this interval and these high frequency modes are more dominant for the small B particles (magenta lines in the inset). They appear for small time-increments as well as for the maximum increment $h = 0.015$, shown in the inset. For $h = 0.015$ these modes are ill-integrated according to the stability criterion (Eq. (B8)) (Appendix B). The energy is not conserved, but they are rare and have a marginal influence on the behavior of the viscous fluid.

The usefulness of MD is that one can determine the time behavior of a complex system, e.g., its transport coefficients. The viscosity η of the KABLJ fluid is inversely

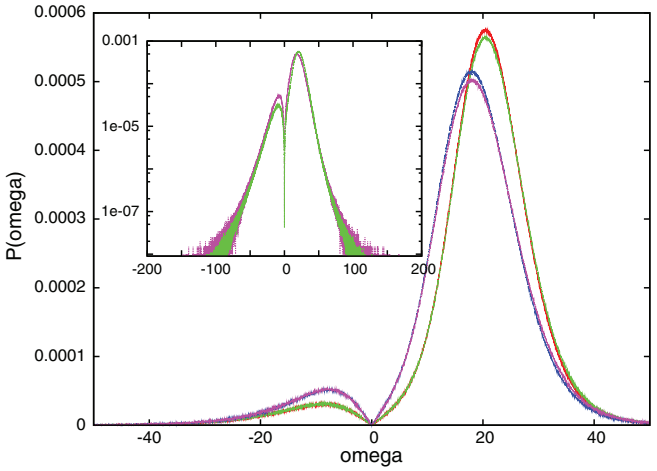

FIG. 5. Distribution $P(\omega)$ of normal modes with frequencies ω in a KABLJ-fluid at $T = 0.45$ and $\rho = 1.20$. The two distributions red and green are for the A-particles; blue and magenta for the B-particles. The red and blue lines are the distributions for a small time-increment $h = 0.0025$; the green and magenta lines that for $h = 0.015$. The distributions in the inset show log $P(\omega)$ for $h = 0.015$; green for A particles; and magenta for B particles.

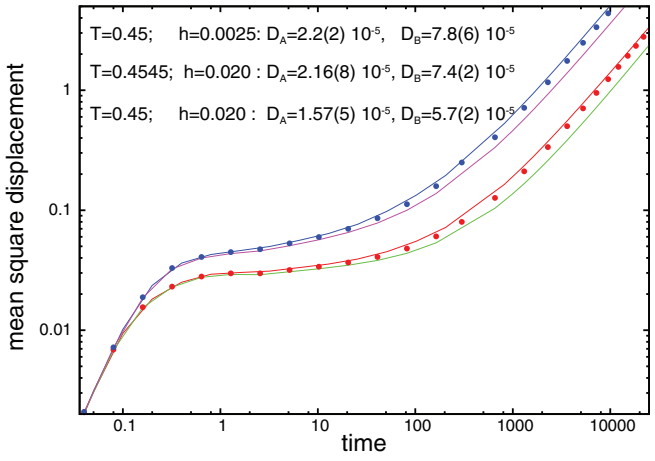

FIG. 6. The mean square displacement of particles in a KABLJ system at $\rho = 1.20$ in a $log - log$ plot. Red and green lines and red points are for A particles, blue and magenta lines and blue points are for B particles. Red line is for $h = 0.0025$ (standard value); green line is for $h = 0.02$, both at $T = 0.45$; red points are for $h = 0.02$ at $T = 0.4545$. Blue line is for $h = 0.0025$; magenta line is for $h = 0.02$, both at $T = 0.45$; blue points are for $h = 0.02$ at $T = 0.4545$. The values of the self-diffusion constants are given in the figure.

proportional to the self diffusion constant D and changes dramatically near the "mode-coupling" transition temperature $T_c = 0.435$.[18] The self diffusion constant D can be obtained as the slope of the mean square displacement (msd) of the particles, which is a linear function of time in the diffusive regime. The mean-square displacement for a particle in a viscous fluid separates into two regimes: the short-time ballistic regime in which the particle vibrates within its first coordination shell and the long-time diffusive regime that reflects the occasional escape from the shell and where the msd is a linear function of time. A crucial requirement for using a larger time-increment is that one obtains not only the same temperature, energy, and pressure, but also the same dynamical behavior.

It requires very long simulations to determine the msd accurately and thereby determine the self diffusion constant and the viscosity of viscous systems. The msd of the particles at $\rho = 1.20$ and $T \approx 0.45$ are shown in Fig. 6. The data shown in the figure is the mean of ten independent simulations, each of 1.6×10^7 steps, with different time-increments h. The values of the self-diffusion constants D_A and D_B are given in the figure together with the uncertainties of the last ciphers (in parentheses). The simulations reveal that not only the thermodynamics but also the mean square displacement can be obtained for a significantly larger time-increment. The radial distribution functions as well as the msd and D_A and D_B are independent of h within the (energy) stability regime $h \leq 0.01$. For $h = 0.015$ a small energy drift can be removed by a thermostat and all mean data are still unaffected, even within the very long simulations; but for $h = 0.02$ there are small changes in the msd (green and magenta lines) and D_A and D_B, and the fluid is more viscous. The MD for $h = 0.025$ breaks down sooner or later.

The higher viscosity of the system with $h = 0.02$ at $T = 0.45$ is, however, not a defect caused by too big a value of h, but a consequence of a shift in the shadow energy. The exact solution of the discrete dynamics of a harmonic mode (DDHO) in Appendix B shows that the shadow Hamiltonian and the shadow energy is (only) a function of h. The simulations with $h = 0.02$ (green and magenta lines in Figure 6) are performed with a thermostat at $T = 0.45$, which results in a slightly smaller shadow energy than obtained for $h = 0.0025$. The shadow energy of the system with $h = 0.02$ was also obtained for (slightly) higher temperatures from which the shadow energy at $T = 0.4545$ was estimated to be equal to the corresponding shadow energy with $h = 0.0025$ at $T = 0.45$. The red and blue points in Figure 6 are the msd for $h = 0.02$ and $T = 0.4545$ and agree with the corresponding msd (red and blue lines) for $h = 0.0025$ at $T = 0.45$ and show that the msd at a given density are governed not by the temperature, but by the shadow energy. For equal shadow energy there are no detectable differences in the equilibrium data as well as the msd, and the investigation indicates that one can gain a factor of eight in computer time by increasing the time-increment from $h = 0.0025$ to $h = 0.020$ without any detectable loss in accuracy.

IV. SUMMARY AND DISCUSSION

The conclusion of the investigation of the two systems (LJ and KABLJ) at various state points is that the equilibrium behavior as well as the mean square displacements for a given density only is a function of the shadow energy within the stability limit of the energy. The stability of the discrete dynamics can be determined from the limit of stability of the shadow energy or estimated from the INM spectrum of the normal frequencies.

The INM spectrum of frequencies depends only on the energy $U(\mathbf{r}^N)$ explored by the particles during the simulation and this spectrum is almost independent of h (Figures 4 and 5). This behavior is very useful when one shall chose an efficient time-increment for a new system: The distribution and the maximum frequency ω_{max} can be obtained from a short simulation with a small time-increment and used to determine the maximum time-increment $h \leq 2/\omega_{max}$[6,11] by which the system can be simulated. Alternatively one can determine the stability of the discrete dynamics from the sensitive measure of the stability of the shadow energy, Eq. (1) or Eq. (4).

The standard-second order (Verlet) algorithm used in MD is reversible and symplectic. There exist many higher order algorithms, e.g., Runge-Kutta and occasionally they are also used in MD of complex systems. Most of these algorithms are not symplectic; but there exist, e.g., symplectic Runge-Kutta algorithms.[25,26] The symplecticity ensures a conserved shadow energy for sufficient small time-increments and the higher-order algorithms give not only a more precise estimate of the energy, but they are also more stable. Investigation of the fourth-order central difference algorithm shows, however, that one only gains a factor of $\sqrt{3}$ in the range of stability and since the computational overhead in this case is at least a factor of two, it is from a computational point of view simpler and faster to use the second order Verlet algorithm. The same must be true also for the fifth-order symplectic Kunge-Kutta algorithm since it requires additional force calculations and almost all computer time is spent on calculating the forces within the many-body system. One might, however, avoid this

overhead of computer time of the forces by using a higher-order force-gradient algorithm.[27]

The limit of stability of the MD simulation can be extended with a factor of 2 behind the limit of stability of the energy by using a thermostat which removes the (energy) errors caused by the few ill-integrated modes. In the case of the KABLJ system the time-increment can be increased in total by a factor of eight from the traditional (conservative) choice $h = 0.0025$[18] to $h = 0.020$ without any significant impact on the equilibrium behavior and the mean square displacements (Figure 6), but the system collapses sooner or later for a larger time-increment. The energy conservation per time step (Figure 2) and the distribution of normal modes reveal, however, that the discrete dynamics might be obtained even for significantly larger time-increments than used here, if one can exclude the unstable modes with $\omega > 2/h$. They appear rarely and they have only a negligible influence on the mean distribution of the particles and the mean square displacements, but they destroy the simulations sooner or later.

ACKNOWLEDGMENTS

The author acknowledges useful discussions with Ole J. Heilmann and Jeppe C. Dyre. The centre for viscous liquid dynamics "Glass and Time" is sponsored by the Danish National Research Foundation (DNRF).

APPENDIX A: HIGHER-ORDER CENTRAL DIFFERENCE ALGORITHMS

Higher order central difference algorithms can be obtained by a Taylor expansion. The algorithm in Ref. 9 was obtained by a forward and backward Taylor expansion at time t. For a time-increment h and for particle i,

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + h\mathbf{r}_i'(t) + \frac{1}{2}h^2\mathbf{r}_i''(t) + \frac{1}{6}h^3\mathbf{r}_i'''(t) + \frac{1}{24}h^4\mathbf{r}_i''''(t) + \ldots, \tag{A1}$$

and

$$\mathbf{r}_i(t-h) = \mathbf{r}_i(t) - h\mathbf{r}_i'(t) + \frac{1}{2}h^2\mathbf{r}_i''(t) - \frac{1}{6}h^3\mathbf{r}_i'''(t) + \frac{1}{24}h^4\mathbf{r}_i''''(t) - \ldots. \tag{A2}$$

We consider a system of N particles with equal mass m. For pair interaction $u(r_{ij})$ with particle no. j the force is

$$\mathbf{f}_i(t) = -\sum_{j\neq i}\nabla u(r_{ij}). \tag{A3}$$

By adding Eqs. (A1) and (A2) one obtains the extended fourth-order central difference algorithm[9] for Newtonian dynamics ($\mathbf{r}_i''(t) = \mathbf{f}_i(t)$ and masses included in the time unit)

$$\mathbf{r}_i(t+h) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-h) + h^2\mathbf{f}_i(t) + \frac{1}{12}h^4\mathbf{f}_i''(t) + \mathcal{O}(h^6), \tag{A4}$$

where the algorithm to second order is the Verlet algorithm. It is computationally convenient to reformulate the algorithm in the "leap-frog" version, which corresponds to Hamilton's formulation of Newtonian dynamics

$$\mathbf{r}_i(t+h) = \mathbf{r}_i(t) + h\mathbf{v}(t+h/2), \tag{A5}$$

$$\mathbf{v}_i(t+h/2) = \mathbf{v}_i(t-h/2) + h\mathbf{f}_i(t) + \frac{1}{12}h^3\mathbf{f}_i''(t), \tag{A6}$$

with $\mathbf{v}_i(t-h/2) \equiv (\mathbf{r}_i(t) - \mathbf{r}_i(t-h))/h$.

The extension (Eq. (A4)) of the Verlet-algorithm is the next term in an infinite central-difference expansion in even powers of h. It is symplectic due to the symmetry and has a shadow Hamiltonian. The algorithm and all the higher order central difference algorithms deviate, however, from the simple second-order by that they depend on the momenta or velocities at time t. This can be seen by determining $\mathbf{f}_i''(t)$. By differentiating Eq. (A3) with respect to time one obtains the algorithm derived in Ref. 9

$$\mathbf{f}_i'(t) = \sum_{j\neq i}^{N} A(r_{ij}(t))\mathbf{r}_{ij}'(t) + B(r_{ij}(t))[\mathbf{r}_{ij}(t)\cdot\mathbf{r}_{ij}'(t)]\mathbf{r}_{ij}(t), \tag{A7}$$

$$\begin{aligned}\mathbf{f}_i''(t) = \sum_{j\neq i}^{N}\{&B(r_{ij}(t))[\mathbf{r}_{ij}(t)\cdot\mathbf{f}_{ij}(t) + \mathbf{r}_{ij}'(t)^2] \\ &+ C(r_{ij}(t))[\mathbf{r}_{ij}(t)\cdot\mathbf{r}_{ij}'(t)]^2\}\mathbf{r}_{ij}(t) \\ &+ 2B(r_{ij}(t))[\mathbf{r}_{ij}(t)\cdot\mathbf{r}_{ij}'(t)]\mathbf{r}_{ij}'(t) + A(r_{ij}(t))\mathbf{f}_{ij}(t),\end{aligned} \tag{A8}$$

where

$$A(r_{ij}) = -\frac{1}{r_{ij}}\frac{du(r_{ij})}{dr_{ij}}, \tag{A9}$$

$$B(r_{ij}) = \frac{1}{r_{ij}}\frac{dA(r_{ij})}{dr_{ij}}, \tag{A10}$$

$$C(r_{ij}) = \frac{1}{r_{ij}}\frac{dB(r_{ij})}{dr_{ij}}. \tag{A11}$$

The second time-derivative of the force $\mathbf{f}_i''(t)$ depends on $\mathbf{r}_i'(t)$. Upon subtracting Eq. (A2) from Eq. (A1) (and using Eq. (A4)) one obtains an expression for $\mathbf{r}_i'(t)$

$$\mathbf{r}_i'(t) = \frac{\mathbf{r}_i(t+h) - \mathbf{r}_i(t-h)}{2h} - \frac{1}{6}h^2\mathbf{f}_i'(t) + \mathcal{O}(h^4), \tag{A12}$$

$$= \frac{\mathbf{r}_i(t) - \mathbf{r}_i(t-h)}{h} + \frac{1}{2}h\mathbf{f}_i(t) - \frac{1}{6}h^2\mathbf{f}_i'(t) + \frac{1}{24}h^3\mathbf{f}_i''(t) + \mathcal{O}(h^4). \tag{A13}$$

The fourth-order algorithm (Eq. (A4)) requires not only the forces, but also its first two time-derivatives at time t in order to obtain the new positions. In Ref. 9 it was obtained by determining the forces at time t from which one then can

construct an estimate of $\mathbf{r}'_{pr,i}(t)$. Then $\mathbf{f}''_i(t)$ was determined by Eq. (A8) as a sum over pair-contributions and using the approximated expression $\mathbf{r}'_{pr,i}(t)$. Since most computer time is spent on determining the pair-contributions, the computer time is increased by a factor of two by including the four-order term. But also the time symmetry and the symplecticity are destroyed, with the consequences that the energy in principle is no longer conserved, and that there does not exist a shadow Hamiltonian. The energy is, however, in practice conserved with high accuracy for a traditional choice of h (see Sec. III).

To first order $\mathbf{r}'_i(t)$ is

$$\mathbf{r}'_{pr,i}(t) = \mathbf{v}_i(t - h/2) + \frac{1}{2}h\mathbf{f}_i(t), \tag{A14}$$

and to second order it can be obtained by approximating $\mathbf{f}'_i(t)$ in Eq. (A13) by $\mathbf{f}'_i(t) \simeq (\mathbf{f}_i(t) - \mathbf{f}_i(t-h))h + \frac{1}{2}h\mathbf{f}''_i(t-h)$,

$$\begin{aligned}\mathbf{r}'_i(t) &\simeq \mathbf{v}_i(t - h/2) + \frac{1}{2}h\mathbf{f}_i(t) - \frac{1}{6}h^2\mathbf{f}'_i(t)\\ &\simeq \mathbf{v}_i(t - h/2) + \frac{1}{6}h(2\mathbf{f}_i(t) + \mathbf{f}_i(t-h)) - \frac{1}{12}h^3\mathbf{f}''_i(t-h)).\end{aligned} \tag{A15}$$

Equation (A15) without the last term was used in Ref. 9. If one also includes the last term in Eq. (A15) one gets a third-order estimate of $\mathbf{r}'(t)$

$$\begin{aligned}\mathbf{r}'_{pr,i}(t) \simeq \mathbf{v}_i(t - h/2) &+ \frac{1}{6}h(2\mathbf{f}_i(t) + \mathbf{f}_i(t-h))\\ &- \frac{1}{24}h^3\mathbf{f}''_i(t-h)),\end{aligned} \tag{A16}$$

which is used in Sec. III to determine the stability of the algorithm.

APPENDIX B: DISCRETE DYNAMICS FOR A HARMONIC MODE

The discrete dynamics of a harmonic mode with the potential energy $\frac{1}{2}\omega_0^2x^2$ can be solved exactly.[6,10] The solution in Ref. 6 was obtained directly from the discrete points and without any use of an expansion of an analytic $H(q, p)$ by noticing that if the Verlet algorithm

$$x(t+h) = 2x(t) - x(t-h) - \omega_0^2h^2x(t) = \alpha x(t) - x(t-h), \tag{B1}$$

with $\alpha = 2 - \omega_0^2h^2$ for the DDHO is started from the two points $x(0) = 0$ and $x(h) = A_0\sin(\omega_0 h)$, the generated discrete points lie on a harmonic curve with the frequency ω and amplitude A given by

$$\begin{aligned}\omega &= \cos^{-1}\left(1 - \frac{(\omega_0 h)^2}{2}\right)\Big/h,\\ A &= \frac{A_0\sin(\omega_0 h)}{\sin(\omega h)},\end{aligned} \tag{B2}$$

i.e., the harmonic shadow Hamiltonian $\tilde{H}(\omega)$ for which the discrete generated positions lie on its analytic trajectory $\frac{1}{2}\omega^2x^2$ is $\tilde{H}(\omega)$ with the energy $\tilde{E} = (A\omega)^2/2$.

In Ref. 6, a conserved energy E^* of the discrete dynamics was obtained directly without the use of the analytic shadow Hamiltonian. It was expressed by scaling the discrete velocities, but it can also be expressed as

$$E^* = \frac{1}{2}\left(\frac{x_{n+1} - x_{n-1}}{2h}\right)^2 + \frac{1}{2}\omega_0^2x_n^2(1 - \omega_0^2h^2/4). \tag{B3}$$

Although there does not exist a well defined expression for the kinetic energy for the traditional discrete MD dynamics, the first term in the central difference expansion (Eq. (A12)), which corresponds to the first term in Eq. (B3), is traditionally used. For a discrete harmonic dynamics, started at $x(0)$, this expression for the kinetic energy corresponds to a temperature

$$kT = \left(\frac{x_1 - x_{-1}}{2h}\right)^2 = \left(\frac{A_0\sin(\omega_0 h)}{h}\right)^2 \tag{B4}$$

proportional to the square of the amplitude as in analytic dynamics.

The exact solution for the discrete dynamics with the higher-order central difference algorithms in Appendix A are easily obtained, since, e.g., the fourth-order algorithm reduces to Eq. (B1) with

$$\alpha = 2 - \omega_0^2h^2\left(1 - \frac{1}{12}\omega_0^2h\right)^2 = 2 - \tilde{\omega}_0^2h^2, \tag{B5}$$

replacing ω_0^2 in Eq. (B1) with

$$\tilde{\omega}_0^2 = \omega_0^2\left(1 - \frac{1}{12}\omega_0^2h^2\right). \tag{B6}$$

According to Eq. (B2) the algorithms are stable for a time-increment h_{max}

$$\left|1 - \frac{\tilde{\omega}_0^2h_{max}^2}{2}\right| \le 1. \tag{B7}$$

For the Verlet algorithm, this gives[6]

$$h_{max} \le \frac{2}{\omega_0}. \tag{B8}$$

For the fourth-order algorithm (Eq. (A4)),

$$h_{max} \le \frac{2\sqrt{3}}{\omega_0}. \tag{B9}$$

The solution $x(nh) = A\sin(\omega nh)$ is close to the analytic solution $x(nh) = A_0\sin(\omega_0 nh)$ if one integrates a harmonic mode with a time-increment $h \ll h_{max}$. But the shadow Hamiltonians and the shadow energies $\tilde{E}$ deviate significantly from the analytic solution when the time-increment is increased toward h_{max}.

The solutions of the DDHO for different values of h are shown in Figure 7. The figure shows the discrete solutions for a DDHO with $A_0 = 1$, $\omega_0 = 1$ and and for three different values of h. Figure 7(a) shows the discrete points for $h = 0.1$. The shadow energy $\tilde{E}$ for the Verlet integrator is $\tilde{E} = 0.5000014$ and for the fourth-order algorithm (Eq. (A4)) the shadow energy is $\tilde{E} = 0.4999999995$, and the differences between the three solutions are not visible on Figure 7(a).

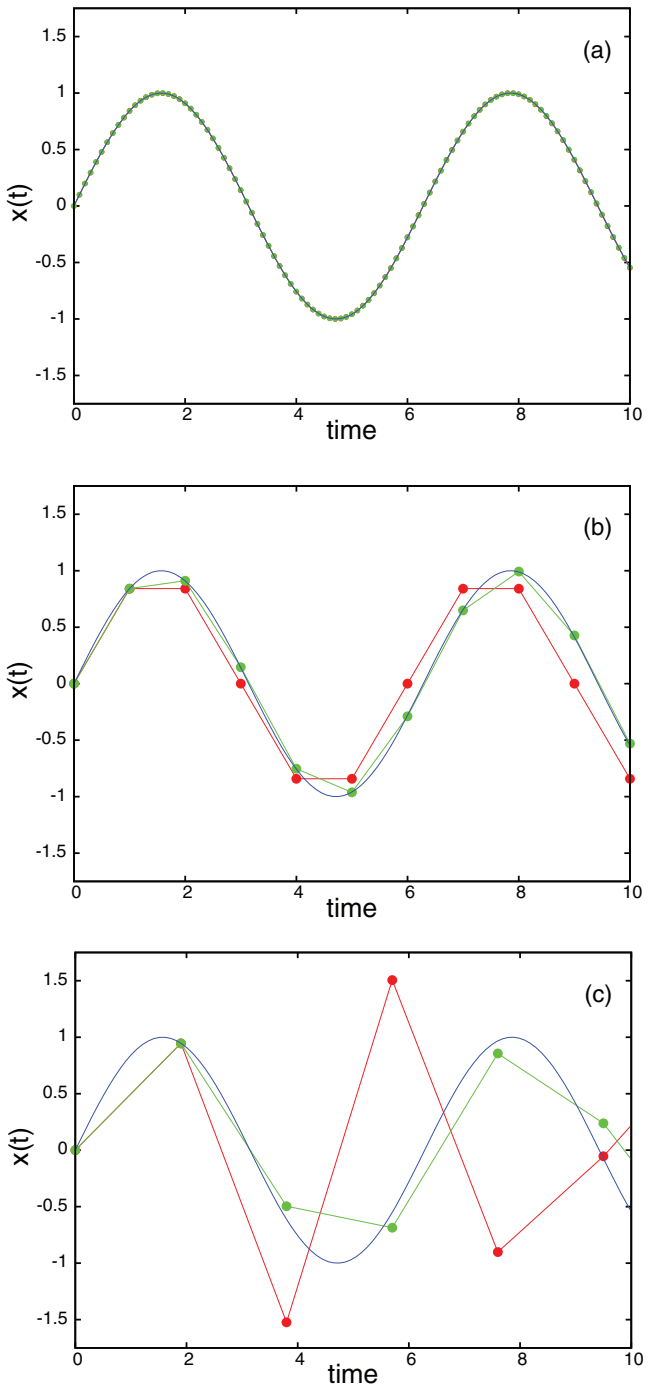

FIG. 7. The discrete points of a DDHO with $A_0 = \omega_0 = 1$. Red: Verlet; green: the fourth-order integrator; blue: the analytic solution. A is for $h = 0.1$; B is for $h = 1$; and C is for $h = 1.9 \approx h_{max}$.

Figure 7(b) is for $h = 1$; the shadow energy for the Verlet algorithm is $\tilde{E} = 0.518$ and for the fourth-order algorithm $\tilde{E} = 0.49942$. Figure 7(c) is for $h = 1.9$, close to the limit of stability $h_{max} = 2$. The shadow energy for the Verlet algorithm is $\tilde{E} = 2.214$, i.e., more than four times higher than the energy $E = 0.5$ for the analytic dynamics, whereas the energy obtained for the fourth-order algorithm is $\tilde{E} = 0.449$.

The shift in $\tilde{E}$ with h for time-increments near the stability limit is used in Sec. III to correct the mean square displacements for the observed shift in $\tilde{E}$ for large h.

[1]M. P. Allen, and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford Science, Oxford, 1987); D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic, New York, 2002).
[2]D. F. Griffiths and J. M. Sanz-Serna, SIAM (Soc. Ind. Appl. Math.) J. Sci. Stat. Comput. **7**, 994 (1986).
[3]J. M. Sanz-Serna, Acta Numerica **1**, 243 (1992).
[4]E. Hairer, Ann. Numer. Math. **1**, 107 (1994).
[5]S. Reich, SIAM (Soc. Ind. Appl. Math.) J. Numer. Anal. **36**, 1549 (1999).
[6]S. Toxvaerd, Phys. Rev. E **50**, 2271 (1994). The word *shadow Hamiltonian* was introduced in this paper, inspired by the terms *a slightly perturbed Hamiltonian* [H. Yoshida, Phys. Lett. A **150**, 262 (1990)] and *shadow trajectories* [C. Grebogi, S. M. Hammel, J. A. Yorke, and T. Saur, Phys. Rev. Lett. **65**, 1527 (1990)].
[7]L. Verlet, Phys. Rev. **159**, 98 (1967).
[8]S. Toxvaerd, O. J. Heilmann, and J. C. Dyre, J. Chem. Phys. **136**, 224106 (2012).
[9]S. Toxvaerd, J. Comput. Phys. **47**, 444 (1982).
[10]G. D. Venneri and W. G. Hoover, J. Comput. Phys. **73**, 468 (1987).
[11]For a review of the DDHO, see B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics* (Cambridge University Press, Cambridge, 2004), Chap. 2.6.
[12]E. Hairer, C. Lubich, and G. Wanner, *Geometrical Numerical Integration* (Springer, 2006), Chap. 1.5.2.
[13]G. Seeley and T. Keyes, J. Chem. Phys. **91**, 5581 (1989).
[14]S. D. Bembenek and B. B. Laird, J. Chem. Phys. **104**, 5199 (1996).
[15]T.-M. Wu, W.-J. Ma, and S. L. Chang, J. Chem. Phys. **113**, 274 (2000).
[16]E. D. Chisolm and D. C. Wallace, J. Phys. Condens. Matter **13**, R739 (2001).
[17]H. C. Andersen, Proc. Natl. Acad. Sci. U.S.A. **102**, 6686 (2005).
[18]W. Kob, and H. C. Andersen, Phys. Rev. E **51**, 4626 (1995); *ibid.* **52**, 4134 (1995).
[19]J. Gans and D. Shalloway, Phys. Rev. E **61**, 4587 (2000).
[20]The different potentials in Ref. 18 were cut at $r_c = 2.5\sigma_{\alpha,\beta}$, i.e., at varying lengths depending on the specific pair interaction.
[21]In MD the mass m is usually included in the time unit. The unit length, energy, and time used for LJ systems are, respectively, σ, ϵ, and $\sigma\sqrt{m/\epsilon}$.
[22]S. Toxvaerd and J. C. Dyre, J. Chem. Phys. **130**, 224501 (2009).
[23]The thermostat is a "Nose-Hoover" thermostat, see, e.g., S. Toxvaerd, Mol. Phys. **72**, 159 (1991).
[24]S. Toxvaerd and J. C. Dyre, J. Chem. Phys. **134**, 081102 (2011).
[25]J. M. Sanz-Serna, BIT **28**, 877 (1988).
[26]S. Reich, J. Comput. Phys. **157**, 473 (2000).
[27]I. P. Omelyan, I. M. Mryglod, and R. Folk, Phys. Rev. E **66**, 026701 (2002).