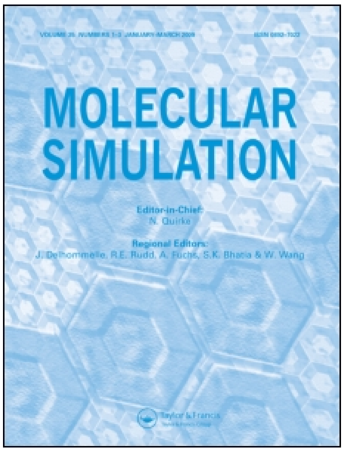

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Molecular Dynamics as a Mathematical Mapping. II. Partial Derivatives in the Microcanonical Ensemble

Jelena Stefanović[a]; Constantinos C. Pantelides[a]
[a] Imperial College of Science, Technology and Medicine, Centre for Process Systems Engineering, London, United Kingdom

To cite this Article Stefanović, Jelena and Pantelides, Constantinos C.(2001) 'Molecular Dynamics as a Mathematical Mapping. II. Partial Derivatives in the Microcanonical Ensemble', Molecular Simulation, 26: 3, 167 — 192

To link to this Article: DOI: 10.1080/08927020108028292

URL: http://dx.doi.org/10.1080/08927020108028292

Molecular Simulation, 2001, Vol. 26, pp. 167–192
Reprints available directly from the publisher
Photocopying permitted by license only

Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

MOLECULAR DYNAMICS AS A MATHEMATICAL MAPPING.
II. PARTIAL DERIVATIVES IN THE MICROCANONICAL ENSEMBLE

JELENA STEFANOVIĆ and CONSTANTINOS C. PANTELIDES*

Centre for Process Systems Engineering, Imperial College of Science, Technology and Medicine, London SW7 2BY, United Kingdom

(Received May 2000; accepted May 2000)

Performing molecular dynamics in a fully continuous and differentiable framework can be viewed as a deterministic mathematical mapping between, on one side, the force field parameters that describe the potential energy interactions and input macroscopic conditions, and, on the other, the calculated corresponding macroscopic properties of the bulk molecular system.

Within this framework, it is possible to apply standard methods of variational calculus for the computation of the partial derivatives of the molecular dynamics mapping based on the integration of either the adjoint equations or the sensitivity equations of the classical Newtonian equations of motion. We present procedures for these computations in the standard microcanonical (N, V, E) ensemble, and compare the computational efficiency of the two approaches. The general formulations developed are applied to the specific example of bulk ethane fluid.

With these procedures in place, it is now possible to compute the partial derivatives of any property determined by molecular dynamics with respect to any input property and any potential parameter. Moreover, these derivatives are computed to essentially the same level of numerical accuracy as the output properties themselves.

Keywords: Molecular dynamics; Partial derivatives; Adjoint equations; Sensitivity equations

---

*Corresponding author.

1. INTRODUCTION

Molecular dynamics solves the Newtonian equations of motion for a system of N interacting particles given by:

$$\dot{\mathbf{r}}_i = \mathbf{v}_i \quad \forall i = 1,\dots,N \tag{1}$$

$$\dot{\mathbf{v}}_i = \frac{\mathbf{F}_i(\mathbf{r},\alpha)}{m_i} \quad \forall i = 1,\dots,N \tag{2}$$

subject to the initial conditions:

$$\mathbf{r}(0) = \mathbf{r}^0(\alpha) \tag{3}$$

$$\mathbf{v}(0) = \mathbf{v}^0(\alpha) \tag{4}$$

where vectors $\mathbf{r} \in \mathbb{R}^{3N}$ and $\mathbf{v} \in \mathbb{R}^{3N}$ denote the particle positions and velocities respectively,[1] $\mathbf{F}_i$ is the total force exerted on particle i and m_i is the mass of particle i. The vector α denotes a set of constant parameters comprising the parameters ϑ that characterize the interparticle potential $U(r,\ \vartheta)$ and/or other quantities q that are directly related to macroscopic input quantities (*e.g.*, the specified system density and energy).

As shown in part I of this paper [1], the total force exerted on particle i in a direction $\gamma \in \{x,y,z\}$ can be expressed as:

$$F_i^{\gamma} = \sum_{\substack{j=1 \\ \neq i}}^{N} \mathcal{F}^{\gamma}(X_{ij}, Y_{ij}, Z_{ij}, L, \vartheta^{\mathrm{NB}}) + \Delta F_i^{\mathrm{B},\gamma}(\mathbf{r}_i; \mathbf{r}_j, j \in \mathcal{B}_i, \vartheta^{\mathrm{B}}_{\tau_i}), \qquad \gamma \in \{x,y,z\} \tag{5}$$

The first term on the right hand side of the above expression describes non-bonded interactions between particle i and all other particles j. Here X_{ij}, Y_{ij}, Z_{ij}, denote normalized interparticle distances, *e.g.*, $X_{ij} = (x_i - x_j)/L$ where L

---

[1] Throughout this paper, **bold** subscripted lower case symbols (*e.g.*, $\mathbf{r}_i$) will denote vectors of length 3 referring to a specific particle. **Bold** unsubscripted lower case symbols (*e.g.*, $\mathbf{r}$) will denote vectors of length $3N$ of the form $\mathbf{r}^T \equiv (\mathbf{r}_1^T, \mathbf{r}_2^T, \dots, \mathbf{r}_N^T)$. Bold doubly superscripted upper case symbols (*e.g.*, $\mathbf{M}^{(i,j)}$) will denote 3×3 matrices referring to a specific pair of particles i and j.

is the reference box size. $\mathcal{F}^\gamma$ is the modified force function [1] given by:

$$\mathcal{F}^x(X,Y,Z) \equiv -\frac{1}{L}\sum_{k=-\infty}^{+\infty}\sum_{k'=-\infty}^{+\infty}\sum_{k''=-\infty}^{+\infty}\frac{X-k}{R_{kk'k''}}\frac{\partial U^{NB}}{\partial R}\bigg|_{R_{ijkk'k''}} \qquad \forall(X,Y,Z)\in\mathbb{R}^3\backslash\mathcal{N}^{[3]} \tag{6}$$

and similarly for $\mathcal{F}^y(\cdot)$ and $\mathcal{F}^z(\cdot)$. The normalized interparticle distance $R_{kk'k''}$ is given by:

$$R_{kk'k''} \equiv \sqrt{(X-k)^2+(Y-k')^2+(Z-k'')^2} \tag{7}$$

The potential function U^{NB} (R) represents non-bonded interactions between pairs of particles at a normalized distance R apart. We note that U^{NB} is also a function of the potential parameters ϑ^{NB} and, as a result of normalization, of the reference box size, L (and consequently the density, ρ[2]). Consequently, the force (6) is also a function of these parameters. All these dependences can clearly be seen in the example of the standard 6–12 Lennard–Jones potential, for which:

$$\frac{\partial U^{\mathrm{NB}}}{\partial R} = -\frac{24k_b\,\varepsilon\sigma^6}{L^6}\frac{1}{R^7}\left[2\left(\frac{\sigma}{LR}\right)^6-1\right] \tag{8}$$

where $\vartheta^{\mathrm{NB}} \equiv \{\varepsilon,\sigma\}$.

The second term on the right hand side of Eq. (5) is a bond correction function that accounts for the fact that, in reality, the interactions of particle i with a subset $\mathcal{B}_i$ of the other particles in the system are due to the existence of chemical bonds. This correction is given by:

$$\begin{aligned}\Delta F_i^{\mathrm{B},\gamma}(\mathbf{r}_i;\mathbf{r}_j,j\in\mathcal{B}_i,\vartheta_{\tau_i}^{\mathrm{B}}) \equiv{}& F_i^{\mathrm{B},\gamma}(\mathbf{r}_i;\mathbf{r}_j,j\in\mathcal{B}_i,\vartheta_{\tau_i}^{\mathrm{B}})\\ &-\sum_{j\in\mathcal{B}_i}F^{\mathrm{NB},\gamma}(x_i-x_j,y_i-y_j,z_i-z_j)\\ &\forall\gamma\in\{x,y,z\}\end{aligned} \tag{9}$$

where $\vartheta_{\tau_i}^{\mathrm{B}}$ represents a set of parameters that describes a category τ_i of bonded interactions, *i.e.*, a collection of various forms of bonded interactions (such as stretching, bending, torsion) in which particle i is involved. Also, $F^{\mathrm{NB},\gamma} = -\partial U^{\mathrm{NB}}/\partial\gamma$.

---

[2] Note that $\rho = N_m M_r / N_A\, L^3$ where N_m is the number of molecules under consideration, M_r is the molecular weight and N_A is Avogadro's constant.

As will be explained in detail in Section 4, the initial conditions of the system (3)–(4) are generally also functions of α.

Most macroscopic properties p of interest determined by molecular dynamics computations involve time-integral expressions of the form:

$$p(\alpha) = \int_0^{t_f} \varphi(\mathbf{r}(t,\alpha), \mathbf{v}(t,\alpha), \alpha)dt \tag{10}$$

where t_f is some final time and φ is a continuous and differentiable function of the positions $\mathbf{r}$, the velocities $\mathbf{v}$ and the parameters α. For example, the computation of the system temperature and pressure involves the following functions φ_T and φ_P respectively [2]:

$$\varphi_T = \frac{1}{3k_bN}\sum_{i=1}^{N} m_i\mathbf{v}_i^T\mathbf{v}_i \tag{11}$$

$$\varphi_P = \frac{1}{3L^3}\sum_{i=1}^{N}(m_i\mathbf{v}_i^T\mathbf{v}_i + \mathbf{r}_i^T\mathbf{F}_i(\mathbf{r},\alpha)) \tag{12}$$

where k_b is Boltzmann's constant.

Equations (1)–(4) describe an initial value problem which, given values of the parameters α, can be solved to determine $\mathbf{r}(t)$ and $\mathbf{v}(t)$ for $t \in [0, t_f]$. This then allows the computation of the quantity p *via* Eq. (10). Thus, p is solely a function of α, as indicated on the left hand side of Eq. (10).

This paper is concerned with the computation of the partial derivatives $\partial p/\partial\alpha$ for the case of molecular dynamics calculations in the microcanonical ensemble, *i.e.*, for specified system energy E and density $\rho(q \equiv \{E, \rho\})$. In Sections 2 and 3, we present two general alternative approaches for achieving this *via* variational calculus techniques involving the solution of, respectively, the adjoint and the sensitivity equations of the Newtonian system (1)–(4). Section 4 considers the partial derivatives of the *initial* particle positions and velocities with respect to the parameters α; these quantities are needed for the computation of $\partial p/\partial\alpha$ by both the adjoint and the sensitivity formulations. Section 5 compares the computational characteristics of these two approaches. Section 6 applies our general formulations to the computation of partial derivatives for the systems of flexible ethane molecules, and presents the results of some numerical experiments performed in order to test our formulations.

2. PARTIAL DERIVATIVES *VIA* THE ADJOINT SYSTEM

One way of obtaining the partial derivatives of a function defined in terms of the solution of a set of ordinary differential equations (ODEs) is *via* the solution of the so-called *adjoint system* [3]. In order to introduce the basic concepts, we consider first the general system of n ordinary differential equations:

$$\dot{x} = f(x, \alpha) \tag{13}$$

subject to the initial conditions:

$$x(0) = x^0(\alpha) \tag{14}$$

Here, x is a vector of length n that depends on both time t and the parameters α. The latter are a vector of length N_α.

Now consider a quantity $\Phi^*(\alpha)$ defined as:

$$\Phi^*(\alpha) \equiv \int_0^{t_f} \Phi(x(t), \alpha)dt \tag{15}$$

We note that Φ^* is solely a function of α since specifying the values of α allows the solution of the ODE system (13) subject to initial conditions (14). Provided that the functions $f(x, \alpha)$, $x^0(\alpha)$ and $\Phi(x, \alpha)$ are continuous and differentiable, we can *define* the so-called "adjoint" set of ordinary differential equations:

$$\dot{\lambda} = -\left(\frac{\partial f}{\partial x}\right)^T \lambda - \left(\frac{\partial \Phi}{\partial x}\right)^T \tag{16}$$

$$\dot{\lambda}^\alpha = -\left(\frac{\partial f}{\partial \alpha}\right)^T \lambda^\alpha - \left(\frac{\partial \Phi}{\partial \alpha}\right)^T \tag{17}$$

subject to the *final* time conditions:

$$\lambda(t_f) = 0; \quad \lambda^\alpha(t_f) = 0 \tag{18}$$

Here, λ and λ^α are vectors of length n and N_α respectively. The above adjoint system is useful primarily because it can be shown that [3] the partial derivatives of Φ^* with respect to α are given by:

$$\left(\frac{\partial \Phi^*}{\partial \alpha}\right)^T = \left(\frac{\partial x^0}{\partial \alpha}\right)^T \lambda(0) + \lambda^\alpha(0) \tag{19}$$

2.1. Adjoint System of Newtonian Equations of Motion

If we apply the general equations (16), (17) to the equations of motion (1)–(4) and (10), we obtain the following adjoint system:

$$\dot{\boldsymbol{\lambda}}_i^r = -\frac{1}{m_i}\sum_{j=1}^{N}\mathbf{M}^{(j,i)}\boldsymbol{\lambda}_j^v - \left(\frac{\partial\varphi}{\partial\mathbf{r}_i}\right)^T \quad \forall i = 1,\ldots,N \tag{20}$$

$$\dot{\boldsymbol{\lambda}}_i^v = -\boldsymbol{\lambda}_i^r - \left(\frac{\partial\varphi}{\partial\mathbf{v}_i}\right)^T \quad \forall i = 1,\ldots,N \tag{21}$$

$$\dot{\lambda}^\alpha = -\sum_{i=1}^{N}\frac{1}{m_i}\left(\frac{\partial\mathbf{F}_i}{\partial\alpha}\right)^T\boldsymbol{\lambda}_i^v - \frac{\partial\varphi}{\partial\alpha} \tag{22}$$

where $\boldsymbol{\lambda}_i^r(t)$, $\boldsymbol{\lambda}_i^v(t) \in \mathbb{R}^3$, $i \in 1,\ldots, N$ represent N vectors of adjoint variables corresponding to Eqs. (1)–(2) respectively, and $\lambda^\alpha(t)$ is an additional variable for each parameter α under consideration.[3] Also, the symbol $\mathbf{M}^{(j,i)}$ represents a 3×3 matrix of the form:

$$\mathbf{M}^{(j,i)} \equiv \begin{pmatrix} \dfrac{\partial F_i^x}{\partial x_j} & \dfrac{\partial F_i^y}{\partial x_j} & \dfrac{\partial F_i^z}{\partial x_j} \\ \dfrac{\partial F_i^x}{\partial y_j} & \dfrac{\partial F_i^y}{\partial y_j} & \dfrac{\partial F_i^z}{\partial y_j} \\ \dfrac{\partial F_i^x}{\partial z_j} & \dfrac{\partial F_i^y}{\partial z_j} & \dfrac{\partial F_i^z}{\partial z_j} \end{pmatrix} \tag{23}$$

The above set of equations is subject to the *final* time conditions (*cf.* Eq. (18)):

$$\boldsymbol{\lambda}_i^r(t_f) = \boldsymbol{\lambda}_i^v(t_f) = 0; \quad \lambda^\alpha(t_f) = 0 \tag{24}$$

Finally, by applying Eq. (19) to the specific case under consideration, we can see that the partial derivatives $\partial p/\partial\alpha$ are determined from:

$$\frac{\partial p}{\partial\alpha} = \left(\frac{\partial\mathbf{r}^0}{\partial\alpha}\right)^T\boldsymbol{\lambda}^r(0) + \left(\frac{\partial\mathbf{v}^0}{\partial\alpha}\right)^T\boldsymbol{\lambda}^v(0) + \lambda^\alpha(0) \tag{25}$$

---

[3] In the interests of notational simplicity, here we prefer to formulate the equations for the case of a single parameter α and a single property p. For the multiparametric case, Eqs. (22), (24), (25) have to be written separately for each parameter α. Moreover, for the case of multiple properties p, Eqs. (20)–(22) and (25) have to be written separately for each property.

2.2. Integration of the Adjoint System

The integration of the adjoint system (20)–(22) involves evaluation of several quantities (namely $\mathbf{M}^{(j,i)}$, $\partial\varphi/\partial\mathbf{r}_i$, $\partial\varphi/\partial\mathbf{v}_i$, $\partial\mathbf{F}_i/\partial\alpha$ and $\partial\varphi/\partial\alpha$), all of which are functions of particle positions $\mathbf{r}(t)$, velocities $\mathbf{v}(t)$ and parameters α. It is, therefore, necessary to know $\mathbf{r}(t)$ and $\mathbf{v}(t)$ $\forall t\in[0, t_f]$ for the integration of the adjoint system. Moreover, although the adjoint system is a set of ODEs, it is subject to *final* (rather than initial) conditions, and therefore it has to be integrated backwards from $t=t_f$ to $t=0$. Hence, the following algorithm is employed to determine the gradients $\partial p/\partial\alpha$:

1. Perform a molecular dynamics simulation from time $t=0$ to time $t=t_f$.
2. Starting from the end system configuration (*i.e.*, $\mathbf{r}(t_f)$ and $\mathbf{v}(t_f)$) at time t_f, integrate the adjoint system of Eqs. (20)–(22) together with the original Newton's equations of motion (1)–(2) from time $t=t_f$ to $t=0$.
3. Use the values of $\boldsymbol{\lambda}^r$, $\boldsymbol{\lambda}^v$ and λ^α at $t=0$ to obtain the required gradients $\partial p/\partial\alpha$ from Eq. (25).

In practice, we are interested in the time-averaged values of the instantaneous system properties which are calculated *after* an initial equilibration period has been completed. Thus, time averaging is performed only during a final part of length τ of the total time horizon t_f, and the property of interest p can be expressed as (*cf.* Eq. (10)):

$$p(\alpha)=\frac{1}{\tau}\int_{t_f-\tau}^{t_f}\varphi(\mathbf{r},\mathbf{v},\alpha)dt \tag{26}$$

To compute the partial derivatives of the quantity p, we express Eq. (26) as:

$$p=\frac{1}{\tau}(\tilde{\tilde{p}}-\tilde{p}) \tag{27}$$

where:

$$\tilde{\tilde{p}}\equiv\int_0^{t_f}\varphi(\mathbf{r},\mathbf{v},\alpha)dt \tag{28}$$

and:

$$\tilde{p}\equiv\int_0^{t_f-\tau}\varphi(\mathbf{r},\mathbf{v},\alpha)dt \tag{29}$$

We can then solve the adjoint systems corresponding to $\tilde{\tilde{p}}$ and $\tilde{p}$. Of course, both of these adjoint systems are of exactly the same form (*i.e.*,

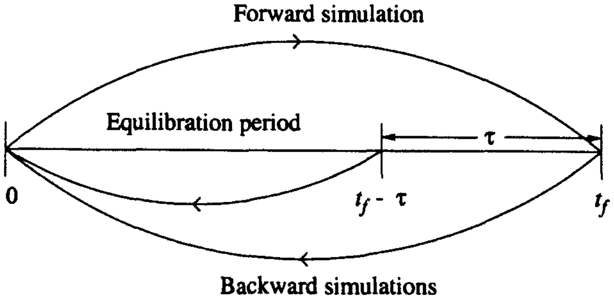

FIGURE 1 Adjoint calculation of partial derivatives of time-averaged quantities.

Eqs. (20)–(22)), but the first one is subject to the final conditions:

$$\tilde{\tilde{\boldsymbol{\lambda}}}^r(t_f) = \tilde{\tilde{\boldsymbol{\lambda}}}^v(t_f) = 0; \quad \tilde{\tilde{\lambda}}^\alpha(t_f) = 0 \tag{30}$$

while the second adjoint system is subject to:

$$\tilde{\boldsymbol{\lambda}}^r(t_f - \tau) = \tilde{\boldsymbol{\lambda}}^v(t_f - \tau) = 0; \quad \tilde{\lambda}^\alpha(t_f - \tau) = 0 \tag{31}$$

Hence we need *two* backward integrations, one starting from t_f and another from $t = t_f - \tau$, as illustrated in Figure 1.

The values of $\boldsymbol{\lambda}^r(0)$, $\boldsymbol{\lambda}^v(0)$ $\lambda^\alpha(0)$ that need to be used in Eq. (25) to determine the partial derivatives $\partial p/\partial\alpha$ of the time-averaged properties are then given by:

$$\boldsymbol{\lambda}^r(0) = \frac{1}{\tau}\left(\tilde{\tilde{\boldsymbol{\lambda}}}^r(0) - \tilde{\boldsymbol{\lambda}}^r(0)\right) \tag{32}$$

$$\boldsymbol{\lambda}^v(0) = \frac{1}{\tau}\left(\tilde{\tilde{\boldsymbol{\lambda}}}^v(0) - \tilde{\boldsymbol{\lambda}}^v(0)\right) \tag{33}$$

$$\lambda^\alpha(0) = \frac{1}{\tau}\left(\tilde{\tilde{\lambda}}^\alpha(0) - \tilde{\lambda}^\alpha(0)\right) \tag{34}$$

3. PARTIAL DERIVATIVES *VIA* THE SENSITIVITY EQUATIONS

An alternative approach to the computation of the partial derivatives $\partial p/\partial\alpha$ is provided by the solution of the *sensitivity equations*. Again assuming continuity and differentiability of the functions $f(\cdot)$, $x^0(\cdot)$ and $\Phi(\cdot)$, these can be derived simply by differentiating Eqs. (1)–(4) with respect to parameters α to obtain:

$$\dot{\mathbf{r}}_{i,\alpha} = \mathbf{v}_{i,\alpha} \quad \forall i = 1,\ldots,N \tag{35}$$

$$\dot{\mathbf{v}}_{i,\alpha} = \frac{1}{m_i}\left(\sum_{j=1}^{N} \mathbf{M}^{(j,i)}\mathbf{r}_{j,\alpha} + \frac{\partial \mathbf{F}_i}{\partial \alpha}\right) \quad \forall i = 1,\ldots,N \tag{36}$$

subject to the initial conditions:

$$\mathbf{r}_{i,\alpha}(0) = \frac{\partial \mathbf{r}_i^0}{\partial \alpha} \quad \forall i = 1,\ldots,N \tag{37}$$

$$\mathbf{v}_{i,\alpha}(0) = \frac{\partial \mathbf{v}_i^0}{\partial \alpha} \quad \forall i = 1,\ldots,N \tag{38}$$

Here, we have introduced the following notation:

$$\mathbf{r}_{i,\alpha}(t) \equiv \frac{\partial \mathbf{r}_i}{\partial \alpha}(t) \quad \text{and} \quad \mathbf{v}_{i,\alpha}(t) \equiv \frac{\partial \mathbf{v}_i}{\partial \alpha}(t)$$

The sensitivity equations (35), (36) form a set of ODEs that can be integrated simultaneously with the original equations (1), (2) to obtain the sensitivities $\mathbf{r}_{i,\alpha}(t)$ and $\mathbf{v}_{i,\alpha}(t)$ for $t \in [0, t_f]$.

By also differentiating Eq. (10) with respect to α, we obtain an expression for the required partial derivatives:

$$\frac{\partial p}{\partial \alpha} = \int_0^{t_f} \left[\sum_{i=1}^{N}\left(\frac{\partial \varphi}{\partial \mathbf{r}_i}\mathbf{r}_{i,\alpha} + \frac{\partial \varphi}{\partial \mathbf{v}_i}\mathbf{v}_{i,\alpha}\right) + \frac{\partial \varphi}{\partial \alpha}\right] dt \tag{39}$$

The integral in Eq. (39) can be evaluated simultaneously with the solution of (35), (36) in the forward direction from $t = 0$ to $t = t_f$. In practice, it is often more convenient to define an additional ODE:

$$\dot{\mu} = \sum_{i=1}^{N}\left(\frac{\partial \varphi}{\partial \mathbf{r}_i}\mathbf{r}_{i,\alpha} + \frac{\partial \varphi}{\partial \mathbf{v}_i}\mathbf{v}_{i,\alpha}\right) + \frac{\partial \varphi}{\partial \alpha} \tag{40}$$

subject to the initial condition:

$$\mu(0) = 0 \tag{41}$$

Thus, $\mu(t)$ is equal to the integral of the right hand side of (40) from the initial time $t = 0$ to time t. Hence, the required gradient is simply the final value of the variable μ:

$$\frac{\partial p}{\partial \alpha} = \mu(t_f) \tag{42}$$

We note that the computation of gradients of integrals of the more complex form (26) using the sensitivity equations still requires just one forward integration, and the partial derivatives of time-averaged quantities are obtained simply as:

$$\frac{\partial p}{\partial \alpha} = \frac{1}{\tau}(\mu(t_f) - \mu(t_f - \tau)) \tag{43}$$

As in the case of adjoint system considered in Section 2, here we have formulated the sensitivity equations for the case of a single parameter α and a single property p. For the multiparametric case, Eqs. (35)–(38) and (40)–(43) have to be written separately for each parameter α. Also, for the case of multiple properties p, Eqs. (40)–(43) have to be written separately for each property.

4. PARTIAL DERIVATIVES OF INITIAL SYSTEM CONDITIONS WITH RESPECT TO SYSTEM PARAMETERS

Both the adjoint and the sensitivity equations make use of the partial derivatives of the initial particle positions and velocities with respect to the system parameters α (see Eqs. (25) and (37), (38) respectively). In this section, we consider how these partial derivatives can be computed in the context of practical implementations of molecular dynamics computations.

At the start of the molecular dynamics computation, the "input" macroscopic conditions of the system, energy E and density ρ, are translated into the microscopic system conditions using the interparticle potential parameters ϑ (*e.g.*, the Lennard–Jones parameters ε and σ). This defines the initial system state in terms of the vector of particle positions $\mathbf{r}^0$, velocities $\mathbf{v}^0$, and the reference box size L.

4.1. Initialization Procedure for Molecular Dynamics Computations

We start by examining in detail the procedure employed to obtain the initial vector of particle positions and velocities for molecular dynamics computations in the microcanonical ensemble. This procedure involves the following steps:

Given density ρ (kg/m^3) and energy E (J/mol):

1. The reference box size L is directly related to the specified density ρ *via*:

$$L = \left(\frac{\sum_{i=1}^{N} m_i}{\rho} \right)^{1/3} \tag{44}$$

2. The particles are placed at *given* normalized positions $\mathbf{r}_i^* \in [0, 1]^3 (i = 1, \ldots, N)$ corresponding to nodes on a grid defined on a unit cube. The initial vector of particle positions within a cube of side length L is, therefore, given by:

$$\mathbf{r}_i^0 = \mathbf{r}_i^* L = \mathbf{r}_i^* \left(\frac{\sum_{i=1}^{N} m_i}{\rho} \right)^{1/3}, \quad \forall i = 1, \ldots, N \tag{45}$$

The above initial particle positions in turn determine the initial potential energy for the system of N particles. We denote this potential energy as $U(\mathbf{r}^0, L, \vartheta)$ recognizing explicitly the fact that in our modified molecular dynamics framework, the potential depends directly on L as well as on the particle positions $\mathbf{r}$ and the potential parameters ϑ (*cf.* discussion in Section 1).

3. The initial particle velocities $\mathbf{v}_i^0$ are set to be proportional to *given* velocity vectors $\mathbf{v}_i^*$.[4] Thus, we have:

$$\mathbf{v}_i^0 = \xi \mathbf{v}_i^*, \quad \forall i = 1, \ldots, N \tag{46}$$

where ξ is a scalar scaling factor. The value of ξ is chosen so that the overall kinetic energy of the system of N particles is equal to:

$$\frac{EN_m}{N_A} - U(\mathbf{r}^0, L, \vartheta) \tag{47}$$

where N_m is the number of molecules in the system[5] and N_A is Avogadro's number.[6] Thus, Eq. (46) becomes:

$$\mathbf{v}_i(0) = \sqrt{2 \frac{EN_m/N_A - U(\mathbf{r}^0, L, \vartheta)}{\sum_{j=1}^{N} m_j \mathbf{v}_j^{*T} \mathbf{v}_j^*}} \mathbf{v}_i^*, \quad \forall i = 1, \ldots, N \tag{48}$$

---

[4] Typically, the vectors $\mathbf{v}_i^*$ are generated using a (pseudo-)random number sequence, and are adjusted so that the overall momentum of the system is zero. In order to preserve the fully deterministic nature of the mapping $\mathcal{P}(q, \vartheta)$ it is essential to use the *same* set of $\mathbf{v}_i^*$ for all molecular dynamics computations performed within a particular application.

[5] Note that N_m may be smaller than the number N of particles in the system.

[6] Since E is specified in J/mol, the quantity EN_m/N_A corresponds to the total energy of the N_m molecules under consideration.

We now consider the partial derivatives of $\mathbf{r}_i^0$ and $\mathbf{v}_i^0$ with respect to the parameters α. We recall that the latter comprise both the input system properties q and the potential parameters ϑ. We deal with each one of these separately.

4.2. Partial Derivatives with Respect to Input System Quantities

In the case of the microcanonical ensemble, the input system quantities are ρ and E. By differentiating (45) with respect to these, we obtain:

$$\frac{\partial \mathbf{r}_i^0}{\partial \rho} = -\frac{\mathbf{r}_i^*\left(\sum_{i=1}^N m_i\right)^{1/3}}{3\rho^{4/3}} = -\frac{\mathbf{r}_i^0}{3\rho}, \quad \forall i = 1, \ldots, N \tag{49}$$

$$\frac{\partial \mathbf{r}_i^0}{\partial E} = 0, \quad \forall i = 1, \ldots, N \tag{50}$$

Also, by differentiating (48) with respect to ρ and taking account of the fact that both $\mathbf{r}^0$ and L are functions of ρ, we obtain:

$$\frac{\partial \mathbf{v}_i^0}{\partial \rho} = -\frac{\left[\sum_{j=1}^N (\partial U/\partial \mathbf{r}_j^0)(\partial \mathbf{r}_j^0/\partial \rho) + (\partial U/\partial L)(\partial L/\partial \rho)\right]\mathbf{v}_i^*}{\sqrt{2[EN_m/N_A - U(\mathbf{r}^0, L, \vartheta)]\sum_{j=1}^N m_j \mathbf{v}_j^{*T}\mathbf{v}_j^*}}, \quad \forall i = 1, \ldots, N \tag{51}$$

which, by virtue of (49) and (44) becomes:

$$\frac{\partial \mathbf{v}_i^0}{\partial \rho} = \frac{\left[\sum_{j=1}^N (\partial U/\partial \mathbf{r}_j^0)\mathbf{r}_j^0 + (\partial U/\partial L)L\right]\mathbf{v}_i^*}{3\rho\sqrt{2[EN_m/N_A - U(\mathbf{r}^0, L, \vartheta)]\sum_{j=1}^N m_j \mathbf{v}_j^{*T}\mathbf{v}_j^*}}, \quad \forall i = 1, \ldots, N \tag{52}$$

Finally, differentiating (48) with respect to E yields:

$$\frac{\partial \mathbf{v}_i^0}{\partial E} = \frac{N_m \mathbf{v}_i^*}{N_A\sqrt{2[EN_m/N_A - U(\mathbf{r}^0, L, \vartheta)]\sum_{j=1}^N m_j \mathbf{v}_j^{*T}\mathbf{v}_j^*}}, \quad \forall i = 1, \ldots, N \tag{53}$$

4.3. Partial Derivatives with Respect to Potential Parameters

By differentiating (45) and (48) with respect to the potential parameters ϑ, we obtain:

$$\frac{\partial \mathbf{r}_i^0}{\partial \vartheta} = 0, \quad \forall i = 1, \ldots, N \tag{54}$$

$$\frac{\partial \mathbf{v}_i^0}{\partial \vartheta} = -\frac{\mathbf{v}_i^*}{\sqrt{2[EN_m/N_A - U(\mathbf{r}^0, L, \vartheta)] \sum_{j=1}^{N} m_j \mathbf{v}_j^{*T} \mathbf{v}_j^*}} \frac{\partial U}{\partial \vartheta}, \quad \forall i = 1, \ldots, N \tag{55}$$

5. COMPUTATIONAL CONSIDERATIONS

Both the adjoint system (20)–(22), and the sensitivity equations (35), (36) and (40) allow the computation of the desired gradients $\partial p/\partial \alpha$. The relative advantages and disadvantages of the two approaches have been discussed widely in the literature in the general context of the optimization of dynamic systems [4, 5]. Here we examine the relative computational efficiency of the two approaches for the specific case of molecular dynamics.

5.1. Size of ODE System

For a system involving N interacting particles, the number, NA, of ODEs in the adjoint system (20)–(22) is given by:

$$NA \equiv (6N + N_\alpha)N_p \tag{56}$$

while the number of sensitivity equations (35), (36) and (40) is given by:

$$NS \equiv (6N + N_p)N_\alpha \tag{57}$$

where N_p and N_α denote the numbers of properties p and parameters α respectively.

From Eqs. (56) and (57), it is clear that the sensitivity system is more compact than the adjoint one (*i.e.*, $NS < NA$) if there are fewer parameters α than properties of interest, p. The reverse is true if $N_\alpha > N_p$. Thus, in general:

$$N_\alpha > N_p \Longleftrightarrow NS > NA \tag{58}$$

5.2. Evaluation of ODE Right Hand Sides

Both the adjoint and the sensitivity equations involve exactly the same functions of the vectors **r**, **v** and α, namely:

$$\mathbf{M}^{(j,i)}, \quad \frac{\partial \mathbf{F}_i}{\partial \alpha}, \quad \frac{\partial \varphi}{\partial \mathbf{r}_i}, \quad \frac{\partial \varphi}{\partial \mathbf{v}_i}, \quad \frac{\partial \varphi}{\partial \alpha}$$

and all of these have to be computed at the current values of **r** and **v** whenever one wishes to evaluate the right hand sides of the ODEs being integrated.

5.2.1. Spatial Partial Derivatives of the Force Functions

From Eqs. (5) and (23), the following formulae can be derived for the elements of matrix $\mathbf{M}^{(j,i)}$:

$$\frac{\partial F_i^\gamma}{\partial \gamma_j'} = \frac{\partial \mathcal{F}_i^\gamma}{\partial \gamma_j'} + \frac{\partial \Delta F_i^{\mathrm{B},\gamma}}{\partial \gamma_j'} \qquad \gamma, \gamma' \in \{x, y, z\} \tag{59}$$

The second term on the right hand side above can be obtained by direct differentiation of the bond correction function $\Delta F_i^{\mathrm{B},\gamma}$ (*cf.* Eq. (9) and examples in Section 6). On the other hand, the force $\mathcal{F}_i^\gamma$ appearing above denotes the first term on the right hand side of Eq. (5) and its partial derivatives are given by:

For $j \neq i$:

$$\frac{\partial \mathcal{F}_i^\gamma}{\partial \gamma_j} = \frac{1}{L^2} \sum_{k,k',k''} \left[\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{(((\gamma_i - \gamma_j)/L) - k)^2}{R_{kk'k''}^2} \left(\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{\partial^2 U^{\mathrm{NB}}}{\partial R^2} \right) \right] \tag{60}$$

$$\frac{\partial \mathcal{F}_i^\gamma}{\partial \gamma_j'} = -\frac{1}{L^2} \sum_{k,k',k''} \left[\frac{(((\gamma_i - \gamma_j)/L) - k)(((\gamma_i' - \gamma_j')/L) - k')}{R_{kk'k''}^2} \left(\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{\partial^2 U^{\mathrm{NB}}}{\partial R^2} \right) \right] \quad \forall \gamma' \neq \gamma \tag{61}$$

For $j = i$:

$$\frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_i} = -\frac{1}{L^2} \sum_{\substack{j=1 \\ \neq i}}^{N} \sum_{k,k',k''} \left[\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{(((\gamma_i - \gamma_j)/L) - k)^2}{R_{kk'k''}^2} \left(\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{\partial^2 U^{\mathrm{NB}}}{\partial R^2} \right) \right] \tag{62}$$

$$\frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_i'} = \frac{1}{L^2} \sum_{\substack{j=1 \\ \neq i}}^{N} \sum_{k,k',k''} \left[\frac{(((\gamma_i - \gamma_j)/L) - k)(((\gamma_i' - \gamma_j')/L) - k')}{R_{kk'k''}^2} \left(\frac{1}{R_{kk'k''}} \frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{\partial^2 U^{\mathrm{NB}}}{\partial R^2} \right) \right] \quad \forall \gamma' \neq \gamma \tag{63}$$

where it is understood that all partial derivatives of the potential U^{NB} are computed at an interparticle distance of $R_{kk'k''}$ as defined by Eq. (7) and the triple summation operator is defined as:

$$\sum_{k,k',k''} (\cdot) \equiv \sum_{k=-\infty}^{+\infty} \sum_{k'=-\infty}^{+\infty} \sum_{k''=-\infty}^{+\infty} (\cdot)$$

Analogously to the proofs given for the generalized modified force function in part I of this paper [1], we can show that the following properties hold:

Property I Invariance under integral shifts

$$\frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_j'}(X + k_x, Y + k_y, Z + k_z) = \frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_j'}(X, Y, Z), \quad \forall k_x, k_y, k_z \in \mathbb{Z} \quad \forall \gamma, \gamma' \in \{x, y, z\} \tag{64}$$

Property II Computation over a limited domain

$$\frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_j'}(X, Y, Z) = \frac{\partial \mathcal{F}_i^{\gamma}}{\partial \gamma_j'}(X - \lfloor X \rfloor, Y - \lfloor Y \rfloor, Z - \lfloor Z \rfloor), \quad \forall \gamma, \gamma' \in \{x, y, z\} \tag{65}$$

Property III Further reduction of the computational domain

$$\frac{\partial \mathcal{F}_i^{x}}{\partial \gamma_j}(1 - X, Y, Z) = (-1)^{(1-\delta_{\gamma x})} \frac{\partial \mathcal{F}_i^{x}}{\partial \gamma_j}(X, Y, Z) \tag{66}$$

$$\frac{\partial \mathcal{F}_i^x}{\partial \gamma_j}(X, 1-Y, Z) = (-1)^{\delta_{\gamma y}} \frac{\partial \mathcal{F}_i^x}{\partial \gamma_j}(X, Y, Z) \tag{67}$$

$$\frac{\partial \mathcal{F}_i^x}{\partial \gamma_j}(X, Y, 1-Z) = (-1)^{\delta_{\gamma z}} \frac{\partial \mathcal{F}_i^x}{\partial \gamma_j}(X, Y, Z) \tag{68}$$

where $\delta_{\gamma\gamma'}$ is the Kronecker delta:

$$\delta_{\gamma\gamma'} = \begin{cases} 1 & \text{if } \gamma = \gamma' \\ 0 & \text{otherwise} \end{cases} \tag{69}$$

and similarly for the partial derivatives of $\mathcal{F}_i^y$ and $\mathcal{F}_i^z$.

Property IV Relations between partial derivatives of $\mathcal{F}^x$, $\mathcal{F}^y$, $\mathcal{F}^z$

$$\frac{\partial \mathcal{F}_i^y}{\partial x_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial y_j}(X, Y, Z) \tag{70}$$

$$\frac{\partial \mathcal{F}_i^y}{\partial y_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial x_j}(Y, X, Z) \tag{71}$$

$$\frac{\partial \mathcal{F}_i^y}{\partial z_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial y_j}(Y, Z, X) \tag{72}$$

$$\frac{\partial \mathcal{F}_i^z}{\partial x_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial z_j}(X, Y, Z) \tag{73}$$

$$\frac{\partial \mathcal{F}_i^z}{\partial y_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial y_j}(Y, Z, X) \tag{74}$$

$$\frac{\partial \mathcal{F}_i^z}{\partial z_j}(X, Y, Z) = \frac{\partial \mathcal{F}_i^x}{\partial x_j}(Z, X, Y) \tag{75}$$

As in the case of the force functions themselves, Properties I–III imply that the domain over which the matrices of force derivatives need to be evaluated can be limited to $X \in [0, 0.5]$, $Y \in [0, 0.5]$, $Z \in [0, 0.5]$. Property IV shows that the only partial derivatives that actually need to be evaluated are those of $\mathcal{F}_i^x$.

5.2.2. *Partial Derivatives of the Force Functions with Respect to Parameters* α

The partial derivatives $\partial \mathrm{F}_i/\partial\alpha$ that appear in Eqs. (22) and (36) may be obtained by differentiating (5) with respect to α:

$$
\begin{aligned}
\frac{\partial F_i^x}{\partial\alpha} = {} & \frac{1}{L^2}\sum_{k,k',k''}\frac{k}{R_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L^2}\sum_{k,k',k''}\frac{(X-k)[k(X-k)+k'(Y-k')+k''(Z-k'')]}{R^3_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L}\sum_{k,k',k''}\frac{X-k}{R_{kk'k''}}\frac{\partial^2 U^{\mathrm{NB}}}{\partial R\partial\alpha}+\frac{\partial\Delta F_i^{\mathrm{B},x}}{\partial\alpha} \quad \forall i=1,\ldots,N
\end{aligned}
\tag{76}
$$

$$
\begin{aligned}
\frac{\partial F_i^y}{\partial\alpha} = {} & \frac{1}{L^2}\sum_{k,k',k''}\frac{k'}{R_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L^2}\sum_{k,k',k''}\frac{(Y-k')[k(X-k)+k'(Y-k')+k''(Z-k'')]}{R^3_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L}\sum_{k,k',k''}\frac{Y-k'}{R_{kk'k''}}\frac{\partial^2 U^{\mathrm{NB}}}{\partial R\partial\alpha}+\frac{\partial\Delta F_i^{\mathrm{B},y}}{\partial\alpha} \quad \forall i=1,\ldots,N
\end{aligned}
\tag{77}
$$

$$
\begin{aligned}
\frac{\partial F_i^z}{\partial\alpha} = {} & \frac{1}{L^2}\sum_{k,k',k''}\frac{k''}{R_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L^2}\sum_{k,k',k''}\frac{(Z-k'')[k(X-k)+k'(Y-k')+k''(Z-k'')]}{R^3_{kk'k''}}\frac{\partial U^{\mathrm{NB}}}{\partial R}\frac{\partial L}{\partial\alpha} \\
& -\frac{1}{L}\sum_{k,k',k''}\frac{Z-k''}{R_{kk'k''}}\frac{\partial^2 U^{\mathrm{NB}}}{\partial R\partial\alpha}+\frac{\partial\Delta F_i^{\mathrm{B},z}}{\partial\alpha} \quad \forall i=1,\ldots,N
\end{aligned}
\tag{78}
$$

We note that, unless a parameter α is either L or a quantity related to it (*e.g.*, density ρ), Eqs. (76)–(78) can be reduced to their last two terms, the other terms on the right hand sides being zero.

5.2.3. *Partial Derivatives of the Instantaneous System Properties*

The quantities $\partial\varphi/\partial\mathbf{r}$, $\partial\varphi/\partial\mathbf{v}$ and $\partial\varphi/\partial\alpha$ corresponding to the temperature and pressure functions (*cf.* Eqs. (11), (12)) are given by:

$$\frac{\partial\varphi_T}{\partial\mathbf{r}_i} = 0 \quad \forall i = 1,\ldots,N \tag{79}$$

$$\frac{\partial\varphi_T}{\partial\mathbf{v}_i} = \frac{2m_i\mathbf{v}_i}{3k_bN} \quad \forall i = 1,\ldots,N \tag{80}$$

$$\frac{\partial\varphi_T}{\partial\alpha} = 0 \tag{81}$$

$$\frac{\partial\varphi_P}{\partial\mathbf{r}_i} = \frac{1}{3L^3}\left(\mathbf{F}_i + \sum_j \mathbf{M}^{(j,i)}\mathbf{r}_j\right) \quad \forall i = 1,\ldots,N \tag{82}$$

$$\frac{\partial\varphi_P}{\partial\mathbf{v}_i} = \frac{2m_i\mathbf{v}_i}{3L^3} \quad \forall i = 1,\ldots,N \tag{83}$$

$$\frac{\partial\varphi_P}{\partial\alpha} = \frac{1}{3L^3}\sum_{i=1}^{N}\mathbf{r}_i^T\frac{\partial\mathbf{F}_i}{\partial\alpha} - \frac{3\varphi_P}{L}\frac{\partial L}{\partial\alpha} \tag{84}$$

Note that the term $\partial L/\partial\alpha$ in Eq. (84) is non-zero only if the parameter α is either L itself or a quantity directly related to it (*e.g.*, the system density ρ).

5.3. Matrix-vector Operations in ODE Right Hand Sides

The evaluation of the right hand sides of the ODEs for the two systems involve different matrix–matrix and matrix–vector operations. Bearing in mind that the multiplication of a $n_1 \times n_2$ matrix by a $n_2 \times n_3$ matrix requires $n_1 \times n_2 \times n_3$ additions and multiplications, we can perform the computational cost analysis presented in Table I.

It is clear that the dominant cost terms are $9N^2$ and $9N_\alpha N^2$ additions-multiplications for the adjoint system and sensitivity equations respectively. Therefore, the sensitivity equations are at a relative disadvantage in this respect for problems involving many parameters α (*i.e.*, $N_\alpha \gg 1$).

TABLE I Additions/multiplications for evaluation of ODE right hand sides

Type of term	*Cost of term (add + multiply)*	*No. of terms computed*	*Total cost (add + multiply)*
Adjoint system			
$\mathbf{M}^{(j,i)}\boldsymbol{\lambda}_j^v$	9	N^2	$9N^2$
$(\partial \mathbf{F}_i/\partial \alpha)^T \boldsymbol{\lambda}_i^v$	3	$N_\alpha N$	$3N_\alpha N$
Total			$3N(3N+N_\alpha)$
Sensitivity equations			
$\mathbf{M}^{(j,i)}\mathbf{r}_{j,\alpha}$	9	$N_\alpha N^2$	$9N_\alpha N^2$
$(\partial\varphi/\partial\mathbf{r}_i)\mathbf{r}_{i,\alpha}$	3	$N_\alpha N$	$3N_\alpha N$
$(\partial\varphi/\partial\mathbf{v}_i)\mathbf{v}_{i,\alpha}$	3	$N_\alpha N$	$3N_\alpha N$
Total			$3N_\alpha N(3N+2)$

5.4. Other Considerations

The integration of both adjoint and sensitivity equations requires knowledge of $\mathbf{r}(t)$ and $\mathbf{v}(t)$, $t \in [0, t_f]$. The sensitivity equations have the advantage of being integrated in the forward direction; consequently, their integration can be carried out together with that of the original Newtonian equations (1) – (4). On the other hand, as the integration of the adjoint system is carried out backwards in time, the original equations also have to be integrated backwards together with the adjoint equations, despite the fact that they have already been integrated in the forward direction (see Step 2 of the adjoint system solution algorithm outlined above). Alternatively, the solution obtained during the forward integration could be stored so that it can be used by the backward adjoint integration, but this may result in rather large storage requirements ($6NN_t$ real numbers where N_t is the number of integration steps).

Moreover, as the values of the sensitivities can be sampled at any point in the trajectory and not just at the final time t_f, the gradients of the averaged system (*i.e.*, ensemble) properties can be obtained directly from that single forward integration (*cf.* Eq. (43)). In contrast, the adjoint system requires two backward integrations starting from t_f and $t_f - \tau$ respectively (*cf.* Fig. 1) to evaluate the gradients of the averaged properties.

Overall, whether adjoints or sensitivities should be used for computation of partial derivatives will depend on the characteristics of the specific application under consideration. For example, consider a standard molecular dynamics computation in the microcanonical ensemble in which we specify values of the energy E and density ρ and calculate the bulk temperature T and pressure P. In addition, we may wish to calculate various partial derivatives:

Computation of partial derivatives with respect to input macroscopic properties Here we wish to compute the partial derivatives of T and P with respect to E and ρ. In this case, we have two system properties $p = \{P, T\}$ and two parameters $\alpha = \{E, \rho\}$, *i.e.*, $N_p = N_\alpha$. Hence, the sizes of the adjoints and the sensitivity systems are identical (*cf.* Eqs. (56), (57)). Therefore, the use of sensitivities may be preferable in view of their other advantages.

Computation of partial derivatives with respect to force field parameters Suppose we require the partial derivatives of T with respect to several potential parameters. A simple Lennard–Jones force field describing the interactions of spherical particles involves only two parameters, $\alpha = \{\varepsilon, \sigma\}$; however, a more complex force field describing flexible molecules may involve a considerable number of parameters required to describe stretching, bending and torsional molecular motion, electrostatic interactions *etc.* In such cases, the use of adjoints may be preferable.

6. APPLICATION: DYNAMICS OF SYSTEMS OF ETHANE MOLECULES

As an illustration of the methodology presented in this paper, we now apply it to the dynamics of systems of flexible ethane molecules.

6.1. Derivation of System Specific Quantities

The non-bonded interactions are described by the Lennard–Jones potential:

$$U^{\mathrm{NB}}(R,L) = 4k_b\varepsilon\left[\left(\frac{\sigma}{LR}\right)^{12} - \left(\frac{\sigma}{LR}\right)^{6}\right] \tag{85}$$

where ε and σ are the energy and distance parameters respectively. The partial derivatives of U^{NB} that need to be inserted in Eqs. (6) and (60)–(63) are given by:

$$\frac{\partial U^{\mathrm{NB}}}{\partial R} = -\frac{24k_b\varepsilon\sigma^6}{L^6}\frac{1}{R^7}\left[2\left(\frac{\sigma}{LR}\right)^6 - 1\right] \tag{86}$$

$$\frac{1}{R}\frac{\partial U^{\mathrm{NB}}}{\partial R} - \frac{\partial^2 U^{\mathrm{NB}}}{\partial R^2} = -\frac{192 k_b \varepsilon \sigma^6}{L^6}\frac{1}{R^8}\left[3.5\left(\frac{\sigma}{LR}\right)^6 - 1\right] \tag{87}$$

while Eqs. (76)–(78) involve the partial derivatives $\partial^2 U^{\mathrm{NB}}/\partial R\, \partial\alpha$ which, in the case considered here, are given by:

$$\frac{\partial^2 U^{\mathrm{NB}}}{\partial R \partial \varepsilon} = -\frac{24 k_b \sigma^6}{L^6}\frac{1}{R^7}\left[2\left(\frac{\sigma}{LR}\right)^6 - 1\right] \tag{88}$$

$$\frac{\partial^2 U^{\mathrm{NB}}}{\partial R \partial \sigma} = -\frac{144 k_b \varepsilon \sigma^5}{L^6}\frac{1}{R^7}\left[4\left(\frac{\sigma}{LR}\right)^6 - 1\right] \tag{89}$$

$$\frac{\partial^2 U^{\mathrm{NB}}}{\partial R \partial L} = \frac{144 k_b \varepsilon \sigma^6}{L^7}\frac{1}{R^7}\left[4\left(\frac{\sigma}{LR}\right)^6 - 1\right] \tag{90}$$

We now turn to the terms in our methodology that are due to bonded interactions. We assume that these are characterized by the bond stretching potential of the form:

$$U^{\mathrm{S}}(d) = k_b \frac{k_d}{2}(d - d_0)^2 \tag{91}$$

where k_b is Boltzmann's constant, k_d is the stretching parameter, d is the distance between the two bonded particles, and d_0 is the corresponding equilibrium bond length. Therefore, the force due to bond stretching exerted on a particle i by a particle j bonded to i is given by:

$$F_i^{\mathrm{B},\gamma} = -\frac{\partial U^{\mathrm{S}}}{\partial \gamma_i} = -k_b k_d \frac{d_{ij} - d_0}{d_{ij}}(\gamma_i - \gamma_j) \quad \forall \gamma \in \{x, y, z\} \tag{92}$$

where d_{ij} is the distance between particles i and j. The bonding correction function is given by the difference of Eq. (92) and the corresponding non-bonded force:

$$\Delta F_i^{\mathrm{B},\gamma} = -k_b k_d \frac{d_{ij} - d_0}{d_{ij}}(\gamma_i - \gamma_j) - \frac{24 k_b \varepsilon}{d_{ij}^2}\left[2\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^6\right](\gamma_i - \gamma_j) \quad \forall \gamma \in \{x, y, z\} \tag{93}$$

The partial derivatives of $\Delta F_i^{\mathrm{B},\gamma}$ appearing in Eqs. (60)–(63) are given by:

$$\begin{aligned}\frac{\partial \Delta F_i^{\mathrm{B},\gamma}}{\partial \gamma_i} = -\frac{\partial \Delta F_i^{\mathrm{B},\gamma}}{\partial \gamma_j} &= -k_b k_d \frac{d_{ij} - d_0}{d_{ij}} - k_b k_d \frac{d_0}{d_{ij}^3} (\gamma_1 - \gamma_2)^2 \\ &\quad - \frac{24 k_b \varepsilon}{d_{ij}^2} \left[2\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^{6} \right] \\ &\quad + \frac{192 k_b \varepsilon}{d_{ij}^4} \left[3.5\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^{6} \right] \\ &\quad \times (\gamma_i - \gamma_j)^2 \quad \forall \gamma \in \{x, y, z\} \end{aligned} \tag{94}$$

$$\begin{aligned}\frac{\partial \Delta F_i^{\mathrm{B},\gamma}}{\partial \gamma_i'} = -\frac{\partial \Delta F_i^{\mathrm{B},\gamma}}{\partial \gamma_j'} &= -k_b k_d \frac{d_0}{d_{ij}^3} (\gamma_i - \gamma_j)(\gamma_i' - \gamma_j') \\ &\quad + \frac{192 k_b \varepsilon}{d_{ij}^4} \left[3.5\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^{6} \right] \\ &\quad \times (\gamma_i - \gamma_j)(\gamma_i' - \gamma_j') \quad \forall \gamma \in \{x, y, z\} \end{aligned} \tag{95}$$

Finally, Eqs. (76)–(78) require the partial derivatives of $\Delta F_i^{\mathrm{B},\gamma}$ with respect to the potential parameters which are given by:

$$\frac{\partial \Delta F_i^{\gamma}}{\partial \varepsilon} = -\frac{24 k_b}{d_{ij}^2} \left[2\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^{6} \right] (\gamma_i - \gamma_j) \quad \forall \gamma \in \{x, y, z\} \tag{96}$$

$$\frac{\partial \Delta F_i^{\gamma}}{\partial \sigma} = -\frac{144 k_b \varepsilon}{\sigma d_{ij}^2} \left[4\left(\frac{\sigma}{d_{ij}}\right)^{12} - \left(\frac{\sigma}{d_{ij}}\right)^{6} \right] (\gamma_i - \gamma_j) \quad \forall \gamma \in \{x, y, z\} \tag{97}$$

$$\frac{\partial \Delta F_i^{\gamma}}{\partial k_d} = -k_b \frac{d_{ij} - d_0}{d_{ij}} (\gamma_i - \gamma_j) \quad \forall \gamma \in \{x, y, z\} \tag{98}$$

$$\frac{\partial \Delta F_i^{\gamma}}{\partial d_0} = k_b k_d \frac{(\gamma_i - \gamma_j)}{d_{ij}} \quad \forall \gamma \in \{x, y, z\} \tag{99}$$

6.2. Numerical Experiments

We simulate a bulk ethane fluid using the potential parameter values shown in Table II [6]. The simulation involves 108 ethane molecules at a density of 360 kg/m^3 and energy of 1000 J/mol. Here the simulation time horizon t_f is 22 ps; the time τ over which the averaging is performed is taken as $t_f/2$. The force was computed using the MOD-3 framework (*cf.* Section 5.1 of part I of this paper [1]), according to which the modified force function

TABLE II Potential parameter values and results of base case simulation for bulk ethane

Simulation of ethane fluid ($N = 216$)			
k_d (K/Å^2)	d_0(Å)	ε(K)	σ(Å)
96500	1.54	100.6	3.825
E (J/mol)	ρ (kg/m^3)	T(K)	P(MPa)
1000	360	348.8	42.5

in the γ coordinate direction is given by:

$$\mathcal{F}^{\gamma}(X_{ij},Y_{ij},Z_{ij}) = -\frac{1}{L}\sum_{k=-3}^{+3}\sum_{k'=-3}^{+3}\sum_{k''=-3}^{+3}\frac{(\gamma_i-\gamma_j)/L-k}{R_{kk'k''}}\frac{\partial U^{NB}}{\partial R}\bigg|_{R_{ijkk'k''}} \qquad \forall (X,Y,Z)\in\mathbb{R}^3\backslash\mathcal{N}^{[3]} \tag{100}$$

The results presented in part I of this paper indicate that this provides a sufficiently accurate approximation to the "true" (MOD-∞) modified function without entailing excessive computation. An Adams–Bashforth integration method [7] of order 3 with a fixed time step of 2 fs was used for the solution of the equations of motion. The energy of the system was found to be conserved within $\pm 0.5\%$.

The bottom row of Table II presents the results of the base case simulation. Table III presents the partial derivatives of temperature and pressure with respect to the input thermodynamic quantities, E and ρ, and compares these with the values obtained using the first-order finite difference formula:

$$\frac{\partial T}{\partial \alpha} \simeq \frac{T(\alpha(1+\delta)) - T(\alpha)}{\delta \times \alpha} \tag{101}$$

where the finite difference perturbation is taken as a fraction δ of α.

Tables IV and V present the partial derivatives of temperature and pressure with respect to the potential parameters $\vartheta \equiv \{k_d, d_0, \varepsilon, \sigma\}$, and compare these with the corresponding finite difference values.

TABLE III Partial derivatives of temperature and pressure with respect to input system quantities for bulk ethane

	$\partial T/\partial E$ ($K\ mol/J$)	$\partial T/\partial \rho$ ($K\ m^3/mol$)	$\partial P/\partial E$ ($MPa\ mol/J$)	$\partial P/\partial \rho$ ($MPa\ m^3/mol$)
Finite Differences:				
$\delta = 1\times10^{-2}$	3.370×10^{-2}	2.160×10^{-2}	1.260×10^{-2}	1.270×10^{-2}
$\delta = 1\times10^{-3}$	3.370×10^{-2}	2.230×10^{-2}	1.260×10^{-2}	1.200×10^{-2}
$\delta = 1\times10^{-4}$	3.400×10^{-2}	2.230×10^{-2}	1.260×10^{-2}	1.190×10^{-2}
$\delta = 1\times10^{-5}$	4.000×10^{-2}	2.250×10^{-2}	1.300×10^{-2}	1.190×10^{-2}
Adjoints:	3.371×10^{-2}	2.245×10^{-2}	1.262×10^{-2}	1.190×10^{-2}
Sensitivities:	3.371×10^{-2}	2.245×10^{-2}	1.262×10^{-2}	1.190×10^{-2}

TABLE IV Bulk ethane: partial derivatives of temperature with respect to potential parameters

	$\partial T/\partial k_d$ ($Å^2$)	$\partial T/\partial d_0$ ($K/Å$)	$\partial T/\partial \varepsilon$ (−)	$\partial T/\partial \sigma$ ($K/Å$)
Finite Differences:				
$\delta = 1 \times 10^{-2}$	-2.135×10^{-3}	-2.216×10^{1}	3.152	2.051×10^{2}
$\delta = 1 \times 10^{-3}$	-2.073×10^{-3}	-2.195×10^{1}	3.152	2.168×10^{2}
$\delta = 1 \times 10^{-4}$	-2.073×10^{-3}	-2.208×10^{1}	3.151	2.180×10^{2}
$\delta = 1 \times 10^{-5}$	0.000	-1.948×10^{1}	3.181	2.196×10^{2}
Adjoints:	-2.130×10^{-3}	-2.196×10^{1}	3.151	2.175×10^{2}
Sensitivities:	-2.130×10^{-3}	-2.196×10^{1}	3.151	2.175×10^{2}

TABLE V Bulk ethane: partial derivatives of pressure with respect to potential parameters

	$\partial P/\partial k_d$ ($MPa\,Å^2/K$)	$\partial P/\partial d_0$ ($MPa/Å$)	$\partial P/\partial \varepsilon$ (MPa/K)	$\partial P/\partial \sigma$ ($MPa/Å$)
Finite Differences:				
$\delta = 1 \times 10^{-2}$	-4.759×10^{-3}	-1.281×10^{1}	1.202×10^{-1}	1.044×10^{2}
$\delta = 1 \times 10^{-3}$	-4.777×10^{-3}	-1.249×10^{1}	1.184×10^{-1}	9.423×10^{1}
$\delta = 1 \times 10^{-4}$	-4.767×10^{-3}	-1.240×10^{1}	1.183×10^{-1}	9.323×10^{1}
$\delta = 1 \times 10^{-5}$	-4.145×10^{-3}	-1.234×10^{1}	1.193×10^{-1}	9.333×10^{1}
Adjoints:	-4.769×10^{-3}	-1.242×10^{1}	1.183×10^{-1}	9.353×10^{1}
Sensitivities:	-4.769×10^{-3}	-1.242×10^{1}	1.183×10^{-1}	9.353×10^{1}

The main observations arising from the results of Tables III–V are as follows:

- Both the adjoint and the sensitivity formulations predict the same values of all partial derivatives within (at least) the four significant digits shown here.
- The above values are generally similar to those obtained from finite difference approximations. Moreover, small perturbations applied to the values of the inputs α of this system in the context of finite difference approximations result in correspondingly small changes in the values of outputs, p. There is no evidence of "chaotic" behavior in the mapping $p(\alpha)$ despite the relatively high density (and, consequently, high frequency of interparticle "collisions") in the system being considered here.
- As is usually the case with complex functions, the selection of an appropriate size for the accurate determination of partial derivatives *via* finite difference perturbations appears to be problematic. The optimal relative perturbation (*i.e.*, the one producing values nearest to the "exact"

ones computed by the adjoint and sensitivity formulations) is generally of the order of $10^{-3}-10^{-4}$, but varies from one partial derivative to another. On the other hand, a relative perturbation of 10^{-5} results in severe loss of accuracy due to the numerical error of integration interfering with the effects of the small perturbation.

Overall, the results presented in this section provide some evidence of the advantages of variational techniques (*i.e.*, adjoints or sensitivities) for the computation of the partial derivatives of molecular dynamics mappings over simpler approaches such as those based on finite differences.

7. CONCLUDING REMARKS

This paper has presented two general methodologies for the computation of partial derivatives $\partial p/\partial \alpha$ of mappings of the form $p=p(\alpha)$ defined by molecular dynamics computations. These involve solving either the adjoint or the sensitivity equations of the classical Newtonian equations of motion. In general, both formulations require the computation of the same functions of the vectors $\mathbf{r}$, $\mathbf{v}$ and α. Amongst these is the matrix of spatial derivatives of the force functions. Hence, for these approaches to be directly applicable, the potential must be a continuous and *twice*-differentiable function of the interparticle distance r. Just like the modified force function itself, the spatial derivatives of the force function need to be evaluated in the limited domain $(X, Y, Z) \in [0, 0.5]^3$. We will exploit this property when we consider efficient ways of computing the force and its spatial derivatives in part III of this work.

In general terms, the adjoint formulation should be used if there are many more parameters α of interest than properties p. Otherwise, the sensitivity formulation is preferable in view of its relative simplicity. Leaving aside efficiency considerations, both approaches lead to the same values for the partial derivatives. Moreover, as our numerical results illustrate, these values are more reliable than those obtained by finite difference perturbations even when the latter are applied to the continuous and differentiable molecular dynamics mappings defined by our modified force functions.

Although our examples have considered only simple properties such as temperature and pressure, the approach presented is completely general and can be applied to obtain the partial derivative of any property computed by molecular dynamics with respect to any potential parameter and/or input condition.

Acknowledgments

The financial support of Mitsubishi Chemical Corporation through a Mitsubishi Fellowship to J. Stefanović is gratefully acknowledged. Partial funding has also been provided by the United Kingdom Engineering and Physical Sciences Research Council (EPSRC) under the IRC grant to the Centre for Process Systems Engineering at Imperial College.

References

[1] Stefanović, J. and Pantelides, C. C. (2000). *Molecular dynamics as a mathematical mapping. I. Differentiable force functions*, To appear in *Mol. Sim.*
[2] Allen, M. P. and Tildesley, D. J., *Computer Simulation of Liquids*. Clarendon Press, Oxford, 1987.
[3] Bryson, A. E. and Ho, Y.-C., *Applied Optimal Control*. Wiley, New York, 1975.
[4] Rosen, O. and Luus, R. (1991). Evaluation of gradients for piecewise constant optimal control. *Comput. Chem. Engng.*, **15**, 273–281.
[5] Vassiliadis, V. S., Sargent, R. W. H. and Pantelides, C. C. (1994). *Solution of a class of multistage dynamic optimization problems.* 1. *Problems without path constraints. I and EC Research*, **33**, 2111–2122.
[6] Nath, S. K., Escobedo, F. A. and de Pablo, J. J. (1998). On the simulation of the vapor–liquid equilibria for alkanes. *J. Chem. Phys.*, **108**, 9905–9911.
[7] Hairer, E., Nørsett, S. P. and Wanner, G. (1993). *Solving Ordinary Differential Equations* I. *Nonstiff Problems*. Springer-Verlag, Berlin, 2nd edition.