## Research

**Author for correspondence:**
Xiaocheng Shang
e-mail: x.shang.1@bham.ac.uk

# Structure-preserving integrators for dissipative systems based on reversible–irreversible splitting

## Xiaocheng Shang[1] and Hans Christian Öttinger[2]

[1]School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[2]Department of Materials, Polymer Physics, ETH Zürich, Leopold-Ruzicka-Weg 4, Zürich CH-8093, Switzerland

XS, 0000-0001-8376-6498; HCÖ, 0000-0003-0096-3176

We study the optimal design of numerical integrators for dissipative systems, for which there exists an underlying thermodynamic structure known as GENERIC (general equation for the nonequilibrium reversible–irreversible coupling). We present a framework to construct structure-preserving integrators by splitting the system into reversible and irreversible dynamics. The reversible part, which is often degenerate and reduces to a Hamiltonian form on its symplectic leaves, is solved by using a symplectic method (e.g. Verlet) with degenerate variables being left unchanged, for which an associated modified Hamiltonian (and subsequently a modified energy) in the form of a series expansion can be obtained by using backward error analysis. The modified energy is then used to construct a modified friction matrix associated with the irreversible part in such a way that a modified degeneracy condition is satisfied. The modified irreversible dynamics can be further solved by an explicit midpoint method if not exactly solvable. Our findings are verified by various numerical experiments, demonstrating the superiority of structure-preserving integrators over alternative schemes in terms of not only the accuracy control of both energy conservation and entropy production but also the preservation of the conformal symplectic structure in the case of linearly damped systems.

# 1. Introduction

As an introduction to this article on structure-preserving integrators for dissipative systems, we first summarize the state-of-the-art of the literature and then provide a description of the GENERIC formulation and its properties. The introduction ends with an outline of the article.

## (a) State-of-the-art in structure-preserving integrators

In the last few decades, considerable effort has been devoted to developing structure-preserving integrators for Hamiltonian systems. It has been demonstrated that the so-called symplectic integrators, which preserve the symplectic structure, have superior long-time behaviour compared to their non-symplectic counterparts, and should be preferred in practice [1–3]. On the other hand, there has been growing interest in designing appropriate numerical methods for gradient flows [4–8] that respect their underlying properties. In contrast to the symplectic structure, the conformal symplectic structure [9–14] for Hamiltonian systems that are perturbed by a linear damping (which can be thought of as a special case of the Rayleigh dissipation) has been less studied. It is also worth mentioning that variational integrators [15] and specialized Runge–Kutta methods [16] have also been used to solve dissipative systems. It turns out that thermodynamically admissible evolution equations for nonequilibrium systems have a more general (including an additional variable known as entropy) and well-defined structure known as GENERIC (general equation for the nonequilibrium reversible–irreversible coupling) [17–20], which possesses the following distinct features:

   (i) conservation of the total energy;
  (ii) separation of the reversible and irreversible dynamics;
 (iii) the reversible dynamics preserves a Poisson structure;
 (iv) entropy production is unaffected by the reversible dynamics;
  (v) non-negative entropy production rate.

## (b) GENERIC formulation

The GENERIC formulation of the time evolution for nonequilibrium systems is given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = L\frac{\partial E}{\partial x} + M\frac{\partial S}{\partial x}, \tag{1.1}$$

where $x$ is the set of independent variables required to describe a given nonequilibrium system, $E$ and $S$ represent, respectively, the total energy and entropy as functions of the independent variables $x$, and $L$ and $M$ denote the antisymmetric Poisson matrix and the positive semidefinite (symmetric) friction matrix, respectively. Note that both $L$ and $M$ can also depend on the independent variables $x$ so that the fundamental time evolution equation (1.1) could be highly nonlinear. We also point out that $\partial/\partial x$ in (1.1) simply implies the partial derivative although it typically denotes the functional derivative when $x$ is a function/field. Moreover, (1.1) is supplemented by two degeneracy conditions

$$L\frac{\partial S}{\partial x} = 0 \tag{1.2}$$

and

$$M\frac{\partial E}{\partial x} = 0. \tag{1.3}$$

Equations (1.2)–(1.3) indicate the conservation of the entropy by the reversible dynamics (i.e. the $L$ contribution) and the conservation of the total energy in a closed system by the irreversible dynamics (i.e. the $M$ contribution), respectively. Note that 'reversible' and 'irreversible' dynamics (in thermodynamics) are simply the names of the two fundamental contributions to the time evolution equation (1.1), and should not be confused with similar terms in other subjects. The rank

of $M$ has the interpretation of the number of dissipative processes taking place in the system. (See more discussions on the formulation of the GENERIC framework in [17–20].)

The usefulness and maturity of the GENERIC framework have been illustrated in a very large number of successful applications in a wide range of areas in appendix E of [19] (see also a most recent review of [21] and references therein). In particular, despite its simple form, we believe that the irreversible dynamics in (1.1) is the most general form of meaningful irreversible equations in nonequilibrium thermodynamics—it is a belief based on both a very large variety of successful examples and statistical mechanics, so that it can be called knowledge (in particular, as this belief is widely accepted in the nonequilibrium thermodynamics community).

In order to further demonstrate the general properties of $L$ and $M$, the respective Poisson and dissipative brackets are often adopted

$$\{\mathcal{A}, \mathcal{B}\} = \frac{\partial \mathcal{A}}{\partial x} \cdot L \frac{\partial \mathcal{B}}{\partial x} \tag{1.4}$$

and

$$[\mathcal{A}, \mathcal{B}] = \frac{\partial \mathcal{A}}{\partial x} \cdot M \frac{\partial \mathcal{B}}{\partial x}, \tag{1.5}$$

where $\mathcal{A}$ and $\mathcal{B}$ are sufficiently regular (and real-valued) functions of the independent variables $x$. With the help of the two brackets and the chain rule, the time evolution equation of an arbitrary function $\mathcal{A}$ can then be written as

$$\frac{\mathrm{d}\mathcal{A}}{\mathrm{d}t} = \{\mathcal{A}, E\} + [\mathcal{A}, S]. \tag{1.6}$$

More specifically, the Poisson bracket (1.4) inherits the antisymmetry of $L$,

$$\{\mathcal{A}, \mathcal{B}\} = -\{\mathcal{B}, \mathcal{A}\}, \tag{1.7}$$

and satisfies the Leibniz rule,

$$\{\mathcal{A}\mathcal{B}, \mathcal{C}\} = \mathcal{A}\{\mathcal{B}, \mathcal{C}\} + \mathcal{B}\{\mathcal{A}, \mathcal{C}\}, \tag{1.8}$$

where $\mathcal{C}$ is another arbitrary sufficiently regular (and real-valued) function of the independent variables $x$. In addition, the Poisson bracket is required to satisfy the Jacobi identity,

$$\{\mathcal{A}, \{\mathcal{B}, \mathcal{C}\}\} + \{\mathcal{B}, \{\mathcal{C}, \mathcal{A}\}\} + \{\mathcal{C}, \{\mathcal{A}, \mathcal{B}\}\} = 0. \tag{1.9}$$

The dissipative bracket (1.5) inherits the symmetry of $M$,

$$[\mathcal{A}, \mathcal{B}] = [\mathcal{B}, \mathcal{A}], \tag{1.10}$$

and also satisfies the Leibniz rule,

$$[\mathcal{A}\mathcal{B}, \mathcal{C}] = \mathcal{A}[\mathcal{B}, \mathcal{C}] + \mathcal{B}[\mathcal{A}, \mathcal{C}]. \tag{1.11}$$

The positive semidefinite nature of $M$ leads to the non-negativeness condition

$$[\mathcal{A}, \mathcal{A}] \geq 0, \tag{1.12}$$

which implies the second law of nonequilibrium thermodynamics (i.e. the entropy production rate is always non-negative),

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\partial S}{\partial x} \cdot M \frac{\partial S}{\partial x} = [S, S] \geq 0. \tag{1.13}$$

This article addresses the long-standing challenge of how to preserve the underlying structures when numerically discretizing GENERIC systems in practice. Although in recent years, this topic has attracted increasing attention [22–25], to the best of our knowledge, there are no such numerical integrators in the literature. Unlike common approaches that are based on exact energy conservation, we propose in this article a framework to construct structure-preserving integrators for dissipative systems, i.e. GENERIC integrators (also known as metriplectic integrators [26–30] in the mathematical literature), based on splitting the reversible and irreversible dynamics. The topic of structure-preserving integrators for GENERIC/metriplectic systems is the counterpart and generalization of the theory of symplectic integrators for Hamiltonian systems.

## (c) Outline of the article

The rest of the article is organized as follows. We give specific definitions of GENERIC integrators and discuss their requirements in numerical discretizations in §2. In §3, we propose a framework to construct split GENERIC integrators based on reversible and irreversible splitting, the generality of the framework is demonstrated in examples of linearly damped systems in §3a as well as in a more challenging (and fully coupled) case of two gas containers exchanging heat and volume in §3b. Section 4 presents various numerical experiments to investigate the performance of the two split GENERIC integrators introduced in this article. Our findings are summarized in §5.

# 2. Definitions of GENERIC integrators

In this section, we provide the definitions of GENERIC integrators and discuss their requirements when numerically discretizing a system in practice.

## (a) Full GENERIC integrators

We recall the definition of (full) GENERIC integrators given in [20]. Analogous to the definition of symplectic integrators for Hamiltonian dynamics [31], a mapping, $x_0 \mapsto x_h$, is said to be a full GENERIC integrator if it corresponds to a continuous time evolution of a modified GENERIC system

$$\frac{\mathrm{d}x}{\mathrm{d}t} = L \frac{\partial \tilde{E}_h}{\partial x} + \tilde{M}_h \frac{\partial S}{\partial x}, \tag{2.1}$$

where $\tilde{E}_h$ and $\tilde{M}_h$ represent the modified energy and friction matrix associated with the integrator, respectively, satisfying a modified degeneracy condition:

$$\tilde{M}_h \frac{\partial \tilde{E}_h}{\partial x} = 0. \tag{2.2}$$

That is, given initial conditions $x(0) = x_0$, the analytical solution of (2.1), $x(t)$, should agree with what we obtain from the integrator at time $h$, i.e. $x(h) = x_h$. A full GENERIC integrator $x \mapsto x_h$, which can be thought of as the formal solution of (2.1), possesses the following structure:

$$x_h = \exp \left\{ h \left( L \frac{\partial \tilde{E}_h}{\partial x} + \tilde{M}_h \frac{\partial S}{\partial x} \cdot \frac{\partial}{\partial x} \right) \right\} x. \tag{2.3}$$

Similar to symplectic integrators for Hamiltonian dynamics, the modified energy, $\tilde{E}_h$, is strictly conserved by a GENERIC integrator. The physical energy $E$ is expected to remain close to the modified energy, $\tilde{E}_h$, even for long integration periods. Additionally, the modified friction matrix, $\tilde{M}_h$, should not introduce any additional dissipative processes not present in the original matrix $M$. We point out that full GENERIC integrators may only be available in special cases, for instance, a full GENERIC integrator in the case of a damped harmonic oscillator, where analytical solutions of the GENERIC system can be obtained, was proposed and discussed in [20]. However, it should be noted that it is highly unlikely that analytical solutions would be available for general GENERIC systems. (Nevertheless, it might be eventually possible to recognize a full GENERIC integrator without exact solutions.) Therefore, in what follows, we introduce a framework to construct 'split' GENERIC integrators.

## (b) Split GENERIC integrators

Inspired by recent developments on splitting methods [3,32–39], we consider to split the reversible and irreversible parts of the GENERIC system in such a way that the reversible dynamics, which is often degenerate but possesses a Hamiltonian form on its symplectic leaves, can be integrated by using a symplectic method (e.g. Verlet) with degenerate variables being left unchanged, while

the irreversible part (gradient flow) can be solved in such a way that as many structure elements as possible can be preserved (see more references on the challenging task of structure preservation on manifolds in [4–8,40,41]).

An interesting question for the split GENERIC integrators is: under what conditions do a modified energy and an associated friction matrix, satisfying the modified degeneracy condition (2.2), exist? If they exist, how much do we know about their respective forms? GENERIC integrators share some common features of GENERIC systems discussed at the beginning of this article, which can also be thought of as the requirements for GENERIC integrators. Denoting the Jacobian matrix of the independent variables $x$ as $\Omega$, we have

(i) preservation of the Poisson structure for the reversible dynamics: $\Omega(x_0)L(x_0)\Omega^{\mathrm{T}}(x_0) = L(x_h)$;
(ii) non-negative entropy production rate: $S(x_h) \geq S(x_0)$;
(iii) the modified degeneracy condition (2.2) is satisfied with the other (1.2) being unchanged;
(iv) preservation of the rank of the friction matrix: $\mathrm{rank}(\tilde{M}_h) = \mathrm{rank}(M)$.

Note that the satisfaction of the modified degeneracy condition (2.2) may be based on a truncated modified energy as discussed in §3a(ii). As pointed out in [42], it has been proved in [43] that there cannot exist an integrator for 'non-integrable' Hamiltonian dynamics that preserves both the symplectic (Poisson) structure and the energy (Hamiltonian). In fact, it has been discussed in [44] that the preservation of either property has its advantages and disadvantages. While previous attempts to construct structure-preserving integrators for dissipative systems have been relying on the exact conservation of energy (i.e. the energy-conserving discrete gradient methods [45–47], see more discussions in §4a(ii)), there is no obvious reason why integrators that preserve the Poisson structure for the reversible dynamics should be ignored.

# 3. Construction of split GENERIC integrators based on reversible–irreversible splitting

In this section, we discuss the construction of GENERIC integrators based on splitting the reversible and irreversible parts of the system. In order to satisfy the modified degeneracy condition (2.2), we explore the possibility of adjusting the irreversible part using a modified friction matrix that corresponds to a modified energy associated with the symplectic integrator used for the reversible part.

## (a) Linearly damped systems

We first consider a linearly damped system that possesses a natural GENERIC structure (1.1) with independent variables $x = (q, p, S)$, where $q$ and $p$ represent the position and momentum of the particle, respectively, and $S$ is the entropy of the surrounding thermal bath. While $S$ is an independent variable and thus $\partial S/\partial x = (0, 0, 1)$, the total energy of the GENERIC system is given by

$$E(q, p, S) = H(q, p) + TS = \frac{p^2}{2m} + U(q) + TS, \tag{3.1}$$

where $H(q, p)$ represents the Hamiltonian of the particle, $U(q)$ denotes the potential energy, and $TS$ is the energy of the thermal bath. Given the antisymmetric Poisson matrix

$$L = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{3.2}$$

and the positive semidefinite (symmetric) friction matrix

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \gamma mT & -\gamma p \\ 0 & -\gamma p & \dfrac{\gamma p^2}{mT} \end{pmatrix} = yy^{\mathrm{T}} \quad \text{and} \quad y = \sqrt{\dfrac{\gamma}{mT}} \begin{pmatrix} 0 \\ mT \\ -p \end{pmatrix}, \tag{3.3}$$

where constant parameters $m$, $\gamma$, and $T$ represent the mass of the particle, the damping rate, and the constant temperature of the thermal bath, respectively, the equations of motion of the GENERIC system can be written as

$$\dot{q} = \frac{p}{m}, \tag{3.4}$$

$$\dot{p} = F(q) - \gamma p \tag{3.5}$$

and

$$\dot{S} = \frac{\gamma p^2}{mT}, \tag{3.6}$$

where $F(q) = -U'(q)$ is the conservative force. Note that in this particular case the symplectic leaves are given by the $(q, p)$ subsystem within the reversible dynamics for constant entropy $S$.

## (i) The YBABY method

Following the discussions in §2b, we suggest to split the GENERIC system (3.4)–(3.6) into reversible and irreversible parts,

$$\mathrm{d} \begin{bmatrix} q \\ p \\ S \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial E}{\partial p} \\ 0 \\ 0 \end{bmatrix}}_{A} \mathrm{d}t + \underbrace{\begin{bmatrix} 0 \\ -\dfrac{\partial E}{\partial q} \\ 0 \end{bmatrix}}_{B} \mathrm{d}t + \underbrace{\begin{bmatrix} 0 \\ -\gamma p \\ \dfrac{\gamma p^2}{mT} \end{bmatrix}}_{Y} \mathrm{d}t. \tag{3.7}$$

Moreover, we can always use a symplectic method (e.g. Verlet) for the reversible dynamics on its symplectic leaves (this is possible in the setting of linearly damped systems (3.4)–(3.6) where $S$ is an independent variable),

$$\mathrm{d} \begin{bmatrix} q \\ p \\ S \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial E}{\partial p} \\ 0 \\ 0 \end{bmatrix}}_{A} \mathrm{d}t + \underbrace{\begin{bmatrix} 0 \\ -\dfrac{\partial E}{\partial q} \\ 0 \end{bmatrix}}_{B} \mathrm{d}t, \tag{3.8}$$

while for linearly damped systems with the total energy (3.1) the irreversible dynamics

$$\mathrm{d} \begin{bmatrix} q \\ p \\ S \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ -\gamma p \\ \dfrac{\gamma p^2}{mT} \end{bmatrix}}_{Y} \mathrm{d}t, \tag{3.9}$$

is exactly solvable (with $q$ being left unchanged)

$$p_h = \exp(-\gamma h)\, p \tag{3.10}$$

and

$$S_h = S + \frac{\gamma p^2}{mT} \int_0^h \exp(-2\gamma t)\, \mathrm{d}t = S + \frac{p^2}{2mT} \left[ 1 - \exp(-2\gamma h) \right]. \tag{3.11}$$

Therefore, we can apply the Verlet method to integrate the reversible part (3.8)

$$\mathrm{e}^{h\hat{\mathcal{L}}_{\text{Verlet}}} = \mathrm{e}^{(h/2)\mathcal{L}_{\text{B}}} \mathrm{e}^{h\mathcal{L}_{\text{A}}} \mathrm{e}^{(h/2)\mathcal{L}_{\text{B}}}, \tag{3.12}$$

and then further split the exact solver (3.10)–(3.11), $e^{h\mathcal{L}_Y}$, for the irreversible part (3.9) to composite a symmetric splitting method, termed 'YBABY', as

$$e^{h\hat{\mathcal{L}}_{YBABY}} = e^{(h/2)\mathcal{L}_Y} e^{h\mathcal{L}_{Verlet}} e^{(h/2)\mathcal{L}_Y} = e^{(h/2)\mathcal{L}_Y} e^{(h/2)\mathcal{L}_B} e^{h\mathcal{L}_A} e^{(h/2)\mathcal{L}_B} e^{(h/2)\mathcal{L}_Y}, \tag{3.13}$$

where $\exp(h\mathcal{L}_f)$ denotes the phase space propagator associated with the corresponding vector field $f$, with $\mathcal{L}_f$ being the corresponding generator. The generators for each part of the GENERIC system may be written out as follows:

$$\mathcal{L}_A = \frac{p}{m} \cdot \nabla_q, \tag{3.14}$$

$$\mathcal{L}_B = F(q) \cdot \nabla_p \tag{3.15}$$

and

$$\mathcal{L}_Y = -\gamma p \cdot \nabla_p + \frac{\gamma p^2}{mT} \cdot \nabla_S. \tag{3.16}$$

Thus, the generator for the GENERIC system can be written as $\mathcal{L}_{GENERIC} = \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_Y$. The integration steps of the YBABY method read:

$$p^{n+1/4} = \exp\left(-\frac{\gamma h}{2}\right) p^n, \tag{3.17}$$

$$S^{n+1/2} = S^n + [p^n]^2 \left[1 - \exp(-\gamma h)\right] / (2mT), \tag{3.18}$$

$$p^{n+2/4} = p^{n+1/4} + \left(\frac{h}{2}\right) F(q^n), \tag{3.19}$$

$$q^{n+1} = q^n + hm^{-1} p^{n+2/4}, \tag{3.20}$$

$$p^{n+3/4} = p^{n+2/4} + \left(\frac{h}{2}\right) F(q^{n+1}), \tag{3.21}$$

$$p^{n+1} = \exp\left(-\frac{\gamma h}{2}\right) p^{n+3/4} \tag{3.22}$$

and

$$S^{n+1} = S^{n+1/2} + \left[p^{n+3/4}\right]^2 \left[1 - \exp(-\gamma h)\right] / (2mT). \tag{3.23}$$

The order of convergence of a splitting method can be determined by using the Baker–Campbell–Hausdorff formula [1–3]. For general operators $A$ and $B$, we have

$$e^{hA} e^{hB} = e^{hZ_1}, \tag{3.24}$$

where

$$Z_1 = A + B + \frac{h}{2} \langle A, B \rangle + O(h^2), \tag{3.25}$$

with $\langle A, B \rangle = AB - BA$ being the commutator. Subsequently, we can work out

$$e^{(h/2)B} e^{hA} e^{(h/2)B} = e^{hZ_2}, \tag{3.26}$$

where

$$Z_2 = A + B + O(h^2). \tag{3.27}$$

Therefore, a symmetric splitting typically gives second-order convergence whereas a non-symmetric one is generally first order. One can then obtain the associated operator of the YBABY method

$$\hat{\mathcal{L}}_{YBABY} = \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_Y + O(h^2), \tag{3.28}$$

which indicates formally second-order convergence for the YBABY method (3.13). Note that the order of convergence can also be demonstrated by using the Taylor series expansion for the solutions, but the procedure is often tedious. Note also that in principle higher-order methods can also be constructed, as in Hamiltonian dynamics [48], by suitably composing the operators.

We would also like to point out that while all three subsystems can be solved exactly in linearly damped systems, in cases where the irreversible part is not exactly solvable (see the example of

two gas containers exchanging heat and volume in §3b) it is important to solve the irreversible part by using a numerical method that is at least second order so that an overall second-order convergence is expected. Alternatively, one could solve the irreversible part by using a numerical method, which could be first order (e.g. the Euler method), and its adjoint method for half a step each, it can be shown that the resulting YBABY$^{\dagger}$ method is self-adjoint (or *symmetric*) and typically has even order (see more discussions in [1–3]). However, such a method could become implicit, for instance, the adjoint method of the Euler method is the implicit backward Euler method.

In the case of $\gamma = 0$, the YBABY method reduces to the Verlet method with degenerate variable $S$ being constant, which is a well-known symplectic method that preserves the Poisson structure for the reversible dynamics [1–3]. Therefore, in order to guarantee the preservation of the Poisson structure for the reversible dynamics, in what follows, we will apply the Verlet method for the reversible part, unless otherwise stated.

For linearly damped systems, it has been demonstrated in [12,14] that numerical methods that preserve the underlying 'conformal symplectic' structure [11] are advantageous over alternative schemes. Moreover, high-order conformal symplectic and ergodic schemes for stochastic Langevin equation have also been investigated [13].

**Definition 3.1.** A numerical method is said to be conformal symplectic if the symplectic two form decays exponentially with a constant decay rate, i.e.

$$\mathrm{d}q_h \wedge \mathrm{d}p_h = \mathrm{e}^{-Kh}\mathrm{d}q \wedge \mathrm{d}p, \tag{3.29}$$

where $\wedge$ represents the wedge product and $K > 0$ is the constant decay rate. Similarly, a numerical method is said to be symplectic if the symplectic two form is preserved, i.e.

$$\mathrm{d}q_h \wedge \mathrm{d}p_h = \mathrm{d}q \wedge \mathrm{d}p. \tag{3.30}$$

We point out that if the prefactor in front of $\mathrm{d}q \wedge \mathrm{d}p$ is initially not in an exponential form, we can always rewrite it into an exponential form as long as it is a constant value between zero and one. Following [12,13], we can show that the YBABY method (3.13) is conformal symplectic:

$$\mathrm{d}q^{n+1} \wedge \mathrm{d}p^{n+1} = \mathrm{e}^{-\gamma h/2}\mathrm{d}q^{n+1} \wedge \mathrm{d}p^{n+2/4},$$
$$= \mathrm{e}^{-\gamma h/2}\mathrm{d}q^n \wedge \mathrm{d}p^{n+2/4},$$
$$= \mathrm{e}^{-\gamma h}\mathrm{d}q^n \wedge \mathrm{d}p^n. \tag{3.31}$$

in which case the decay rate is the physical damping rate, i.e. $K = \gamma$.

We have so far verified the second-order convergence for the YBABY method, and its preservation of the Poisson structure for the reversible dynamics as well as the conformal symplecticity. However, it is unclear under what conditions there exist a modified energy and an associated friction matrix as in (2.2). To this end, in what follows, we modify the irreversible part of the system as discussed at the beginning of this section.

## (ii) The mYBABY method

It is well known that if a symplectic method is used for the reversible dynamics (3.8), there exists a modified Hamiltonian, $\tilde{H}_h$, in the form of a (typically infinite) series expansion obtained by using backward error analysis [49], which is exactly preserved by the symplectic integrator [1–3]. In the example of the Verlet method, the modified Hamiltonian is given by

$$\tilde{H}_h = \frac{p^2}{2m} + U(q) + h^2 \left( \frac{U''(q)p^2}{12m^2} - \frac{[U'(q)]^2}{24m} \right) + O(h^4). \tag{3.32}$$

In order to identify a modified energy conserved by a GENERIC integrator, we can replace the original energy $E$ (3.1) by a modified energy, $\tilde{E}_h = \tilde{H}_h + TS$, and then try to explore whether we can construct an associated friction matrix, $\tilde{M}_h$, in such a way that the modified degeneracy condition (2.2) is satisfied. However, it is unlikely that we can find such a friction matrix due to the

infiniteness of the series expansion (and often complicated higher-order terms) in the modified energy. Nevertheless, we can truncate the series expansion of the modified energy to certain order in practice, which will introduce some perturbations to the modified energy. For instance, we can use the Verlet method for the reversible part, and then truncate the modified energy up to second order, introducing a perturbation of order four to the modified energy, to obtain

$$\tilde{E}_h = \frac{p^2}{2m} + U(q) + TS + h^2 \left( \frac{U''(q)p^2}{12m^2} - \frac{[U'(q)]^2}{24m} \right). \tag{3.33}$$

Subsequently, we can construct the associated modified friction matrix in the fashion of backward error analysis [1,2,49]:

$$\tilde{M}_h = \tilde{y}_h \tilde{y}_h^{\mathrm{T}}, \tag{3.34}$$

where $\tilde{y}_h$ is assumed to be a truncated series expansion up to second order with $y_i = [0, a_i, b_i]^{\mathrm{T}}$, $i = 1, 2$:

$$\tilde{y}_h = y + hy_1 + h^2 y_2 = \sqrt{\frac{\gamma}{mT}} \begin{pmatrix} 0 \\ mT + ha_1 + h^2 a_2 \\ -p + hb_1 + h^2 b_2 \end{pmatrix}. \tag{3.35}$$

In order to satisfy the modified degeneracy condition

$$\tilde{M}_h \frac{\partial \tilde{E}_h}{\partial x} = \tilde{y}_h \tilde{y}_h^{\mathrm{T}} \frac{\partial \tilde{E}_h}{\partial x} = 0, \tag{3.36}$$

which leads to

$$\tilde{y}_h^{\mathrm{T}} \frac{\partial \tilde{E}_h}{\partial x} = 0, \tag{3.37}$$

the following condition has to be satisfied

$$\left( mT + ha_1 + h^2 a_2 \right) \frac{\partial \tilde{E}_h}{\partial p} + \left( -p + hb_1 + h^2 b_2 \right) \frac{\partial \tilde{E}_h}{\partial S} = 0, \tag{3.38}$$

which has a solution

$$a_1 = a_2 = b_1 = 0 \quad \text{and} \quad b_2 = -\frac{U''(q)p}{6m}. \tag{3.39}$$

Thus, the modified friction matrix can be written as

$$\tilde{M}_h = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \gamma mT & -\gamma p\alpha(q) \\ 0 & -\gamma p\alpha(q) & \frac{\gamma p^2 \alpha^2(q)}{mT} \end{pmatrix}, \tag{3.40}$$

where the 'modifying factor' is given by

$$\alpha(q) = 1 + \frac{h^2 U''(q)}{6m}. \tag{3.41}$$

Moreover, the modified friction matrix induces a small (second-order) perturbation of the physical entropy production

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\partial S}{\partial x} \cdot \tilde{M}_h \frac{\partial S}{\partial x} = \frac{\gamma p^2 \alpha^2(q)}{mT} \geq 0. \tag{3.42}$$

As a result, the irreversible part, incorporating the modified friction matrix (3.40), becomes

$$\mathrm{d} \begin{bmatrix} q \\ p \\ S \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ -\gamma p\alpha(q) \\ \frac{\gamma p^2 \alpha^2(q)}{mT} \end{bmatrix}}_{Y_m} \mathrm{d}t, \tag{3.43}$$

which can be solved exactly (with $q$ being left unchanged)

$$p_h = \exp\left(-\gamma\alpha(q)h\right)p \tag{3.44}$$

and

$$S_h = S + \frac{\gamma p^2\alpha^2(q)}{mT}\int_0^h \exp\left(-2\gamma\alpha(q)t\right)\,dt = S + \frac{p^2\alpha(q)}{2mT}\left[1 - \exp\left(-2\gamma\alpha(q)h\right)\right]. \tag{3.45}$$

In this case, the generator for the modified irreversible dynamics becomes

$$\mathcal{L}_{Y_m} = -\gamma p\alpha(q)\cdot\nabla_p + \frac{\gamma p^2\alpha^2(q)}{mT}\cdot\nabla_S. \tag{3.46}$$

By replacing the Y piece by $Y_m$ in the YBABY method (3.13), we can similarly define a symmetric splitting method, termed '$Y_m$BABY$_m$' or 'mYBABY', as

$$e^{h\hat{\mathcal{L}}_{mYBABY}} = e^{\frac{h}{2}\mathcal{L}_{Y_m}}e^{h\mathcal{L}_{Verlet}}e^{\frac{h}{2}\mathcal{L}_{Y_m}} = e^{\frac{h}{2}\mathcal{L}_{Y_m}}e^{\frac{h}{2}\mathcal{L}_B}e^{h\mathcal{L}_A}e^{\frac{h}{2}\mathcal{L}_B}e^{\frac{h}{2}\mathcal{L}_{Y_m}}, \tag{3.47}$$

where the associated operator can be worked out by applying the Baker–Campbell–Hausdorff formula [1–3] as

$$\hat{\mathcal{L}}_{mYBABY} = \mathcal{L}_A + \mathcal{L}_B + \mathcal{L}_{Y_m} + O(h^2), \tag{3.48}$$

which indicates formally second-order convergence for the mYBABY method (3.47). It can be easily shown that all four requirements listed in §2b are satisfied for the mYBABY method. The integration steps of the mYBABY method read:

$$p^{n+1/4} = \exp\left(-\gamma\alpha\left(q^n\right)\frac{h}{2}\right)p^n, \tag{3.49}$$

$$S^{n+1/2} = S^n + \left[p^n\right]^2\alpha\left(q^n\right)\left[1 - \exp\left(-\gamma\alpha\left(q^n\right)h\right)\right]/(2mT), \tag{3.50}$$

$$p^{n+2/4} = p^{n+1/4} + \left(\frac{h}{2}\right)F(q^n), \tag{3.51}$$

$$q^{n+1} = q^n + hm^{-1}p^{n+2/4}, \tag{3.52}$$

$$p^{n+3/4} = p^{n+2/4} + \left(\frac{h}{2}\right)F(q^{n+1}), \tag{3.53}$$

$$p^{n+1} = \exp\left(-\gamma\alpha\left(q^{n+1}\right)\frac{h}{2}\right)p^{n+3/4} \tag{3.54}$$

and $$S^{n+1} = S^{n+1/2} + \left[p^{n+3/4}\right]^2\alpha\left(q^{n+1}\right)\left[1 - \exp\left(-\gamma\alpha\left(q^{n+1}\right)h\right)\right]/(2mT). \tag{3.55}$$

Note that in the case of the 'modifying factor' (3.41) being unity, the mYBABY method (3.47) reduces exactly to the YBABY method (3.13).

It can be shown that the truncated energy $\tilde{E}_h$ (3.33) is the truncated modified energy, up to second order, for the mYBABY method (3.47), based on the fact that: (i) the Verlet method for the reversible dynamics preserves $\tilde{E}_h$ (3.33) at second order; (ii) the exact solver for the irreversible dynamics preserves $\tilde{E}_h$ (3.33) exactly. In principle, we could truncate the modified energy $\tilde{E}_h$ at higher orders (e.g. fourth, sixth, ...) than that of (3.33), which would lead to higher orders for the overall methods if the irreversible dynamics can be solved exactly. Moreover, it might be more appropriate to refer those GENERIC integrators that incorporate the truncation of the modified
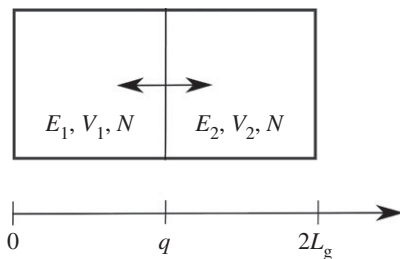
**Figure 1.** Schematic illustration of the two gas containers exchanging heat and volume.

energy to 'pseudo-GENERIC integrators' (in a sense similar to pseudo-symplectic integrators that preserve the symplectic structure only to certain orders [50]).

It can be further shown that the mYBABY method (3.47) preserves the conformal symplectic structure if the Hessian of the potential energy is a constant, i.e. $U''(q) = C$. That is, following (3.31), we have

$$
\begin{aligned}
\mathrm{d}q^{n+1} \wedge \mathrm{d}p^{n+1} &= \mathrm{e}^{-\gamma_{\mathrm{m}} h/2} \mathrm{d}q^{n+1} \wedge \mathrm{d}p^{n+2/4}, \\
&= \mathrm{e}^{-\gamma_{\mathrm{m}} h/2} \mathrm{d}q^{n} \wedge \mathrm{d}p^{n+2/4}, \\
&= \mathrm{e}^{-\gamma_{\mathrm{m}} h} \mathrm{d}q^{n} \wedge \mathrm{d}p^{n},
\end{aligned}
\tag{3.56}
$$

where

$$
\gamma_{\mathrm{m}} = \gamma \left( 1 + \frac{h^2 C}{6m} \right).
\tag{3.57}
$$

which can be thought of as a modified decay rate compared to the damping rate in the YBABY method (3.31).

Note that the preservation of the conformal symplectic structure is, in the literature, often associated with a decay rate of exactly the damping rate as in the YBABY method. Therefore, we may interpret that the YBABY method preserves the conformal symplectic structure in a 'strong' sense whereas the mYBABY method preserves the conformal symplectic structure in a 'weak' sense.

## (b) Two gas containers exchanging heat and volume

In order to demonstrate the generality of our framework introduced in §3a(ii), we also consider an example of two (ideal) gas containers exchanging heat and volume (see figure 1 and Exercises 3 & 9 in [19] for more details) with independent variables $x = (q, p, S_1, S_2)$, where $q$ and $p$, respectively, represent the position and momentum of the separating wall of mass $m$, while $S_1$ and $S_2$ are, respectively, the entropies of the two subsystems. In this case, the total energy is given by

$$
E(x) = \frac{p^2}{2m} + E_1 + E_2,
\tag{3.58}
$$

where $E_1$ and $E_2$ are, respectively, the internal energies of the two subsystems with the following relationships to their associated entropies and volumes (i.e. the Sackur–Tetrode equation for ideal gases)

$$
\frac{S_i(x)}{N k_{\mathrm{B}}} = \ln \left[ \hat{c} V_i \, (E_i)^{3/2} \right], \quad i = 1, 2,
\tag{3.59}
$$

11

royalsocietypublishing.org/journal/rspa  *Proc. R. Soc. A* **476:** 20190446

where $k_B$ is the Boltzmann constant, $\hat{c}$ is another constant that is needed to ensure the argument of the logarithm dimensionless, and it is assumed that the two subsystems contain the same number of particles, $N$. The volumes of the two subsystems are given by

$$V_1 = qA_c \quad \text{and} \quad V_2 = (2L_g - q)A_c,$$ (3.60)

where $A_c$ is the area of the cross section and $2L_g$ is the length of the container. Given the antisymmetric Poisson matrix

$$L = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$ (3.61)

and the positive semidefinite (symmetric) friction matrix

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{\alpha}{T_1^2} & -\dfrac{\alpha}{T_1 T_2} \\ 0 & 0 & -\dfrac{\alpha}{T_1 T_2} & \dfrac{\alpha}{T_2^2} \end{pmatrix} = yy^T \quad \text{and} \quad y = \begin{pmatrix} 0 \\ 0 \\ \dfrac{\sqrt{\alpha}}{T_1} \\ -\dfrac{\sqrt{\alpha}}{T_2} \end{pmatrix},$$ (3.62)

where the positive constant parameter $\alpha$ determines the strength of the heat exchange, and $T_1$ and $T_2$ are, respectively, the temperatures of the two subsystems, related to the associated internal energies by

$$E_i = \frac{3}{2} N k_B T_i, \quad i = 1, 2,$$ (3.63)

the resulting equations of motion of the GENERIC system can be written as

$$\dot{q} = \frac{\partial E}{\partial p} = \frac{p}{m},$$ (3.64)

$$\dot{p} = -\frac{\partial E}{\partial q} = \frac{2}{3}\left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right),$$ (3.65)

$$\dot{S}_1 = \frac{\alpha}{T_1}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) = \frac{9N^2 k_B^2 \alpha}{4E_1}\left(\frac{1}{E_1} - \frac{1}{E_2}\right)$$ (3.66)

and

$$\dot{S}_2 = -\frac{\alpha}{T_2}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) = -\frac{9N^2 k_B^2 \alpha}{4E_2}\left(\frac{1}{E_1} - \frac{1}{E_2}\right).$$ (3.67)

Although the motion of the wall is assumed to be frictionless (i.e. there is no explicit damping term as in linearly damped systems in §3a), all oscillations of the separating wall have to be damped since they induce (time dependent) temperature differences and thus a heat flux with entropy production. Alternatively, an analysis of the equations (3.64)–(3.67) linearized around equilibria indicates that the system would always relax to equilibrium. We would also like to point out that, unlike linearly damped systems considered in §3a where the $(q, p)$ dynamics may be viewed as being independent of the entropy, the $(q, p)$ dynamics in this case 'strongly' depends on the dynamics of $(S_1, S_2)$, and vice versa—it is a fully coupled GENERIC system.

As in §3a(i), we could also split the system (3.64)–(3.67) into reversible and irreversible parts:

$$d\begin{bmatrix} q \\ p \\ S_1 \\ S_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \dfrac{\partial E}{\partial p} \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{A} dt + \underbrace{\begin{bmatrix} 0 \\ -\dfrac{\partial E}{\partial q} \\ 0 \\ 0 \end{bmatrix}}_{B} dt + \underbrace{\begin{bmatrix} 0 \\ 0 \\ \dfrac{\alpha}{T_1}\left(\dfrac{1}{T_1} - \dfrac{1}{T_2}\right) \\ -\dfrac{\alpha}{T_2}\left(\dfrac{1}{T_1} - \dfrac{1}{T_2}\right) \end{bmatrix}}_{Y} dt,$$ (3.68)

for which we can use the Verlet method for the reversible dynamics, with degenerate variables $S_1$ and $S_2$ being constants, while a suitable method can be used to solve the irreversible dynamics. We would again like to construct a modified energy and an associated friction matrix as in (2.2). To this end, following the procedures in §3a(ii), we first identify the modified energy associated with the Verlet method used for the reversible dynamics

$$\tilde{E}_h = \frac{p^2}{2m} + E_1 + E_2 + \frac{h^2}{54m}\left[\frac{5p^2}{m}\left(\frac{E_1}{q^2} + \frac{E_2}{(2L_g - q)^2}\right) - \left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right)^2\right] + O(h^4), \qquad (3.69)$$

based on which we can subsequently work out the derivatives of the truncated modified energy up to second order

$$\frac{\partial \tilde{E}_h}{\partial S_1} = T_1 + \frac{h^2 T_1}{54mq}\left[\frac{5p^2}{mq} - 2\left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right)\right], \qquad (3.70)$$

and

$$\frac{\partial \tilde{E}_h}{\partial S_2} = T_2 + \frac{h^2 T_2}{54m(2L_g - q)}\left[\frac{5p^2}{m(2L_g - q)} + 2\left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right)\right]. \qquad (3.71)$$

We can then construct the associated modified friction matrix in the fashion of backward error analysis as

$$\tilde{M}_h = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{\alpha\alpha_2^2}{T_1^2} & -\dfrac{\alpha\alpha_1\alpha_2}{T_1 T_2} \\ 0 & 0 & -\dfrac{\alpha\alpha_1\alpha_2}{T_1 T_2} & \dfrac{\alpha\alpha_1^2}{T_2^2} \end{pmatrix} = \tilde{y}_h \tilde{y}_h^T \quad \text{and} \quad \tilde{y}_h = \begin{pmatrix} 0 \\ 0 \\ \dfrac{\alpha_2\sqrt{\alpha}}{T_1} \\ -\dfrac{\alpha_1\sqrt{\alpha}}{T_2} \end{pmatrix}, \qquad (3.72)$$

where the modifying factors are given by

$$\alpha_1 = 1 + h^2\beta_1 \quad \beta_1 = \frac{1}{54mq}\left[\frac{5p^2}{mq} - 2\left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right)\right], \qquad (3.73)$$

and

$$\alpha_2 = 1 + h^2\beta_2, \quad \beta_2 = \frac{1}{54m(2L_g - q)}\left[\frac{5p^2}{m(2L_g - q)} + 2\left(\frac{E_1}{q} - \frac{E_2}{2L_g - q}\right)\right], \qquad (3.74)$$

respectively. The modified irreversible part, incorporating the modified friction matrix (3.72), is now given by

$$\mathrm{d}\begin{bmatrix} q \\ p \\ S_1 \\ S_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \dfrac{\alpha\alpha_2}{T_1}\left(\dfrac{\alpha_2}{T_1} - \dfrac{\alpha_1}{T_2}\right) \\ -\dfrac{\alpha\alpha_1}{T_2}\left(\dfrac{\alpha_2}{T_1} - \dfrac{\alpha_1}{T_2}\right) \end{bmatrix}}_{Y_m} \mathrm{d}t. \qquad (3.75)$$

The modified friction matrix again induces a small (second-order) perturbation of the physical entropy production

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\partial S}{\partial x} \cdot \tilde{M}_h \frac{\partial S}{\partial x} = \alpha\left(\frac{\alpha_2}{T_1} - \frac{\alpha_1}{T_2}\right)^2 \geq 0. \qquad (3.76)$$

Both YBABY and mYBABY methods are similarly defined in this setting as for linearly damped systems in §3a. However, we are unable to solve the modified irreversible dynamics exactly here, thus a second-order explicit midpoint method is suggested to approximate the modified irreversible dynamics (3.75) while the Verlet method is still used for the reversible dynamics, with degenerate variables $S_1$ and $S_2$ being constants. Overall, the two split GENERIC integrators are both expected to be second order.

We would like to point out that in some cases it might be beneficial to replace the explicit midpoint method by alternative (higher order and/or higher accuracy) methods. Moreover, inspired by the subsampling techniques popular in large-scale Bayesian sampling [38,51], it might be computationally highly advantageous (especially in high dimension) to decompose the positive semidefinite modified friction matrix into non-overlapping principal submatrices (a principal submatrix can be obtained by selecting a subset of rows and the same subset of columns) that are still positive semidefinite. Having avoided directly solving a high dimensional gradient flow, we could instead solve each of the decomposed and much smaller subsystems with a significantly reduced computational overhead (even with high accuracy). A thorough investigation of this direction is beyond the scope of this article, and will be left for future work.

It is also worth mentioning that when the modified irreversible dynamics has to be approximated by certain numerical methods, the truncated modified energy is expected to be preserved in an 'approximation' sense. A detailed analysis of the effect of the approximation is also beyond the scope of this article, and will be left for future work.

# 4. Numerical experiments

In this section, we conduct various numerical experiments to examine the performance of the two split GENERIC integrators introduced in this article.

## (a) Simulation details

In the case of linearly damped systems, we consider one-dimensional examples of a damped harmonic oscillator (i.e. $U(q) = kq^2/2$), for which an analytical solution can be obtained [20], as well as a damped nonlinear oscillator (i.e. $U(q) = -k\cos(q)$) where the argument of the cosine function should be dimensionless and this is achieved by fixing the unit of length via the initial position $q_0$. The equations of motion of both linearly damped systems can be simplified by dimensional analysis [52]. Without loss of generality, in both cases, we choose the basic units (mass, time, temperature, and length, respectively) as $m = k = T = 1$ and $q_0 = 2$, where the initial position was particularly chosen to demonstrate the nonlinear effects in the damped nonlinear oscillator. Subsequently, the equations of motion of both linearly damped systems involve only the single dimensionless parameter of $\gamma \geq 0$. Moreover, in both cases, we chose $p_0 = 0$ as more general values of the initial momentum essentially correspond to a shift of the initial time. Since we are more interested in the deviation from the initial entropy than its absolute value, we set the initial entropy to be zero in both cases.

In the other case of two gas containers (where $A_c = L_g^2$) exchanging heat and volume, we chose the basic units of mass and length as $m = L_g = 1$, respectively. We further set $Nk_B = 1$, which fixes a characteristic macroscopic unit of entropy, the counterpart of $T = 1$ (i.e. the thermodynamic unit) in the previous examples of linearly damped systems. In order to fix the fourth unit of time, $\alpha = 0.5$ was chosen so that: (i) the period of the oscillation is of order one; (ii) there are enough oscillations to collect statistical data (larger values of $\alpha$ lead to faster decay of the amplitude of the oscillation). Furthermore, initial conditions of $(q, p, E_1, E_2) = (1, 2, 2, 2)$ were used (i.e. the separating wall is initially in the middle of the container with an initial velocity).

In all three cases, the positions appeared to be oscillating with the amplitudes decaying exponentially. The total simulation time $T_s$ in each case was thus chosen so that $t = T_s$ is the time at which the amplitude of the oscillation was reduced to approximately $1/e$ times its initial value.

Denoting $h$ as the integration stepsize and subsequently $\hat{N} = T_s/h$ the number of integration steps, the root-mean-square error (RMSE) of observable $\phi$ is defined as follows:

$$\text{RMSE}(\phi) = \sqrt{\frac{1}{\hat{N}} \sum_{i=1}^{\hat{N}} \left( \hat{\phi}_i - \phi_i \right)^2}, \tag{4.1}$$

where $\hat{\phi}_i$ and $\phi_i$ represent the numerical approximation at time $ih$ and its corresponding exact (reference) value, respectively.

In order to demonstrate the superiority of structure-preserving integrators over alternative schemes, we compare the two split GENERIC integrators introduced in this article with the explicit third-order Runge–Kutta (RK3) method used also in [12] as well as the average discrete gradient (ADG) method [53]. The choice of the RK3 method is clearly arbitrary, while other methods are typically second order, it serves as a good example of a higher-order method that is not structure-preserving.

## (i) The third-order Runge–Kutta method

Rewriting GENERIC systems in a compact form as $\dot{x}(t) = f(t, x)$ with initial conditions $x(0) = x_0$, the RK3 method is given by

$$x_{n+1} = x_n + \frac{h}{6} (k_1 + 4k_2 + k_3), \tag{4.2}$$

where

$$k_1 = f(t_n, x_n), \tag{4.3}$$

$$k_2 = f\left(t_n + \frac{h}{2}, x_n + \frac{hk_1}{2}\right) \tag{4.4}$$

and

$$k_3 = f(t_n + h, x_n - hk_1 + 2hk_2), \tag{4.5}$$

with $t_n = nh$, $n = 0, 1, 2, \ldots$. Note that the RK3 method is neither symplectic nor conformal symplectic. Thus, it does not preserve the Poisson structure for the reversible dynamics. It is also worth mentioning that the two split GENERIC integrators introduced in this article at each step typically require only one force calculation, which often dominates the computational cost per step especially for large-scale simulations, whereas three force calculations are needed for the RK3 method.

## (ii) The average discrete gradient method

The so-called discrete gradient methods [45–47], which are also known as discrete derivative methods [54], have often been used for the time integration of dissipative systems [23–25,55–60]. For instance, they have recently been suggested to temporally discretize the Landau collision operator in an attempt to preserve its metriplectic/GENERIC structure [30]. However, as stated in [61,62], discrete gradient methods are generally not symplectic for the symplectic leaves and thus the Poisson structure of the reversible dynamics is not preserved. Therefore, those discrete gradient methods do not belong to either of the GENERIC integrators defined in §2.

Moreover, discrete gradient methods are typically implicit, in which case iterative methods (e.g. Newton's method) are needed to approximate the solutions at each step. Therefore, discrete gradient methods could be considerably more time-consuming than alternative explicit methods depending on not only the stopping criterion for the iterating procedure [61] but also the size of the linear system that needs to be solved at each iteration. However, in the special case of a

damped harmonic oscillator (i.e. $U(q) = kq^2/2$), we can work out the integration steps without the iterating procedure. More precisely, we rewrite the GENERIC system (3.4)–(3.6) as

$$\frac{dx}{dt} = S(x)\nabla E(x) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -\gamma m & 0 \\ 0 & 0 & \frac{\gamma p^2}{mT^2} \end{pmatrix} \begin{pmatrix} kq \\ p \\ m \\ T \end{pmatrix}, \tag{4.6}$$

which is discretized by

$$\frac{x_{n+1} - x_n}{h} = \bar{S}(x_n, x_{n+1})\bar{\nabla}E(x_n, x_{n+1}), \tag{4.7}$$

where the matrix $\bar{S}(x_n, x_{n+1})$ approaches $S(x)$ in the limits of $x_{n+1} \to x_n$ and $h \to 0$, while the discrete gradient $\bar{\nabla}E(x_n, x_{n+1})$ satisfies the following conditions:

$$(x_{n+1} - x_n) \cdot \bar{\nabla}E(x_n, x_{n+1}) = E(x_{n+1}) - E(x_n) \tag{4.8}$$

and

$$\bar{\nabla}E(x_n, x_n) = \nabla E(x_n). \tag{4.9}$$

We consider a midpoint discretization for $\bar{S}$, i.e.

$$\bar{S}(x_n, x_{n+1}) = S(x_{n+1/2}) \quad \text{and} \quad x_{n+1/2} = \frac{x_n + x_{n+1}}{2}, \tag{4.10}$$

and the ADG [53] for $\bar{\nabla}E$, i.e.

$$\bar{\nabla}E(x_n, x_{n+1}) = \int_0^1 \nabla E\left((1 - \xi)x_n + \xi x_{n+1}\right) d\xi, \tag{4.11}$$

in which case the ADG method reduces to the implicit midpoint method, which is second order and symplectic (for the symplectic leaf with $\gamma = 0$) [1,63]

$$q^{n+1} = q^n + \frac{hp^{n+1/2}}{m}, \tag{4.12}$$

$$p^{n+1} = p^n - hkq^{n+1/2} - h\gamma p^{n+1/2} \tag{4.13}$$

and

$$S^{n+1} = S^n + h\gamma \left[p^{n+1/2}\right]^2 /(mT), \tag{4.14}$$

where $q^{n+1/2} = (q^n + q^{n+1})/2$ and $p^{n+1/2} = (p^n + p^{n+1})/2$. One might be surprised how, with the irreversible dynamics being switched off (i.e. $\gamma = 0$), the energy-conserving ADG method (or the implicit midpoint method) can also be symplectic for the symplectic leaves, which seems to 'contradict' the findings of [43] (see discussions in §2b). However, we point out that in the case of a harmonic oscillator, the corresponding Hamiltonian subsystem is in fact integrable, in such a special case the ADG method preserves not only the energy but also the Poisson structure. Moreover, we can easily solve (4.12)–(4.13) to obtain

$$q^{n+1} = \frac{\left(4m + 2mh\gamma - h^2k\right)q^n + 4hp^n}{4m + 2mh\gamma + h^2k} \tag{4.15}$$

and

$$p^{n+1} = \frac{-4mhkq^n + \left(4m - 2mh\gamma - h^2k\right)p^n}{4m + 2mh\gamma + h^2k}, \tag{4.16}$$

and subsequently (if $4m + h^2k > 2mh\gamma$)

$$dq^{n+1} \wedge dp^{n+1} = e^{-\gamma_{ADG}h}dq^n \wedge dp^n, \tag{4.17}$$

where

$$\gamma_{ADG} = -\frac{1}{h}\ln\left(\frac{4m - 2mh\gamma + h^2k}{4m + 2mh\gamma + h^2k}\right) = \gamma + O(h^2). \tag{4.18}$$

We can see from (4.17)–(4.18) that the 'symplectic two form' of the ADG method decays exponentially with a constant decay rate, thus the ADG method in this special case preserves the

conformal symplectic structure in a 'weak' sense. Furthermore, when the irreversible dynamics is switched off (i.e. $\gamma = 0$), (4.17) reduces to $dq^{n+1} \wedge dp^{n+1} = dq^n \wedge dp^n$, which indicates that the ADG method preserves the Poisson structure for the reversible dynamics. We emphasize here that if the force is nonlinear or alternative discrete gradient approximations (e.g. the midpoint discrete gradient [54], which was used in several methods compared in [25]) are used, the preservation of the conformal symplectic structure and the Poisson structure for the reversible dynamics is expected to be violated while the iterating procedure seems to be unavoidable, which could result in a substantial computational overhead.

The ADG method is unsurprisingly implicit in both cases of the damped nonlinear oscillator (i.e. $U(q) = -k\cos(q)$) and two gas containers exchanging heat and volume. While the former is similar to the damped harmonic oscillator case except replacing $kq$ in (4.6) by $k\sin(q)$, whose ADG is still analytically integrable, the latter is more involved. To be more precise, we rewrite the GENERIC system (3.64)–(3.67) as

$$\frac{dx}{dt} = S(x)\nabla E(x) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & \dfrac{\alpha}{T_1^2} & -\dfrac{\alpha}{T_1 T_2} \\ 0 & 0 & -\dfrac{\alpha}{T_1 T_2} & \dfrac{\alpha}{T_2^2} \end{pmatrix} \begin{pmatrix} \nabla E_q(x) \\ \dfrac{p}{m} \\ 1 \\ 1 \end{pmatrix}, \tag{4.19}$$

where

$$\nabla E_q(x) = \frac{2}{3}\left( \frac{E_2}{2L_g - q} - \frac{E_1}{q} \right) = \frac{2\hat{B}}{3}\left[ (2L_g - q)^{-\frac{5}{3}} e^{\frac{2S_2}{3Nk_B}} - q^{-\frac{5}{3}} e^{\frac{2S_1}{3Nk_B}} \right], \tag{4.20}$$

with the constant $\hat{B}$ being defined as

$$\hat{B} = \left( \hat{c} A_c \right)^{-\frac{2}{3}}. \tag{4.21}$$

In this case, the ADG for $\bar{\nabla}E_q(x_n, x_{n+1})$ is no longer analytically integrable, and thus approximated by using the trapezoidal rule

$$\bar{\nabla}E_q(x_n, x_{n+1}) = \int_0^1 \nabla E_q\left( (1-\xi)x_n + \xi x_{n+1} \right) d\xi$$

$$\approx \frac{\hat{B}}{3}\left[ (2L_g - q_n)^{-\frac{5}{3}} e^{\frac{2S_{2,n}}{3Nk_B}} + (2L_g - q_{n+1})^{-\frac{5}{3}} e^{\frac{2S_{2,n+1}}{3Nk_B}} \right]$$

$$- \frac{\hat{B}}{3}\left[ q_n^{-\frac{5}{3}} e^{\frac{2S_{1,n}}{3Nk_B}} + q_{n+1}^{-\frac{5}{3}} e^{\frac{2S_{1,n+1}}{3Nk_B}} \right]. \tag{4.22}$$

As a result of the approximation, the exact total energy conservation of the ADG method is expected to be violated. (Note that one may rewrite the GENERIC system (3.64)–(3.67) with independent variables $x = (q, p, E_1, E_2)$. However, the same issue of violating the exact total energy conservation for the ADG method is still expected.)

## (b) Damped harmonic oscillator

We first consider the damped harmonic oscillator (i.e. $U(q) = kq^2/2$) example, where the analytical solution is available using the same set of parameters (except $q_0$) in [20]. It is of great importance that numerical approximations of GENERIC systems have (i) a good conservation of the total energy and (ii) a faithful production of the physical entropy, both of which were compared in figure 2. We compare the performance of the two split GENERIC integrators with that of the RK3 method and the ADG method. Since the ADG method conserves the total energy exactly (i.e. up to machine precision) in this setting [42], its results will not be included in comparisons of the energy conservation. According to the dashed order lines, the RK3 method shows third-order convergence whereas other methods are all second order as expected. Among the second-order methods, the mYBABY method in all the cases we tested outperforms the YBABY method in
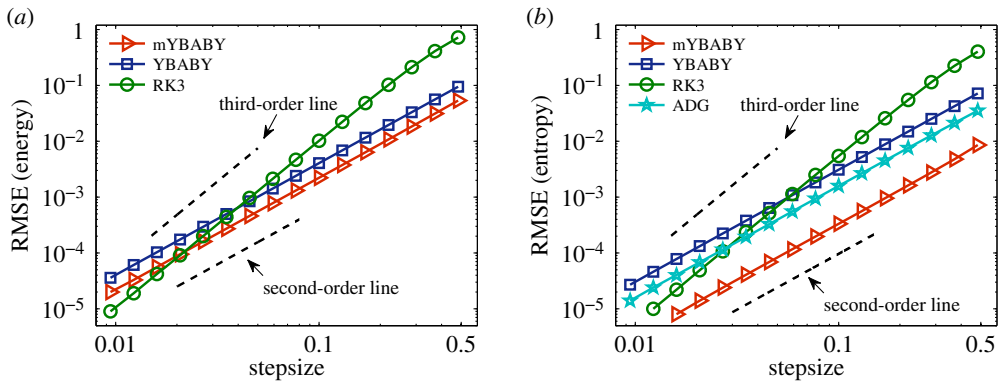
**Figure 2.** Double logarithmic plot of the root-mean-square error (RMSE) (4.1) in the total energy (3.1) (*a*) and entropy (*b*) against stepsize by comparing the two split GENERIC integrators introduced in this article with the third-order Runge–Kutta (RK3) method and the average discrete gradient (ADG) method, which conserves the total energy exactly (i.e. up to machine precision) and thus is only included in comparisons of the entropy production, with a damping rate of $\gamma = 0.01$ and a total simulation time of $T_s = 200$ in a standard setting of a damped harmonic oscillator as described in §4a. The stepsizes tested began at $h = 0.0094$ and were increased incrementally by 30% until around $h = 0.5$. Note that, with this set of parameters, the damped harmonic oscillator is 'underdamped' (i.e. the position of the particle oscillates around zero with the amplitude exponentially decreasing to zero) and the associated period is $T_p \approx 2\pi$. Dashed black lines represent the second- and third-order convergence as indicated. (Online version in colour.)
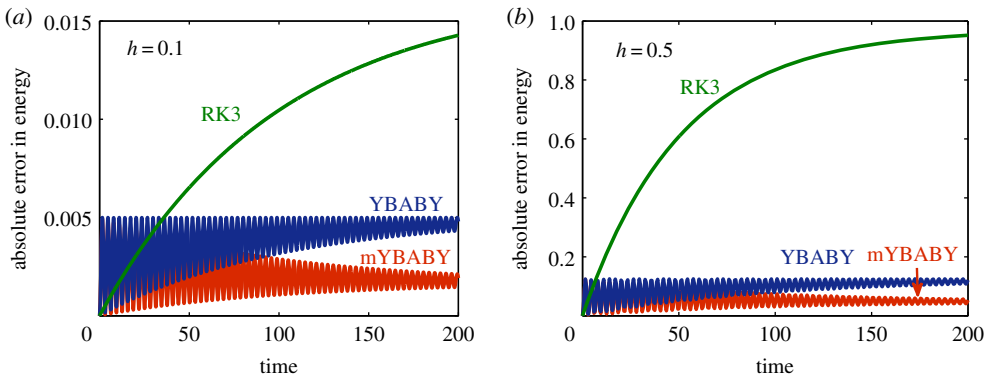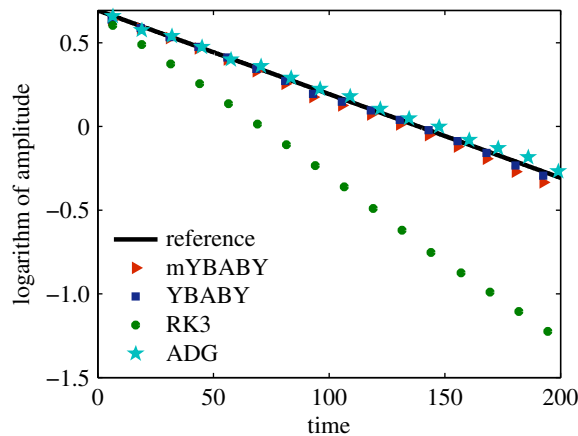


**Figure 3.** Evolutions of the absolute error in the total energy (3.1) obtained from various numerical methods in a standard setting of a damped harmonic oscillator with a damping of $\gamma = 0.01$, a total simulation time of $T_s = 200$, and a fixed stepsize of $h = 0.1$ (*a*) and $h = 0.5$ (*b*). (Online version in colour.)

terms of the RMSE in both quantities. Particularly, in terms of the entropy production, mYBABY is remarkably one order of magnitude more accurate than YBABY, which is slightly outperformed by the ADG method. In both cases, despite its higher computational overhead, the higher-order RK3 method is only more accurate than either of the split GENERIC integrators when the stepsize is relatively small, especially for the mYBABY method.

The evolutions of the absolute error in the total energy from various methods were compared (against the exact value of $E_0 = 2$) and plotted in figure 3. We can see from the figure that, with a stepsize of $h = 0.1$ (figure 3*a*), the absolute error of the RK3 method rises quickly before eventually settling down, while the absolute errors of the two split GENERIC integrators oscillate strongly with the amplitudes decreasing. Consistent with our findings in figure 2, the absolute error of mYBABY is largely smaller than that of YBABY. Nevertheless, both mYBABY and YBABY
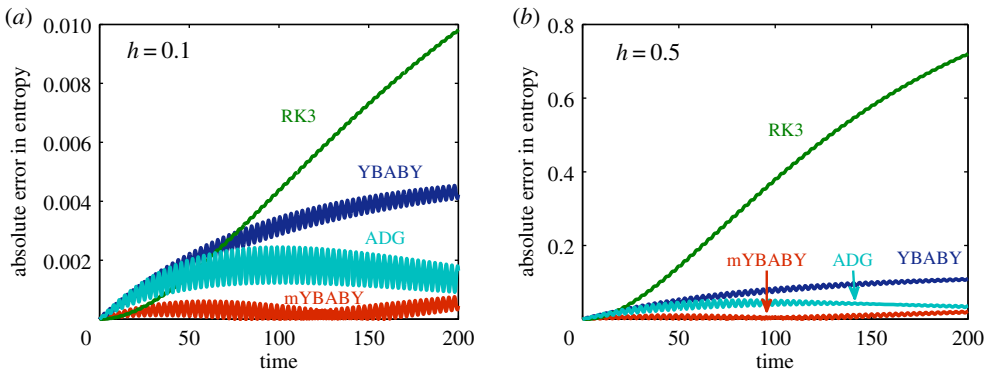
**Figure 4.** Evolutions of the absolute error in the entropy obtained from various numerical methods in a standard setting of a damped harmonic oscillator with a damping rate of $\gamma = 0.01$, a total simulation time of $T_s = 200$, and a fixed stepsize of $h = 0.1$ (a) and $h = 0.5$ (b). (Online version in colour.)



**Figure 5.** Evolutions of the logarithm of the oscillation amplitude obtained from various numerical methods in a standard setting of a damped harmonic oscillator with a damping rate of $\gamma = 0.01$, a total simulation time of $T_s = 200$, and a fixed stepsize of $h = 0.5$. Note that only the numerical solution of $q$ is presented since it is very similar to that of $p$ with a slight shift. (Online version in colour.)

methods are more accurate than the RK3 method with a relatively large stepsize. The behaviour is rather similar with a larger stepsize of $h = 0.5$ (figure 3b) except the magnitude of the error obtained from each method is considerably larger than that with a smaller stepsize.

Figure 4 compares the control of the entropy production from a variety of methods. The behaviour of the YBABY, mYBABY, and RK3 methods are similar to that of figure 3. Interestingly, with a stepsize of $h = 0.1$ (figure 4a), the absolute error of the ADG method, while oscillating, initially grows before decreasing while the absolute error of mYBABY, also oscillating, is constantly smaller than that of ADG. The behaviour is again very similar with a larger stepsize of $h = 0.5$ (figure 4b) except the magnitude of the errors.

We also compare in figure 5 the decay of the oscillation amplitude represented by the 'local maximum' (in logarithm) of the numerical solution of the position, which characterizes the preservation of the conformal symplectic structure. It can be seen from the figure that while the decay rate of the YBABY method is preserved (almost indistinguishable from the reference decay rate of the damping rate $\gamma$), the RK3 method, which is not conformal symplectic, exhibits a clear
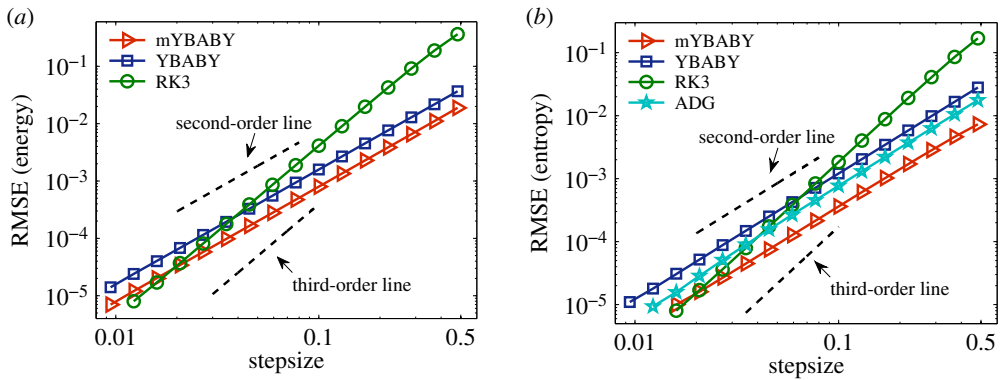
**Figure 6.** Double logarithmic plot of the root-mean-square error (RMSE) (4.1) in the total energy (3.1) (*a*) and entropy (*b*) against stepsize by comparing the two split GENERIC integrators introduced in this article with the third-order Runge–Kutta (RK3) method and the average discrete gradient (ADG) method with a damping rate of $\gamma = 0.01$ and a total simulation time of $T_s = 180$ in a standard setting of a damped nonlinear oscillator as described in §4a. In this case, the position of the damped nonlinear oscillator also oscillates with an associated period of $T_p \approx 8.4$. The format of the plots is the same as in figure 2. (Online version in colour.)

drift. It can be also observed (and verified) that both mYBABY and ADG decay at slightly different rates of $\gamma_m$ (3.57) and $\gamma_{ADG}$ (4.18), respectively, compared to the reference decay. This indicates that both mYBABY and ADG in this particular case preserve the conformal symplectic structure in a 'weak' sense (see discussions at the end of §3a(ii)).

## (c) Damped nonlinear oscillator

We also investigate the performance of various methods with the damped nonlinear oscillator (i.e. $U(q) = -k\cos(q)$), where the reference solution was obtained by using the RK3 method with a very small stepsize of $h = 0.001$. It turns out that the performance of those methods is very similar to that in the case of the damped harmonic oscillator in §4b. Therefore, we only present the results of the accuracy control of both energy conservation and entropy production as in figure 2. According to the dashed order lines in figure 6, the RK3 method again exhibits third-order convergence whereas the other methods appear to be second order. Moreover, we can see from the figure that in both cases the mYBABY method again comfortably outperforms the YBABY method. The RK3 method, with a higher computational overhead, is again in both cases only more accurate than either of the split GENERIC integrators when the stepsize is relatively small. As mentioned in §4a(ii), the time-consuming iterating procedure had to be adopted for the ADG method in this nonlinear case. Since the ADG method conserves the total energy up to machine precision, we only include it for comparisons of the entropy production. As can be seen from figure 6b that the ADG method is more accurate than the YBABY method but is outperformed by the mYBABY method despite its higher computational overhead.

## (d) Two gas containers

We further examine the performance of various methods in the case of two gas containers exchanging heat and volume described in §3b, where the reference solution was again obtained by using the RK3 method with a very small stepsize of $h = 0.001$. While the RK3 method still shows third-order convergence, the other methods appear to be second order as expected, according to the dashed order lines in figure 7. The performance of the two split GENERIC integrators and the RK3 method is largely similar to that in the previous two examples. More precisely, in both cases the mYBABY method still clearly outperforms the YBABY method, and the two split GENERIC
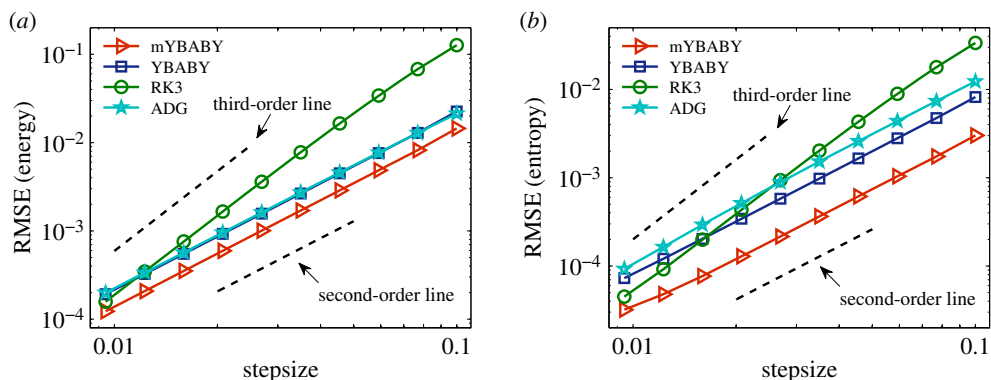
**Figure 7.** Double logarithmic plot of the root-mean-square error (RMSE) (4.1) in the total energy (3.58) (*a*) and total entropy (*b*) against stepsize by comparing the two split GENERIC integrators introduced in this article with the third-order Runge–Kutta (RK3) method and the average discrete gradient (ADG) method with a total simulation time of $T_s = 30$ in a standard setting of two gas containers exchanging heat and volume as described in §4c. In this case, the separating wall oscillates around its equilibrium position (i.e. $q = 1$) with an associated period of $T_p \approx 2.0$. The format of the plots is the same as in figure 2. (Online version in colour.)

integrators appear to be more accurate than the RK3 method unless the stepsize is relatively small. Unlike the damped nonlinear oscillator example in §4c, the ADG is not analytically integrable and had to be approximated, resulting in errors in the total energy. As a result, we can see from figure 7 that the performance of the ADG method is almost indistinguishable from that of the YBABY method in terms of the accuracy control of the energy conservation in figure 7*a*, while the former is outperformed by the latter in terms of the accuracy control of the entropy production in figure 7*b*. In both cases, the ADG method is clearly outperformed by the mYBABY method, although the latter only preserves the truncated modified energy in an 'approximation' sense. Moreover, we would like to point out that while the two split GENERIC integrators are both explicit, the implicit ADG method is computationally much more time-consuming (in this particular case, the evolution of the system was obtained by using the iterative Newton's method at each step in which a linear system associated with a $4 \times 4$ Jacobian matrix was repeatedly solved).

## 5. Conclusions

We have given specific definitions of GENERIC integrators that preserve the underlying thermodynamic structures. In order to construct such integrators, we have presented a framework by splitting a GENERIC system into reversible and irreversible parts. The former, which is often degenerate and reduces to a Hamiltonian form on its symplectic leaves, is solved by a symplectic (Verlet) method (with degenerate variables being left unchanged) for which an associated modified Hamiltonian (and subsequently a modified energy) can be obtained by using backward error analysis. The modified energy is subsequently used to construct a modified friction matrix associated with the irreversible part in such a way that the modified degeneracy condition (2.2) is satisfied. Following the framework, the mYBABY method has been proposed, which, along with another split GENERIC integrator of the YBABY method, is expected to be second order and typically require only one force calculation at each step. Between the two split GENERIC integrators, we have observed that mYBABY clearly outperforms YBABY in all the cases tested, indicating the importance of satisfying the modified degeneracy condition (2.2).

We have demonstrated by conducting a variety of numerical experiments (including linearly damped systems and two gas containers exchanging heat and volume) that, in terms of the accuracy control of both energy conservation and entropy production, the two split GENERIC integrators (particularly the mYBABY method) are more accurate than the higher-order RK3

method unless the stepsizes are relatively small, not to mention the latter requires three force calculations at each step. While the two split GENERIC integrators preserve the conformal symplectic structure for linearly damped systems, RK3 fails and exhibits a clear drift in the decay of the oscillation amplitude of the numerical solutions.

Since the ADG method conserves the total energy up to machine precision, we do not include it in comparisons of the energy conservation for linearly damped systems. It turns out that in both examples of linearly damped systems, the ADG method appears to be more accurate than YBABY, but (despite the use of the time-consuming iterating procedure in the damped nonlinear oscillator case) outperformed by mYBABY in terms of the accuracy control of the entropy production. The ADG is not analytically integrable and had to be approximated in the case of two gas containers exchanging heat and volume, leading to errors in the total energy. As a result of that approximation, the ADG method is as accurate as the YBABY method in terms of the accuracy control of the energy conservation, while the former is outperformed by the latter in terms of the accuracy control of the entropy production. In both cases, although preserving the truncated modified energy in an 'approximation' sense, the mYBABY method is clearly more accurate than the ADG method. This indicates that in cases where approximations have to be made in the ADG method, it could lose its 'built-in' advantage of exact conservation of the total energy and be outperformed by alternative methods (especially mYBABY). Moreover, we would like to emphasize again that the implicit ADG method is considerably more time-consuming than the two split GENERIC integrators due to the use of the iterative Newton's method where a linear system associated with a $4 \times 4$ Jacobian matrix was repeatedly solved at each step. It is anticipated that the computational overhead of discrete gradient methods could be substantially increased for large systems, making them unfavourable compared to explicit structure-preserving integrators (especially mYBABY) in practice.

It is worth mentioning that it might be possible to design GENERIC integrators without an explicit construction of the modified energy $\tilde{E}_h$. This is related to the question whether the irreversible dynamics (i.e. a vector field) is 'compatible' with the canonical transformation associated with the time step $h$. Just as the canonical transformation guarantees that a modified energy $\tilde{E}_h$ does exist, there might be a criterion for 'compatibility' of vector fields with a canonical transformation. The next question would be whether such a 'compatibility' holds only for the physical entropy or for all possible entropies (which would be the original degeneracy). Alternatively, as symplectic integrators can be obtained most easily from a variational principle [64–67], it might be worth looking at irreversible equations with a variational principle for GENERIC integrators.

# References

1. Hairer E, Lubich C, Wanner G. 2006 *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*. Berlin, Germany: Springer.
2. Leimkuhler B, Reich S. 2005 *Simulating Hamiltonian dynamics*. Cambridge, UK: Cambridge University Press.
3. Leimkuhler B, Matthews C. 2015 *Molecular dynamics: with deterministic and stochastic numerical methods*. Berlin, Germany: Springer.

4. Stuart AM, Humphries AR. 1996 *Dynamical systems and numerical analysis*. Cambridge Monographs on Applied and Computational Mathematics, vol. 2. Cambridge, UK: Cambridge University Press.

5. Jordan R, Kinderlehrer D, Otto F. 1998 The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29**, 1–17. (doi:10.1137/S0036141096303359)

6. Otto F. 2001 The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26**, 101–174. (doi:10.1081/PDE-100002243)

7. Ambrosio L, Gigli N, Savaré G. 2008 *Gradient flows: in metric spaces and in the space of probability measures*. Berlin, Germany: Springer Science & Business Media.

8. Hairer E, Lubich C. 2014 Energy-diminishing integration of gradient systems. *IMA J. Numer. Anal.* **34**, 452–461. (doi:10.1093/imanum/drt031)

9. Dressler U. 1988 Symmetry property of the Lyapunov spectra of a class of dissipative dynamical systems with viscous damping. *Phys. Rev. A* **38**, 2103. (doi:10.1103/PhysRevA.38.2103)

10. Moser J. 1994 On quadratic symplectic mappings. *Math. Z.* **216**, 417–430. (doi:10.1007/BF02572331)

11. McLachlan R, Perlmutter M. 2001 Conformal Hamiltonian systems. *J. Geom. Phys.* **39**, 276–300. (doi:10.1016/S0393-0440(01)00020-1)

12. Bhatt A, Floyd D, Moore BE. 2016 Second order conformal symplectic schemes for damped Hamiltonian systems. *J. Sci. Comput.* **66**, 1234–1259. (doi:10.1007/s10915-015-0062-z)

13. Hong J, Sun L, Wang X. 2017 High order conformal symplectic and ergodic schemes for the stochastic Langevin equation via generating functions. *SIAM J. Numer. Anal.* **55**, 3006–3029. (doi:10.1137/17M111691X)

14. Bhatt A, Moore BE. 2017 Structure-preserving exponential Runge–Kutta methods. *SIAM J. Sci. Comput.* **39**, A593–A612. (doi:10.1137/16M1071171)

15. Kane C, Marsden JE, Ortiz M, West M. 2000 Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *Int. J. Numer. Methods Eng.* **49**, 1295–1325. (doi:10.1002/1097-0207(20001210)49:10<1295::aid-nme993>3.3.co;2-n)

16. Jay LO. 2003 Solution of index 2 implicit differential-algebraic equations by Lobatto Runge–Kutta methods. *BIT* **43**, 93–106. (doi:10.1023/A:1023696822355)

17. Grmela M, Öttinger HC. 1997 Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. *Phys. Rev. E* **56**, 6620. (doi:10.1103/PhysRevE.56.6620)

18. Öttinger HC, Grmela M. 1997 Dynamics and thermodynamics of complex fluids. II. Illustrations of a general formalism. *Phys. Rev. E* **56**, 6633. (doi:10.1103/PhysRevE.56.6633)

19. Öttinger HC. 2005 *Beyond equilibrium thermodynamics*. Hoboken, NJ: John Wiley & Sons.

20. Öttinger HC. 2018 GENERIC integrators: structure preserving time integration for thermodynamic systems. *J. Non-Equilib. Thermodyn.* **43**, 89–100. (doi:10.1515/jnet-2017-0034)

21. Öttinger HC. 2017 GENERIC: review of successful applications and a challenger for the future. In *14th Joint European Thermodynamics Conf. JETC 2017, Budapest, Hungary, 21–25 May*. Budapest, Hungary: Department of Energy Engineering, Budapest University of Technology and Engineering.

22. Morrison PJ. 2017 Structure and structure-preserving algorithms for plasma physics. *Phys. Plasmas* **24**, 055502. (doi:10.1063/1.4982054)

23. Portillo D, García Orden JC, Romero I. 2017 Energy–entropy–momentum integration schemes for general discrete non-smooth dissipative problems in thermomechanics. *Int. J. Numer. Methods Eng.* **112**, 776–802. (doi:10.1002/nme.5532)

24. Krüger M, Groß M, Betsch P. 2016 An energy-entropy-consistent time stepping scheme for nonlinear thermo-viscoelastic continua. *ZAMM Z. Angew. Math. Mech.* **96**, 141–178. (doi:10.1002/zamm.201300268)

25. Krüger M, Groß M, Betsch P. 2011 A comparison of structure-preserving integrators for discrete thermoelastic systems. *Comput. Mech.* **47**, 701–722. (doi:10.1007/s00466-011-0570-0)

26. Morrison PJ. 1986 A paradigm for joined Hamiltonian and dissipative systems. *Physica D* **18**, 410–419. (doi:10.1016/0167-2789(86)90209-5)

27. Morrison PJ. 1984 Bracket formulation for irreversible classical fields. *Phys. Lett. A* **100**, 423–427. (doi:10.1016/0375-9601(84)90635-2)

28. Kaufman AN. 1984 Dissipative Hamiltonian systems: a unifying principle. *Phys. Lett. A* **100**, 419–422. (doi:10.1016/0375-9601(84)90634-0)

29. Grmela M. 1984 Bracket formulation of dissipative fluid mechanics equations. *Phys. Lett. A* **102**, 355–358. (doi:10.1016/0375-9601(84)90297-4)

30. Kraus M, Hirvijoki E. 2017 Metriplectic integrators for the Landau collision operator. *Phys. Plasmas* **24**, 102311. ([doi:10.1063/1.4998610](https://doi.org/10.1063/1.4998610))
31. Moser J. 1968 Lectures on Hamiltonian systems. *Mem. Am. Math. Soc.* 1–60.
32. Leimkuhler B, Matthews C. 2013a Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express* **2013**, 34–56. ([doi:10.1093/amrx/abs010](https://doi.org/10.1093/amrx/abs010))
33. Leimkuhler B, Matthews C. 2013b Robust and efficient configurational molecular sampling via Langevin dynamics. *J. Chem. Phys.* **138**, 174102. ([doi:10.1063/1.4802990](https://doi.org/10.1063/1.4802990))
34. Leimkuhler B, Matthews C, Stoltz G. 2016 The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.* **36**, 13–79. ([doi:10.1093/imanum/dru056](https://doi.org/10.1093/imanum/dru056))
35. Abdulle A, Vilmart G, Zygalakis KC. 2015 Long time accuracy of Lie–Trotter splitting methods for Langevin dynamics. *SIAM J. Numer. Anal.* **53**, 1–16. ([doi:10.1137/140962644](https://doi.org/10.1137/140962644))
36. Leimkuhler B, Shang X. 2015 On the numerical treatment of dissipative particle dynamics and related systems. *J. Comput. Phys.* **280**, 72–95. ([doi:10.1016/j.jcp.2014.09.008](https://doi.org/10.1016/j.jcp.2014.09.008))
37. Leimkuhler B, Shang X. 2016a Pairwise adaptive thermostats for improved accuracy and stability in dissipative particle dynamics. *J. Comput. Phys.* **324**, 174–193. ([doi:10.1016/j.jcp.2016.07.034](https://doi.org/10.1016/j.jcp.2016.07.034))
38. Leimkuhler B, Shang X. 2016b Adaptive thermostats for noisy gradient systems. *SIAM J. Sci. Comput.* **38**, A712–A736. ([doi:10.1137/15M102318X](https://doi.org/10.1137/15M102318X))
39. Shang X, Kröger M, Leimkuhler B. 2017 Assessing numerical methods for molecular and particle simulation. *Soft Matter* **13**, 8565–8578. ([doi:10.1039/C7SM01526G](https://doi.org/10.1039/C7SM01526G))
40. Matthes D, Osberger H. 2014 Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.* **48**, 697–726. ([doi:10.1051/m2an/2013126](https://doi.org/10.1051/m2an/2013126))
41. Matthes D, Plazotta S. 2019 A variational formulation of the BDF2 method for metric gradient flows. *ESAIM Math. Model. Numer. Anal.* **53**, 145–172. ([doi:10.1051/m2an/2018045](https://doi.org/10.1051/m2an/2018045))
42. Quispel GRW, McLaren DI. 2008 A new class of energy-preserving numerical integration methods. *J. Phys. A: Math. Theor.* **41**, 045206. ([doi:10.1088/1751-8113/41/4/045206](https://doi.org/10.1088/1751-8113/41/4/045206))
43. Zhong G, Marsden JE. 1988 Lie–Poisson Hamilton–Jacobi theory and Lie–Poisson integrators. *Phys. Lett. A* **133**, 134–139. ([doi:10.1016/0375-9601(88)90773-6](https://doi.org/10.1016/0375-9601(88)90773-6))
44. Simo JC, Tarnow N, Wong KK. 1992 Exact energy-momentum conserving algorithms and symplectic schemes for nonlinear dynamics. *Comput. Methods Appl. Mech. Eng.* **100**, 63–116. ([doi:10.1016/0045-7825(92)90115-Z](https://doi.org/10.1016/0045-7825(92)90115-Z))
45. Quispel GRW, Turner GS. 1996 Discrete gradient methods for solving ODEs numerically while preserving a first integral. *J. Phys. A: Math. Gen.* **29**, L341. ([doi:10.1088/0305-4470/29/13/006](https://doi.org/10.1088/0305-4470/29/13/006))
46. Cohen D, Hairer E. 2011 Linear energy-preserving integrators for Poisson systems. *BIT* **51**, 91–101. ([doi:10.1007/s10543-011-0310-z](https://doi.org/10.1007/s10543-011-0310-z))
47. McLachlan RI, Quispel GRW, Robidoux N. 1999 Geometric integration using discrete gradients. *Phil. Trans. R. Soc. A* **357**, 1021–1045. ([doi:10.1098/rsta.1999.0363](https://doi.org/10.1098/rsta.1999.0363))
48. Yoshida H. 1990 Construction of higher order symplectic integrators. *Phys. Lett. A* **150**, 262–268. ([doi:10.1016/0375-9601(90)90092-3](https://doi.org/10.1016/0375-9601(90)90092-3))
49. Reich S. 1999 Backward error analysis for numerical integrators. *SIAM J. Numer. Anal.* **36**, 1549–1570. ([doi:10.1137/S0036142997329797](https://doi.org/10.1137/S0036142997329797))
50. Aubry A, Chartier P. 1998 Pseudo-symplectic Runge–Kutta methods. *BIT* **38**, 439–461. ([doi:10.1007/BF02510253](https://doi.org/10.1007/BF02510253))
51. Shang X, Zhu Z, Leimkuhler B, Storkey AJ. 2015 Covariance-controlled adaptive Langevin thermostat for large-scale Bayesian sampling. In *Advances in Neural Information Processing Systems 28* pp. 37–45. Curran Associates.
52. Venerus DC, Öttinger HC. 2018 *A modern course in transport phenomena*. Cambridge, UK: Cambridge University Press.
53. Harten A, Lax PD, Van Leer B. 1983 On upstream differencing and Godunov-type scheme for hyperbolic conservation laws. *SIAM Rev.* **25**, 35–61. ([doi:10.1137/1025002](https://doi.org/10.1137/1025002))
54. Gonzalez O. 1996 Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.* **6**, 449–467. ([doi:10.1007/BF02440162](https://doi.org/10.1007/BF02440162))
55. Romero I. 2009 Thermodynamically consistent time-stepping algorithms for non-linear thermomechanical systems. *Int. J. Numer. Methods Eng.* **79**, 706–732. ([doi:10.1002/nme.2588](https://doi.org/10.1002/nme.2588))
56. Hesch C, Betsch P. 2011 Energy-momentum consistent algorithms for dynamic thermomechanical problems—Application to mortar domain decomposition problems. *Int. J. Numer. Methods Eng.* **86**, 1277–1302. ([doi:10.1002/nme.3095](https://doi.org/10.1002/nme.3095))

57. Groß M. 2009 Higher-order accurate and energy-momentum consistent discretisation of dynamic finite deformation thermo-viscoelasticity. Habilitation Thesis, University of Siegen.

58. Suzuki Y, Ohnawa M. 2016 GENERIC formalism and discrete variational derivative method for the two-dimensional vorticity equation. *J. Comput. Appl. Math.* **296**, 690–708. (doi:10.1016/j.cam.2015.10.018)

59. Romero I. 2010a Algorithms for coupled problems that preserve symmetries and the laws of thermodynamics: part I: monolithic integrators and their application to finite strain thermoelasticity. *Comput. Methods Appl. Mech. Eng.* **199**, 1841–1858. (doi:10.1016/j.cma.2010.02.014)

60. Romero I. 2010b Algorithms for coupled problems that preserve symmetries and the laws of thermodynamics: part II: fractional step methods. *Comput. Methods Appl. Mech. Eng.* **199**, 2235–2248. (doi:10.1016/j.cma.2010.03.016)

61. Cieśliński JL, Ratkiewicz B. 2010 Improving the accuracy of the discrete gradient method in the one-dimensional case. *Phys. Rev. E* **81**, 016704. (doi:10.1103/PhysRevE.81.016704)

62. McLachlan RI, Quispel GRW. 2014 Discrete gradient methods have an energy conservation law. *Discrete Contin. Dyn. Syst.* **34**, 1099–1104. (doi:10.3934/dcds.2014.34.1099)

63. Hernandez DM, Bertschinger E. 2018 Time-symmetric integration in astrophysics. *Mon. Not. R. Astron. Soc.* **475**, 5570–5584. (doi:10.1093/mnras/sty184)

64. Bateman H. 1931 On dissipative systems and related variational principles. *Phys. Rev.* **38**, 815–819. (doi:10.1103/PhysRev.38.815)

65. Dodin IY, Zhmoginov AI, Ruiz DE. 2017 Variational principles for dissipative (sub) systems, with applications to the theory of linear dispersion and geometrical optics. *Phys. Lett. A* **381**, 1411–1430. (doi:10.1016/j.physleta.2017.02.023)

66. Merker J, Krüger M. 2013 On a variational principle in thermodynamics. *Contin. Mech. Thermodyn.* **25**, 779–793. (doi:10.1007/s00161-012-0277-2)

67. Wendlandt JM, Marsden JE. 1997 Mechanical integrators derived from a discrete variational principle. *Physica D* **106**, 223–246. (doi:10.1016/S0167-2789(97)00051-1)