Computational Physics

Hessian-free force-gradient integrators ☆

Kevin Schäfers [a], [iD], [*], Jacob Finkenrath [b,c], Michael Günther [a], Francesco Knechtli [b]

[a] *Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM), Chair of Applied and Computational Mathematics, Bergische Universität Wuppertal, Gaußstraße 20, Wuppertal, 42119, Germany*
[b] *Department of Physics, Bergische Universität Wuppertal, Gaußstraße 20, Wuppertal, 42119, Germany*
[c] *Department of Theoretical Physics, European Organization for Nuclear Research, CERN, Geneve, 1211, Switzerland*

A R T I C L E   I N F O

A B S T R A C T

We propose a new framework of Hessian-free force-gradient integrators that do not require the analytical expression of the force-gradient term based on the Hessian of the potential. Due to that the new class of decomposition algorithms for separable Hamiltonian systems with quadratic kinetic energy may be particularly useful when applied to Hamiltonian systems where an evaluation of the Hessian is significantly more expensive than an evaluation of its gradient, e.g. in molecular dynamics simulations of classical systems. Numerical experiments of an N-body problem, as well as applications to the molecular dynamics step in the Hybrid Monte Carlo (HMC) algorithm for lattice simulations of the Schwinger model and Quantum Chromodynamics (QCD) verify these expectations.

1. Introduction

The numerical integration of Hamiltonian systems imposes challenging demands on the numerical integration scheme. Particularly, the phase space (p, q) with generalized coordinates q and conjugate momenta p is a symplectic manifold, giving rise to the need of *geometric numerical integration* [1]. In this paper, we consider separable Hamiltonian systems $\mathcal{H}(p, q) = \mathcal{T}(p) + \mathcal{V}(q)$ with kinetic energy of the form

$$\mathcal{T}(p) = \tfrac{1}{2} p^\top M^{-1} p, \tag{1}$$

where M is a constant symmetric positive definite matrix. This particular structure has a wide range of applications in the field of many-body problems, e.g. classical mechanics [2], quantum mechanics [3], statistical mechanics [4] and lattice field theories [5]. In the latter case, the phase space is the cotangent bundle over a base space that is a Lie group manifold. Moreover, any system of second order ODEs $\ddot{y} = f(y)$ is reduced to the equations of motion under consideration in this work.

There exist many approaches for geometric numerical integration of Hamiltonian systems. Recently, symplectic generalized additive Runge–Kutta (GARK) schemes [6,7] have been developed. An alternative approach based on the variational principle has been discussed in [8]. Both approaches are restricted to Euclidean space. An extension of symplectic partitioned Runge–Kutta methods to the Lie group setting has been presented in [9]. However, the general framework does not include time-reversibility, another important feature of Hamiltonian systems. For separable Hamilto-

nian systems, splitting methods allow for the derivation of explicit geometric numerical integration schemes of arbitrarily high convergence via composition techniques [10–12].

By exploiting the special structure of the kinetic energy $\mathcal{T}(p)$, *force-gradient integrators* (FGIs) [13,14] allow for an efficient computational process. FGIs can be regarded as splitting methods, applied to a modified potential, resulting in more accurate numerical approximations with respect to the original system. One drawback of FGIs is the necessity of deriving the so-called *force-gradient term* (FG-term), containing the Hessian of the potential $\mathcal{V}(q)$. To our knowledge, the idea of using the Hessian of the potential $\mathcal{V}(q)$ to enhance splitting methods dates back to [15,16] and has been further investigated in [17]. Recently, force-gradient integrators have been successfully developed in the more general context of port-Hamiltonian (pH) systems to derive higher order schemes fulfilling the dissipation inequality of pH systems and thus breaking the order-two limit for splitting methods [18]. Another possible drawback is the evaluation cost of the FG-term. In molecular dynamics simulations of classical systems, for example, the evaluation typically is 2-3 times more expensive than a usual force evaluation [12]. One can overcome this issue by approximating the FG-term, as proposed in [19] and further investigated in [20] for a particular FGI. In the context of lattice quantum chromodynamics (QCD), the approximation of the FG-term was initially applied to another FGI in [21]. The adapted FGI from [21] has been applied to the two-dimensional Schwinger model where we implemented the FG-term analytically [22] and in large scale lattice QCD simulations, see e.g. [23–25]. Numerical results highlight that the Hessian-free variant results in a more efficient computational process.

In this paper we will generalize the idea of approximating the FG-term for the entire class of FGIs introduced in [13]. This will be applied not only in case of lattice field theories, but in general for all separable Hamiltonian systems with kinetic energy of the form (1), where the new class of *Hessian-free force-gradient integrators* can be utilized. We will discuss the geometric properties, a refined error analysis of the approximation, resulting in explicit formulae for the error terms, as well as a backward error analysis giving insight into the long-time behavior of Hessian-free FGIs.

The paper is organized as follows. In Section 2, we will briefly introduce Hamiltonian mechanics and its demands on the numerical integration scheme. Section 3 introduces the new class of Hessian-free force-gradient integrators. In Section 4, Hessian-free force-gradient integrators with up to eleven stages are derived. Here, (weighted) norms of the leading error coefficients are defined as aggregated functions of the principal error term. By performing a global minimization of this aggregated function, we obtain optimal sets for the integrator coefficients. In Section 5, numerical results for three different test examples are discussed: a) the outer solar system, a N-body problem with $\boldsymbol{M} \neq \mathrm{Id}$, b) the two-dimensional Schwinger model where we can compare the performance of exact FGIs and Hessian-free FGIs as the analytical expression of the FG-term is available [22], and c) four-dimensional gauge field simulations in lattice QCD with two heavy Wilson fermions. Here, the physical degrees of freedoms are elements of the non-Abelian matrix Lie group SU(3) and all tests are performed using the openQCD implementation [26,27]. The paper concludes with a summary and outlook for future research.

2. Hamiltonian mechanics and geometric numerical integration

In this section, we will briefly introduce Hamiltonian mechanics on matrix Lie groups. Moreover we will state the demands on the numerical integration scheme for separable Hamiltonian systems

$$\mathcal{H}(\boldsymbol{p},\boldsymbol{q}) = \mathcal{T}(\boldsymbol{p}) + \mathcal{V}(\boldsymbol{q}) = \tfrac{1}{2}\boldsymbol{p}^\top \boldsymbol{M}^{-1}\boldsymbol{p} + \mathcal{V}(\boldsymbol{q}) \tag{2}$$

with constant symmetric positive definite matrix $\boldsymbol{M}$. Without loss of generality,[1] we consider a separable Hamiltonian system of the form

$$\mathcal{H}(\boldsymbol{p},\boldsymbol{q}) = \mathcal{T}(\boldsymbol{p}) + \mathcal{V}(\boldsymbol{q}) = \tfrac{1}{2}\langle \boldsymbol{p}, \boldsymbol{p}\rangle + \mathcal{V}(\boldsymbol{q}). \tag{3}$$

2.1. Hamiltonian mechanics

We consider a phase space $(\boldsymbol{p}, \boldsymbol{q}) \in T^*\mathcal{G}$ where $T^*\mathcal{G}$ denotes the cotangent bundle over a base space that is a d-dimensional matrix Lie group manifold $\mathcal{G}$ and whose fibers are isomorphic to its Lie algebra $\mathfrak{g}$. The linear space $\mathfrak{g}$ has a basis, consisting of *generators* $\boldsymbol{T}_i$, $i = 1, \ldots, d$. For matrix Lie groups, there exists a matrix representation $\boldsymbol{Q}$ of the Lie group element $q \in \mathcal{G}$ and $\boldsymbol{P} := p^i \boldsymbol{T}_i \in \mathfrak{g}$ of the momentum p. The generators $\boldsymbol{T}_i$ are linked to the Lie group elements $\boldsymbol{Q}$ via the right-invariant linear differential operator e_i whose action on $\boldsymbol{Q}$ is defined by

$$e_i(\boldsymbol{Q}) = -\boldsymbol{T}_i \boldsymbol{Q}. \tag{4}$$

The operator (4) can be regarded as a generalization of the vector fields $\partial/\partial \boldsymbol{q}$ in the Lie group space. The cotangent space has a natural symplectic structure $\boldsymbol{\omega} = -\mathrm{d}\boldsymbol{p}$ that is closed, $\mathrm{d}\boldsymbol{\omega} = 0$. The fundamental two-form $\boldsymbol{\omega}$ defines the Poisson bracket of two arbitrary Hamiltonian vector fields $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}$ corresponding to zero-forms $\boldsymbol{A}, \boldsymbol{B}$ as $\{\boldsymbol{A}, \boldsymbol{B}\} = -\boldsymbol{\omega}(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}})$. Hamilton's equations are most naturally expressed in terms of the Lie derivative operators

$$\hat{\mathcal{V}} := \{\mathcal{V}, \cdot\} \quad \text{and} \quad \hat{\mathcal{T}} := \{\mathcal{T}, \cdot\},$$

resulting in the additively partitioned system of ordinary differential equations (ODEs)

$$\begin{pmatrix} \dot{\boldsymbol{P}} \\ \dot{\boldsymbol{Q}} \end{pmatrix} = \begin{pmatrix} \{\mathcal{V}, \boldsymbol{P}\} \\ \{\mathcal{V}, \boldsymbol{Q}\} \end{pmatrix} + \begin{pmatrix} \{\mathcal{T}, \boldsymbol{P}\} \\ \{\mathcal{T}, \boldsymbol{Q}\} \end{pmatrix} = \begin{pmatrix} \{\mathcal{V}, \boldsymbol{P}\} \\ \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0} \\ \{\mathcal{T}, \boldsymbol{Q}\} \end{pmatrix} = \begin{pmatrix} \hat{\mathcal{V}}\boldsymbol{P} \\ \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} \boldsymbol{0} \\ \hat{\mathcal{T}}\boldsymbol{Q} \end{pmatrix}.$$

For matrix Lie groups, there exist structure constants c_{jk}^i satisfying $[\boldsymbol{T}_j, \boldsymbol{T}_k] = c_{jk}^i \boldsymbol{T}_i$. They occur in the Lie derivative operator

$$\hat{\mathcal{T}} = p^i e_i + c_{ki}^j p_j p^k \frac{\partial}{\partial p_i} \tag{5}$$

---

[1] With the Cholesky decomposition $\boldsymbol{M} := \boldsymbol{L}^\top \boldsymbol{L}$ and transformed momenta $\tilde{p} := \boldsymbol{L}p$, any system (2) can be written as a system of the form (3).

of the kinetic part. For Abelian Lie groups (e.g. U(1) and SO(2)), it holds $c^i_{jk} = 0$ for all i, j, k. In case of semisimple[2] Lie groups [28] like SU(n) ($n \geq 2$), SO(n) ($n \geq 3$), SL(n) and Sp(n) ($n \geq 1$), it holds the total antisymmetry of the structure constants $c^j_{ki} p_j p^k = 0$. Consequently, for this wide range of matrix Lie groups, the Lie derivative operator (5) of the kinetic part simplifies so that explicit forms of the Lie derivative operators read

$$\hat{\mathcal{V}} = -e_i(\mathcal{V})\frac{\partial}{\partial p_i} \quad \text{and} \quad \hat{\mathcal{T}} = p^i e_i, \tag{6}$$

so that Hamilton's equations become

$$\dot{\boldsymbol{P}} = -e_i(\mathcal{V})\frac{\partial \boldsymbol{P}}{\partial p_i} = -e_i(\mathcal{V})\boldsymbol{T}^i, \qquad \boldsymbol{P}(0) = \boldsymbol{P}_0,$$

$$\dot{\boldsymbol{Q}} = p^i e_i(\boldsymbol{Q}) = -p^i \boldsymbol{T}_i \boldsymbol{Q} = -\boldsymbol{P}\boldsymbol{Q}, \quad \boldsymbol{Q}(0) = \boldsymbol{Q}_0. \tag{7}$$

The formal solution of the subsystems

$$\begin{pmatrix} \dot{\boldsymbol{P}} \\ \dot{\boldsymbol{Q}} \end{pmatrix} = \begin{pmatrix} \hat{\mathcal{V}}\boldsymbol{P} \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} -e_i(\mathcal{V})\boldsymbol{T}^i \\ \boldsymbol{0} \end{pmatrix}, \qquad \begin{pmatrix} \dot{\boldsymbol{P}} \\ \dot{\boldsymbol{Q}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \hat{\mathcal{T}}\boldsymbol{Q} \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ -\boldsymbol{P}\boldsymbol{Q} \end{pmatrix},$$

can be expressed in terms of the matrix exponential $\exp(t\mathcal{A}) = \sum_{k=0}^{\infty}(t\mathcal{A})^k/k!$ and the Lie derivative operators via

$$\varphi_t^{\{1\}}(\boldsymbol{P}_0, \boldsymbol{Q}_0) = \begin{pmatrix} \exp(t\hat{\mathcal{V}})\boldsymbol{P}_0 \\ \boldsymbol{Q}_0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_0 - te_i(\mathcal{V}(\boldsymbol{Q}_0))\boldsymbol{T}^i \\ \boldsymbol{Q}_0 \end{pmatrix}, \qquad \varphi_t^{\{2\}}(\boldsymbol{P}_0, \boldsymbol{Q}_0) = \begin{pmatrix} \boldsymbol{P}_0 \\ \exp(t\hat{\mathcal{T}})\boldsymbol{Q}_0 \end{pmatrix} = \begin{pmatrix} \boldsymbol{P}_0 \\ \exp(-t\boldsymbol{P}_0)\boldsymbol{Q}_0 \end{pmatrix},$$

for the potential and the kinetic part, respectively.

Remark 1 (*System updates*). When having many components $(\boldsymbol{p}_\ell, \boldsymbol{q}_\ell) \in T^*\mathcal{G}$, $\ell = 1, \ldots, L$, one can trivially generalize the fundamental two-form, resulting in sums over all components $\boldsymbol{\omega} = -\sum_{\ell=1}^{L} \mathrm{d}\boldsymbol{p}_\ell$. Analogously, the locally acting kinetic energy generalizes to $\mathcal{T}(\boldsymbol{p}) = \sum_{\ell=1}^{L} \frac{1}{2}\langle \boldsymbol{p}_\ell, \boldsymbol{p}_\ell \rangle$. Consequently, Hamilton's equations become

$$\dot{\boldsymbol{P}}_\ell = -e_i(\mathcal{V})_\ell \boldsymbol{T}^i, \quad \boldsymbol{P}_\ell(0) = \boldsymbol{P}_{\ell,0},$$

$$\dot{\boldsymbol{Q}}_\ell = -\boldsymbol{P}_\ell \boldsymbol{Q}_\ell, \qquad \boldsymbol{Q}_\ell(0) = \boldsymbol{Q}_{\ell,0}, \, \ell = 1, \ldots, L.$$

Remark 2 (*Classical mechanics*). In classical Hamiltonian mechanics, the phase space is $(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^d \times \mathbb{R}^d$. With the fundamental two-form $\boldsymbol{\omega} = \mathrm{d}\boldsymbol{q} \wedge \mathrm{d}\boldsymbol{p}$, one obtains the well-known Hamiltonian equations of motion

$$\begin{pmatrix} \dot{\boldsymbol{p}} \\ \dot{\boldsymbol{q}} \end{pmatrix} = \begin{pmatrix} -\frac{\partial \mathcal{H}}{\partial \boldsymbol{q}} \\ \frac{\partial \mathcal{H}}{\partial \boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} -\mathcal{V}_q(\boldsymbol{q}) \\ \mathcal{T}_p(\boldsymbol{p}) \end{pmatrix}, \quad \begin{pmatrix} \boldsymbol{p}(0) \\ \boldsymbol{q}(0) \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_0 \\ \boldsymbol{q}_0 \end{pmatrix}.$$

For the sake of clarity and coherence, we will keep the notation (7) for Hamiltonian mechanics on matrix Lie groups. However, it is important to emphasize that the results in this paper also hold for system updates, as well as for classical mechanics in Euclidean space.

2.2. Geometric integration

The Hamiltonian flow $\varphi_t(\boldsymbol{p}_0, \boldsymbol{q}_0)$ is characterized by the following properties.

- **Energy conservation:** The Hamiltonian is an invariant of the flow, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(\varphi_t(\boldsymbol{p}_0, \boldsymbol{q}_0)) = 0.$$

- **Time-reversibility:** The Hamiltonian flow is time-reversible, i.e.,

$$\rho \circ \varphi_t \circ \rho \circ \varphi_t(\boldsymbol{p}_0, \boldsymbol{q}_0) = (\boldsymbol{p}_0, \boldsymbol{q}_0), \quad \rho(\boldsymbol{p}, \boldsymbol{q}) = (-\boldsymbol{p}, \boldsymbol{q}).$$

- **Symplecticity:** The Hamiltonian flow is symplectic, i.e., $\mathrm{d}\boldsymbol{\omega} = 0$. A direct consequence of the symplecticity of the Hamiltonian flow is the preservation of volume,

$$\left| \det \frac{\partial \varphi_t(\boldsymbol{p}_0, \boldsymbol{q}_0)}{\partial(\boldsymbol{p}_0, \boldsymbol{q}_0)} \right| = 1.$$

- **Closure property:** The Hamiltonian flow satisfies $\varphi_t(\boldsymbol{p}_0, \boldsymbol{q}_0) \in T^*\mathcal{G}$ for all $t > 0$ provided that $(\boldsymbol{p}_0, \boldsymbol{q}_0) \in T^*\mathcal{G}$.

We demand the numerical scheme $\Phi_h(\boldsymbol{p}_0, \boldsymbol{q}_0) = (\boldsymbol{p}_1, \boldsymbol{q}_1) \approx \varphi_h(\boldsymbol{p}_0, \boldsymbol{q}_0)$ to preserve the time-reversibility, volume-preservation, as well as the closure property of the Hamiltonian flow. For *time-reversibility*, one gets the criterion

$$\rho \circ \Phi_h \circ \rho \circ \Phi_h(\boldsymbol{p}_0, \boldsymbol{q}_0) = (\boldsymbol{p}_0, \boldsymbol{q}_0). \tag{8}$$

The numerical integration scheme is *volume-preserving* if it holds

---

[2] A semisimple Lie group is a non-Abelian Lie group whose Lie algebra is semisimple, i.e., it is a direct sum of simple Lie algebras.

$$\left| \det \frac{\partial \Phi_h(\boldsymbol{p}_0, \boldsymbol{q}_0)}{\partial(\boldsymbol{p}_0, \boldsymbol{q}_0)} \right| = 1. \tag{9}$$

The *closure property* demands

$$\Phi_h(\boldsymbol{p}_0, \boldsymbol{q}_0) = (\boldsymbol{p}_1, \boldsymbol{q}_1) \in T^*\mathcal{G}, \tag{10}$$

provided $(\boldsymbol{p}_0, \boldsymbol{q}_0) \in T^*\mathcal{G}$.

3. Hessian-free force-gradient integrators

An efficient approach for geometric numerical integration of Hamiltonian systems of the form (3) is given by force-gradient integrators (FGIs) [13,14]. In this section, we will introduce the new class of Hessian-free FGIs. Before introducing the adaption, we will start with a brief recapitulation of FGIs based on their introduction in [13].

3.1. Force-gradient integrators

As we have seen in the previous section, it is possible to compute the flows of the subsystems using the exponential map and the Lie derivative operators (6). Hence it is possible to compute an approximation to Hamilton's equations (7) by composing evaluations of these exact flows, resulting in splitting methods [29]

$$\Psi^{(0)} = \mathrm{Id}, \quad \Psi^{(j)} = \Psi^{(j-1)} \circ \varphi_{a_j h}^{\{2\}} \circ \varphi_{b_j h}^{\{1\}}, \quad j = 1, \dots, P.$$

Since we demand the numerical integration scheme to preserve the time-reversibility (8), it has to hold either $a_1 = 0$ (velocity version) or $b_P = 0$ (position version). The remaining composition of $2P - 1$ exponentials has to be symmetric, i.e., it has to hold $a_{j+1} = a_{P-j+1}$ and $b_j = b_{P-j+1}$ for the velocity version, $a_j = a_{P-j+1}$ and $b_j = b_{P-j}$ for the position version. Applying the Baker–Campbell–Hausdorff (BCH) formula, the overall integrator $\Psi_h := \Psi_h^{(P)}$ can be written as an exponential [13]

$$\exp\left((\nu\hat{\mathcal{T}} + \sigma\hat{\mathcal{V}})h + \mathcal{O}_3 h^3 + \mathcal{O}_5 h^5 + \mathcal{O}_7 h^7 + \mathcal{O}(h^9) \right) \tag{11a}$$

with $\nu := \sum_{j=1}^P a_j$, $\sigma := \sum_{j=1}^P b_j$, and

$$\mathcal{O}_3 = \alpha \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right] + \beta \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right], \tag{11b}$$

$$\mathcal{O}_5 = \gamma_1 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right] \right] \right] + \gamma_2 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right] \right] \right] + \gamma_3 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right] \right] \right] + \gamma_4 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \hat{\mathcal{V}} \right] \right] \right] \right], \tag{11c}$$

$$\begin{aligned}
\mathcal{O}_7 &= \zeta_1 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_2 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_3 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] \\
&\quad + \zeta_4 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_5 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_6 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] \\
&\quad + \zeta_7 \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_8 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] + \zeta_9 \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right] \\
&\quad + \zeta_{10} \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{T}}, \left[\hat{\mathcal{V}}, \hat{\mathcal{T}} \right] \right] \right] \right] \right] \right],
\end{aligned} \tag{11d}$$

where $[\cdot, \cdot]$ denotes the commutator $[\mathcal{A}, \mathcal{B}] := \mathcal{AB} - \mathcal{BA}$.[3] To obtain splitting methods of convergence order $p > 2$, one has to cancel the order-three error term $\mathcal{O}_3$, consisting of the two commutators $[\hat{\mathcal{T}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]$ and $[\hat{\mathcal{V}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]$. Without using negative time steps, that are used for example in composition techniques [10–12], it is not possible to cancel both commutators at once [11]. For Hamiltonian systems of the form (3), the second commutator $\hat{C} := [\hat{\mathcal{V}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]$ exhibits a special structure

$$\hat{C} = 2\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}} = 2e^j(\mathcal{V})e_j e_i(\mathcal{V}) \frac{\partial}{\partial p_i} \tag{12}$$

that only depends on the generalized coordinates q_j, and is called *force-gradient term* (FG-term). Extending the momentum updates $\varphi_{b_j h}^{\{1\}}$ by including evaluations of the force-gradient term, $\exp(b_j \hat{\mathcal{V}} + c_j \hat{C})$, results in a FGI. This allows to choose the coefficients a_j, b_j such that the first commutator $[\hat{\mathcal{T}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]$ vanishes and then the coefficients c_j can be chosen to remove the remaining order-3 error term $[\hat{\mathcal{V}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]$, resulting in higher-order integrators. A symmetric FGI consists of $2P - 1$ exponentials that are either momentum updates

$$\exp\left(b^{(n)} h\hat{\mathcal{V}} + c^{(n)} h^3 \hat{C} \right) \tag{13}$$

or position updates

$$\exp\left(a^{(n)} h\hat{\mathcal{T}} \right) \tag{14}$$

with coefficients $a^{(n)}, b^{(n)}, c^{(n)}$ that are related to the coefficients a_j, b_j, c_j via

---

[3] The error term $\mathcal{O}_5$ neglects two zero-valued commutators $[\hat{\mathcal{V}}, [\hat{\mathcal{V}}, [\hat{\mathcal{V}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]]]$ and $[\hat{\mathcal{T}}, [\hat{\mathcal{V}}, [\hat{\mathcal{V}}, [\hat{\mathcal{T}}, \hat{\mathcal{V}}]]]]$.

$$j = \begin{cases} \frac{P+2}{2} - n, & P \text{ even (velocity version)}, \\ \frac{P+1}{2} + n, & P \text{ odd (velocity version)}, \\ \frac{P}{2} + n, & P \text{ even (position version)}, \\ \frac{P+1}{2} - n, & P \text{ odd (position version)}, \end{cases}$$

for $n = 1, \ldots, \lfloor \frac{P}{2} \rfloor$. A FGI is constructed as follows. Starting from a central single-exponential operator

$$\Psi^{(0)} = \begin{cases} \exp(a_{(P+2)/2} h \hat{T}), & P \text{ even (velocity version)}, \\ \exp(b_{(P+1)/2} h \hat{V} + c_{(P+1)/2} h^3 \hat{C}), & P \text{ odd (velocity version)}, \\ \exp(b_{P/2} h \hat{V} + c_{P/2} h^3 \hat{C}), & P \text{ even (position version)}, \\ \exp(a_{(P+1)/2} h \hat{T}), & P \text{ odd (position version)}, \end{cases} \tag{15a}$$

the integrator is obtained by consecutively applying the following symmetric transformations

$$\Psi^{(n)} = \exp\left(b^{(n)} h \hat{V} + c^{(n)} h^3 \hat{C}\right) \circ \exp\left(a^{(n)} h \hat{T}\right) \circ \Psi^{(n-1)} \circ \exp\left(a^{(n)} h \hat{T}\right) \circ \exp\left(b^{(n)} h \hat{V} + c^{(n)} h^3 \hat{C}\right), \tag{15b}$$

for P odd (velocity version) and P even (position version), and

$$\Psi^{(n)} = \exp\left(a^{(n)} h \hat{T}\right) \circ \exp\left(b^{(n)} h \hat{V} + c^{(n)} h^3 \hat{C}\right) \circ \Psi^{(n-1)} \circ \exp\left(b^{(n)} h \hat{V} + c^{(n)} h^3 \hat{C}\right) \circ \exp\left(a^{(n)} h \hat{T}\right), \tag{15c}$$

for P even (velocity version) and P odd (position version), $n = 1, \ldots, \lfloor \frac{P}{2} \rfloor$. By applying the BCH formula, the overall FGI $\Psi_h := \Psi_h^{(\lfloor P/2 \rfloor)}$ again takes the form (11). Thanks to the recursive definition (15) of the integrator, the multipliers $\nu, \sigma, \alpha, \beta, \gamma_1, \ldots, \gamma_4$, and $\zeta_1, \ldots, \zeta_{10}$ can be determined using recursive formulations stated in [13].

Order conditions. The FGI (15) has convergence order p if its representation as an exponential (11) satisfies

$$\exp\left((\nu \hat{T} + \sigma \hat{V})h + \mathcal{O}_3 h^3 + \mathcal{O}_5 h^5 + \mathcal{O}_7 h^7 + \mathcal{O}(h^9)\right) = \exp\left(h(\hat{T} + \hat{V})\right) + \mathcal{O}(h^{p+1}).$$

Hence the order conditions up to order seven are given by the recursive formulations of the multipliers $\nu, \sigma, \alpha, \beta, \gamma_1, \ldots, \gamma_4$, and $\zeta_1, \ldots, \zeta_{10}$ in [13].

Geometric integration. The momentum updates (13) of the FGI with $c^{(n)} \neq 0$ can be regarded as an evaluation of the exact flow, corresponding to a modified potential

$$\mathcal{V}_{\text{FG}}(\boldsymbol{q}) = \mathcal{V}(\boldsymbol{q}) - \frac{c^{(n)} h^2}{b^{(n)}} \mathcal{V}_{\hat{C}}(\boldsymbol{q}),$$

where $\mathcal{V}_{\hat{C}}(\boldsymbol{q})$ is solved by $e_i(\mathcal{V}_{\hat{C}}) = 2e^j(\mathcal{V}) e_j e_i(\mathcal{V})$, that is again Hamiltonian. Due to the symmetric construction of (15), the integrator is time-reversible (8). As the integrator is a composition of exact flows that are symplectic, the overall integration scheme is symplectic and thus volume-preserving (9). The integrator also satisfies the closure property (10) as a) the position updates (14) remain unchanged and b) the momentum updates (13) still yield momenta $\boldsymbol{P} \in \mathfrak{g}$.

Backward error analysis. It is a well-known quantity of symplectic integration schemes that they preserve a nearby *shadow Hamiltonian* $\tilde{\mathcal{H}}$ exactly [30]. Given a symplectic integrator of order p, it holds $\tilde{\mathcal{H}} = \mathcal{H} + \mathcal{O}(h^p)$. As the FGI can be written as an exponential (11), it is straight-forward to determine an explicit expression for $\tilde{\mathcal{H}}$. Replacing the commutators of the vector fields $\hat{V}, \hat{T}$ by Poisson brackets of the zero-forms $\mathcal{V}, \mathcal{T}$ gives the shadow Hamiltonian

$$\begin{aligned} \tilde{\mathcal{H}}_{\text{FG}} = &\, \nu \mathcal{T} + \sigma \mathcal{V} + h^2 \left(\alpha \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\} + \beta \{\mathcal{V}, \{\mathcal{T}, \mathcal{V}\}\}\right) \\ &+ h^4 \left(\gamma_1 \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \gamma_2 \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \gamma_3 \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \gamma_4 \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\}\right) \\ &+ h^6 \left(\zeta_1 \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_2 \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_3 \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\}\right. \\ &\quad + \zeta_4 \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_5 \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_6 \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} \\ &\quad + \zeta_7 \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_8 \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} + \zeta_9 \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\} \\ &\quad \left. + \zeta_{10} \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \mathcal{T}\}\}\}\}\}\}\right) + \mathcal{O}(h^8) \end{aligned} \tag{16}$$

that is exactly preserved by the FGI (15).

Example 1 (*Fourth-order FGI*). One example of a fourth-order FGI that demands only five stages and a single evaluation of the FG-term is given by

$$\exp\left(\frac{h}{6} \hat{V}\right) \exp\left(\frac{h}{2} \hat{T}\right) \exp\left(\frac{2h}{3} \hat{V} + \frac{h^3}{72} \hat{C}\right) \exp\left(\frac{h}{2} \hat{T}\right) \exp\left(\frac{h}{6} \hat{V}\right). \tag{17}$$

This integrator exactly preserves the shadow Hamiltonian

$$\begin{aligned} \tilde{\mathcal{H}}_{\text{FGI4}} = &\, \mathcal{T} + \mathcal{V} + h^4 \left(\frac{1}{2880} \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \frac{1}{2880} \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{V}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \frac{1}{2160} \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\} + \frac{1}{4320} \{\mathcal{V}, \{\mathcal{V}, \{\mathcal{T}, \{\mathcal{T}, \mathcal{V}\}\}\}\}\right) \\ &+ \mathcal{O}(h^6). \end{aligned}$$

In general, one has to derive the FG-term $\hat{C}$ for each particular Hamiltonian system by computing the Hessian of the potential. An approximation of the FG-term proposed in [19] enables the use of the force-gradient integrator (17) by replacing the FG-term by an additional force evaluation. We will generalize this idea to the entire class of FGIs.

3.2. Approximation of the force-gradient step

The derivation of the FG-term (12) can be quite tedious. In [22], for example, the FG-term for the two-dimensional Schwinger model has been computed. In other applications, e.g. the application for Hybrid Monte Carlo (HMC) [5] simulations in lattice QCD on four-dimensional lattices [31], it becomes even more complicated to derive the FG-term. It is possible to approximate the force-gradient step (13) by a composed evaluation of the force $e_i(\mathcal{V})$ as it has been shown for Example 1 in [21]. A generalization to this approximation for arbitrary FGIs reads

$$
\begin{aligned}
\exp\left(b^{(n)}h\hat{\mathcal{V}} + c^{(n)}h^3\hat{C}\right) &= \exp\left(-b^{(n)}he_i(\mathcal{V})\frac{\partial}{\partial p_i} + 2c^{(n)}h^3 e^j(\mathcal{V})e_j e_i(\mathcal{V})\frac{\partial}{\partial p_i}\right) = \exp\left(-b^{(n)}h\left(\mathrm{Id} - \frac{2c^{(n)}h^2}{b^{(n)}}e^j(\mathcal{V})e_j\right)e_i(\mathcal{V})\frac{\partial}{\partial p_i}\right) \\
&= \exp\left(-b^{(n)}h\exp\left(-\frac{2c^{(n)}h^2}{b^{(n)}}F^j e_j\right)e_i(\mathcal{V})\frac{\partial}{\partial p_i}\right) + \mathcal{O}(h^5)
\end{aligned}
\tag{18}
$$

Here, $F^j e_j = e^j(\mathcal{V})_q e_j$ denotes that the vector field is frozen at (p, q) [32] so that e_j acting on F^j is defined to be zero. Moreover, performing a Taylor expansion verifies the equality

$$
\exp\left(-\frac{2c^{(n)}h^2}{b^{(n)}}F^j e_j\right)e_i(\mathcal{V})(Q) = e_i(\mathcal{V})\left(\exp\left(-\frac{2c^{(n)}h^2}{b^{(n)}}F^j T_j\right)Q\right).
$$

Consequently, this approximation allows one to replace the momentum update (13) by the following two-step procedure:

1. Compute a temporary position update via $Q' = \exp\left(-\frac{2c^{(n)}h^2}{b^{(n)}}F^j T_j\right)Q$.
2. Compute a usual momentum update of step size $b^{(n)}h$ by using the temporary position update for evaluating the force term $-e_i(\mathcal{V})$, i.e., $P_{b^{(n)}h} = P_0 - b^{(n)}he_i(\mathcal{V})(Q')T^i$.

Although this two-step procedure is used in implementations, the following equivalent expression is of importance for performing the error analysis.

Theorem 1 (Hessian-free force-gradient step). *The two-step procedure is equivalent to replacing the momentum update* (13) *by*

$$
\exp\left(b^{(n)}h\hat{D}(h, b^{(n)}, c^{(n)})\right) := \exp\left(b^{(n)}h\sum_{k=0}^{\infty}\frac{2^k c^{(n)^k}h^{2k}}{k! \cdot b^{(n)^k}}\left(\widetilde{\mathcal{V}}\hat{\mathcal{T}}\right)^k\hat{\mathcal{V}}\right),
\tag{19}
$$

with $\widetilde{\mathcal{V}} := -e_i(\mathcal{V})_q\frac{\partial}{\partial p_i}$.

Proof. When applying the exponential (13) to (p, q), one obtains

$$
\begin{aligned}
\exp\left(b^{(n)}h\hat{\mathcal{V}} + c^{(n)}h^3\hat{C}\right)\begin{pmatrix}p\\q\end{pmatrix} &= \exp\left(b^{(n)}h\hat{\mathcal{V}} + 2c^{(n)}h^3\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\right)\begin{pmatrix}p\\q\end{pmatrix} = \exp\left(b^{(n)}h\left(\mathrm{Id} + \frac{2c^{(n)}h^2}{b^{(n)}}\hat{\mathcal{V}}\hat{\mathcal{T}}\right)\hat{\mathcal{V}}\right)\begin{pmatrix}p\\q\end{pmatrix} \\
&= \exp\left(b^{(n)}h\left(\mathrm{Id} + \frac{2c^{(n)}h^2}{b^{(n)}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\right)\hat{\mathcal{V}}\right)\begin{pmatrix}p\\q\end{pmatrix} = \exp\left(b^{(n)}h\exp\left(\frac{2c^{(n)}h^2}{b^{(n)}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\right)\hat{\mathcal{V}}\right)\begin{pmatrix}p\\q\end{pmatrix} + \mathcal{O}(h^5).
\end{aligned}
$$

Applying the definition of the matrix exponential to the inner exponential concludes the proof. $\square$

Consequently, the symmetric transformations (15b) and (15c) become

$$
\Psi^{(n)} = \exp\left(b^{(n)}h\hat{D}(h, b^{(n)}, c^{(n)})\right) \circ \exp\left(a^{(n)}h\hat{\mathcal{T}}\right) \circ \Psi^{(n-1)} \circ \exp\left(a^{(n)}h\hat{\mathcal{T}}\right) \circ \exp\left(b^{(n)}h\hat{D}(h, b^{(n)}, c^{(n)})\right),
\tag{20a}
$$

$$
\Psi^{(n)} = \exp\left(a^{(n)}h\hat{\mathcal{T}}\right) \circ \exp\left(b^{(n)}h\hat{D}(h, b^{(n)}, c^{(n)})\right) \circ \Psi^{(n-1)} \circ \exp\left(b^{(n)}h\hat{D}(h, b^{(n)}, c^{(n)})\right) \circ \exp\left(a^{(n)}h\hat{\mathcal{T}}\right),
\tag{20b}
$$

respectively. We refer to FGIs that utilize the approximation (18) as *Hessian-free FGIs*. From now on, we will refer to FGIs evaluating the FG-term (12) as *exact FGIs*.

Order conditions. As the approximation (19) introduces an error of $\mathcal{O}(h^5)$, the order conditions up to order four do not change. Based on the expression (19), it is possible to explicitly state the formulae of the error terms for Hessian-free FGIs. Of particular interest are the changes in the error terms (11c) and (11d). The new error terms read

$$
\begin{aligned}
\mathcal{O}'_5 &= \mathcal{O}_5 + \gamma_5\,\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}, \\
\mathcal{O}'_7 &= \mathcal{O}_7 + \zeta_{11}\,\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}} + \zeta_{12}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}} + \zeta_{13}\hat{\mathcal{T}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}},
\end{aligned}
\tag{21}
$$

where the new multipliers $\gamma_5, \zeta_{11}, \zeta_{12}, \zeta_{13}$ can be computed based on recursive relations. Starting from

$$
\gamma_5^{(0)} := \begin{cases} 0, & P \text{ even (velocity version)}, \\ 2c_{(P+1)/2}^2/b_{(P+1)/2}, & P \text{ odd (velocity version)}, \\ 2c_{P/2}^2/b_{P/2}, & P \text{ even (position version)}, \\ 0, & P \text{ odd (position version)}, \end{cases}
\qquad
\zeta_{11}^{(0)} := \begin{cases} 0, & P \text{ even (velocity version)}, \\ 4c_{(P+1)/2}^3/(3b_{(P+1)/2}^2), & P \text{ odd (velocity version)}, \\ 4c_{P/2}^3/(3b_{P/2}^2), & P \text{ even (position version)}, \\ 0, & P \text{ odd (position version)}, \end{cases}
$$

and $\zeta_{12}^{(0)} = \zeta_{13}^{(0)} = 0$, the recursive relations for the symmetric transformations (20a) become

$$\gamma_5^{(n)} = \gamma_5^{(n-1)} + 4c^{(n)2}/b^{(n)}, \qquad\qquad \zeta_{11}^{(n)} = \zeta_{11}^{(n-1)} + \left(8c^{(n)3}/b^{(n)} + 2\sigma^{(n)}\nu^{(n)}c^{(n)2}\right)/(3b^{(n)}),$$

$$\zeta_{12}^{(n)} = \zeta_{12}^{(n-1)} + a^{(n)2}\gamma_5^{(n-1)}/3 - 2\nu^{(n)2}c^{(n)2}/(3b^{(n)}), \qquad \zeta_{13}^{(n)} = \zeta_{13}^{(n-1)} - a^{(n)2}\gamma_5^{(n-1)}/6 + \nu^{(n)2}c^{(n)2}/(3b^{(n)}),$$

and for the symmetric transformations (20b)

$$\gamma_5^{(n)} = \gamma_5^{(n-1)} + 4c^{(n)2}/b^{(n)}, \qquad\qquad \zeta_{11}^{(n)} = \zeta_{11}^{(n-1)} + \left(8c^{(n)3}/b^{(n)} + 2\sigma^{(n)}\nu^{(n)}c^{(n)2}\right)/(3b^{(n)}),$$

$$\zeta_{12}^{(n)} = \zeta_{12}^{(n-1)} + a^{(n)2}\gamma_5^{(n)}/3 - 2\nu^{(n)2}c^{(n)2}/(3b^{(n)}), \qquad \zeta_{13}^{(n)} = \zeta_{13}^{(n-1)} - a^{(n)2}\gamma_5^{(n)}/6 + \nu^{(n)2}c^{(n)2}/(3b^{(n)}),$$

for $n = 1, \ldots, \lfloor \frac{P}{2} \rfloor$.

Geometric integration. The Hessian-free force-gradient step (19) yields updated momenta $\boldsymbol{P} \in \mathfrak{g}$, whereas the position updates (14) remain unchanged. Consequently, the closure property (10) is still satisfied. The approximation does not affect the time-reversibility of the integrator. Furthermore, Hessian-free FGIs are a composition of shears that are volume-preserving so that the overall integrator preserves the volume. Hessian-free FGIs are no longer symplectic. The modified momentum updates (19) are only symplectic if the Hessian of the potential commutes with itself when evaluated at (in general, different) $\boldsymbol{Q}$. Generally, this is not the case.

Backward error analysis. Due to the additional error terms, the modified differential equation is not Hamiltonian, i.e. it does not preserve a nearby shadow Hamiltonian. Instead, the modified differential equation of the Hessian-free FGI reads

$$\begin{pmatrix} \dot{\boldsymbol{P}} \\ \dot{\boldsymbol{Q}} \end{pmatrix} = \begin{pmatrix} \{\tilde{\mathcal{H}}_{\mathrm{FG}}, \boldsymbol{P}\} \\ \{\tilde{\mathcal{H}}_{\mathrm{FG}}, \boldsymbol{Q}\} \end{pmatrix} + \begin{pmatrix} h^4\gamma_5\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\boldsymbol{P} + h^6\zeta_{11}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\boldsymbol{P} + h^6\zeta_{13}\hat{\mathcal{T}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\boldsymbol{P} \\ h^6\zeta_{12}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\boldsymbol{Q} \end{pmatrix} + \mathcal{O}(h^8), \tag{23}$$

where the first term corresponds to the shadow Hamiltonian (16) of the exact FGI using the same set of coefficients. The remaining terms introduce an energy drift. In general, without making any further assumption on the potential $\mathcal{V}$ of the Hessian-free FGI, the energy will have a linear drift of size $\mathcal{O}(th^{\max\{4,p\}})$ for a Hessian-free FGI of order p and a time interval of length t. Note that one may choose the coefficients of the Hessian-free FGI to cancel drift terms so that one ends up with Hessian-free FGIs of order p where the energy will have a linear drift of size $\mathcal{O}(th^r)$ with $r > p$.

Example 2 *(Hessian-free fourth-order FGI).* Based on the fourth-order exact FGI (17), one gets the Hessian-free FGI

$$\exp\left(\frac{h}{6}\hat{\mathcal{V}}\right)\exp\left(\frac{h}{2}\hat{\mathcal{T}}\right)\exp\left(\frac{2h}{3}\hat{D}(h,\tfrac{2}{3},\tfrac{1}{72})\right)\exp\left(\frac{h}{2}\hat{\mathcal{T}}\right)\exp\left(\frac{h}{6}\hat{\mathcal{V}}\right) \tag{24}$$

of convergence order four. With $\gamma_5 = 1/1728 \neq 0$, the numerical energy has a linear drift of size $\mathcal{O}(th^4)$.

Despite the lack of symplecticity, Hessian-free FGIs are interesting integrators and are particularly useful in the following situations.

a) The evaluation of the FG-term $\hat{C}$ is significantly more expensive than an evaluation of the force $e_i(\mathcal{V})$,
b) An analytical expression for the FG-term $\hat{C}$ is not available, e.g. because its derivation is tedious.

In certain applications, such as particles moving in gravitational or coulomb fields, evaluations of the FG-term are less expensive than force evaluations due to the specific structure of the potential. Here, Hessian-free FGIs are less efficient and not of practical relevance unless the exact FG-term is not available. In case of expensive evaluations (e.g. in molecular dynamics simulations), it demands a refined analysis of Hessian-free FGIs to clarify whether Hessian-free FGIs allow for a more efficient computational process. For this purpose, the next section provides a classification and derivation of Hessian-free FGIs with up to 11 stages, similar to the investigations made in [13].

4. Derivation of Hessian-free force-gradient integrators

In this section, we will provide a complete classification of Hessian-free FGIs with up to eleven stages. The coefficients are chosen so that the maximum possible convergence order p is achieved. If the solution is not unique, the free parameters can be optimized by minimizing a (weighted) norm of the multipliers in $\mathcal{O}'_{p+1}$. By expanding the commutators in (21), one ends up with the following contributions containing the new error terms:

- $4\gamma_4\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}} + \gamma_5\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}$,
- $-8\zeta_1\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}} + \zeta_{11}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}$,
- $8\zeta_6\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}} + \zeta_{12}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}$,
- $(-3\zeta_4 - 4\zeta_6)\hat{\mathcal{T}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}} + \zeta_{13}\hat{\mathcal{T}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\widetilde{\mathcal{V}}\hat{\mathcal{T}}\hat{\mathcal{V}}$.

Whereas we choose to minimize the norm

$$\mathrm{Err}_3 := \sqrt{\alpha^2 + \beta^2} \tag{25}$$

in case of $p = 2$, we use weighted norms for $p = 4$ and $p = 6$ incorporating the aforementioned relations. For $p = 4$, we choose the norm

$$\mathrm{Err}_5 := \sqrt{\gamma_1^2 + \gamma_2^2 + \gamma_3^2 + \gamma_4^2 + \left(\tfrac{1}{4}\gamma_5\right)^2}, \tag{26}$$

which, in case of non-gradient schemes ($\gamma_5 = 0$), results in the usual Euclidean norm used in [13]. The weighted norm incorporates that the contribution of γ_5 is four times less than γ_4. Analogously, we choose for $p = 6$

$$\mathrm{Err}_7 := \sqrt{\sum_{j=1}^{10} \zeta_j^2 + \left(\frac{1}{8}\zeta_{11}\right)^2 + \left(\frac{1}{8}\zeta_{12}\right)^2 + \left(\frac{7}{24}\zeta_{13}\right)^2}. \tag{27}$$

Here, the weight of 7/24 for ζ_{13} is derived by taking the mean value of 1/3 and 1/4, so that for non-gradient schemes ($\zeta_{11} = \zeta_{12} = \zeta_{13} = 0$) the weighted norm again reduces to the norm used in [13]. Moreover, the efficiency of the integrators has been measured using the formula

$$\mathrm{Eff}^{(p)} = \frac{1}{(n_f)^p \cdot \mathrm{Err}_{p+1}}. \tag{28}$$

Note that in order to reduce the number of unknowns, we will take into account in advance the symmetry of the FGI, as well as the order-1 conditions $\sum_{j=1}^{P} a_j = \sum_{j=1}^{P} b_j = 1$ throughout this section. For any number of stages s, we obtain different variants based on the coefficients c_j set to zero in order to reduce the number n_f of force evaluations. All variants are summarized by the total number of force evaluations n_f per time step, the achieved convergence order p, and the number of degrees of freedom v, followed by the set of coefficients that is minimizing the (weighted) norm of the leading error term Err_{p+1}. In addition, the corresponding efficiency according to (28) is stated. Note that for velocity algorithms, the number of force recalculations n_f is reduced by one since the last modified momentum update of time step n and the first one of time step $n+1$ coincide so that the force does not have to be re-evaluated.

4.1. Hessian-free force-gradient integrators with $s = 3$ stages (P=2)

For $P = 2$, it is only possible to achieve convergence order $p = 2$. The free parameters are chosen so that the norm (25) is minimized.

4.1.1. Velocity versions
The velocity version of a convergent three-stage Hessian-free FGI reads

$$\Psi_h = \exp\left(\frac{h}{2}\hat{D}(h, 1/2, c_1)\right) \exp\left(h\hat{T}\right) \exp\left(\frac{h}{2}\hat{D}(h, 1/2, c_1)\right)$$

with free parameter c_1. The method has convergence order two with leading error terms $\alpha = 1/12$ and $\beta = 1/24 + 2c_1$.
Variant 1 ($n_f = 2$, $p = 2$, $v = 1$).

$$c_1 = -1/48, \qquad \mathrm{Err}_3 = 1/12, \qquad \mathrm{Eff} = 3. \tag{29}$$

Variant 2 ($n_f = 1$, $p = 2$, $v = 0$). For $c_1 = 0$, one obtains the well-known velocity version of the Störmer–Verlet scheme [1], resulting in

$$\mathrm{Err}_3 = \sqrt{5}/24 \approx 0.0932, \quad \text{and} \quad \mathrm{Eff} = 24/\sqrt{5} \approx 10.73. \tag{30}$$

4.1.2. Position versions
The position version for $P = 2$ reads

$$\Psi_h = \exp\left(\frac{h}{2}\hat{T}\right) \exp\left(h\hat{D}(h, 1, c_1)\right) \exp\left(\frac{h}{2}\hat{T}\right)$$

with free parameter c_1. The scheme has convergence order two and leading error terms $\alpha = -1/24$ and $\beta = -1/12 + c_1$.
Variant 1 ($n_f = 2$, $p = 2$, $v = 1$).

$$c_1 = 1/12, \qquad \mathrm{Err}_3 = 1/24, \qquad \mathrm{Eff} = 6. \tag{31}$$

Variant 2 ($n_f = 1$, $p = 2$, $v = 0$). Setting $c_1 = 0$ yields the position version of the Störmer–Verlet method [1] with

$$\mathrm{Err}_3 = \sqrt{5}/24 \approx 0.0932, \quad \text{and} \quad \mathrm{Eff} = 24/\sqrt{5} \approx 10.7. \tag{32}$$

4.1.3. Remarks
As all schemes for $s = 3$ cannot achieve a higher convergence order than two, the optimal values for c_1, as well as the principal error term do not change compared to the results in [13]. However, the efficiency of the Hessian-free FGIs with $c_1 \neq 0$ increases since in [13] evaluating the FG-term is assumed to be twice as expensive as the force.

4.2. Hessian-free force-gradient integrators with $s = 5$ stages (P=3)

Considering $P = 3$, it is possible to obtain Hessian-free FGIs of convergence order $p = 4$. For fourth-order methods, we minimize Err_5 defined in (26) and denote its minimum by $\gamma_{\min}$.

4.2.1. Velocity versions
The velocity version

$$\Psi_h = \exp\left(b_1 h\hat{D}(h, b_1, c_1)\right) \exp\left(\frac{h}{2}\hat{T}\right) \exp\left((1 - 2b_1)h\hat{D}(h, 1 - 2b_1, c_2)\right) \exp\left(\frac{h}{2}\hat{T}\right) \exp\left(b_1 h\hat{D}(h, b_1, c_1)\right)$$

has three free parameters b_1, c_1, c_2. Solving the order-three conditions while keeping c_1 as a free parameter gives

$$b_1 = 1/6, \qquad c_2 = 1/72 - 2c_1. \tag{33}$$

Variant 1 ($n_f = 4$, $p = 4$, $v = 1$). The minimization of Err_5 yields

$$c_1 = -0.000881991367333, \quad c_2 = 0.015652871623554, \quad \gamma_{\min} \approx 0.000625, \quad \mathrm{Eff} \approx 6.25. \tag{34}$$

Variant 2 ($n_f = 3$, $p = 4$, $v = 0$). Setting $c_1 = 0$, (33) yields the unique solution

$$c_2 = 1/72, \quad \gamma_{\min} \approx 0.000728, \quad \text{Eff} \approx 16.96. \tag{35}$$

Variant 3 ($n_f = 3$, $p = 4$, $v = 0$). For $c_2 = 0$, (33) yields

$$c_1 = 1/144, \quad \gamma_{\min} \approx 0.00335, \quad \text{Eff} \approx 3.68. \tag{36}$$

Variant 4 ($n_f = 2$, $p = 2$, $v = 1$). For $c_1 = c_2 = 0$, we obtain a non-gradient algorithm, i.e., the approximation of the FG-term does not apply. Hence we obtain exactly the same coefficients and principal error term as in [13], namely

$$b_1 = \frac{1}{2} - \frac{(2\sqrt{326}+36)^{1/3}}{12} + \frac{1}{6(2\sqrt{326}+36)^{1/3}} \approx 0.1931833275037836, \qquad \text{Err}_3 \approx 0.00855, \qquad \text{Eff} \approx 29.24. \tag{37}$$

4.2.2. Position versions

The position version

$$\Psi_h = \exp\left(a_1 h\hat{\mathcal{T}}\right)\exp\left(\frac{h}{2}\hat{D}(h,1/2,c_1)\right)\exp\left((1-2a_1)h\hat{\mathcal{T}}\right)\exp\left(\frac{h}{2}\hat{D}(h,1/2,c_1)\right)\exp\left(a_1 h\hat{\mathcal{T}}\right)$$

has two degrees of freedom, namely a_1, c_1.

Variant 1 ($n_f = 4$, $p = 4$, $v = 0$). We obtain two real solutions where the optimal one reads

$$a_1 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right), \quad c_1 = \frac{1}{48}(2 - \sqrt{3}), \quad \gamma_{\min} \approx 0.000718, \quad \text{Eff} \approx 5.44. \tag{38}$$

Variant 2 ($n_f = 2$, $p = 2$, $v = 1$). Setting $c_1 = 0$, the Hessian-free FGI reduces to a non-gradient scheme that does not make use of the approximation to the FG-term $\hat{C}$. The global minimum of Err_3 is achieved at

$$a_1 = \frac{1}{2} - \frac{(2\sqrt{326}+36)^{1/3}}{12} + \frac{1}{6(2\sqrt{326}+36)^{1/3}} \approx 0.1931833275037836, \qquad \text{Err}_3 \approx 0.00855, \qquad \text{Eff} \approx 29.24. \tag{39}$$

4.2.3. Remarks

All variants except (34) have the same optimal solutions as the exact FGIs in [13]. However, their principal error term and efficiency change due to the additional error term γ_5 (except for the non-gradient versions). In case of (34), the additional error term also results in a difference in the coefficients, particularly in c_1 and thus in c_2. The Hessian-free FGI (35) coincides with (24) which was the first Hessian-free FGI derived in [21] and has been applied to the Schwinger model [22] in a nested integrator. It is the most efficient Hessian-free FGI with $s = 5$ stages. The non-gradient schemes (37) and (39) define the most efficient integrators of order two within this framework.

4.3. Hessian-free force-gradient integrators with $s = 7$ stages ($P = 4$)

For $P = 4$, one can again derive Hessian-free FGIs up to convergence order four. However, the additional number of stages allows to reduce the principal error term Err_5. Furthermore, it is now possible to derive non-gradient schemes of order $p = 4$.

4.3.1. Velocity versions

The velocity version with $P = 4$ reads

$$\Psi_h = \exp\left(b_1 h\hat{D}(h,b_1,c_1)\right)\exp\left(a_2 h\hat{\mathcal{T}}\right)\exp\left((\tfrac{1}{2}-b_1)h\hat{D}(h,\tfrac{1}{2}-b_1,c_2)\right)\exp\left((1-2a_2)h\hat{\mathcal{T}}\right)\exp\left((\tfrac{1}{2}-b_1)h\hat{D}(h,\tfrac{1}{2}-b_1,c_2)\right)$$

$$\times \exp\left(a_2 h\hat{\mathcal{T}}\right)\exp\left(hb_1\hat{D}(h,b_1,c_1)\right)$$

where a_2, b_1, c_1, c_2 are free parameters. The order-three conditions have the solutions

$$b_1 = \frac{1}{12}\left(6 + \frac{1}{a_2(a_2-1)}\right), \quad c_1 = -\frac{1}{288}\left(6 + 288c_2 - \frac{1}{a_2(a_2-1)^2}\right). \tag{40}$$

Variant 1 ($n_f = 6$, $p = 4$, $v = 2$). Minimizing Err_5 with respect to the two free parameters a_2, c_2 results in

$$
\begin{aligned}
&a_2 = 0.273005515864808, &&b_1 = 0.080128674198082, &&c_1 = 0.000271601364672, \\
&c_2 = 0.002959399979707, &&\gamma_{\min} \approx 0.0000275, &&\text{Eff} \approx 28.09.
\end{aligned}
\tag{41}
$$

Variant 2 ($n_f = 5$, $p = 4$, $v = 1$). At $c_1 = 0$, the global minimum is reached at

$$
\begin{aligned}
&a_2 = 0.281473422092232, &&b_1 = 0.087960811032557, &&c_2 = 0.003060423791562, \\
&\gamma_{\min} \approx 0.0000498, &&\text{Eff} \approx 32.12.
\end{aligned}
\tag{42}
$$

Variant 3 ($n_f = 4$, $p = 4$, $v = 1$). Setting $c_2 = 0$, one finds

$$
\begin{aligned}
&a_2 = 0.258529167713908, &&b_1 = 0.065274481323251, &&c_1 = 0.003595899064589, \\
&\gamma_{\min} \approx 0.000891, &&\text{Eff} \approx 4.38.
\end{aligned}
\tag{43}
$$

Variant 4 ($n_f = 3$, $p = 4$, $v = 0$). For $c_1 = 0$ and $c_2 = 0$, one obtains the velocity version of the Forest-Ruth algorithm [33] with

$$a_1 = 1/(2 - 2^{1/3}), \qquad b_1 = 1/(2 \cdot (2 - 2^{1/3})), \qquad \gamma_{\min} \approx 0.0383, \qquad \text{Eff} \approx 0.32. \tag{44}$$

4.3.2. Position versions

The position counterpart has the form

$$\Psi_h = \exp\left(a_1 h \hat{\mathcal{T}}\right) \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left((\tfrac{1}{2} - a_1) h \hat{\mathcal{T}}\right) \exp\left((1 - 2b_1) h \hat{D}(h, 1 - 2b_1, c_2)\right) \exp\left((\tfrac{1}{2} - a_1) h \hat{\mathcal{T}}\right)$$
$$\times \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left(a_1 h \hat{\mathcal{T}}\right),$$

where a_1, b_1, c_1, c_2 are free parameters. Analogously to the velocity counterpart, one can solve the third-order conditions for a_1 and c_1, leading to

$$a_1 = \frac{1}{2} \pm \frac{1}{\sqrt{24 b_1}}, \quad c_1 = \frac{1}{24}\left(1 - 12 c_2 \pm \sqrt{6 b_1}(1 - b_1)\right). \tag{45}$$

Variant 1 ($n_f = 6$, $p = 4$, $v = 2$). One reaches the minimum of Err_5 at

$$a_1 = 0.116438749543126, \qquad b_1 = 0.283216992495952, \qquad c_1 = 0.001247201195115,$$
$$c_2 = 0.002974030329635, \qquad \gamma_{\min} \approx 0.0000200, \qquad \text{Eff} \approx 38.57. \tag{46}$$

Variant 2 ($n_f = 5$, $p = 4$, $v = 1$). Putting $c_2 = 0$, one achieves the minimum at

$$a_1 = 0.136458051118946, \qquad b_1 = 0.315267858070664, \qquad c_1 = 0.002427032834125,$$
$$\gamma_{\min} \approx 0.0000844, \qquad \text{Eff} \approx 18.95. \tag{47}$$

Variant 3 ($n_f = 4$, $p = 4$, $v = 1$). Setting $c_1 = 0$, the minimum is achieved at

$$a_1 = 0.089775972994422, \qquad b_1 = 0.247597680043986, \qquad c_2 = 0.006911440413815,$$
$$\gamma_{\min} \approx 0.000149, \qquad \text{Eff} \approx 26.19. \tag{48}$$

Variant 4 ($n_f = 3$, $p = 4$, $v = 0$). The non-gradient version, obtained by setting $c_1 = c_2 = 0$, is the position version of the algorithm by Forest and Ruth [33] which is given by

$$a_1 = 1/(2 \cdot (2 - 2^{1/3})), \qquad b_1 = 1/(2 - 2^{1/3}), \qquad \gamma_{\min} \approx 0.0283, \qquad \text{Eff} \approx 0.44. \tag{49}$$

4.3.3. Remarks

The most efficient seven-stage algorithm is given by the position version (46) with an efficiency of approximately 38.57. Compared to the most efficient five-stage Hessian-free FGI of order four, namely the velocity version (35) with $\text{Eff} \approx 16.96$. Consequently, using the same number of total force recalculations, the seven-stage algorithm (46) will reduce the global errors approximately by a factor of $38.57/16.96 \approx 2.27$ compared to the five-stage Hessian-free FGI (35). The non-gradient versions (44) and (49) now define methods of order four. These methods are equivalent to applying Yoshida's triple-jump composition [11] to (30) and (32), respectively. The principal error term of the non-gradient schemes is relatively large, resulting in less efficient algorithms. In practice, one should either choose force-gradient schemes (e.g., the five-stage Hessian-free FGI (35) reduces the global errors approximately by a factor of $16.96/0.44 \approx 38.55$ while using the same number of total force evaluations) or non-gradient schemes with more stages which will be presented in the proceeding sections.

4.4. Hessian-free force-gradient integrators with $s = 9$ stages ($P = 5$)

Taking $P = 5$ enables the derivation of a Hessian-free FGI of convergence order $p = 6$. In this case, we aim for minimizing the norm (27) and denote its minimum by $\zeta_{\min}$, otherwise we remain at minimizing Err_5 according to (26).

4.4.1. Velocity versions

The velocity version of a nine-stage Hessian-free FGI takes the form

$$\Psi_h = \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left(a_2 h \hat{\mathcal{T}}\right) \exp\left(b_2 h \hat{D}(h, b_2, c_2)\right) \exp\left((\tfrac{1}{2} - a_2) h \hat{\mathcal{T}}\right) \exp\left((1 - 2(b_1 + b_2)) h \hat{D}(h, 1 - 2(b_1 + b_2), c_3)\right)$$
$$\times \exp\left((\tfrac{1}{2} - a_2) h \hat{\mathcal{T}}\right) \exp\left(b_2 h \hat{D}(h, b_2, c_2)\right) \exp\left(a_2 h \hat{\mathcal{T}}\right) \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right)$$

and is characterized by six degrees of freedom $a_2, b_1, b_2, c_1, c_2, c_3$.

Variant 1 ($n_f = 7$, $p = 6$, $v = 0$). Despite there is a new order condition due to γ_5, we find the same unique solution as in [13], namely

$$a_2 = \frac{1}{2} + \frac{\sqrt[3]{675 + 75\sqrt{6}}}{30} + \frac{5}{2\sqrt[3]{675 + 75\sqrt{6}}}, \quad b_1 = \frac{a_2}{3}, \quad b_2 = -\frac{5 a_2}{3}(a_2 - 1), \quad c_1 = 0,$$

$$c_2 = -\frac{5 a_2^2}{144} + \frac{a_2}{36} - \frac{1}{288}, \quad c_3 = \frac{1}{144} - \frac{a_2}{36}\left(\frac{a_2}{2} + 1\right), \quad \zeta_{\min} \approx 0.00154, \quad \text{Eff} \approx 0.0055. \tag{50}$$

Since the unique solution satisfies $c_1 = 0$, this variant coincides with the variant of demanding $c_1 = 0$.

Variant 2 ($n_f = 7$, $p = 4$, $v = 3$). For $c_3 = 0$, minimizing the norm of the leading error term results in

$$a_2 = 0.227758000273404, \qquad b_1 = 0.070935378258660, \qquad b_2 = 0.322911610232109,$$
$$c_1 = 0.000067752132787, \qquad c_2 = 0.001597508440746, \qquad \gamma_{\min} \approx 0.0000101, \text{ Eff} \approx 41.06. \tag{51}$$

Variant 3 ($n_f = 6$, $p = 4$, $v = 3$). By demanding $c_2 = 0$, one gets

$$a_2 = 0.197279141794602, \qquad b_1 = 0.060885008530668, \qquad b_2 = 0.288579639891554,$$
$$c_1 = 0.000429756946246, \qquad c_3 = 0.002373498029145, \qquad \gamma_{\min} \approx 0.0000130, \text{ Eff} \approx 59.33. \tag{52}$$

Variant 4 ($n_f = 6$, $p = 4$, $v = 2$). At $c_1 = 0$ and $c_3 = 0$, the minimization of Err_5 yields

$$a_2 = 0.219039425103133, \qquad b_1 = 0.068466565514186, \qquad b_2 = 0.311000565033563,$$
$$c_2 = 0.001602470431500, \qquad \gamma_{\min} \approx 0.0000105, \qquad \text{Eff} \approx 73.45. \tag{53}$$

Variant 5 ($n_f = 5$, $p = 4$, $v = 2$). For $c_1 = 0$ and $c_2 = 0$, we achieve

$$a_2 = 0.200395293638238, \qquad b_1 = 0.073943321445602, \qquad b_2 = 0.258244950046509,$$
$$c_3 = 0.003147048491590, \qquad \gamma_{\min} \approx 0.0000651, \qquad \text{Eff} \approx 24.57. \tag{54}$$

Variant 6 ($n_f = 5$, $p = 4$, $v = 2$). Putting $c_2 = 0$ and $c_3 = 0$, one finds the minimum

$$a_2 = 0.190585159174513, \qquad b_1 = 0.036356798097337, \qquad b_2 = 0.340278911234329,$$
$$c_1 = 0.002005691094612, \qquad \gamma_{\min} \approx 0.000336, \qquad \text{Eff} \approx 4.76. \tag{55}$$

Variant 7 ($n_f = 4$, $p = 4$, $v = 1$). Finally, choosing $c_1 = 0$, $c_2 = 0$, $c_3 = 0$ and minimizing Err_5 results in the non-gradient scheme [12]

$$a_2 = 0.520943339103990, \qquad b_1 = 0.164498651557576, \qquad b_2 = 1.235692651138917,$$
$$\gamma_{\min} \approx 0.000654, \qquad \text{Eff} \approx 5.97. \tag{56}$$

4.4.2. Position versions

The position version

$$\Psi_h = \exp\left(a_1 h \hat{\mathcal{T}}\right) \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left(a_2 h \hat{\mathcal{T}}\right) \exp\left((\tfrac{1}{2} - b_1) h \hat{D}(h, \tfrac{1}{2} - b_1, c_2)\right) \exp\left((1 - 2(a_1 + a_2)) h \hat{\mathcal{T}}\right)$$
$$\times \exp\left((\tfrac{1}{2} - b_1) h \hat{D}(h, \tfrac{1}{2} - b_1, c_2)\right) \exp\left(a_2 h \hat{\mathcal{T}}\right) \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left(a_1 h \hat{\mathcal{T}}\right)$$

comes with five degrees of freedom a_1, a_2, b_1, c_1, c_2, only allowing for the derivation of Hessian-free FGIs up to convergence order $p = 4$. The order conditions up to order four yield the equations

$$a_1 = \left(3 - 6a_2 + 12a_2 b_1 \mp \sqrt{3}\sqrt{1 - 24a_2^2 b_1 + 48a_2^2 b_1^2}\right)/6,$$
$$c_2 = \left(2 - 12a_2 b_1 + 24a_2 b_1^2 \mp \sqrt{3}\sqrt{1 - 24a_2^2 b_1 + 48a_2^2 b_1^2} - 48c_1\right)/48, \tag{57}$$

so that three degrees of freedom a_2, b_1, c_1 are remaining.

Variant 1 ($n_f = 8$, $p = 4$, $v = 3$). The norm of the leading error term is minimized at sign minus with

$$a_1 = 0.094471605659163, \qquad a_2 = 0.281057227947299, \qquad b_1 = 0.227712700174579,$$
$$c_1 = 0.000577062053569, \qquad c_2 = 0.000817399268485, \qquad \gamma_{\min} \approx 0.00000501, \text{ Eff} \approx 48.71. \tag{58}$$

Variant 2 ($n_f = 6$, $p = 4$, $v = 2$). Setting $c_1 = 0$ results in a minimum achieved at sign minus, given by

$$a_1 = 0.047802682977081, \qquad a_2 = 0.265994592108478, \qquad b_1 = 0.143282503449494,$$
$$c_2 = 0.002065558490728, \qquad \gamma_{\min} \approx 0.0000346, \qquad \text{Eff} \approx 22.32. \tag{59}$$

Variant 3 ($n_f = 6$, $p = 4$, $v = 2$). For $c_2 = 0$, we obtain (at sign minus)

$$a_1 = 0.118030603246046, \qquad a_2 = 0.295446189611111, \qquad b_1 = 0.273985556386628,$$
$$c_1 = 0.001466561305710, \qquad \gamma_{\min} \approx 0.0000471, \qquad \text{Eff} \approx 16.39. \tag{60}$$

Variant 4 ($n_f = 4$, $p = 4$, $v = 1$). With $c_1 = 0$ and $c_2 = 0$, one finds the non-gradient scheme

$$a_1 = 0.178617895844809, \qquad a_2 = -0.066264582669818, \qquad b_1 = 0.712341831062606,$$
$$\gamma_{\min} \approx 0.000610, \qquad \text{Eff} \approx 6.40. \tag{61}$$

4.4.3. Remarks

For $s = 9$, we obtain the first (and at the same time unique) Hessian-free FGI of convergence order $p = 6$, namely the velocity version (50). It is the same unique solution as in [13], although one has to satisfy the new order condition due to γ_5. Since some variants of eleven-stage Hessian-free

FGIs will be affected by order reduction, this can be regarded as a lucky coincidence. All other variants define order-four integrators. The most efficient one is given by the velocity version (53) with an efficiency of $Eff \approx 73.45$. Compared to the most efficient seven-stage variant (46), it allows to reduce the global errors approximately by a factor of $73.45/38.57 \approx 1.9$ while keeping the total number of force recalculations fixed.

4.5. Hessian-free force-gradient integrators with $s = 11$ stages (P = 6)

We finally consider the case of Hessian-free FGIs with $P = 6$. For the cases $P \le 5$, each variant of Hessian-free FGIs has the same convergence order as the common FGIs in [13], i.e., the additional order conditions did not result in an order reduction. For $P = 6$, however, some variants are affected by order reduction.

4.5.1. Velocity versions
In velocity form, the eleven-stage propagation takes the form

$$\Psi_h = \exp\left(b_1 h\hat{D}(h, b_1, c_1)\right) \exp\left(a_2 h\hat{\mathcal{T}}\right) \exp\left(b_2 h\hat{D}(h, b_2, c_2)\right) \exp\left(a_3 h\hat{\mathcal{T}}\right) \exp\left((\tfrac{1}{2} - (b_1 + b_2))h\hat{D}(h, \tfrac{1}{2} - (b_1 + b_2), c_3)\right)$$

$$\times \exp\left((1 - 2(a_2 + a_3))h\hat{\mathcal{T}}\right) \exp\left((\tfrac{1}{2} - (b_1 + b_2))h\hat{D}(h, \tfrac{1}{2} - (b_1 + b_2), c_3)\right) \exp\left(a_3 h\hat{\mathcal{T}}\right) \exp\left(b_2 h\hat{D}(h, b_2, c_2)\right)$$

$$\times \exp\left(a_2 h\hat{\mathcal{T}}\right) \exp\left(b_1 h\hat{D}(h, b_1, c_1)\right)$$

with seven degrees of freedom $a_2, a_3, b_1, b_2, c_1, c_2, c_3$.

Variant 1 ($n_f = 9$, $p = 6$, $v = 0$). The only real solution is given by

$$
\begin{aligned}
&a_2 = 0.270990466773838, \quad && a_3 = 0.635374358266882, \quad && b_1 = 0.090330155591279, \\
&b_2 = 0.430978044876253, \quad && c_1 = 0, \quad && c_2 = 0.002637435980472, \\
&c_3 = -0.000586445610932, \quad && \zeta_{\min} \approx 0.00000699, \quad && Eff \approx 0.27.
\end{aligned}
\tag{62}
$$

Note that (62) also covers the variant demanding $c_1 = 0$.

Variant 2 ($n_f = 8$, $p = 4$, $v = 4$). For $c_2 = 0$, the best solution is achieved at

$$
\begin{aligned}
&a_2 = 0.068597474282941, \quad && a_3 = 0.284851197274498, \quad && b_1 = -0.029456704762871, \\
&b_2 = 0.228751459942521, \quad && c_1 = 0.000410146066173, \quad && c_3 = 0.001249935251564, \\
&\gamma_{\min} \approx 0.00000355, \quad && Eff \approx 68.84.
\end{aligned}
\tag{63}
$$

Variant 3 ($n_f = 8$, $p = 4$, $v = 4$). Putting $c_3 = 0$ yields

$$
\begin{aligned}
&a_2 = 0.203263079324187, \quad && a_3 = 0.200698071607808, \quad && b_1 = 0.066202529912271, \\
&b_2 = 0.267856111220228, \quad && c_1 = 0.000012570620797, \quad && c_2 = 0.001042408779514, \\
&\gamma_{\min} \approx 0.00000519, \quad && Eff \approx 47.08.
\end{aligned}
\tag{64}
$$

Variant 4 ($n_f = 7$, $p = 4$, $v = 3$). Setting $c_1 = 0$ and $c_2 = 0$, the optimal solution looks as

$$
\begin{aligned}
&a_2 = 0.122268182901557, \quad && a_3 = 0.203023211433263, \quad && b_1 = 0.055200549768959, \\
&b_2 = 0.127408150658963, \quad && c_3 = 0.001487834491987, \quad && \gamma_{\min} \approx 0.0000189,\ Eff \approx 21.98.
\end{aligned}
\tag{65}
$$

Variant 5 ($n_f = 7$, $p = 4$, $v = 3$). Putting $c_1 = 0$ and $c_3 = 0$, one finds

$$
\begin{aligned}
&a_2 = 0.201110227930330, \quad && a_3 = 0.200577842713366, \quad && b_1 = 0.065692416344302, \\
&b_2 = 0.264163604920340, \quad && c_2 = 0.001036943019757, \quad && \gamma_{\min} \approx 0.00000520,\ Eff \approx 80.13.
\end{aligned}
\tag{66}
$$

Variant 6 ($n_f = 6$, $p = 4$, $v = 3$). For $c_2 = 0$ and $c_3 = 0$, minimization of Err_5 results in

$$
\begin{aligned}
&a_2 = 0.282918304065611, \quad && a_3 = -0.002348009438292, \quad && b_1 = 0.080181913812571, \\
&b_2 = -1.372969015964262, \quad && c_1 = 0.000325098077953, \quad && \gamma_{\min} \approx 0.0000166,\ Eff \approx 46.47.
\end{aligned}
\tag{67}
$$

Variant 7 ($n_f = 5$, $p = 4$, $v = 2$). Finally, letting $c_1 = c_2 = c_3 = 0$, we obtain the optimal eleven-stage non-gradient scheme (in velocity form) by choosing

$$
\begin{aligned}
&a_2 = 0.253978510841060, \quad && a_3 = -0.032302867652700, \quad && b_1 = 0.083983152628767, \\
&b_2 = 0.682236533571909, \quad && \gamma_{\min} \approx 0.0000270, \quad && Eff \approx 59.26.
\end{aligned}
\tag{68}
$$

4.5.2. Position versions
When starting with a position update (14), the eleven-stage Hessian-free FGI reads

$$\Psi_h = \exp\left(a_1 h\hat{\mathcal{T}}\right) \exp\left(b_1 h\hat{D}(h, b_1, c_1)\right) \exp\left(a_2 h\hat{\mathcal{T}}\right) \exp\left(b_2 h\hat{D}(h, b_2, c_2)\right) \exp\left((\tfrac{1}{2} - (a_1 + a_2))h\hat{\mathcal{T}}\right)$$

$$\times \exp\left((1 - 2(b_1 + b_2))h\hat{D}(h, 1 - 2(b_1 + b_2), c_3)\right) \exp\left((\tfrac{1}{2} - (a_1 + a_2))h\hat{\mathcal{T}}\right)$$

$$\times \exp\left(b_2 h \hat{D}(h, b_2, c_2)\right) \exp\left(a_2 h \hat{\mathcal{T}}\right) \exp\left(b_1 h \hat{D}(h, b_1, c_1)\right) \exp\left(a_1 h \hat{\mathcal{T}}\right),$$

where $a_1, a_2, b_1, b_2, c_1, c_2, c_3$ are free parameters.

Variant 1 ($n_f = 10$, $p = 6$, $v = 0$). In general, there exist four real solutions to obtain a Hessian-free FGI of order $p = 6$. The best solution is

$$
\begin{aligned}
&a_1 = 0.109534125980058, &&a_2 = 0.426279051773841, &&b_1 = 0.268835839917653,\\
&b_2 = 0.529390037396794, &&c_1 = 0.000806354602850, &&c_2 = 0.007662601517364,\\
&c_3 = -0.011627206142396, &&\zeta_{\min} \approx 0.00000603, &&\mathrm{Eff} \approx 0.17.
\end{aligned}
\tag{69}
$$

All four solutions have coefficients $c_j \neq 0$, i.e., an order reduction appears as soon as one c_j is supposed to be zero. The variant with $c_1 = 0$ is omitted since order $p = 6$ can be achieved by setting $a_1 = 0$, resulting in the nine-stage velocity algorithm (50).

Variant 2 ($n_f = 8$, $p = 4$, $v = 4$). Putting $c_2 = 0$ yields

$$
\begin{aligned}
&a_1 = 0.083684971641549, &&a_2 = 0.225966488946428, &&b_1 = 0.199022868372193,\\
&b_2 = 0.197953981691206, &&c_1 = 0.000437056543403, &&c_3 = 0.000870457820984,\\
&\gamma_{\min} \approx 0.00000318, &&\mathrm{Eff} \approx 76.79.
\end{aligned}
\tag{70}
$$

Variant 3 ($n_f = 9$, $p = 4$, $v = 4$). For $c_3 = 0$, the minimum is achieved at

$$
\begin{aligned}
&a_1 = 0.082541033171754, &&a_2 = 0.228637847036999, &&b_1 = 0.196785139280847,\\
&b_2 = 0.206783248777282, &&c_1 = 0.000317260402502, &&c_2 = 0.000555360763892,\\
&\gamma_{\min} \approx 0.00000235, &&\mathrm{Eff} \approx 64.99.
\end{aligned}
\tag{71}
$$

Variant 4 ($n_f = 6$, $p = 4$, $v = 3$). Setting $c_1 = 0$ and $c_2 = 0$ results in

$$
\begin{aligned}
&a_1 = 0.134257092137626, &&a_2 = -0.007010267216916, &&b_1 = -0.485681409840328,\\
&b_2 = 0.767464037573892, &&c_3 = 0.002836723107629, &&\gamma_{\min} \approx 0.0000154, \;\; \mathrm{Eff} \approx 50.09.
\end{aligned}
\tag{72}
$$

Variant 5 ($n_f = 7$, $p = 4$, $v = 3$). At $c_1 = 0$ and $c_3 = 0$, the best solution reads

$$
\begin{aligned}
&a_1 = 0.062702644098210, &&a_2 = 0.193174566017780, &&b_1 = 0.149293739165427,\\
&b_2 = 0.220105234408407, &&c_2 = 0.000966194415594, &&\gamma_{\min} \approx 0.00000445, \;\; \mathrm{Eff} \approx 93.60.
\end{aligned}
\tag{73}
$$

Variant 6 ($n_f = 7$, $p = 4$, $v = 3$). Letting $c_2 = 0$ and $c_3 = 0$, the minimization of Err_5 leads to

$$
\begin{aligned}
&a_1 = 0.115889910143319, &&a_2 = 0.388722377182381, &&b_1 = 0.282498420841510,\\
&b_2 = -0.625616553474143, &&c_1 = 0.001208219887746, &&\gamma_{\min} \approx 0.0000128, \;\; \mathrm{Eff} \approx 32.64.
\end{aligned}
\tag{74}
$$

Variant 7 ($n_f = 5$, $p = 4$, $v = 2$). Demanding $c_1 = c_2 = c_3 = 0$, we arrive at the non-gradient scheme

$$
\begin{aligned}
&a_1 = 0.275008121233242, &&a_2 = -0.134795009910679, &&b_1 = -0.084429619507071,\\
&b_2 = 0.354900057157426, &&\gamma_{\min} \approx 0.0000518, &&\mathrm{Eff} \approx 30.89.
\end{aligned}
\tag{75}
$$

4.5.3. Remarks

The variants include two additional Hessian-free FGIs of order $p = 6$, a velocity version (62) with efficiency $\mathrm{Eff} \approx 0.27$ and a position version (69) with $\mathrm{Eff} \approx 0.17$. Both versions are significantly more efficient than the nine-stage algorithm (50) with $\mathrm{Eff} \approx 0.0055$: using the same number of total force evaluations, the eleven-stage algorithms reduce the global errors approximately by a factor of $0.27/0.0055 \approx 49.09$ and $0.17/0.0055 \approx 30.91$, respectively.

The most efficient eleven-stage Hessian-free FGI of order $p = 4$ is given by the position algorithm (73) with $\mathrm{Eff} \approx 93.60$. Compared to the most efficient nine-stage algorithm (53) of order four ($\mathrm{Eff} \approx 73.45$), the eleven-stage algorithm reduces the global errors approximately by a factor of $93.60/73.45 \approx 1.27$. The most efficient velocity algorithm is (66) with an efficiency of approximately 80.13. The non-gradient algorithm (68) ($\mathrm{Eff} \approx 59.3$) became a state-of-the-art scheme in lattice quantum chromodynamics [31] due to its high efficiency compared to lower-stage algorithms. Based on the efficiency measure, both the position algorithm (73) and the velocity algorithm (68) are able to reduce the global errors while keeping the number of total force recalculations fixed, namely by a factor of approximately 1.58 and 1.35, respectively.

4.6. Complete classification

A summary of all schemes derived in this section is provided in Table 1. The table contains the schemes in abbreviated forms, i.e., A and B correspond to the exponential operators $\exp(a_j h \hat{\mathcal{T}})$ and $\exp(b_j h \hat{V})$, respectively, and D denotes approximated FG-updates of the form $\exp(b_j h \hat{D}(h, b_j, c_j))$. For any number of stages $s = 3, 5, 7, 9, 11$, the integrators are grouped according to their order of convergence. Within each group, the schemes are sorted in ascending order with respect to the number of force evaluations n_f. For the same number of evaluations, more efficient schemes are listed first. While some of the new schemes have only slightly different time coefficients compared to the results in [13], there are also variants where one can observe large differences in the time coefficients.

Table 1

Collection of Hessian-free force-gradient integrators with up to eleven stages. For all algorithms, the convergence order p, the number of force evaluations n_f per time step, the leading error term Err_{p+1}, its efficiency Eff, as well as a remark on the earlier appearance of the time coefficients, are stated. For completeness, the final column refers to the ID of the respective variant in [13].

Algorithm	Equation	p	n_f	Err_{p+1}	Efficiency	Remarks	ID in [13]
BAB	(30)	2	1	0.0932	10.73	[34]	1
ABA	(32)	2	1	0.0932	10.73	[34]	2
DAD	(29)	2	2	0.0833	3.00	[13,20]	3
ADA	(31)	2	2	0.0417	6.00	[13]	4
BABAB	(37)	2	2	0.00855	29.24	[13,35]	5
ABABA	(39)	2	2	0.00855	29.24	[13,35]	6
BADAB	(35)	4	3	0.000728	16.96	[36,37]	8
DABAD	(36)	4	3	0.00335	3.68	[13]	7
DADAD	(34)	4	4	0.000625	6.25	New	9
ADADA	(38)	4	4	0.000718	5.44	[36,37]	10
ABABABA	(49)	4	3	0.0283	0.44	[11,38]	12
BABABAB	(44)	4	3	0.0383	0.32	[11,38]	11
ABADABA	(48)	4	4	0.000149	26.19	New	14
DABABAD	(43)	4	4	0.000891	4.38	New	13
BADADAB	(42)	4	5	0.0000498	32.12	New	15
ADABADA	(47)	4	5	0.0000844	18.95	New	16
ADADADA	(46)	4	6	0.0000200	38.57	New	18
DADADAD	(41)	4	6	0.0000275	28.09	New	17
ABABABABA	(61)	4	4	0.000610	6.40	[13]	20
BABABABAB	(56)	4	4	0.000654	5.97	[13]	19
BABADABAB	(54)	4	5	0.0000651	24.57	New	21
DABABABAD	(55)	4	5	0.000336	4.76	New	22
BADABADAB	(53)	4	6	0.0000105	73.45	New	24
DABADABAD	(52)	4	6	0.0000130	59.33	New	23
ABADADABA	(59)	4	6	0.0000346	22.32	New	25
ADABABADA	(60)	4	6	0.0000471	16.39	New	26
DADABADAD	(51)	4	7	0.0000101	41.06	New	27
ADADADADA	(58)	4	8	0.00000501	48.71	New	29
BADADADAB	(50)	6	7	0.00154	0.0055	[13]	28
BABABABABAB	(68)	4	5	0.0000270	59.26	[13]	30
ABABABABABA	(75)	4	5	0.0000518	30.89	[13]	31
ABABADABABA	(72)	4	6	0.0000154	50.09	New	33
DABABABABAD	(67)	4	6	0.0000166	46.47	New	32
ABADABADABA	(73)	4	7	0.00000445	93.60	New	36
BADABABADAB	(66)	4	7	0.00000520	80.13	New	34
ADABABABADA	(74)	4	7	0.0000128	32.64	New	37
BABADADABAB	(65)	4	7	0.0000189	21.98	New	35
ADABADABADA	(70)	4	8	0.00000318	76.79	New	40
DABADADABAD	(63)	4	8	0.00000355	68.84	New	38
DADABABABADAD	(64)	4	8	0.00000519	47.08	New	39
ADADABADADA	(71)	4	9	0.00000235	64.99	New	43
BADADADADAB	(62)	6	9	0.00000699	0.27	New	42
ADADADADADA	(69)	6	10	0.00000603	0.17	New	45

4.7. Hessian-free force-gradient integrators of higher orders

In case of very accurate integrations, it can be useful to consider numerical time integration schemes of even higher convergence order p. The derivation of Hessian-free and exact FGIs via direct decompositions will result in the most efficient schemes within the framework. However, the derivation becomes a more challenging numerical problem when further increasing the number of stages.

An alternative approach of deriving numerical time integration schemes of arbitrarily high convergence order is given by means of composition techniques [10–12]. Using symmetric compositions, the time-reversibility, volume-preservation and closure property of the underlying base methods is preserved. Since the derivation based on direct decomposition will always lead to more efficient algorithms, composition techniques should only be applied for very high orders for which no methods based on direct decomposition are known. For example, Yoshida's triple jump [11] based on the non-gradient scheme (30) yields the method (44) which is the least effective decomposition scheme ($\text{Eff} \approx 0.32$) of order four under investigation. Starting from a base method of order p, Yoshida's triple jump requires 3^k applications of the base method to obtain a composition scheme of order $p + 2k$. The alternative approach of Suzuki's fractals [10] demands even more applications, namely 5^k. Advanced composition schemes [12] provide a possibility to derive composition schemes with significantly less stages if one wants to increase the order by more than two orders. Advanced composition schemes starting from a base scheme of order $p \in \{4, 6, 8\}$ have been discussed in [12]. Advanced composition schemes based on base schemes of order $p = 2$ have been investigated in [39].

5. Numerical results

This section provides numerical tests of Hessian-free FGIs for three different test examples. In Section 5.1, numerical tests for the outer solar system, a N-body problem with mass matrix $\boldsymbol{M} \neq \mathrm{Id}$, are performed. Then, in Section 5.2, we consider the Hybrid Monte Carlo (HMC) algorithm [5] in two different applications. The two-dimensional Schwinger model is discussed in Section 5.2.1. Here, we have implemented the analytical FG-term [22]. This enables a comparison of the exact FGIs [13] to the new class of Hessian-free FGIs. Secondly, in Section 5.2.2, four-dimensional gauge field simulations in lattice QCD with two heavy Wilson fermions are discussed. Since the links are elements of the special unitary group SU(3), this provides an example in a non-Abelian setting.

5.1. The outer solar system

We consider the separable Hamiltonian system

$$\mathcal{H}(\boldsymbol{p}, \boldsymbol{q}) = \frac{1}{2} \sum_{i=0}^{5} \frac{1}{m_i} p_i^\top p_i - G \sum_{i=1}^{5} \sum_{j=0}^{i-1} \frac{m_i m_j}{\|q_i - q_j\|},$$

describing the motion of the four outer planets (and Pluto) relative to the sun. Here, the momenta $\boldsymbol{p}$ and positions $\boldsymbol{q}$ are supervectors composed by the vectors $p_i, q_i \in \mathbb{R}^3$. We have taken the parameters from [1] and also took the initial data from [40] corresponding to September 5, 1994 at midnight. Solutions of the outer solar system using the Hessian-free FGI (73) with step size $h = 200$ over a time period of 200,000 days are depicted in Fig. 1, showing the correct behavior. Numerical results for the other variants of Hessian-free FGIs show the same behavior, as expected.
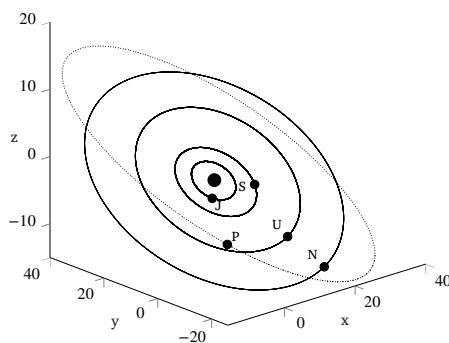
Fig. 1. Outer solar system. Numerical solution using the Hessian-free FGI (73) and $h = 200$ over a time period of 200,000 days.
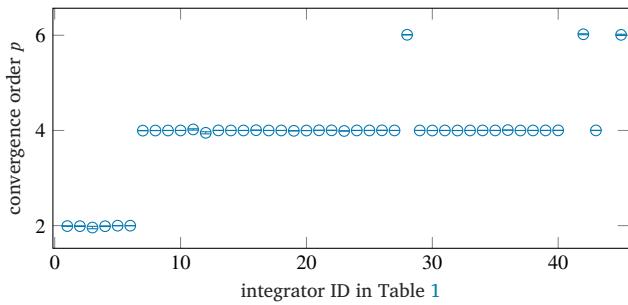
Fig. 2. Outer solar system. Numerical verification of the convergence order of Hessian-free FGIs. For each variant of Hessian-free FGIs, the scaling has been measured performing simulations using different step sizes. The plot displays the mean, as well as the standard deviation of the numerical measurement of the convergence order.

Fig. 2 shows that all Hessian-free FGIs achieve the desired convergence order. These results are more or less just a reference that the derivation and solution of the order conditions has been performed correctly and that Hessian-free FGIs are also applicable to Hamiltonian systems (2) with $\boldsymbol{M} \neq \mathrm{Id}$. Furthermore, the efficiency of the integrators for the outer solar system is depicted in Fig. 3 by plotting the global error vs. the total number of force evaluations for different step sizes. For the sake of clarity, only a selection of integrators is shown: the most efficient non-gradient scheme for any number of stages, the frequently used Hessian-free FGI (35), as well as the three best performing Hessian-free FGIs of order four. Compared to the most efficient fourth-order non-gradient scheme (68), Hessian-free FGIs are able to reduce the error by more than one order of magnitude at the same number of force evaluations, emphasizing the efficiency of the proposed class of integrators. For accuracy demands up to $1e - 4$, the Hessian-free FGI (73) is even able to compete with the most efficient sixth-order integrator under investigation, the 15-stage non-gradient velocity algorithm introduced in [13]. In addition, the results show that the practical efficiency of the integrators might diverge from the theoretical findings summarized in Table 1. For example, (35) ($\mathrm{Eff} \approx 16.96$) is more efficient than (68) ($\mathrm{Eff} \approx 59.26$) and (52) ($\mathrm{Eff} \approx 59.33$) is approximately as efficient as (73) ($\mathrm{Eff} \approx 93.60$). The theoretical efficiency values are derived by equating all brackets to one; therefore, this outcome is not contradictory. Nevertheless, the efficiency measure provides a valuable heuristic, indicating which integrators are likely to be efficient without requiring additional assumptions about the specific system.

Another important feature is the long-time energy conservation of Hessian-free FGIs. As an example, Fig. 4 depicts the relative energy error $(\mathcal{H}(\boldsymbol{p}_n, \boldsymbol{q}_n) - \mathcal{H}(\boldsymbol{p}_0, \boldsymbol{q}_0))/\mathcal{H}(\boldsymbol{p}_0, \boldsymbol{q}_0)$ for the Hessian-free FGI (73). For this problem, no drift is visible, i.e. the Hessian-free FGIs behave like a symplectic integrator and thus show excellent long-term energy behavior. Note that this observation coincides with the observations made in [20] for another N-body problem.

5.2. Hybrid Monte Carlo simulations

The most common approach of simulating quantum field theories on the lattice is given by Hybrid Monte Carlo (HMC) simulations [5]. Within the HMC algorithm, the lattice quantum field theory is embedded in a higher-dimensional classical system by introducing a fictitious simulation time. It is described by a separable Hamiltonian system

$$\mathcal{H}([\boldsymbol{P}], [\boldsymbol{Q}]) = \mathcal{T}([\boldsymbol{P}]) + \mathcal{V}([\boldsymbol{Q}]),$$

reaching the equilibrium distribution of the desired expectation values in the large time limit. The square brackets denote a configuration, i.e. the set of all variables arranged on the lattice. For a given configuration $[\boldsymbol{Q}]$, a new configuration $[\boldsymbol{Q}']$ is generated by performing an HMC update consisting of a molecular dynamics (MD) step and a Metropolis step.
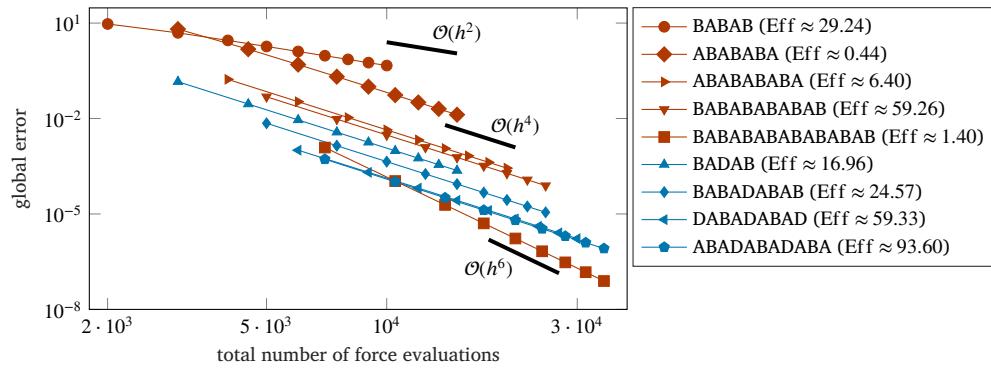
Fig. 3. Outer solar system. Global error vs. number of force evaluations for a selected number of Hessian-free FGIs (blue lines) and non-gradient schemes (red lines). The simulation has been performed over a time period of 200,000 days using different number of steps, nsteps $\in \{1000, 1500, \ldots, 5000\}$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)
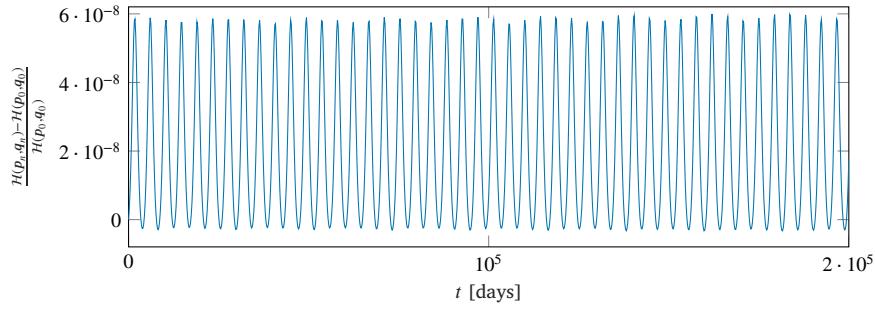

Fig. 4. Outer solar system. Relative energy error of the Hessian-free FGI (73) using $h = 200$ over a time period of 200,000 days.

- **MD step.** For a trajectory of length τ, solve the Hamiltonian equations of motion given by

$$\dot{\boldsymbol{P}}_{x,\mu} = -\frac{\partial \mathcal{H}([\boldsymbol{P}], [\boldsymbol{Q}])}{\partial \boldsymbol{Q}_{x,\mu}}, \qquad \dot{\boldsymbol{Q}}_{x,\mu} = \frac{\partial \mathcal{H}([\boldsymbol{P}], [\boldsymbol{Q}])}{\partial \boldsymbol{P}_{x,\mu}}$$

for each position (x, μ) on the lattice. We denote the new configuration by $([\boldsymbol{P}'], [\boldsymbol{Q}'])$.

- **Metropolis step.** Accept the new configuration $([\boldsymbol{P}'], [\boldsymbol{Q}'])$ with probability

$$\mathcal{P}([\boldsymbol{Q}] \to [\boldsymbol{Q}']) = \min\left(1, \exp(-\Delta\mathcal{H})\right),$$

where $\Delta\mathcal{H} = \mathcal{H}([\boldsymbol{P}'], [\boldsymbol{Q}']) - \mathcal{H}([\boldsymbol{P}], [\boldsymbol{Q}])$. Depending on the acceptance step, either the new or old configuration is added to the Markov chain and used as the initial configuration for the next MD step.

The Markov process will converge to a unique fixed point distribution provided that it is ergodic and satisfies the detailed balance condition [5]. The detailed balance condition is satisfied, if the numerical integration scheme used in the MD step is time-reversible and volume-preserving [31]. Good energy conservation is required to ensure a high acceptance rate in the Metropolis step. At the same time, one is interested in minimizing the computational cost of the HMC algorithm. Consequently, the choice of the integrator in the MD step plays a crucial role. The optimal acceptance rate for a numerical integration scheme of order p is [41]

$$\langle P_{\text{acc}} \rangle_{\text{opt}} = \exp(-1/p). \tag{76}$$

Equation (76) indicates that the optimal acceptance rate for order $p = 2$ is 61%, 78% for order $p = 4$ and 85% for order $p = 6$. In practice, however, achieving the optimal acceptance rate is often challenging, as instabilities emerge as a limiting factor in the step size h. Hence one frequently aims at achieving an acceptance rate of $P_{\text{acc}} \approx 90\%$. For log-normal distributed ΔH, the expected acceptance rate is given by [42]

$$\langle P_{\text{acc}} \rangle = \text{erfc}\left(\sqrt{\sigma^2/8}\right), \tag{77}$$

where σ^2 denotes the (empirical) variance in ΔH and erfc the complementary error function. All in all, we are interested in minimizing the number of force evaluations per unit trajectory n_f/h while achieving a desired acceptance rate $P_{\text{acc}} \in [0, 1]$, i.e., we aim at maximizing the step size h such that $\sigma^2 \leq 8(\text{erfc}^{-1}(P_{\text{acc}}))^2$ is satisfied.

### 5.2.1. The two-dimensional Schwinger model

The two-dimensional Schwinger model is defined by the Hamiltonian

$$\mathcal{H}(p,q) = \frac{1}{2} \sum_{n=1,\mu=1}^{V,2} p_{n,\mu}^2 + \mathcal{S}_G([Q]) + \mathcal{S}_F([Q])$$

with $V = L \times T$ the volume of the lattice and $\mathcal{S}_G([Q])$ the pure gauge action and $\mathcal{S}_F([Q])$ the fermion action. The links Q on the lattice are elements of the Abelian Lie group U(1). Consequently the links $Q_\mu(n)$ connecting the sites n and $n + \hat{\mu}$ and can be written in the form $Q_\mu(n) = \exp(iq_\mu(n)) \in \mathrm{U}(1)$, $q_\mu(n) \in [-\pi, \pi]$, $\mu \in \{x, t\}$ the respective space and time directions. The momenta $p_{n,\mu}$ are elements of the corresponding Lie algebra $\mathbb{R}$. Details on the action $\mathcal{S}$ can be found here Appendix A. For this Hamiltonian system, the force-gradient term $\hat{C}$ has been derived in [22], allowing for a comparison of the exact FGIs in [13] and our Hessian-free ones. On a lattice of size 16×16 at pure gauge coupling $\beta = 1$ and at bare mass parameter $m_0 = 0.352443$ [43], we computed 2000 trajectories of length $\tau = 1$ with $N = 20, 25, 30, 35, 40$ number of steps using any set of coefficients from Table 1 and [13, Tab. 2]. Note that we chose for the simulations a large bare quark mass corresponding to $z = 0.8$ in [43, Tab. 2] to guarantee a stable integration for our setup without nested integration and without Hasenbusch mass-preconditioning in a large volume. For all variants, we consider both the exact FG-term (12) and the approximation (18) resulting in Hessian-free FGIs. By measuring the dependence of the variance σ^2 of ΔH on the number of steps N and using the relation to the acceptance rate (77), we estimated for each integrator the number of required force steps to reach 90% acceptance rate. In Fig. 5, a pair-wise comparison to our proposed integrators (Hessian-free FGIs with coefficients from Table 1) is depicted. Fig. 5a shows that, for almost all variants, Hessian-free FGIs allow for a more efficient computational process than the FGIs from [13]. A major difference can be observed for those variants who are affected by an order reduction, e.g. (70). This can be explained by the fact that in the relevant region of $P_{\mathrm{acc}} \approx 90\%$, second and fourth order integrators turn out to be more efficient for this lattice size. For increasing lattice sizes (see e.g. Section 5.2.2), higher-order schemes become more efficient. For simplicity, exact FGIs are often implemented using the approximation (18), see e.g. [44]. The coefficients derived in [13] do not incorporate the new error terms (21) that occur due to the approximation. Fig. 5b highlights that Hessian-free FGIs with coefficients from Table 1 usually perform better than Hessian-free FGIs with coefficients from [13]. In contrast, Fig. 5c underlines that the coefficients derived in this paper are only recommended for Hessian-free FGIs. Using the coefficients for exact FGIs leads to less efficient schemes, as expected. In this case, the coefficients from [13] should be used since they are optimized for exact FGIs.
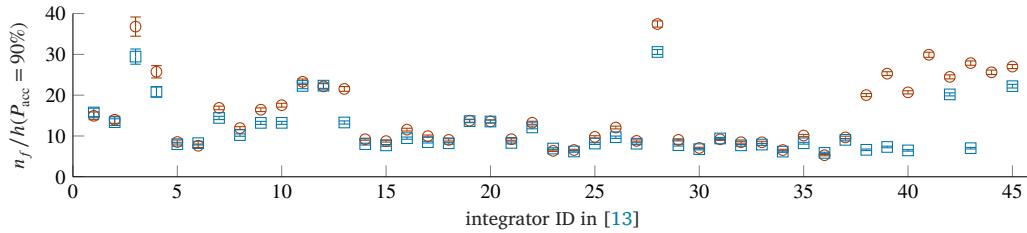
(a) Exact FGIs with coefficients from [13] (red circles) vs. Hessian-free FGIs with coefficients from Table 1 (blue squares).
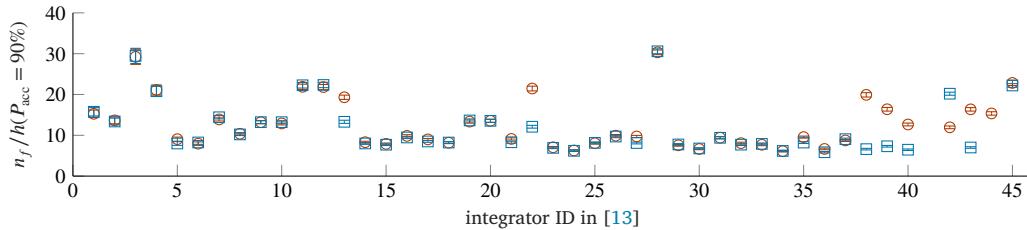
(b) Hessian-free FGIs with coefficients from [13] (red circles) vs. Hessian-free FGIs with coefficients from Table 1 (blue squares). The coefficients from [13] with ID = 41 yield $n_f/h(P_{\mathrm{acc}} = 90\%) \approx 63$ and are not displayed for visualization purposes.
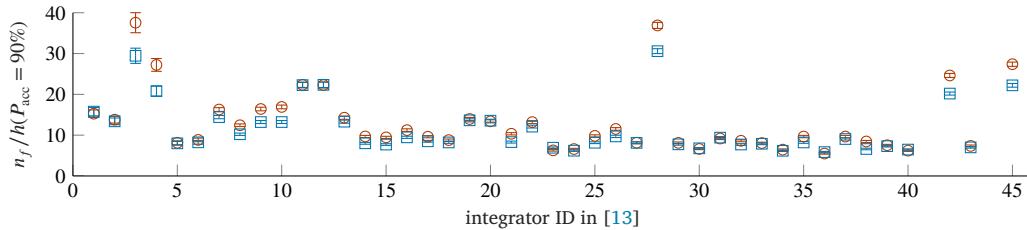
(c) Exact FGIs with coefficients from Table 1 (red circles) vs. Hessian-free FGIs with coefficients from Table 1 (blue squares).

Fig. 5. Schwinger model. Estimation of the required number of force steps per unit trajectory n_f/h to reach $P_{\mathrm{acc}} = 90\%$. Evaluations of the FG-term are counted as two force evaluations. The subfigures show a pair-wise comparison to Hessian-free FGIs with coefficients from Table 1 (blue squares) with (a) exact FGIs with coefficients from [13], (b) Hessian-free FGIs with coefficients from [13], (c) exact FGIs with coefficients from Table 1 (red circles).

5.2.2. Four-dimensional lattice QCD simulations with Wilson fermions

We consider gauge field simulations in lattice quantum chromodynamics (QCD) on a four-dimensional lattice of size $V = T \times L^3$ with lattice spacing a. Here, the Hamiltonian reads

$$\mathcal{H}([\boldsymbol{P}], [\boldsymbol{Q}]) = \frac{1}{2} \sum_{x,\mu} \text{tr}(P_{x,\mu}^2) + S_G([\boldsymbol{Q}]) + S_F([\boldsymbol{Q}]), \tag{78}$$

where $S_G([\boldsymbol{Q}])$ denotes the Wilson gauge action, $S_F([\boldsymbol{Q}])$ denotes the fermion action (for details, see Appendix B). Here, the links $Q_{x,\mu}$, connecting the sites x and $x + a\hat{\mu}$, are elements of the non-Abelian matrix Lie group SU(3) of unitary matrices with determinant one. The scaled momenta $iP_{x,\mu}$ are elements of the associated Lie algebra $\mathfrak{su}(3)$ of traceless and anti-Hermitian matrices and occur in the formulation of the equations of motion. As an initial test in lattice QCD simulations, we consider an ensemble used in [45] generated with two dynamical non-perturbatively $\mathcal{O}(a)$ improved Wilson quarks at a mass equal to half of the physical charm quark. Starting from a thermalized configuration, we computed 100 trajectories for $\tau = 2.0$ and varying step size $h = \tau/N$ for different values of N on a 48×24^3 lattice with gauge coupling $\beta = 5.3$ and hopping parameter $\kappa = 0.1327$ for all Hessian-free FGIs discussed in this paper. For the simulations, we have put all forces on a single time-scale of integration.[4] The results of selected integrators are depicted in Fig. 6. The results emphasize that the use of fourth-order integrators and in particular the Hessian-free FGI (48) results in the most efficient computational process with $n_f \cdot N \approx 34$ to achieve the optimal acceptance rate $\langle P_{\text{acc}} \rangle_{\text{opt}} = 78\%$. In contrast, the most efficient non-gradient scheme (68) demands $n_f \cdot N \approx 38$. Moreover, the Hessian-free FGI (35) is a very efficient choice ($n_f \cdot N \approx 37$). The integrators (48) (Eff ≈ 26.19) and (35) (Eff ≈ 16.96) are not among the most efficient integrators according to the efficiency measure (28). In lattice QCD, however, we are only interested in the region of $0.01 \leq \sigma^2 \leq 1$. In this region, the limiting factor in the step size h usually comes from instabilities [48] so that the performance of the integrators strongly depends on their stability properties. For the non-gradient schemes, a linear stability analysis is already available [49]. First results on the linear stability of FGIs support the claim that Hessian-free FGIs with a larger stability domain are beneficial for application in lattice QCD simulations. Moreover, we emphasize that Hessian-free FGIs are not symplectic and thus will in general have a linear energy drift of size $\mathcal{O}(\tau h^{\max\{4,p\}})$ while exact FGIs are symplectic, resulting in a bounded energy error without any drift. Since the trajectory length τ is rather short, the energy drift is negligible. The linear stability and energy conservation of Hessian-free FGIs is subject of a following paper in progress.
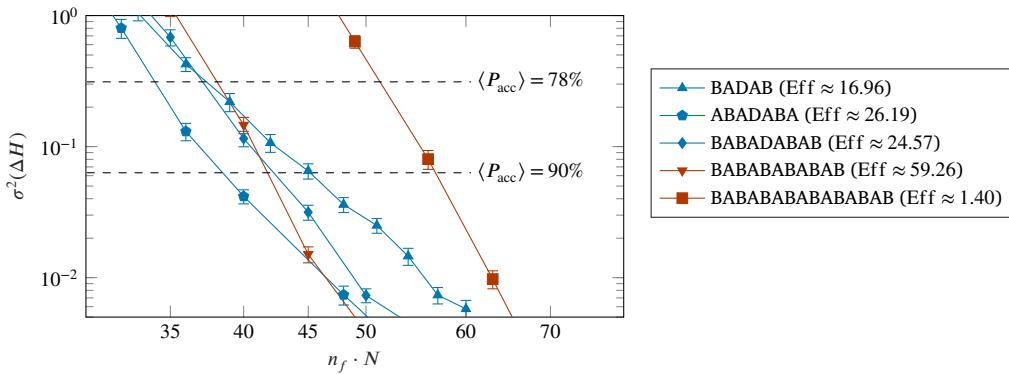

Fig. 6. Lattice QCD. Variance of ΔH vs. number of force evaluations per trajectory $n_f \cdot N$ for a selected number of Hessian-free FGIs, namely the three best performing Hessian-free FGIs (blue lines), as well as the most efficient non-gradient scheme for order $p \in \{4, 6\}$ (red lines). The most efficient non-gradient scheme of order $p = 2$ is not included for visualization purposes as it achieves $\langle P_{\text{acc}} \rangle_{\text{opt}}$ at $n_f \cdot N \approx 140$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

For the best-performing Hessian-free FGI (48) and the most efficient non-gradient integrator (68), we tuned the step size h to achieve an acceptance rate of $P_{\text{acc}} \geq 90\%$, resulting in the step size $h = 0.2$ for both integrators, and computed 2000 trajectories. The results of both setups are summarized in Table 2, including the cost measure

$$\text{cost} = \frac{n_f \cdot N}{P_{\text{acc}} \cdot \tau},$$

as well as estimates of the integrated autocorrelation times of the topological charge Q_0 measured at Wilson flow time t_0, and the Wilson flow reference scale t_0/a^2. For both quantities we observe similar autocorrelation times, despite the lower acceptance rate of the Hessian-free force-gradient integrator. By investigating the cost measure, it turns out that the Hessian-free FGI only demands approximately 84% of the computational cost compared to the non-gradient integrator. Despite the simplicity of the ensemble, the results show the potential of Hessian-free FGIs for lattice QCD simulations. In more complicated ensembles, the performance usually benefits from applying nested integration techniques [46,47] to employ different step sizes for different parts of the forces. The investigation of Hessian-free FGIs inside nested integrators is subject to future research focusing on multiple time stepping techniques.

---

[4] We also performed simulations using nested integrators that employ a smaller step size to the gauge force. At these lattice parameters, however, we were not able to observe any significant improvement in the acceptance probability by applying nested integration techniques, see e.g. [46,47].

Table 2

Lattice QCD. Comparison of the performance of two tuned setups based on 2000 trajectories of length $\tau = 2.0$. Both setups use a step size of $h = 0.2$. The results contain the number of force evaluations per trajectory, the acceptance rate P_{acc}, integrated autocorrelation times $\tau_{\mathrm{int}}(t_0)$ and $\tau_{\mathrm{int}}(Q_0)$ in molecular dynamics units (MDUs), and a measure for the computational cost.

ID	$n_f \cdot N$	P_{acc}	$\tau_{\mathrm{int}}(t_0)$ [MDU]	$\tau_{\mathrm{int}}(Q_0)$ [MDU]	cost
BABABABABAB	50	97.5%	37.40(12.66)	22.91(6.52)	25.64
ABADABA	40	92.3%	28.15(9.24)	22.05(6.31)	21.67

6. Conclusion and outlook

In this work, we generalized the approximation proposed in [19,21], resulting in a new class of Hessian-free FGIs. The new integrators are applicable to separable Hamiltonian systems of the form (2), as well as to general second order ODEs $\ddot{y} = f(y)$. In contrast to exact FGIs [13], the Hessian-free variants do not need for the Hessian of the potential energy $\mathcal{V}$ and replace it by another force evaluation. At the price of additional error terms, that have been derived explicitly in this work, one is able to save costly evaluations of the force-gradient term, e.g. in molecular dynamics simulations. Hessian-free FGIs are time-reversible, volume-preserving and satisfy the closure property, but they are no longer symplectic. Consequently, they do not preserve a nearby shadow Hamiltonian. In general, this results in a linear energy drift of size $\mathcal{O}(th^p)$. A complete classification of Hessian-free FGIs up to $s = 11$ stages shows that some variants undergo only small changes in the time coefficients compared to the respective variants in [13], whereas the time coefficients of other variants change significantly, e.g. if an order reduction appears due to the additional error terms. Numerical tests emphasize the computational efficiency of the proposed numerical integration schemes for different applications, making them a valuable class of integration schemes for a wide range of applications.

The efficiency of higher order integration schemes strongly depends on the system size so that Hessian-free FGIs of order $p > 4$ may become efficient for simulating lattice quantum field theories of larger volumes [50]. Hence a next natural step is to investigate Hessian-free FGIs with $s > 11$ stages, aiming for more efficient sixth-order integrators.

In the context of lattice QCD, the numerical results indicate that the stability properties of the integrators are crucial for an efficient computational process. Thus we will perform a linear stability analysis of Hessian-free and exact FGIs to validate which variants are superior in terms of stability. Furthermore, a refined analysis on the energy conservation of Hessian-free FGIs in lattice QCD will contribute in designing more efficient numerical integration schemes for simulating quantum field theories on the lattice.

The efficiency measure used in this work assumes that the error terms are equally dominant. However, this simplifying assumption does not hold in general. Minimizing a problem-specific weighted norm results in even more efficient integrators [50]. Therefore, we will develop a tuning model that numerically measures the error terms to obtain appropriate weights for the weighted norm. Since we derived the additional error terms for Hessian-free FGIs in this work, this enables tuning procedures similar to [50] and may enhance other existing tuning models, e.g. [44].

In many applications, separable Hamiltonian systems with a potential splitting $\mathcal{H}(p, q) = \mathcal{T}(p) + \sum_{m=1}^{N} \mathcal{V}^{\{m\}}(q)$ occur. Here, the potential part is split with respect to, for example, stiffness, nonlinearity, dynamical behavior, and evaluation cost. In lattice QCD, for example, the fermion part of the action is characterized by slow dynamics and expensive evaluation cost, whereas the gauge part is fast and cheap to evaluate. This multirate potential can be exploited by means of nested integration, see e.g. [22,46,47]. We aim at designing a general framework of nested integration schemes and providing an order theory including explicit formulae of the error terms.

CRediT authorship contribution statement

Kevin Schäfers: Writing – original draft, Validation, Software, Methodology, Investigation, Conceptualization. **Jacob Finkenrath:** Writing – original draft, Validation, Software, Investigation. **Michael Günther:** Writing – review & editing, Supervision, Funding acquisition. **Francesco Knechtli:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Details on the action for the two-dimensional Schwinger model

The action S is given by the sum of the gauge part

$$S_G([\boldsymbol{Q}]) = \beta \sum_{n=1}^{V} (1 - \mathfrak{Re}(\mathbb{P}(n)))$$

summing up over all plaquettes

$$\mathbb{P}(n) = Q_1(n)Q_2(n+\hat{1})Q_1^\dagger(n+\hat{2})Q_2^\dagger(n),$$

scaled by a coupling constant β, and the fermion part

$$S_F([\boldsymbol{Q}]) = \eta^\dagger \left(D^\dagger D\right)^{-1} \eta,$$

with pseudofermion field η and Wilson-Dirac operator

$$D_{n,m} = (2 + m_0)\delta_{n,m} - \frac{1}{2}\sum_{\mu=1}^{2}\left((1 - \sigma_\mu)Q_\mu(n)\delta_{n,m-\hat{\mu}} + (1 + \sigma_\mu)Q_\mu^\dagger(n - \hat{\mu})\delta_{n,m+\hat{\mu}}\right),$$

where m_0 denotes the mass parameter, $\delta_{n,m}$ the Kronecker delta acting on the pseudofermion field via

$$\sum_{m=1}^{V}\delta_{n,m}\eta(m) = \eta(n),$$

and σ_μ the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

which build a basis of the Lie group U(1).

Appendix B. Details on the Hamiltonian for lattice QCD

Any matrix representation $P_{x,\mu}$ can be written as $P_{x,\mu} = p^i T_i$ where the generators T_i, $i = 1, \ldots, 8$, are given by the Gell-Mann matrices [51]. With $\boldsymbol{p}_{x,\mu} := (p_1, \ldots, p_8)^\top$, a straight-forward computation shows that

$$\mathrm{tr}(P_{x,\mu}^2) = \langle \boldsymbol{p}_{x,\mu}, \boldsymbol{p}_{x,\mu}\rangle,$$

holds. Consequently, the kinetic energy in (78) can be written as

$$\mathcal{T}(\boldsymbol{p}) = \frac{1}{2}\sum_{x,\mu}\langle \boldsymbol{p}_{x,\mu}, \boldsymbol{p}_{x,\mu}\rangle$$

which is of the form (1) with $\boldsymbol{M} = \mathrm{Id}$. The action $S = S_G + S_F$ is defined on a four dimensional lattice with volume V and consists of the gauge part

$$S_G([\boldsymbol{Q}]) = \sum_x \sum_{\mu < \nu} \beta \left(1 - \frac{1}{3}\mathfrak{Re}\left(\mathrm{tr}\left(\mathbb{P}_{x,\mu\nu}([\boldsymbol{Q}])\right)\right)\right)$$

with gauge coupling β and plaquettes

$$\mathbb{P}_{x,\mu\nu}([\boldsymbol{Q}]) = Q_{x,\mu}Q_{x+a\hat{\mu}}Q_{x+a\hat{\nu},\mu}^\dagger Q_{x,\nu}^\dagger,$$

and the fermion part

$$S_F([\boldsymbol{Q}]) = \langle \phi_1, (D_{eo}D_{eo}^\dagger + \mu^2)^{-1}\phi_1\rangle + \langle \phi_2, (1 + \mu^2(D_{eo}D_{eo}^\dagger)^{-1})\phi_2\rangle + \sum_x \mathrm{tr}\log D_{x,x}^2$$

with complex-valued pseudofermion fields ϕ_1 and ϕ_2. The fermion part is decomposed using even-odd reduction in combination with one Hasenbusch mass preconditioning term [52] with shift parameter μ, such that the determinant is given by

$$\det D^2 = \frac{\det(D_{eo}D_{eo}^\dagger + \mu^2 \cdot \mathrm{Id})}{\det(\mathrm{Id} + \mu^2(D_{eo}D_{eo}^\dagger)^{-1})} \cdot \prod_x \det D_{x,x}^2$$

with the even-odd reduced Dirac operator $D_{eo} = \mathrm{Id} - D_{e,e}^{-1}D_{e,o}D_{o,o}^{-1}D_{o,e}$ and the block diagonal terms $D_{x,x}$. The clover improved Wilson Dirac operator can be represented as a complex matrix of size $12 \times V$ with next neighbor interactions. For a more detailed discussion, we refer to [53].

Data availability

An extended version of the openQCD package [26] including the proposed class of integrators is opensource and available online at https://github.com/KevinSchaefers/openQCD_force-gradient.

References

[1] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, 2nd edition, Springer Ser. Comput. Math., vol. 31, Springer-Verlag, Berlin, 2006.
[2] H. Goldstein, C. Poole, J. Safko, Classical Mechanics, Addison Wesley, 2002.
[3] D.J. Griffiths, D.F. Schroeter, Introduction to Quantum Mechanics, Cambridge University Press, 2018.
[4] J.W. Gibbs, Elementary Principles in Statistical Mechanics, C. Scribner's Sons, New York, 1902.
[5] S. Duane, A. Kennedy, B.J. Pendleton, D. Roweth, Hybrid Monte Carlo, Phys. Lett. B 195 (2) (1987) 216–222, https://doi.org/10.1016/0370-2693(87)91197-X.
[6] M. Günther, A. Sandu, K. Schäfers, A. Zanna, Symplectic GARK methods for partitioned Hamiltonian systems, arXiv preprint, arXiv:2103.04110, 2023.
[7] K. Schäfers, M. Günther, A. Sandu, Symplectic multirate generalized additive Runge–Kutta methods for Hamiltonian systems, arXiv preprint, arXiv:2306.04389, 2023.
[8] A. Zanna, Discrete variational methods and symplectic generalized additive Runge–Kutta methods, arXiv preprint, arXiv:2001.07185, 2020.
[9] G. Bogfjellmo, H. Marthinsen, High-order symplectic partitioned Lie group methods, Found. Comput. Math. 16 (2016) 493–530, https://doi.org/10.1007/s10208-015-9257-9.

[10] M. Suzuki, Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations, Phys. Lett. A 146 (6) (1990) 319–323, https://doi.org/10.1016/0375-9601(90)90962-N.

[11] H. Yoshida, Construction of higher order symplectic integrators, Phys. Lett. A 150 (5–7) (1990) 262–268, https://doi.org/10.1016/0375-9601(90)90092-3.

[12] I.P. Omelyan, I.M. Mryglod, R. Folk, Construction of high-order force-gradient algorithms for integration of motion in classical and quantum systems, Phys. Rev. E 66 (2002) 026701, https://doi.org/10.1103/PhysRevE.66.026701.

[13] I. Omelyan, I. Mryglod, R. Folk, Symplectic analytically integrable decomposition algorithms: classification, derivation, and application to molecular dynamics, quantum and celestial mechanics simulations, Comput. Phys. Commun. 151 (3) (2003) 272–314, https://doi.org/10.1016/S0010-4655(02)00754-3.

[14] A.D. Kennedy, M.A. Clark, P.J. Silva, Force gradient integrators, PoS LATTICE 2009 (2010), https://doi.org/10.22323/1.091.0021.

[15] M. Takahashi, M. Imada, Monte Carlo calculation of quantum systems. II. Higher order correction, J. Phys. Soc. Jpn. 53 (11) (1984) 3765–3769, https://doi.org/10.1143/JPSJ.53.3765.

[16] G. Rowlands, A numerical algorithm for Hamiltonian systems, J. Comput. Phys. 97 (1) (1991) 235–239, https://doi.org/10.1016/0021-9991(91)90046-N.

[17] M.A. López-Marcos, J.M. Sanz-Serna, R.D. Skeel, Explicit symplectic integrators using Hessian–vector products, SIAM J. Sci. Comput. 18 (1) (1997) 223–238, https://doi.org/10.1137/S1064827595288085.

[18] M. Mönch, N. Marheineke, Commutator-based operator splitting for linear port-Hamiltonian systems, Appl. Numer. Math. 210 (2025) 25–38, https://doi.org/10.1016/j.apnum.2024.12.007.

[19] J. Wisdom, M. Holman, J. Touma, Symplectic correctors, in: Integration Algorithms and Classical Mechanics, Amer. Math. Soc., 1996, pp. 217–244.

[20] E. Hairer, R.I. McLachlan, R.D. Skeel, On energy conservation of the simplified Takahashi-Imada method, ESAIM: Math. Model. Numer. Anal. 43 (4) (2009) 631–644, https://doi.org/10.1051/m2an/2009019.

[21] H. Yin, R.D. Mawhinney, Improving DWF simulations: the force gradient integrator and the Möbius accelerated DWF solver, PoS LATTICE 2011 (2012), https://doi.org/10.22323/1.139.0051.

[22] D. Shcherbakov, M. Ehrhardt, J. Finkenrath, M. Günther, F. Knechtli, M. Peardon, Adapted nested force-gradient integrators: the Schwinger model case, Commun. Comput. Phys. 21 (4) (2017) 1141–1153, https://doi.org/10.4208/cicp.OA-2016-0048.

[23] C. Jung, C. Kelly, R.D. Mawhinney, D.J. Murphy, Domain wall fermion QCD with the exact one flavor algorithm, Phys. Rev. D 97 (5) (2018) 054503, https://doi.org/10.1103/PhysRevD.97.054503.

[24] J. Finkenrath, et al., Twisted mass gauge ensembles at physical values of the light, strange and charm quark masses, PoS LATTICE 2021 (2022) 284, https://doi.org/10.22323/1.396.0284.

[25] J. Finkenrath, Review on algorithms for dynamical fermions, PoS LATTICE 2022 (2023) 227, https://doi.org/10.22323/1.430.0227.

[26] M. Lüscher, S. Schaefer, J. Bulava, I. Campos, A. Rago, L. Del Debbio, L. Giusti, B. Leder, F. Palombi, openQCD v2.4, https://luscher.web.cern.ch/luscher/openQCD/, 2022.

[27] M. Lüscher, S. Schaefer, Lattice QCD with open boundary conditions and twisted-mass reweighting, Comput. Phys. Commun. 184 (3) (2013) 519–528, https://doi.org/10.1016/j.cpc.2012.10.003.

[28] B.C. Hall, Lie Groups, Lie Algebras, and Representations, Springer New York, New York, 2013, pp. 333–366, Ch. 16.

[29] R.I. McLachlan, G.R.W. Quispel, Splitting methods, Acta Numer. 11 (2002) 341–434, https://doi.org/10.1017/S0962492902000053.

[30] A.D. Kennedy, P.J. Silva, M.A. Clark, Shadow Hamiltonians, Poisson brackets, and gauge theories, Phys. Rev. D 87 (2013) 034511, https://doi.org/10.1103/PhysRevD.87.034511.

[31] F. Knechtli, M. Günther, M. Peardon, Lattice Quantum Chromodynamics: Practical Essentials, SpringerBriefs in Physics, Springer, Dordrecht, 2017.

[32] P.E. Crouch, R. Grossman, Numerical integration of ordinary differential equations on manifolds, J. Nonlinear Sci. 3 (1) (1993) 1–33, https://doi.org/10.1007/BF02429858.

[33] É. Forest, Sixth-order Lie group integrators, J. Comp. Physiol. 99 (2) (1992) 209–213, https://doi.org/10.1016/0021-9991(92)90203-B.

[34] L. Verlet, Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules, Phys. Rev. 159 (1967) 98–103, https://doi.org/10.1103/PhysRev.159.98.

[35] R.I. McLachlan, On the numerical integration of ordinary differential equations by symmetric composition methods, SIAM J. Sci. Comput. 16 (1) (1995) 151–168, https://doi.org/10.1137/0916010.

[36] M. Suzuki, New scheme of hybrid exponential product formulas with applications to quantum Monte-Carlo simulations, in: D.P. Landau, K.K. Mon, H.-B. Schüttler (Eds.), Computer Simulation Studies in Condensed-Matter Physics VIII, Springer Berlin Heidelberg, Berlin, Heidelberg, 1995, pp. 169–174.

[37] S.A. Chin, Symplectic integrators from composite operator factorizations, Phys. Lett. A 226 (6) (1997) 344–348, https://doi.org/10.1016/S0375-9601(97)00003-0.

[38] E. Forest, R.D. Ruth, Fourth-order symplectic integration, Phys. D 43 (1) (1990) 105–117, https://doi.org/10.1016/0167-2789(90)90019-L.

[39] W. Kahan, R.-C. Li, Composition constants for raising the orders of unconventional schemes for ordinary differential equations, Math. Comput. 66 (219) (1997) 1089–1099, https://doi.org/10.1090/S0025-5718-97-00873-9.

[40] L. Burckhardt, et al., Ahnerts Kalender für Sternfreunde 1994, Astron. Nachr. 314 (1993) 398.

[41] T. Takaishi, Choice of integrator in the hybrid Monte Carlo algorithm, Comput. Phys. Commun. 133 (1) (2000) 6–17, https://doi.org/10.1016/S0010-4655(00)00161-2.

[42] S. Gupta, A. Irbäc, F. Karsch, B. Petersson, The acceptance probability in the hybrid Monte Carlo method, Phys. Lett. B 242 (3) (1990) 437–443, https://doi.org/10.1016/0370-2693(90)91790-I.

[43] N. Christian, K. Jansen, K. Nagai, B. Pollakowski, Scaling test of fermion actions in the Schwinger model, Nucl. Phys. B 739 (1) (2006) 60–84, https://doi.org/10.1016/j.nuclphysb.2006.01.029.

[44] J.C. Osborn, Tuning HMC parameters with gradients, PoS LATTICE 2023 (2023) 023, arXiv:2402.04976.

[45] F. Knechtli, T. Korzec, M. Peardon, J.A. Urrea-Niño, Optimizing creation operators for charmonium spectroscopy on the lattice, Phys. Rev. D 106 (2022) 034501, https://doi.org/10.1103/PhysRevD.106.034501.

[46] J. Sexton, D. Weingarten, Hamiltonian evolution for the hybrid Monte Carlo algorithm, Nucl. Phys. B 380 (1992) 665–677, https://doi.org/10.1016/0550-3213(92)90263-B.

[47] C. Urbach, K. Jansen, A. Shindler, U. Wenger, Hmc algorithm with multiple time scale integration and mass preconditioning, Comput. Phys. Commun. 174 (2) (2006) 87–98, https://doi.org/10.1016/j.cpc.2005.08.006.

[48] B. Joó, B. Pendleton, A.D. Kennedy, A.C. Irving, J.C. Sexton, S.M. Pickles, S.P. Booth, Instability in the molecular dynamics step of a hybrid Monte Carlo algorithm in dynamical fermion lattice QCD simulations, Phys. Rev. D 62 (2000) 114501, https://doi.org/10.1103/PhysRevD.62.114501.

[49] S. Blanes, F. Casas, A. Murua, On the linear stability of splitting methods, Found. Comput. Math. 8 (3) (2008) 357–393, https://doi.org/10.1007/s10208-007-9007-8.

[50] T. Takaishi, P. de Forcrand, Testing and tuning symplectic integrators for the hybrid Monte Carlo algorithm in lattice QCD, Phys. Rev. E 73 (2006) 036706, https://doi.org/10.1103/PhysRevE.73.036706.

[51] H. Georgi, Lie Algebras in Particle Physics: from Isospin to Unified Theories, Taylor & Francis, 2000.

[52] M. Hasenbusch, Speeding up the hybrid Monte Carlo algorithm for dynamical fermions, Phys. Lett. B 519 (2001) 177–182, https://doi.org/10.1016/S0370-2693(01)01102-9.

[53] A. Frommer, K. Kahl, S. Krieg, B. Leder, M. Rottmann, Adaptive aggregation-based domain decomposition multigrid for the lattice Wilson–Dirac operator, SIAM J. Sci. Comput. 36 (4) (2014) A1581–A1608, https://doi.org/10.1137/130919507.