ELSEVIER

Computer Physics Communications 151 (2003) 272–314

Computer Physics Communications

www.elsevier.com/locate/cpc

Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations

I.P. Omelyan[a,b], I.M. Mryglod[a,b], R. Folk[b,*]

[a] *Institute for Condensed Matter Physics, 1 Svientsitskii Street, UA-79011 Lviv, Ukraine*
[b] *Institute for Theoretical Physics, Linz University, A-4040 Linz, Austria*

Received 10 June 2002; accepted 16 September 2002

Abstract

A complete classification of all explicit self-adjoint decomposition algorithms with up to 11 stages is given and the derivation process used is described. As a result, we have found 37 (6 non-gradient plus 31 force-gradient) new schemes of orders 2 to 6 in addition to 8 (4 non- and 4 force-gradient) integrators known earlier. It is shown that the derivation process proposed can be extended, in principle, to arbitrarily higher stage numbers without loss of generality. In practice, due to the restricted capabilities of modern computers, the maximal number of stages, which can be handled within the direct decomposition approach, is limited to 23. This corresponds to force-gradient algorithms of order 8. Combining the decomposition method with an advanced composition technique allows to increase the overall order up to a value of 16. The implementation and application of the introduced algorithms to numerical integration of the equations of motion in classical and quantum systems is considered as well. As is predicted theoretically and confirmed in molecular dynamics and celestial mechanics simulations, some of the new algorithms are particularly outstanding. They lead to much superior integration in comparison with known decomposition schemes such as the Verlet, Forest–Ruth, Chin, Suzuki, Yoshida, and Li integrators.

PACS: 02.60.Cb; 05.10.-a; 95.10.Ce; 95.75.Pq

Keywords: Decomposition method; Symplectic integration; Order conditions; Molecular dynamics; Celestial mechanics; Time-dependent potentials; Time-dependent Schrödinger equation

1. Introduction

The construction of decomposition algorithms for integration of motion in classical and quantum systems has been the subject of many investigations [1–16]. Within the decomposition approach the exponential operator of time evolution is split into analytically solvable parts. As a result, the symplectic map of flow of the particles in

[*] Corresponding author.
E-mail address: folk@tphys.uni-linz.ac.at (R. Folk).

0010-4655/02/$ – see front matter © 2002 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0010-4655(02)00754-3

phase space appears to be exactly conserved, despite an approximate character of the produced trajectories. The symplecticity leads in turn to the fact that the numerical uncertainties are bounded even for large sizes of the time step [17,18]. This is contrary to traditional predictor-corrector and Runge–Kutta schemes [19,20], where the errors in energy conservation increase linearly with increasing the time of the integration [15,21–24]. The decomposition algorithms are, therefore, ideal for long-duration molecular dynamics [17] and astrophysical [25] simulations.

Much of the previous work on the construction of decomposition integrators has been of a heuristic nature. Moreover, any freedom in the choice of schemes has been eliminated by restricting attention to particular methods, mainly to those which lead to a minimal number of stages at a given order. The most notorious example is the widely used Verlet algorithm [26,27], which corresponds to a three-stage decomposition scheme of the second order. The fourth-order algorithm by Forest and Ruth [2] relates to a scheme with seven single-exponential stages. Sixth-order propagations are available [4,9] beginning from fifteen such stages. With further increasing the order of decompositions, the number of stages increases drastically. This makes impossible to represent standard (non-gradient) algorithms of order eighth and higher by direct decompositions in an explicit form [9]. The reason is the impossibility to solve the resulting systems of cumbersome non-linear equations with respect to the required single-exponential time coefficients. The difficulties on the derivation of higher-order integrators have been obviated in part by composing schemes of lower (actually second) orders [1,3,5,9,28–31]. In such a way, the second-order-based composition algorithms have been explicitly obtained up to the tenth order in the time step [5,9,30].

Recently, a deeper analysis of the exponential factorization process has allowed to extent the family of analytically integrable decomposition algorithms by new members from the so-called force-gradient class [8,10, 12]. But again, as in the non-gradient case, the study has been restricted to particular solutions of the fourth order only. For example, Chin [12] introduced a set of five- and seven-stage force-gradient solutions. In order to derive the algorithms of higher orders, it has been proposed to apply a crude iterative procedure [32]. Advantages of the force-gradient integrators over their standard non-gradient (including Verlet and Forest–Ruth) counterparts have been demonstrated in celestial [12,32], stochastic [33,34], and quantum [35–37] dynamics simulations.

The question of how to improve the efficiency of decomposition schemes employing specific properties of the system has also been considered. In particular, when very slow and very fast dynamical processes are present simultaneously, the equations of motion can be integrated more productively using multiple time scale propagators [27,38]. For perturbed systems, the decomposition integration can also be carried out in a specific way, applying a smaller number of stages than usually to achieve a given order of accuracy [39–41]. When the investigated quantities need to be evaluated not after every time step of the integration but less frequently, a processing technique can be utilized to reduce the number of stages [42]. Splitting methods for non-autonomous systems with time-dependent potentials have been proposed as well [43–46]. Decomposition maps for polynomial Hamiltonian systems have been introduced too [16]. In addition, it has been shown that the decomposition approach can be adapted to integrate not only the translational motion in atomic liquids, but also simulate more complicated molecular and spin systems at the presence of orientational degrees of freedom [13–15] (for the historical development of this approach and its applications to various fields of physics, chemistry, biology and mathematics, see an excellent survey by McLachlan and Quispel [47]).

Despite these advances, there has been very little work done on a systematic construction of decomposition integrators. Some results in this direction have been obtained by Koseleff [6,11] and Li [9], but they restricted themselves to the most obvious schemes. For instance, Koseleff [6,11] focused on exhaustive search of solutions to schemes with minimal numbers of stages only. Li [9] limited his consideration to non-gradient integrators, omitting the case of gradient decompositions. Note that a full systematization of decomposition integrators presents not only a methodological interest, but may have an important practical meaning because the new solutions found can be more efficient than the already known ones. In particular, as has been shown quite recently [48–50], the decomposition propagators with minimal number of stages do not necessarily lead to optimal performance.

The primary goal of this paper is to provide a comprehensive study of a family of symplectic self-adjoint decomposition integrators. Note that the self-adjointness should be fulfilled to reproduce the property of time reversibility in solutions for non-dissipative systems (for stochastic dynamics the same is needed to recover the principle of detailed balance). In this context we should mention another family of symplectic self-adjoint schemes developed within a modified Runge–Kutta approach [51]. These schemes, however, are implicit and require the systems of non-linear equations be solved by expensive iterations at each step of the integration process. Since such equations cannot be solved exactly (because the number of the iterations is finite), the symplecticity and reversibility will be violated. These disadvantages are absent in the present decomposition method, where the time propagation is performed explicitly by a set of canonical transformations. Another benefit of our approach is that we are working with the most general solutions by considering extended decomposition schemes. Thus we are able to easily impose additional constraints, such as minimizing principal error coefficients, to choose the best integrator for each given order.

The paper is organized as follows. The basic equations of motion for classical and quantum systems, as well as a general formulation of the decomposition approach for solving these equations are presented in Section 2. The derivation of all possible decomposition algorithms with up to eleven stages and the optimization of the extended schemes are carried out in Section 3. A complete classification of the algorithms, comparison of them and choosing the best ones for orders 2, 4, and 6 are also given there. Schemes with higher numbers of stages and orders are considered in Section 4. The next Section 5 is devoted to applications of the most outstanding algorithms obtained to molecular dynamics and celestial mechanics simulations. Concluding remarks are highlighted in the final Section 6.

2. Theoretical background

2.1. Basic equations of motion in classical and quantum systems

For classical systems, the equations of motion can be written in the following compact form

$$\frac{\mathrm{d}\boldsymbol{\rho}}{\mathrm{d}t} = [\boldsymbol{\rho} \circ H] \equiv L(t)\boldsymbol{\rho}(t), \tag{1}$$

where $\boldsymbol{\rho}$ is the set of phase variables, $[\,\circ\,]$ denotes the Poisson bracket, and H represents the Hamiltonian function. In the case of N particles, placed in a spatially inhomogeneous time-dependent external field $u(\mathbf{r}_i, t)$ and interacting between themselves through the pair-wise potential $\varphi(r_{ij}) \equiv \varphi(|\mathbf{r}_i - \mathbf{r}_j|)$, the explicit expression for the Hamiltonian is

$$H = \sum_{i=1}^{N} \frac{m_i \mathbf{v}_i^2}{2} + \frac{1}{2} \sum_{i \neq j}^{N} \varphi(r_{ij}) + \sum_{i=1}^{N} u(\mathbf{r}_i, t) \equiv K(\mathbf{v}) + U(\mathbf{r}, t). \tag{2}$$

Here $\mathbf{r}_i$ denotes the position of particle i $(i = 1, 2, \ldots, N)$ moving with velocity $\mathbf{v}_i = \mathrm{d}\mathbf{r}_i/\mathrm{d}t$ and carrying mass m_i, so that K and U are the total kinetic and potential energies, respectively. Then $\boldsymbol{\rho} = \{\mathbf{r}_i, \mathbf{v}_i\} \equiv \{\mathbf{r}, \mathbf{v}\}$, and the Liouville operator of the system takes the form

$$L(t) = \sum_{i=1}^{N} \left(\mathbf{v}_i \cdot \frac{\partial}{\partial \mathbf{r}_i} + \frac{\mathbf{f}_i(t)}{m_i} \cdot \frac{\partial}{\partial \mathbf{v}_i} \right), \tag{3}$$

where $\mathbf{f}_i(t) = -\sum_{j(j\neq i)}^{N} \varphi'(r_{ij})\mathbf{r}_{ij}/r_{ij} - \partial u(\mathbf{r}_i, t)/\partial \mathbf{r}_i$ are the forces acting on the particles due to the interactions.

For quantum systems, the state evolution can be described by the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = \mathcal{H}(t)\psi \equiv \bigl(\mathcal{K} + \mathcal{U}(\mathbf{r}, t)\bigr)\psi(t), \tag{4}$$

where $\mathcal{K} = -\frac{1}{2}\sum_{i=1}^{N} \hbar^2 \boldsymbol{\nabla}_i^2/m_i$ and $\mathcal{U}$ are the kinetic and potential energy operators, respectively, and ψ is the wave function. Stochastic [33,34] and hybrid quantum-classical [52] dynamics models can also be introduced (in the latter case we come to a coupled system of Newtonian (1) and Schrödinger (4) equations). But, for the sake of simplicity, we will deal only with the above purely classic and quantum considerations.

If an initial configuration $\boldsymbol{\rho}(0)$ or $\psi(0)$ is specified, the unique solution to Eq. (1) or (4) can formally be presented as

$$\boldsymbol{\mathcal{R}}(t) = \left[\mathrm{T}\exp\left(\int_0^{\Delta t} \mathcal{L}(s)\,\mathrm{d}s\right)\right]^l \boldsymbol{\mathcal{R}}(0) \equiv \left[\exp\left[(\mathcal{D}+\mathcal{L})\Delta t\right]\right]^l \boldsymbol{\mathcal{R}}(0), \tag{5}$$

where T is the time ordering operator, Δt the size of the time step, $l = t/\Delta t$ the total number of steps, $\boldsymbol{\mathcal{R}}$ denotes either $\boldsymbol{\rho}$ or ψ, and $\mathcal{L}$ corresponds to $L(t)$ or $-i\mathcal{H}(t)/\hbar$. Suzuki proved [53] that the ordered exponential can be expressed in terms of usual exponential propagators by introducing the time derivative operator $\mathcal{D} = \overleftarrow{\partial}/\partial t$ (see the second line of Eq. (5)). Note that this operator is specific and should act on the left on time-dependent functions. In particular, $F(t)\mathrm{e}^{\mathcal{D}\Delta t} = F(t+\Delta t)$ for any function F of t. When the Liouville operator $\mathcal{L}$ does not depend explicitly on time, we are thus entitled to put $\mathcal{D} \equiv 0$. It is well known that the time propagation of many-particle systems ($N > 2$) cannot be performed exactly even in the absence of time-dependent potentials. So that the only way in handling the problem (5) is to carry out the calculations by numerical methods.

2.2. Symplectic self-adjoint decomposition schemes

The basic idea of a decomposition approach is to factor out the exponential propagator $\mathrm{e}^{(\mathcal{D}+\mathcal{L})\Delta t}$ on such suboperators which can be presented analytically or at least in quadratures. This is achieved by splitting the full operator $\mathcal{D}+\mathcal{L} = \mathcal{A}+\mathcal{B}$ into its kinetic $\mathcal{A}$ and potential $\mathcal{B}$ parts, where $\mathcal{A} = \mathcal{D} + \mathbf{v}\cdot\partial/\partial\mathbf{r}$ or $\mathcal{A} = \mathcal{D} - i\mathcal{K}/\hbar$ and $\mathcal{B} = \mathbf{a}\cdot\partial/\partial\mathbf{v}$ with $\mathbf{a} \equiv \{\mathbf{a}_i\} = \{\mathbf{f}_i/m_i\}$ being the acceleration or $\mathcal{B} = -i\mathcal{U}/\hbar$ for the cases of classical or quantum mechanics, respectively. Then the total propagator can be decomposed as

$$\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t + \mathcal{O}(\Delta t^{K+1})} = \prod_{p=1}^{P} \mathrm{e}^{\mathcal{A}a_p\Delta t}\mathrm{e}^{\mathcal{B}b_p\Delta t + \mathcal{C}c_p\Delta t^3}, \tag{6}$$

where $\mathcal{C} = [\mathcal{B},[\mathcal{A},\mathcal{B}]]$ and [,] denotes the commutator of two operators. The coefficients a_p, b_p, and c_p have to be chosen in such a way to provide the highest possible value for the order $K \geqslant 1$ at a given integer number $P \geqslant 1$. As a result, taking into account the smallness of Δt, the integration (5) can be performed approximately with the help of Eq. (6) by neglecting truncation terms $\mathcal{O}(\Delta t^{K+1})$. The precision of the integration will improve with increasing the order K and decreasing the size Δt of the time step. Formula (6) represents the decomposition approach in its most general form [50]. For $c_p \equiv 0$, Eq. (6) reduces to the standard non-gradient factorization [1–3,5,7,8,49].

In order to show that the exponential subpropagators arising on the right-hand side of Eq. (6) are indeed analytically integrable, let us first consider in more detail the structure of the third-order operator $\mathcal{C}$. In view of the expressions for operators $\mathcal{A}$ and $\mathcal{B}$, it can be obtained readily in the case of classical systems that

$$\mathcal{C} \equiv \left[\mathcal{B},[\mathcal{A},\mathcal{B}]\right] = \sum_{i=1}^{N} \frac{\mathbf{g}_i}{m_i}\cdot\frac{\partial}{\partial\mathbf{v}_i} \equiv \mathbf{G}\cdot\frac{\partial}{\partial\mathbf{v}}, \tag{7}$$

where $\mathbf{g}_{i\alpha} = 2\sum_{j\beta}\mathbf{f}_{j\beta}/m_j\partial\mathbf{f}_{i\alpha}/\partial\mathbf{r}_{j\beta}$ (α and β denote the cartesian components of vectors) and the property $[\mathcal{B},[\mathcal{D},\mathcal{B}]] = 0$ has been used. The force-gradient evaluations $\partial\mathbf{f}_{i\alpha}/\partial\mathbf{r}_{j\beta}$ can also be explicitly represented taking into account that $\mathbf{f}_{i\alpha} = m_i\mathbf{w}_{i\alpha} - \partial u(\mathbf{r}_i,t)/\partial\mathbf{r}_{i\alpha}$, where $\mathbf{w}_{i\alpha} = -\frac{1}{m_i}\sum_{j(j\neq i)}\varphi'(r_{ij})(\mathbf{r}_{i\alpha}-\mathbf{r}_{j\alpha})/r_{ij}$ is the interparticle part of the acceleration. The result is

$$\mathbf{g}_i = -2\sum_{j(j\neq i)}^{N}\left[(\mathbf{w}_i-\mathbf{w}_j)\frac{\varphi'_{ij}}{r_{ij}} + \frac{\mathbf{r}_{ij}}{r_{ij}^3}\left(r_{ij}\varphi''_{ij}-\varphi'_{ij}\right)\left(\mathbf{r}_{ij}\cdot(\mathbf{w}_i-\mathbf{w}_j)\right)\right] + \mathbf{h}_i, \tag{8}$$

where $\mathbf{h}_{i\alpha} = \frac{2}{m_i}\sum_\beta \partial u/\partial \mathbf{r}_{i\beta}\partial^2 u/\partial \mathbf{r}_{i\alpha}\partial \mathbf{r}_{i\beta}$ denotes the external-field contribution, $\varphi'_{ij} \equiv \varphi'(r_{ij}) = \mathrm{d}\varphi(r_{ij})/\mathrm{d}r_{ij}$ and $\varphi''_{ij} = \mathrm{d}\varphi'_{ij}/\mathrm{d}r_{ij}$. As can be seen, the function **G** like **a** does not depend on velocity and, thus, the operator $\mathcal{C}$ will commute with $\mathcal{B}$. Then, taking also into account the independence of **v** on **r** and the commutation of $\mathcal{D}$ with $\mathbf{v}\cdot\partial/\partial\mathbf{r}$ yields the following two equalities

$$\begin{aligned}
&\mathrm{e}^{\mathcal{A}a_p\Delta t}\{\mathbf{r},\mathbf{v}\} = \mathrm{e}^{\mathcal{D}a_p\Delta t}\{\mathbf{r}+a_p\mathbf{v}\Delta t,\mathbf{v}\},\\
&\mathrm{e}^{\mathcal{B}b_p\Delta t+\mathcal{C}c_p\Delta t^3}\{\mathbf{r},\mathbf{v}\} = \bigl\{\mathbf{r},\mathbf{v}+b_p\mathbf{a}\Delta t+c_p\mathbf{G}\Delta t^3\bigr\}
\end{aligned} \tag{9}$$

that represent simple shifts in position and velocity spaces, respectively [48–50]. Note also that the quantities **a** and **G** (being functions of **r**) may depend on time explicitly. So that any appearance of them on the left of the operator $\mathrm{e}^{\mathcal{D}a_p\Delta t}$ will shift their time coordinates, i.e. $\{\mathbf{a}(t),\mathbf{G}(t)\}\mathrm{e}^{\mathcal{D}a_p\Delta t} = \{\mathbf{a}(t+a_p\Delta t),\mathbf{G}(t+a_p\Delta t)\}$. For quantum systems, where $\mathcal{C} = i\sum_i|\nabla_i\mathcal{U}|^2/(\hbar m_i)$, the calculation of $\mathrm{e}^{\mathcal{B}b_p\Delta t+\mathcal{C}c_p\Delta t^3}$ also presents no difficulties, because this requires only the knowledge of the potential and its gradient. The kinetic part $\mathrm{e}^{\mathcal{A}a_p\Delta t}$ will require carrying out two, one direct and one inverse, spatial Fourier transforms [35].

An important feature of the decomposition integration (6) is its exact conservation of the symplectic map of flow of the particles in phase space. This follows from the fact that separate shifts (9) of positions and velocities do not change the phase volume. The property $\mathcal{S}^{-1}(t) = \mathcal{S}(-t)$ of time reversibility of evolution operators $\mathcal{S}(t) = \mathrm{e}^{(\mathcal{D}+\mathcal{L})t}$ can also be reproduced perfectly by imposing self-adjoint conditions on the coefficients a_p, b_p, and c_p. There are two kinds of such conditions [49], namely, $a_1 = 0$, $a_{p+1} = a_{P-p+1}$, $b_p = b_{P-p+1}$, $c_p = c_{P-p+1}$, as well as $a_p = a_{P-p+1}$, $b_p = b_{P-p}$, $c_p = c_{P-p}$ with $b_P = 0$ and $c_P = 0$. Then single-exponential subpropagators will enter symmetrically into the decomposition (6), providing automatically the required reversibility. The existence of these two kinds of conditions is caused by the presence of two kinds of single-exponential operators (9) and the fact that the symmetry is invariant with respect to the replacement of these operators between themselves.

2.3. Order conditions

The symmetry just mentioned will lead in turn to the automatic disappearance of all even-order terms in the error function of decomposition transformation (6), i.e.

$$\mathcal{O}\bigl(\Delta t^{K+1}\bigr) = \mathcal{O}_1\Delta t + \mathcal{O}_3\Delta t^3 + \mathcal{O}_5\Delta t^5 + \mathcal{O}_7\Delta t^7 + \cdots + \mathcal{O}_{K+1}\Delta t^{K+1} + \cdots. \tag{10}$$

We see, therefore, that the order K of self-adjoint algorithms may accept only even numbers ($K = 2, 4, 6, 8, \ldots$). The cancellation of the remaining odd-order terms (up to $\mathcal{O}_{K-1}\Delta t^{K-1}$ for order K) have to be provided by fulfilling a set of basic conditions for a_p, b_p, and c_p. Let us write down explicit expressions for the functions $\mathcal{O}_1$, $\mathcal{O}_3$, $\mathcal{O}_5$, and $\mathcal{O}_7$ (this will be enough to derive algorithms up to the eighth order). Expanding both sides of Eq. (6) into Taylor's series, and collecting the terms with the same powers of Δt one finds:

$$\mathcal{O}_1 = (\nu-1)\mathcal{A} + (\sigma-1)\mathcal{B}, \qquad \mathcal{O}_3 = \alpha\bigl[\mathcal{A},[\mathcal{A},\mathcal{B}]\bigr] + \beta\bigl[\mathcal{B},[\mathcal{A},\mathcal{B}]\bigr], \tag{11}$$

$$\begin{aligned}
\mathcal{O}_5 = {} & \gamma_1\bigl[\mathcal{A},\bigl[\mathcal{A},[\mathcal{A},[\mathcal{A},\mathcal{B}]]\bigr]\bigr] + \gamma_2\bigl[\mathcal{A},\bigl[\mathcal{A},[\mathcal{B},[\mathcal{A},\mathcal{B}]]\bigr]\bigr] + \gamma_3\bigl[\mathcal{B},\bigl[\mathcal{A},[\mathcal{A},[\mathcal{A},\mathcal{B}]]\bigr]\bigr]\\
& + \gamma_4\bigl[\mathcal{B},\bigl[\mathcal{B},[\mathcal{A},[\mathcal{A},\mathcal{B}]]\bigr]\bigr],
\end{aligned} \tag{12}$$

$$\begin{aligned}
\mathcal{O}_7 = {} & \zeta_1\bigl[\mathcal{B},\bigl[\mathcal{B},[\mathcal{A},[\mathcal{B},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr] + \zeta_2\bigl[\mathcal{B},\bigl[\mathcal{B},[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr]\\
& + \zeta_3\bigl[\mathcal{B},\bigl[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr] + \zeta_4\bigl[\mathcal{B},\bigl[\mathcal{A},[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr]\\
& + \zeta_5\bigl[\mathcal{A},\bigl[\mathcal{B},[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr] + \zeta_6\bigl[\mathcal{A},\bigl[\mathcal{B},[\mathcal{A},[\mathcal{B},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr]\\
& + \zeta_7\bigl[\mathcal{B},\bigl[\mathcal{A},[\mathcal{A},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr] + \zeta_8\bigl[\mathcal{A},\bigl[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr]\\
& + \zeta_9\bigl[\mathcal{A},\bigl[\mathcal{A},[\mathcal{B},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr] + \zeta_{10}\bigl[\mathcal{A},\bigl[\mathcal{A},[\mathcal{A},[\mathcal{A},[\mathcal{A},[\mathcal{B},\mathcal{A}]]]]\bigr]\bigr].
\end{aligned} \tag{13}$$

Here we take into account the equality $[\mathcal{B},\mathcal{C}]=0$ following from the commutation of operators $\mathcal{B}$ and $\mathcal{C}$, so that any occurrence of combinations containing the chain $[\mathcal{B},[\mathcal{B},[\mathcal{A},\mathcal{B}]]]$ has been ignored (in particular, for the fifth-order truncation term $\mathcal{O}_5$, this has allowed us to exclude two zero-valued commutators, $[\mathcal{B},[\mathcal{B},[\mathcal{B},[\mathcal{A},\mathcal{B}]]]]$ and $[\mathcal{A},[\mathcal{B},[\mathcal{B},[\mathcal{A},\mathcal{B}]]]]$).

The multipliers ν, σ, α, β, γ_{1-4}, and ζ_{1-10}, arising in Eqs. (11)–(13), will depend on the coefficients a_p, b_p, and c_p, where $p=1,2,\ldots,P$. The most simple way to obtain explicit expressions for these multipliers consists in the following. As far as we are dealing with self-adjoint schemes, the total number of single-exponential operators (stages) in Eq. (6) is actually equal to $S=2P-1$, i.e. it accepts only odd values (remember that one set of boundary coefficients is set to zero, $a_1=0$ or $b_P=c_P=0$). Then we can always choose a central single-exponential operator and apply $P-1$ times the following two types of symmetric transformations

$$\begin{aligned}
\mathrm{e}^{\mathcal{W}^{(n+1)}} &= \mathrm{e}^{\mathcal{A}a^{(n)}\Delta t}\mathrm{e}^{\mathcal{W}^{(n)}}\mathrm{e}^{\mathcal{A}a^{(n)}\Delta t},\\
\mathrm{e}^{\mathcal{W}^{(n+1)}} &= \mathrm{e}^{\mathcal{B}b^{(n)}\Delta t+\mathcal{C}c^{(n)}\Delta t^3}\mathrm{e}^{\mathcal{W}^{(n)}}\mathrm{e}^{\mathcal{B}b^{(n)}\Delta t+\mathcal{C}c^{(n)}\Delta t^3},
\end{aligned} \tag{14}$$

with

$$\mathcal{W}=(\nu\mathcal{A}+\sigma\mathcal{B})\Delta t+\mathcal{O}_3\Delta t^3+\mathcal{O}_5\Delta t^5+\mathcal{O}_7\Delta t^7+\cdots.$$

As a result, we again come to the factorization (6), where the time coefficients a_p, b_p, and c_p will be related to the quantities $a^{(n)}$, $b^{(n)}$, and $c^{(n)}$, respectively (the relationship between the subscripts n and p is determined below).

For the first kind of self-adjoint constraints (when $a_1=0$) with even P or the second one (when $b_P=c_P=0$) with odd P, the central operator of the symmetric decompositions will be correspondingly $\mathrm{e}^{\mathcal{A}a_{(P-2)/2+1}\Delta t}$ or $\mathrm{e}^{\mathcal{A}a_{(P-1)/2+1}\Delta t}$. So that here we must put $\sigma^{(0)}=0$, $\alpha^{(0)}=\beta^{(0)}=\gamma_{1-4}^{(0)}=\zeta_{1-10}^{(0)}=0$ as well as either $\nu^{(0)}=a_{(P-2)/2+1}$ or $\nu^{(0)}=a_{(P-1)/2+1}$ on the very beginning ($n=0$) of the recursive procedure (14). Moreover, we should start from the second line of Eq. (14) at $b^{(0)}=b_{(P-2)/2}$ and $c^{(0)}=c_{(P-2)/2}$ or $b^{(0)}=b_{(P-1)/2}$ and $c^{(0)}=c_{(P-1)/2}$ with further consecutive decreasing the index p by unity with increasing the number $n=1,2,\ldots,P-1$ at $a^{(n)}\equiv a_p$, $b^{(n)}\equiv b_p$, and $c^{(n)}\equiv c_p$ on both lines of transformation (14). For the first kind of constraints with odd P or the second kind with even P, the central operator is $\mathrm{e}^{\mathcal{B}b_{(P-1)/2+1}\Delta t+\mathcal{C}c_{(P-1)/2+1}\Delta t^3}$ or $\mathrm{e}^{\mathcal{B}b_{(P-2)/2+1}\Delta t+\mathcal{C}c_{(P-2)/2+1}\Delta t^3}$ that corresponds to $\sigma^{(0)}=b_{(P-1)/2+1}$ and $\beta^{(0)}=c_{(P-1)/2+1}$ or $\sigma^{(0)}=b_{(P-2)/2+1}$ and $\beta^{(0)}=c_{(P-2)/2+1}$, respectively, with $\nu^{(0)}=0$ and $\alpha^{(0)}=\gamma_{1-4}^{(0)}=\zeta_{1-10}^{(0)}=0$. In this case, the procedure should be started from the first line of Eq. (14) at $a^{(0)}=a_{(P-1)/2+1}$ or $a^{(0)}=a_{(P-2)/2+1}$ with decreasing p at increasing n for $b^{(n)}\equiv b_p$, $c^{(n)}\equiv c_p$, and $a^{(n)}\equiv a_p$.

The recursive relations between the multipliers ν, σ, α, β, and γ_{1-4} corresponding to the first line of Eq. (14) are:

$$\nu^{(n+1)}=\nu^{(n)}+2a^{(n)},\qquad \sigma^{(n+1)}=\sigma^{(n)}, \tag{15}$$

$$\alpha^{(n+1)}=\alpha^{(n)}-a^{(n)}\sigma^{(n)}\big(a^{(n)}+\nu^{(n)}\big)/6, \tag{16}$$

$$\beta^{(n+1)}=\beta^{(n)}-a^{(n)}\sigma^{(n)2}/6, \tag{17}$$

$$\begin{aligned}
\gamma_1^{(n+1)} &= \gamma_1^{(n)}+a^{(n)}\big(a^{(n)}+\nu^{(n)}\big)\big(\big(7a^{(n)2}+7a^{(n)}\nu^{(n)}+\nu^{(n)2}\big)\sigma^{(n)}-60\alpha^{(n)}\big)/360,\\
\gamma_2^{(n+1)} &= \gamma_2^{(n)}+a^{(n)}\big(30\alpha^{(n)}\sigma^{(n)}-30a^{(n)}\beta^{(n)}-30\beta^{(n)}\nu^{(n)}+3a^{(n)2}\sigma^{(n)2}\\
&\qquad +2a^{(n)}\nu^{(n)}\sigma^{(n)2}+\nu^{(n)2}\sigma^{(n)2}\big)/180,\\
\gamma_3^{(n+1)} &= \gamma_3^{(n)}+a^{(n)}\sigma^{(n)}\big(\big(8a^{(n)2}+12a^{(n)}\nu^{(n)}+\nu^{(n)2}\big)\sigma^{(n)}-120\alpha^{(n)}\big)/360,\\
\gamma_4^{(n+1)} &= \gamma_4^{(n)}+a^{(n)}\sigma^{(n)}\big(\big(6a^{(n)}+\nu^{(n)}\big)\sigma^{(n)2}-60\beta^{(n)}\big)/180.
\end{aligned} \tag{18}$$

For the second line of transformation (14) the relations read:

$$\nu^{(n+1)} = \nu^{(n)}, \qquad \sigma^{(n+1)} = \sigma^{(n)} + 2b^{(n)}, \tag{19}$$

$$\alpha^{(n+1)} = \alpha^{(n)} + b^{(n)}\nu^{(n)2}/6, \tag{20}$$

$$\beta^{(n+1)} = \beta^{(n)} + \big(12c^{(n)} + b^{(n)}\nu^{(n)}\big(b^{(n)} + \sigma^{(n)}\big)\big)/6, \tag{21}$$

$$\begin{aligned}
\gamma_1^{(n+1)} &= \gamma_1^{(n)} - b^{(n)}\nu^{(n)4}/360, \\
\gamma_2^{(n+1)} &= \gamma_2^{(n)} - \nu^{(n)}\big(60\alpha^{(n)}b^{(n)} - \nu^{(n)}\big(30c^{(n)} - b^{(n)}\nu^{(n)}\big(6b^{(n)} + \sigma^{(n)}\big)\big)\big)/180, \\
\gamma_3^{(n+1)} &= \gamma_3^{(n)} + b^{(n)}\nu^{(n)}\big(60\alpha^{(n)} + \nu^{(n)2}\big(4b^{(n)} - \sigma^{(n)}\big)\big)/360, \qquad (22)\\
\gamma_4^{(n+1)} &= \gamma_4^{(n)} - \big(30\alpha^{(n)}b^{(n)}\big(b^{(n)} + \sigma^{(n)}\big) - \nu^{(n)}\big(30\beta^{(n)}b^{(n)} + 60b^{(n)}c^{(n)} - 3b^{(n)3}\nu^{(n)} + 30c^{(n)}\sigma^{(n)} \\
&\quad - 2b^{(n)2}\nu^{(n)}\sigma^{(n)} - b^{(n)}\nu^{(n)}\sigma^{(n)2}\big)\big)/180.
\end{aligned}$$

The relations for ζ_{1-10} are presented in Appendix A. Thus, the multipliers can readily be obtained at the end (i.e. after $P-1$ steps) of the recursive process.

The forms of the first two multipliers are particularly simple and look as $\nu = \sum_{p=1}^{P} a_p$ and $\sigma = \sum_{p=1}^{P} b_p$. So that putting $\nu = 1$ and $\sigma = 1$ will cancel the first-order truncation uncertainties, $\mathcal{O}_1 = 0$. The next multipliers α, β, γ_{1-4}, ζ_{1-10} should be set to zero to kill higher-order truncation terms, namely, $\mathcal{O}_3$, $\mathcal{O}_5$, and $\mathcal{O}_7$ (see Eqs. (11)–(13)). In such a way, we come to a set of the required order conditions. They form, in fact, a system of non-linear equations which should be solved with respect to unknown time coefficients a_p, b_p, and c_p. The number M of the conditions as well as the minimal numbers $S_{\min}^{(\mathrm{f})} = 2M-1$ and $S_{\min}^{(\mathrm{g})}$ of stages, which are necessary to obtain at least the same number M of unknowns within non-gradient (f) and gradient (g) self-adjoint decomposition schemes, are displayed in Table 1 up to order $K = 8$ (the number of unknowns has been calculated taking into account the symmetry of time coefficients).

As can be seen, the minimal number of stages corresponding to a gradient scheme of an order K can be considerably smaller (especially for high K) than that of a non-gradient ($c_p \equiv 0$) integrator of the same order. Moreover, there exists a wide group of extended schemes with intermediate numbers of stages. In these schemes, the number of unknowns exceeds the number of order conditions, so that one or more parameters will be free. The free parameters can be optimized by minimizing the norm of principal error coefficients, in order to obtain the best solutions. The extended schemes will require extra computational efforts in comparison with algorithms which have minimal numbers of stages. However, at each given order one can always find such an optimal extended scheme, where these efforts will be significantly overcompensated by the increased precision of the calculations. So that the resulting efficiency will be much higher with respect to the efficiency of usual non-optimized algorithms. The proof of this hypothesis constitutes one of the main aims of the present paper.

3. Derivation and classification of schemes through 11 stages

During our derivation and classification processes, for each self-adjoint algorithm introduced we will point out the number $S = 2P - 1$ of stages, the order K of the integration as well as the numbers n_{f} and n_{g} of force and

Table 1
The minimal numbers of stages for self-adjoint decomposition schemes within up to order 8

Order	Number of conditions	Number of non-gradient stages	Number of gradient stages
2	2	3	3
4	4	7	5
6	8	15	9
8	18	35	23

force-gradient evaluations per time step. Note that the number $n_{\rm f} = P - 1 = (S+1)/2 - 1$ of force recalculations is directly connected with S. The maximal possible value of $n_{\rm g}$ is, thus, equal to $n_{\rm f}$. The minimum is equal to zero and can be achieved when all the coefficients $c_p = 0$ are equal to zero as well. The latter case will correspond to a scheme belonging to the non-gradient class. The force-gradient class will be filled up by algorithms with nonzero numbers $n_{\rm g}$. Different numbers $n_{\rm g}$ will lead to different decomposition variants within the same value of S. In addition, we will distinguish two versions, namely velocity and position ones, which are related to the first ($a_1 = 0$) and second ($b_P = c_P = 0$) kinds of boundary self-adjoint constraints, respectively.

3.1. $S = 3$ stages, order $K = 2$

3.1.1. Velocity versions

Putting $P = 2$ in Eq. (6) with $a_1 = 0$, $b_1 = b_2 = 1/2$, $a_2 = 1$, and $c_1 = c_2 \equiv \xi$, one obtains the following three-stage force-gradient scheme of the second order in velocity form

$$\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^3)} = \mathrm{e}^{\mathcal{B}\frac{\Delta t}{2}+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\Delta t}\mathrm{e}^{\mathcal{B}\frac{\Delta t}{2}+\xi\mathcal{C}\Delta t^3}, \tag{23}$$

where ξ should be treated as a free parameter. The multipliers at the leading $\mathcal{O}_3$-term (see Eq. (11)) of truncation uncertainties $\mathcal{O}(\Delta t^3)$ for this scheme are $\alpha = 1/12$ and $\beta = 1/24 + 2\xi$. Note that throughout this section, for reducing the number of unknowns, we will take into account in advance not only the symmetry of coefficients a_p, b_p, c_p, but also their fulfilling of the first two order conditions $\sum_{p=1} a_p = 1$ and $\sum_{p=1} b_p = 1$, when writing the decomposition formulas.

Variant 1. $n_g = 1$. For $\xi \neq 0$, the best choice will be to minimize the norm $(\alpha^2 + \beta^2)^{1/2}$ of the leading error coefficient $\mathcal{O}_3$ to the value $|\alpha| = 1/12$ by letting $\beta = 0$. Solving the last equation yields $\xi = -1/48$.

Variant 2. $n_g = 0$. When $\xi = 0$, Eq. (23) immediately reduces to the well-known velocity-Verlet algorithm [26, 27] with the error function $(\alpha^2 + \beta^2)^{1/2} = ((1/12)^2 + (1/24)^2)^{1/2} = \sqrt{5}/24$.

3.1.2. Position versions

The position counterpart of Eq. (23) is obtained from Eq. (6) at $a_1 = 0$, $a_1 = a_2 = 1/2$, $b_1 = 1$, $b_2 = 0$, $c_1 \equiv \xi$ and $c_2 = 0$. This gives

$$\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^3)} = \mathrm{e}^{\mathcal{A}\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\frac{\Delta t}{2}} \tag{24}$$

for which $\alpha = -1/24$ and $\beta = -1/12 + \xi$.

Variant 1. $n_g = 1$. Letting $\xi = 1/12$, to reduce β to zero, will minimize the norm $(\alpha^2 + \beta^2)^{1/2}$ of $\mathcal{O}_3$ to the level $|\alpha| = 1/24$ which is twice smaller than that corresponding to the velocity version (23) (when $|\alpha| = 1/12$ at $\xi = -1/48$).

Variant 2. $n_g = 0$. For $\xi = 0$, Eq. (24) leads to the Verlet algorithm in position form [27] with the same value $\sqrt{5}/24$ for the error function $(\alpha^2 + \beta^2)^{1/2}$ as in the case of velocity-Verlet version.

3.1.3. Remarks

It is worth remarking that within both new force-gradient integrators ($\xi \neq 0$) introduced, the order of integration does not increase, generally speaking, with respect to that of the usual Verlet algorithms ($\xi = 0$), but the norm $(\alpha^2 + \beta^2)^{1/2}$ of truncation uncertainties can be reduced from $\sqrt{5}/24$ to $|\alpha| = 1/24$. On the other hand, the new versions require one force ($n_{\rm f} = 1$) plus one force-gradient ($n_{\rm g} = 1$) evaluations per time step, whereas the Verlet approach needs only in one force recalculation ($n_{\rm f} = 1$, $n_{\rm g} = 0$). Note also that the evaluation of forces and their gradients is the most time-consuming part of the computations. In view of this, the application of the gradient approach to a particular case of $S = 3$ can especially be justified for strongly interacting systems, when the weight of kinetic part $\mathcal{A}$ in the Liouville operator $\mathcal{L}$ is much smaller than that of the potential part $\mathcal{B}$, i.e. when $\mathcal{L} = \varepsilon\mathcal{A} + \mathcal{B}$ with $\varepsilon \ll 1$. Then the remaining portion $\alpha[\mathcal{A}, [\mathcal{A}, \mathcal{B}]]\Delta t^3$ of local uncertainties $\mathcal{O}_3$ will behave like $\propto \varepsilon^2$ and can be neglected (that will correspond to a pseudo fourth-order integration).

3.2. $S = 5$ stages, order $K = 4$ or 2

3.2.1. Velocity versions

Further increasing P by unity allows us to kill exactly both the multipliers α and β, that is needed for obtaining genuine fourth-order integrators. So that choosing $P = 3$ leads to the five-stage velocity-like propagation

$$e^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^5)} = e^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3} e^{\mathcal{A}\frac{\Delta t}{2}} e^{\mathcal{B}(1-2\lambda)\Delta t+\chi\mathcal{C}\Delta t^3} e^{\mathcal{A}\frac{\Delta t}{2}} e^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3} \tag{25}$$

following from Eq. (6) at $a_1 = 0$, $a_2 = a_3 = 1/2$, $b_1 = b_3 = \lambda$, $b_2 = 1 - 2\lambda$, $c_1 = c_3 = \xi$, and $c_2 = \chi$. Here relations (16), (17), (20), and (21) result in the two order conditions

$$\alpha = -\frac{1-6\lambda}{24} = 0, \qquad \beta = -\frac{1}{12} + \frac{\lambda}{2} - \frac{\lambda^2}{2} + 2\xi + \chi = 0,$$

with the three unknowns λ, ξ, and χ. The first unknown is immediately obtained satisfying the first condition,

$$\lambda = \frac{1}{6}. \tag{26}$$

The second equality then reduces to $2\xi + \chi = 1/72$, resulting in a whole series of five-stage algorithms of the fourth order. In general, such algorithms will require two force ($n_{\mathrm{f}} = 2$) and two force-gradient ($n_{\mathrm{g}} = 2$) recalculations per time step. Remembering that we are interested in the derivation of most efficient integrators, four variants deserve to be considered.

Variant 1. $n_{\mathrm{g}} = 2$, order 4. The first obvious choice is to minimize the norm

$$\gamma = \sqrt{\gamma_1^2 + \gamma_2^2 + \gamma_3^2 + \gamma_4^2} \tag{27}$$

of the leading term $\mathcal{O}_5$ (see Eq. (12)) of fifth-order truncation errors $\mathcal{O}(\Delta t^5)$ arising when performing the time propagation with the help of scheme (25). In view of recursive relations (18) and (22), explicit expressions for the components of $\mathcal{O}_5$ are:

$$\gamma_1 = \frac{7-30\lambda}{5760}, \qquad \gamma_2 = \frac{1}{480} - \frac{\chi}{24} - \frac{\lambda^2}{24} + \frac{\xi}{6}, \qquad \gamma_3 = \frac{1}{360} - \frac{\lambda}{48} + \frac{\lambda^2}{24},$$
$$\gamma_4 = \frac{1}{120} - \frac{\lambda}{16} + \frac{7\lambda^2}{48} - \frac{\lambda^3}{8} + \frac{\xi}{6} - \frac{\chi}{2}\left(\frac{1}{3} - \lambda\right).$$

Then taking into account Eq. (26) and the equality $\chi = 1/72 - 2\xi$, where ξ is considered as a free parameter, one finds the function $\gamma = [(19 + 12240\xi + 6480000\xi^2)/2048]^{1/2}/135$ with the minimum $\gamma_{\min} = \sqrt{661}/43200 \approx 0.000595$ at

$$\xi = -\frac{17}{18000}, \qquad \chi = \frac{71}{4500}. \tag{28}$$

Variants 2 and 3. $n_{\mathrm{g}} = 1$, *order 4.* The two other cases are aimed to reduce the number of force-gradient recalculations from two to one. This is possible by choosing either

$$\xi = 0, \qquad \chi = \frac{1}{72} \tag{29}$$

or

$$\chi = 0, \qquad \xi = \frac{1}{144}. \tag{30}$$

The values of γ corresponding to the algorithms (29) and (30) are $\sqrt{19/2048}/135 \approx 0.000713$ and $7\sqrt{17}/8640 \approx 0.00334$, respectively.

Variant 4. $n_{\mathrm{g}} = 0$, *order 2.* A special case of the five-stage integration (25) is to put $\xi = 0$ as well as $\chi = 0$ that will result in a non-gradient ($n_{\mathrm{g}} = 0$) algorithm. Note that here the order decreases from 4 to 2, because it is impossible to simultaneously reduce the two functions α and β to zero having only one parameter λ. The evident strategy of the optimization in this case is to choose λ in such a way to minimize the norm $(\alpha^2 + \beta^2)^{1/2}$. Calculations show that the global minimum of that norm is equal to ≈ 0.00855 and achieved at

$$\lambda = \frac{1}{2} - \frac{(2\sqrt{326} + 36)^{1/3}}{12} + \frac{1}{6(2\sqrt{326} + 36)^{1/3}} \approx 0.1931833275037836. \tag{31}$$

3.2.2. Position versions

Position version of (25) reads

$$\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t + \mathcal{O}(\Delta t^5)} = \mathrm{e}^{\mathcal{A}\lambda\Delta t}\mathrm{e}^{\mathcal{B}\frac{\Delta t}{2} + \xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2\lambda)\Delta t}\mathrm{e}^{\mathcal{B}\frac{\Delta t}{2} + \xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\lambda\Delta t} \tag{32}$$

and can be obtained from Eq. (6) at $P = 3$ with $a_1 = a_3 = \lambda$, $a_2 = (1 - 2\lambda)$, $b_1 = b_2 = 1/2$, $c_1 = c_2 = \xi$, and $b_3 = c_3 = 0$. Now, the number of unknowns coincides with the number of the following order conditions

$$\alpha = \frac{1}{12} - \frac{\lambda}{2} + \frac{\lambda^2}{2} = 0, \qquad \beta = \frac{1}{24} - \frac{\lambda}{4} + 2\xi = 0.$$

Variants 1 and 2. $n_{\mathrm{g}} = 2$, *order 4.* Solving of these conditions yields two solutions,

$$\lambda = \frac{1}{2}\left(1 \mp \frac{1}{\sqrt{3}}\right), \qquad \xi = \frac{1}{48}(2 \mp \sqrt{3}), \tag{33}$$

for which the norm of the leading term $\mathcal{O}_5$ of truncation uncertainties $\mathcal{O}(\Delta t^5)$ appearing in Eq. (32) is $\gamma = (1873 \mp 40\sqrt{2187})^{1/2}/2160$. So that the preference should be given to sign "−", because this leads to a smaller value, $\gamma_- \approx 0.000715$, of γ, whereas $\gamma_+ \approx 0.0283$. The position algorithm (33) needs, like the velocity version (28), two force and the same number of force-gradient evaluations per time step.

Variant 3. $n_{\mathrm{g}} = 0$, *order 2.* A special case of the position-type integration (32) is the choice $\xi = 0$. Then we come to a non-gradient scheme requiring only two force recalculations. But, the price we pay for the decreasing of computational efforts is the reduction of the order from 4 to 2. Here, the minimization of $(\alpha^2 + \beta^2)^{1/2}$ with respect to λ leads to the same result (31) as in the case of the non-gradient velocity version.

3.2.3. Remarks

The first attempt of building gradient five-stage decompositions like (25) and (32) has been reported by Koseleff [6]. But, in fact, he considered only velocity-like scheme (25) and, in addition, wrote incorrectly the order conditions. Integrators (29) and (33) have been previously derived by Suzuki [8] basing on McLachlan's method of small perturbation [39] and have been referred by Chin [12] as schemes A and B, respectively. Algorithms (28) and (30) are new. While scheme (30) seems to have no advantages over the integrator A, the new algorithm (28) corresponds to the best accuracy of the integration (among the above five-stage solutions), because it leads to the smallest value of γ. It requires, however, one extra force-gradient evaluation and, thus, can be recommended for systems when this evaluation does not present significant difficulties.

It is interesting also to compare the efficiency of our new non-gradient five-stage algorithms of the second order (see variants 4 and 3 within velocity and position versions, respectively) with that of the original three-stage Verlet integrators (Eqs. (23) and (24), $\xi = 0$) of the same order. This efficiency can be measured in terms of the norm Γ_2 of global errors. Note that during a whole integration over a fixed time interval t, the total number l of single steps is proportional to Δt^{-1} (see Eq. (5)). As a result, the norm of local third-order uncertainties $\|o_3\| \equiv \|\mathcal{O}_3\|\Delta t^3 = (\alpha^2+\beta^2)^{1/2}\Delta t^3$ will accumulate step by step during the integration process and lead at $t \gg \Delta t$ to the second-order global errors $\Gamma_2(\Delta t) = \|o_3\|\Delta t^{-1} \equiv (\alpha^2+\beta^2)^{1/2}\Delta t^2$. Since the non-gradient five-stage propagations require two force evaluation per time step Δt, they should be performed with the double step size $2\Delta t$ with respect to that of the Verlet algorithms, in order to provide the same number of total force recalculations. Thus, these extended propagations will be more efficient if the following inequality $\Gamma_2^{(S=5)}(2\Delta t) < \Gamma_2^{(S=3)}(\Delta t)$ is satisfied. Taking into account that $(\alpha^2+\beta^2)^{1/2}$ is equal to $\sqrt{5}/24 \approx 0.0932$ or ≈ 0.00855 within the Verlet ($S=3$) or extended ($S=5$) schemes, respectively, one obtains $\Gamma_2^{(S=5)}(2\Delta t)/\Gamma_2^{(S=3)}(\Delta t) \approx 0.367$. So that the extended second-order integrators, being applied even with double sizes of the time step, will reduce the global errors approximately in three times. We see, therefore, that propagations with minimal number of stages may not lead to optimal performance. A similar pattern will be observed for higher-order decomposition schemes.

3.3. $S = 7$ *stages, order* $K = 4$

3.3.1. *Velocity versions*

The seven-stage propagation in velocity form is

$$\begin{aligned}&\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^5)}\\ &\quad = \mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}(1-2\lambda)\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2\theta)\Delta t}\mathrm{e}^{\mathcal{B}(1-2\lambda)\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\end{aligned} \tag{34}$$

that can be reproduced from the general formula (6) at $P = 4$. Note that further we will not present the relationship between the coefficients a_p, b_p, c_p of Eq. (6) and reduced variables (such as, for example, λ, ξ, θ, and χ in Eq. (34)) in view of its evidence. The order conditions for scheme (34) read

$$\begin{aligned}\alpha &= \frac{1}{12} - \frac{\theta}{2} + \lambda\theta(1-\theta) + \frac{\theta^2}{2} = 0,\\ \beta &= \frac{1}{24} + \lambda\theta(1-\lambda) + 2\xi + 2\chi - \frac{\theta}{4} = 0\end{aligned}$$

with the solutions

$$\begin{aligned}\lambda &= \frac{1}{12}\left(6 + \frac{1}{\theta(\theta-1)}\right),\\ \xi &= -\frac{1}{288}\left(6 + 288\chi - \frac{1}{\theta(\theta-1)^2}\right)\end{aligned} \tag{35}$$

at the existence of two free parameters θ and χ. Taking into account these solutions yields the following γ-components

$$\gamma_1 = -\frac{1}{720}\left(1 - 5\theta + 5\theta^2\right),$$

$$\gamma_2 = \frac{1}{2880(\theta-1)^2}\big(1 - 2880\chi\theta\big[1 - 3\theta + 3\theta^2 - \theta^3\big] - 2\theta + 6\theta^2\big),$$
$$\gamma_3 = -\frac{1}{4320(\theta-1)^2}\big(3 - 16\theta + 18\theta^2\big),$$
$$\gamma_4 = -\frac{1}{17280\,\theta(\theta-1)^3}\big(5 - 2880\chi\theta\big[1 - 2\theta + \theta^2\big] - 36\,\theta\big[1 - 3\theta + 3\theta^2 - \theta^3\big]\big)$$

of $\mathcal{O}_5$, so that $\gamma \equiv \gamma(\theta, \chi)$ (see Eqs. (12) and (27)).

Variant 1. $n_{\mathrm{g}} = 3$. Performing the total minimization of the error function $\gamma(\theta, \chi)$ with respect to the two parameters θ and χ, one finds that the global minimum is $\gamma_{\min} \approx 0.0000167$ and achieved at

$$\begin{aligned} &\theta = 0.2728983001988755\mathrm{E}{+}00, \\ &\chi = 0.2960781208329478\mathrm{E}{-}02, \quad \text{so that} \\ &\lambda = 0.8002565306418866\mathrm{E}{-}01, \\ &\xi = 0.2725753410753895\mathrm{E}{-}03. \end{aligned} \tag{36}$$

Variant 2. $n_{\mathrm{g}} = 2$. Letting $\xi = 0$, to decrease the number of gradient recalculations from 3 to 2, allows to express χ in terms of θ, namely, $\chi = (1 - 6\theta + 12\theta^2 - 6\theta^3)/(288(\theta-1)^2\theta))$ (see the second line of Eq. (35)). Then $\gamma(\theta, \chi(\theta)) \equiv \gamma(\theta)$ and the task reduces to the minimization of γ with respect to one parameter θ. The result is

$$\begin{aligned} &\theta = 0.2813980611667719\mathrm{E}{+}00, \quad \text{so that} \\ &\lambda = 0.8789368601680709\mathrm{E}{-}01, \\ &\chi = 0.3061810122369770\mathrm{E}{-}02 \end{aligned} \tag{37}$$

with $\gamma_{\min} \approx 0.0000443$.

Variant 3. $n_{\mathrm{g}} = 1$. At $\chi = 0$, when the number gradient recalculations decreases to 1, the global minimum of $\gamma(\theta, \chi = 0) \equiv \gamma(\theta)$ is reached at

$$\begin{aligned} &\theta = 0.2409202729169543\mathrm{E}{+}00, \quad \text{so that} \\ &\lambda = 0.4432204907934768\mathrm{E}{-}01, \\ &\xi = 0.4179297897540420\mathrm{E}{-}02 \end{aligned} \tag{38}$$

and is equal to $\gamma_{\min} \approx 0.000855$.

Variant 4. $n_{\mathrm{g}} = 0$. Putting $\xi = 0$ and $\chi = 0$, one finds from Eq. (35) that

$$\theta = \frac{1}{2 - \sqrt[3]{2}}, \qquad \lambda = \frac{\theta}{2}. \tag{39}$$

Non-gradient solution (39) corresponds to the usual Forest–Ruth algorithm [2] in velocity form. For this algorithm, the error function takes the value $\gamma \approx 0.0383$.

3.3.2. Position versions

The seven-stage propagation in position form (that can also be reproduced at $P = 4$ from Eq. (6)) is

$$\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t + \mathcal{O}(\Delta t^5)} = \mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t + \xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2\theta)\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}(1-2\lambda)\Delta t + \chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2\theta)\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}\lambda\Delta t + \xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t} \tag{40}$$

and the order conditions are

$$\alpha = -\frac{1}{24} + \lambda\left(\frac{1}{4} - \theta + \theta^2\right) = 0,$$

$$\beta = -\frac{1}{12} + \frac{\lambda}{2} - \frac{\lambda^2}{2} - \lambda\theta(1-\lambda) + 2\xi + \chi = 0.$$

Solving of these conditions by treating λ and χ as free parameters gives

$$\xi = \frac{1}{24}\bigl(1 - 12\chi \pm \sqrt{6\lambda}(1-\lambda)\bigr), \qquad \theta = \frac{1}{2} \pm \frac{1}{\sqrt{24\lambda}}. \tag{41}$$

This leads to

$$\gamma_1 = -\frac{1}{1920} + \frac{1}{6912\lambda}, \qquad \gamma_2 = \frac{1}{2880\lambda}\bigl(5 + 6\lambda \pm 5\sqrt{6\lambda} - 60\chi\bigr),$$

$$\gamma_3 = -\frac{1}{360}\left(\frac{3}{2} \pm \frac{5}{\sqrt{96\lambda}} \pm 5\sqrt{\frac{\lambda}{24}}\right), \qquad \gamma_4 = -\frac{11}{480} \pm \left(\frac{1}{\sqrt{\lambda}}\left[\frac{\chi}{\sqrt{96}} - \frac{1}{\sqrt{13824}}\right] - \chi\sqrt{\frac{\lambda}{24}}\right) + \frac{\lambda}{32} - \frac{\lambda^2}{48}$$

with $\gamma \equiv \gamma_\pm(\lambda,\chi)$.

Variant 1. $n_{\mathrm{g}} = 3$. Considering $\gamma_\pm(\lambda,\chi)$ as two functions of the two parameters each, one finds that the global minimum is equal to $\gamma_{\min} \approx 0.0000123$ and achieved at sign minus when

$$\begin{aligned} \lambda &= 0.2825633404177051\mathrm{E}{+}00, \\ \chi &= 0.3035236056708454\mathrm{E}{-}02, \quad \text{so that} \\ \theta &= 0.1159953608486416\mathrm{E}{+}00, \\ \xi &= 0.1226088989536361\mathrm{E}{-}02. \end{aligned} \tag{42}$$

Variant 2. $n_{\mathrm{g}} = 2$. At $\chi = 0$, the global minimum of $\gamma_\pm(\lambda, \chi = 0) \equiv \gamma_\pm(\lambda)$ is $\gamma_{\min} \approx 0.0000823$ and can be obtained at sign minus when

$$\begin{aligned} \lambda &= 0.3152315246820299\mathrm{E}{+}00, \quad \text{so that} \\ \theta &= 0.1364371009136296\mathrm{E}{+}00, \\ \xi &= 0.2427211985487600\mathrm{E}{-}02. \end{aligned} \tag{43}$$

Variant 3. $n_{\mathrm{g}} = 1$. For $\xi = 0$, when $\chi = (1 \pm \sqrt{6\lambda}(1-\lambda))/12$ (see the first line of Eq. (41)), the global minimum of $\gamma_\pm(\lambda, \chi(\lambda)) \equiv \gamma_\pm(\lambda)$ is equal to $\gamma_{\min} \approx 0.000141$ that can be observed at sign minus and

$$\begin{aligned} \lambda &= 0.2470939580390842\mathrm{E}{+}00, \quad \text{so that} \\ \theta &= 0.8935804763220157\mathrm{E}{-}01, \\ \chi &= 0.6938106540706989\mathrm{E}{-}02. \end{aligned} \tag{44}$$

Variant 4. $n_{\mathrm{g}} = 0$. Putting $\xi = 0$ and $\chi = 0$, one obtains from Eq. (41) that the real solution is possible at sign plus and it is

$$\lambda = \frac{1}{2 - \sqrt[3]{2}}, \qquad \theta = \lambda/2. \tag{45}$$

This leads to the usual non-gradient algorithm by Forest and Ruth [2] in position form for which $\gamma \approx 0.0283$.

3.3.3. Remarks

Earlier, Chin [12] introduced an algorithm similar to the position-like scheme (40) when $\xi = 0$. But he derived it in a somewhat different way, namely, using a symmetric concatenation of two third-order schemes (to increase the order from three to four). This resulted only in one particular set of time coefficients, called by Chin as scheme C, which can be reproduced from Eq. (40) at

$$\lambda = \frac{3}{8}, \quad \theta = \frac{1}{6}, \quad \chi = \frac{1}{192}, \tag{46}$$

with $\gamma = \sqrt{87817}/414720 \approx 0.000715$. Within our general approach, we have shown that even at $\xi = 0$, there exists a whole λ-series of fourth-order solutions (41). Moreover, it can be seen easily that the scheme C is not optimal. Indeed, the new solution found by us (see Eq. (44)) reduces the norm of truncation errors to the value $\gamma \approx 0.000141$ that is approximately in 5 times smaller than that corresponding to scheme C (and such a reduction is achieved without spending any additional computational efforts).

Nearly the same situation happens with the case of velocity-like scheme (34). Previously Chin and Chen [35], using again the symmetric product of two third-order integrators, have obtained at $\chi = 0$ the following set

$$\lambda = \frac{1}{8}, \quad \theta = \frac{1}{3}, \quad \xi = \frac{1}{384} \tag{47}$$

of time coefficients and referred it to scheme D. We have realized that this is not the only possible set and we found a whole series of solutions (35) which includes (47). The optimal set (see Eq. (38)), obtained within this series, differs from (47) and leads to a lower level, $\gamma \approx 0.000855$, of truncation errors with respect to that of scheme D for which $\gamma \approx 0.00117$. The velocity version (38) may be more preferable than its position counterpart (44) when performing quantum mechanics simulations, because it requires a fewer number of spatial Fourier transforms.

A distinguishable feature of the seven- as well as five-stage gradient schemes of the fourth order (introduced in Sections 3.2 and 3.3) is the fact that all their time coefficients $\{a_p, b_p\}$, arising at basic operators $\mathcal{A}$ and $\mathcal{B}$ under exponentials, are positive. This is in a sharp contrast with the non-gradient decomposition approach, where beyond second order (as has been rigorously proved by Suzuki [3]) any scheme expressed in terms of only force evaluations must produce some negative time coefficients. For example, in the case of velocity Forest–Ruth integration (Eqs. (34) and (39)), three of eight coefficients, namely, $a_3 = (1 - 2\theta)$, and $b_2 = b_3 = (1 - 2\lambda)/2 = (1 - \theta)/2$, are less than zero. In this context it should be emphasized that the propagation with negative times is impossible, in principle, in such important fields as stochastic dynamics, non-equilibrium statistical mechanics, quantum statistics, etc., because one cannot simulate diffusion or stochastic processes backward in time nor sample configurations with negative temperatures. Therefore, contrary to the usual non-gradient Forest–Ruth schemes, the force-gradient algorithms can simulate dynamical processes in all areas of physics and chemistry without any principal restrictions.

Another advantage of force-gradient algorithms is the possibility to significantly improve the precision of the computations. For instance, during the gradient evaluation (44), when $\gamma \approx 0.000141$ as well as $n_{\mathrm{f}} = 3$ and $n_{\mathrm{g}} = 1$, the norm of truncation uncertainties decreases by more than 200 times with respect to that of the non-gradient ($n_{\mathrm{g}} \equiv 0$) algorithm (45) by Forest and Ruth for which $\gamma \approx 0.0283$ as well as $n_{\mathrm{f}} = 3$. The resulting efficiency increases too, despite an increased computational work spent on the evaluations of force gradients. Indeed, in view of expression (8), it can be estimated that for many-particle systems this work takes a processor time which is approximately a factor of $g = 2$ to 3 (depending on the structure of interparticle potentials φ) higher than that related to the calculation of forces themselves (note that when we are proceeding to evaluate the gradients $\mathbf{g}_i$, the interparticle forces and, thus, the accelerations $\mathbf{w}_i$, appearing in Eq. (8), are already known quantities). So that at $g = 2$, the overall computational efforts will increase by a factor of $(n_{\mathrm{f}} + 2n_{\mathrm{g}})/n_{\mathrm{f}} = (3 + 2 \times 1)/3 = 5/3$. Taking into account that the norm $\Gamma_4(\Delta t) = \gamma \Delta t^4$ of global errors is proportional now to the 4th power of Δt, it can be obtained that the efficiency ratio is equal to $200/(5/3)^4 \approx 26$ (the efficiency is defined according to Eq. (81), see Section 3.6).

Some words on the comparison of seven- and five-stage gradient algorithms of the fourth order should also be added. As has already been demonstrated in the case of non-gradient second-order schemes (see remarks to Section 3.2), increasing of the number of stages within the same order may pay dividends. In the case of gradient schemes, a typical example follows when comparing five- and seven-stage integrators which are described by Eqs. (29) and (44), respectively (it seems that such fourth-order integrators are the most efficient within their own number of stages, see also Table 2 at the end of this section). Both integrators require the same number $n_{\mathrm{g}} = 1$ of gradient evaluations, but differ in the numbers of force recalculations, which are equal to $n_{\mathrm{f}} = 2$ and 3 for $S = 5$ and $S = 7$. Thus, increasing S from 5 to 7 will lead to the increase in the computational efforts by a factor of $(3 + 2 \times 1)/(2 + 2 \times 1) = 5/4$. On the other hand, the norm of errors decreases by $0.000713/0.000141 \approx 5$ times. Then one obtains that the extended ($S = 7$) force-gradient scheme (44) is more efficient than the shortened five-stage ($S = 5$) integrator, because of $5/(5/4)^4 \approx 2$. The efficiency of fourth-order algorithms can be improved in one order yet by further increasing the number of stages to a value of $S = 11$ (see next subsections).

3.4. *$S = 9$ stages, order $K = 6$ or 4*

3.4.1. *Velocity versions*

Beginning from $P = 5$, the force-gradient factorization, being written in velocity representation, allows us to eliminate the components of truncation uncertainties up to the sixth order inclusively. In view of Eq. (6), such a representation reads

$$\begin{aligned} \mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^7)} = {} & \mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2\theta)\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}(1-2(\lambda+\vartheta))\Delta t+\chi\mathcal{C}\Delta t^3} \\ & \times \mathrm{e}^{\mathcal{A}(1-2\theta)\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3} \end{aligned} \tag{48}$$

and the order conditions take the form

$$\begin{aligned} \alpha &= \lambda\left(\frac{1}{4} - \theta + \theta^2\right) - \frac{1}{24}(1 - 6\vartheta) = 0, \\ \beta &= \chi - \frac{1}{12}\bigl(1 - 24\mu - 6\lambda^2(2\theta - 1) - 6\vartheta + 6\vartheta^2 - 6\lambda(2\theta - 1)(2\vartheta - 1) - 24\xi\bigr) = 0, \\ \gamma_1 &= \frac{1}{5760}\bigl(7 - 30\lambda(2\theta - 1)^2\bigl(1 + 4\theta - 4\theta^2\bigr) - 30\vartheta\bigr) = 0, \\ \gamma_2 &= \frac{1}{480}\bigl(1 - 20\chi + 80\mu - 20\lambda^2\bigl(1 - 8\theta + 18\theta^2 - 12\theta^3\bigr) - 20\vartheta^2 \\ &\quad + 20\lambda(2\theta - 1)(\theta + 2\vartheta - 6\theta\vartheta) + 80\xi - 480\theta\xi + 480\theta^2\xi\bigr) = 0, \\ \gamma_3 &= \frac{1}{720}\bigl(2 - 30\lambda^2(2\theta - 1)^3 - 15\vartheta + 30\vartheta^2 - 15\lambda(2\theta - 1)^2\bigl(1 - 4\vartheta - \theta(4\vartheta - 2)\bigr)\bigr) = 0, \\ \gamma_4 &= \frac{1}{240}\bigl(2 + 40\mu - 30\lambda^3(2\theta - 1)^2 - 40\chi\bigl(1 + \lambda(6\theta - 3) - 3\vartheta\bigr) - 15\vartheta + 35\vartheta^2 - 30\vartheta^3 \\ &\quad - 5\lambda^2(2\theta - 1)\bigl(7 - 18\vartheta + 6\theta(2\vartheta - 1)\bigr) + 5\lambda(2\theta - 1)\bigl(3 - 14\vartheta + 18\vartheta^2 + 2\theta(1 - 6\vartheta + 6\vartheta^2)\bigr) \\ &\quad + 40\xi - 240\theta\xi + 480\theta\vartheta\xi\bigr) = 0. \end{aligned} \tag{49}$$

Variant 1. $n_{\mathrm{g}} = 3$, order 6. The number of unknowns in propagation (48) is the same as the number of order conditions, and the unique real solution to system (49) is

$$\begin{aligned} &\theta = \frac{1}{2} + \frac{\sqrt[3]{675 + 75\sqrt{6}}}{30} + \frac{5}{2\sqrt[3]{675 + 75\sqrt{6}}}, \qquad \vartheta = \frac{\theta}{3}, \\ &\lambda = -\frac{5\theta}{3}(\theta - 1), \qquad \xi = -\frac{5\theta^2}{144} + \frac{\theta}{36} - \frac{1}{288}, \end{aligned} \tag{50}$$

$$\chi = \frac{1}{144} - \frac{\theta}{36}\left(\frac{\theta}{2} + 1\right), \qquad \mu = 0.$$

Solution (50) constitutes a velocity-force-gradient algorithm of the sixth order with four force and three (not four, since $\mu = 0$) force-gradient evaluations per time step ($n_{\rm f} = 4$ and $n_{\rm g} = 3$). The norm

$$\zeta = \sqrt{\sum_{k=1}^{10} \zeta_k^2} \tag{51}$$

of the leading term $\mathcal{O}_7$ (see Eq. (13)) of seventh-order truncation errors $\mathcal{O}(\Delta t^7)$ corresponding to solution (50) is equal to $\zeta \approx 0.00150$ (the components ζ_{1-10} can be calculated with the help of recursive relations presented in Appendix A).

Variant 2. $n_{\rm g} = 3$, order 4. Putting $\chi = 0$ decreases the number of independent variables from 6 to 5, so that the entire set of order conditions (49) with 6 equations cannot be longer satisfied. In such a situation, we should exclude two variables (λ and ξ, say) by solving a system composed of the first two equations $\alpha = 0$ and $\beta = 0$, substitute the solutions in the remaining four equalities of Eq. (49) and try to find a global minimum of the norm γ with respect to other of three variables (θ, ϑ, and μ). The result is $\gamma_{\min} \approx 0.00000605$ and

$$\begin{aligned} \theta &= 0.1705755127786631\text{E}{+}00, \\ \vartheta &= 0.4775180236616381\text{E}{-}01, \\ \lambda &= 0.2739456420927671\text{E}{+}00, \\ \xi &= 0.2464531166166595\text{E}{-}02, \\ \mu &= -0.6175944713542174\text{E}{-}03. \end{aligned} \tag{52}$$

Note that scheme (52) corresponds to a fourth order algorithm, because now $\gamma \neq 0$.

Variant 3. $n_{\rm g} = 2$, order 4. At $\xi = 0$, the minimization of γ (again with respect to three free parameters) yields $\gamma_{\min} \approx 0.00000368$ and

$$\begin{aligned} \theta &= 0.1921125277429464\text{E}{+}00, \\ \vartheta &= 0.5851872613455621\text{E}{-}01, \\ \lambda &= 0.2852162240687091\text{E}{+}00, \\ \chi &= 0.2427475259663050\text{E}{-}02, \\ \mu &= 0.4339598806816256\text{E}{-}03. \end{aligned} \tag{53}$$

Variant 4. $n_{\rm g} = 2$, order 4. For $\mu = 0$ and $\chi = 0$, the minimization (with respect to two parameters) results in the solution

$$\begin{aligned} \theta &= 0.2189286596427438\text{E}{+}00, \\ \vartheta &= 0.6840805970727767\text{E}{-}01, \\ \lambda &= 0.3109406355938166\text{E}{+}00, \\ \xi &= 0.1602503681334363\text{E}{-}02 \end{aligned} \tag{54}$$

for which $\gamma_{\min} \approx 0.00000649$.

Variant 5. $n_{\rm g} = 1$, order 4. At $\mu = 0$ and $\xi = 0$, the minimization (with respect to two parameters) gives the result $\gamma_{\min} \approx 0.0000634$ and

$$\begin{aligned}
\theta &= 0.1987553828429444\text{E}{+}00,\\
\vartheta &= 0.7332763128496152\text{E}{-}01,\\
\lambda &= 0.2571370908951839\text{E}{+}00,\\
\chi &= 0.3175049859241442\text{E}{-}02.
\end{aligned} \tag{55}$$

Variant 6. $n_{\mathrm{g}} = 1$, order 4. Choosing $\xi = 0$ and $\chi = 0$ and minimizing γ (with respect to two free parameters) leads to $\gamma_{\min} \approx 0.000294$ and

$$\begin{aligned}
\theta &= 0.1658018263462486\text{E}{+}00,\\
\vartheta &= 0.1799550304949169\text{E}{-}02,\\
\lambda &= 0.3690335968449654\text{E}{+}00,\\
\mu &= 0.2530872904428947\text{E}{-}02.
\end{aligned} \tag{56}$$

Variant 7. $n_{\mathrm{g}} = 0$, order 4. Finally, setting all the time coefficients at operator $\mathcal{C}$ to be equal to zero, namely, $\mu = 0$, $\xi = 0$ and $\chi = 0$, and minimizing γ (with respect to one parameter), one finds the following non-gradient scheme

$$\begin{aligned}
\theta &= 0.5209433391039899\text{E}{+}00,\\
\vartheta &= 0.1644986515575760\text{E}{+}00,\\
\lambda &= 1.2356926511389169\text{E}{+}00
\end{aligned} \tag{57}$$

with $\gamma_{\min} \approx 0.000654$. This scheme can be considered as an extended ($S = 9$) counterpart of the non-gradient Forest–Ruth ($S = 7$) integrators.

3.4.2. Position versions

The nine-stage factorization in position form can be reproduced from Eq. (6) at $P = 5$ as

$$\begin{aligned}
\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^5)} = {}& \mathrm{e}^{\mathcal{A}\boldsymbol{\rho}\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}(1-2\lambda)\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2(\theta+\rho))\Delta t}\\
&\times \mathrm{e}^{\mathcal{B}(1-2\lambda)\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\boldsymbol{\rho}\Delta t}
\end{aligned} \tag{58}$$

with the order conditions

$$\begin{aligned}
\alpha &= \frac{1}{12} - \frac{\rho}{2}(1-\rho) - \theta\left(\frac{1}{2} - \lambda - \rho + 2\rho\lambda\right) + \theta^2\left(\frac{1}{2} - \lambda\right) = 0,\\
\beta &= \frac{1}{24} - \frac{\rho}{4} - \theta\left(\frac{1}{4} - \lambda + \lambda^2\right) + 2(\xi + \chi) = 0.
\end{aligned}$$

These conditions can be solved with respect to two appropriate time coefficients (λ and χ, say) and the solutions found can be inserted in the corresponding expression for the γ-function. As a result, this function will depend on the remaining three coefficients (ρ, θ, and ξ). Note that the position nine-stage algorithm of the sixth order does not exist, since we cannot reduce the four components γ_{1-4} of γ having three parameters. Thus, scheme (58) will produce a series of extended algorithms of order four, where up to three time coefficients can be free. The actual values for these coefficients will be found using our general procedure on the minimization of the norm γ. In order to simplify notations, the concrete number n of free parameters will be pointed out now additionally. The final results for possible position variants of the nine-stage propagator (58) are presented below.

Variant 1. $n_{\mathrm{g}} = 4$, *order 4*, $n = 3$.

$$\begin{aligned}
&\rho = 0.9325912861071900\mathrm{E}{-01},\\
&\theta = 0.2791634819768266\mathrm{E}{+00},\\
&\xi = 0.5562281089130940\mathrm{E}{-03},\\
&\lambda = 0.2247800288685984\mathrm{E}{+00},\\
&\chi = 0.8405927247441154\mathrm{E}{-03},\\
&\gamma_{\min} \approx 0.00000312.
\end{aligned} \tag{59}$$

Variant 2. $n_{\mathrm{g}} = 2$, *order 4*, $n = 2$.

$$\begin{aligned}
&\xi = 0,\\
&\rho = 0.4418173708072988\mathrm{E}{-01},\\
&\theta = 0.2658951191619568\mathrm{E}{+00},\\
&\lambda = 0.1376315482160252\mathrm{E}{+00},\\
&\chi = 0.2146846818235837\mathrm{E}{-02},\\
&\gamma_{\min} \approx 0.0000323.
\end{aligned} \tag{60}$$

Variant 3. $n_{\mathrm{g}} = 2$, *order 4*, $n = 2$.

$$\begin{aligned}
&\chi = 0,\\
&\rho = 0.1181905290564645\mathrm{E}{+00},\\
&\theta = 0.2960939018274884\mathrm{E}{+00},\\
&\lambda = 0.2744544305971990\mathrm{E}{+00},\\
&\xi = 0.1471750203280252\mathrm{E}{-02},\\
&\gamma_{\min} \approx 0.0000464.
\end{aligned} \tag{61}$$

Variant 4. $n_{\mathrm{g}} = 0$, *order 4*, $n = 1$.

$$\begin{aligned}
&\xi = 0,\\
&\chi = 0,\\
&\rho = 0.1786178958448091\mathrm{E}{+00},\\
&\theta = -0.6626458266981843\mathrm{E}{-01},\\
&\lambda = 0.7123418310626056\mathrm{E}{+00},\\
&\gamma_{\min} \approx 0.000610.
\end{aligned} \tag{62}$$

The latter variant corresponds to a non-gradient integration which like scheme (57) can be classified as an extended Forest–Ruth-like integrator.

3.4.3. Remarks

First of all it is necessary to point out that the introduction of nine-stage propagation (48) has allowed us to derive a sixth-order ($K = 6$) integrator (Eq. (50)) using a considerably smaller number of stages ($S = 9$, instead of 15, see Table 1) than that inherent in existing (non-gradient) schemes [4,9] of the same order. A sixth-order scheme like (48) has been analyzed earlier by Koseleff [6], but again as in the case of five-stage propagations (see at the top of Section 3.2.3), we have realized that his relations for time coefficients do not lead to the correct result (50) (they do lead only to similar expressions which however are not exactly the same as ours).

Secondly, the advantage of nine-stage algorithms of order four ($K = 4$) with respect to the fourth-order seven-stage integrators is further significant improvement of the precision of the integration in a little additional computational cost. For example, the nine-stage scheme (53) (with $n_{\mathrm{f}} = 4$, $n_{\mathrm{g}} = 2$, and $\gamma \approx 0.00000368$) reduces the numerical uncertainties by a factor of 40 with respect to the prominent seven-stage integrator (44) (for which $n_{\mathrm{f}} = 3$, $n_{\mathrm{g}} = 1$, and $\gamma \approx 0.000141$). A similar situation is observed within the non-gradient ($n_{\mathrm{g}} = 0$) class, where the precision of the fourth-order integration can be improved in $0.0283/0.000610 \approx 50$ times by increasing n_{f} only by unity from 3 to 4 (please compare the values of γ for the seven-stage Forest–Ruth algorithm (45) and its extended nine-stage counterpart (62)). This leads to the increase in the resulting efficiency by a factor of $50/(4/3)^4 \approx 15$.

3.5. $S = 11$ *stages, order* $K = 6$ *or* 4

3.5.1. *Velocity versions*

Taking $P = 6$ and using the self-adjoint boundary conditions in velocity form, one comes to the following eleven-stage propagation

$$\begin{aligned} \mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^7)} = {} & \mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\boldsymbol{\rho}\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}(1-2(\lambda+\vartheta))\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2(\theta+\rho))\Delta t} \\ & \times \mathrm{e}^{\mathcal{B}(1-2(\lambda+\vartheta))\frac{\Delta t}{2}+\chi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\boldsymbol{\rho}\Delta t}\mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3}. \end{aligned} \tag{63}$$

Note that beginning from $P = 6$, the expressions for order conditions and error functions become to be cumbersome enough. In view of this we do not present them explicitly and restrict ourselves to final results only with brief comments to each variant introduced.

Variant 1. $n_{\mathrm{g}} = 5$, *order 6,* $n = 1$. The total number of independent time coefficients (ϑ, ν, ρ, λ, ξ, θ, and χ) for Eq. (63) is equal to 7, i.e. it exceeds by unity the number of the sixth-order conditions $\alpha = 0$, $\beta = 0$, and $\gamma_{1-4} = 0$. This will lead to a whole continuity of solutions with one free parameter. It is quite natural to choose an optimal value for such a parameter by minimizing the norm ζ (see Eq. (51)) of the leading term $\mathcal{O}_7$ (see Eq. (13)) of seventh-order truncation uncertainties $\mathcal{O}(\Delta t^7)$ appearing in propagation (63). The result is

$$\begin{aligned} &\rho = 0.5309910490348568\text{E}{+}00, \\ &\theta = -0.2573883543804353\text{E}{+}00, \\ &\vartheta = 0.8281492492827128\text{E}{-}01, \\ &\lambda = 0.8354543940755644\text{E}{-}02, \\ &\xi = -0.2401600937577623\text{E}{-}03, \\ &\chi = 0.4267631995107088\text{E}{-}02, \\ &\mu = -0.1633190022736910\text{E}{-}03, \\ &\zeta_{\min} \approx 0.00000249. \end{aligned} \tag{64}$$

Variant 2. $n_{\mathrm{g}} = 4$, *order 6,* $n = 0$. Other variants are obtained by putting one or all time coefficients at operator $\mathcal{C}$ to zero. Thus, for $\mu = 0$ (then the number of unknowns coincides with the number of order conditions) one finds a discrete set of up to fourteen real solutions and the optimal among them which leads to the smallest value 0.00000366 of ζ can be presented analytically,

$$
\begin{aligned}
&\rho = \frac{1}{2}\left(1+\frac{1}{\sqrt{5}}\right), \qquad \theta = -\frac{1}{\sqrt{5}},\\
&\vartheta = \frac{1}{12}, \qquad \lambda = \frac{5}{12} - \frac{1}{24}\left(\sqrt{50+22\sqrt{5}}\right)\\
&\xi = \frac{15+5\sqrt{5}}{1152} - \sqrt{50+22\sqrt{5}}\left(\frac{1}{2880}+\frac{\sqrt{5}}{1152}\right),\\
&\chi = -\frac{11+5\sqrt{5}}{1152} + \sqrt{50+22\sqrt{5}}\left(\frac{1}{2880}+\frac{\sqrt{5}}{1152}\right).
\end{aligned} \tag{65}
$$

Variant 3. $n_\mathrm{g} = 3$, order 6, $n = 0$. At $\xi = 0$, there are eight real solutions and the optimal, which minimizes ζ to the value 0.0000147, is

$$
\begin{aligned}
&\rho = 0.2742082240034209\mathrm{E}{+}00,\\
&\theta = 0.4812780570021632\mathrm{E}{+}00,\\
&\vartheta = 0.8350330494925359\mathrm{E}{-}01,\\
&\lambda = 0.4474919773539384\mathrm{E}{+}00,\\
&\chi = 0.3435650653755542\mathrm{E}{-}02,\\
&\mu = -0.2544189176362832\mathrm{E}{-}03.
\end{aligned} \tag{66}
$$

Variant 4. $n_\mathrm{g} = 3$, order 6, $n = 0$. For $\chi = 0$, there exist six real solutions and the optimal one is

$$
\begin{aligned}
&\rho = 0.1667381233476491\mathrm{E}{+}00,\\
&\theta = 0.3800389344302596\mathrm{E}{+}00,\\
&\vartheta = 0.4800136993352096\mathrm{E}{-}01,\\
&\lambda = 0.2633957069935350\mathrm{E}{+}00,\\
&\xi = 0.4668083730519805\mathrm{E}{-}02,\\
&\mu = -0.1709693171449844\mathrm{E}{-}02,\\
&\zeta_\mathrm{min} \approx 0.0000264.
\end{aligned} \tag{67}
$$

Variant 5. $n_\mathrm{g} = 2$, order 4, $n = 3$. Further increasing the number of zero-valued coefficients at force-gradient operator $\mathcal{C}$, the number of remaining unknowns reduces to 5 or 4 (in dependence on the variant considered), i.e. it becomes smaller than six. As a result, we come to extended fourth-order schemes by satisfying only two order conditions ($\alpha = 0$ and $\beta = 0$) and minimizing the γ-function with respect to $n = 3$ or 2 free parameters. So that when $\mu = 0$ and $\chi = 0$, one obtains

$$
\begin{aligned}
&\rho = 0.2029270564692829\mathrm{E}{+}00,\\
&\theta = 0.1926052063353027\mathrm{E}{+}00,\\
&\vartheta = 0.6668876199434440\mathrm{E}{-}01,\\
&\lambda = 0.2620356629687677\mathrm{E}{+}00,\\
&\xi = 0.1042387551227681\mathrm{E}{-}02,\\
&\gamma_\mathrm{min} \approx 0.00000320.
\end{aligned} \tag{68}
$$

Variant 6. $n_\mathrm{g} = 2$, order 4, $n = 3$. For $\mu = 0$ and $\xi = 0$, one finds the following result

$$\begin{aligned}
&\rho = -0.3936043328394478\text{E}{-01},\\
&\theta = 0.3268925828232685\text{E}{+00},\\
&\vartheta = 0.1540533458110347\text{E}{+00},\\
&\lambda = -0.5071304262389421\text{E}{-01},\\
&\chi = 0.2527818460124813\text{E}{-02},\\
&\gamma_{\min} \approx 0.0000132.
\end{aligned} \tag{69}$$

Variant 7. $n_{\mathrm{g}} = 1$, order 4, $n = 3$. At $\xi = 0$ and $\chi = 0$, the optimal solution is

$$\begin{aligned}
&\rho = 0.2797644436188271\text{E}{+00},\\
&\theta = -0.1180329820696323\text{E}{-02},\\
&\vartheta = 0.8010998355755116\text{E}{-01},\\
&\lambda = -2.0220148671481104\text{E}{+00},\\
&\mu = 0.3098750751031143\text{E}{-03},\\
&\gamma_{\min} \approx 0.0000165.
\end{aligned} \tag{70}$$

Variant 8. $n_{\mathrm{g}} = 0$, order 4, $n = 2$. Putting, finally, $\mu = 0$, $\xi = 0$ and $\chi = 0$ will lead to the non-gradient algorithm

$$\begin{aligned}
&\rho = 0.2539785108410595\text{E}{+00},\\
&\theta = -0.3230286765269967\text{E}{-01},\\
&\vartheta = 0.8398315262876693\text{E}{-01},\\
&\lambda = 0.6822365335719091\text{E}{+00},\\
&\gamma_{\min} \approx 0.0000270.
\end{aligned} \tag{71}$$

3.5.2. *Position versions*

The position counterpart of eleven-stage propagation (63) has the form

$$\begin{aligned}
\mathrm{e}^{(\mathcal{A}+\mathcal{B})\Delta t+\mathcal{O}(\Delta t^7)} = {} & \mathrm{e}^{\mathcal{A}\boldsymbol{\rho}\Delta t}\mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}(1-2(\theta+\rho))\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}(1-2(\lambda+\vartheta))\Delta t+\chi\mathcal{C}\Delta t^3}\\
& \times \mathrm{e}^{\mathcal{A}(1-2(\theta+\rho))\frac{\Delta t}{2}}\mathrm{e}^{\mathcal{B}\lambda\Delta t+\xi\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\theta\Delta t}\mathrm{e}^{\mathcal{B}\vartheta\Delta t+\mu\mathcal{C}\Delta t^3}\mathrm{e}^{\mathcal{A}\rho\Delta t}
\end{aligned} \tag{72}$$

Variant 1. $n_{\mathrm{g}} = 5$, order 6, $n = 1$. The total number of time coefficients entering in Eq. (72) also exceeds by one (as in the case of velocity propagation (63)) the number of sixth-order conditions. So that the most general optimal solution can be obtained here performing the minimization of ζ with respect to an arbitrary coefficient (free parameter), and the result is

$$\begin{aligned}
&\rho = 0.1098059301577147\text{E}{+00},\\
&\theta = 0.4828099940251012\text{E}{+00},\\
&\vartheta = 0.2693816517677854\text{E}{+00},\\
&\lambda = 0.7611936345860829\text{E}{-01},\\
&\xi = -0.1803378129376054\text{E}{-02},\\
&\chi = 0.1083650107661986\text{E}{-01},\\
&\mu = 0.1011249349033012\text{E}{-02},\\
&\zeta_{\min} \approx 0.00000299.
\end{aligned} \tag{73}$$

Variant 2. $n_{\mathrm{g}} = 4$, order 6, $n = 0$. The next three variants are constructed by putting one (of three) force-gradient coefficient to zero (reducing in such a way the number of unknowns to the number of order conditions). So that at $\chi = 0$, we have found a discrete set with four real solutions and the optimal one within this set which leads to the lowest value 0.0000139 of ζ is

$$\begin{aligned}
&\rho = 0.1558931576791768\mathrm{E}{+}00,\\
&\theta = -0.1270876254528190\mathrm{E}{+}00,\\
&\vartheta = 0.2446016254916385\mathrm{E}{+}00,\\
&\lambda = 0.5721675541779425\mathrm{E}{-}01,\\
&\xi = -0.6695276810842814\mathrm{E}{-}03,\\
&\mu = 0.4431288445550721\mathrm{E}{-}02.
\end{aligned}\tag{74}$$

Variant 3. $n_{\mathrm{g}} = 3$, order 6, $n = 0$. For $\mu = 0$ there are four real solutions and the optimal one looks as

$$\begin{aligned}
&\rho = 0.1094983141824115\mathrm{E}{+}00,\\
&\theta = 0.4393783304256709\mathrm{E}{+}00,\\
&\vartheta = 0.2687036338598425\mathrm{E}{+}00,\\
&\lambda = 0.2895143963458680\mathrm{E}{+}00,\\
&\xi = 0.5612760102273875\mathrm{E}{-}01,\\
&\chi = -0.1032070838153398\mathrm{E}{+}00,\\
&\zeta_{\min} \approx 0.000146.
\end{aligned}\tag{75}$$

Variant 4. $n_{\mathrm{g}} = 3$, order 6, $n = 0$. At $\xi = 0$ there exist two sets of real solutions and the best is

$$\begin{aligned}
&\rho = 0.1097059723948682\mathrm{E}{+}00,\\
&\theta = 0.4140632267310831\mathrm{E}{+}00,\\
&\vartheta = 0.2693315848935301\mathrm{E}{+}00,\\
&\lambda = 1.1319803486515564\mathrm{E}{+}00,\\
&\chi = -0.1324638643416052\mathrm{E}{-}01,\\
&\mu = 0.8642161339706166\mathrm{E}{-}03,\\
&\zeta_{\min} \approx 0.00000607.
\end{aligned}\tag{76}$$

Variant 5. $n_{\mathrm{g}} = 2$, order 4, $n = 3$. In the remaining four variants, when more than one gradient coefficient is put to zero, the order decreases from 6 to 4, so that the minimization should be carried out for the γ-function with respect to $n = 3$ or 2 parameters. Then at $\xi = 0$ and $\chi = 0$ we find

$$\begin{aligned}
&\rho = 0.1159989388152167\mathrm{E}{+}00,\\
&\theta = 0.3885522942527583\mathrm{E}{+}00,\\
&\vartheta = 0.2826569520375214\mathrm{E}{+}00,\\
&\lambda = -0.6289171779553212\mathrm{E}{+}00,\\
&\mu = 0.1214053476775188\mathrm{E}{-}02,\\
&\gamma_{\min} \approx 0.0000117.
\end{aligned}\tag{77}$$

Variant 6. $n_{\mathrm{g}} = 2$, order 4, $n = 3$. For $\mu = 0$ and $\chi = 0$ the global minimum is $\gamma_{\min} \approx 0.00000127$ and achieved at

$$\begin{aligned}
\rho &= 0.6419108866816235\text{E}{-}01,\\
\theta &= 0.1919807940455741\text{E}{+}00,\\
\vartheta &= 0.1518179640276466\text{E}{+}00,\\
\lambda &= 0.2158369476787619\text{E}{+}00,\\
\xi &= 0.9628905212024874\text{E}{-}03.
\end{aligned} \tag{78}$$

Variant 7. $n_{\mathrm{g}} = 1$, order 4, $n = 3$. At $\mu = 0$ and $\xi = 0$ one finds the global minimum $\gamma_{\min} \approx 0.0000121$ at

$$\begin{aligned}
\rho &= 0.1255768596433302\text{E}{+}00,\\
\theta &= -0.2407093745014925\text{E}{-}02,\\
\vartheta &= -0.8938074259467744\text{E}{+}00,\\
\lambda &= 1.1758501877269955\text{E}{+}00,\\
\chi &= 0.2952744354631969\text{E}{-}02.
\end{aligned} \tag{79}$$

Variant 8. $n_{\mathrm{g}} = 0$, order 4, $n = 2$. Finally, letting $\mu = 0$, $\xi = 0$ and $\chi = 0$ yields the non-gradient scheme

$$\begin{aligned}
\rho &= 0.2750081212332419\text{E}{+}00,\\
\theta &= -0.1347950099106792\text{E}{+}00,\\
\vartheta &= -0.8442961950707149\text{E}{-}01,\\
\lambda &= 0.3549000571574260\text{E}{+}00,
\end{aligned} \tag{80}$$

with $\gamma_{\min} \approx 0.0000518$.

3.5.3. Remarks

The introduction of eleven-stage propagations has confirmed our previous conclusion that the best algorithm of a given order should be found among extended schemes. Indeed, as was shown in Section 3.4, the norm of the leading term of truncation errors, corresponding to the sixth-order decomposition algorithm (see Eq. (50)) with the minimal possible number ($S = 9$) of stages as well as with $n_{\mathrm{f}} = 4$ and $n_{\mathrm{g}} = 3$, is equal to $\zeta \approx 0.00150$. On the other hand, using, for example, the eleven-stage algorithm (76) (when $n_{\mathrm{f}} = 5$ and $n_{\mathrm{g}} = 3$) reduces that norm to the value $\zeta \approx 0.00000607$, i.e. by more than 200 times. Note that such a reduction is achieved by applying only one additional force evaluation which leads to an increased computational work by a factor of $(5 + 2 \times 3)/(4 + 2 \times 3) = 11/10$, so that the efficiency increases by $200/(11/10)^6 \approx 110$ times.

3.6. A complete classification and general remarks

All the algorithms obtained by us in Section 3 are collected below in Table 2. Here, Err_3, Err_5, and Err_7 designate the norms $(\alpha^2 + \beta^2)^{1/2}$, γ, and ζ of third-, fifth-, and seventh-order truncation errors, respectively (see Eqs. (11)–(13), (27), and (51)). The total numbers of force and force-gradient evaluations per time step, namely n_{f} and n_{g}, are shown in the table as well. The algorithms have been presented in abbreviated forms, so that the letters A and B correspond to the exponential operators $\mathrm{e}^{\mathcal{A}a_p\Delta t}$ and $\mathrm{e}^{\mathcal{B}b_p\Delta t}$, respectively, whereas the letter C denotes $\mathrm{e}^{\mathcal{B}b_p\Delta t + \mathcal{C}c_p\Delta t^3}$. We refer also to the determining equations, where the explicit forms of the decomposition algorithms and the numerical values for their time coefficients $\{a_p, b_p, c_p\}$ can be found. Each group of algorithms corresponding to the same number ($S = 3$, 5, 7, 9, or 11) of stages is separated by horizontal lines. Within the same group, the algorithms are allocated in the order of increasing computational efforts. At the same values of n_{f} and n_{g}, the velocity versions (where the letters C or B, but not A, appear first) are written first as well. When all these characteristics coincide between themselves, the preference in the ordering of location is given to more efficient schemes.

Table 2
The complete family of self-adjoint decomposition algorithms with up to 11 stages

Algorithm	Order	n_f	n_g	Err$_3$	Err$_5$	Err$_7$	Efficiency	Equations	Remarks	No
BAB	2	1	0	0.0932	0.00913	0.00132	10.7**	(23), $\xi=0$	[26,27]	1
ABA	2	1	0	0.0932	0.00911	0.00134	10.7**	(24), $\xi=0$	[27]	2
CAC	2	1	1	0.0833	0.0134	0.00224	1.3	(23), $\xi=\frac{-1}{48}$	New	3
ACA	2	1	1	0.0417	0.00648	0.000725	2.7*	(24), $\xi=\frac{1}{12}$	New	4
BABAB	2	2	0	0.00855	0.00103	0.000110	29.2***	(25), (31)	**New**	5
ABABA	2	2	0	0.00855	0.00106	0.000106	29.2***	(32), (31)	**New**	6
CABAC	4	2	1	0	0.00334	0.000272	1.2*	(25), (30)	New	7
BACAB	4	2	1	0	0.000713	0.0000630	5.5**	(25), (29)	[8,12]	8
CACAC	4	2	2	0	0.000595	0.0000483	1.3*	(25), (28)	New	9
ACACA	4	2	2	0	0.000715	0.0000559	1.1*	(32), (33)	[8,12]	10
BABABAB	4	3	0	0	0.0383	0.0126	0.32	(34), (39)	[2]	11
ABABABA	4	3	0	0	0.0283	0.00630	0.44	(40), (45)	[2]	12
CABABAC	4	3	1	0	0.000855(a)	0.0000224(a)	1.4*/1.9*	(34), (47/38)	[35]/New	13
ABACABA	4	3	1	0	0.000141(a)	0.0000104(a)	2.2*/11.3**	(40), (46/44)	[12]/New	14
BACACAB	4	3	2	0	0.0000443	0.00000371	9.4**	(34), (37)	New	15
ACABACA	4	3	2	0	0.0000823	0.00000912	5.1**	(40), (43)	New	16
CACACAC	4	3	3	0	0.0000167	0.00000574	9.1**	(34), (36)	New	17
ACACACA	4	3	3	0	0.0000123	0.00000491	12.4**	(40), (42)	New	18
BABABABAB	4	4	0	0	0.000654	0.0000645	6.0**	(48), (57)	New	19
ABABABABA	4	4	0	0	0.000610	0.0000456	6.4**	(58), (62)	New	20
BABACABAB	4	4	1	0	0.0000634	0.00000511	12.2**	(48), (55)	New	21
CABABABAC	4	4	1	0	0.000294	0.0000129	2.6*	(48), (56)	New	22
CABACABAC	4	4	2	0	0.00000368	0.00000679	66.3***	(48), (53)	**New**	23
BACABACAB	4	4	2	0	0.00000649	0.00000123	37.6***	(48), (54)	New	24
ABACACABA	4	4	2	0	0.0000323	0.00000171	7.6**	(58), (60)	New	25
ACABABACA	4	4	2	0	0.0000464	0.00000541	5.3**	(58), (61)	New	26
CACABACAC	4	4	3	0	0.00000605	0.0000104	16.5**	(48), (52)	New	27
≡ CACACACAC BACACACAB	6	4	3	0	0	0.00150	0.00067	(48), (50)	New	28
ACACACACA	4	4	4	0	0.00000312	0.00000227	15.5**	(58), (59)	New	29
BABABABABAB	4	5	0	0	0.0000270	0.00000272	59.3***	(63), (71)	**New**	30
ABABABABABA	4	5	0	0	0.0000518	0.0000173	30.9***	(72), (80)	New	31
CABABABABAC	4	5	1	0	0.0000165	0.00000631	25.2***	(63), (70)	New	32
ABABACABABA	4	5	1	0	0.0000121	0.00000537	34.4***	(72), (79)	New	33
BACABABACAB	4	5	2	0	0.00000320	0.00000156	47.6***	(63), (68)	New	34
BABACACABAB	4	5	2	0	0.0000132	0.00000351	11.5**	(63), (69)	New	35
ABACABACABA	4	5	2	0	0.00000127	0.00000224	120.0***	(72), (78)	**New**	36
ACABABABACA	4	5	2	0	0.0000117	0.00000483	13.0**	(72), (77)	New	37
CABACACABAC	6	5	3	0	0	0.0000147	0.0384**	(63), (66)	New	38
CACABABACAC	6	5	3	0	0	0.0000264	0.0214*	(63), (67)	New	39
ACABACABACA	6	5	3	0	0	0.00000607	0.0930***	(72), (76)	New	40
ABACACACABA	6	5	3	0	0	0.000146	0.0039	(72), (75)	New	41
BACACACACAB	6	5	4	0	0	0.00000366	0.0566**	(63), (65)	New	42
ACACABACACA	6	5	4	0	0	0.0000139	0.0149*	(72), (74)	New	43
CACACACACAC	6	5	5	0	0	0.00000249	0.0353**	(63), (64)	New	44
ACACACACACA	6	5	5	0	0	0.00000299	0.0294*	(72), (73)	New	45

(a) Values corresponding to new algorithms.
*** Particularly outstanding algorithms.

The efficiency of the algorithms has been measured using the formula

$$\mathrm{Eff}^{(K)} = \frac{1}{(n_{\mathrm{f}} + g n_{\mathrm{g}})^K \mathrm{Err}_{K+1}} \tag{81}$$

at $g = 2$, where Err_{K+1} is the leading error coefficient (i.e. Err_3, Err_5, or Err_7, depending on the order $K = 2, 4$, or 6, respectively). The weight g reflects the fact that one gradient evaluation requires more (a factor g) processor (CPU) time than that necessary for one force calculation (see Eq. (8)). In molecular dynamics simulations of classical systems, typical values for this weight are of order 2 to 3 depending on the structure of interaction potentials (see Ref. [50] and remarks in Section 3.3.3 of this paper). For some cases, such as particles moving in gravitational or coulomb fields, the force-gradient evaluations practically do not lead to an increase in the computational work due to a specific structure of the potential, i.e. $g \ll 1$. So that the actual efficiency of force-gradient schemes in these cases will be higher than that pointed out in Table 2. For quantum systems, the quantities n_{f} and n_{g} should be referred to the numbers of potential and its gradient evaluations, respectively. In addition, we should take into account that here any appearance of the kinetic subpropagator $\mathrm{e}^{\mathcal{A} a_p \Delta t}$ requires performing two spatial Fourier transforms. This operation may take a processor time which is comparable with the time needed for the evaluation of the potential subpropagator $\mathrm{e}^{\mathcal{B} b_p \Delta t + \mathcal{C} c_p \Delta t^3}$ (contrary to classical systems, where the kinetic part of evolution reduces to simple shifts of position and time coordinates, see first line of Eq. (9)). Thus, in the case of quantum mechanics simulations, preference should be given to velocity versions of the decomposition algorithms, where the number of Fourier transforms is smaller (namely $S - 1$, instead of $S + 1$ as for position counterparts).

It can be seen from Table 2, that for $S \leqslant 11$ there are 46 possible self-adjoint decomposition variants which lead to 45 (10 non-gradient plus 35 force-gradient) different algorithms (the scheme CACACACAC actually reduces to BACACACAB, because of $\mu = 0$, see Eqs. (48) and (50)). Among them, 4 non- plus 4 force-gradient integrators were known earlier (see the references under remarks in the table). All other 37 schemes are new. The algorithms have been categorized according to efficiency using a three-star classification (the worst with no star and the best with 3 stars). Some of the integrators (marked by the bold font and three bold stars) are particularly outstanding because they result in the most efficient computations within a given order. For instance, the five-stage non-gradient algorithms under Nos 5 and 6 (see Table 2) are the best among second-order decomposition schemes. The well-known Verlet integrators (Nos 1 and 2), which are used in the majority of molecular dynamics simulations, are clearly worse (see also remarks on this at the end of Section 3.2.3). An interesting situation happens with fourth-order decomposition schemes. As we can see, the previously known seven-stage integrators by Forest and Ruth (Nos 11 and 12) result in the worst (with no star) integration. On the other hand, the best fourth-order integrator, belonging to the gradient class with eleven stages and marked under No 36, allows us to reduce the numerical uncertainties by up to $120/0.32 \approx 375$ times at the same overall computational costs. The algorithm No 23 exhibits a similar equivalence in the reduction of numerical errors and appears to be the best within a nine-stage group. The seven-stage scheme by Chin (Eq. (46)) and even its improved version (Eq. (44)) under No 14 are both evidently worse. The eleven-stage scheme No 30 leads to nearly the same efficiency as that of the gradient algorithm No 23, and should be considered as the best within the non-gradient class. Finally, for order six, the best integration can be carried out with the help of the eleven-stage algorithm under No 40. Its efficiency is higher by a factor of $0.0930/0.00067 \approx 140$ with respect to the worst sixth-order algorithm under No 28. We see, therefore, that the best algorithms at each order considered belong to extended schemes, where the number of stages exceeds the necessary minimum. Moreover, the efficiency of these new algorithms is higher with respect to previously known schemes by factors from 3 up to 375 for orders 2 to 6.

The formula (81) has been obtained by us in view of the following arguments. For algorithms of order K, the local errors corresponding to one-step propagations are equal to $o_{K+1} = \mathrm{Err}_{K+1} \Delta t^{K+1}$. When performing a whole propagation over a given time interval t, such errors will heap together step by step during the integration process and lead to the global errors $\mathrm{O}_K = o_{K+1} l = t\, \mathrm{Err}_{K+1} \Delta t^K$, where $l = t/\Delta t$ denotes the total number of steps. On the other hand, the CPU time required for carrying out the above whole propagation is equal to $\mathcal{T}_{\mathrm{CPU}} = c\, l(n_{\mathrm{f}} + g n_{\mathrm{g}}) = c\, t(n_{\mathrm{f}} + g n_{\mathrm{g}})/\Delta t$, where c denotes the proportionality constant depending on the speed of the computer used. It is quite natural to introduce the efficiency of algorithms as inverse to their numerical uncertainties O_K, i.e. as $\mathrm{Eff}^{(K)} = 1/\mathrm{O}_K$, provided the quantities c, t, and $\mathcal{T}_{\mathrm{CPU}}$ are fixed. Then the time step can be cast in terms of these quantities as $\Delta t = ct(n_{\mathrm{f}} + gn_{\mathrm{g}})/\mathcal{T}_{\mathrm{CPU}}$ and substituted into the expression for O_K. As a result, one finds $\mathrm{Eff}^{(K)} = (1/t)(\mathcal{T}_{\mathrm{CPU}}/(ct))^K/[(n_{\mathrm{f}} + gn_{\mathrm{g}})^K \mathrm{Err}_{K+1}]$. For fixed values of c, t, and $\mathcal{T}_{\mathrm{CPU}}$, the quantity $(1/t)(\mathcal{T}_{\mathrm{CPU}}/(ct))^K$ remains constant. Thus it can be omitted in the definition of $\mathrm{Eff}^{(K)}$ and we come to Eq. (81).

It should be pointed out also that formula (81) can be used for comparison of the efficiencies of algorithms of the same orders K only. Indeed, in view of the structure of this formula, it follows that the relative efficiency of two algorithms is determined at a given K as the ratio of precisions which can be achieved in the integration by the same overall computational costs. The precision is introduced as $1/\mathrm{Err}$, where Err denotes the numerical uncertainties. The leading term of these uncertainties behaves like $\mathrm{Err}^{(K)} = C_K \mathrm{Err}_{K+1} \Delta t^K$, where C_K is the parameter which depends on the order K, the system under consideration and the measured quantity. For the same value of K, such a parameter is cancelled when determining the relative efficiency of one integrator with respect to another (for this reason and to simplify notations, this parameter was not considered when deriving Eq. (81)). When analyzing algorithms of different orders, K and $K + M$ say, one obtains that the relative efficiency is $\frac{C_{K+M}}{C_K} \mathrm{Eff}^{(K)}/\mathrm{Eff}^{(K+M)}$. It depends not only on the norms Err_{K+1} and Err_{K+M+1} of numerical uncertainties (see Eq. (81)), but also on the factor C_{K+M}/C_K. The last factor is not universal and cannot be evaluated theoretically in advance. It should be investigated directly in simulations for each concrete system, where the final decision on the preference of higher-order algorithms over lower-order ones or vice versa can be given.

4. Schemes with higher stage numbers and orders

Increasing the number of stages above 11, the number of decomposition variants increases too rapidly. In view of this, the performance of a complete classification for $S > 11$ becomes difficult. Moreover, sufficiently high stage numbers will result in very cumbersome systems of non-linear order conditions. So that we may come to an unresolvable technical problem when trying to handle such conditions. In this respect it should be noted that all the results presented in the preceding Section 3 have been obtained analytically or in quadratures with the help of computer algebra packages, Maple 6 and Mathematica 4 (the two packages have been applied to ensure the correctness of the results and to minimize transcription mistakes). By quadratures we mean that the problem has been reduced to finding of real zeros for a one-dimensional polynomial of a certain order. So that we could identify exactly the number of solutions and their locations. For $S > 11$, the situation is complicated by the impossibility of such a reduction in view of the restricted capabilities of supercomputers. Thus, we are forced to solve the order conditions using purely numerical approaches, such as the well-recognized Newton method. For this reason, one cannot guarantee that we will found all possible solutions. However, solving the system on a computer during significantly long time, one can say with a great probability that we have found almost all physically interesting solutions and chosen among them nearly optimal sets.

In this section we will concentrate attention on limiting cases with maximal possible stage numbers and orders which can still be explicitly considered. Our computations (performed with the help of a high speed digital Fortran compiler) have shown that within the direct decomposition approach, the maximal order of explicit non-gradient algorithms is limited to 6. Adding the force-gradient evaluations, this order can be increased to 8 that corresponds to a decomposition algorithm with $S = 23$ stages. Schemes of extremely high orders (up to $K = 16$) can be derived by composing lower-order algorithms.

4.1. Non-gradient schemes with 15 stages, order 6

4.1.1. Velocity version

Choosing $P = 8$ and putting $c_p = 0$ $(p = 1, \ldots, P)$ as well as $a_1 = 0$, one comes from general decomposition formula (6) to a fifteen-stage scheme in velocity form,

$$\text{BABABABABABABAB}. \tag{82}$$

There are eight independent variables, $b_1 = b_8$, $a_2 = a_8$, $b_2 = b_7$, $a_3 = a_7$, $b_3 = b_6$, $a_4 = a_6$, $b_4 = b_5$, and a_5, which must be determined by solving the same number (see Table 1) of the sixth-order conditions $\nu = \sigma = 1$, $\alpha = \beta = 0$, and $\gamma_{1-4} = 0$ (see Eqs. (11) and (12)). Here, we have found eleven real solutions (it seems that no other solutions exist) and the optimal set, which reduces the norm ζ of sixth-order truncation errors (see Eqs. (13) and (51)) to a minimum, is

$$\begin{aligned}
b_1 &= 0.8333333333333333\text{E}{-01},\\
a_2 &= 0.2465881872786138\text{E}{+00},\\
b_2 &= 0.3977675859548440\text{E}{+00},\\
a_3 &= 0.6047073875057809\text{E}{+00},\\
b_3 &= -0.3933369314462574\text{E}{-01},\\
a_4 &= -0.4009869039788007\text{E}{+00},\\
b_4 &= 0.5823277385644840\text{E}{-01},\\
a_5 &= 0.9938265838881204\text{E}{-01},
\end{aligned} \tag{83}$$

with $\zeta_{\text{min}} \approx 0.00000609$ and the efficiency Eff ≈ 1.40.

4.1.2. Position version

The position counterpart of propagation (82) reads

$$\text{ABABABABABABABA} \tag{84}$$

and can be reproduced from Eq. (6) at $P = 8$ by putting $c_p = 0$ $(p = 1, \ldots, P)$ as well as $b_8 = 0$. Solving the sixth-order conditions with respect to the unknowns $a_1 = a_8$, $b_1 = b_7$, $a_2 = a_7$, $b_2 = b_6$, $a_3 = a_6$, $b_3 = b_5$, $a_4 = a_5$, and b_4, we have realized that there exist sixteen real solutions (it can be stated with a great probability that there are no others) and the best one, which minimizes ζ to the value 0.0000109, looks as

$$\begin{aligned}
a_1 &= -1.0130879789171747\text{E}{+00},\\
b_1 &= 1.6600692650009894\text{E}{-04},\\
a_2 &= 1.1874295737325427\text{E}{+00},\\
b_2 &= -0.3796242142637736\text{E}{+00},\\
a_3 &= -1.8335852096460590\text{E}{-02},\\
b_3 &= 0.6891374118518106\text{E}{+00},\\
a_4 &= 0.3439942572810926\text{E}{+00},\\
b_4 &= 0.3806415909709257\text{E}{+00}
\end{aligned} \tag{85}$$

that corresponds to the efficiency Eff ≈ 0.780.

4.1.3. Remarks

The scheme (84) was investigated previously in the works by Forest [4] and Li [9]. But they presented solutions which are not optimal. For instance, Li used a different criteria for the optimization, namely, he tried to minimize the quantities $\delta_{\text{I}} = \sum_{p=1}^{P}(|a_p| + |b_p|)$ and $\delta_{\text{II}} = \max_{p=1}^{P}\{|a_p|, |b_p|\}$ rather than the true error-function ζ. As a result, he found a solution for which $\delta_{\text{I}} = 4.093$ and $\delta_{\text{II}} = 0.534$ (whereas $\delta_{\text{I}} = 7.644$ and $\delta_{\text{II}} = 1.187$ for the optimal algorithm (85)), but $\zeta_{\text{Li}} \approx 0.0000696$ that is approximately seven times worse with respect to our value $\zeta \approx 0.0000109$. We see, therefore, that the Li's criteria may have nothing to do with the best performance of the integration. Forest presented two solutions for position propagation (84), which correspond to the values $\zeta = 0.00505$ and 0.00141 that are even much worse than the result by Li.

Within the velocity sixth-order propagation (82) we were able to derive an algorithm (83) which, as can be seen, is more than ten times better with respect to the efficiency of the Li's integrator. Moreover, the new algorithm leads to a better value of $\delta_{\rm I} = 3.761$ (with $\delta_{\rm II} = 0.605$). It is also better by nearly two times with respect to our position counterpart (85). In addition, the velocity-like propagation (Eqs. (82) and (83)) has an advantage over the position version (Eqs. (84) and (85)) in that all the intermediate ($q \leqslant P$) states in velocity and position space stay during the integration within a given interval $[0, \Delta t]$, i.e. $0 \leqslant \sum_{p=1}^{q} a_p \leqslant 1$ and $0 \leqslant \sum_{p=1}^{q} b_p \leqslant 1$. This property may be important when solving ordinary differential equations which exhibit singularities beyond the interval of the integration. Equations of such a type may arise in pure mathematics applications or when writing the equations of motion for specific physical systems (note that any system of differential equations of the form $\mathrm{d}^2\mathbf{x}/\mathrm{d}t^2 = \mathbf{f}(\mathbf{x})$ is reduced to the equations of motion under consideration in this paper). It is worth remarking also that the velocity-like propagation (82) was not considered at all in the work [9] by Li. Forest [4] has presented one velocity-like scheme. However, it is far from the optimal solution (83), because the Forest's value for ζ is 0.000207, i.e. more than 30 times worse than our value 0.00000609. Moreover, the efficiency of our optimal fifteen-stage algorithm (83) is higher by a factor of 15 with respect to the three-star eleven-stage integrator No 40 (see Table 2). From the aforesaid, this fifteen-stage algorithm should be considered as the best among all decomposition integrators of the sixth order.

4.2. Force-gradient schemes with 23 stages, order 8

In the case when $K = 8$, we must satisfy eighteen order conditions, namely, $\nu = 1$, $\sigma = 1$, $\alpha = 0$, $\beta = 0$, $\gamma_{1-4} = 0$, and $\zeta_{1-10} = 0$. Taking into account the symmetry of time coefficients a_p, b_p, and c_p, this can be achieved at least at $P = 12$, i.e. using $S = 2P - 1 = 23$ single exponential operators (see Table 1). For $P = 12$, the decomposition formula (6) transforms into the following velocity-like

$$\text{CACACACACACACACACACACAC} \tag{86}$$

and position-like

$$\text{ACACACACACACACACACACACA} \tag{87}$$

force-gradient schemes. The number of unknowns for both the schemes is also equal to eighteen and we can try to solve the system of order conditions with respect to these unknowns.

As in the case of schemes (82) and (84), the values of non-linear functions (which constitute the system of order conditions) in the current points of the eighteen-dimensional space were calculated numerically using recursive relations (15)–(22), (A.1), and (A.2), but not explicit analytical expressions for them, to save the processor time and increase the precision of the computations (note that such expressions being written in Mathematica create a file of 0.5 Mb in length!). The initial guessed values of unknowns needed for the Newton solver were generated at random within the interval $[-2.5, 2.5]$ in each of the eighteenth directions. If the Newton's iterations diverge at a particular guess or during the calculations, a next random point was employed to repeat the process. In such a way, after several days of continuous attacking the systems of equations on a Silicon Graphics Origin 3800 workstation, we have found two and five solutions for schemes (86) and (87), respectively. The optimal solutions are presented in Appendix B.

In view of a complicated structure of the ninth-order truncation uncertainties $\mathcal{O}(\Delta t^9)$, these optimal solutions have been determined in a somewhat other way than above, namely, by providing a minimum for the function $\delta = \max_{p=1}^{P}\{|a_p|, |b_p|\}$ [as was mentioned above, this simplified criterion was used by Li [9] (and Kahan [30]), when optimizing usual non-gradient algorithms]. As a result, we have obtained $\delta_{\min} \equiv |a_5| = |a_9| \approx 0.721$ for scheme (86) and $\delta_{\min} \equiv |a_5| = |a_8| \approx 0.569$ for scheme (87). Since $\delta_{\min}$ is smaller in the latter case, the position-like integration should be considered as more preferable. Moreover, here all the intermediate time states remain within the interval of integration (that is contrary to the velocity-like scheme). The time coefficients for the position-like integration have been presented (see Appendix B) with thirty two significant digits (instead of sixteen, as above) to be used for very accurate integration. In order to ensure that all the digits shown are correct in both the cases, we have carried out a few additional Newton's iterations in Maple with up to 200 digits for the internal computations, taking as initial guesses the solutions obtained already in Fortran (within the quadruple precision).

Note also that in order to construct eighth-order decomposition schemes without involving any force gradients (i.e. letting $c_p \equiv 0$), it could be necessary to apply up to $2 \times 18 - 1 = 35$ (instead of 23) single exponential propagators (see Table 1). Such schemes cannot be obtained explicitly within the decomposition method because of the serious technical difficulties (but can be presented by composing second-order schemes, see the next Section 4.3). Instead, using the generalized force-gradient approach (6) has allowed us to derive the eighth-order algorithms by direct decompositions for the first time.

4.3. Composition schemes up to order 16

Increasing the order of integration to 10 and higher, the number of the order conditions and unknown time coefficients becomes too large. As a result, the construction of algorithms by direct decompositions (6) leads to a very difficult numerical problem, which cannot be solved at present because of the restricted capabilities of modern computers. However, having the already derived integrators of a lower order K, we can try to compose them as

$$S_Q(\Delta t) = S_K(d_1 \Delta t) \ldots S_K(d_P \Delta t) \ldots S_K(d_1 \Delta t) \tag{88}$$

for obtaining an algorithm of order $Q > K$. The composition constants d_p, where $p = 1, 2, \ldots, P$, should be chosen in such a way to provide the maximal possible value of Q at a given number $P \geqslant 2$. Note that the lower-order propagations $S_K(d_p \Delta t)$ enter symmetrically in composition (88), and their total number $2P - 1$ accepts odd values. So that if the basic integrator S_K is self-adjoint, the resulting algorithm S_Q will be self-adjoint as well.

The idea of using compositions like (88) is not new and has been applied by different authors in previous investigations [1,3,5,9,28–31]. But these investigations were limited, in fact, to the compositions of usual (non-gradient) second-order ($K = 2$) schemes. For instance, the fourth-order composition scheme of minimal length can be obtained [1] as a product of 3 second-order integrators S_2. If S_2 are represented by the 3-stage Verlet constructions (see algorithms Nos 1 and 2 in Table 2) this fourth-order scheme appears to be identical to that derived by Forest and Ruth [2] (algorithms Nos 11 and 12) using direct decompositions. The extended fourth-order algorithm by Suzuki [5] corresponds to a concatenation of 5 second-order Verlet integrators. Such a concatenation can be reduced to a decomposition scheme with $5 \times 3 - (5 - 1) = 11$ single-exponential stages. As has been recently shown [49], the norm of fifth-order truncation terms (see Eqs. (12) and (27)) of the Suzuki algorithm is $\gamma_{\mathrm{S}} \approx 0.0014$. This leads to the efficiency $\mathrm{Eff}_{\mathrm{S}} = 1/(5^4 \times 0.0014) \approx 1.14$ that is $0.0014/0.000027 \approx 52$ times worse than the efficiency of our best (within the non-gradient class) eleven-stage algorithm No 30 (see Table 2), obtained by the optimization of direct decompositions. The sixth-order scheme by Yoshida [1] relates to a composition with 7 integrators of the second order. It can be reduced to a decomposition scheme with $7 \times 3 - (7 - 1) = 15$ stages. Yoshida found three solutions [1] to his composition scheme and the best one leads, as has been realized by us, to the norm $\zeta_{\mathrm{Y}} \approx 0.00135$ of seventh-order truncation terms (see Eqs. (13) and (51)). Thus the maximal efficiency $\mathrm{Eff}_{\mathrm{Y}} = 1/(7^6 \times 0.00135) \approx 0.00630$ of the Yoshida integrators is in $0.00135/0.00000609 \approx 222$ times worse than that of our best fifteen-stage algorithm presented by Eqs. (82) and (83). We see, therefore, that the second-order-based compositions present, in fact, particular and very inefficient solutions of a more general direct decomposition approach. For this reason, the composition method should be recommended for the construction of algorithms of extremely high orders only, where the direct decompositions result in an unresolvable numerical problem.

Note that using the second-order-based compositions allows one to introduce algorithms to the tenth order [5,9, 30]. To our knowledge, no higher-order integrators, explicitly obtained within the composition approach (88), have been reported. Of course, employing the standard method by Creutz and Gocksch [54], an algorithm of arbitrary high order Q can be derived on the basis of S_2-integrators by iteratively applying $3^{(Q-2)/2}$ times the following triplet concatenation

$$S_{K+2}(\Delta t) = S_K(D_K \Delta t) S_K\big((1 - 2D_K)\Delta t\big) S_K(D_K \Delta t) \tag{89}$$

with $D_K = 1/(2 - 2^{1/(K+1)})$, where $K = 2, 4, \dots, Q - 2$. Chin and Kidwell [32] proposed to begin the iterations (89) immediately from $K = 4$ (not from $K = 2$) taking as the basis a force-gradient algorithm of order four (Eq. (46)), and indicated a visible increase in efficiency of the computations with respect to the second-order-iterated schemes. The main problem with the iterative method (89) is that when increasing the overall order Q, the total number of basic (lower-order) constructions increases much more rapidly than in the case of advanced compositions (88) (and such an increase does not lead to any improvement in the precision of the calculation). In particular, in order to obtain a tenth-order ($Q = 10$) algorithm using (88) it is necessary to build a chain composed of 31 S_2-elements [9,30], whereas $3^{(Q-2)/2} = 81$ such elements are needed within the crude iteration (89). For higher values of Q, the relative efficiency of the iterative approach will decrease yet more significantly (see below).

Quite recently [50], a general theory of construction of order conditions within the advanced composition approach (88) has been proposed. As a result, four-, sixth-, and eighth-order-based algorithms have been explicitly considered and the composition constants have been presented up to the overall order of 16. In particular, it has been established that the twelfth-order algorithm ($Q = 12$) can be composed with 23 (instead of 81 as follows from Eq. (89)) fourth-order integrators ($K = 4$). The fourteenth- and sixteenth-order composition algorithms ($Q = 14$ and $Q = 16$) can be cast as a product of 21 (instead of 81) sixth- and eighth-order integrators, respectively ($K = 6$ and $K = 8$). In actual calculations, for the basis of these advanced compositions we can choose our best integrators of order $K = 4$, 6, and 8, obtained in the preceding Section 3 by direct decomposition method (6). For instance, in the case of $Q = 16$, we can use the 23-stage gradient integrator of order $K = 8$ in position form (see Eq. (87) and Appendix B). This corresponds to a composition scheme with $21 \times 23 - (21 - 1) = 463$ single-exponential stages. On the other hand, in order to derive an algorithm of order $Q = 16$ on the basis of triplet concatenation (89) of a gradient scheme of order $K = 4$, it is necessary to apply $3^{(Q-K)/2} = 729$ fourth-order constructions. Each such a construction may consist of 5 to 11 stages (see Table 2). Even at a moderate length of 7 stages, one obtains that their total number is $729 \times 7 - (729 - 1) = 4375$. So that the advanced sixteenth-order algorithm will require nearly one order ($4375/463 \approx 9.45$) less operations than its crude counterpart. In addition, the precision of this algorithm appears to be several orders better [50] (see also Section 5). It can be useful in very accurate integrations to identify or confirm very subtle dynamical effects. The composition constants of the sixteenth-order algorithm are presented in Appendix C.

5. Applications

In order to verify our theoretical predictions presented in Sections 3 and 4, we have performed molecular dynamics (MD) and celestial mechanics (CM) simulations. The most outstanding algorithms of orders 2, 4, and 6 are tested in MD runs, whereas the schemes of orders 4, 6, 8, and 16 are considered in CM evaluations.

5.1. Molecular dynamics simulations

In MD simulations, the system under consideration was a collection of $N = 256$ identical ($m_i \equiv m$) particles interacting via a modified Lennard–Jones (LJ) potential, $\varphi(r) = \Phi(r) - \Phi(r_\text{c}) - (r - r_\text{c})\Phi'(r_\text{c})$ for $r < r_\text{c}$ and $\varphi(r) = 0$ otherwise, where $\Phi(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ is the genuine LJ potential. The particles were placed in a basic cubic box of volume $V = L^3$ in the absence of external fields ($u \equiv 0$). Periodic boundary conditions as well as the above modification of $\Phi(r)$ with $r_\text{c} = L/2 \approx 3.36\sigma$ have been used to reduce finite-size effects. It is worth pointing out that the modified potential φ and its first-order derivative $\varphi' = \mathrm{d}\varphi/\mathrm{d}r$ are continuous functions of r. In other words, the functions $\varphi'(r)$ and $\varphi''(r)$ (and thus the forces and force-gradients, see Eq. (8)) are free of singularities, despite the truncation at $r = r_\text{c}$. This avoids the energy drift (caused by the passage of particles through the surface of truncation sphere) when integrating the equations of motion with up to the fourth-order accuracy in the time step. For sixth-order algorithms, a higher-fold correcting term, $-\frac{1}{2}(r - r_\text{c})^2\Phi''(r_\text{c})$, should be included in $\varphi(r)$ additionally (note that for the system under consideration the finite-size corrections

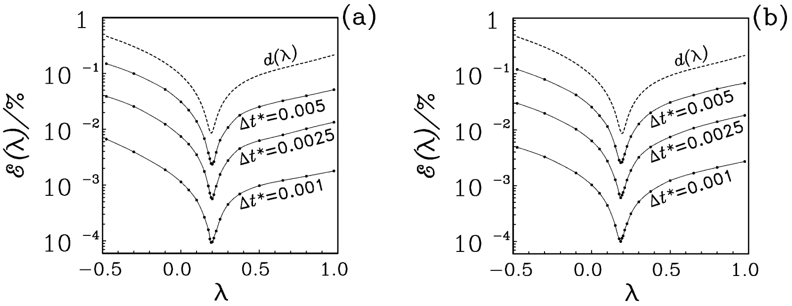

Fig. 1. The total energy fluctuations obtained in the MD simulations for different values of free parameter λ at three reduced time steps, $\Delta t^* = 0.005$, 0.0025, and 0.001, using the velocity (subset (a), Eq. (25), $\xi = \chi = 0$) and position (subset (b), Eq. (32), $\xi = 0$) five-stage decompositions of the second order. The simulation results are presented by circles connected by the solid curves. The error function $d(\lambda) = (\alpha^2 + \beta^2)^{1/2}$ (see Eq. (11)) is plotted in both subsets by the dashed curve.

are small, $\Phi(r_c) \approx -3 \times 10^{-3}$, $\Phi'(r_c) \approx 5 \times 10^{-3}$ and $\Phi''(r_c) \approx -10^{-2}$, i.e. $\varphi(r) \approx \Phi(r)$, and they disappear completely for $r_c/\sigma \gg 1$). The simulations were carried out in a microcanonical ensemble at a reduced density of $n^* = \frac{N}{V}\sigma^3 = 0.845$ and a reduced temperature of $T^* = k_B T/\epsilon = 1.7$ (a typical thermodynamic point of LJ fluids). All the runs were started from an identical well equilibrated initial configuration, and have been continued during $l = 10\,000$ time steps. The truncation uncertainties of the algorithms were determined by measuring the relative total energy fluctuations $\mathcal{E} = \langle (E - \langle E \rangle)^2 \rangle^{1/2} / |\langle E \rangle|$, where $E \equiv H$ (see Eq. (2)) and $\langle\,\rangle$ denotes the microcanonical averaging. Note that if the equations of motion could be solved exactly, the above fluctuations should vanish, because in microcanonical ensembles the total energy is an integral of motion, $E(t) = E(0)$. Therefore, in approximate MD simulations, smaller values of $\mathcal{E}$ will correspond to a better precision of the integration.

First of all, we would like to demonstrate that actual numerical uncertainties of the algorithm can indeed be expressed in terms of the norm of leading error coefficients, so that the minimization of that norm do leads to the best solutions. To simplify the presentation, let us consider extended decomposition schemes with one free parameter. In the case of order 2, the most notorious examples are the velocity $\mathrm{B}_\lambda\mathrm{ABAB}_\lambda$ and position $\mathrm{A}_\lambda\mathrm{BABA}_\lambda$ versions of the non-gradient ($\xi = 0$ and $\chi = 0$ in Eqs. (25) and (32)) scheme with five stages. For both these versions, the leading error coefficients are α and β, so that the norm $(\alpha^2 + \beta^2)^{1/2} \equiv d(\lambda)$ will depend on free parameter λ. The total energy fluctuations obtained in the simulations at the end of the runs for three (fixed within each run) dimensionless time steps, $\Delta t^* = \Delta t(\epsilon/m\sigma^2)^{1/2} = 0.005$, 0.0025, and 0.001, are shown in Fig. 1 as functions of λ (this parameter, being unchanged within each run, varied from one run to another). The subsets (a) and (b) of this figure correspond to the velocity (Eq. (25)) and position (Eq. (32)) versions, respectively.

As can be seen readily, all the functions $\mathcal{E}(\lambda, \Delta t)$ have one minimum located at the same point $\lambda \approx 0.19$ independent of the size Δt of the time step. The λ-location of this point coincides completely with that of the minimum (see Eq. (31)) of function $d(\lambda)$ that is plotted (by dashed curves) in the subsets of Fig. 1 as well. Moreover, the energy fluctuations $\mathcal{E}(\lambda, \Delta t)$ appear to be proportional to the norm $\Gamma_2(\lambda, \Delta t) = d(\lambda)\Delta t^2$ of global errors (see Section 3.2.3), and the coefficient of this proportionality almost does not depend on λ and Δt. In addition, at each step size considered the energy fluctuations decrease at the minimum by more than ten times with respect to those at $\lambda = 0$. This is in agreement with the ratio $0.0932/0.00855 \approx 11$ of truncation errors Err_3 predicted for the optimized algorithms Nos 5 and 6 and their original Verlet counterparts Nos 1 and 2 (see Table 2). Note that at the minimum of $d(\lambda)$, the schemes $\mathrm{B}_\lambda\mathrm{ABAB}_\lambda$ and $\mathrm{A}_\lambda\mathrm{BABA}_\lambda$ reduce by definition to the algorithms Nos 5 and 6, whereas $\lim_{\lambda\to 0}\mathrm{B}_\lambda\mathrm{ABAB}_\lambda = \mathrm{ABA}$ (integrator No 2) and $\lim_{\lambda\to 0}\mathrm{A}_\lambda\mathrm{BABA}_\lambda = \mathrm{BAB}$ (integrator No 1).

A similar pattern has been observed in the case of fourth-order integration, where the equations of motion were solved with the help of the one-parameterization seven-stage gradient $\mathrm{AB}_\lambda\mathrm{ACAB}_\lambda\mathrm{A}$ (Eq. (40), $\xi = 0$) and nine-

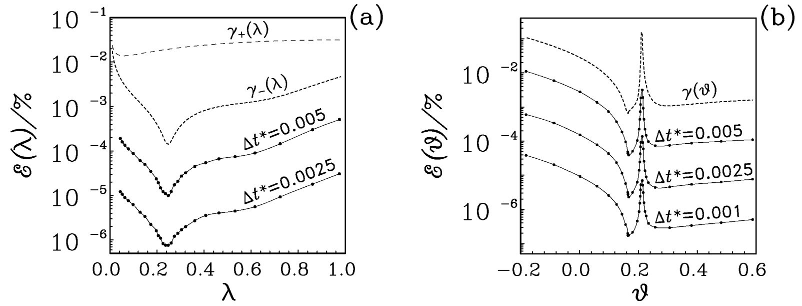

Fig. 2. The total energy fluctuations obtained in the MD simulations at different values of free parameter λ (subset (a)) and ϑ (subset (b)) for a set of time steps, $\Delta t^* = 0.005$, 0.0025, and 0.001, using the (a) gradient seven-stage (Eq. (44), $\xi = 0$) and (b) non-gradient nine-stage (Eq. (48), $\mu = \xi = \chi = 0$) schemes of the fourth order. The simulation results are presented by circles connected by the solid curves. The error functions $\gamma_\pm(\lambda)$ and $\gamma(\vartheta)$ (see the text) are plotted by the dashed curves.

stage non-gradient $\mathrm{B}_\vartheta\mathrm{ABABABAB}_\vartheta$ (Eq. (48), $\mu = \xi = \chi = 0$) schemes. The functions $\mathcal{E}(\lambda, \Delta t)$ and $\mathcal{E}(\vartheta, \Delta t)$ obtained for these schemes in our simulations are presented in subsets (a) and (b) of Fig. 2, respectively. In the former scheme, the norm of principal error coefficients (see Eqs. (12) and (27)) is described by the two-valued function $\gamma_\pm(\lambda)$ (see Eq. (41) and that just following it, where the two signs "±" correspond to two different sets for the remaining time coefficients χ and θ at the same value of λ). For the latter scheme, the norm is determined by the one-parameterization function $\gamma(\vartheta)$ (Eq. (49) with $\mu = \xi = \chi = 0$). The functions $\gamma_\pm(\lambda)$ and $\gamma(\vartheta)$ are drawn in Fig. 2 (a) and (b), respectively, too (note that the MD results shown in subset (a) relates to sign "$-$" in Eq. (41)).

Again, as for the case of order two (Fig. 1), it can be seen in Fig. 2 that each of the functions $\mathcal{E}(\lambda, \Delta t)$ and $\mathcal{E}(\vartheta, \Delta t)$ has one global minimum, namely, at $\lambda \approx 0.247$ and $\vartheta \approx 0.164$, respectively. The position of this minimum (on λ- or ϑ-axis) practically does not depend on Δt and coincides with that of the minimum given by Eqs. (44) and (57) correspondingly for the functions $\gamma_-(\lambda)$ and $\gamma(\vartheta)$. Note that at the above minima, the schemes $\mathrm{AB}_\lambda\mathrm{ACAB}_\lambda\mathrm{A}$ and $\mathrm{B}_\vartheta\mathrm{ABABABAB}_\vartheta$ transform into the optimized algorithms marked in Table 2 under Nos 14 and 19, respectively. On the other hand, the depth of these minima depends on the size Δt of the time step in a characteristic way and is proportional to the norm $\Gamma_4 = \gamma \Delta t^4$ of fourth-order global errors. It is worth remarking that for $\lambda = 3/8 = 0.375$, the decomposition $\mathrm{AB}_\lambda\mathrm{ACAB}_\lambda\mathrm{A}$ reduces to the C-integrator by Chin [12]. The level of truncation uncertainties $\mathcal{E}(\lambda, \Delta t)$ for this integrator is approximately 5 times higher than that related to our optimized ($\lambda \approx 0.247$) algorithm No 14 (see Fig. 2(a)). This is in agreement with the predicted value $\gamma_-(\lambda = 3/8)/\gamma_-(\lambda \approx 0.247) \approx 5 \approx 11.3/2.2$ for the ratio of efficiencies of algorithm No 14 and C-integrator (see Table 2 and Section 3.3.3). Further, for $\vartheta = 0$, the nine-stage decomposition $\mathrm{B}_\vartheta\mathrm{ABABABAB}_\vartheta$ reduces to the seven-stage algorithm ABABABA by Forest and Ruth [2] (No 12 in Table 2). Here, the fluctuations $\mathcal{E}(\vartheta, \Delta t)$ decrease at the minimum by 40 to 50 times with respect to those at $\vartheta = 0$ (see Fig. 2). This is also very close to our prediction $\gamma(\vartheta = 0)/\gamma(\vartheta \approx 0.164) \approx 0.0283/0.000654 \approx 43$ based on the ratio of γ-functions (look at lines under Nos 12 and 19 in Table 2). We see, therefore, that the criterion used to measure the precision of the integration in terms of the norm of local truncated uncertainties is in excellent accord with our simulation results, and the algorithms derived are indeed the best within a given decomposition group.

And now we are in a position to test properties of the most outstanding algorithms introduced in more detail and compare them with those corresponding to the best previously known ones, such as the second-order Verlet integrators [26,27] (Nos 1 and 2, see Table 2), the fourth-order Forest–Ruth [2] (No 12), Chin [12] (Eqs. (40) and (46)) and Suzuki [5] algorithms, as well as the sixth-order schemes by Yoshida [1] and Li [9]. According to the conclusion given in Section 3.6, the non-gradient velocity- and position-Verlet-like algorithms Nos 5 and 6 are the most outstanding within order 2 (and within a 5-stage group). For order 4, the best algorithms proposed appear in Table 2 under No 23 (gradient, $S = 9$ stages), No 30 (non-gradient, $S = 11$), and No 36 (gradient, $S = 11$). For completeness of comparison, the algorithm No 20 (as the best within the non-gradient class at $S = 9$) and

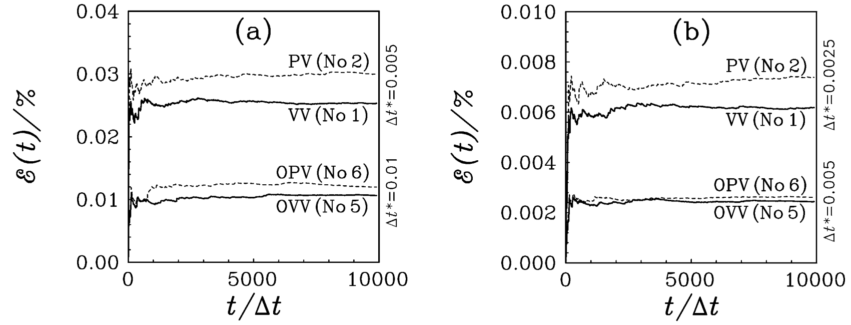

Fig. 3. The total energy fluctuations as functions of the length of the MD simulations performed using the optimized velocity-Verlet- (solid curve marked as OVV) and position-Verlet- (dashed curve, OPV) like algorithms (Nos 5 and 6, see Table 2) as well as the original VV (solid curve, VV) and PV (dashed curve, PV) integrators (Nos 1 and 2) of the second order. The results corresponding to different values of the time step, namely, $\Delta t^* = 0.01$ and 0.005 as well as 0.005 and 0.0025 are presented in subsets (a) and (b), respectively.

the gradient algorithm No 14 (as the most prominent within a 7-stage group) will be considered also. In the case of order 6, we will employ the gradient scheme No 28 of minimal length with nine stages and our most efficient non-gradient integrator with fifteen stages described by Eqs. (82) and (83).

The result for the total energy fluctuations as functions of the length of the simulations corresponding to the optimized velocity- (OVV) and position-Verlet (OPV) algorithms Nos 5 and 6 (see Table 2) is presented in Fig. 3 for two fixed time steps, $\Delta t^* = 0.01$ (subset (a)) and $\Delta t^* = 0.005$ (subset (b)). The functions related to the original VV [26] and PV [26,27] integrators Nos 1 and 2 are drawn there as well. Note that for the original versions, the time step within each subset was chosen to be always twice smaller than that of the optimized counterparts, namely, $\Delta t^* = 0.005$ and 0.0025 (see subsets (a) and (b), respectively). This condition is necessary to provide the same total number of force recalculations during the same observation time (Section 3.2.3).

As we can see, both the original (VV and PV) and optimized (OVV and OPV) algorithms exhibit excellent stability properties (that can be explained [17,18] by the symplecticity and time reversibility of the produced solutions). No systematic deviations in the total energy fluctuations can be observed for all the integrators. Instead, in each of the cases the amplitude of these deviations tends to its own value which does not increase with further increasing the length of the simulations. However, this value appears to be significantly larger for the original versions VV and PV. On the other hand, using the optimized OVV and OPV algorithms even with double sizes of the time step allows us to decrease the unphysical energy fluctuations approximately by a factor of three. This is in good agreement with our theoretical prediction for the ratio $29.2/10.7 \approx 2.7$ of efficiencies (which are presented in Table 2 under lines No 1/2 and 5/6). Note also that the step size of $\Delta t^* = 0.005$ should be considered as an upper limit for precise MD simulations of the LJ fluid within the original Verlet algorithms of the second order. This is so because at $\Delta t^* > 0.005$, the higher-order truncation uncertainties become too big. In particular, then the ratio of the total energy fluctuations to the fluctuations in potential energy appears to be more than a few per cent. Applying the optimized algorithms makes it possible to increase this limit approximately by a factor of $(\mathrm{Err}_{K+1}/\mathrm{Err}^\star_{K+1})^{1/K}$, where Err_{K+1} and $\mathrm{Err}^\star_{K+1}$ are the norms of leading error coefficients of usual and optimized schemes, respectively. For instance, in the case of $K = 2$, such a factor is equal to $(0.0932/0.00855)^{1/2} \approx 3.3$.

In practice, however, the upper limit for Δt may be somewhat lower than that we just predicted theoretically. The reason is that we considered only the leading term $C_K \mathrm{Err}_{K+1} \Delta t^K$ of truncation uncertainties and did not take into account the influence $\sum_{M=2,4,\ldots}^{\infty} C_{K+M} \mathrm{Err}_{K+M+1} \Delta t^{K+M}$ of higher-order terms. This influence (being small at small Δt) increases with increasing the size of the time step and can be important in the regime of maximal possible (in view of an acceptable level of precision) values of Δt. It is not such a simple matter to handle analytically the infinite sum of the truncated terms, because the factors C_{K+M} are not universal (see the remarks at the end of Section 3) and depend on the form of the interaction potential φ. Another related and more difficult question is how to determine the Δt-region of stability of an algorithm. To answer this question we need to consider

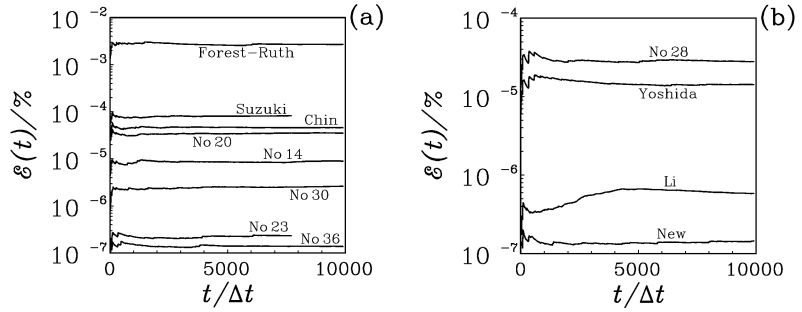

Fig. 4. The total energy fluctuations versus the length of the MD simulations carried out within the previously known integrators by Forest–Ruth, Suzuki, Chin, Li, and Yoshida as well as our most outstanding new algorithms of orders four (subset (a), Nos 14, 20, 23, 30, and 36) and six (subset (b), see the text). All the results correspond to a time step of $\Delta t^* = 0.005$.

not only a possible divergence of the sum $\sum_{M=2,4,\ldots}^{\infty} C_{K+M}\mathrm{Err}_{K+M+1}\Delta t^{K+M}$ at sufficiently large values of Δt, but also a complicated process of accumulation of local truncation uncertainties $C_{K+M}\mathrm{Err}_{K+M+1}\Delta t^{K+M}$ during the step-by-step integration. Again, no general answer can be given here, because even the ratio of stability regions of two algorithms corresponding to the same system will depend on details of the interactions in that system. The most reliable way to determine the Δt-stability is empirical and lies in performing direct measurements in the simulations for each particular situation.

The total energy fluctuations obtained in the simulations when integrating the motion using the most outstanding new algorithms of orders 4 and 6 are presented in subsets (a) and (b) of Fig. 4, respectively. As was recognized above, in the case of order 4 such algorithms are those marked under Nos 14, 20, 23, 30, and 36. For order 6, the most efficient integrator (marked simply as “New” in Fig. 4(b)) is that described by Eqs. (82) and (83). For the purpose of comparison the best previously known algorithms by Forest and Ruth [2] (No 12), Chin [12] (Eqs. (40) and (46)), Suzuki [5], Yoshida [1] and Li [9] as well as the sixth-order scheme No 28 are also included in the corresponding sets of Fig. 4. In order to simplify the presentation, all the results shown in this figure correspond to the same value $\Delta t^* = 0.005$ of the time step.

As can be seen in subset (a) of Fig. 4, the level $\mathcal{E}$ of truncation errors can be reduced by more than 4 orders by increasing the number of stages from 7 (Forest–Ruth integrator No 12) to 11 (algorithm No 36). This is in agreement with the predicted value for the ratio $0.00000127/0.0283 \approx 4 \times 10^{-5}$ of the corresponding fifth-order uncertainties Err_5 (look at Table 2). Moreover, as it follows from Table 2, the relative levels in energy conservation $\mathcal{E}$ related to each of the other algorithms presented in Fig. 4 can also be expressed in terms of Err_5 (subset (a), order 4) or seventh-order truncation uncertainties Err_7 (subset (b), order 6). Another conclusion, which can be made when looking at Fig. 4 (and which is in agreement with our theoretical predictions too) is that the previously known integrators are clearly worse than the new algorithms introduced. It is interesting to remark also that the best fourth-order integrator No 36 leads to approximately the same level in energy conservation as that of the best sixth-order integrator (please look at curves labeled as No 36 and “New” in subsets (a) and (b), respectively). In other words, $C_4\mathrm{Err}_5\Delta t^4 \approx C_6\mathrm{Err}_7\Delta t^6$ at $\Delta t^* \sim 0.005$, because of the smallness of C_4. Note, however, that the integrator No 36 cannot be classified as a pure sixth-order scheme, given that the coefficient C_4 does not vanish completely. In addition, this integrator will lead to better (worse) results at $\Delta t^* > 0.005$ ($\Delta t^* < 0.005$), because with increasing (decreasing) Δt the fourth-order term $C_4\mathrm{Err}_5\Delta t^4$ increases (decreases) less rapidly than the sixth-order counterpart $C_6\mathrm{Err}_7\Delta t^6$.

The CPU time used per step by each of the algorithms considered in our MD simulations was proportional, as expected, to the factor $(n_{\mathrm{f}} + g_{\mathrm{MD}}n_{\mathrm{g}})$ with $g_{\mathrm{MD}} \approx 2$. The coefficient C_{CPU} of this proportionality was found to be almost independent on the type of the algorithm. For this reason, the actual MD efficiency and the efficiency predicted by Eq. (81) (and presented in Table 2) nearly coincided between themselves. On the other hand, the quantity C_{CPU} varied strongly on the type of the computer employed. For instance, in the case of the Silicon

Graphics Origin 3800 workstation, we have obtained that $C_{\mathrm{CPU}} \approx 5 \times 10^{-3}$ sec (for one processor). Note also that this result corresponds to the LJ system with $N = 256$ particles, and the factor C_{CPU} will increase with increasing N as $\propto N^2$ (at $r_{\mathrm{c}} = L/2$).

5.2. Celestial mechanics simulations

In CM simulations, we dealt with a motion of a particle (planet) of mass m_1 in the (gravitational) field $u(r) = -c/r$ of the central body (sun) with mass $m_2 \gg m_1$, where $c > 0$ is the constant responsible for the intensity of the interaction. For simplifying the calculations, the influence of all other ($i = 3, 4, \ldots, N$) particles (planets, for which $m_i \ll m_2$) in the (solar) system have been neglected ($\varphi \equiv 0$). Then the equations of motion become particularly simple and look as

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}, \qquad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\mathbf{r}}{r^3}, \tag{90}$$

where $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$, and we have used units in which the reduced mass $m_1 m_2/(m_1 + m_2)$ as well as the interaction constant c are equal to unity (so that $\mathbf{a} \equiv \mathbf{f} = -\mathbf{r}/r^3$ as well as $\mathbf{g} \equiv \mathbf{h} = -4\mathbf{r}/r^6$, see Eq. (8)). Since the quantity $E = v^2/2 - 1/r$ (which is associated with the total energy) presents an integral of motion for Eq. (90), it should be conserved during the integration. However, this will be so if these equations are solved exactly. In numerical simulations, the local truncation uncertainties $\mathcal{O}(\Delta t^{K+1})$ accumulate step by step of the integration process, leading at $t \gg \Delta t$ to the global errors $\mathcal{O}(\Delta t^K)$, where K denotes the order of precision. So that for self-adjoint algorithms, the total energy fluctuations $\mathcal{E}(t) = \langle [E(t) - E(0)]^2 \rangle^{1/2}/|E(0)|$ can be presented as a function of time as

$$\mathcal{E}(t) = \mathcal{E}_K(t)\Delta t^K + \mathcal{O}\bigl(\Delta t^{K+2}\bigr), \tag{91}$$

where $\mathcal{E}_K$ is the main step-size independent error coefficient, and $\langle\,\rangle$ denotes the averaging along phase trajectories.

In our simulations we solved a two-dimensional Kepler problem (90) with the same initial conditions $\mathbf{r}(0) = (10, 0)$ and $\mathbf{v}(0) = (0, 1/10)$ as those used by previous authors [12,32]. Note that the resulting highly eccentric ($e = 0.9$) orbit provides a nontrivial testing ground for trajectory integration. The numerical effectiveness of each algorithm was gauged in terms of the leading error coefficient $\mathcal{E}_K = \lim_{\Delta t \to 0} \mathcal{E}/\Delta t^K$ (see Eq. (91)). It can actually be extracted from the fraction $\mathcal{E}/\Delta t^K$ by choosing smaller and smaller time steps Δt to be entitled to ignore completely the higher-order uncertainties $\mathcal{O}(\Delta t^{K+2})$. This can be observed in the neighborhood of $\Delta t \sim P/5000$, where $P = \pi/(2|E(0)|^3)^{1/2}$ is the period of the elliptical orbit. All the calculations have been carried out in Fortran using double (for orders 4 and 6, as for MD simulations) or quadruple (for orders 8 and 16) precision arithmetics.

The leading error coefficients $\mathcal{E}_K$ obtained in the CM simulations using various schemes of orders $K = 4$, 6, 8, and 16 are plotted in subsets (a), (b), (c), and (d) of Fig. 5, respectively, as functions of the observation time t. Here, for orders 4 and 6 we have employed practically the same set of algorithms as in Section 5.1 for the case of MD simulations (adding only one sixth-order integrator No 40, see comments below). The higher-order schemes used will be described later. Note that the changes in $\mathcal{E}_K$ are substantial only near the first mid period ($t/P = 1/2$), when the particle is at its closest position to the attractive center and when the process of accumulation of energy deviations just started. With further increasing t, the function $\mathcal{E}_K(t)$ only slightly corrects itself (especially at next mid-period points, $t/P = 3/2, 5/2, \ldots$) and already begins to tend to a limiting value $\lim_{t\to\infty} \mathcal{E}_K$ after 10 periods. Such a behavior of $\mathcal{E}_K(t)$ is caused by the fact that within a symplectic integration, the non-conservation $(E(t) - E(0))$ of energy for periodic orbits is periodic and the fluctuations $\mathcal{E}$ are bounded and independent on t when averaged over long times $t \gg P$ (this independence of averaged energy fluctuations for MD simulations has already been demonstrated at $t \gg \Delta t$ in Fig. 4).

As can be seen when comparing the results of Fig. 4 (a) and (b) with those of Fig. 5 (a) and (b), the main conclusions made in the preceding MD subsection are valid for the CM evaluations as well. Namely, the gradient algorithm No 36 is definitely the best within the fourth-order integration. The non-gradient integrator of order six

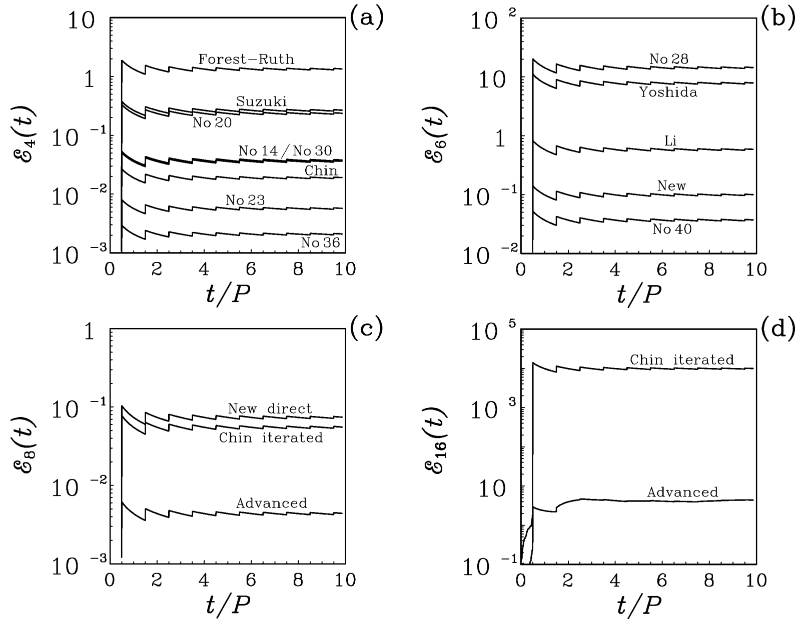

Fig. 5. The normalized energy deviations of a particle in a Keplerian orbit as functions of the length of the CM simulations. The results obtained within fourth-, sixth-, eighth-, and sixteenth-order schemes are shown in subsets (a), (b), (c), and (d), respectively. The schemes used are the previously known integrators by Forest–Ruth, Suzuki, Chin, Li, and Yoshida as well as our most outstanding new algorithms (see the text).

marked as “New” in Fig. 5(b) (and described by Eqs. (82) and (83)) leads also to much superior integration with respect to the previously known schemes of the corresponding order. In addition, the new sixth-order gradient scheme No 40 results even in a more accurate evaluation of phase trajectory. It is worth remarking that this scheme was not included into the consideration when performing the MD runs. The reason is that it requires a relatively large number ($n_{\rm g} = 3$) of force-gradient recalculations per time step (in MD simulations, the calculation of force gradients is the most time-consuming part of the computations). This decreases the MD efficiency of the algorithm No 40 to a value of Eff ≈ 0.0930 (see Table 2). Although this efficiency is the highest one among those corresponding to all other sixth-order gradient schemes introduced, it appeared to be significantly lower than that of the best non-gradient algorithm of order 6 (Eff ≈ 1.40, see Eqs. (82) and (83)).

As was pointed out earlier (see Section 3.6), in CM simulations the pattern with efficiency is somewhat different, because of a specific form of the interaction potential and a small number ($N \sim 10$) of particles (planets), whereas $N \sim 10^2$ to 10^5 for MD systems. In other words, the factor g (which reflects the relative work spent on the calculation of force gradients) will take here smaller values, $g_{\rm CM} < g_{\rm MD}$. This increases the CM efficiency of gradient schemes by $[(n_{\rm f} + g_{\rm MD} n_{\rm g})/(n_{\rm f} + g_{\rm CM} n_{\rm g})]^K$ times (see Eq. (81)) with respect to that presented in Table 2 for the MD case. For example, by decreasing g from $g_{\rm MD} = 2$ to $g_{\rm CM} = 1/2$, the efficiency of the algorithm No 40 (when $n_{\rm f} = 5$, $n_{\rm g} = 3$, and $K = 6$) will increase from ${\rm Eff}_{\rm MD} \approx 0.0930$ to ${\rm Eff}_{\rm CM} = [(5+2\times 3)/(5+1/2\times 3)]^6 {\rm Eff}_{\rm MD} \approx 2.18$ that is even slightly higher than the efficiency Eff ≈ 1.40 of the above non-gradient ($n_{\rm g} = 0$) scheme (Eqs. (82) and (83)).

We see, therefore, that the best algorithms predicted do lead to the most precise integration in both MD and CM simulations. Nevertheless, as in the case of determining the MD and CM efficiencies (see the preceding paragraph), there is also a difference when comparing relative levels of precision (corresponding to some algorithms with respect to each other) within these two types of simulations. This difference is not visible for sixth-order integrators, where the order of horizontal locations of energy conservation levels related to schemes labeled by (from top to bottom) No 28, Yoshida, Li, and New is the same as in MD as well as in CM computations (please compare Figs. 4(b) and 5(b)). But for fourth-order algorithms with moderate levels of precision (with respect to the best integrator No 36), namely for those marked as No 14, No 20, No 30, and Chin, the relative horizontal locations of $\mathcal{E}_{(\mathrm{MD})}$ and $\mathcal{E}_{4(\mathrm{CM})}$ are not the same (compare the corresponding curves of Figs. 4(a) and 5(a)). For instance, in the case of MD simulations, the algorithm No 14 is evidently more accurate (as was pointed out in Section 5.1) than the integrator by Chin. On the other hand, such a relative advantage in accuracy disappears in our CM simulations, where both schemes become to be comparable in efficiency (the algorithm No 14 leads here even to slightly worse results). Another example is the algorithm No 30 which produces more precise solutions in MD simulations than those of No 14 (see Fig. 4(a)), but within the CM runs these two algorithms exhibit an equivalence in energy conservation (the corresponding curves coincide in Fig. 5(a)).

In order to understand such a situation, it should be taken into account that in the CM simulations we actually deal with one particle moving in an effective external field. Moreover, such a motion is periodic and covers only a small part of phase space corresponding to one particular set of initial conditions. In contrast, for statistical many-body systems, the phase point may visit considerably wider regions of phase space due to the interparticle interactions and due to an extended set of initial conditions chosen at random for each particle. Then, for instance in the case of fourth-order integration, all the four fifth-fold commutators $\mathcal{G}_1 = [\mathcal{A}, [\mathcal{A}, [\mathcal{A}, [\mathcal{A}, \mathcal{B}]]]]$, $\mathcal{G}_2 = [\mathcal{A}, [\mathcal{A}, [\mathcal{B}, [\mathcal{A}, \mathcal{B}]]]]$, $\mathcal{G}_3 = [\mathcal{B}, [\mathcal{A}, [\mathcal{A}, [\mathcal{A}, \mathcal{B}]]]]$, and $\mathcal{G}_4 = [\mathcal{B}, [\mathcal{B}, [\mathcal{A}, [\mathcal{A}, \mathcal{B}]]]]$, being averaged over the phase points of all particles, will enter approximately with equal weights when forming the total error vector $\mathcal{O}_5 = \sum_{k=1}^{4} \gamma_k \mathcal{G}_k$ of local fifth-order uncertainties (see Eq. (12)). This has been tentatively assumed when writing the norm of that vector with respect to such commutators in terms of the sum $\gamma^2 = \sum_{k=1}^{4} \gamma_k^2$ of squares of all $(k = 1, 2, 3, 4)$ the error-function components γ_k (see Eq. (27)) which appear as multipliers at $\mathcal{G}_k$. For one-particle dynamics (especially with periodic highly eccentric orbits) the above weights may differ considerably with respect to one another depending on the current values of position $\mathbf{r}(t)$ and velocity $\mathbf{v}(t)$ of the particle. This complicates an analysis of the truncation terms and makes it impossible to find an exact global minimum for them within any analytical approach.

Note, however, that even for the specific CM system, the assumption on uniform contribution of different components γ_k into the error-vector norm γ works relatively well. So that the minimization of the sum $\sum_{k=1}^{4} \gamma_k^2$ should be considered as a procedure that allows one to obtain analytically the best algorithm in its most general consideration. Of course, one can try to reach empirically an extra optimization for each given system with particular initial conditions by performing trial runs corresponding to different sets of time decomposition coefficients. Further, one can choose such optimal coefficients which minimize actual observable uncertainties of the concrete investigated quantity, rather than those which minimize the sum $\sum_{k=1}^{4} \gamma_k^2$. But the additional computational efforts spent on such runs will not be compensated, as a rule (especially in the case of decompositions with two and three free parameters), by a slightly increased precision of the decompositions.

A special case is the decomposition integration of motion for perturbed and weakly interacting systems, when the potential part of the Liouville operator can be cast in the form $\mathcal{B} = \varepsilon B$ with $\varepsilon \ll 1$, i.e. when $\|\mathcal{B}\| \ll \|\mathcal{A}\|$. Then, for instance, $\mathcal{G}_2$, $\mathcal{G}_3$, and $\mathcal{G}_4$ will behave like ε^2 or ε^3 and can be neglected with respect to $\mathcal{G}_1 \sim \varepsilon$. Therefore, having at least one free parameter it is quite natural here to reduce the first component γ_1 to zero exactly, rather than to find a non-zero global minimum for the total sum $\sum_{k=1}^{4} \gamma_k^2$ (because the contribution of $g_k \mathcal{G}_k$ at $k = 2, 3$, and 4 will be negligibly small compared to the main term $g_1 \mathcal{G}_1$). As was predicted by McLachlan [39] and rigorously proved quite recently by Laskar and Robutel [41], the time constants a_p and b_p of decomposition (6) can be expressed at $\varepsilon \ll 1$ for arbitrary order K in terms of nodes and weights of the well-known Gaussian quadrature (widely used for usual integration of one-dimensional functions). Then at $c_p \equiv 0$, the leading truncation terms in Δt and ε of the decomposition formula will behave like $\mathcal{O}(\varepsilon \Delta t^{K+1})$ and $\mathcal{O}(\varepsilon^2 \Delta t^2)$. When ε is not too small, it is desirable also to kill the latter term, and this can be achieved by appropriately choosing non-zero values for c_p [41].

Consider, finally, the results of the CM simulations obtained within eighth- and sixteenth-order integrating schemes (Figs. 5 (c) and (d), respectively). Until recently, the only way to construct integrators of such high orders was to apply the crude iterative concatenation (89) starting with lower-order integrators. In particular, it has been shown [32] that (see also Section 4.3) using the fourth-order gradient integrator by Chin as a basis for the iterations allows one to produce much more efficient higher-order algorithms with respect to those based on the non-gradient integrators of the second order. The leading error coefficients $\mathcal{E}_8$ and $\mathcal{E}_{16}$ corresponding to the Chin iterated algorithms of order 8 and 16 are, thus, plotted in subsets (c) and (d) of Fig. 5 for comparison with our new algorithms. In the case of order 8, the new algorithms employed were the direct decomposition scheme (see Eq. (87) and Appendix B) as well as an algorithm based on advanced composition (88) of the fourth-order integrator No 14. For order 16, we have tested an advanced composition algorithm (see Appendix C) based on the above direct decomposition scheme of order 8.

As can be seen in Fig. 5(c), the direct eighth-order decomposition (marked in the figure as "New direct") and Chin iterated algorithms are comparable in precision of energy conservation. However, the former scheme uses $n_{\rm f} = 11$ force and the same number $n_{\rm g} = 11$ of gradient evaluations per time step, whereas in the latter case we have that $n_{\rm f} = 9 \times 3 = 27$ and $n_{\rm g} = 9 \times 1 = 9$. So that the CM efficiency (when $g \equiv g_{\rm CM} = 1/2$) of the direct scheme will be $[(27 + 1/2 \times 9)/(11 + 1/2 \times 11)]^8 \approx 175$ times higher than that of the Chin's approach. On the other hand, the relative level of accuracy of the advanced composition scheme (marked simply "Advanced") is a factor of 17 higher with respect to that of the direct decomposition, but (being composed of 7 integrators No 14) it requires $[(7 \times 3 + 1/2 \times 7)/(11 + 1/2 \times 11)]^8 \approx 1.48$ times more computational time. As a result, its efficiency is lower by a factor of $1.48^8/17 \approx 1.35$, and the direct decomposition approach appears to be the best. In the case of sixteenth-order integration (see Fig. 5(d)), the advanced composition algorithm is tree orders better in precision than the Chin iterated scheme. In addition, this advanced algorithm requires approximately one order less operations (see the end of Section 4.3) on the evaluation of forces and their gradients.

As in the case of MD simulations, we have realized that the CPU time spent on performing our CM runs can be described with a great precision by the same (in form) function $C_{\rm CPU}(n_{\rm f} + g n_{\rm g})$, but with smaller values of $g = g_{\rm CM} \approx 1/2 < g_{\rm MD} = 2$ and $C_{\rm CPU} \approx 5 \times 10^{-6}$ sec per time step (the total number of steps used in CM simulations within each run was equal to 50 000). Therefore, as was predicted earlier in this section, the actual CM efficiency of a decomposition algorithm is a factor of $[(n_{\rm f} + g_{\rm MD} n_{\rm g})/(n_{\rm f} + g_{\rm CM} n_{\rm g})]^K$ better with respect to the efficiency of that algorithm within MD simulations.

6. Conclusion

In this work we have presented a complete classification and consequent derivation of all possible self-adjoint decomposition algorithms with up to 11 stages and orders 2 to 6. As a result, it has been established that the well-known integrators by Verlet, Forest and Ruth, Chin, and Suzuki appear to be particular members of the same 11-stage decomposition family. Schemes with higher numbers of stages, which extend the Yoshida, Li, and Chin integrators, have also been studied up to order 16. Explicitly analyzing the structure of truncation uncertainties for each algorithm introduced and estimating the computational efforts needed, we have come to an important conclusion. Namely, the previously known integrators just mentioned, which are used in the majority of molecular and stochastic dynamics simulations of classical and quantum systems, are far to be perfect. There are much more efficient decomposition algorithms for each given order of precision in the time step. The most outstanding algorithms have been identified by us and tested in actual molecular dynamics and celestial mechanics simulations. It has been demonstrated that with the help of these new algorithms the efficiency of the integration can be increased by several times to several orders depending on the required level of precision and the system under consideration.

The decomposition algorithms exactly reproduce such relevant properties of conservative systems as the symplectic map and time reversibility of trajectories in phase space. Moreover, they, being used for the integration of motion in dissipative systems, are able to recover the principle of detailed balance. The satisfaction of these properties guarantees a high stability of the generated solutions and provides also the fulfilling of ergodicity for ergodic systems. In addition, the decomposition algorithms are cheap in implementation because they require performing only simple shifts of position, velocity, and time during current stages of the decomposed propagation. The chief advantage of our approach is that it has allowed us to obtain the most optimal algorithms within each order, stage, and class (non-gradient or force-gradient) considered. The optimum is ensured by the fact that these algorithms (being built within the most general description at the presence of free parameters) minimize (by appropriately adjusting these free parameters) the numerical uncertainties to the lowest possible level at given computational costs. Such algorithms should be treated as the best for the integration of motion in systems admitting the separation of the Liouville operator into two analytically solvable parts.

The presented approach is also directly adaptable to a more complicated case, when the number of analytically integrable parts of the Liouville operator may be greater than two. Examples are systems with spin or orientational degrees of freedom, collections of particles at the presence of virtual velocity-dependent forces describing various extended statistical ensembles, etc. For such systems, the single two-part separation can be iteratively repeated the necessary number of times. Since these repetitions can be performed in many different ways, an additional optimization should be employed to find the best solutions. An open and interesting question remains on the possibility to construct decomposition algorithms of order six and higher with only positive basic time coefficients. In the case of order four, such a positiveness (being unpermitted within the non-gradient class) has been achieved (see Section 3.3.3) within force-gradient decompositions. For higher orders, it seems that some higher-order gradient operators should be explicitly included into the decomposed exponential propagators. But the question arises whether such propagators will be analytically integrable and symplectic or not. These and other related problems could be the subject of a future investigation.

Acknowledgement

This work was supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung under Project No. 15247. The calculations have been performed on the Silicon Graphics Origin 3800 workstation of the University of Linz.

Appendix A. Recursive relations for ζ-components

The recursive relations for the higher-order multipliers ζ_{1-10} (see Eqs. (6), (10), and (13)) corresponding to the first type of self-adjoint transformations given by the first line of Eq. (14) are:

$$\begin{aligned}
\zeta_1^{(n+1)} &= \zeta_1^{(n)} + a^{(n)}\big(630\beta^{(n)^2} + 1260\gamma_4^{(n)}\sigma^{(n)} - 63\beta^{(n)}\big(6a^{(n)} + \nu^{(n)}\big)\sigma^{(n)^2} \\
&\quad + \sigma^{(n)^3}\big(21\alpha^{(n)} + \big(27a^{(n)^2} + 9a^{(n)}\nu^{(n)} + \nu^{(n)^2}\big)\sigma^{(n)}\big)\big)/3780, \\
\zeta_2^{(n+1)} &= \zeta_2^{(n)} + a^{(n)}\big(336\beta^{(n)}\big(6a^{(n)} + \nu^{(n)}\big)\sigma^{(n)^2} - 5040\beta^{(n)^2} - 5040\gamma_4^{(n)}\sigma^{(n)} \\
&\quad - \sigma^{(n)^3}\big(336\alpha^{(n)} + \big(120a^{(n)^2} + 12a^{(n)}\nu^{(n)} - \nu^{(n)^2}\big)\sigma^{(n)}\big)\big)/45360, \\
\zeta_3^{(n+1)} &= \zeta_3^{(n)} - a^{(n)}\big(5040\alpha^{(n)}\beta^{(n)} + \sigma^{(n)}\big(5040\gamma_2^{(n)} - 84\beta^{(n)}\nu^{(n)^2} + 72a^{(n)^3}\sigma^{(n)^2} + \nu^{(n)^3}\sigma^{(n)^2} \\
&\quad + 24a^{(n)}\nu^{(n)}\big(\nu^{(n)}\sigma^{(n)^2} - 42\beta^{(n)}\big) + a^{(n)^2}\big(88\nu^{(n)}\sigma^{(n)^2} - 672\beta^{(n)}\big)\big)\big)/15120, \\
\zeta_4^{(n+1)} &= \zeta_4^{(n)} + a^{(n)}\big(168\alpha^{(n)}\big(60\beta^{(n)} - \big(6a^{(n)} + \nu^{(n)}\big)\sigma^{(n)^2}\big) + \sigma^{(n)}\big(10080\gamma_2^{(n)} + 5040\gamma_3^{(n)} - 168\beta^{(n)}\nu^{(n)^2} \\
&\quad + 192a^{(n)^3}\sigma^{(n)^2} + 5\nu^{(n)^3}\sigma^{(n)^2} + 6a^{(n)}\nu^{(n)}\big(13\nu^{(n)}\sigma^{(n)^2} - 336\beta^{(n)}\big) \\
&\quad + a^{(n)^2}\big(272\nu^{(n)}\sigma^{(n)^2} - 1344\beta^{(n)}\big)\big)\big)/15120, \\
\zeta_5^{(n+1)} &= \zeta_5^{(n)} - a^{(n)}\big(2520\gamma_4^{(n)}\nu^{(n)} + 7560\gamma_3^{(n)}\sigma^{(n)} - 294\beta^{(n)}\nu^{(n)^2}\sigma^{(n)} + 180a^{(n)^3}\sigma^{(n)^3} - \nu^{(n)^3}\sigma^{(n)^3} \\
&\quad + 84\alpha^{(n)}\big(120\beta^{(n)} + \big(3\nu^{(n)} - 22a^{(n)}\big)\sigma^{(n)^2}\big) + a^{(n)^2}\big(234\nu^{(n)}\sigma^{(n)^3} - 1512\beta^{(n)}\sigma^{(n)}\big) \\
&\quad + 6a^{(n)}\big(420\gamma_4^{(n)} - 308\beta^{(n)}\nu^{(n)}\sigma^{(n)} + 3\nu^{(n)^2}\sigma^{(n)^3}\big)\big)/45360,
\end{aligned} \qquad \text{(A.1)}$$

$$
\begin{aligned}
\zeta_6^{(n+1)} = {} & \zeta_6^{(n)} + a^{(n)}(18a^{(n)3}\sigma^{(n)3} - 84\alpha^{(n)}(15\beta^{(n)} - (a^{(n)} + \nu^{(n)})\sigma^{(n)2}) \\
& + a^{(n)2}(15\nu^{(n)}\sigma^{(n)3} - 252\beta^{(n)}\sigma^{(n)}) + 6a^{(n)}(210\gamma_4^{(n)} - 28\beta^{(n)}\nu^{(n)}\sigma^{(n)} + \nu^{(n)2}\sigma^{(n)3}) \\
& + 2(630\gamma_4^{(n)}\nu^{(n)} - 630\gamma_2^{(n)}\sigma^{(n)} - 42\beta^{(n)}\nu^{(n)2}\sigma^{(n)} + \nu^{(n)3}\sigma^{(n)3}))/7560, \\
\zeta_7^{(n+1)} = {} & \zeta_7^{(n)} + a^{(n)}(2520\alpha^{(n)2} - 84\alpha^{(n)}(8a^{(n)2} + 12a^{(n)}\nu^{(n)} + \nu^{(n)2})\sigma^{(n)} + \sigma^{(n)}(5040\gamma_1^{(n)} \\
& + (48a^{(n)4} + 120a^{(n)3}\nu^{(n)} + 92a^{(n)2}\nu^{(n)2} + 18a^{(n)}\nu^{(n)3} + \nu^{(n)4})\sigma^{(n)}))/15120, \\
\zeta_8^{(n+1)} = {} & \zeta_8^{(n)} - a^{(n)}(5040\alpha^{(n)2} + 2520\gamma_2^{(n)}\nu^{(n)} - 42\beta^{(n)}\nu^{(n)3} + 2520\gamma_1^{(n)}\sigma^{(n)} \\
& - 420\alpha^{(n)}a^{(n)}(a^{(n)} + 2\nu^{(n)})\sigma^{(n)} + 69a^{(n)4}\sigma^{(n)2} + \nu^{(n)4}\sigma^{(n)2} \\
& + 2a^{(n)2}\nu^{(n)}(53\nu^{(n)}\sigma^{(n)2} - 294\beta^{(n)}) + a^{(n)3}(148\nu^{(n)}\sigma^{(n)2} - 294\beta^{(n)}) \\
& + 6a^{(n)}(420\gamma_2^{(n)} - 56\beta^{(n)}\nu^{(n)2} + 3\nu^{(n)3}\sigma^{(n)2}))/15120, \\
\zeta_9^{(n+1)} = {} & \zeta_9^{(n)} + a^{(n)}(2520\alpha^{(n)2} - 42\alpha^{(n)}(8a^{(n)2} + 12a^{(n)}\nu^{(n)} + \nu^{(n)2})\sigma^{(n)} + 114a^{(n)4}\sigma^{(n)2} \\
& - 4a^{(n)3}(147\beta^{(n)} - 59\nu^{(n)}\sigma^{(n)2}) + a^{(n)2}\nu^{(n)}(173\nu^{(n)}\sigma^{(n)2} - 1176\beta^{(n)}) \\
& + 24a^{(n)}(210\gamma_2^{(n)} + 105\gamma_3^{(n)} - 28\beta^{(n)}\nu^{(n)2} + 2\nu^{(n)3}\sigma^{(n)2}) \\
& + \nu^{(n)}(5040\gamma_2^{(n)} + 2520\gamma_3^{(n)} - 84\beta^{(n)}\nu^{(n)2} + 5\nu^{(n)3}\sigma^{(n)2}))/15120, \\
\zeta_{10}^{(n+1)} = {} & \zeta_{10}^{(n)} + a^{(n)}(a^{(n)} + \nu^{(n)})(2520\gamma_1^{(n)} - 42\alpha^{(n)}(7a^{(n)2} + 7a^{(n)}\nu^{(n)} + \nu^{(n)2}) \\
& + (31a^{(n)4} + 62a^{(n)3}\nu^{(n)} + 42a^{(n)2}\nu^{(n)2} + 11a^{(n)}\nu^{(n)3} + \nu^{(n)4})\sigma^{(n)})/15120.
\end{aligned}
$$

For the transformation of the second type (see the second line of Eq. (14)) we have obtained:

$$
\begin{aligned}
\zeta_1^{(n+1)} = {} & \zeta_1^{(n)} - (18b^{(n)4}\nu^{(n)3} + 15b^{(n)3}\nu^{(n)3}\sigma^{(n)} + 42c^{(n)}\nu^{(n)}(30\beta^{(n)} + 30c^{(n)} - \nu^{(n)}\sigma^{(n)2}) \\
& - 84\alpha^{(n)}(15\beta^{(n)}b^{(n)} + 30b^{(n)}c^{(n)} - 3b^{(n)3}\nu^{(n)} + 15c^{(n)}\sigma^{(n)} - 2b^{(n)2}\nu^{(n)}\sigma^{(n)} - b^{(n)}\nu^{(n)}\sigma^{(n)2}) \\
& - 6b^{(n)2}(210\gamma_2^{(n)} + \nu^{(n)2}(14\beta^{(n)} + 63c^{(n)} - \nu^{(n)}\sigma^{(n)2})) + b^{(n)}(1260\gamma_4^{(n)}\nu^{(n)} \\
& - 2\sigma^{(n)}(630\gamma_2^{(n)} + \nu^{(n)2}(42\beta^{(n)} + 84c^{(n)} - \nu^{(n)}\sigma^{(n)2}))))/7560, \\
\zeta_2^{(n+1)} = {} & \zeta_2^{(n)} + (12b^{(n)4}\nu^{(n)3} - 39b^{(n)3}\nu^{(n)3}\sigma^{(n)} + 42c^{(n)}\nu^{(n)}(120\beta^{(n)} + 120c^{(n)} - \nu^{(n)}\sigma^{(n)2}) \\
& - 252\alpha^{(n)}(20\beta^{(n)}b^{(n)} + 40b^{(n)}c^{(n)} - 3b^{(n)3}\nu^{(n)} + 20c^{(n)}\sigma^{(n)} - 2b^{(n)2}\nu^{(n)}\sigma^{(n)} - b^{(n)}\nu^{(n)}\sigma^{(n)2}) \\
& + 24b^{(n)2}(315\gamma_3^{(n)} - \nu^{(n)2}(21\beta^{(n)} + 42c^{(n)} + \nu^{(n)}\sigma^{(n)2})) + b^{(n)}(2520\gamma_4^{(n)}\nu^{(n)} \\
& + \sigma^{(n)}(7560\gamma_3^{(n)} - \nu^{(n)2}(294\beta^{(n)} + 168c^{(n)} + \nu^{(n)}\sigma^{(n)2}))))/45360, \\
\zeta_3^{(n+1)} = {} & \zeta_3^{(n)} - (2520\alpha^{(n)2}b^{(n)} + 57b^{(n)3}\nu^{(n)4} - 840\alpha^{(n)}\nu^{(n)}(3c^{(n)} - b^{(n)2}\nu^{(n)}) + 42c^{(n)}\nu^{(n)3}\sigma^{(n)} \\
& - 12b^{(n)2}(210\gamma_1^{(n)} - \nu^{(n)4}\sigma^{(n)}) - b^{(n)}(2520\gamma_2^{(n)}\nu^{(n)} - 42\beta^{(n)}\nu^{(n)3} + 336c^{(n)}\nu^{(n)3} \\
& + 2520\gamma_1^{(n)}\sigma^{(n)} + \nu^{(n)4}\sigma^{(n)2}))/15120, \\
\zeta_4^{(n+1)} = {} & \zeta_4^{(n)} + (5040\alpha^{(n)2}b^{(n)} - 42\alpha^{(n)}\nu^{(n)}(120c^{(n)} - b^{(n)}\nu^{(n)}(36b^{(n)} + \sigma^{(n)})) \\
& + \nu^{(n)}(96b^{(n)3}\nu^{(n)3} + 84c^{(n)}\nu^{(n)2}\sigma^{(n)} + 18b^{(n)2}\nu^{(n)3}\sigma^{(n)} - b^{(n)}(5040\gamma_2^{(n)} + 2520\gamma_3^{(n)} \\
& - \nu^{(n)2}(84\beta^{(n)} - 672c^{(n)} - 5\nu^{(n)}\sigma^{(n)2}))))/15120, \qquad \text{(A.2)}
\end{aligned}
$$

$$
\begin{aligned}
\zeta_5^{(n+1)} &= \zeta_5^{(n)} - \bigl(2520\alpha^{(n)^2} b^{(n)} - 36 b^{(n)^3} \nu^{(n)^4} + 42 c^{(n)} \nu^{(n)^3} \sigma^{(n)} + 30 b^{(n)^2} \nu^{(n)^4} \sigma^{(n)} + 168\alpha^{(n)} \nu^{(n)} \bigl(15 c^{(n)} \\
&\quad - b^{(n)} \nu^{(n)} \bigl(6 b^{(n)} + \sigma^{(n)}\bigr)\bigr) - b^{(n)} \bigl(15120\gamma_3^{(n)} \nu^{(n)} - \nu^{(n)^3} \bigl(252\beta^{(n)} + 504 c^{(n)} + \nu^{(n)} \sigma^{(n)^2}\bigr)\bigr)\bigr)/45360, \\
\zeta_6^{(n+1)} &= \zeta_6^{(n)} - \bigl(630\alpha^{(n)^2} b^{(n)} + 27 b^{(n)^3} \nu^{(n)^4} - 21 c^{(n)} \nu^{(n)^3} \sigma^{(n)} + 9 b^{(n)^2} \nu^{(n)^4} \sigma^{(n)} - 63\alpha^{(n)} \nu^{(n)} \bigl(20 c^{(n)} \\
&\quad - b^{(n)} \nu^{(n)} \bigl(6 b^{(n)} + \sigma^{(n)}\bigr)\bigr) - b^{(n)} \bigl(1260\gamma_2^{(n)} \nu^{(n)} + \nu^{(n)^3} \bigl(21\beta^{(n)} + 252 c^{(n)} - \nu^{(n)} \sigma^{(n)^2}\bigr)\bigr)\bigr)/3780, \\
\zeta_7^{(n+1)} &= \zeta_7^{(n)} - b^{(n)} \nu^{(n)} \bigl(2520\gamma_1^{(n)} - 42\alpha^{(n)} \nu^{(n)^2} - \nu^{(n)^4} \bigl(6 b^{(n)} - \sigma^{(n)}\bigr)\bigr)/15120, \\
\zeta_8^{(n+1)} &= \zeta_8^{(n)} + \bigl(5040 b^{(n)} \gamma_1^{(n)} \nu^{(n)} - 42 c^{(n)} \nu^{(n)^4} - 6 b^{(n)^2} \nu^{(n)^5} + b^{(n)} \nu^{(n)^5} \sigma^{(n)}\bigr)/15120, \\
\zeta_9^{(n+1)} &= \zeta_9^{(n)} - \nu^{(n)^3} \bigl(84\alpha^{(n)} b^{(n)} - \nu^{(n)} \bigl(84 c^{(n)} - b^{(n)} \nu^{(n)} \bigl(12 b^{(n)} + 5\sigma^{(n)}\bigr)\bigr)\bigr)/15120, \\
\zeta_{10}^{(n+1)} &= \zeta_{10}^{(n)} - b^{(n)} \nu^{(n)^6}/15120.
\end{aligned}
$$

Appendix B. Time coefficients for 23-stage gradient algorithms of order 8

In the case of velocity-like integration (86), the optimal values for the time coefficients are:

$$
\begin{aligned}
&a_1 = 0, \quad b_1 = b_{12} = 0.1839699354244402\text{E}{+}00, \quad c_1 = c_{12} = 0, \\
&a_2 = a_{12} = 0.6922517172738832\text{E}{+}00, \quad b_2 = b_{11} = 0.7084389757230299\text{E}{+}00, \\
&\quad c_2 = c_{11} = 0.3976209968238716\text{E}{-}01, \\
&a_3 = a_{11} = -0.3183450347119991\text{E}{+}00, \quad b_3 = b_{10} = 0.1981440445033534\text{E}{+}00, \\
&\quad c_3 = c_{10} = 0.2245403440322733\text{E}{-}01, \\
&a_4 = a_{10} = 0.6766724088765565\text{E}{+}00, \quad b_4 = b_9 = -0.6409380745116974\text{E}{-}01, \\
&\quad c_4 = c_9 = 0.9405266232181224\text{E}{-}03, \\
&a_5 = a_9 = -0.7207972470858706\text{E}{+}00, \quad b_5 = b_8 = -0.6887429532761409\text{E}{+}00, \\
&\quad c_5 = c_8 = -0.7336500519635302\text{E}{-}01, \\
&a_6 = a_8 = 0.3580316862350045\text{E}{+}00, \quad b_6 = b_7 = 0.1622838050764871\text{E}{+}00, \\
&\quad c_6 = c_7 = 0.2225664796363730\text{E}{-}01, \\
&a_7 = -0.3756270611751488\text{E}{+}00.
\end{aligned}
$$

For position-like counterpart (87), the optimal time coefficients read

$$
\begin{aligned}
&b_{12} = 0, \\
&c_{12} = 0, \\
&a_1 = a_{12} = 0.41009674738801111928784693005080\text{E}{+}00, \\
&b_1 = b_{11} = 0.48249309817414952912695842664785\text{E}{-}02, \\
&c_1 = c_{11} = 0.14743936907797528364717244760736\text{E}{-}03, \\
&a_2 = a_{11} = -0.34123345756052780489101697378499\text{E}{+}00, \\
&b_2 = b_{10} = 0.17492394861090375603419001374207\text{E}{+}00, \\
&c_2 = c_{10} = 0.23288450531932545357194967600155\text{E}{-}03, \\
&a_3 = a_{10} = 0.25644714021068150492361761631743\text{E}{+}00,
\end{aligned}
$$

$$
\begin{aligned}
&b_3 = b_9 = 0.29304366370957066164364546204288\mathrm{E}{+}00,\\
&c_3 = c_9 = 0.61648659635535962497705619884752\mathrm{E}{-}02,\\
&a_4 = a_9 = 0.27765273975812438394100476242641\mathrm{E}{+}00,\\
&b_4 = b_8 = 0.47448940168459770284238136482511\mathrm{E}{-}01,\\
&c_4 = c_8 = -0.12307516860831240716732016960034\mathrm{E}{-}01,\\
&a_5 = a_8 = -0.56926266869753773902939657321159\mathrm{E}{+}00,\\
&b_5 = b_7 = -0.15299863411743974499219652320477\mathrm{E}{-}02,\\
&c_5 = c_7 = -0.73296648559126385387017161643798\mathrm{E}{-}04,\\
&a_6 = a_7 = 0.46629949890124853576794423820194\mathrm{E}{+}00,\\
&b_6 = -0.37422994259002571606842462603791\mathrm{E}{-}01,\\
&c_6 = 0.15295860994523744731993293847001\mathrm{E}{-}01.
\end{aligned}
$$

The number of force evaluations per times step is $n_{\mathrm{f}} = P - 1 = 11$ for both these schemes, whereas the numbers of force-gradient recalculations are $n_{\mathrm{g}} = 10$ (since $c_1 = 0$, and thus two boundary letters C in formula (86) should actually be replaced by B) and $n_{\mathrm{g}} = 11$, respectively.

Appendix C. Composition constants of the 16th-order algorithm, base 8

For $Q = 16$ and $K = 8$, acting in the spirit of the advanced theory [50] of composed integration (88), one comes to $P = 11$ order conditions. These conditions can be satisfied using the same number of independent composition constants d_p ($p = 1, 2, \ldots, P$). They lead to $2P - 1 = 21$ eighth-order elements S_K in the resulting sixteenth-order algorithm S_Q. Handling the order conditions, we have found more than two hundreds real solutions for $\{d_p\}$. The optimal set, which minimizes $\delta = \max_{p=1}^{P}\{|d_p|\}$, is

$$
\begin{aligned}
&d_1 = 0.29642254891413070953312450213071,\\
&d_2 = 0.55268563185301488324882994018746,\\
&d_3 = -0.58134339535553339331560554430994 0,\\
&d_4 = 0.23403665265420481243563202333267,\\
&d_5 = -0.51788958989817055303978658827453,\\
&d_6 = -0.43983975477992920522811970527874,\\
&d_7 = -0.20137078150942169957468111993444,\\
&d_8 = 0.34412872002528894622975927197416,\\
&d_9 = 0.03072591760996558798895428309765,\\
&d_{10} = 0.48652953960727041281280535031455,
\end{aligned}
$$

with $d_{11} = 1 - 2\sum_{p=1}^{10} d_p$ and $\delta_{\min} \equiv |d_{11}| = 0.592$.

References

[1] H. Yoshida, Phys. Lett. A 150 (1990) 262.
[2] E. Forest, R.D. Ruth, Physica D 43 (1990) 105.

[3] M. Suzuki, Phys. Lett. A 165 (1992) 387.
[4] E. Forest, J. Comput. Phys. 99 (1992) 209.
[5] M. Suzuki, K. Umeno, in: D.P. Landau, K.K. Mon, H.-B. Schüttler (Eds.), Computer Simulation Studies in Condensed Matter Physics VI, Springer, Berlin, 1993.
[6] P.-V. Koseleff, Lecture Notes in Comput. Sci. 673 (1993) 213.
[7] J.M. Sanz-Serna, M.P. Calvo, Numerical Hamiltonian Problems, Chapman & Hall, London, 1994.
[8] M. Suzuki, in: D.P. Landau, K.K. Mon, H.-B. Schüttler (Eds.), Computer Simulation Studies in Condensed Matter Physics VIII, Springer-Verlag, Berlin, 1995.
[9] R.-C. Li, PhD thesis, Department of Mathematics, University of California at Berkeley, CA, 1995.
[10] M. Suzuki, Phys. Lett. A 201 (1995) 425.
[11] P.-V. Koseleff, Fields Inst. Commun. 10 (1996) 103.
[12] S.A. Chin, Phys. Lett. A 226 (1997) 344.
[13] A. Dullweber, B. Leimkuhler, R. McLachlan, J. Chem. Phys. 107 (1997) 5840.
[14] M. Krech, A. Bunker, D.P. Landau, Comput. Phys. Comm. 111 (1998) 1.
[15] I.P. Omelyan, I.M. Mryglod, R. Folk, Phys. Rev. Lett. 86 (2001) 898.
[16] S. Blanes, Phys. Rev. E 65 (2002) 056703.
[17] D. Frenkel, B. Smit, Understanding Molecular Simulation: from Algorithms to Applications, Academic Press, New York, 1996.
[18] S. Reich, SIAM J. Numer. Anal. 36 (1999) 1549.
[19] C.W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Engelwood Cliffs, NJ, 1971.
[20] R.L. Burden, J.D. Faires, Numerical Analysis, 5th edn., PWS Publishing, Boston, 1993.
[21] M.P. Allen, D.J. Tildesley, Computer Simulation of Liquids, Clarendon, Oxford, 1987.
[22] I.P. Omelyan, Phys. Rev. E 58 (1998) 1169.
[23] H. Kinoshita, H. Yoshida, H. Nakai, Celest. Mech. 50 (1991) 59.
[24] B. Gladman, M. Duncan, J. Candy, Celest. Mech. 52 (1991) 221.
[25] J. Wisdom, M. Holman, Astrophys. J. 102 (1991) 1528.
[26] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, J. Chem. Phys. 76 (1982) 637.
[27] M. Tuckerman, B.J. Berne, G.J. Martyna, J. Chem. Phys. 97 (1992) 1990.
[28] M. Qin, W.J. Zhu, Computing 47 (1992) 309.
[29] R.I. McLachlan, SIAM J. Sci. Comput. 16 (1995) 151.
[30] W. Kahan, R.-C. Li, Math. Comput. 66 (1997) 1089.
[31] A. Murua, J.M. Sanz-Serna, Phil. Trans. Roy. Soc. A 357 (1999) 1079.
[32] S.A. Chin, D.W. Kidwell, Phys. Rev. E 62 (2000) 8746.
[33] A.N. Drozdov, J.J. Brey, Phys. Rev. E 57 (1998) 1284.
[34] H.A. Forbert, S.A. Chin, Phys. Rev. E 63 (2000) 016703.
[35] S.A. Chin, C.R. Chen, J. Chem. Phys. 114 (2001) 7338.
[36] H.A. Forbert, S.A. Chin, Phys. Rev. B 63 (2001) 144518.
[37] J. Auer, E. Krotscheck, S.A. Chin, J. Chem. Phys. 115 (2001) 6841.
[38] S.J. Stuart, R. Zhou, B.J. Berne, J. Chem. Phys. 105 (1996) 1426.
[39] R. McLachlan, BIT 35 (2) (1995) 258.
[40] J.E. Chambers, M.A. Murison, Astronom. J. 119 (2000) 425.
[41] J. Laskar, P. Robutel, Celest. Mech. Dyn. Astr. 80 (2001) 39.
[42] S. Blanes, Appl. Numer. Math. 37 (2001) 289.
[43] J.M. Sanz-Serna, A. Portillo, J. Chem. Phys. 104 (1996) 2349.
[44] S. Blanes, P.C. Moan, Phys. Lett. A 265 (2000) 35.
[45] S. Blanes, P.C. Moan, J. Comput. Phys. 170 (2001) 205.
[46] S.A. Chin, C.R. Chen, J. Chem. Phys. 117 (2002) 1409.
[47] R.I. McLachlan, G.R.W. Quispel, Acta Numer. 11 (2002) 341.
[48] I.P. Omelyan, I.M. Mryglod, R. Folk, Phys. Rev. E 65 (2002) 056706.
[49] I.P. Omelyan, I.M. Mryglod, R. Folk, Comput. Phys. Comm. 146 (2002) 188.
[50] I.P. Omelyan, I.M. Mryglod, R. Folk, Phys. Rev. E 66 (2002) 026701.
[51] M. Sofroniou, W. Oevel, SIAM J. Numer. Anal. 34 (1997) 2063.
[52] F.A. Bornemann, P. Nettesheim, Ch. Schütte, J. Chem. Phys. 105 (1996) 1074.
[53] M. Suzuki, Proc. Japan Acad. B 69 (1993) 161.
[54] M. Creutz, A. Gocksch, Phys. Rev. Lett. 63 (1989) 9.