科学通报 2018 年 第 63 卷 第 33 期: 3467 ~ 3483
专题: 理论计算化学 论 文
《中国科学》杂志社
SCIENCE CHINA PRESS

一种基于蛙跳算法的高效分子动力学控温方法的统一方案

张志军, 严康宇, 刘歆子建, 刘剑*

北京大学化学与分子工程学院, 北京分子科学国家研究中心, 北京 100871
* 联系人, E-mail: jianliupku@pku.edu.cn

2018-09-01 收稿, 2018-10-11 修回, 2018-10-16 接受, 2018-11-08 网络版发表
国家重点基础研究发展计划(2016YFC0202803, 2017YFA0204901)、国家自然科学基金(21373018, 21573007)、中组部"青年千人计划"和NSFC-广东联合基金(第二期)超级计算科学应用研究专项资助(U1501501)资助

摘要 现代分子模拟的一个重要挑战是如何发展新的分子动力学算法保证演化稳定性的同时使得时间步长能尽量大, 在提高采样效率的同时还能保证采样所得分布精度. 本文基于蛙跳算法提出了蛙跳"居中"方案. 这是一种新的统一方案, 可以用来发展新的控温算法对正则系综高效地同时得到精确的位置边缘分布和精确的动量边缘分布. 由于蛙跳算法是一种常用算法, 在一些分子动力学模拟软件中常被默认采用, 这一统一控温方案能为这些软件提供方便高效的新算法来对实际分子体系做经典(或量子)正则系综的模拟.

关键词 正则系综, "居中"方案, 控温算法, 位置空间采样, 动量边缘分布

自从Fermi等人[1]在1955年的开创性工作以来, 分子动力学(molecular dynamics, MD)已经成为用来模拟和预测物理、化学、生物、材料、环境科学中各种复杂体系的各种热力学和动力学物理量的一个有力工具[2,3]. 一种常用的热力学平衡系综是正则系综(canonical ensemble). 它是粒子数N, 体积V和温度T这3个物理参数保持不变的热力学平衡体系的统称, 又称NVT系综.

当原子核可以用经典力学来描述时, 依据牛顿力学方程的分子动力学结合合适的控温算法可以得到经典的正则系综分布, 即NVT三个物理量保持不变的经典Boltzmann分布:

$$\frac{1}{Z}\exp\left[-\beta U(\boldsymbol{x})\right]\rho_{\mathrm{MB}}(\boldsymbol{p}), \tag{1}$$

其中

$$\rho_{\mathrm{MB}}(\boldsymbol{p})=\left(\frac{\beta}{2\pi}\right)^{3N/2}\left|\boldsymbol{M}\right|^{-1/2}\exp\left[-\beta\left(\frac{1}{2}\boldsymbol{p}^{T}\boldsymbol{M}^{-1}\boldsymbol{p}\right)\right], \tag{2}$$

是Maxwell动量分布. 式(1)或(2)中, $\beta=1/k_{\mathrm{B}}T$(k_{B}是Boltzmann常数), $\boldsymbol{M}$是(对角)质量矩阵, Z是配分函数, $\boldsymbol{p}$和$\boldsymbol{x}$分别是动量和位置, 而$U(\boldsymbol{x})$是这个系统的势能函数. 常用的控温方法分为随机控温方法(如Andersen控温方法[4]和Langevin(动力学)控温方法[5~17]等)和决定性控温方法(如Nosé-Hoover控温方法[18,19]和Nosé-Hoover Chain(NHC)控温方法[20,21]等).

当原子核的量子效应(nuclear quantum effects)变得不可忽略甚至极为重要时, 基于虚时间(imaginary time)费曼路径积分(Feynman's path integral)量子力学框架[22,23]下的同构(isomorphism)关系可以将一个量子体系映射到一个经典环形高聚物(ring polymer)模型[24]. 分子动力学可以用来模拟这个环形高聚物模型, 从而得到对应的真实分子体系的量子正则系综分布的采样[25]. 这种理论方法称为路径积分分子动力学(path integral molecular dynamics, PIMD)[26,27].

然而计算机模拟分子动力学的时间步长Δt总是

引用格式: 张志军, 严康宇, 刘歆子建, 等. 一种基于蛙跳算法的高效分子动力学控温方法的统一方案. 科学通报, 2018, 63: 3467–3483
Zhang Z J, Yan K Y, Liu X Z J, et al. A leap-frog algorithm-based efficient unified thermostat scheme for molecular dynamics (in Chinese). Chin Sci Bull, 2018, 63: 3467–3483, doi: 10.1360/N972018-00908

© 2018《中国科学》杂志社

有限小的. 过小的Δt会导致分子动力学对正则分布(Boltzmann分布)的采样效率过低, 而过大的Δt又会使得采样所得到的分布精度过低, 甚至破坏分子动力学轨线的演化. 一个重要挑战无疑是如何发展新的分子动力学算法使得时间步长Δt能尽量大, 在提高采样效率的同时能保证采样所得分布精度. 在经典力学框架下, 所有结构性质和大部分热力学物理量都只和位置(分子构型)有关. 在PIMD中, 所有物理量都只和路径积分的对应环形高聚物构型有关. 由此可见, 在MD或PIMD模拟中, 高效地得到精确的位置分布往往是一个关键[15,28].

1 控温算法的统一理论框架

在近期的一系列工作[15,28~31]中, 我们提出了一个理论框架来统一以前提出的大部分控温算法, 并发展了一种新的统一控温方案——"居中"控温方案来实现位置空间的高效精确采样. 很多控温算法将一个时间步长内的分子动力学演化分割为3个部分: 更新位置、更新动量和控温步骤. "运动方程"可以记为

$$\begin{pmatrix} \mathrm{d}\boldsymbol{x} \\ \mathrm{d}\boldsymbol{p} \end{pmatrix} = \underbrace{\begin{pmatrix} \boldsymbol{M}^{-1}\boldsymbol{p} \\ 0 \end{pmatrix}\mathrm{d}t}_{x} + \underbrace{\begin{pmatrix} 0 \\ -\dfrac{\partial U(\boldsymbol{x})}{\partial \boldsymbol{x}} \end{pmatrix}\mathrm{d}t}_{p} + \underbrace{\text{控温}}_{T}. \tag{3}$$

用式(3)来做理论分析并不方便. 一个更合适的做法是写出和式(3)对应的相空间分布函数$\rho(\boldsymbol{x}, \boldsymbol{p})$的Kolmogorov向前方程(forward equation):

$$\frac{\partial \rho}{\partial t} = \mathcal{L}\rho = \left(\mathcal{L}_x + \mathcal{L}_p + \mathcal{L}_T\right)\rho, \tag{4}$$

其中, $\mathcal{L}_x$和$\mathcal{L}_p$分别是与式(3)中右边第1和第2项对应的Kolmogorov算符:

$$\mathcal{L}_x\rho = -\boldsymbol{p}^T\boldsymbol{M}^{-1}\frac{\partial \rho}{\partial \boldsymbol{x}}, \tag{5}$$

$$\mathcal{L}_p\rho = \left(\frac{\partial U}{\partial \boldsymbol{x}}\right)^T\frac{\partial \rho}{\partial \boldsymbol{p}}, \tag{6}$$

而$\mathcal{L}_T$是式(3)中右边控温方法对应的Kolmogorov算符. $\mathcal{L}_T$取决于具体的控温方法, 例如, Andersen控温方法有:

$$\mathcal{L}_T\rho = \nu\left[\rho_{\mathrm{MB}}(\boldsymbol{p})\int_{-\infty}^{\infty}\rho(\boldsymbol{x}, \boldsymbol{p})\mathrm{d}\boldsymbol{p} - \rho(\boldsymbol{x}, \boldsymbol{p})\right], \tag{7}$$

式中, ν是粒子与热库的碰撞频率[28]. Langevin控温方法有:

$$\mathcal{L}_T\rho = \frac{\partial}{\partial \boldsymbol{p}}\cdot(\gamma\boldsymbol{p}\rho) + \frac{1}{\beta}\frac{\partial}{\partial \boldsymbol{p}}\cdot\left(\gamma\boldsymbol{M}\frac{\partial \rho}{\partial \boldsymbol{p}}\right), \tag{8}$$

式中, γ是摩擦系数[13,15,28]. 利用式(5)~(8)中的Kolmogorov算符, 可以写下对应的分布函数相空间演化算符:

$$e^{\mathcal{L}_x\Delta t}\rho(\boldsymbol{x}, \boldsymbol{p}) = \rho(\boldsymbol{x} - \boldsymbol{M}^{-1}\boldsymbol{p}\Delta t, \boldsymbol{p}), \tag{9}$$

$$e^{\mathcal{L}_p\Delta t}\rho(\boldsymbol{x}, \boldsymbol{p}) = \rho\left(\boldsymbol{x}, \boldsymbol{p} + \frac{\partial U}{\partial \boldsymbol{x}}\Delta t\right). \tag{10}$$

以及对Andersen控温方法有:

$$e^{\mathcal{L}_T\Delta t}\rho = (1 - e^{-\nu\Delta t})\rho_{\mathrm{MB}}(\boldsymbol{p})\int_{-\infty}^{\infty}\rho(\boldsymbol{x}, \boldsymbol{p})\mathrm{d}\boldsymbol{p} + e^{-\nu\Delta t}\rho(\boldsymbol{x}, \boldsymbol{p}). \tag{11}$$

或对Langevin控温方法有:

$$e^{\mathcal{L}_T\Delta t}\rho = \left[\frac{\beta}{2\pi(1 - e^{-2\gamma\Delta t})}\right]^{3N/2}|\boldsymbol{M}|^{-1/2}\int \mathrm{d}\boldsymbol{p}_0\rho(\boldsymbol{x}, \boldsymbol{p}_0)\exp\left[-\frac{\beta}{2(1 - e^{-2\gamma\Delta t})}\left(\boldsymbol{p} - e^{-\gamma\Delta t}\boldsymbol{p}_0\right)^T\boldsymbol{M}^{-1}\left(\boldsymbol{p} - e^{-\gamma\Delta t}\boldsymbol{p}_0\right)\right]. \tag{12}$$

从式(11)或(12)容易导出, 无论时间步长Δt多大,

$$e^{\mathcal{L}_T\Delta t}\rho_{\mathrm{MB}}(\boldsymbol{p}) = \rho_{\mathrm{MB}}(\boldsymbol{p}), \tag{13}$$

都成立, 即式(3)的控温步骤不改变正确的Maxwell动量分布[15,28]. 这是本课题组[28]指出的一个好的控温方法应当满足的条件.

利用Kolmogorov算符可以把大部分文献中的分子动力学控温算法(和PIMD算法)统一归为下面两类:

(1) "居后"方案:

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}^{\mathrm{End}}\Delta t} = e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_p\Delta t/2}, \tag{14}$$

例如, 1980年提出的Andersen分子动力学控温算法[4]和1984年提出的PIMD算法[32].

(2) "两边"方案:

$$e^{\mathcal{L}\Delta t} \approx e^{\mathcal{L}^{\mathrm{Side}}\Delta t} = e^{\mathcal{L}_T\Delta t/2}e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_T\Delta t/2}, \tag{15}$$

例如, 基于NHC控温方法的MD算法[20,21]和PIMD算法[33]; 基于Langevin控温方法的MD算法[10]和PIMD算法[14].

以往研究中符合"居后"方案(式(14))的控温算法比符合"两边"方案(式(15))的控温算法提出要更早, 也通常认为后者要比前者要更精确. 有趣的是, 本课题组[28]在近期的工作中证明对于一般体系, 两者实际上会导致完全相同的位置边缘分布, 而且"居后"方案(式(14))要比"两边"方案(式(15))得到更准确的动量边缘分布. 对于谐振子体系:

$$U(\boldsymbol{x}) = \left(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{eq}}\right)^T\boldsymbol{A}\left(\boldsymbol{x} - \boldsymbol{x}_{\mathrm{eq}}\right)/2. \tag{16}$$

两种方案会得到同样的稳态相空间分布:

$$\rho_{\mathrm{eq}}^{\mathrm{Side/End}}=\frac{1}{Z_N}\exp\left[-\beta\left(\frac{1}{2}\boldsymbol{p}^T\boldsymbol{M}^{-1}\boldsymbol{p}+\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)^T\left(\boldsymbol{1}-\boldsymbol{A}\boldsymbol{M}^{-1}\frac{\Delta t^2}{4}\right)\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)\right)\right],\tag{17}$$

其中Z_N是归一化常数[28,29].

本课题组[28]提出了一种新的统一控温方案, 即"居中"方案:

$$e^{\mathcal{L}\Delta t}\approx e^{\mathcal{L}_{\mathrm{VV}}^{\mathrm{Middle}}\Delta t}=e^{\mathcal{L}_p\Delta t/2}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t/2},\tag{18}$$

"居中"方案比前两种控温方案在位置空间的采样要更高效. 其稳态相空间分布满足

$$\rho_{\mathrm{eq}}^{\mathrm{VV\text{-}Middle}}=e^{\mathcal{L}_{\mathrm{VV}}^{\mathrm{Middle}}\Delta t}\rho_{\mathrm{eq}}^{\mathrm{VV\text{-}Middle}}.\tag{19}$$

对于谐振子体系(式(16)), "居中"方案所得到的稳态相空间分布是

$$\rho_{\mathrm{eq}}^{\mathrm{VV\text{-}Middle}}=\frac{1}{\bar{Z}_N}\exp\left[-\beta\left(\frac{1}{2}\boldsymbol{p}^T\left(\boldsymbol{M}-\boldsymbol{A}\frac{\Delta t^2}{4}\right)^{-1}\boldsymbol{p}+\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)^T\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)\right)\right],\tag{20}$$

其中, $\bar{Z}_N$是归一化常数[15,28,29]. 式(20)表明谐振子极限下无论时间步长Δt多大, 只要动力学演化稳定(矩阵$\boldsymbol{M}-\boldsymbol{A}\frac{\Delta t^2}{4}$是正定的, 或者$\Delta t<2/\omega_{\max}$, $\omega_{\max}$为体系最大特征频率), 在控温算法满足式(13)的情况下, "居中"方案都能给出真实的位置空间分布[15,28,29]. 本课题组[15,28~31]近期一系列文章的各种实际分子体系(非谐体系)的测试结果也表明, 无论是随机还是决定性控温方法, "居中"方案无疑可以导出更优越的分子动力学控温算法来模拟经典力学中只与位置(构型)有关的热力学物理量和量子统计力学中的所有热力学物理量. "居中"方案[28~30]自然地涵盖了最近基于Langevin控温方法的高效分子动力学算法[11,13]和PIMD高效算法[15], 而且更重要的是提供了一个统一理论方案来发展其他控温方法的高效算法[28,30]. 例如, 我们指出了基于Langevin控温方法的Grønbech-Jensen/Farago分子动力学算法[11]和Leimkuhler/Matthews分子动力学算法(BAOAB)[13]对于实动力学(real dynamics)其实有等价关系[29], 而且前者更广泛还包括了虚动力学(virtual dynamics)的情况[29]; 无论是实动力学[11,13,29]还是虚动力学[11,29]的算法都可以在"居中"方案中推广到其他控温方法[30].

式(14), (15)和(18)这3种方案建立在速度-Verlet (velocity-Verlet)算法之上. 本课题组[28,29]还比较了建立在位置-Verlet(position-Verlet)算法上的各种控温方案, 结果表明基于速度-Verlet算法的"居中"方案(式(18))是这些控温方案之中能高效地得到精确位置分布的最佳方案. 在同一种方案中, 各种控温方法得到的结果都极其接近. 这清楚地表明, 控温方案是更具有决定性的因素[28,30].

本文在上述采用Kolmogorov算符建立的统一理论框架下继续对控温算法进行探讨. 第二部分先介绍分子动力学中常用的蛙跳(leap-frog)算法, 然后提出与蛙跳算法相容的一种新的控温方案用以导出高效控温算法, 使得既能得到精确的位置边缘分布又能得到精确的动量边缘分布. 最后给出本文结论.

2 基于蛙跳算法的高效统一控温方案

2.1 蛙跳"居中"控温方案

2.1.1 无约束系统

和速度-Verlet算法同样常用的另外一种分子动力学的保辛(symplectic)算法是蛙跳算法. 在用于微正则系综(NVE系综, E是体系总能量)的蛙跳算法中, 其第n+1个时间步长Δt内, 动量和位置的更新步骤如下:

$$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\leftarrow\boldsymbol{p}\left(\left(n-\frac{1}{2}\right)\Delta t\right)-U'\left(\boldsymbol{x}(n\Delta t)\right)\Delta t,\tag{21}$$

$$\boldsymbol{x}\left((n+1)\Delta t\right)\leftarrow\boldsymbol{x}(n\Delta t)+\boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\Delta t.\tag{22}$$

给定t=0时刻的相空间分布函数$\rho_0(\boldsymbol{x},\boldsymbol{p})$, 那么第$n$+1个时间步长$\Delta t$后得到的位置$\boldsymbol{x}((n+1)\Delta t)$和动量$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$所构成的相空间分布函数$\rho_{n+1}(\boldsymbol{x},\boldsymbol{p})$满足:

$$\rho_{n+1}(\boldsymbol{x},\boldsymbol{p})=\left(e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_p\Delta t}\right)^n e^{\mathcal{L}_x\Delta t}e^{\mathcal{L}_p\Delta t/2}\rho_0(\boldsymbol{x},\boldsymbol{p}).\tag{23}$$

对于正则系综(NVT系综), 我们构建蛙跳"居中"方案如下:

$$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\leftarrow\boldsymbol{p}\left(\left(n-\frac{1}{2}\right)\Delta t\right)-U'\left(\boldsymbol{x}(n\Delta t)\right)\Delta t,\tag{24}$$

$$\boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\leftarrow\boldsymbol{x}(n\Delta t)+\boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2},\tag{25}$$

控温整个时间步长Δt:改变动量$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$,　(26)

$$\boldsymbol{x}((n+1)\Delta t) \leftarrow \boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right)+\boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2}, \tag{27}$$

其中控温步骤(式(26))依赖于具体的控温方法. 例如, 对于Langevin控温方法, 式(26)的具体形式为

$$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\leftarrow e^{-\gamma\Delta t}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)+\sqrt{\frac{1-e^{-2\gamma\Delta t}}{\beta}}\boldsymbol{M}^{1/2}\boldsymbol{\eta}, \tag{28}$$

其中$\boldsymbol{\eta}$是个维度为3N的向量. 这个向量的每一个元素都是一个独立的标准高斯分布随机数, 每一个维度都不同, 每一次调用式(28)时也不同. 对于Andersen控温方法, 式(26)的具体形式则是

$$\boldsymbol{p}^{(j)}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\leftarrow\sqrt{\frac{1}{\beta}}\boldsymbol{M}_j^{1/2}\boldsymbol{\theta}_j,$$
$$\text{如果 } \mu_j<1-e^{-\nu\Delta t}\ (j=1,2,\cdots,N), \tag{29}$$

其中 $\boldsymbol{p}^{(j)}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$和$\boldsymbol{M}_j$是第$j$个粒子的三维动量向量和对角质量矩阵. μ_j是满足范围(0,1)内的均匀分布的一个随机数, 对于不同粒子, 每次调用式(29)也不同. $\boldsymbol{\theta}_j$是个三维向量, 它的每一个元素都是一个独立的标准高斯分布随机数, 每一个维度都不同, 每一次调用式(29)时也不同. 需要着重强调的是控温步骤(式(26))中的控温算法可以是任意一种(解析上或数值精度上)满足式(13)的随机或决定性控温算法, 所以式(24)~(27)提供了一个基于蛙跳算法上的统一控温方案. 这本质上其实是我们采用Kolmogorov算符建立的统一理论框架[28~30]中的一种新的控温方案形式.

考虑由位置$\boldsymbol{x}((n+1)\Delta t)$和动量 $\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$ 所构成的相空间. 容易证明由式(24)~(27)组成的算法所导致的相空间分布函数的演化满足:

$$\rho_{n+1}(\boldsymbol{x},\boldsymbol{p})=\left(e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t}\right)^n e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t/2}\rho_0(\boldsymbol{x},\boldsymbol{p}). \tag{30}$$

定义

$$\rho_1(\boldsymbol{x},\boldsymbol{p})=e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t/2}\rho_0(\boldsymbol{x},\boldsymbol{p}), \tag{31}$$

和

$$e^{\mathcal{L}_{\mathrm{LF}}^{\mathrm{Middle}}\Delta t}=e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_T\Delta t}e^{\mathcal{L}_x\Delta t/2}e^{\mathcal{L}_p\Delta t}. \tag{32}$$

式(30)变为

$$\rho_{n+1}(\boldsymbol{x},\boldsymbol{p})=\left(e^{\mathcal{L}_{\mathrm{LF}}^{\mathrm{Middle}}\Delta t}\right)^n\rho_1(\boldsymbol{x},\boldsymbol{p}). \tag{33}$$

该算法的稳态相空间分布函数 $\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p})$ 可以由式(33)在$n\to\infty$的极限下求得. 式(33)可直接导出

$$\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p})=e^{\mathcal{L}_{\mathrm{LF}}^{\mathrm{Middle}}\Delta t}\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p}). \tag{34}$$

由式(18), (19), (32)和(34)可得

$$e^{\mathcal{L}_p\Delta t/2}\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p})=\rho_{\mathrm{eq}}^{\mathrm{VV\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p}). \tag{35}$$

由式(35)和(10)可得

$$\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x},\boldsymbol{p})=\rho_{\mathrm{eq}}^{\mathrm{VV\text{-}Middle}}\left(\boldsymbol{x},\boldsymbol{p}-U'(\boldsymbol{x})\frac{\Delta t}{2}\right). \tag{36}$$

对式(36)两边对动量空间做积分得到

$$\rho_{\mathrm{eq},\,x}^{\mathrm{LF\text{-}Middle}}(\boldsymbol{x})=\rho_{\mathrm{eq},\,x}^{\mathrm{VV\text{-}Middle}}(\boldsymbol{x}). \tag{37}$$

即对于任何体系, 速度-Verlet“居中”方案和蛙跳“居中”方案各自所得稳态分布的位置边缘分布完全相同. 注意对式(36)右边对动量的积分可以引入变量替换 $\boldsymbol{p}\leftarrow\boldsymbol{p}-U'(\boldsymbol{x})\frac{\Delta t}{2}$.

特别地, 对于谐振子体系(式(16)), 蛙跳“居中”方案所得到的稳态相空间分布是

$$\rho_{\mathrm{eq}}^{\mathrm{LF\text{-}Middle}}=\frac{1}{\overline{\overline{Z}}_N}\exp\left[-\beta\left(\frac{1}{2}\left(\boldsymbol{p}-\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)\frac{\Delta t}{2}\right)^T\left(\boldsymbol{M}-\boldsymbol{A}\frac{\Delta t^2}{4}\right)^{-1}\left(\boldsymbol{p}-\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)\frac{\Delta t}{2}\right)+\frac{1}{2}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)^T\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}_{\mathrm{eq}}\right)\right)\right], \tag{38}$$

其中, $\overline{\overline{Z}}_N$ 是归一化常数. 容易验证式(38)所得到的位置边缘分布是准确的, 与此同时式(38)所得到的动量边缘分布也是准确的. 例如, 由式(38)易得, 在谐振子极限下蛙跳“居中”方案得到准确的动能平均值, 即

$$\left\langle\frac{1}{2}\boldsymbol{p}^T\boldsymbol{M}^{-1}\boldsymbol{p}\right\rangle_{\mathrm{LF\text{-}Middle}}=\frac{3N}{2\beta}. \tag{39}$$

同时, 所得势能平均值和哈密顿量(总能量)平均值也是准确的. 与之相对比, 速度-Verlet“居中”方案所得稳态分布(式(20))可以导出准确的位置边缘分布, 却不能得到准确的动量边缘分布. 在谐振子极限下, 速度-Verlet“居中”方案所得势能平均值是准确的, 但动能(或哈密顿量)平均值的误差随时间步长Δt的平方增长.

考虑哈密顿量的涨落 $f_H\equiv\sqrt{\langle H^2\rangle-\langle H\rangle^2}$ (与正则系综的比热相关). 这个物理量包含位置和动量的交叉项, 其计算精度依赖于相空间分布的精度. 其势能

为式(16)的一维形式:

$$U(x)=\frac{1}{2}m\omega^2(x-x_{\mathrm{eq}})^2. \tag{40}$$

由式(20)或(38)可得蛙跳或速度-Verlet“居中”方案所得的哈密顿量的涨落为

$$f_H^{\text{LF-Middle}}=\frac{1}{\beta}\left(1+\frac{\omega^2\Delta t^2}{4}\right)^{1/2}, \tag{41}$$

$$f_H^{\text{VV-Middle}}=\frac{1}{\beta}\left(1+\frac{\omega^2\Delta t^2}{2}-\frac{\omega^4\Delta t^4}{64}\right)^{1/2}, \tag{42}$$

比较式(41)和(42), 由稳定性条件$\omega\Delta t<2$可以得到$f_H^{\text{LF-Middle}}<f_H^{\text{VV-Middle}}$, 即对于一维谐振子体系的哈密顿量的涨落的计算, 蛙跳“居中”方案比速度-Verlet“居中”方案导出更精确的结果.

再考虑能流(energy current)的涨落(与导热系数有关), 这个量同样包含位置和动量的交叉项. 对于两粒子体系, 能流为

$$\boldsymbol{S}=\frac{\boldsymbol{p}_1}{m_1}\left(\frac{|\boldsymbol{p}_1|^2}{2m_1}+\frac{1}{2}V(\boldsymbol{x}_1-\boldsymbol{x}_2)\right)+\frac{\boldsymbol{p}_2}{m_2}\left(\frac{|\boldsymbol{p}_2|^2}{2m_2}+\frac{1}{2}V(\boldsymbol{x}_1-\boldsymbol{x}_2)\right)$$
$$-\frac{1}{2}(\boldsymbol{x}_1-\boldsymbol{x}_2)\frac{\partial V}{\partial \boldsymbol{x}_1}\cdot\left(\frac{\boldsymbol{p}_1}{m_1}+\frac{\boldsymbol{p}_2}{m_2}\right). \tag{43}$$

考虑两个一维粒子相互作用势$V(x_1-x_2)$是一个平衡相对位置为r_{eq}的谐振子, 即

$$V(x_1-x_2)=\frac{1}{2}\mu\omega^2(x_1-x_2-r_{\mathrm{eq}})^2, \tag{44}$$

其中μ是约化质量$\mu=m_1m_2/(m_1+m_2)$. 能流(式(43))则可写为

$$S=\frac{p_1^3}{2m_1^2}+\frac{p_2^3}{2m_2^2}-\frac{1}{4}\mu\omega^2(x_1-x_2+r_{\mathrm{eq}})(x_1-x_2-r_{\mathrm{eq}})\left(\frac{p_1}{m_1}+\frac{p_2}{m_2}\right). \tag{45}$$

采用简正坐标

$$X_1=\frac{m_1x_1+m_2x_2}{\sqrt{m_1+m_2}},\ X_2=\sqrt{\mu}(x_1-x_2-r_{\mathrm{eq}}), \tag{46}$$

和简正动量

$$P_1=\frac{p_1+p_2}{\sqrt{m_1+m_2}},\ P_2=\sqrt{\mu}\left(\frac{p_1}{m_1}-\frac{p_2}{m_2}\right). \tag{47}$$

这个体系的哈密顿量变为

$$H=\frac{P_1^2}{2}+\frac{P_2^2}{2}+\frac{\omega^2}{2}X_2^2. \tag{48}$$

体系可以分解为相互独立的1个谐振子和1个自由粒子, 能流相应地变为

$$S=\frac{P_1^3}{2\sqrt{m_1+m_2}}+\frac{3P_1P_2^2}{2\sqrt{m_1+m_2}}+\frac{\sqrt{\mu}}{2}\frac{m_2-m_1}{m_1m_2}P_2^3$$
$$-\frac{\omega^2}{2\sqrt{m_1+m_2}}X_2\left(X_2+2\sqrt{\mu}r_{\mathrm{eq}}\right)P_1$$
$$-\frac{\sqrt{\mu}\omega^2}{4}\frac{m_2-m_1}{m_1m_2}X_2\left(X_2+2\sqrt{\mu}r_{\mathrm{eq}}\right)P_2. \tag{49}$$

容易证明采用蛙跳“居中”方案所得的能流的平均值为

$$\langle S\rangle^{\text{LF-Middle}}=\frac{\omega^2r_{\mathrm{eq}}}{4\beta}\frac{m_1-m_2}{m_1+m_2}\Delta t. \tag{50}$$

能流平方的平均值是

$$\langle S^2\rangle^{\text{LF-Middle}}=\frac{51}{16\mu\beta^3}+\frac{\omega^2r_{\mathrm{eq}}^2}{4\beta^2}$$
$$+\left[-\frac{9}{16\mu\beta^3}+\frac{3}{2\beta^3}\frac{1}{m_1+m_2}+\frac{\omega^2r_{\mathrm{eq}}^2}{8\beta^2}\left(\frac{m_2-m_1}{m_1+m_2}\right)^2\right]\omega^2\Delta t^2, \tag{51}$$

于是能流涨落的平方是

$$\left(f_S^{\text{LF-Middle}}\right)^2=\frac{51}{16\mu\beta^3}+\frac{\omega^2r_{\mathrm{eq}}^2}{4\beta^2}$$
$$+\left[-\frac{9}{16\mu\beta^3}+\frac{3}{2\beta^3}\frac{1}{m_1+m_2}+\frac{\omega^2r_{\mathrm{eq}}^2}{16\beta^2}\left(\frac{m_2-m_1}{m_1+m_2}\right)^2\right]\omega^2\Delta t^2. \tag{52}$$

而采用速度-Verlet“居中”方案所得的能流的平均值为0, 能流的涨落的平方为

$$\left(f_S^{\text{VV-Middle}}\right)^2=\langle S^2\rangle^{\text{VV-Middle}}-\left(\langle S\rangle^{\text{VV-Middle}}\right)^2=\frac{3}{\beta^3}\frac{1}{m_1+m_2}+\frac{\omega^2\mu^2r_{\mathrm{eq}}^2}{\beta^2m_1m_2}+\left[\frac{9}{4\beta^3}\frac{1}{m_1+m_2}+\frac{3}{16\mu\beta^3}+\frac{\omega^2r_{\mathrm{eq}}^2}{4\beta^2}\left(\frac{m_2-m_1}{m_1+m_2}\right)^2\right]\left(1-\frac{\omega^2\Delta t^2}{4}\right)$$
$$+\left[\frac{39}{4\beta^3}\frac{1}{m_1+m_2}-\frac{3}{4\mu\beta^3}\right]\left(1-\frac{\omega^2\Delta t^2}{4}\right)^2+\left[-\frac{15}{\beta^3}\frac{1}{m_1+m_2}+\frac{15}{4\mu\beta^3}\right]\left(1-\frac{\omega^2\Delta t^2}{4}\right)^3$$
$$\approx\frac{51}{16\mu\beta^3}+\frac{\omega^2r_{\mathrm{eq}}^2}{4\beta^2}+\left[-\frac{159}{16\mu\beta^3}+\frac{93}{4\beta^3}\frac{1}{m_1+m_2}-\frac{\omega^2r_{\mathrm{eq}}^2}{4\beta^2}\left(\frac{m_2-m_1}{m_1+m_2}\right)^2\right]\frac{\omega^2\Delta t^2}{4}+O\left(\omega^4\Delta t^4\right). \tag{53}$$

可见, 对于能流的估计值, 蛙跳“居中”方案的误差比速度-Verlet“居中”方案的要大. 而对于能流的涨落的估计值, 蛙跳“居中”方案和速度-Verlet“居中”方案哪种表现更好与所研究的具体体系有关.

采用“两边”方案所得的能流的平均值为0, 能流的涨落的平方为

$$\begin{aligned}\left(f_S^{\text{Side}}\right)^2 &= \left\langle S^2\right\rangle^{\text{Side}} - \left(\left\langle S\right\rangle^{\text{Side}}\right)^2 \\ &= \frac{15}{4\mu\beta^3} + \left[-\frac{3}{4\mu\beta^3} + \frac{\omega^2 r_{\text{eq}}^2}{4\beta^2}\right]\left(1-\frac{\omega^2\Delta t^2}{4}\right)^{-1} \\ &\quad + \frac{3}{16\mu\beta^3}\left(1-\frac{\omega^2\Delta t^2}{4}\right)^{-2}.\end{aligned} \tag{54}$$

当时间步长$\Delta t \to 2/\omega$, 即趋于稳定性条件边界值时, 能流的涨落趋于无穷大. 这表明“两边”方案对于这个物理量的计算表现不佳.

综合而言, 对于包含位置和动量的交叉项的物理量, 两种“居中”方案中的何种方案表现更佳取决于要研究的体系和具体物理量.

2.1.2 含部分构型约束的系统

本文中, 对于分子体系含有部分构型约束条件(例如, 固定一部分或全部键长、键角、或二面角)时的NVT系综分子动力学, 蛙跳“居中”方案(式(24)~(27))只需做如下调整: 在式(24)和(25)之间加入RATTLE算法[34]来更新动量使之与约束条件相匹配; 在式(27)之后加入SHAKE算法[35]和RATTLE算法来分别更新位置和动量使之与约束条件相匹配. 即有

$$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right) \leftarrow \boldsymbol{p}\left(\left(n-\frac{1}{2}\right)\Delta t\right) - U'\left(\boldsymbol{x}(n\Delta t)\right)\Delta t, \tag{55}$$

RATTLE算法 (56)

$$\boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right) \leftarrow \boldsymbol{x}(n\Delta t) + \boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2}, \tag{57}$$

控温整个时间步长 Δt: 改变动量 $\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$, (58)

$$\boldsymbol{x}\left((n+1)\Delta t\right) \leftarrow \boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right) + \boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2}, \tag{59}$$

SHAKE和RATTLE算法 (60)

与之对应的速度-Verlet“居中”方案如下:

$$\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right) \leftarrow \boldsymbol{p}(n\Delta t) - U'\left(\boldsymbol{x}(n\Delta t)\right)\frac{\Delta t}{2}, \tag{61}$$

RATTLE算法 (62)

$$\boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right) \leftarrow \boldsymbol{x}(n\Delta t) + \boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2}, \tag{63}$$

控温整个时间步长 Δt: 改变动量 $\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)$, (64)

$$\boldsymbol{x}\left((n+1)\Delta t\right) \leftarrow \boldsymbol{x}\left(\left(n+\frac{1}{2}\right)\Delta t\right) + \boldsymbol{M}^{-1}\boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right)\frac{\Delta t}{2}, \tag{65}$$

SHAKE算法 (66)

$$\boldsymbol{p}\left((n+1)\Delta t\right) \leftarrow \boldsymbol{p}\left(\left(n+\frac{1}{2}\right)\Delta t\right) - U'\left(\boldsymbol{x}\left((n+1)\Delta t\right)\right)\frac{\Delta t}{2}, \tag{67}$$

RATTLE算法 (68)

特别地, 对于三原子体系可以用SETTLE算法[36]解析地求解约束方程, 分别在蛙跳“居中”方案的式(56)和(60)中或在速度-Verlet“居中”方案的式(62), (66)和(68)中用该算法替代RATTLE和SHAKE算法.

本文仅考虑上述这种策略来构造蛙跳“居中”方案. 在后续的工作中对含约束的NVT系综分子动力学算法构建的更多策略进行更详细讨论.

本文后面部分的图表中分别用“Side”, “LF-Middle”与“VV-Middle”来标记“两边”方案、蛙跳“居中”方案与速度-Verlet“居中”方案. 由于分子动力学模拟软件AMBER[37]中默认的含部分构型约束的控温算法[38]是基于蛙跳算法建立的, 本文标记为“Old”, 并与之比较来研究实际分子体系.

下面我们测试非谐模型和实际分子体系来检验蛙跳“居中”方案的效果.

2.2 蛙跳“居中”方案的数值计算精度

用于数值测试的非谐体系有一维四次势模型和一个凝聚相体系: 室温(298.15 K)液态水. 我们计算得到体系的势能、动能、哈密顿量及这些物理量的涨落.

2.2.1 一维四次势

势能形式为$U(x)=x^4/4$. 体系质量设为1 au(原子单位), 温度参数设为 $\beta \equiv 1/k_{\text{B}}T=8$ au. 由于势能不含任何二次项(即谐振作用项), 数值模拟的结果有重要的参考价值. 每次模拟中的时间步长Δt在0.1~0.8 au之间选取一个固定值. 采用Langevin控温方法, 选取摩擦系数γ=1 au. 计算20条NVT系综分子动力学轨迹用于取样. 每条轨迹先取10^6 au的时间长度用于达到平衡, 之后时间长度10^8 au的轨迹数据用于取样.

图1(a)和(b)显示, 对于一维四次势体系的势能

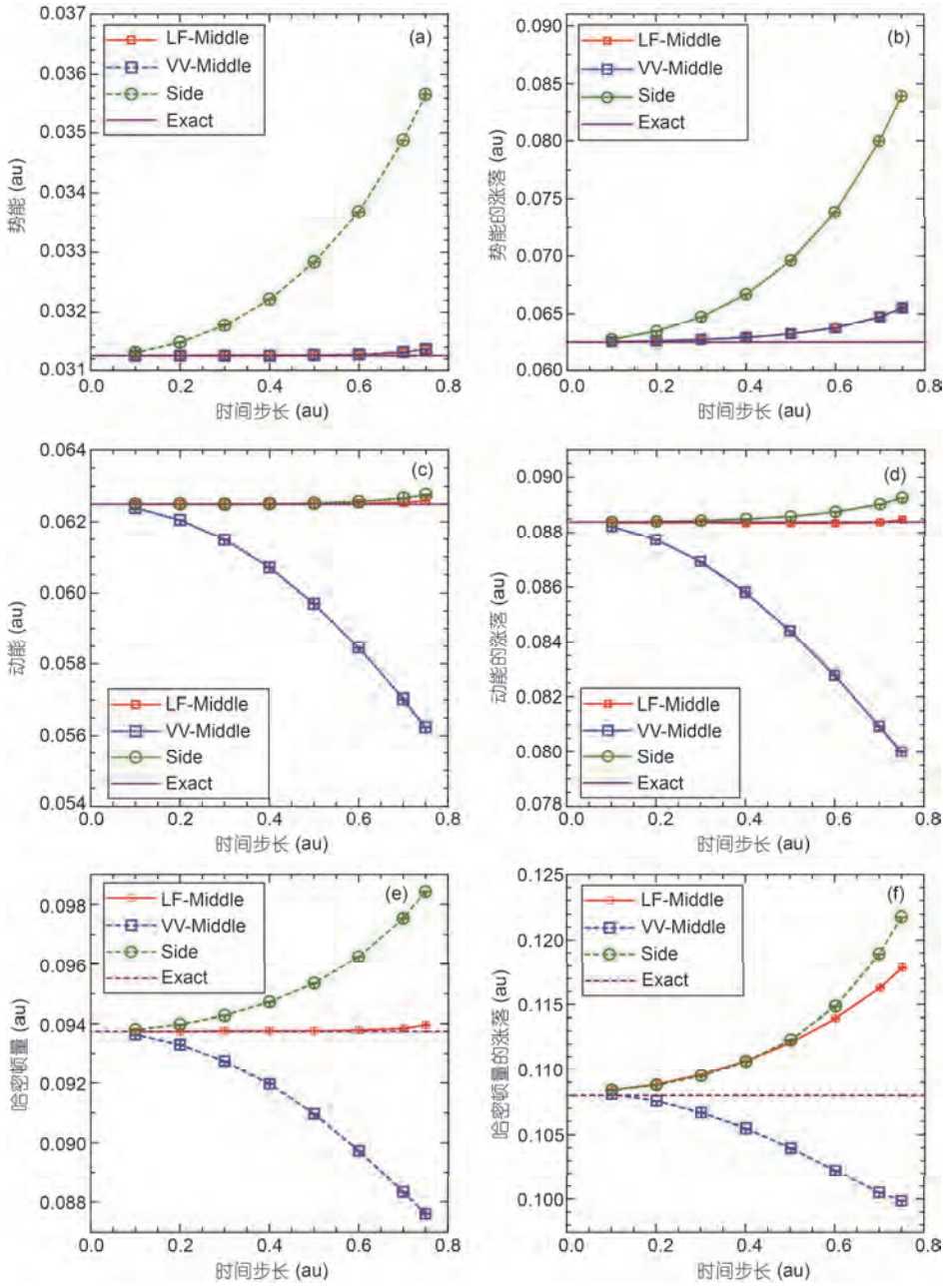


图 1 (网络版彩色)一维四次势模型体系. (a) 势能平均值; (b) 势能的涨落; (c) 动能平均值; (d) 动能的涨落; (e) 哈密顿量平均值; (f) 哈密顿量的涨落

Figure 1 (Color online) Averaged values of the one dimensional quartic potential. (a) Potential energy; (b) fluctuation of the potential energy; (c) kinetic energy; (d) fluctuation of the kinetic energy; (e) Hamiltonian; (f) fluctuation of the Hamiltonian

(及其涨落), 蛙跳"居中"方案与速度-Verlet"居中"方案表现相同, 比"两边"方案要更加稳定, 误差更小. 例如, 蛙跳"居中"方案可以用比"两边"方案大6~7倍的时间步长Δt来取得势能的收敛结果. 另一方面, 图1(c)和(d)显示, 对于这个体系的动能(及其涨落), 蛙跳"居中"方案可以得到比"两边"方案更精确的结果, 大大优于速度-Verlet"居中"方案. 图1(e)显示, 蛙跳"居中"方案即便在较大的时间步长Δt=0.6 au时得到的哈密顿量数值上仍然与准确值非常接近, 而这一时间步长是其他两种方案得到收敛值时所用时间步长Δt=0.1 au的6倍. 图1(f)表明, 对于哈密顿量的涨落的计算, 两种"居中"方案误差大小相近, 但偏离准

确值的方向相反; 两者都比“两边”方案更稳定.

2.2.2 液态水体系

考虑室温(298.15 K)液态水这个实际凝聚相体系. 模拟中使用TIP3P力场模型[39], 设置一个含216个水分子的具有周期性边界条件的盒子. 将各个方案加入分子模拟软件AMBER中(版本AMBER18)来进行计算. 模拟中使用Langevin控温方法, 摩擦系数γ=5 ps^{-1}, 时间步长Δt在0.2~1.7 fs之间取一个固定值. 每条轨迹的平衡时间取为500 ps, 之后的时间长度5 ns的数据用于取样. 我们计算了100条分子动力学轨迹来得到体系的势能、动能、哈密顿量的平均值及涨落.

图2显示, “两边”方案与AMBER默认方案[38]的势

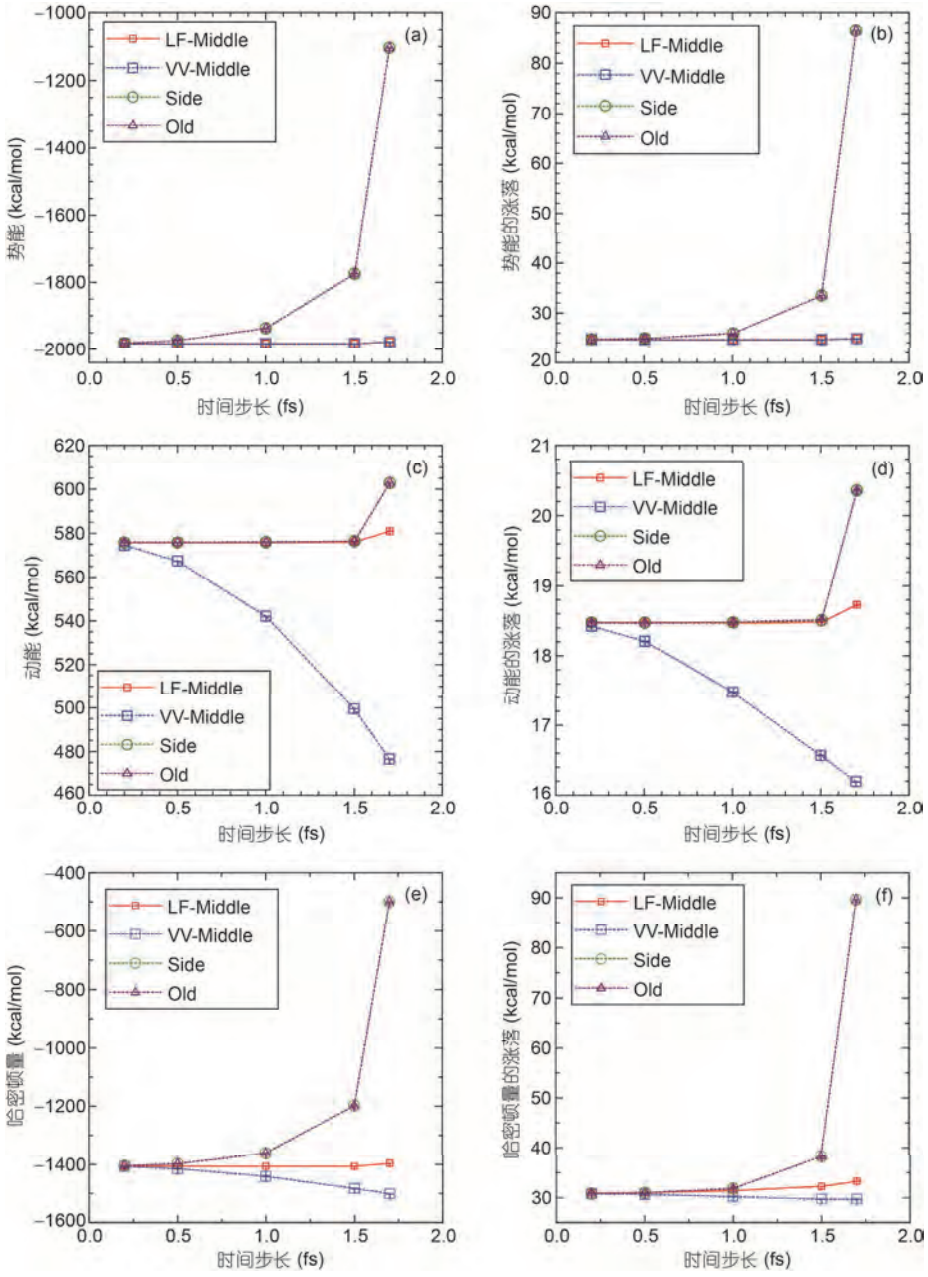


图 2 (网络版彩色)液态水体系, TIP3P力场模型, 温度为 298.15 K. (a) 势能平均值; (b) 势能的涨落; (c) 动能平均值; (d) 动能的涨落; (e) 哈密顿量的平均值; (f) 哈密顿量的涨落

Figure 2 (Color online) Averaged values of liquid water with the TIP3P force field model at 298.15 K. (a) Potential energy; (b) fluctuation of the potential energy; (c) kinetic energy; (d) fluctuation of the kinetic energy; (e) Hamiltonian; (f) fluctuation of the Hamiltonian

能、动能、哈密顿量(以及它们的涨落)结果基本相同. 图2(a)和(b)表明, 对于液态水这一实际大体系的势能(及其涨落)的计算, 蛙跳“居中”方案与速度-Verlet“居中”方案表现相同, 比“两边”方案要更加稳定, 误差更小. 例如, 为了取得收敛的势能结果, 蛙跳“居中”方案可以使用比“两边”方案大7~8倍的时间步长Δt. 图2(c)和(d)显示了动能(及其涨落)计算值随时间步长增大时的误差变化. 对于这些和动量有关的物理量的计算, 蛙跳“居中”方案表现比“两边”方案更好, 远胜于速度-Verlet“居中”方案. 图2(e)显示, 对于哈密顿量的计算, 蛙跳“居中”方案比其它方案要更加稳定, 误差更小. 图2(f)显示, 对于哈密顿量的涨落的模拟, 两种居中方案随着时间步长的增大均会逐渐偏离收敛值, 偏离方向相反, 误差大小接近; 两者都比“两边”方案更加稳定和接近收敛值.

2.2.3 在键长约束条件下的液态水体系

将216个水分子设置在一个具有周期性边界条件的盒子里, 采用TIP3P力场模型, 并加上O–H键长约束, 使用修改过的Amber18版本软件进行模拟. 模拟中仍然使用Langevin控温方法, 设置摩擦系数γ=5 ps^{-1}. 模拟中使用SETTLE算法[36]求解约束方程, 修正位置和速度. 每次模拟中取一个固定时间步长Δt, 其取值范围为0.5~7 fs. 每条分子动力学轨迹的平衡时间取为500 ps, 之后的时间长度5 ns的数据用于取样. 每次模拟用100条这样的轨迹来得到体系的势能动能平均值及涨落.

图3显示, 对于这个含有键长约束条件的体系, “两边”方案与AMBER默认方案的势能、动能、哈密顿量以及3个物理量的涨落的结果大致相同, 动能(及其涨落)结果在时间步长Δt较大时稍微更接近收敛值. 对于势能(及其涨落)的计算, 蛙跳“居中”方案与速度-Verlet“居中”方案表现基本相同, 比“两边”方案要更加稳定, 精度更高. 例如, 以取得收敛的势能结果来衡量, 蛙跳“居中”方案可以采用比“两边”方案大7~14倍的时间步长Δt. 另一方面, 对于动能(及其涨落)的计算, 蛙跳“居中”方案的结果比速度-Verlet“居中”方案更接近收敛值. 对于哈密顿量的计算, 蛙跳“居中”方案数值表现最佳, 而对于哈密顿量的涨落的计算, 速度-Verlet“居中”方案数值表现最好.

2.2.4 三维Morse势的OH分子

我们考察一个三维Morse势的OH分子体系的能流的涨落 $f_S \equiv \sqrt{\left\langle |\mathbf{S}|^2 \right\rangle - |\langle \mathbf{S} \rangle|^2}$, 能流的表达式见式(43), Morse势的参数见文献[40], 体系温度设为T=300 K. 模拟采用Langevin控温方法时其摩擦系数设为γ=0.017 au, 而采用Andersen控温方法时其碰撞频率设为$\nu = 0.017\sqrt{2}$ au. 采样基于20条经过10^6 au的时间长度平衡后的NVT系综分子动力学轨迹, 每条轨迹平衡后再演化10^8 au的时间长度用于采样.

图4表明, 对于三维Morse势体系的能流及其涨落的估计, 同一方案中Andersen和Langevin控温方法得到的结果彼此比较接近, 而无论哪种控温方法, 对于能流及其涨落的估计, 速度-Verlet“居中”方案的误差都是最小的. 对于能流的估计, 蛙跳“居中”方案误差最大, 而对于能流的涨落的估计, “两边”方案误差最大.

本文2.1和2.2节中的解析或计算结果表明蛙跳“居中”方案和速度-Verlet“居中”方案在位置空间的采样精度在无约束条件下完全相同, 而且在有约束条件下几乎一致, 与此同时, 蛙跳“居中”方案比速度-Verlet“居中”方案在只含动量的物理量的计算上精度更高. 然而, 这并不意味着蛙跳“居中”方案一定比速度-Verlet“居中”方案更优越. 对于含有位置和动量交叉项的物理量, 两种“居中”方案中何种方案表现更佳随体系不同或考量物理量不同所得结果也不相同. 从两个方案的相空间演化算符可知, 对于普遍体系而言, 蛙跳“居中”方案导出对有限时间步长Δt是一阶精度的相空间稳态平衡分布, 而速度-Verlet“居中”方案则得到对Δt是二阶精度的相空间稳态平衡分布. 对于一般性含有位置和动量交叉项的物理量, 考虑速度-Verlet“居中”方案更合理.

下面我们集中比较两种“居中”方案的采样效率.

2.3 蛙跳“居中”方案的采样效率

势能关联函数的特征时间常用来表征位置空间采样效率, 而哈密顿量关联函数的特征时间常用来表征相空间采样效率. 在时间步长Δt有限的情况下, 势能关联函数的特征时间为

$$\tau_{UU} = \Delta t \sum_{n=0}^{\infty} \frac{\langle U(0)U(n\Delta t)\rangle - \langle U\rangle^2}{\langle U^2\rangle - \langle U\rangle^2}, \tag{69}$$

而哈密顿量关联函数的特征时间为

$$\tau_{HH} = \Delta t \sum_{n=0}^{\infty} \frac{\langle H(0)H(n\Delta t)\rangle - \langle H\rangle^2}{\langle H^2\rangle - \langle H\rangle^2}. \tag{70}$$

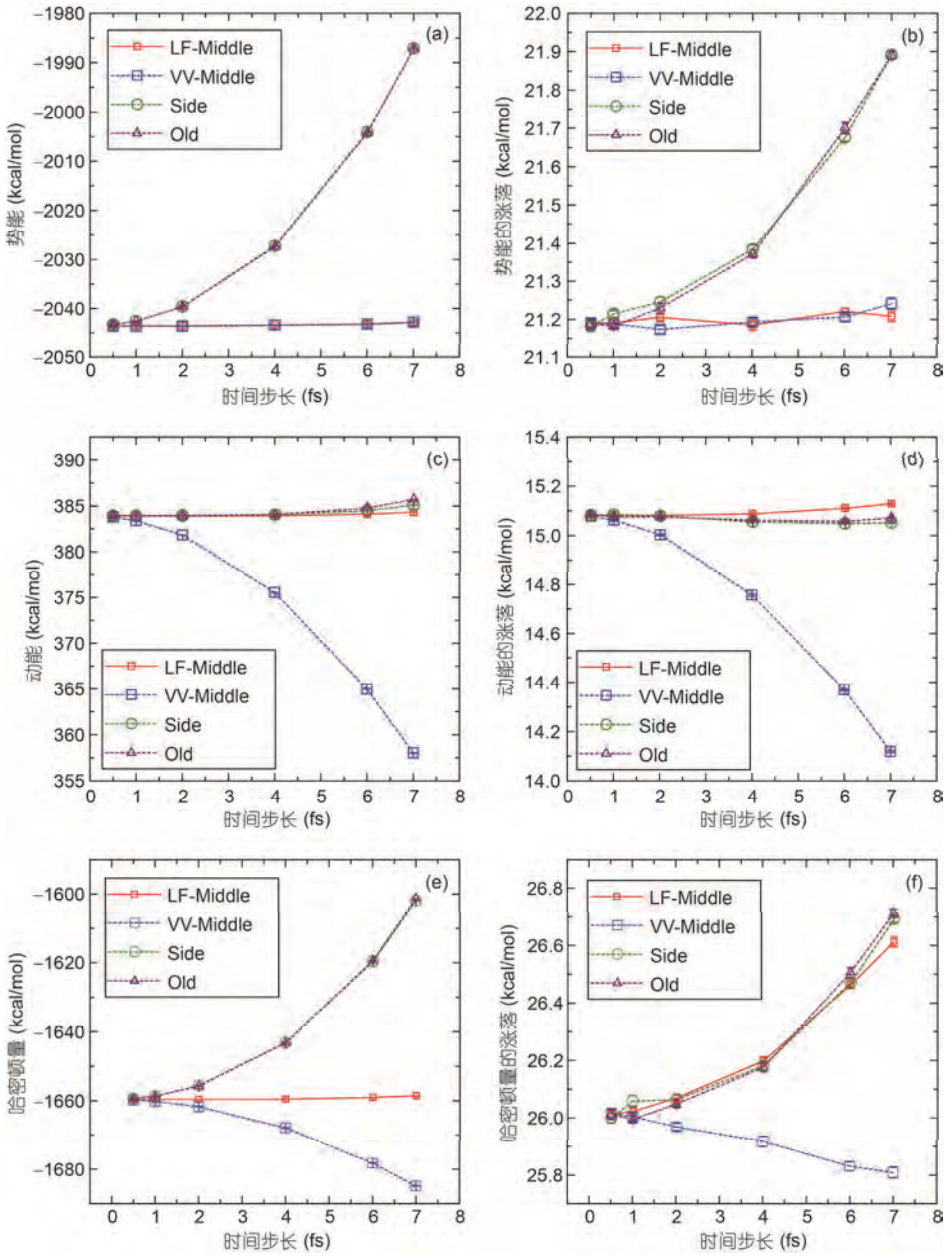


图 3 (网络版彩色)O–H键长约束下的液态水体系, TIP3P力场模型, 温度为 298.15 K. (a) 势能平均值; (b) 势能的涨落; (c) 动能平均值; (d) 动能的涨落; (e) 哈密顿量平均值; (f) 哈密顿量的涨落

Figure 3 (Color online) Averaged values of liquid water with O–H bond-length constraints with the TIP3P force field model at temperature 298.15 K. (a) Potential energy; (b) fluctuation of the potential energy; (c) kinetic energy; (d) fluctuation of the kinetic energy; (e) Hamiltonian; (f) fluctuation of the Hamiltonian

文献[29,30]表明特征关联(时间)步数$\tau_{UU}/\Delta t$(或$\tau_{HH}/\Delta t$)是可以用来衡量控温算法对位置空间(或相空间)采样效率的指标. 需要重点强调的是, 在蛙跳"居中"方案中, 第n个时间步长Δt后得到的相空间由位置$\boldsymbol{x}(n\Delta t)$和动量$\boldsymbol{p}\left(\left(n-\frac{1}{2}\right)\Delta t\right)$所构成, 所以$n\Delta t$这个时刻的任何物理量函数都采用$\boldsymbol{x}(n\Delta t)$和$\boldsymbol{p}\left(\left(n-\frac{1}{2}\right)\Delta t\right)$作为变量来计算.

2.3.1 谐振子体系

首先考虑一维谐振子体系(式(40)). 利用文献[29,30]中的针对有限时间步长Δt条件下的相空间

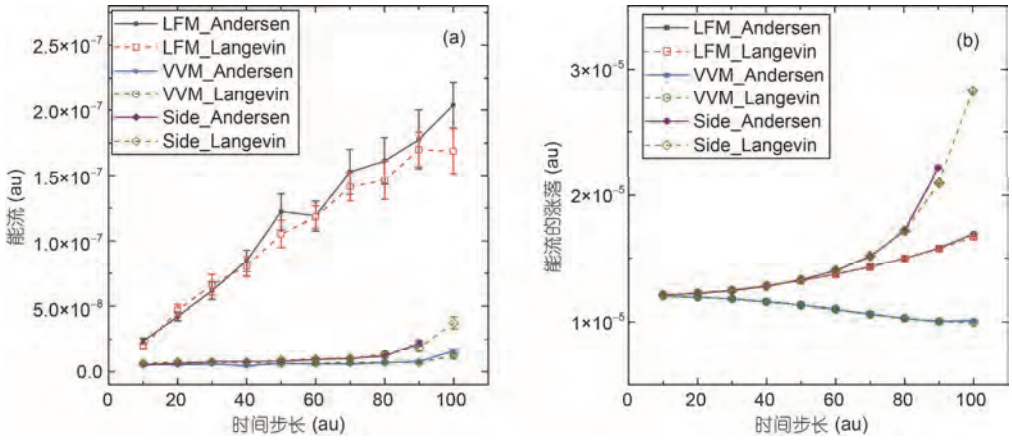


图 4 (网络版彩色)三维Morse势的OH分子体系, 温度为 300 K. (a) 能流平均值; (b) 能流的涨落. “LFM_Andersen”表示蛙跳“居中”方案采用Andersen控温方法, “VVM_Langevin”表示速度-Verlet方案采用Langevin控温方法, 以此类推

Figure 4 (Color online) Averaged values of the OH molecular described by the 3-D Morse potential at 300 K. (a) Energy current; (b) fluctuation of the energy current. “LFM_Andersen” represents the leap-frog “middle” scheme using the Andersen thermostat, “VVM_Langevin” denotes the velocity-Verlet “middle” scheme using the Langevin thermostat, etc.

演化子(phase space propagator)分析方法, 容易推导得到蛙跳“居中”控温方案在有限时间步长Δt情况下的势能或哈密顿量关联函数的特征时间. 对于势能关联函数特征时间, 蛙跳“居中”控温方案采用Andersen控温方法时有

$$\tau_{UU}^{\text{Andersen}}=\frac{(1-e^{-\nu\Delta t})^2+\left(3+6e^{-\nu\Delta t}-e^{-2\nu\Delta t}\right)\dfrac{\omega^2\Delta t^2}{4}}{\omega^2\Delta t(1-e^{-2\nu\Delta t})}. \quad (71)$$

而采用Langevin控温方法有

$$\tau_{UU}^{\text{Langevin}}=\frac{\left(1-e^{-\gamma\Delta t}\right)^2+\left(3+2e^{-\gamma\Delta t}-e^{-2\gamma\Delta t}\right)\dfrac{\omega^2\Delta t^2}{4}}{\omega^2\Delta t\left(1+e^{-\gamma\Delta t}\right)\left(1-e^{-\gamma\Delta t}\right)}. \quad (72)$$

与文献[29]和[30]中速度-Verlet“居中”方案的对应结果比较可知, 蛙跳“居中”方案与速度-Verlet “居中”方案所得势能关联函数特征时间完全相同. 因此, 蛙跳“居中”方案对于谐振子体系的势能关联函数特征时间的分析结果与文献[30]中对应结论完全一致. 两种“居中”方案导出相同的势能关联函数特征时间说明两者对位置空间的采样效率相同.

但两种“居中”方案所得的哈密顿量关联函数特征时间并不相同. 蛙跳“居中”控温方案采用控温方法时有

$$\tau_{HH}^{\text{Andersen}}=\frac{4(1-e^{-\nu\Delta t})^2+\left(11+18e^{-\nu\Delta t}+3e^{-2\nu\Delta t}\right)\omega^2\Delta t^2+(1-e^{-2\nu t})\omega^4\Delta t^4}{2\omega^2\Delta t(1-e^{-2\nu\Delta t})(\omega^2\Delta t^2+4)}, \quad (73)$$

这与文献[30]中式(78)所得的速度-Verlet“居中”方案的对应结果不同.

由式(73)可得两个极限:

$$\tau_{HH}^{\text{Andersen}}\xrightarrow{\nu\to0^+}\infty$$
$$\tau_{HH}^{\text{Andersen}}\xrightarrow{\nu\to\infty}\frac{4+11\omega^2\Delta t^2+\omega^4\Delta t^4}{2\omega^2\Delta t(4+\omega^2\Delta t^2)}. \quad (74)$$

当$0<\omega\Delta t<2/3$时, $\tau_{HH}^{\text{Andersen}}$有最小值

$$\tau_{HH}^{\text{Andersen,min}}=\frac{\omega^3\Delta t^3+4\omega\Delta t+4\sqrt{2}\sqrt{4-\omega^2\Delta t^2}}{2\omega\left(\omega^2\Delta t^2+4\right)}, \quad (75)$$

对应的最优碰撞频率参数为

$$\nu^{\text{opt}}=\frac{1}{\Delta t}\ln\left[\frac{4+7\omega^2\Delta t^2+4\sqrt{2}\omega\Delta t\sqrt{4-\omega^2\Delta t^2}}{4-9\omega^2\Delta t^2}\right] \quad (76)$$
$$(0<\omega\Delta t<2/3).$$

当$2/3\leqslant\omega\Delta t<2$时, $\tau_{HH}^{\text{Andersen}}$无最小值, 随着$\nu$增大而单调减小, 逐渐趋近于平台值$\tau_{HH}^{\text{Andersen},\nu\to\infty}$.

蛙跳“居中”控温方案采用Langevin控温方法时的哈密顿量关联函数特征时间为

$$\tau_{HH}^{\text{Langevin}}=\frac{4(1-e^{-\gamma\Delta t})^2+\left(11+2e^{-\gamma\Delta t}+3e^{-2\gamma\Delta t}\right)\omega^2\Delta t^2+(1-e^{-2\gamma\Delta t})\omega^4\Delta t^4}{2\omega^2\Delta t(1-e^{-2\gamma\Delta t})(\omega^2\Delta t^2+4)}, \quad (77)$$

这与文献[29]中式(187)所得的速度-Verlet“居中”方案的对应结果不同. 式(77)的两个极限是:

$$\tau_{HH}^{\text{Langevin}}\xrightarrow{\gamma\to0^+}\infty$$
$$\tau_{HH}^{\text{Langevin}}\xrightarrow{\gamma\to\infty}\frac{4+11\omega^2\Delta t^2+\omega^4\Delta t^4}{2\omega^2\Delta t(4+\omega^2\Delta t^2)}. \quad (78)$$

注意式(78)中的第2个极限(平台值)和式(74)中的第2个极限(平台值)完全相同.

由式(77)可得 $\tau_{HH}^{\text{Langevin}}$ 的最小值为

$$\tau_{HH}^{\text{Langevin,min}}=\frac{4\omega\Delta t+\omega^3\Delta t^3+4\sqrt{4+3\omega^2\Delta t^2}}{2\omega\left(4+\omega^2\Delta t^2\right)}, \quad (79)$$

对应的最优摩擦系数是:

$$\gamma^{\text{opt}}=\frac{1}{\Delta t}\ln\left[\frac{4+7\omega^2\Delta t^2+4\omega\Delta t\sqrt{4+3\omega^2\Delta t^2}}{4-\omega^2\Delta t^2}\right] \quad (80)$$

$(0<\omega\Delta t<2)$.

因为$0<\omega\Delta t<2$是这个体系数值模拟稳定的条件, 所以 $\tau_{HH}^{\text{Langevin}}$ 总是有最小值. 即 $\tau_{HH}^{\text{Langevin}}$ 随着摩擦系数γ的增大呈现先递减到最小值然后再逐渐增大并趋近于平台值 $\tau_{HH}^{\text{Langevin},\gamma\to\infty}$.

文献[30]建议可以定义统一的控温参数 $\xi=v=\sqrt{2}\gamma$ 来刻画同一控温方案中这两种随机控温方法的表现. 由前面分析可知, 在蛙跳"居中"方案中Andersen和Langevin这两种随机控温方法的所得的势能(或哈密顿量)关联函数特征时间, 在控温参数ξ趋于0时都趋于无穷, 而当控温参数ξ趋于无穷时都趋于同一个常数平台. 因为两种"居中"方案导出不同的哈密顿量关联函数特征时间, 所以我们以此来比较这两种方案对相空间的采样效率的差别.

图5(a)显示, 对于哈密顿量关联函数特征时间的最小值, 在采用相同控温方法(无论是Andersen或是Langevin方法)的情况下, 蛙跳"居中"方案都比速度-Verlet"居中"方案给出稍微更小的结果. 图5(b)则显示, 对于控温参数ξ趋于无穷时所得到的哈密顿量关联函数特征时间的平台值, 蛙跳"居中"方案给出的结果比速度-Verlet"居中"方案导出的结果要更小. 图5表明, 对于谐振子极限下的相空间采样效率, 蛙跳"居中"方案比速度-Verlet"居中"方案要略微高效.

图6(a)显示当$\omega\Delta t\to 0$时, 两种"居中"方案在谐振子极限下的哈密顿量关联函数特征时间完全相同. 图6(b)~(d)显示两种"居中"方案的 $\tau_{HH}^{\text{Langevin}}$ 随着摩擦系数γ的增大, 先递减到最小值, 再逐渐增大并趋近于平台值 $\tau_{HH}^{\text{Langevin},\gamma\to\infty}$. 不论采用Andersen还是Langevin控温方法, 蛙跳"居中"方案都比速度-Verlet"居中"方案的相空间采样效率要略高但很接近.

下面我们数值计算检验蛙跳"居中"方案关于非谐体系的相空间采样效果.

2.3.2 非谐体系

使用一维四次势$U(x)=x^4/4$进行数值测试, 设置质量为1 au, 温度参数为β=1 au. 我们分别取时间步长Δt为0.1, 0.3和0.4 au进行模拟. 每次模拟计算20条体系的NVT系综分子动力学轨迹. 每条轨迹设置平

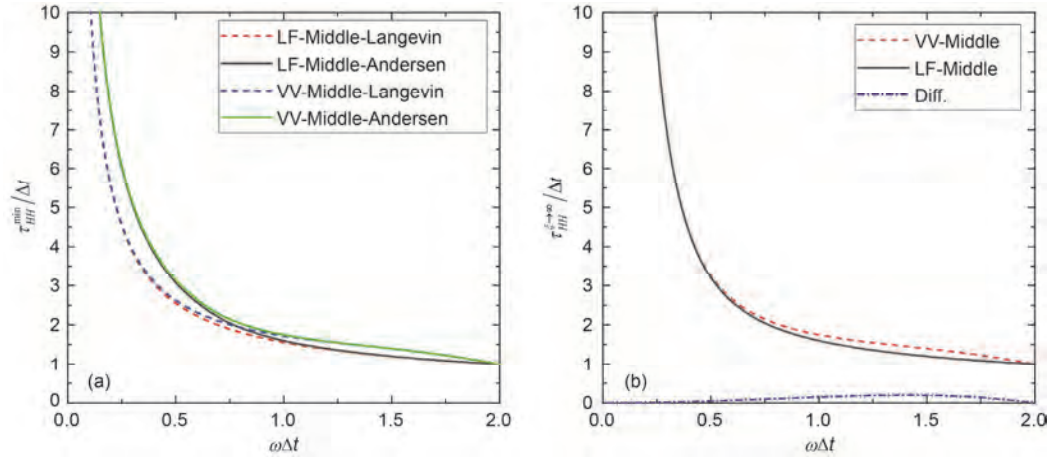


图 5 (网络版彩色)谐振子体系蛙跳"居中"和速度-Verlet"居中"两种方案的哈密顿量关联函数特征时间. (a) 最小值 $\tau_{HH}^{\min}/\Delta t$ 随$\omega\Delta t$的变化; (b) 平台值 $\tau_{HH}^{\xi\to\infty}/\Delta t$ 随$\omega\Delta t$的变化

Figure 5 (Color online) Analytic results for both the leap-frog and the velocity-Verlet "middle" schemes for the harmonic system. (a) $\tau_{HH}^{\min}/\Delta t$ as a function of $\omega\Delta t$, where $\tau_{HH}^{\min}/\Delta t$ is the step number of the minimal characteristic correlation time of the Hamiltonian. (b) $\tau_{HH}^{\xi\to\infty}/\Delta t$ as a function of $\omega\Delta t$, where $\tau_{HH}^{\xi\to\infty}$ is the plateau value of the characteristic correlation time of the Hamiltonian (here ξ stands for $\sqrt{2}\gamma$ for Langevin dynamics and v for the Andersen thermostat), and $\tau_{HH}^{\xi\to\infty}/\Delta t$ represents the step number of the value of the plateau

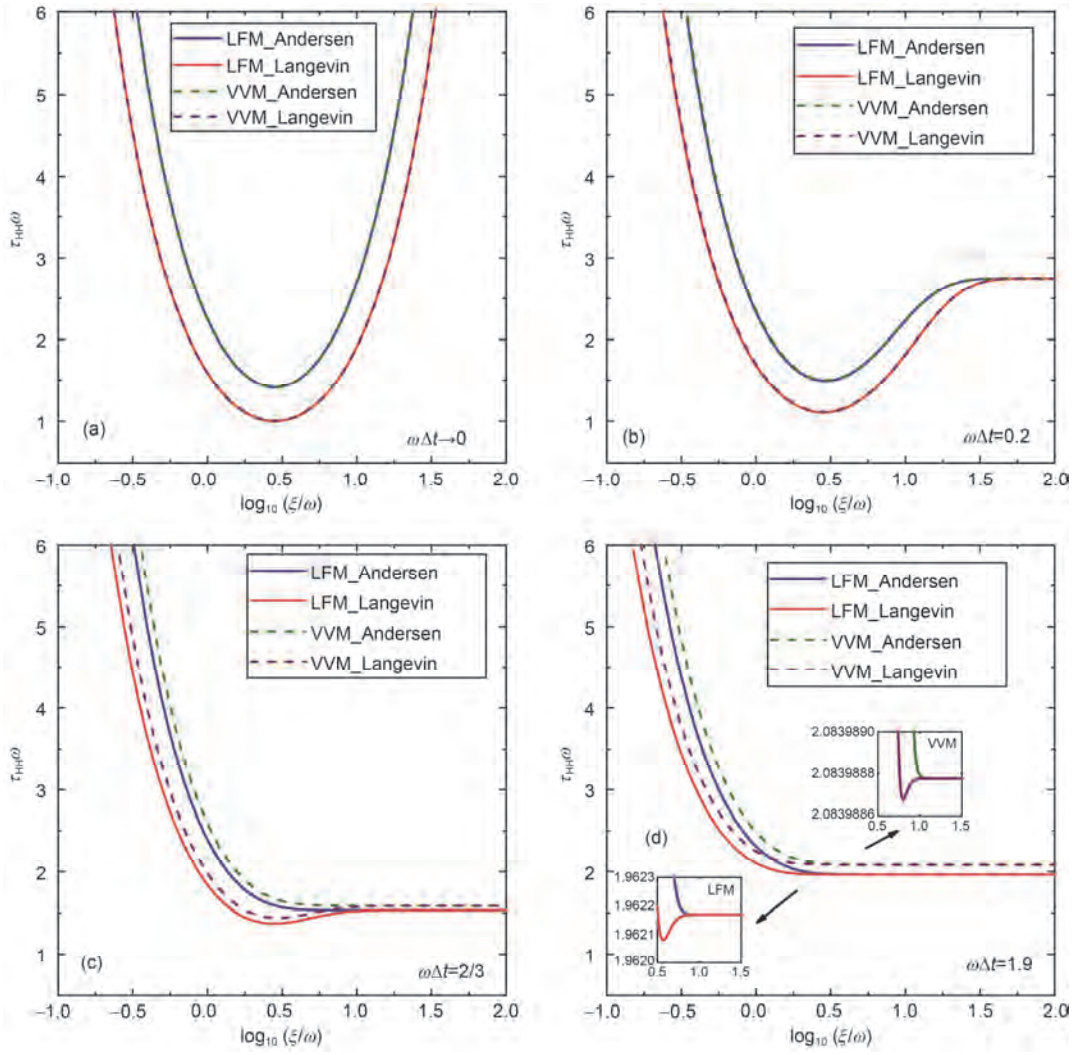


图 6 (网络版彩色)一维谐振子体系的哈密顿量关联函数特征时间的解析结果. (a) $\omega\Delta t\to 0$; (b) $\omega\Delta t$=0.2; (c) $\omega\Delta t$=2/3; (d) $\omega\Delta t$=1.9. 控温参数ξ对应于Andersen控温方法的v或Langevin控温方法的$\sqrt{2}\gamma$. "LFM_Andersen"表示蛙跳"居中"方案采用Andersen控温方法, "VVM_Langevin"表示速度-Verlet"居中"方案采用Langevin控温方法, 以此类推

Figure 6 (Color online) Analytic results of the characteristic time of the correlation function for the Hamiltonian for one-dimensional harmonic system. The curves depict the equations that τ_{HH}^{ω} and ξ/ω satisfy for different values for the parameter $\omega\Delta t$. Here ξ stands for $\sqrt{2}\gamma$ for Langevin dynamics and v for the Andersen thermostat. (a) $\omega\Delta t\to 0$; (b) $\omega\Delta t$=0.2; (c) $\omega\Delta t$=2/3; (d) $\omega\Delta t$=1.9. "LFM_Andersen" represents the leap-frog "middle" scheme using the Andersen thermostat, "VVM_Langevin" denotes the velocity-Verlet "middle" scheme using the Langevin thermostat, etc.

衡时间长度为10^6 au, 之后时间长度为10^8 au的轨迹数据用以采样. 对每条轨线做时间平均得到哈密顿量关联函数来计算其特征关联时间. 再对20条轨迹的特征关联时间计算结果求平均值和统计误差.

图7显示对于一维四次势体系, 采用蛙跳"居中"方案, 当控温参数$\xi=v=\sqrt{2}\gamma$大于100 au时, 哈密顿量关联函数特征时间取到平台值, 且Langevin和Andersen控温方法的平台值相同. 图7(a)显示, 在采用相同控温方法的情况下, 时间步长Δt=0.1 au比Δt=0.3 au的特征时间最小值更小, 但特征时间平台值更大. 最优的控温参数区间大约在1~50 au. 当控温参数为5~7 au时, 哈密顿量关联函数的特征时间取到最小值. 与Andersen控温方法相比, Langevin控温方法的特征时间要更小. 图7(b)~(d)分别比较了蛙跳"居中"方案与速度-Verlet"居中"方案在Δt=0.1, 0.3, 0.4 au时的数值表现. 当采用相同控温方法时, 蛙跳"居中"方案与速度-Verlet"居中"方案结果在不同控温参数下均非常接近.

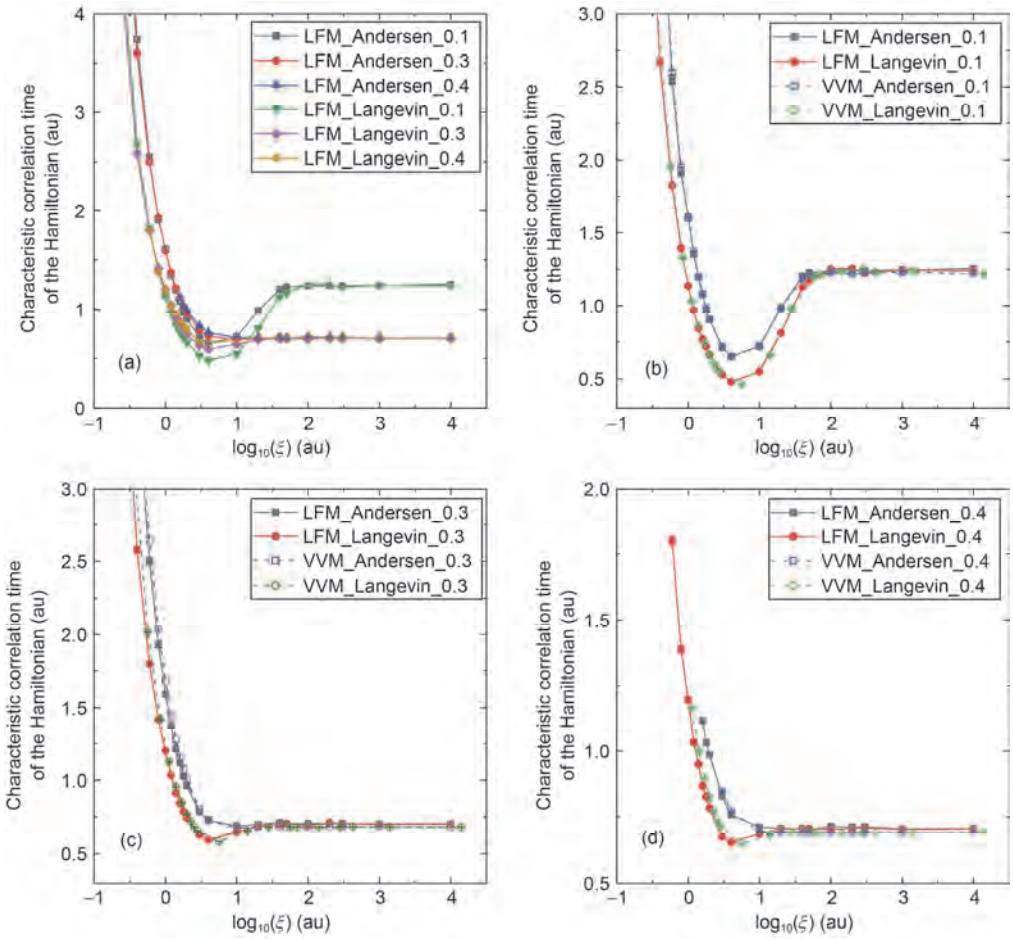


图 7 (网络版彩色)一维四次势体系哈密顿量的特征关联时间测试结果. (a) 随控温参数的变化; (b) Δt=0.1 au; (c) Δt=0.3 au; (d) Δt=0.4 au. “LFM_Andersen_0.1”表示蛙跳“居中”方案采用Andersen控温方法, 时间步长为 0.1 au, “VVM_Langevin_0.3”表示速度-Verlet“居中”方案采用Langevin控温方法, 时间步长为 0.3 au, 以此类推

Figure 7 (Color online) Numerical results for the characteristic time of the correlation function for the Hamiltonian for the one-dimensional quartic potential system. (a) Thermostat parameter, (b) Δt=0.1 au; (c) Δt=0.3 au; (d) Δt=0.4 au. “LFM_Andersen_0.1” represents the leap-frog “middle” scheme using the Andersen thermostat with Δt=0.1 au, “VVM_Langevin_0.3” denotes the velocity-Verlet “middle” scheme using the Langevin thermostat with Δt=0.3 au, etc.

总而言之, 无论对谐振子体系还是非谐体系, 两种“居中”方案对相空间的采样效率都很接近.

3 总结

本文提出了一种基于蛙跳算法的新控温方案(蛙跳“居中”方案)用于构建高效且精确的分子动力学控温算法. 在之前采用Kolmogorov算符建立的统一理论框架下[28~31], 我们重点比较了蛙跳“居中”方案和前期提出的速度-Verlet“居中”方案. 理论推导和数值测试结果均表明两种“居中”方案对位置空间采样的精度和效率相同, 都比传统控温方案(“两边”方案等)更优越, 即对于和位置(构型)相关的物理量, 可以采用大4倍以上乃至一个量级的时间步长Δt来得到相同精度的结果. 更重要的是, 与速度-Verlet“居中”方案相比, 蛙跳“居中”方案可以给出更精确的动量空间边缘分布, 对只依赖动量的物理量的计算精度能够与传统方案(“两边”方案和AMBER默认方案等)媲美. 对于有部分构型约束条件下的NVT系综的分子动力学采样, 与无约束的情况相似, 两种“居中”方案对位置空间的采样效果基本相同, 而且蛙跳“居中”方案比速度-Verlet“居中”方案对于只依赖动量的物理量的计算精度要更高. 因为两种“居中”方案对相空间的采样效率十分接近, 对于需要同时计算只依赖位置(构型)的物理量和只依赖动量的物理量的场合, 蛙跳“居中”方案

是个更佳方案. 对于计算包含位置与动量交叉项的物理量, 两种“居中”方案各有优胜, 这与具体体系和具体物理量有关. 考虑到蛙跳“居中”方案导出的稳态相空间分布只有一阶精度而速度-Verlet“居中”方案是二阶精度, 用速度-Verlet“居中”方案来计算普遍的包含位置与动量交叉项的物理量会更合理.

由于蛙跳算法在一些分子模拟软件(例如AMBER)中作为默认的分子动力学算法, 蛙跳“居中”方案无疑可以很方便地为这些软件提供新的高效的分子动力学控温算法和PIMD算法.

致谢 本工作所用计算资源得到国家超级计算广州中心、北京市计算中心、北京大学“未名一号”高性能计算平台和“北极星”高性能计算平台的支持, 特此感谢.

参考文献

1 Fermi E, Pasta J R, Ulam S M. Studies of nonlinear problems. Los Alamos Report LA-1940, 1955
2 Allen M P, Tildesley D J. Computer Simulation of Liquids. New York: Clarendon Press, 1989
3 Frenkel D, Smit B. Understanding Molecular Simulation. 2nd ed. San Diego: Academic Press, 2002
4 Andersen H C. Molecular dynamics simulations at constant pressure and/or temperature. J Chem Phys, 1980, 72: 2384–2393
5 Schneider T, Stoll E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. Phys Rev B, 1978, 17: 1302–1322
6 Brünger A, Brooks III C L, Karplus M. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. Chem Phys Lett, 1984, 105: 495–500
7 Grest G S, Kremer K. Molecular dynamics simulation for polymers in the presence of a heat bath. Phys Rev A, 1986, 33: 3628–3631
8 Mishra B, Schlick T. The notion of error in Langevin dynamics. 1. Linear analysis. J Chem Phys, 1996, 105: 299–318
9 Ricci A, Ciccotti G. Algorithms for Brownian dynamics. Mol Phys, 2003, 101: 1927–1931
10 Bussi G, Parrinello M. Accurate sampling using Langevin dynamics. Phys Rev E, 2007, 75: 056707
11 Grønbech-Jensen N, Farago O. A simple and effective Verlet-type algorithm for simulating Langevin dynamics. Mol Phys, 2013, 111: 983–991
12 Leimkuhler B, Matthews C. Robust and efficient configurational molecular sampling via Langevin dynamics. J Chem Phys, 2013, 138: 174102
13 Leimkuhler B, Matthews C. Rational construction of stochastic numerical methods for molecular sampling. Appl Math Res Express, 2013, 2013: 34–56
14 Ceriotti M, Parrinello M, Markland T E, et al. Efficient stochastic thermostatting of path integral molecular dynamics. J Chem Phys, 2010, 133: 124104
15 Liu J, Li D, Liu X. A simple and accurate algorithm for path integral molecular dynamics with the Langevin thermostat. J Chem Phys, 2016, 145: 024103
16 Liu J, Li D, Liu X. Further study of path integral Liouville dynamics. Sci China Chem, 2016, 46: 27–37
17 Liu J, Zhang Z. Path integral Liouville dynamics: Applications to infrared spectra of OH, water, ammonia, and methane. J Chem Phys, 2016, 144: 034307
18 Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Mol Phys, 1984, 52: 255–268
19 Hoover W G. Canonical dynamics: Equilibrium phase-space distributions. Phys Rev A, 1985, 31: 1695–1697
20 Martyna G J, Klein M L, Tuckerman M. Nosé-Hoover chains: The canonical ensemble via continuous dynamics. J Chem Phys, 1992, 97: 2635–2643
21 Martyna G J, Tuckerman M E, Tobias D J, et al. Explicit reversible integrators for extended systems dynamics. Mol Phys, 1996, 87: 1117–1157
22 Feynman R P. Atomic theory of the λ transition in helium. Phys Rev, 1953, 91: 1291–1301
23 Feynman R P, Hibbs A R. Quantum Mechanics and Path Integrals. New York: McGraw-Hill, 1965
24 Chandler D, Wolynes P G. Exploiting the isomorphism between quantum theory and the classical statistical mechanics of polyatomic fluids. J Chem Phys, 1981, 74: 4078–4095
25 Parrinello M, Rahman A. Study of an F center in molten KCl. J Chem Phys, 1984, 80: 860–867
26 Berne B J, Thirumalai D. On the simulation of quantum systems: Path integral methods. Annu Rev Phys Chem, 1986, 37: 401–424

27 Tuckerman M E. Statistical Mechanics: Theory and Molecular Simulation. New York: Oxford University Press, 2010
28 Zhang Z, Liu X, Chen Z, et al. A unified thermostat scheme for efficient configurational sampling for classical/quantum canonical ensembles via molecular dynamics. J Chem Phys, 2017, 147: 034109
29 Li D, Han X, Chai Y, et al. Stationary state distribution and efficiency analysis of the Langevin equation via real or virtual dynamics. J Chem Phys, 2017, 147: 184104
30 Li D, Chen Z, Zhang Z, et al. Understanding molecular dynamics with stochastic processes via real or virtual dynamics. Chin J Chem Phys, 2017, 30: 735–760
31 Liu X, Liu J. Path integral molecular dynamics for exact quantum statistics of multi-electronic-state systems. J Chem Phys, 2018, 148: 102319
32 Hall R, Berne B J. Nonergodicity in path integral molecular dynamics. J Chem Phys, 1984, 81: 3641–3643
33 Tuckerman M E, Marx D, Klein M L, et al. Efficient and general algorithms for path integral Car-Parrinello molecular dynamics. J Chem Phys, 1996, 104: 5579–5588
34 Andersen H C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. J Comp Phys, 1983, 52: 24–34
35 Ryckaert J P, Ciccotti G, Berendsen H J C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. J Comp Phys, 1977, 23: 327–341
36 Miyamoto S, Kollman P A. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. J Comp Chem, 1992, 13: 952–962
37 Case D A, Ben-Shalom I Y, Brozell S R, et al. AMBER 2018. San Francisco: University of California, 2018
38 Loncharich R J, Brooks B R, Pastor R W. Langevin dynamics of peptides: The frictional dependence of isomerization rates of *N*-acetylalanyl-*N*′-methylamide. Biopolymers, 1992, 32: 523–535
39 Jorgensen W L, Chandrasekhar J, Madura J D, et al. Comparison of simple potential functions for simulating liquid water. J Chem Phys, 1983, 79: 926–935
40 Liu J, Miller W H, Fanourgakis G S, et al. Insights in quantum dynamical effects in the infrared spectroscopy of liquid water from a semiclassical study with an *ab initio*-based flexible and polarizable force field. J Chem Phys, 2011, 135: 244503

Summary for “一种基于蛙跳算法的高效分子动力学控温方法的统一方案”

A leap-frog algorithm-based efficient unified thermostat scheme for molecular dynamics

Zhijun Zhang, Kangyu Yan, Xinzijian Liu & Jian Liu*

Beijing National Laboratory for Molecular Sciences, Institute of Theoretical and Computational Chemistry, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
* Corresponding author, E-mail: jianliupku@pku.edu.cn

An important challenge for modern molecular simulations is to develop robust algorithms for molecular dynamics (MD), which employ a time interval (i.e., step size) of evolution as large as possible to achieve high efficiency while maintaining accuracy for sampling the distribution of configuration and momentum, e.g., for the canonical ensemble where the number of particles (N), the volume (V), and the temperature (T) are constant. Starting from the velocity-Verlet algorithm, in the recent paper we have proposed a unified scheme—the “middle” scheme for constructing efficient thermostatting integrators for configurational sampling of the canonical ensemble.

In the present paper we extend the “middle” scheme to derive new thermostatting integrators based on the leap-frog algorithm. The leap-frog “middle” scheme is able to recover the exact marginal distribution of configuration and that of momentum in the harmonic limit, regardless of the time interval as long as the evolution of the MD trajectory is stable. It is also proved that the leap-frog “middle” scheme leads to the same marginal distribution of configuration as the velocity-Verlet “middle” scheme does even for general anharmonic systems. As illustrated in numerical examples (such as a 1-dimensional quartic potential, liquid water at room temperature with or without bond-length constraints, etc.), while both “middle” schemes perform better than conventional approaches (e.g., the “side” scheme and the default Langevin dynamics method in AMBER) and can calculate configuration-dependent properties with the same accuracy by using a larger time interval for from at least 4 times to even an order of magnitude, the leap-frog “middle” scheme produces a more accurate marginal distribution of momentum than that the velocity-Verlet “middle” scheme does. The sampling efficiency analysis by employing the characteristic correlation time step number (of the potential or of the Hamiltonian) suggests that both “middle” schemes share the same efficiency in sampling the configuration space and lead to similar efficiency in sampling the whole phase space (i.e., both configuration and momentum). When both only-configuration-dependent and only-momentum-dependent physical observables are considered, we suggest that the leap-frog “middle” scheme should in principle offer a better approach.

When we study properties involving cross-terms of configuration and momentum, e.g., the fluctuation of the Hamiltonian (related to the heat capacity) and the energy current and its fluctuation (related to the thermal conductivity), analytic results in the harmonic limit and numerical data for anharmonic systems indicate that which of the two “middle” schemes performs better depends on the system and on the specific property. Note that the leap-frog “middle” scheme is of first-order accuracy for sampling the whole phase space while the velocity-Verlet “middle” scheme is of second-order accuracy. It is reasonable to choose the velocity-Verlet “middle” scheme to perform simulations for such properties for general molecular systems.

Because either the leap-frog or the velocity-Verlet algorithm is the default integrator in molecular dynamics simulation software packages, the two unified “middle” schemes can offer novel, simple, accurate, and efficient thermostatting integrators for performing MD (or path integral MD (PIMD)) for the classical (or quantum) canonical ensemble.

canonical ensemble, “middle” scheme, thermostatting integrators, configurational sampling, marginal distribution of momentum

doi: 10.1360/N972018-00908