



# Order theory for discrete gradient methods

Sølve Eidnes<sup>1,2</sup>

Received: 20 August 2021 / Accepted: 17 January 2022  
© The Author(s) 2022

## Abstract

The discrete gradient methods are integrators designed to preserve invariants of ordinary differential equations. From a formal series expansion of a subclass of these methods, we derive conditions for arbitrarily high order. We derive specific results for the average vector field discrete gradient, from which we get P-series methods in the general case, and B-series methods for canonical Hamiltonian systems. Higher order schemes are presented, and their applications are demonstrated on the Hénon–Heiles system and a Lotka–Volterra system, and on both the training and integration of a pendulum system learned from data by a neural network.

**Keywords** Geometric integration · Energy preservation · B-series · Neural networks · Order of accuracy · High-order methods

**Mathematics Subject Classification** 65L05 · 65P10 · 37M15 · 05C05

## 1 Energy preservation and discrete gradient methods

For an ordinary differential equation (ODE)

$$\dot{x} = f(x), \quad x \in \mathbb{R}^d, \quad f : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad (1)$$

a first integral, or invariant, is a function  $H : \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $H(x(t)) = H(x(t_0))$  along the solution curves of (1). If we can write

---

Communicated by Christian Lubich.

---

The research was supported by the Research Council of Norway, through the projects SPIRIT (No. 231632) and BigDataMine (No. 309691).

---

✉ Sølve Eidnes  
solve.eidnes@sintef.no

<sup>1</sup> Department of Mathematical Sciences, NTNU, 7491 Trondheim, Norway

<sup>2</sup> Department of Mathematics and Cybernetics, SINTEF Digital, 0373 Oslo, Norway

$$f(x) = S(x)\nabla H(x), \quad (2)$$

where  $S(x) : \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^d$  is a skew-symmetric matrix, then (1) preserves  $H$ : this follows from the skew-symmetry of  $S(x)$ , which yields

$$\frac{d}{dt}H(x) = \nabla H(x)^T \dot{x} = \nabla H(x)^T S(x)\nabla H(x) = 0. \quad (3)$$

The converse is also true: McLachlan et al. showed in [30] that, whenever (1) has a first integral  $H$ , there exists a skew-symmetric matrix  $S(x)$ , bounded near every non-degenerate critical point of  $H$ , such that (1) can be written on what is called the *skew-gradient form*:

$$\dot{x} = S(x)\nabla H(x). \quad (4)$$

The proof provided in [30] for this is based on presenting a general form of one such  $S(x)$ , the so-called default formula

$$S(x) = \frac{f(x)\nabla H(x)^T - \nabla H(x)f(x)^T}{\nabla H(x)^T \nabla H(x)}. \quad (5)$$

Unless  $d = 2$ , this is generally not a unique choice of  $S(x)$ , as e.g.

$$S(x) = \frac{f(x)g(x)^T - g(x)f(x)^T}{g(x)^T \nabla H(x)}$$

will satisfy (2) for any non-vanishing function  $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ . Many ODEs with first integrals have a well-known skew-gradient form (4). This includes Poisson systems, and the important class consisting of canonical Hamiltonian ODEs. For the latter,  $S$  will be constant, so that we may write

$$\dot{x} = S\nabla H(x). \quad (6)$$

A numerical integrator preserving a first integral  $H$  exactly is called an integral-preserving, or *energy-preserving*, method. Starting in the late 1970s, a few energy-preserving methods were proposed which relied on some discrete analogue of the property (3), see e.g. [9, 23, 25, 26]. Most prominent among these is the class of methods called discrete gradient methods, defined formally by Gonzalez in [18] and given their current name in [30].

Given the first integral  $H$ , a discrete gradient  $\bar{\nabla}H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a function satisfying the conditions

$$\bar{\nabla}H(x, y)^T(y - x) = H(y) - H(x), \quad (7)$$

$$\bar{\nabla}H(x, x) = \nabla H(x), \quad (8)$$

for all  $x, y \in \mathbb{R}^d$ . Introducing also the discrete approximation  $\bar{S}(x, y, h)$  to  $S(x)$ , skew-symmetric and satisfying  $\bar{S}(x, x, 0) = S(x)$ , the corresponding discrete gradient

method is given by

$$\frac{\hat{x} - x}{h} = \bar{S}(x, \hat{x}, h) \bar{\nabla} H(x, \hat{x}). \tag{9}$$

This scheme satisfies a discrete analogue to (3):

$$H(\hat{x}) - H(x) = h \bar{\nabla} H(x, \hat{x})^T \bar{S}(x, \hat{x}, h) \bar{\nabla} H(x, \hat{x}) = 0.$$

We say that (9) is consistent to the skew-gradient system (4), since  $\bar{S}(x, \hat{x}, h)$  is a consistent approximation of  $S(x)$  and  $\bar{\nabla} H(x, \hat{x})$  is a consistent approximation of  $\nabla H(x)$ .

If  $d \geq 2$ , there are in general infinitely many functions satisfying (7)–(8). Many explicit definitions of concrete discrete gradients have been suggested, and we will discuss the most prominent among them in Sect. 2.1. One of these is the *average vector field (AVF) discrete gradient*, first introduced in [22] and sometimes called the mean value discrete gradient [30]. For a given  $H$ , it is given by the average of  $\nabla H$  on the segment  $[x, y]$ :

$$\bar{\nabla}_{\text{AVF}} H(x, y) = \int_0^1 \nabla H((1 - \xi)x + \xi y) \, d\xi. \tag{10}$$

When applied to the constant  $S$  system (6), the discrete gradient method with  $\bar{S}(x, y, h) = S$  and  $\bar{\nabla} H = \bar{\nabla}_{\text{AVF}} H$  coincides with the scheme

$$\frac{\hat{x} - x}{h} = \int_0^1 f((1 - \xi)x + \xi \hat{x}) \, d\xi. \tag{11}$$

This is sometimes viewed as a method by itself, applicable to any system (1), in which case it is called the average vector field (AVF) method [37]. This was shown in [6] to be a B-series method.

As pointed out in [30], the discrete gradient is restricted by its definition to be at best a second order approximation to point values of  $\nabla H$ . In much of the literature on discrete gradient methods, see e.g. [18, 21], the approximation  $\bar{S}$  is defined as being independent of  $h$ . In that case, the discrete gradient scheme (9) can at best guarantee second order convergence towards the exact solution. Over the last two decades, there have been published some notable papers on higher order discrete gradient methods. McLaren and Quispel were first out with their bootstrapping technique derived in [31, 32]. Given any discrete gradient  $\bar{\nabla} H$  and an approximation to  $S(x)$  given by  $\bar{S}(x, y, h)$ , they compare the Taylor expansion of the corresponding discrete gradient scheme to that of the exact solution, and thus find a new approximation  $\tilde{S}(x, y, h)$  to  $S(x)$  which yields higher order. This quickly becomes a very involved procedure, but by using a symmetric discrete gradient, they derive fourth order methods. A downside of this method is that the schemes of order higher than two require the calculation of tensors of order three or higher at every time step.

A fourth order generalization of the AVF method is proposed by the same authors in [37]. This can be viewed as a fourth-order discrete gradient method for all skew-gradient systems where  $S$  is constant. Also worth mentioning in this setting is the collocation-like method introduced by Hairer [20] and then generalized to Poisson

systems by Cohen and Hairer [10]. This is a multi-stage extension of the AVF discrete gradient method. To get higher than second order, more than one stage is required. In that case the method is not a discrete gradient method, although it is energy-preserving.

Norton et al. show in [35] that linear projection methods can be viewed as a class of discrete gradient methods for skew-gradient systems with  $S(x)$  given by the default formula (5). In connection to this, Norton and Quispel suggest in [36] the class of approximations to (5) given by

$$\bar{S}(x, y, h) = \frac{\tilde{f}(x, y, h)\tilde{g}(x, y, h)^T - \tilde{g}(x, y, h)\tilde{f}(x, y, h)^T}{\hat{g}(x, y, h)^T\check{g}(x, y, h)}, \quad (12)$$

where  $\tilde{f}(x, y, h)$  is a consistent approximation to  $f(x)$ , and  $\tilde{g}(x, y, h)$ ,  $\hat{g}(x, y, h)$  and  $\check{g}(x, y, h)$  are all consistent approximations to  $\nabla H(x)$ . The corresponding discrete gradient method then inherits the order of the method  $\hat{x} = x + h\tilde{f}(x, \hat{x}, h)$ .

The use of the discrete gradient method expands beyond the numerical integration of ODEs. It can be applied to the time-integration of partial differential equations (PDEs) which has constants of motion, guaranteeing preservation of a discrete approximation of an integral [5, 14, 16]. Higher-order methods have been studied in this context, but then only schemes that are expanding on the AVF method [3, 24]. Furthermore, building on the recently introduced Hamiltonian neural networks [8, 19], Matsubara et al. have shown how discrete gradients can be used to learn energy-preserving systems from data [27]. Since the energy is given by a neural network, none of the higher-order methods reviewed above are applicable in that case, and the authors suggest to use multi-step methods to get higher than second order.

To the best of our knowledge, no one has so far suggested higher than fourth order discrete gradient methods for a general skew-gradient system (4). Furthermore, for this general case, all discrete gradient methods suggested of higher than second order involve tensors of order three or higher. In this paper we present general theory for achieving methods of arbitrary order, using any discrete gradient and not depending on tensors of order higher than two. Largely inspired by the above mentioned references, especially [31, 32, 37], we present here a general form giving a class of approximations  $\bar{S}(x, y, h)$  to any  $S(x)$  in (4), with corresponding conditions for achieving an arbitrary order of the discrete gradient method (9). We do this step by step. In the next section, we derive some useful properties of a general discrete gradient and discuss the most common specific discrete gradients. Then we consider the AVF method and use order theory for B-series methods to obtain a generalization of this, with corresponding order conditions. In Sect. 4, we build on this to develop higher order discrete gradient methods for a general skew-gradient system, using the AVF discrete gradient. Then, in Sect. 5, we generalize this further to allow for a free choice of the discrete gradient, thus arriving at the general form  $\bar{S}(x, y, h)$  mentioned above, and a formal series expansion of the corresponding discrete gradient methods. Throughout the paper, we present several examples of higher order schemes for the different cases. In the final section, we apply some of these schemes to numerical examples: the Hénon–Heiles system with a constant  $S$ , a Lotka–Volterra system with a non-constant  $S$ , and a pendulum system learned by a neural network.

## 2 A preliminary analysis of discrete gradients

To simplify notation in the following derivations, we define  $g := \nabla H$ . Furthermore, we suppress the first argument of  $\bar{\nabla}H$  and define  $\bar{g}(y) := \bar{\nabla}H(x, y)$ . We use Einstein summation convention, with a comma in the subscript to differentiate between components of covectors and derivatives. That is, we write  $\bar{g}(y)^i_{,j} := \frac{\partial \bar{g}(y)^i}{\partial y^j}$ ,  $\bar{g}(y)_{i,j} := \frac{\partial \bar{g}(y)_i}{\partial y^j}$  and so forth. Taylor expanding  $\bar{g}(y)$  around  $x$ , we get

$$\begin{aligned} \bar{g}(y)^i &= \bar{g}(x)^i + \bar{g}(x)^i_{,j}(y^j - x^j) + \frac{1}{2}\bar{g}(x)^i_{,jk}(y^j - x^j)(y^k - x^k) \\ &\quad + \frac{1}{6}\bar{g}(x)^i_{,jkl}(y^j - x^j)(y^k - x^k)(y^l - x^l) + \mathcal{O}(|y - x|^4), \end{aligned} \tag{13}$$

or

$$\bar{g}(y) = \sum_{\kappa=0}^{\infty} \frac{1}{\kappa!} \bar{g}^{(\kappa)}(x)(y - x)^\kappa. \tag{14}$$

By the consistency criterion (8), we have  $\bar{g}(x) = g(x)$ . However, if we require the discrete gradient to be a differentiable function in its second argument, (8) follows directly from (7). To see this, we write (7) as

$$H(y) - H(x) = \bar{g}(y)_i(y^i - x^i). \tag{15}$$

Differentiating this with respect to  $y^j$ , we get

$$g(y)_j = H(y)_{,j} = \bar{g}(y)_{i,j}(y^i - x^i) + \bar{g}(y)_j, \tag{16}$$

The case  $y = x$  immediately gives  $g(x)_j = \bar{g}(x)_j$ , or (8). Assuming further that  $\bar{\nabla}H \in C^2(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ , we can differentiate once more to get

$$g(y)_{j,k} = H(y)_{,jk} = \bar{g}(y)_{i,jk}(y^i - x^i) + \bar{g}(y)_{j,k} + \bar{g}(y)_{k,j}, \tag{17}$$

which means that

$$g(x)_{j,k} = H(x)_{,jk} = \bar{g}(x)_{j,k} + \bar{g}(x)_{k,j},$$

or

$$D^2H(x) = D_2\bar{\nabla}H(x, x) + (D_2\bar{\nabla}H(x, x))^T, \tag{18}$$

where  $D^2H := D\nabla H$  denotes the Hessian of  $H$ , and  $D_2\bar{\nabla}H$  denotes the Jacobian of  $\bar{\nabla}H$  with respect to its second argument.

**Lemma 1** *If the discrete gradient  $\bar{\nabla}H$  is symmetric, i.e.  $\bar{\nabla}H(x, y) = \bar{\nabla}H(y, x)$  for all  $x, y \in \mathbb{R}^d$ , then*

$$D_2\bar{\nabla}H(x, x) = \frac{1}{2}D^2H(x). \tag{19}$$

**Proof** Disclosing the suppressed argument  $x$  in (16), we have

$$g(y)_j = \frac{\partial}{\partial y^j} (\bar{g}(x, y)_i)(y^i - x^i) + \bar{g}(x, y)_j,$$

which we can differentiate by  $x^k$  to get

$$0 = \frac{\partial^2}{\partial x^k \partial y^j} (\bar{g}(x, y)_i)(y^i - x^i) - \frac{\partial}{\partial y^j} \bar{g}(x, y)_k + \frac{\partial}{\partial x^k} \bar{g}(x, y)_j.$$

If  $\bar{\nabla}H$  is symmetric,

$$\frac{\partial}{\partial x^k} \bar{g}(x, y)_j = \frac{\partial}{\partial x^k} \bar{g}(y, x)_j.$$

Thus, for  $y = x$  we get  $\bar{g}(x)_{k,j} = \bar{g}(x)_{j,k}$ , or  $(D_2 \bar{\nabla}H(x, x))^T = D_2 \bar{\nabla}H(x, x)$ . Inserting that in (18), we obtain (19).

**Definition 1** Given a discrete gradient  $\bar{\nabla}H \in C^1(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ , we define the function  $Q : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$  by

$$Q(x, y) := \frac{1}{2} \left( (D_2 \bar{\nabla}H(x, y))^T - D_2 \bar{\nabla}H(x, y) \right). \quad (20)$$

Note that  $Q(x, y)$  is a skew-symmetric matrix. From (18), we see that  $Q(x, x) = \frac{1}{2} D^2 H(x) - D_2 \bar{\nabla}H(x, x)$ . The following result is crucial for the the main result of this paper: the development of a general theory for higher order discrete gradient methods.

**Lemma 2** For a discrete gradient  $\bar{\nabla}H \in C^p(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$  and the corresponding  $Q$  given by (20),

$$\begin{aligned} D_2^\kappa \bar{\nabla}H(x, x)v^\kappa &= \frac{1}{\kappa + 1} D^\kappa \nabla H(x)v^\kappa \\ &\quad - \frac{2\kappa}{\kappa + 1} D_2^{\kappa-1} Q(x, x)v^\kappa \quad \text{for any } \kappa \in [1, p], v \in \mathbb{R}^d. \end{aligned}$$

**Proof** Differentiating (17)  $\kappa - 1$  times by  $y$  and setting  $y = x$ , we find that the  $\kappa$ th derivatives of  $g(x)$  can be expressed by the  $\kappa$ th derivatives of  $\bar{g}(x)$  through the relation

$$g(x)_{j,I} = \bar{g}(x)_{j,I} + \sum_{m=1}^{\kappa} \bar{g}(x)_{i_m, \{j, I_m\}}, \quad \text{for all } j, I, \kappa, \quad (21)$$

where  $I = \{i_1, i_2, \dots, i_\kappa\}$  is an ordered set of  $\kappa$  indices, and

$$I_m = I \setminus \{i_m\} = \{i_1, i_2, \dots, i_{m-1}, i_{m+1}, \dots, i_\kappa\},$$

i.e.  $I$  with the  $m$ th element excluded. Similarly, by continued differentiation of (20), we obtain

$$(D_2^{\kappa-1} Q(x, x))_{j,I} = \frac{1}{2} \bar{g}(x)_{i_1, \{j, I_1\}} - \frac{1}{2} \bar{g}(x)_{j,I}.$$

Thus

$$\begin{aligned} (D_2^{\kappa-1} Q(x, x)v^\kappa)_j &= (D_2^{\kappa-1} Q(x, x))_{j,I} v^I = \frac{1}{2} \bar{g}(x)_{i_1, \{j, I_1\}} v^{\{j, I_1\}} - \frac{1}{2} \bar{g}(x)_{j,I} v^I \\ &= \frac{1}{2\kappa} \sum_{m=1}^{\kappa} \bar{g}(x)_{i_m, \{j, I_m\}} v^{\{j, I_m\}} + \frac{1}{2\kappa} \bar{g}(x)_{j,I} v^I - \frac{1}{2\kappa} \bar{g}(x)_{j,I} v^I - \frac{1}{2} \bar{g}(x)_{j,I} v^I \\ &= \frac{1}{2\kappa} g(x)_{j,I} v^I - \left(\frac{1}{2\kappa} + \frac{1}{2}\right) \bar{g}(x)_{j,I} v^I = \frac{1}{2\kappa} g(x)_{j,I} v^I - \frac{\kappa+1}{2\kappa} \bar{g}(x)_{j,I} v^I. \end{aligned}$$

### 2.1 Properties of different discrete gradients

While introducing the discrete gradient methods in [18], Gonzalez also gave an example of a discrete gradient satisfying (7)–(8): the *midpoint discrete gradient* is given by

$$\bar{\nabla}_M H(x, y) := \nabla H\left(\frac{x+y}{2}\right) + \frac{H(y) - H(x) - \nabla H\left(\frac{x+y}{2}\right)^T (y-x)}{(y-x)^T (y-x)} (y-x).$$

Even when  $H$  is analytic, this discrete gradient is often not; the second order partial derivatives are in general singular in  $y = x$ . For that reason, it is not suited for achieving higher order methods by the techniques we consider in this paper.

The *Itoh–Abe discrete gradient method*, introduced in [23], notably does not require evaluation of the gradient. The corresponding discrete gradient, which has also been called the coordinate increment discrete gradient [30], is defined by

$$\bar{\nabla}_{IA} H(x, y) := \sum_{j=1}^d \alpha_j e_j, \tag{22}$$

where  $e_j$  is the  $j^{\text{th}}$  canonical unit vector and

$$\alpha_j = \begin{cases} \frac{H(w_j) - H(w_{j-1})}{y^j - x^j} & \text{if } y^j \neq x^j, \\ \frac{\partial H}{\partial x^j}(w_{j-1}) & \text{if } y^j = x^j, \end{cases}$$

$$w_j = \sum_{i=1}^j y^i e_i + \sum_{i=j+1}^n x^i e_i.$$

While the other discrete gradients we consider in this paper are symmetric and thus second order approximations to  $\nabla H$ , the Itoh–Abe discrete gradient is only of first

order. However, a second order discrete gradient, which we call the *symmetrized Itoh–Abe (SIA) discrete gradient*, is given by

$$\bar{\nabla}_{\text{SIA}} H(x, y) := \frac{1}{2} \left( \bar{\nabla}_{\text{IA}} H(x, y) + \bar{\nabla}_{\text{IA}} H(y, x) \right). \quad (23)$$

Furihata presented the discrete variational derivative method for a class of PDEs in [16], a method which has been developed further by Furihata, Matsuo and co-authors in a series of papers, e.g. [29, 38], as well as the monograph [17]. As shown in [14], these schemes can also be obtained by semi-discretizing the PDE in space and then applying a discrete gradient method on the resulting system of ODEs. We give here the definition of the specific discrete gradient that gives the schemes of Furihata and co-author, defined for a class of invariants that includes all polynomial functions:

**Definition 2** Assume that we can write the first integral as

$$H(x) = \sum_l c_l \prod_{j=1}^d f_j^l(x^j), \quad (24)$$

for functions  $f_j^l : \mathbb{R} \rightarrow \mathbb{R}$ . The *Furihata discrete gradient*  $\bar{\nabla}_{\text{F}} H(x, y)$  is defined by

$$\bar{\nabla}_{\text{F}} H(x, y) := \sum_{j=1}^d \alpha_j e_j, \quad (25)$$

where  $e_j$  is the  $j^{\text{th}}$  canonical unit vector and

$$\alpha_j = \begin{cases} \sum_l \frac{c_l}{2} \frac{f_j^l(y^j) - f_j^l(x^j)}{y^j - x^j} \left( \prod_{k=1}^{j-1} f_k^l(x^k) + \prod_{k=1}^{j-1} f_k^l(y^k) \right) \prod_{k=j+1}^d \frac{f_k^l(x^k) + f_k^l(y^k)}{2} & \text{if } y^j \neq x^j, \\ \sum_l \frac{c_l}{2} \frac{d f_j^l(x^j)}{d x^j} \left( \prod_{k=1}^{j-1} f_k^l(x^k) + \prod_{k=1}^{j-1} f_k^l(y^k) \right) \prod_{k=j+1}^d \frac{f_k^l(x^k) + f_k^l(y^k)}{2} & \text{if } y^j = x^j. \end{cases}$$

The discrete gradient introduced by Matsubara and co-authors in [27] shares some relation to the Furihata discrete gradient, in that they are both relying on a discrete analogue to the product rule for derivatives. But the former discrete gradient, obtained by what the authors call the *automatic discrete differentiation algorithm*, more importantly depends on an analogue to the chain rule. By using this algorithm instead of standard automatic differentiation in a neural network that learns a dynamical system from data, the discrete gradient is obtained together with the preserved energy. Thus the learned system can be integrated in an energy-preserving manner, which distinguishes the approach of Matsubara et al. from comparable studies [8, 19, 39].

Lastly we consider the AVF discrete gradient (10), which has several noteworthy characteristics.

**Lemma 3** *The  $Q(x, y)$  corresponding to the AVF discrete gradient is the zero matrix, since  $(D_2 \bar{\nabla}_{\text{AVF}} H(x, y))^T = D_2 \bar{\nabla}_{\text{AVF}} H(x, y)$ .*



**Proof** For  $\bar{g}(y) := \bar{\nabla}_{AVF}H(x, y)$ , we have

$$\begin{aligned} \bar{g}(y)_{i,j} &= \frac{\partial}{\partial y^j} \int_0^1 g((1-\xi)x + \xi y)_i \, d\xi = \int_0^1 \frac{\partial}{\partial y^j} g((1-\xi)x + \xi y)_i \, d\xi \\ &= \int_0^1 \xi g((1-\xi)x + \xi y)_{i,j} \, d\xi = \int_0^1 \xi g((1-\xi)x + \xi y)_{j,i} \, d\xi = \bar{g}(y)_{j,i}. \end{aligned}$$

From the Integrability Lemma (see e.g. [21, Lemma VI.2.7]) and the above, we have that the AVF discrete gradient defines a gradient vector field:

**Corollary 1** *The AVF discrete gradient is the gradient with respect to the second argument of a function  $\tilde{H}(x, y)$ . That is,*

$$\bar{\nabla}_{AVF}H(x, y) = \nabla_2 \tilde{H}(x, y),$$

for some  $\tilde{H} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$  and all  $x, y \in \mathbb{R}^d$ .

Any other discrete gradient will have these properties only for functions  $H$  for which it coincides with the AVF discrete gradient:

**Proposition 1** *The AVF discrete gradient is the unique discrete gradient satisfying  $(D_2 \bar{\nabla}H(x, y))^T = D_2 \bar{\nabla}H(x, y)$  for all  $H, x$  and  $y$ , and it has the formal expansion*

$$\bar{\nabla}_{AVF}H(x, y) = \sum_{\kappa=0}^{\infty} \frac{1}{(\kappa + 1)!} D^\kappa \nabla H(x)(y - x)^\kappa. \tag{26}$$

**Proof** Assume that  $\nabla H$  is an analytic function. As in the proof of Lemma 2, let  $I = \{i_1, i_2, \dots, i_\kappa\}$  be an ordered set of  $\kappa$  indices, and let  $I_m$  be  $I$  with the  $m^{\text{th}}$  element excluded. If  $\bar{g}(y)_{i,j} = \bar{g}(y)_{j,i}$  for all  $i, j$ , then also

$$\bar{g}(y)_{i,I} = \bar{g}(y)_{i_m, \{i, I_m\}} \quad \text{for all } i, I, m. \tag{27}$$

Inserting (27) in (21) we get  $g^{(\kappa)}(x) = (1 + \kappa)\bar{g}^{(\kappa)}(x)$ . Then inserting this for  $\bar{g}^{(\kappa)}(x)$  in (14), we get (26), which uniquely defines the AVF discrete gradient.

A consequence of the above result is that the AVF discrete gradient is the unique discrete gradient for which the scheme (9) with  $\bar{S}(x, \hat{x}, h) = S$  is a B-series method when applied to the system (6).

As we see from the above definitions and discussion, each of the discrete gradients have their advantages and disadvantages. Gonzalez’ midpoint discrete gradient is easily calculated from the energy  $H$  and the gradient  $\nabla H$ , but it is in general only once differentiable. The Itoh–Abe and Furihata discrete gradient methods do not require knowledge of the gradient, but the former is only a first order method and the latter is only defined for  $H$  of the form (24). The AVF discrete gradient is the unique discrete gradient whose series expansion is given by the differentials of the gradient. It does however require an integral to be calculated. The discrete gradient of Matsubara and

co-authors has a different area of use than the others: it is only defined when  $H$  is a neural network. In that case neither Gonzelez' midpoint, the AVF or the Furihata discrete gradients can be obtained.

In the next section we consider the specific case where the AVF discrete gradient is chosen and  $S$  is constant, so that we can build on established results for B-series methods to develop new theory.

### 3 A generalization of the AVF method

Let us recall the concept of B-series. Referring to the definitions in [21, Section III.1], we let  $T$  be the set of rooted trees, built recursively from starting with  $\tau = \bullet$  and letting  $\tau = [\tau_1, \dots, \tau_m]$  be the tree obtained by grafting the roots of the trees  $\tau_1, \dots, \tau_m$  to a new root. Furthermore,  $F(\tau)$  is the elementary differential associated with the tree  $\tau$ , defined by  $F(\bullet)(x) = f(x)$  and

$$F(\tau)(x) = f^{(m)}(x)(F(\tau_1)(x), \dots, F(\tau_m)(x)),$$

and  $\sigma(\tau)$  is the symmetry coefficient for  $\tau$ , defined by  $\sigma(\bullet) = 1$  and

$$\sigma(\tau) = \sigma(\tau_1) \cdots \sigma(\tau_m) \cdot \mu_1! \mu_2! \cdots, \quad (28)$$

where the integers  $\mu_1, \mu_2, \dots$  count equal trees among  $\tau_1, \dots, \tau_m$ . Then, if  $\phi : T \cup \{\emptyset\} \rightarrow \mathbb{R}$  is an arbitrary map, a *B-series* is a formal series

$$B(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in T} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau) F(\tau)(x). \quad (29)$$

The exact solution of (1) can be written as the B-series  $B(\frac{1}{\gamma}, x)$ , where the coefficient  $\gamma$  satisfies  $\gamma(\emptyset) = \gamma(\bullet) = 1$  and

$$\gamma(\tau) = |\tau| \gamma(\tau_1) \cdots \gamma(\tau_m) \quad \text{for } \tau = [\tau_1, \dots, \tau_m], \quad (30)$$

where  $|\tau|$  is the order, i.e. the number of nodes, of  $\tau$ .

**Definition 3** The *generalized AVF method* is given by

$$\begin{aligned} \frac{\hat{x} - x}{h} &= \left( I + \sum_{n=2}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^n f'(z_{njk}) + (-1)^n \prod_{k=1}^n f'(z_{nj(n-k+1)}) \right) \right) \\ &\cdot \int_0^1 f((1-\xi)x + \xi\hat{x}) d\xi, \end{aligned} \quad (31)$$

where each  $z_{njk} := z_{njk}(x, \hat{x}, h) = B(\phi_{njk}, x)$  can be written as a B-series with  $\phi(\emptyset) = 1$ .

**Theorem 1** When applied to (1) with  $f(x) = S\nabla H(x)$ , where  $S$  is a constant skew-symmetric matrix, the scheme (31) preserves  $H$ , in that  $H(\hat{x}) = H(x)$ .

**Proof** With  $f(x) = S\nabla H(x)$ , (31) becomes

$$\frac{\hat{x} - x}{h} = \bar{S}(x, \hat{x}, h)\bar{\nabla}_{AVF}H(x, \hat{x}),$$

with

$$\bar{S}(x, \hat{x}, h) = S + \sum_{n=2}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^n SD^2H(z_{nj k}) + (-1)^n \prod_{k=1}^n SD^2H(z_{nj(n-k+1)}) \right) S.$$

We have

$$\begin{aligned} & \left( \prod_{k=1}^n SD^2H(z_{nj k}) \cdot S + (-1)^n \prod_{k=1}^n SD^2H(z_{nj(n-k+1)}) \cdot S \right)^T \\ &= S^T \prod_{k=1}^n (D^2H(z_{nj(n-k+1)})^T S^T) + (-1)^n S^T \prod_{k=1}^n (D^2H(z_{nj k})^T S^T) \\ &= (-1)^{i+1} S \prod_{k=1}^n D^2H(z_{nj(n-k+1)}) S - S \prod_{k=1}^n D^2H(z_{nj k}) S \\ &= - \left( \prod_{k=1}^n SD^2H(z_{nj k}) \cdot S + (-1)^n \prod_{k=1}^n SD^2H(z_{nj(n-k+1)}) \cdot S \right), \end{aligned}$$

and thus  $\bar{S}(x, \hat{x}, h)$  is a skew-symmetric matrix.

Before considering the order conditions of the generalized AVF method, let us recall a couple of results from the literature on B-series.

**Lemma 4** ([21, Lemma III.1.9]) Let  $B(a, x)$  be a B-series with  $a(\emptyset) = 1$ . Then  $hf(B(a, x)) = B(a', x)$  is also a B-series, with  $a'(\emptyset) = 0$ ,  $a'(\bullet) = 1$  and otherwise

$$a'(\tau) = a(\tau_1) \cdots a(\tau_m) \quad \text{for } \tau = [\tau_1, \dots, \tau_m].$$

**Lemma 5** ([34, Theorem 2.2]) Let  $B(a, x)$  and  $B(b, x)$  be two B-series with  $a(\emptyset) = 1$  and  $b(\emptyset) = 0$ . Then  $hf'(B(a, x))B(b, x) = B(a \times b, x)$ , i.e. a B-series, with  $(a \times b)(\emptyset) = (a \times b)(\bullet) = 0$  and otherwise

$$(a \times b)(\tau) = \sum_{i=1}^m \prod_{j=1, j \neq i}^m a(\tau_j) b(\tau_i) \quad \text{for } \tau = [\tau_1, \dots, \tau_m].$$

Proposition 1 in [6] states that the standard AVF method is a B-series method. We build on the proof of that proposition to prove the following result.

**Proposition 2** *The generalized AVF method (31) is a B-series method.*

**Proof** First we define  $\hat{e} : T \cup \{\emptyset\} \rightarrow \mathbb{R}$  by  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ . Then, assuming that the solution  $\hat{x}$  of (31) can be written as the B-series  $\hat{x} = B(\Phi, x)$ , we find the B-series

$$\begin{aligned} h \int_0^1 f((1 - \xi)x + \xi \hat{x}) \, d\xi &= h \int_0^1 f(B((1 - \xi)\hat{e} + \xi\Phi, x)) \, d\xi \\ &= \int_0^1 B(((1 - \xi)\hat{e} + \xi\Phi)', x) \, d\xi \\ &= B\left(\int_0^1 ((1 - \xi)\hat{e} + \xi\Phi)' \, d\xi, x\right). \end{aligned}$$

Setting  $\theta := \int_0^1 ((1 - \xi)\hat{e} + \xi\Phi)' \, d\xi = \int_0^1 ((1 - \xi)\hat{e})' \, d\xi + \int_0^1 (\xi\Phi)' \, d\xi = \int_0^1 (\xi\Phi)' \, d\xi$ , we get

$$\theta(\emptyset) = 0, \quad \theta(\bullet) = 1, \quad \theta([\tau_1, \dots, \tau_m]) = \frac{1}{m + 1} \Phi(\tau_1) \cdots \Phi(\tau_m). \tag{32}$$

Then we may rewrite (31) as

$$\begin{aligned} \hat{x} &= x + \left( I + \sum_{n=2}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^n f'(B(\phi_{njk}, x)) + (-1)^n \prod_{k=1}^n f'(B(\phi_{nj(n-k+1)}, x)) \right) \right) B(\theta, x) \\ &= x + B(\theta, x) \\ &\quad + \sum_{n=2}^{p-1} \sum_j b_{nj} (B(\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta, x) + (-1)^n B(\phi_{njn} \times \cdots \times \phi_{nj1} \times \theta, x)) \\ &= B(\Phi, x), \end{aligned}$$

with

$$\Phi = \hat{e} + \theta + \sum_{n=2}^{p-1} \sum_j b_{nj} (\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta + (-1)^n \phi_{njn} \times \cdots \times \phi_{nj1} \times \theta). \tag{33}$$

Comparing the B-series of the exact solution and the B-series of the solution of (31), and noting that the elementary differentials are independent, we immediately get the following result.

**Theorem 2** *The generalized AVF method (31) is of order  $p$  if and only if*

$$\Phi(\tau) = \frac{1}{\gamma(\tau)} \quad \text{for } |\tau| \leq p, \tag{34}$$

**Table 1** Elementary differentials and their coefficients in the B-series of the solution of (31), up to fourth order

$ \tau $	$F(\tau)^i$	$\tau$	$\sigma(\tau)$	$\gamma(\tau)$	$\Phi(\tau)$
1	$f^i$	$\bullet$	1	1	1
2	$f_j^i f^j$	$\begin{array}{c} \bullet \\   \\ \bullet \end{array}$	1	2	$\frac{1}{2}$
3	$f_{jk}^i f^j f^k$	$\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}$	2	3	$\frac{1}{3}$
4	$f_j^i f_k^j f^k$	$\begin{array}{c} \bullet \\   \\ \bullet \end{array}$	1	6	$\frac{1}{4} + 2 \sum_j b_{2j}$
	$f_{jkl}^i f^j f^k f^l$	$\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}$	6	4	$\frac{1}{4}$
	$f_{jk}^i f^j f_l^k f^l$	$\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}$	1	8	$\frac{1}{6} + \sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet)$
	$f_j^i f_{kl}^j f^k f^l$	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	2	12	$\frac{1}{6} + 2 \sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet)$
	$f_j^i f_k^j f_l^k f^l$	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	1	24	$\frac{1}{8} + 2 \sum_j b_{2j}$

where  $\Phi$  is given by (33) and  $\gamma$  is given by (30).

The terms  $\Phi(\tau)$  can be found from (33) by applying Lemma 5 recursively, as illustrated by the following example.

**Example 1** Consider  $\tau = \begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}$ , and assume we have found  $\Phi$  for all trees up to and including order four already, as given in Table 1. We have

$$\theta(\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}) = \frac{1}{3} \Phi(\bullet) \Phi(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) = \frac{1}{3} (\frac{1}{4} + 2 \sum_j b_{2j}) = \frac{1}{12} + \frac{2}{3} \sum_j b_{2j}.$$

Then we calculate

$$\begin{aligned} (\phi_{2j_1} \times \phi_{2j_2} \times \theta)(\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}) &= \phi_{2j_1}(\bullet)(\phi_{2j_2} \times \theta)(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) + \phi_{2j_1}(\begin{array}{c} \bullet \\ | \\ \bullet \end{array})(\phi_{2j_2} \times \theta)(\bullet) \\ &= \phi_{2j_1}(\bullet)(\phi_{2j_2} \times \theta)(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) = \phi_{2j_1}(\bullet) \phi_{2j_2}(\emptyset) \theta(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) = \frac{1}{2} \phi_{2j_1}(\bullet), \end{aligned}$$

where we have used in the second equality that  $(\phi_{2j_2} \times \theta)(\bullet) = \phi_{2j_2}(\emptyset) \theta(\emptyset) = 0$ .

Similarly we find  $(\phi_{2j_2} \times \phi_{2j_1} \times \theta)(\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}) = \frac{1}{2} \phi_{2j_2}(\bullet)$ . Furthermore,

$$\begin{aligned} (\phi_{3j_1} \times \phi_{3j_2} \times \phi_{3j_3} \times \theta)(\begin{array}{c} \bullet \\ \vee \\ \bullet \quad \bullet \end{array}) &= \phi_{3j_1}(\bullet)(\phi_{3j_2} \times \phi_{3j_3} \times \theta)(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) \\ &= \phi_{3j_1}(\bullet) \phi_{3j_2}(\emptyset) (\phi_{3j_3} \times \theta)(\begin{array}{c} \bullet \\ | \\ \bullet \end{array}) \\ &= \phi_{3j_1}(\bullet) \phi_{3j_3}(\emptyset) \theta(\bullet) = \phi_{3j_1}(\bullet), \end{aligned}$$

and  $(\phi_{3j3} \times \phi_{3j2} \times \phi_{3j1} \times \theta)(\mathfrak{V}) = \phi_{3j3}(\bullet)$ . Hence,

$$\begin{aligned} \Phi(\mathfrak{V}) &= \frac{1}{12} + \frac{2}{3} \sum_j b_{2j} + \frac{1}{2} \sum_j b_{2j}(\phi_{2j1}(\bullet) + \phi_{2j2}(\bullet)) \\ &\quad + \sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)). \end{aligned}$$

Now, if we assume the order condition (34) to be satisfied for all trees up to and including order four, we can replace  $\sum_j b_{2j} = -\frac{1}{24}$  and  $\sum_j b_{2j}(\phi_{2j1}(\bullet) + \phi_{2j2}(\bullet)) = -\frac{1}{24}$  in the above expression, use that  $\gamma(\mathfrak{V}) = 30$ , and get that (34) is satisfied for  $\mathfrak{V}$  if and only if

$$\sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = -\frac{1}{720}. \tag{35}$$

### 3.1 Construction of higher order schemes

As the size of the trees grows, finding  $\Phi(\tau)$  from (33) can become quite a cumbersome operation. Furthermore, we observe from Table 1 that there are some equivalent order conditions for different trees. Before presenting more convenient techniques for finding order conditions for the generalized AVF method, let us define some more concepts related to B-series and trees.

First, recall that the Butcher product of two trees  $u = [u_1, \dots, u_m]$  and  $v = [v_1, \dots, v_n]$  is given by  $u \circ v = [u_1, u_2, \dots, u_m, v]$ . This operation is neither associative nor commutative, and in contrast to the practice in [21], we here take the product of several factors without parentheses to mean evaluation from right to left:

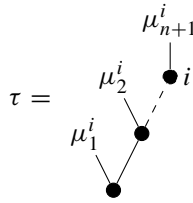
$$u_1 \circ u_2 \circ \dots \circ u_k := u \circ (u_2 \circ (\dots \circ u_k)).$$

Given a forest  $\mu = (\tau_1, \dots, \tau_m)$ , the tree obtained by grafting the roots of every tree in  $\mu$  to a new root is denoted by  $[\mu] = [\tau_1, \dots, \tau_m]$ . Moreover,  $\mu^{-1}(\tau)$  denotes the forest such that  $[\mu^{-1}(\tau)] = \tau$ . We extend the maps  $\phi : T \cup \{\emptyset\} \rightarrow \mathbb{R}$  and  $\gamma : T \cup \{\emptyset\} \rightarrow \mathbb{R}$  to forests by the letting  $\phi(\mu) = \prod_{i=1}^m \phi(\tau_i)$  and  $\gamma(\mu) = \prod_{i=1}^m \gamma(\tau_i)$  for  $\mu = (\tau_1, \dots, \tau_m)$ .

Consider now a tree  $\tau$  consisting of  $|\tau|$  nodes. We may number every tree from 1 to  $|\tau|$ , starting at the root and going from left to right on the increasing levels above. For a given node  $i \in \{1, \dots, |\tau|\}$  on level  $n + 1$ , there exists a unique set of forests  $\hat{\tau}^i = \{\mu_1^i, \dots, \mu_{n+1}^i\}$  such that

$$\tau = [\mu_1^i] \circ [\mu_2^i] \circ \dots \circ [\mu_{n+1}^i].$$

That is, labeling node  $i$ ,



**Proposition 3** The  $\Phi$  of (34) can alternatively be found by

$$\Phi(\tau) = \hat{e}(\tau) + \theta(\tau) + \sum_{i \text{ s.t. } n \geq 2} \Lambda(\hat{\tau}^i) \tag{36}$$

where  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ ,  $\theta(\emptyset) = 0$ ,  $\theta(\bullet) = 1$ ,

$$\theta([\tau_1, \dots, \tau_m]) = \frac{1}{m + 1} \Phi(\tau_1) \cdots \Phi(\tau_m),$$

and

$$\Lambda(\hat{\tau}^i) = \theta([\mu_{n+1}^i]) \sum_j b_{nj} \left( \phi_{nj1}(\mu_1^i) \cdots \phi_{njn}(\mu_n^i) + (-1)^n \phi_{njn}(\mu_1^i) \cdots \phi_{nj1}(\mu_n^i) \right). \tag{37}$$

**Proof** Define  $n_i$  so that  $n_i + 1$  is the level of node  $i$ . Collect the children of node  $i$  in the set  $C_i$ . We have

$$[\mu_{n_i+1}^i] = [\mu_{n_k}^k] \circ [\mu_{n_k+1}^k] \quad \text{for all } k \in C_i,$$

and thus

$$(a \times b)([\mu_{n_i+1}^i]) = \sum_{k \in C_i} a(\mu_{n_k}^k) b([\mu_{n_k+1}^k]).$$

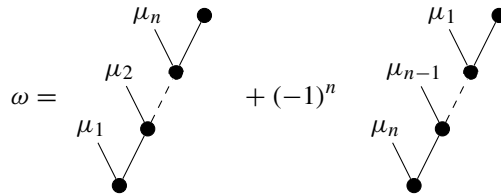
Note also that  $\mu_{n_i}^i = \mu_{n_i}^k = \mu_{n_k-1}^k$  if  $k \in C_i$ . Then we get

$$\begin{aligned} (\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta)(\tau) &= (\phi_{nj1} \times \cdots \times \phi_{njn} \times \theta)([\mu_1^1]) \\ &= \sum_{i_1 \in C_1} \phi_{nj1}(\mu_{i_1}^1) (\phi_{nj2} \times \cdots \times \phi_{njn} \times \theta)([\mu_2^{i_1}]) \\ &= \sum_{i_1 \in C_1} \phi_{nj1}(\mu_{i_1}^1) \sum_{i_2 \in C_{i_1}} \phi_{nj2}(\mu_2^{i_2}) (\phi_{nj3} \times \cdots \times \phi_{njn} \times \theta)([\mu_3^{i_2}]) \\ &= \sum_{i_1 \in C_1} \sum_{i_2 \in C_{i_1}} \phi_{nj1}(\mu_{i_1}^{i_2}) \phi_{nj2}(\mu_2^{i_2}) (\phi_{nj3} \times \cdots \times \phi_{njn} \times \theta)([\mu_3^{i_2}]) \\ &\vdots \\ &= \sum_{i_1 \in C_1} \sum_{i_2 \in C_{i_1}} \cdots \sum_{i_n \in C_{i_{n-1}}} \phi_{nj1}(\mu_{i_1}^{i_n}) \cdots \phi_{njn}(\mu_n^{i_n}) \theta([\mu_{n+1}^{i_n}]) \end{aligned}$$

$$= \sum_{i \text{ on level } n+1} \phi_{nj1}(\mu_1^i) \cdots \phi_{njn}(\mu_n^i) \theta([\mu_{n+1}^i]).$$

Inserting this and the corresponding result for  $(\phi_{njn} \times \cdots \times \phi_{nj1} \times \theta)(\tau)$  in (33), we get (37).

In [7, 15], conditions are derived for a B-series method to be energy preserving when applied to the system (6). In [37], while giving the AVF method as one such method, Quispel and McLaren present a general form of what they call energy-preserving linear combinations of rooted trees:



Here we give their result as a lemma, which is proved later by the proof of the more general Theorem 5.

**Lemma 6** *Let  $\mu_1, \dots, \mu_n$  be  $n$  arbitrary forests. Then, if  $f(x) = S\nabla H(x)$  for some skew-symmetric constant matrix  $S$ , we have that  $F(\omega)(x) \cdot \nabla H(x) = 0$  for*

$$\omega = [\mu_1] \circ [\mu_2] \circ \cdots [\mu_n] \circ [\emptyset] + (-1)^n [\mu_n] \circ [\mu_{n-1}] \circ \cdots [\mu_1] \circ [\emptyset]. \tag{38}$$

There is a connection between (36) and Lemma 6 such that instead of order conditions for every tree, we can calculate order conditions for every energy-preserving linear combination. To see this we start by collecting the leaf nodes, i.e. nodes with no children, of the tree  $\tau$  in a set  $I_l$  and the other nodes in the set  $I_n$ . If node  $i \in I_n$ , we may then use the relation

$$\Lambda(\{\mu_1^i, \dots, \mu_n^i, \mu_{n+1}^i\}) = \theta([\mu_{n+1}^i])\Lambda(\{\mu_1^i, \dots, \mu_n^i, \emptyset\})$$

to find  $\Lambda(\hat{\tau}^i)$  from the previously calculated  $\Lambda$  for a smaller tree. Then if lower order conditions are satisfied, we have numerical values for these  $\Lambda$ . The leaf nodes on the other hand, with their corresponding  $\hat{\tau}^i = \{\mu_1^i, \dots, \mu_n^i, \emptyset\}$ , gives an energy-preserving linear combination (38) which  $\tau$  belongs to. If  $i$  is on level two, this combination is simply  $\tau - \tau = 0$ , and accordingly  $\Lambda$  is not calculated for these nodes in (36). Moreover, leaves on the same level have identical  $\hat{\tau}^i$ . Thus, a tree with leaves on  $m$  different levels above level two will belong to at most  $m$  non-zero energy-preserving linear combinations (38). For each of these combinations there is a corresponding order condition, with the left hand side given by (37). The right hand side can be found by considering the individual trees.

If we assume the conditions for order  $< p$  to be satisfied, we may replace (34) by

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i) = \frac{1}{\gamma(\tau)} - \hat{e}(\tau) - \sum_{i \in I_n} \frac{\Lambda(\{\mu_1^i, \dots, \mu_n^i, \emptyset\})}{(|\mu_{n+1}^i| + 1)\gamma(\mu_{n+1}^i)}, \tag{39}$$



where  $|\mu|$  denotes the number of trees in the forest  $\mu$ . Note that  $\Lambda(\{\emptyset\}) = 1$  and hence  $\Lambda(\hat{\tau}^1) = \theta(\tau)$ . Then we can calculate the numerical value for the right hand side and, if  $\tau$  has leaves on only one level  $> 2$ , find an order condition for both  $\tau$  and the other tree in the combination (38). This warrants an example.

**Example 2** Consider again the tree  $\tau = \mathfrak{V}$ , which is part of the energy-preserving linear combination  $\omega = \mathfrak{V} - \mathfrak{Y}$ . Ignoring the two nodes on level 2, there are three nodes to calculate  $\Lambda$  for:  $i = 1, i = 4$  and  $i = 5$ . We find

$$\begin{aligned} \Lambda(\hat{\tau}^1) &= \frac{1}{(2+1)\gamma(\bullet)\gamma(\mathfrak{I})} = \frac{1}{3 \cdot 1 \cdot 6} = \frac{1}{18} \\ \Lambda(\hat{\tau}^4) &= \frac{1}{(1+1)\gamma(\bullet)} \Lambda(\{\bullet, \emptyset, \emptyset\}) \\ &= \frac{1}{2\gamma(\bullet)} \left( \frac{1}{\gamma(\mathfrak{V})} - \frac{1}{3\gamma(\bullet)\gamma(\mathfrak{I})} \right) = \frac{1}{2} \left( \frac{1}{8} - \frac{1}{6} \right) = -\frac{1}{48}, \\ \Lambda(\hat{\tau}^5) &= \Lambda(\{\bullet, \emptyset, \emptyset, \emptyset\}) = \sum_j b_{3j}(\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)). \end{aligned}$$

The right hand side of (39) becomes

$$\frac{1}{\gamma(\tau)} - \frac{1}{18} - \left(-\frac{1}{48}\right) = \frac{1}{30} - \frac{1}{18} + \frac{1}{48} = -\frac{1}{720},$$

and we have the order condition (35) for the linear combination  $\mathfrak{V} - \mathfrak{Y}$ .

If there are leaves on  $r > 1$  different levels above level two, things get slightly more complicated. Then we get  $r$  different terms on the left hand side of (39) and we need to consider the order condition for  $\tau$  and the  $r$  trees it forms energy-preserving linear combinations with, so that we get an equation for every energy-preserving combination of these trees, also those not including  $\tau$ . This is illustrated by the following example.

**Example 3** The tree  $\mathfrak{V}$  forms energy-preserving combinations with both  $\mathfrak{Y}$  and  $\mathfrak{W}$ . Thus we have to calculate (39) for all three trees to find order conditions for the corresponding linear combinations. Starting with  $\tau = \mathfrak{V}$ , which has three nodes above level two, two leaves and one non-leaf, we get

$$\begin{aligned} \Lambda(\hat{\tau}^4) &= \Lambda(\{\bullet, \mathfrak{I}, \emptyset\}) = \sum_j b_{2j}(\phi_{2j1}(\bullet)\phi_{2j2}(\mathfrak{I}) + \phi_{2j2}(\bullet)\phi_{2j1}(\mathfrak{I})), \\ \Lambda(\hat{\tau}^5) &= \frac{1}{(1+1)\gamma(\bullet)} \Lambda(\{\bullet, \bullet, \emptyset\}) = \frac{1}{2\gamma(\bullet)} \frac{1}{2} \left( \frac{1}{\gamma(\mathfrak{V})} - \frac{1}{3\gamma(\bullet)\gamma(\mathfrak{W})} \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \frac{1}{2} \left( \frac{1}{15} - \frac{1}{9} \right) = -\frac{1}{90}, \\
\Lambda(\hat{\tau}^6) &= \Lambda(\{\bullet, \bullet, \emptyset, \emptyset\}) = \sum_j b_{3j} (\phi_{3j1}(\bullet) \phi_{3j2}(\bullet) - \phi_{3j3}(\bullet) \phi_{3j2}(\bullet)) \\
&= \sum_j b_{3j} \phi_{3j2}(\bullet) (\phi_{3j1} - \phi_{3j3})(\bullet).
\end{aligned}$$

For the right hand side of (39), we get

$$\frac{1}{\gamma(\tau)} - \frac{1}{(2+1)\gamma(\bullet)\gamma(\mathfrak{V})} - \left(-\frac{1}{90}\right) = \frac{1}{48} - \frac{1}{3 \cdot 1 \cdot 8} + \frac{1}{90} = -\frac{7}{720},$$

and hence the order condition for  $\mathfrak{V}$  is

$$\sum_j b_{2j} (\phi_{2j1}(\bullet) \phi_{2j2}(\mathfrak{I}) + \phi_{2j2}(\bullet) \phi_{2j1}(\mathfrak{I})) + \sum_j b_{3j} \phi_{3j2}(\bullet) (\phi_{3j1} - \phi_{3j3})(\bullet) = -\frac{7}{720}. \quad (40)$$

Similarly we calculate (39) for  $\mathfrak{Y}$ ,

$$\sum_{jk} b_{2j} \phi_{2jk}(\mathfrak{V}) - 2 \sum_j b_{3j} \phi_{3j2}(\bullet) (\phi_{3j1} - \phi_{3j3})(\bullet) = -\frac{1}{120}, \quad (41)$$

and for  $\mathfrak{W}$ ,

$$\sum_{jk} b_{2j} \phi_{2jk}(\mathfrak{V}) + 2 \sum_j b_{2j} (\phi_{2j1}(\bullet) \phi_{2j2}(\mathfrak{I}) + \phi_{2j2}(\bullet) \phi_{2j1}(\mathfrak{I})) = -\frac{1}{36}. \quad (42)$$

Combining (40), (41) and (42), we get the equivalent system of equations

$$\sum_j b_{3j} \phi_{3j2}(\bullet) (\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = \frac{1}{240} + \alpha, \quad (43)$$

$$\sum_j b_{2j} (\phi_{2j1}(\bullet) \phi_{2j2}(\mathfrak{I}) + \phi_{2j1}(\mathfrak{I}) \phi_{2j2}(\bullet)) = -\frac{1}{72} - \alpha, \quad (44)$$

$$\sum_{j,k} b_{2j} \phi_{2jk}(\mathfrak{V}) = 2\alpha, \quad (45)$$

where the choice of  $\alpha \in \mathbb{R}$  is arbitrary. The order conditions (43)–(45) can be associated

to the linear combinations  $\mathfrak{V} - \mathfrak{Y}$ ,  $\mathfrak{V} + \mathfrak{W}$  and  $\mathfrak{Y} + \mathfrak{W}$ , respectively.

**Table 2** Energy-preserving linear combinations of elementary differentials, and their associated order conditions for the scheme (31), up to sixth order. The coefficients  $\alpha_1, \alpha_2 \in \mathbb{R}$  are arbitrary

$ \tau $	$\omega$	Order condition
1	$\bullet$	-
2	-	-
3	$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array}$	$\sum_j b_{2j} = -\frac{1}{24}$
4	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet) = -\frac{1}{24}$
5	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet)^2 = -\frac{1}{40}$
	$\begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{2j} \phi_{2j1}(\bullet) \phi_{2j2}(\bullet) = -\frac{1}{90}$
	$\begin{array}{c} \vee \\ \vee \end{array} - \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{3j} (\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = -\frac{1}{720}$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\mathfrak{I}) = -\frac{1}{60}$
	$\begin{array}{c} \vdots \\ \vdots \\ \vdots \end{array}$	$\sum_j b_{4j} = \frac{1}{240}$
6	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet)^3 = -\frac{1}{60}$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{2j} (\phi_{2j1}(\bullet)^2 \phi_{2j2}(\bullet) + \phi_{2j1}(\bullet) \phi_{2j2}(\bullet)^2) = -\frac{1}{2}$
	$\begin{array}{c} \vee \\ \vee \end{array} - \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{3j} (\phi_{3j1}(\bullet)^2 - \phi_{3j3}(\bullet)^2) = -\frac{1}{720}$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\bullet) \phi_{2jk}(\mathfrak{I}) = -\frac{1}{96}$
	$\begin{array}{c} \vee \\ \vee \end{array} - \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{3j} \phi_{3j2}(\bullet) (\phi_{3j1}(\bullet) - \phi_{3j3}(\bullet)) = \frac{1}{240} + \alpha_1$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{4j} (\phi_{4j1}(\bullet) + \phi_{4j4}(\bullet)) = \frac{1}{240}$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{2j} (\phi_{2j1}(\bullet) \phi_{2j2}(\mathfrak{I}) + \phi_{2j1}(\mathfrak{I}) \phi_{2j2}(\bullet)) = -\frac{1}{72} - \alpha_1$
	$\begin{array}{c} \vee \\ \vee \end{array} - \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{3j} (\phi_{3j1}(\mathfrak{I}) - \phi_{3j3}(\mathfrak{I})) = -\frac{1}{180} - \alpha_2$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\mathfrak{V}) = 2\alpha_1$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_{j,k} b_{2j} b_{2k} \phi_{2jk}(\mathfrak{I}) = \alpha_2$
	$\begin{array}{c} \vee \\ \vee \end{array} + \begin{array}{c} \vee \\ \vee \end{array}$	$\sum_j b_{4j} (\phi_{4j2}(\bullet) + \phi_{4j3}(\bullet)) = -\frac{1}{1440} - \alpha_2$

By considering the order conditions in Table 2, we find a fifth order scheme of the form (31) given by

$$\begin{aligned} \frac{\hat{x} - x}{h} = & \left( I - \frac{5}{136} h^2 (f'(z_2) f'(z_3) + f'(z_3) f'(z_2)) - \frac{1}{102} h^2 f'(x) f'(x) \right. \\ & + \frac{1}{288} h^3 (f'(x) f'(x) f'(z_1) + f'(z_1) f'(x) f'(x)) \\ & \left. + \frac{1}{120} h^4 f'(x) f'(x) f'(x) f'(x) \right) \int_0^1 f((1 - \xi)x + \xi \hat{x}) d\xi, \quad (46) \end{aligned}$$

where

$$z_1 = x + \frac{2}{5} h f(x), \quad z_2 = x + \frac{17 + \sqrt{17}}{30} h f(z_1), \quad z_3 = x + \frac{17 - \sqrt{17}}{30} h f(z_1).$$

A symmetric sixth order scheme is given by

$$\begin{aligned} \frac{\hat{x} - x}{h} = & \left( I - \frac{13}{360} h^2 f'(\bar{x} + \frac{\sqrt{13}}{26} h f(\bar{x} - \frac{3\sqrt{13}}{26} h f(\bar{x}))) f'(\bar{x} - \frac{\sqrt{13}}{26} h f(\bar{x} + \frac{3\sqrt{13}}{26} h f(\bar{x}))) \right. \\ & - \frac{13}{360} h^2 f'(\bar{x} - \frac{\sqrt{13}}{26} h f(\bar{x} + \frac{3\sqrt{13}}{26} h f(\bar{x}))) f'(\bar{x} + \frac{\sqrt{13}}{26} h f(\bar{x} - \frac{3\sqrt{13}}{26} h f(\bar{x}))) \\ & - \frac{1}{180} h^2 f'(x) f'(x) - \frac{1}{180} h^2 f'(\hat{x}) f'(\hat{x}) \\ & + \frac{1}{720} h^3 f'(\bar{x} - \frac{1}{2} h f(\bar{x})) f'(\bar{x}) f'(\bar{x} + \frac{1}{2} h f(\bar{x})) \\ & - \frac{1}{720} h^3 f'(\bar{x} + \frac{1}{2} h f(\bar{x})) f'(\bar{x}) f'(\bar{x} - \frac{1}{2} h f(\bar{x})) \\ & \left. + \frac{1}{120} h^4 f'(\bar{x}) f'(\bar{x}) f'(\bar{x}) f'(\bar{x}) \right) \int_0^1 f((1 - \xi)x + \xi \hat{x}) d\xi, \quad (47) \end{aligned}$$

where  $\bar{x} = \frac{x + \hat{x}}{2}$ . If we wish to calculate the matrix in front of the integral explicitly, we have a non-symmetric sixth order scheme given by

$$\begin{aligned} \frac{\hat{x} - x}{h} = & \left( I - \frac{13}{360} h^2 (f'(z_6) f'(z_7) + f'(z_7) f'(z_6)) \right. \\ & - \frac{1}{180} h^2 (f'(x) f'(x) + f'(z_1) f'(z_1)) \\ & + \frac{1}{720} h^3 (f'(x) f'(z_2) f'(z_3) - f'(z_3) f'(z_2) f'(x)) \\ & \left. + \frac{1}{120} h^4 f'(z_2) f'(z_2) f'(z_2) f'(z_2) \right) \int_0^1 f((1 - \xi)x + \xi \hat{x}) d\xi, \quad (48) \end{aligned}$$

with

$$\begin{aligned}
 z_1 &= x + \frac{1}{4}hf(x) + \frac{3}{4}hf\left(x + \frac{2}{3}hf\left(x + \frac{1}{3}hf(x)\right)\right), & z_2 &= x + \frac{1}{2}hf(x), \\
 z_3 &= x + hf(z_2), & z_4 &= \frac{1}{2}(x + z_3) - \frac{3\sqrt{13}}{26}hf(z_2), \\
 z_5 &= \frac{1}{2}(x + z_3) + \frac{3\sqrt{13}}{26}hf(z_2), \\
 z_6 &= \frac{1}{2}(x + z_1) + \frac{\sqrt{13}}{26}hf(z_4), & z_7 &= \frac{1}{2}(x + z_1) - \frac{\sqrt{13}}{26}hf(z_5).
 \end{aligned}$$

### 4 AVF discrete gradient methods for general skew-gradient systems

We will now build on the results of the previous section by generalizing the results to the situation where  $S(x)$  in the skew-gradient system (4) is not necessarily constant. Consider therefore now an ODE of the form (4), and set again  $g := \nabla H$ . By Taylor expansion of  $x$  around  $t = t_0$  we get

$$\begin{aligned}
 x(t_0 + h) &= x + hSg + \frac{h^2}{2}(S'gSg + Sg'Sg) + \frac{h^3}{6}(S''g(Sg, Sg) + 2S'g'(Sg, Sg) \\
 &\quad + Sg''(Sg, Sg) + S'gS'gSg + S'gSg'Sg \\
 &\quad + Sg'S'gSg + Sg'Sg'Sg) + \mathcal{O}(h^4),
 \end{aligned}$$

where  $x := x(t_0)$ , and  $S, g$  and their derivatives are evaluated in  $x$ . Introducing the notation  $f^\circ := S'g$  and  $f^\bullet := Sg'$ , we can write this in the abbreviated form

$$\begin{aligned}
 x(t_0 + h) &= x + hf + \frac{h^2}{2}(f^\circ f + f^\bullet f) + \frac{h^3}{6}(f^{\circ\circ}(f, f) + 2f^{\circ\bullet}(f, f) + f^{\bullet\bullet}(f, f) \\
 &\quad + f^\circ f^\circ f + f^\circ f^\bullet f + f^\bullet f^\circ f + f^\bullet f^\bullet f) + \mathcal{O}(h^4).
 \end{aligned} \tag{49}$$

#### 4.1 Skew-gradient systems and P-series

A *P-series* is given by

$$P(\phi, (x, y)) = \left( \phi(\emptyset)x + \sum_{\tau \in TP_\bullet} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau)F(\tau)(x, y), \right. \tag{50}$$

$$\left. \phi(\emptyset)y + \sum_{\tau \in TP_\circ} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau)F(\tau)(x, y) \right),$$

where  $TP$  is the set of rooted bi-colored trees and  $TP_\bullet$  and  $TP_\circ$  are the subsets of  $TP$  whose roots are black and white, respectively [21,Section III.2]. The bi-colored trees are built recursively; starting with  $\bullet$  and  $\circ$ , we let  $\tau = [\tau_1, \dots, \tau_m]_\bullet$  be the tree you get by grafting the roots of  $\tau_1, \dots, \tau_m$  to a black root and  $\tau = [\tau_1, \dots, \tau_m]_\circ$  the tree you get by grafting  $\tau_1, \dots, \tau_m$  to a white root. No subscript, i.e.  $\tau = [\tau_1, \dots, \tau_m]$ , means grafting to a black root.

The exact solution of a partitioned system

$$\begin{aligned}\dot{x} &= f(x, y), & x(t_0) &= x_0, \\ \dot{y} &= g(x, y), & y(t_0) &= y_0,\end{aligned}\tag{51}$$

can be written as  $(x(t_0 + h), y(t_0 + h)) = P(1/\gamma, (x_0, y_0))$ , where the coefficient  $\gamma$  is given by  $\gamma(\emptyset) = \gamma(\bullet) = \gamma(\circ) = 1$  and (30). As noted in [10], setting  $f(x, y) := S(y)\nabla H(x)$ , the skew-gradient system (4) can be written as (51) with  $g = f$ . When  $g = f$ , all coefficients and the elementary differentials  $F(\tau)$  in (50) are given independently of the color of the root. Thus for the system (4), it suffices to consider

$$P(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in TP_\bullet} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau) F(\tau)(x),\tag{52}$$

and we have that the exact solution of (4) can be written as  $x(t_0 + h) = P(1/\gamma, x_0)$ . Breaking slightly with convention, we define a P-series to be the single row version (52) in the remainder of this paper. Denoting black-rooted subtrees by  $\tau_i$  and white-rooted subtrees by  $\bar{\tau}_i$ , the elementary differentials  $F(\tau)$  for the skew-gradient system are given recursively by  $F(\bullet)(x) = F(\circ)(x) = S(x)\nabla H(x)$ , and

$$F(\tau)(x) = S^{(l)} D^m \nabla H(F(\tau_1)(x), \dots, F(\tau_m)(x), F(\bar{\tau}_1)(x), \dots, F(\bar{\tau}_l)(x))\tag{53}$$

for both  $\tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]_\bullet$  and  $\tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]_\circ$ . The bi-colored trees in  $TP_\bullet$  and their corresponding elementary differentials  $F$  are given up to order three in Table 3. The number of trees grows very quickly with the order; see <https://oeis.org/A000151>.

The following lemma is Lemma III.2.2 in [21] amended to fit our setting.

**Lemma 7** *Let  $P(a, x)$  and  $P(b, x)$  be two P-series with  $a(\emptyset) = b(\emptyset) = 1$ . Then*

$$hS(P(a, x))\nabla H(P(b, x)) = P(a \vee b, x),$$

where  $(a \vee b)(\emptyset) = 0$ ,  $(a \vee b)(\bullet) = 1$ , and

$$(a \vee b)(\tau) = a(\tau_1) \cdots a(\tau_m) b(\bar{\tau}_1) \cdots b(\bar{\tau}_l) \quad \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]_\bullet.$$

**Proposition 4** *The AVF discrete gradient scheme*

$$\frac{\hat{x} - x}{h} = S\left(\frac{x + \hat{x}}{2}\right) \int_0^1 \nabla H((1 - \xi)x + \xi\hat{x}) d\xi\tag{54}$$

is a second order P-series method.

**Table 3** Bi-colored trees and their elementary differentials up to third order

$ \tau $	$F(\tau)^i$	$F(\tau)$	$\tau$	$\alpha(\tau)$	$\gamma(\tau)$	$\sigma(\tau)$
1	$S_j^i g^j$	$f$	$\bullet$	1	1	1
2	$S_{jk}^i g^j S_l^k g^l$	$f^\circ f$	$\begin{array}{c} \circ \\   \\ \bullet \end{array}$	1	2	1
	$S_j^i g_k^j S_l^k g^l$	$f^\bullet f$	$\begin{array}{c} \bullet \\   \\ \circ \end{array}$	1	2	1
3	$S_{jkm}^i g^j S_l^k g^l S_n^m g^n$	$f^{\circ\circ}(f, f)$	$\begin{array}{c} \circ \\ / \quad \backslash \\ \bullet \quad \bullet \end{array}$	1	3	2
	$S_{jk}^i g_m^j S_l^k g^l S_n^m g^n$	$f^{\circ\bullet}(f, f)$	$\begin{array}{c} \circ \\ / \quad \backslash \\ \bullet \quad \circ \end{array}$	2	3	1
	$S_j^i g_{km}^j S_l^k g^l S_n^m g^n$	$f^{\bullet\bullet}(f, f)$	$\begin{array}{c} \bullet \\ / \quad \backslash \\ \circ \quad \circ \end{array}$	1	3	2
	$S_j^i g^j S_{lm}^k S_n^m g^n$	$f^\circ f^\circ f$	$\begin{array}{c} \circ \\   \\ \circ \\   \\ \bullet \end{array}$	1	6	1
	$S_j^i g_k^j S_{lm}^k S_n^m g^n$	$f^\bullet f^\circ f$	$\begin{array}{c} \bullet \\   \\ \circ \\   \\ \bullet \end{array}$	1	6	1
	$S_{jk}^i g^j S_l^k S_m^l S_n^m g^n$	$f^\circ f^\bullet f$	$\begin{array}{c} \circ \\   \\ \bullet \\   \\ \bullet \end{array}$	1	6	1
	$S_j^i g_k^j S_l^k S_m^l S_n^m g^n$	$f^\bullet f^\bullet f$	$\begin{array}{c} \bullet \\   \\ \bullet \\   \\ \bullet \end{array}$	1	6	1

**Proof** As in the proof of Proposition 2, we define  $\hat{e}$  by  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ . Now, assume that the solution  $\hat{x}$  of (54) can be written as the P-series  $\hat{x} = P(\Phi, x)$ . Then, using Lemma 7, we find the P-series

$$\begin{aligned} & h S\left(\frac{x + \hat{x}}{2}\right) \int_0^1 \nabla H((1 - \xi)x + \xi \hat{x}) \, d\xi \\ &= h S\left(P\left(\frac{1}{2}\hat{e} + \frac{1}{2}\Phi, x\right)\right) \int_0^1 \nabla H(P((1 - \xi)\hat{e} + \xi\Phi, x)) \, d\xi \\ &= \int_0^1 h S\left(P\left(\frac{1}{2}\hat{e} + \frac{1}{2}\Phi, x\right)\right) \nabla H(P((1 - \xi)\hat{e} + \xi\Phi, x)) \, d\xi \\ &= P\left(\int_0^1 \left(\left(\frac{1}{2}\hat{e} + \frac{1}{2}\Phi\right) \vee ((1 - \xi)\hat{e} + \xi\Phi)\right) \, d\xi, x\right). \end{aligned}$$

Thus we get  $\Phi = \hat{e} + \int_0^1 \left(\left(\frac{1}{2}\hat{e} + \frac{1}{2}\Phi\right) \vee ((1 - \xi)\hat{e} + \xi\Phi)\right) \, d\xi = \hat{e} + \int_0^1 \left(\left(\frac{1}{2}\Phi\right) \vee (\xi\Phi)\right) \, d\xi$ . That is,  $\Phi(\emptyset) = 1$ ,  $\Phi(\bullet) = 1$ , and

$$\Phi([\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]) = \frac{1}{(m + 1)2^l} \Phi(\tau_1) \cdots \Phi(\tau_m) \Phi(\bar{\tau}_1) \cdots \Phi(\bar{\tau}_l).$$

Writing out the first few terms of the series, we have

$$\begin{aligned} \hat{x} &= x + hf + \frac{h^2}{2}(f^\circ f + f^\bullet f) + h^3\left(\frac{1}{8}f^{\circ\circ}(f, f) + \frac{1}{4}f^{\circ\bullet}(f, f) + \frac{1}{6}f^{\bullet\bullet}(f, f)\right) \\ &\quad + \frac{1}{4}f^\circ f^\circ f + \frac{1}{4}f^\circ f^\bullet f + \frac{1}{4}f^\bullet f^\circ f + \frac{1}{4}f^\bullet f^\bullet f + \mathcal{O}(h^4), \end{aligned}$$

which, after comparing with the expanded exact solution (49), we see is of order two.

The following lemma is obtained in a manner similar to Lemma 5, i.e. Theorem 2.2 in [34], and hence we present it without its proof.

**Lemma 8** *Let  $P(a, x)$ ,  $P(b, x)$  and  $P(c, x)$  be three P-series with  $a(\emptyset) = b(\emptyset) = 1$  and  $c(\emptyset) = 0$ . Then*

$$h S(P(a, x))D^2 H(P(b, x))P(c, x) = P((a, b) \times c, x)$$

with  $((a, b) \times c)(\emptyset) = ((a, b) \times c)(\bullet) = 0$  and otherwise

$$((a, b) \times c)(\tau) = \sum_{i=1}^m \prod_{j=1, j \neq i}^m \prod_{k=1}^l a(\bar{\tau}_k)b(\tau_j)c(\tau_i) \quad \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]. \tag{55}$$

Note that  $\{\emptyset\}$  counts as both a black-rooted and a white-rooted tree. Hence we have e.g.

$$((a, b) \times c)(\heartsuit) = a(\circ)b(\emptyset)c(\bullet) = a(\bullet)c(\bullet),$$

where we also use that  $a(\circ) = a(\bullet)$ .

We now present a subclass of the AVF discrete gradient method, for which we will find order conditions using Lemma 7 and Lemma 8. This subclass is every AVF discrete gradient method for which the approximation of  $S(x)$  can be written in the form

$$\begin{aligned} \bar{S}(x, \hat{x}, h) = & \sum_{n=0}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^n S(\bar{z}_{nj k}) D^2 H(z_{nj k}) \cdot S(\bar{z}_{nj(n+1)}) \right. \\ & \left. + (-1)^n S(\bar{z}_{nj(n+1)}) \prod_{k=1}^n D^2 H(z_{nj(n-k+1)}) S(\bar{z}_{nj(n-k+1)}) \right), \end{aligned} \tag{56}$$

where, if  $\hat{x}$  is the solution of

$$\frac{\hat{x} - x}{h} = \bar{S}(x, \hat{x}, h) \bar{\nabla}_{\text{AVF}} H(x, \hat{x}),$$

each  $z_{nj k} := z_{nj k}(x, \hat{x}, h) = P(\phi_{nj k}, x)$  and each  $\bar{z}_{nj k} := \bar{z}_{nj k}(x, \hat{x}, h) = P(\psi_{nj k}, x)$  can be written as a P-series with  $\phi_{nj k}(\emptyset) = \psi_{nj k}(\emptyset) = 1$  for all  $n, j, k$ . We require that  $\sum_j b_{0j} = \frac{1}{2}$ , which ensures that (56) is a consistent approximation of  $S(x)$ .

**Theorem 3** *The discrete gradient scheme (9) with the AVF discrete gradient (10) and the approximation of  $S(x)$  given by (56) is a P-series method.*

**Proof** Generalizing the argument in the proof of Proposition 4, we find the P-series

$$h S(P(a, x)) \int_0^1 \nabla H((1 - \xi)x + \xi \hat{x}) d\xi = P\left(\int_0^1 (a \vee ((1 - \xi)\hat{e} + \xi\Phi)) d\xi, x\right),$$



where  $\bar{\theta}(a) := \int_0^1 (a \vee ((1 - \xi)\hat{e} + \xi\Phi)) \, d\xi = \int_0^1 (a \vee \xi\Phi) \, d\xi$ , so that  $\bar{\theta}(a)(\emptyset) = 0$ ,  $\bar{\theta}(a)(\bullet) = 1$ , and

$$\bar{\theta}(a)([\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]) = \frac{1}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m) a(\bar{\tau}_1) \cdots a(\bar{\tau}_l). \tag{57}$$

Thus we may write the solution  $\hat{x}$  found from applying the scheme (9) with the AVF discrete gradient (10) and  $\bar{S}(x, \hat{x}, h)$  given by (56) as

$$\begin{aligned} \hat{x} &= x + \sum_{n=0}^{p-1} h^n \sum_j b_{nj} \left( \prod_{k=1}^n S(P(\psi_{nj k}, x)) D^2 H(P(\phi_{nj k}, x)) \cdot P(\bar{\theta}(\psi_{nj(n+1)}, x)) \right. \\ &\quad \left. + (-1)^n \prod_{k=1}^n S(P(\psi_{nj(n-k+2)}, x)) D^2 H(P(\phi_{nj(n-k+1)}, x)) \cdot P(\bar{\theta}(\psi_{nj1}, x)) \right) \\ &= x + \sum_{n=0}^{p-1} \sum_j b_{nj} (P((\psi_{nj1}, \phi_{nj1}) \times (\psi_{nj2}, \phi_{nj2}) \times \cdots \times (\psi_{njn}, \phi_{njn}) \times \bar{\theta}(\psi_{nj(n+1)}, x)) \\ &\quad + (-1)^n P((\psi_{nj(n+1)}, \phi_{njn}) \times (\psi_{njn}, \phi_{nj(n-1)}) \times \cdots \times (\psi_{nj2}, \phi_{nj1}) \times \bar{\theta}(\psi_{nj1}, x))) \\ &= P(\Phi, x), \end{aligned} \tag{58}$$

with

$$\begin{aligned} \Phi &= \hat{e} + \sum_{n=0}^{p-1} \sum_j b_{nj} ((\psi_{nj1}, \phi_{nj1}) \times \cdots \times (\psi_{njn}, \phi_{njn}) \times \bar{\theta}(\psi_{nj(n+1)}) \\ &\quad + (-1)^n (\psi_{nj(n+1)}, \phi_{njn}) \times \cdots \times (\psi_{nj2}, \phi_{nj1}) \times \bar{\theta}(\psi_{nj1})). \end{aligned} \tag{59}$$

**Theorem 4** *The AVF discrete gradient method with  $\bar{S}$  given by (56) is of order  $p$  if and only if*

$$\Phi(\tau) = \frac{1}{\gamma(\tau)} \text{ for } |\tau| \leq p. \tag{60}$$

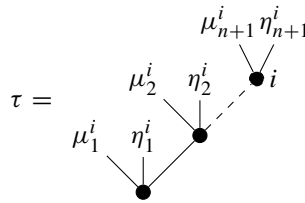
The values  $\Phi(\tau)$  can be found from (59) using (55) recursively and then (57). However, a more convenient approach is derived in the next subsection.

### 4.2 Order conditions

This subsection is devoted to generalizations of the results in Sect. 3.1 to the cases where  $S(x)$  is not necessarily constant. To that end, for a tree  $\tau \in TP_\bullet$ , we cut off all branches between black and white nodes and denote the mono-colored tree we are left with by  $\tau^b$ . We number the nodes in that tree as before, from 1 to  $|\tau^b|$ , and reattach the cut-off parts to the tree to get  $\tau$  again. Let  $\mu$  denote a forest of black-rooted trees and  $\eta$  a forest of white-rooted trees. Then, for a given node  $i \in [1, \dots, |\tau^b|]$  on level  $n + 1$ , there exists a unique set of forests  $\hat{\tau}^i = \{(\mu_1^i, \eta_1^i), \dots, (\mu_{n+1}^i, \eta_{n+1}^i)\}$  such that

$$\tau = [(\mu_1^i, \eta_1^i)] \circ [(\mu_2^i, \eta_2^i)] \circ \cdots \circ [(\mu_{n+1}^i, \eta_{n+1}^i)].$$

That is,



Now we can generalize Proposition 3 as follows.

**Proposition 5** *The  $\Phi$  of (59) can be found by*

$$\Phi(\tau) = \hat{e}(\tau) + \sum_{i=1}^{|\tau^b|} \Lambda(\hat{\tau}^i) \tag{61}$$

where  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ , and

$$\begin{aligned} \Lambda(\hat{\tau}^i) = & \theta([\mu_{n+1}^i]) \sum_j b_{nj} \left( \psi_{nj1}(\eta_1^i) \phi_{nj1}(\mu_1^i) \cdots \psi_{njn}(\eta_n^i) \phi_{njn}(\mu_n^i) \psi_{nj(n+1)}(\eta_{n+1}^i) \right. \\ & \left. + (-1)^n \psi_{nj(n+1)}(\eta_1^i) \phi_{njn}(\mu_1^i) \psi_{njn}(\eta_2^i) \cdots \phi_{nj1}(\mu_n^i) \psi_{nj1}(\eta_{n+1}^i) \right), \end{aligned} \tag{62}$$

with

$$\theta([\tau_1, \dots, \tau_m]) = \frac{1}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m).$$

**Proof** Defining  $n_i$  and  $C_i$  as in the proof of Proposition 3, we have

$$\begin{aligned} [(\mu_{n_i+1}^i, \eta_{n_i+1}^i)] &= [(\mu_{n_k}^k, \eta_{n_k}^k)] \circ [(\mu_{n_k+1}^k, \eta_{n_k+1}^k)] \quad \text{for all } k \in C_i, \\ ((a, b) \times c)[(\mu_{n_i+1}^i, \eta_{n_i+1}^i)] &= \sum_{k \in C_i} a(\eta_{n_k}^k) b(\mu_{n_k}^k) c([\mu_{n_k+1}^k, \eta_{n_k+1}^k]). \end{aligned}$$

Observe that  $\bar{\theta}(a)([\mu, \eta]) = a(\eta)\theta([\mu])$ . For  $n = 0$  we have

$$\bar{\theta}(\psi_{0j1})(\tau) = \bar{\theta}(\psi_{0j1})([\mu_1^1, \eta_1^1]) = \psi_{0j1}(\eta_1^1)\theta([\mu_1^1]),$$

and for  $n > 0$  we get

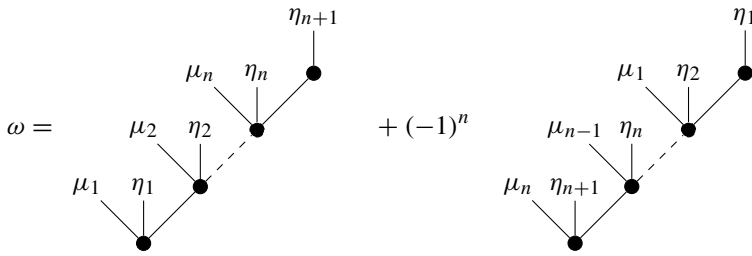
$$\begin{aligned} & ((\psi_{nj1}, \phi_{nj1}) \times \cdots \times (\psi_{njn}, \phi_{njn}) \times \bar{\theta}(\psi_{nj(n+1)}))(\tau) \\ &= ((\psi_{nj1}, \phi_{nj1}) \times \cdots \times (\psi_{njn}, \phi_{njn}) \times \bar{\theta}(\psi_{nj(n+1)}))([\mu_1^1, \eta_1^1]) \\ &= \sum_{i_1 \in C_1} \psi_{nj1}(\eta_1^{i_1}) \phi_{nj1}(\mu_1^{i_1}) (\psi_{nj2}, \phi_{nj2}) \\ & \quad \times \cdots \times (\psi_{njn}, \phi_{njn}) \times \bar{\theta}(\psi_{nj(n+1)})([\mu_2^{i_1}, \eta_2^{i_1}]) \\ & \quad \vdots \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i_1 \in C_1} \cdots \sum_{i_n \in C_{i_{n-1}}} \psi_{nj_1}(\eta_1^{i_1}) \phi_{nj_1}(\mu_1^{i_1}) \cdots \psi_{nj_n}(\eta_n^{i_n}) \phi_{nj_n}(\mu_n^{i_n}) \bar{\theta}(\psi_{nj(n+1)})([\mu_{n+1}^{i_n}, \eta_{n+1}^{i_n}]) \\
 &= \sum_{i \text{ on level } n+1} \psi_{nj_1}(\eta_1^i) \phi_{nj_1}(\mu_1^i) \cdots \psi_{nj_n}(\eta_n^i) \phi_{nj_n}(\mu_n^i) \psi_{nj(n+1)}(\eta_{n+1}^i) \theta([\mu_{n+1}^i]).
 \end{aligned}$$

Inserting this and the corresponding result for  $((\psi_{nj(n+1)}, \phi_{nj_n}) \times \cdots \times (\psi_{nj_2}, \phi_{nj_1}) \times \bar{\theta}(\psi_{nj_1}))(\tau)$  in (59), we get (62).

Note that if  $\tau$  only has black nodes, we have  $\Lambda(\hat{\tau}^1) = \theta(\tau) \sum_j b_{0j}(\psi_{0j_1}(\emptyset) + \psi_{0j_1}(\emptyset)) = \theta(\tau)$ , and also  $\Lambda(\hat{\tau}^i) = 0$  for all nodes  $i$  on level 2. Thus (61) simplifies to (36).

Like for the constant  $S$  case, the order conditions can be given for energy-preserving linear combinations of elementary differentials instead for each elementary differential. In the following generalization of Lemma 6, we state that the energy-preserving linear combinations of bi-colored rooted trees are given by



**Theorem 5** Let  $\mu_1, \mu_2, \dots, \mu_n$  be arbitrary forests of black-rooted trees and  $\eta_1, \eta_2, \dots, \eta_{n+1}$  arbitrary forests of white-rooted trees. Given  $f(x) = S(x)\nabla H(x)$ , where  $S(x)$  is a skew-symmetric matrix, and elementary differentials defined by (53), the linear combinations of trees given by

$$\omega = [(\mu_1, \eta_1)] \circ \cdots \circ [(\mu_n, \eta_n)] \circ [\eta_{n+1}] + (-1)^n [(\mu_n, \eta_{n+1})] \circ \cdots \circ [(\mu_1, \eta_2)] \circ [\eta_1] \tag{63}$$

are energy-preserving in the sense that  $F(\omega)(x) \cdot \nabla H(x) = 0$ .

**Proof** For any forest of black-rooted trees  $\mu_j$ , we have  $F([\mu_j] \circ [\emptyset]) = SB_jS\nabla H$  for some symmetric matrix  $B_j$ , suppressing the argument  $x$ . Similarly, for a forest of white-rooted trees  $\eta_j$ , we have  $F([\eta_j]) = W_j\nabla H$  for some skew-symmetric matrix  $W_j$ . Note that the empty forest is considered both a black-rooted and a white-rooted forest, and accordingly we have  $F([\emptyset] \circ [\emptyset]) = F(\blacklozenge) = S(D^2H)S\nabla H$  and  $F([\emptyset]) = F(\bullet) = S\nabla H$ . For these matrices  $B_j$  and  $W_j$  corresponding to the forests  $\mu_j$  and  $\eta_j$ , we get

$$F([\mu_1, \eta_1] \circ \cdots \circ [(\mu_n, \eta_n)] \circ [\eta_{n+1}]) = W_1 B_1 W_2 B_2 \cdots B_n W_{n+1} \nabla H.$$

We have

$$(W_1 B_1 W_2 B_2 \cdots B_n W_{n+1})^T = \begin{cases} -W_{n+1} B_n W_n B_{n-1} \cdots B_1 W_1 & \text{if } n \text{ even,} \\ W_{n+1} B_n W_n B_{n-1} \cdots B_1 W_1 & \text{if } n \text{ odd.} \end{cases}$$

Thus  $F(\omega)(x)$  is a skew-symmetric matrix times  $\nabla H(x)$ , and the statement in the above theorem follows directly.

For bi-colored trees, we define a node on the tree  $\tau$  to be a leaf if it is a leaf on the corresponding cut tree  $\tau^b$  by the definition of leaves given in the previous section. We let  $I_l$  be the set of leaves and  $I_n$  the set of non-leaf nodes which are also in  $\tau^b$ , so that  $I_l \cup I_n = [1, \dots, |\tau^b|]$ . In contrast to the case with mono-colored trees, a leaf  $i$  on level one or two of a bi-colored tree may give rise to a non-zero energy-preserving linear combination; it does so if and only if  $\eta_k^i \neq \emptyset$  for any  $k = 1, 2$ . Accordingly,  $\Lambda(\hat{\tau}^i)$  is calculated in (61) also when  $n = 0, 1$ . Furthermore, two leaves  $i$  and  $j$  on the same level will belong to two different energy-preserving combinations if  $\eta_{n+1}^i \neq \eta_{n+1}^j$ . Therefore we now simply state that a tree with  $r$  leaves, also including the lower two levels, belong to at most  $r$  non-zero linear combinations. We thus get  $r$  terms on the left hand side of

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i) = \frac{1}{\gamma(\tau)} - \hat{e}(\tau) - \sum_{i \in I_n} \frac{\Lambda(\{(\mu_1^i, \eta_1^i), \dots, (\mu_n^i, \eta_n^i), (\emptyset, \eta_{n+1}^i)\})}{(|\mu_{n+1}^i| + 1)\gamma(\mu_{n+1}^i)}, \quad (64)$$

which is equivalent to (60) if we assume the conditions for lower order to be satisfied.

**Example 4** Consider the tree  $\tau = \mathfrak{V}^{\circ}$ , which is part of the energy-preserving linear combination  $\mathfrak{V}^{\circ} - \mathfrak{V}^{\circ}$ . Assume that the order conditions up to and including order three are all satisfied. The cut tree  $\tau^b = \mathfrak{V}$  has three nodes of which two are leaves. Node number 2 is a leaf on level 2 with  $\eta_1^2 = \eta_2^2 = \emptyset$ , and thus gives  $\Lambda(\hat{\tau}^2) = 0$ . We find for the other two,

$$\begin{aligned} \Lambda(\hat{\tau}^1) &= \Lambda(\{((\bullet, \circ), \emptyset)\}) = \frac{1}{(|\mu_1^1| + 1)\gamma(\mu_1^1)} \Lambda(\{(\emptyset, \emptyset)\}) = \frac{1}{(2 + 1)\gamma(\bullet)\gamma(\circ)} \frac{1}{\gamma(\bullet)} = \frac{1}{6}, \\ \Lambda(\hat{\tau}^3) &= \Lambda(\{(\bullet, \emptyset), (\emptyset, \circ)\}) = \sum_j b_{1j} \phi_{1j1}(\bullet)\psi_{1j2}(\circ) - \psi_{1j1}(\circ)\phi_{1j1}(\bullet) \\ &= \sum_j b_{1j} \phi_{1j1}(\bullet)(\psi_{1j2} - \psi_{1j1})(\circ). \end{aligned}$$

For the right hand side of (64) we get

$$\frac{1}{\gamma(\tau)} - \Lambda(\hat{\tau}^1) = \frac{1}{8} - \frac{1}{6} = -\frac{1}{24},$$

and thus the order condition


$$\sum_j b_{1j} \phi_{1j1}(\bullet)(\psi_{1j2} - \psi_{1j1})(\circ) = -\frac{1}{24}$$

for the energy-preserving linear combination  $\mathfrak{V}^{\circ} - \mathfrak{V}^{\circ}$ .

Even though the number of black-rooted bi-colored trees grows very quickly, e.g. to 26 for  $|\tau| = 4$  and 107 for  $|\tau| = 5$ , finding and satisfying the order conditions is not as daunting a task as it might first appear. First of all, it suffices to find order conditions for the non-zero linear combinations given by (63). Moreover, a couple key observations simplifies the process further:

- The large number of trees  $\tau$  for which  $\tau^b = \bullet$ , i.e. trees with no black nodes on level 2, are all energy-preserving. They can be written  $\tau = [\eta_1^1]$ , and their order condition is given by

$$2 \sum_j b_{0j} \psi_{0j1}(\eta_1^1) = \frac{1}{\gamma(\tau)}.$$

- For trees that are identical except for the colors of the descendants of white nodes, it suffices to calculate one order condition. E.g. for  we have the order condition  $2b_{0j} \psi_{0j1}(\mathcal{V}) = \frac{1}{12}$ , where each of the gray nodes may be black or white. To satisfy these conditions, it is natural to require that  $\bar{z}_{0j1}$  in (56) is a B-series up to order  $p - 1$ .

From the order conditions displayed in Table 4 we find that one second order scheme is given by (9) using the AVF discrete gradient (10) and an explicit skew-symmetric approximation of  $S$  given by  $\bar{S}(x, \cdot, h) = S(x + \frac{1}{2}hf(x))$ . A third order scheme is obtained if we instead use the skew-symmetric approximation of  $S$  explicitly given by

$$\begin{aligned} \bar{S}(x, \cdot, h) = & \frac{1}{4}S(x) + \frac{3}{4}S(z_2) + \frac{1}{4}h \left( S(z_1)D^2H(x)S(x) - S(x)D^2H(x)S(z_1) \right) \\ & - \frac{1}{12}h^2 S(x)D^2H(x)S(x)D^2H(x)S(x), \end{aligned} \tag{65}$$

where  $z_1 = x + \frac{1}{3}hf(x)$ ,  $z_2 = x + \frac{2}{3}hf(z_1)$ .

A symmetric fourth order scheme is given by (9) using the AVF discrete gradient (10) and the skew-symmetric approximation of  $S$

$$\begin{aligned} \bar{S}(x, \hat{x}, h) = & \frac{1}{2}S\left(\bar{x} - \frac{1}{\sqrt{12}}hf(\bar{x}) + \frac{1}{\sqrt{12}}hf(\bar{x})\right) \\ & + \frac{1}{2}S\left(\bar{x} + \frac{1}{\sqrt{12}}hf(\bar{x}) - \frac{1}{\sqrt{12}}hf(\bar{x})\right) \\ & + \frac{1}{2}h S\left(\bar{x} + \frac{1}{12}hf(\bar{x})\right)D^2H(\bar{x})S\left(\bar{x} - \frac{1}{12}hf(\bar{x})\right) \\ & - \frac{1}{2}h S\left(\bar{x} - \frac{1}{12}hf(\bar{x})\right)D^2H(\bar{x})S\left(\bar{x} + \frac{1}{12}hf(\bar{x})\right) \\ & - \frac{1}{12}h^2 S(\bar{x})D^2H(\bar{x})S(\bar{x})D^2H(\bar{x})S(\bar{x}), \end{aligned} \tag{66}$$

**Table 4** Linear combinations  $\omega$  of bi-colored black-rooted trees corresponding to energy-preserving elementary differentials of  $f(x) = S(x)\nabla H(x)$ , where  $S(x)$  is a skew-symmetric matrix, as well as their associated order conditions for the discrete gradient method (9) with the AVF discrete gradient (10) and  $\bar{S}(x, \hat{x}, h)$  given by (56)

$ \tau $	$\omega$	Order condition
1		$2 \sum_j b_{0j} = 1$
2		$2 \sum_j b_{0j} \psi_{0j1}(\bullet) = \frac{1}{2}$
3		$2 \sum_j b_{0j} \psi_{0j1}(\bullet)^2 = \frac{1}{3}$
4		$\sum_j b_{2j} = -\frac{1}{24}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\bullet}) = \frac{1}{6}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\bullet}}) = \frac{1}{6}$
		$\sum_j b_{1j} (\psi_{1j2} - \psi_{1j1})(\bullet) = -\frac{1}{12}$
		$2 \sum_j b_{0j} \psi_{0j1}(\bullet)^3 = \frac{1}{4}$
		$2 \sum_j b_{2j} \psi_{2j2}(\bullet) = -\frac{1}{24}$
		$2 \sum_j b_{0j} \psi_{0j1}(\mathbf{V}) = \frac{1}{12}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\mathbf{V}}) = \frac{1}{12}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\mathbf{V}}}) = \frac{1}{12}$
		$2 \sum_j b_{0j} \psi_{0j1}(\bullet) \psi_{0j1}(\hat{\bullet}) = \frac{1}{8}$
		$2 \sum_j b_{0j} \psi_{0j1}(\bullet) \psi_{0j1}(\hat{\hat{\bullet}}) = \frac{1}{8}$
		$\sum_j b_{2j} (\phi_{2j1} + \phi_{2j2})(\bullet) = -\frac{1}{24}$
		$\sum_j b_{1j} (\psi_{1j2}(\bullet)^2 - \psi_{1j1}(\bullet)^2) = -\frac{1}{12}$
		$\sum_j b_{1j} \phi_{1j1}(\bullet) (\psi_{1j1} - \psi_{1j0})(\bullet) = -\frac{1}{24}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\hat{\bullet}}}) = \frac{1}{24}$
		$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\hat{\hat{\bullet}}}}) = \frac{1}{24}$
	$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\hat{\hat{\hat{\bullet}}}}}) = \frac{1}{24}$	
	$2 \sum_j b_{0j} \psi_{0j1}(\hat{\hat{\hat{\hat{\hat{\hat{\bullet}}}}}}) = \frac{1}{24}$	
	$\sum_j b_{2j} (\psi_{2j1} + \psi_{2j3})(\bullet) = -\frac{1}{24}$	
	$\sum_j b_{1j} (\psi_{1j2} - \psi_{1j1})(\hat{\bullet}) = -\frac{1}{24}$	
	$\sum_j b_{1j} (\psi_{1j2} - \psi_{1j1})(\hat{\hat{\bullet}}) = -\frac{1}{24}$	

where  $\bar{x} = \frac{x+\hat{x}}{2}$ . Another fourth order scheme is obtained if we instead use the explicit skew-symmetric approximation of  $S$  found by

$$\begin{aligned} \bar{S}(x, \cdot, h) &= \frac{1}{2}(S(z_5 + z_6) + S(z_5 - z_6)) \\ &+ \frac{1}{12}h (S(z_2)D^2H(z_1)S(x) - S(x)D^2H(z_1)S(z_2)) \\ &- \frac{1}{12}h^2 S(z_1)D^2H(z_1)S(z_1)D^2H(z_1)S(z_1), \end{aligned} \tag{67}$$

where

$$\begin{aligned} z_1 &= x + \frac{1}{2}hf(x), \quad z_3 = x + hf(z_2), \quad z_5 = \frac{1}{3}(x + z_1 + z_2) + \frac{1}{12}(-z_3 + z_4), \\ z_2 &= x + hf(z_1), \quad z_4 = x + hf(z_3), \quad z_6 = \frac{\sqrt{3}}{36}(7x - 2z_1 - 4z_2 + z_3 - 2z_4). \end{aligned}$$

### 5 Order conditions for general discrete gradient methods

We will now generalize the results of the two previous sections to discrete gradient methods with a general discrete gradient, as defined by (7)–(8). To that end, we introduce two new series in the vein of B- and P-series, as well as related tree structures.

#### 5.1 The constant S case

Consider mono-colored rooted trees whose nodes can have two different shapes: the circle shape of the nodes in trees of B-series, but also a triangle shape. Let  $TG$  be the set of such trees whose leaves are always circles. That is, from the first tree  $\bullet$ , every tree  $\tau \in TG$  can be built recursively through

$$[\tau_1, \dots, \tau_m]_{\bullet}, \quad [\tau_1, \dots, \tau_m]_{\blacktriangle}, \quad \tau_1, \dots, \tau_m \in TG,$$

which denotes the grafting of the trees  $\tau_1, \dots, \tau_m$  to a root  $\bullet$  or  $\blacktriangle$ , respectively. The elementary differentials  $F(\tau)$  corresponding to a tree  $\tau \in TG$  are likewise defined recursively by  $F(\bullet)(x) = f(x) = S\nabla H(x)$  and

$$F(\tau)(x) = \begin{cases} SD^m \nabla H(x)(F(\tau_1)(x), \dots, F(\tau_m)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m]_{\bullet}, \\ SD_2^{m-1} Q(x, x)(F(\tau_1)(x), \dots, F(\tau_m)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m]_{\blacktriangle}. \end{cases}$$

We can then define a generalization of B-series which includes these elementary differentials.

**Definition 4** A G-series is a formal series of the form

$$G(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in TG} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau)F(\tau)(x), \tag{68}$$





To every stem  $s \in SG$  of height  $n + 1 = |s|$ , we associate coefficients  $b_{sj}$  and  $\phi_{sjk}$ . Letting  $s_k$  be the  $k^{\text{th}}$  node of  $s$ , we define the function

$$R(\phi_{sjk}, x) := \begin{cases} D^2 H(G(\phi_{sjk}, x)) & \text{if } s_k = \bullet, \\ Q(x, G(\phi_{sjk}, x)) & \text{if } s_k = \blacktriangle. \end{cases}$$

Then we have  $hSR(\phi_{sjk}, x)G(b, x) = G(\phi_{sjk} \diamond b)$ , with  $(\phi_{sjk} \diamond b)(\emptyset) = (\phi_{sjk} \diamond b)(\bullet) = 0$  and

$$(\phi_{sjk} \diamond b)(\tau) = \begin{cases} \sum_{i=1}^m \prod_{j=1, j \neq i}^m \phi_{sjk}(\tau_j) b(\tau_i) & \text{for } \tau = [\tau_1, \dots, \tau_m]_{s_k}, \\ 0 & \text{if root of } \tau \neq s_k. \end{cases}$$

Consider now the class of skew-symmetric and consistent approximations to  $S$  that can be written in the form

$$\bar{S}(x, \hat{x}, h) = \sum_{s \in SG} h^n \sum_j b_{sj} \left( \prod_{k=1}^n SR(\phi_{sjk}, x) + (-1)^{|s| \bullet - 1} \prod_{k=1}^n SR(\phi_{sj(n-k+1)}, x) \right) S \tag{70}$$

whenever  $\hat{x}$  is the solution of

$$\frac{\hat{x} - x}{h} = \bar{S}(x, \hat{x}, h) \bar{\nabla} H(x, \hat{x}),$$

with  $\phi_{sjk}(\emptyset) = 1$  for every  $s, j, k$ , and with  $\sum_j b_{\bullet j} = \frac{1}{2}$ .

**Lemma 10** *The discrete gradient method (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70) and  $\bar{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  is a G-series method when applied to a constant  $S$  skew-gradient system (6).*

**Proof** Assume that the solution  $\hat{x}$  of (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70) can be written as the G-series  $\hat{x} = G(\Phi, x)$ . Then, using Lemma 2 and  $\bar{\nabla} H(x, x) = \nabla H(x)$ ,

$$\begin{aligned} hS \bar{\nabla} H(x, \hat{x}) &= hS \sum_{m=0}^\infty \frac{1}{m!} D_2^m \bar{\nabla} H(x, x) (G(\Phi, x) - x)^m \\ &= hS \sum_{m=0}^\infty \frac{1}{(m+1)!} D^m \nabla H(x) (G(\Phi, x) - x)^m \\ &\quad - hS \sum_{m=1}^\infty \frac{2m}{(m+1)!} D_2^{m-1} Q(x, x) (G(\Phi, x) - x)^m. \end{aligned}$$

Arguing as in the proof of Lemma III.1.9 in [21], we get  $h S \bar{\nabla} H(x, \hat{x}) = G(\theta, x)$ , with  $\theta(\emptyset) = 0, \theta(\bullet) = 1$ , and

$$\begin{aligned} \theta([\tau_1, \dots, \tau_m]_{\bullet}) &= \frac{1}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m), \\ \theta([\tau_1, \dots, \tau_m]_{\blacktriangle}) &= \frac{-2m}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m). \end{aligned} \tag{71}$$

Then we can write (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70) as

$$\begin{aligned} \hat{x} &= x + \sum_{s \in SG} h^n \sum_j b_{sj} \left( \prod_{k=1}^n SR(\phi_{sjk}, x) + (-1)^{|s|_{\bullet}-1} \prod_{k=1}^n SR(\phi_{sj(n-k+1)}, x) \right) G(\theta, x) \\ &= x + G(\theta, x) \\ &\quad + \sum_{s \in SG, n>0} \sum_j b_j^s (G(\phi_{sj1} \diamond \cdots \diamond \phi_{sjn} \diamond \theta, x) \\ &\quad + (-1)^{|s|_{\bullet}-1} G(\phi_{sjn} \diamond \cdots \diamond \phi_{sj1} \diamond \theta, x)) \\ &= G(\Phi, x), \end{aligned}$$

with

$$\Phi = \hat{e} + \theta + \sum_{s \in SG, n>0} \sum_j b_{sj} \left( \phi_{sj1} \diamond \cdots \diamond \phi_{sjn} \diamond \theta + (-1)^{|s|_{\bullet}-1} \phi_{sjn} \diamond \cdots \diamond \phi_{sj1} \diamond \theta \right). \tag{72}$$

**Theorem 6** *The discrete gradient method (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70) and  $\bar{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  is of order  $p$  if and only if*

$$\Phi(\tau) = \xi(\tau) \quad \text{for } |\tau| \leq p, \tag{73}$$

where  $\Phi$  is given by (72) and the  $\xi$  is given by (69).

We remark that  $\bar{\nabla} H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  is a necessary condition for the method to be a G-series method for all  $S$  and  $H$ , but not for its order;  $\bar{\nabla} H \in C^{p-1}(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  is sufficient for the scheme to be of order  $p$ . The following proposition is presented without its proof, which follows along the lines of the proof of Proposition 3.

**Proposition 6** *The  $\Phi$  of (73) satisfies*

$$\Phi(\tau) = \hat{e}(\tau) + \theta(\tau) + \sum_{i \text{ s.t. } n \geq 1} \Lambda(\hat{\tau}^i, s^i) \tag{74}$$

where  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ ,  $\theta$  is given by (71), and

$$\begin{aligned} \Lambda(\hat{\tau}^i, s^i) &= \theta([\mu_{n+1}^i]_{s_{n+1}^i}) \left( \sum_j b_{s^i j} \phi_{s^i j 1}(\mu_1^i) \cdots \phi_{s^i j n}(\mu_n^i) \right. \\ &\quad \left. + (-1)^{|s^i|_{\bullet}-1} \sum_j b_{\hat{s}^i j} \phi_{\hat{s}^i j n}(\mu_1^i) \cdots \phi_{\hat{s}^i j 1}(\mu_n^i) \right), \end{aligned} \tag{75}$$

with  $\hat{s}^i$  given by  $\hat{s}_k^i = s_{n-k+1}^i$  for  $k = 1, \dots, n$ , and  $\hat{s}_{n+1}^i = s_{n+1}^i$ .

As for the AVF method, one does not need to find the order conditions for every tree; it suffices to find the order condition for each energy-preserving linear combination of the form

$$\omega = [\mu_1]_{s_1} \circ [\mu_2]_{s_2} \circ \dots \circ [\mu_n]_{s_n} \circ [\emptyset]_{\bullet} + (-1)^n [\mu_n]_{s_n} \circ [\mu_{n-1}]_{s_{n-1}} \circ \dots \circ [\mu_1]_{s_1} \circ [\emptyset]_{\bullet}. \tag{76}$$

The above does not give every energy-preserving linear combination of the elementary differentials of G-series; it gives the combinations one gets in the scheme (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70). Now, let again  $I_l$  and  $I_n$  denote the sets of leaf nodes and non-leaf nodes, respectively. If we assume the conditions for order  $< p$  to be satisfied, we have an equivalent order condition to (73) by

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i, s^i) = \xi(\tau) - \hat{e}(\tau) - \sum_{i \in I_n} \Lambda(\hat{\tau}^i, s^i), \tag{77}$$

where we may use the relation

$$\Lambda(\{\mu_1^i, \dots, \mu_n^i, \mu_{n+1}^i\}, s^i) = \hat{\theta}([\mu_{n+1}^i]_{s_{n+1}^i}) \Lambda(\{\mu_1^i, \dots, \mu_n^i, \emptyset\}, \bar{s}^i)$$

to calculate  $\Lambda(\hat{\tau}^i)$  for  $i \in I_n$ . Here  $\bar{s}^i$  is  $s^i$  with  $s_{n+1}^i$  replaced by  $\bullet$ , and  $\hat{\theta}(\emptyset) = 0$ ,  $\hat{\theta}(\bullet) = 1$ , and

$$\begin{aligned} \hat{\theta}([\tau_1, \dots, \tau_m]_{\bullet}) &= \frac{1}{m+1} \xi(\tau_1) \cdots \xi(\tau_m), \\ \hat{\theta}([\tau_1, \dots, \tau_m]_{\blacktriangle}) &= \frac{-2m}{m+1} \xi(\tau_1) \cdots \xi(\tau_m). \end{aligned} \tag{78}$$

Note that  $\Lambda(\hat{\tau}^1, s^1) = \hat{\theta}(\tau)$ .

From the order conditions in Table 5, we can find an  $\bar{S}(x, \hat{x}, h)$  so that (9) becomes a fourth order scheme for any  $\bar{\nabla}H \in C^3(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$ . For instance, the stem  $s = \mathbf{1}$  has the related order conditions  $\sum_j b_{sj} = \frac{1}{2}$  and  $\sum_{j,k} b_{sj} \phi_{sjk}(\bullet) = \frac{2}{3}$ , which sets the requirements for the term

$$h^2 \sum_j b_{sj} (SQ(x, z_{1j})SQ(x, z_{2j}) + SQ(x, z_{2j})SQ(x, z_{1j}))S.$$

**Table 5** Energy-preserving linear combinations of the form (76) and their associated order conditions for the discrete gradient method (9) with  $\bar{S}(x, \hat{x}, h)$  given by (70)

$ \tau $	$\omega$	$s$	Order condition
1	•	•	$\sum_j b_{sj} = \frac{1}{2}$
2	I	I	$\sum_j b_{sj} = \frac{1}{2}$
3	Y	I	$\sum_j b_{sj} \phi_{sj1}(\bullet) = \frac{1}{3}$
	I	I	$\sum_j b_{sj} = \frac{1}{2}$
	I-I	I	$\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$
4	I	I	$\sum_j b_{sj} = -\frac{1}{24}$
	Y	I	$\sum_j b_{sj} \phi_{sj1}(\bullet)^2 = \frac{1}{4}$
	Y <sup>I</sup>	I	$\sum_j b_{sj} \phi_{sj1}(\mathbf{I}) = 0$
	Y <sup>I</sup>	I	$\sum_j b_{sj} \phi_{sj1}(\mathbf{I}) = \frac{1}{6}$
	Y + Y <sup>I</sup>	I	$\sum_{j,k} b_{sj} \phi_{sjk}(\bullet) = \frac{2}{3}$
	Y - Y <sup>I</sup>	I	$\sum_j b_{sj} \phi_{sj2}(\bullet) - \sum_j b_{\bar{s}j} \phi_{\bar{s}j1}(\bullet) = 0$
	Y - Y <sup>I</sup>	I	$\sum_j b_{sj} \phi_{sj2}(\bullet) - \sum_j b_{\bar{s}j} \phi_{\bar{s}j1}(\bullet) = 0$
	Y + Y <sup>I</sup>	I	$\sum_{j,k} b_{sj} \phi_{sjk}(\bullet) = -\frac{1}{24}$
	I	I	$\sum_j b_{sj} = \frac{1}{2}$
	I-I	I	$\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$
4	I+I	I	$\sum_j b_{sj} - \sum_j b_{\bar{s}j} = 0$
	I	I	$\sum_j b_{sj} = 0$

Choosing  $b_{s1} = \frac{1}{2}$  and  $z_{11} = z_{21} = x + \frac{2}{3}hf(x)$ , we have fulfilled these conditions. Likewise, finding terms that satisfy the other order conditions, we get e.g.

$$\begin{aligned}
 \bar{S}(x, h) = & S + hS\left(\frac{8}{9}Q(x, z_3) + \frac{1}{9}Q(x, x)\right)S \\
 & + h^2S(Q(x, z_2)SQ(x, z_2) - \frac{1}{12}D^2H(z_1)SD^2H(z_1))S \\
 & + h^3S(Q(x, x)SQ(x, x)SQ(x, x) \\
 & - \frac{1}{12}D^2H(x)SD^2H(x)SQ(x, x) - \frac{1}{12}Q(x, x)SD^2H(x)SD^2H(x))S,
 \end{aligned}
 \tag{79}$$

where

$$z_1 = x + \frac{1}{2}hf(x), \quad z_2 = x + \frac{2}{3}hf(x), \quad z_3 = x + \frac{3}{4}hf(z_1).$$

### 5.2 The general case

Allowing for  $S$  to be a function of the solution, we define now the set  $TV$  of bi-colored trees whose nodes are either circles or triangles, and whose leaves on the cut tree  $\tau^b$ , defined as the mono-colored tree left when all branches between black and white nodes are cut off, are always circles. Denoting as before black-rooted subtrees by  $\tau_i$  and white-rooted subtrees by  $\bar{\tau}_i$ , the elementary differentials of trees  $\tau \in TV$  are given by  $F(\bullet)(x) = F(\circ)(x) = f(x) = S\nabla H(x)$  and

$$F(\tau)(x) = \begin{cases} S^{(|\tau|)} D^m \nabla H(x)(F(\tau_1)(x), \dots, F(\bar{\tau}_l)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l, ]_{\circ} \\ S^{(|\tau|)} D_2^{m-1} Q(x, x)(F(\tau_1)(x), \dots, F(\bar{\tau}_l)(x)) & \text{for } \tau = [\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l, ]_{\Delta} \end{cases}$$

where  $\circ$  can be either  $\bullet$  or  $\circ$  and  $\Delta$  can be either  $\blacktriangle$  or  $\triangle$ . Let  $TV_{\bullet}$  denote the set of trees in  $TV$  with black roots, either of the shape  $\bullet$  or  $\blacktriangle$ .

**Definition 5** A  $V$ -series is a formal series of the form

$$V(\phi, x) = \phi(\emptyset)x + \sum_{\tau \in TV_{\bullet}} \frac{h^{|\tau|}}{\sigma(\tau)} \phi(\tau) F(\tau)(x), \tag{80}$$

where  $\phi : TV_{\bullet} \cup \{\emptyset\} \rightarrow \mathbb{R}$  is an arbitrary mapping, and the symmetry coefficient  $\sigma$  is given by (28).

Proofs of the theorems in this subsection can be obtained similarly to the proofs in Sects. 4 and 5.1, and are therefore omitted.

We redefine

$$R(\phi_{sjk}, x) := \begin{cases} D^2 H(G(\phi_{sjk}, x)) & \text{if } s_k = \bullet, \\ Q(x, V, (\phi_{sjk}, x)) & \text{if } s_k = \blacktriangle. \end{cases}$$

and consider now approximations of  $S(x)$  that can be written as

$$\begin{aligned} \bar{S}(x, \hat{x}, h) = & \sum_{s \in SG} h^n \sum_j b_{sj} \left( \prod_{k=1}^n S(V(\psi_{sjk}, x)) R(\phi_{sjk}, x) \cdot S(V(\psi_{sj(n+1)}, x)) \right. \\ & \left. + (-1)^{|s|_{\bullet}-1} S(V(\psi_{sj(n+1)}, x)) \prod_{k=1}^n R(\phi_{sj(n-k+1)}, x) S(V(\psi_{sj(n-k+1)}, x)) \right) \end{aligned} \tag{81}$$

whenever  $\hat{x}$  is the solution of

$$\frac{\hat{x} - x}{h} = \bar{S}(x, \hat{x}, h) \bar{\nabla} H(x, \hat{x}),$$

with  $\phi_{sjk}(\emptyset) = \psi_{sjk}(\emptyset) = 1$  for every  $s, j, k$ , and with  $\sum_j b_{\bullet j} = \frac{1}{2}$ .

**Theorem 7** *The discrete gradient scheme (9) with the approximation of  $S(x)$  given by (81) and  $\bar{\nabla}H \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  is a V-series method. It can be written  $\hat{x} = V(\Phi, x)$ , with*

$$\begin{aligned} \Phi = \hat{e} + \sum_{s \in SG} \sum_j b_{sj} & ((\psi_{sj1}, \phi_{sj1}) \diamond \cdots \diamond (\psi_{sjn}, \phi_{sjn}) \diamond \hat{\theta}(\psi_{sj(n+1)})) \\ & + (-1)^n (\psi_{sj(n+1)}, \phi_{sjn}) \diamond \cdots \diamond (\psi_{sj2}, \phi_{sj1}) \diamond \hat{\theta}(\psi_{sj1}). \end{aligned} \tag{82}$$

where

$$\begin{aligned} \hat{\theta}(a)([\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]_{\mathbf{a}}) &= \frac{1}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m) a(\bar{\tau}_1) \cdots a(\bar{\tau}_l), \\ \hat{\theta}(a)([\tau_1, \dots, \tau_m, \bar{\tau}_1, \dots, \bar{\tau}_l]_{\mathbf{a}}) &= \frac{-2m}{m+1} \Phi(\tau_1) \cdots \Phi(\tau_m) a(\bar{\tau}_1) \cdots a(\bar{\tau}_l). \end{aligned} \tag{83}$$

The scheme is of order  $p$  if and only if

$$\Phi(\tau) = \xi(\tau) \quad \text{for } |\tau| \leq p, \tag{84}$$

where

$$\xi(\tau) = \begin{cases} \frac{1}{\gamma(\tau)} & \text{if } \tau \in TP, \\ 0 & \text{otherwise.} \end{cases} \tag{85}$$

As in Sect. 4.2, we cut the branches between black and white nodes, regardless of the shape of the nodes, and denote this tree by  $\tau^b$ . Number the nodes and reattach the cut-off parts. For the node  $i$  and the corresponding stem  $s^i$ , there exists a unique set of forests  $\hat{\tau}^i = \{(\mu_1^i, \eta_1^i), \dots, (\mu_{n+1}^i, \eta_{n+1}^i)\}$  such that

$$\tau = [(\mu_1^i, \eta_1^i)]_{s_1^i} \circ \cdots \circ [(\mu_n^i, \eta_n^i)]_{s_n^i} \circ [(\mu_{n+1}^i, \eta_{n+1}^i)]_{s_{n+1}^i}$$

**Proposition 7** *The  $\Phi$  of (82) satisfies*

$$\Phi(\tau) = \hat{e}(\tau) + \sum_{i=1}^{|\tau^b|} \Lambda(\hat{\tau}^i, s^i) \tag{86}$$

where  $\hat{e}(\emptyset) = 1$  and  $\hat{e}(\tau) = 0$  for all  $\tau \neq \emptyset$ , and

$$\begin{aligned} \Lambda(\hat{\tau}^i, s^i) = \theta([\mu_{n+1}^i]_{s_{n+1}^i}) & \left( \sum_j b_{s^i j} \psi_{s^i j 1}(\eta_1^i) \phi_{s^i j 1}(\mu_1^i) \cdots \phi_{s^i j n}(\mu_n^i) \psi_{s^i j (n+1)}(\eta_{n+1}^i) \right. \\ & \left. + (-1)^{|s^i|-1} \sum_j b_{s^i j} \psi_{s^i j (n+1)}(\eta_1^i) \phi_{s^i j n}(\mu_1^i) \cdots \phi_{s^i j 1}(\mu_n^i) \psi_{s^i j 1}(\eta_n^i) \right), \end{aligned} \tag{87}$$

with  $\theta$  given by (71) and  $\hat{s}^i$  given by  $\hat{s}_k^i = s_{n-k+1}^i$  for  $k = 1, \dots, n$ , and  $\hat{s}_{n+1}^i = s_{n+1}^i$ .

The number of trees in  $TV$  grows very quickly. However, in our task of finding higher order schemes we may use the lessons of the previous sections, and require that the arguments of  $S$ ,  $D^2H$  and  $Q$  in (81) are B-series up to order  $p - 1$ . Then we only need to find order conditions for energy-preserving linear combinations of the form

$$\begin{aligned} \omega &= [(\mu_1, \eta_1)]_{s_1} \circ \cdots \circ [(\mu_n, \eta_n)]_{s_n} \circ [\eta_{n+1}]_{\bullet} \\ &+ (-1)^n [(\mu_n, \eta_{n+1})]_{s_n} \circ \cdots \circ [(\mu_1, \eta_2)]_{s_1} \circ [\eta_1]_{\bullet}, \end{aligned} \tag{88}$$

where  $\mu_i$  and  $\eta_i$  are forests of trees in  $TP_{\bullet}$  and  $TP_{\circ}$  respectively, for  $i = 1, \dots, n + 1$ . Thus we can disregard any tree with  $\Delta$  in it. Furthermore, we may color all nodes of the trees in  $\mu_i$  and  $\eta_i$  except the roots gray, and let the elementary differentials corresponding to these trees be the same as the elementary differentials of B-trees.

We find the order conditions

$$\sum_{i \in I_l} \Lambda(\hat{\tau}^i, s^i) = \xi(\tau) - \hat{\tau}(\tau) - \sum_{i \in I_n} \Lambda(\hat{\tau}^i, s^i),$$

by using the relation

$$\begin{aligned} &\Lambda(\{(\mu_1^i, \eta_1^i), \dots, (\mu_{n+1}^i, \eta_{n+1}^i)\}, s^i) \\ &= \hat{\theta}([(\mu_{n+1}^i]_{s_{n+1}^i}) \Lambda(\{(\mu_1^i, \eta_1^i), \dots, (\emptyset, \eta_{n+1}^i)\}, \bar{s}^i) \end{aligned}$$

to calculate  $\Lambda(\hat{\tau}^i)$  for  $i \in I_n$ . The  $\hat{\theta}$  is given by (78), and  $\bar{s}^i$  is  $s^i$  with  $s_{n+1}^i$  replaced by  $\bullet$ .

**Example 5** Consider  $\mathfrak{Y}$ , which is part of the energy-preserving linear combination  $\mathfrak{I} + \mathfrak{Y}$ . We have two black nodes, and calculate

$$\begin{aligned} \Lambda(\hat{\tau}^1, s^1) &= \Lambda(\{(\bullet, \mathfrak{I})\}, \Delta) = \hat{\theta}(\mathfrak{I}) \Lambda(\{(\emptyset, \mathfrak{I})\}, \bullet) = -\xi(\bullet) \xi(\mathfrak{I}) = -\frac{1}{6}, \\ \Lambda(\hat{\tau}^2, s^2) &= \Lambda(\{(\emptyset, \mathfrak{I}), (\emptyset, \emptyset)\}, \mathfrak{I}) = \sum_j b_{s^2 j} \psi_{s^2 j_1}(\mathfrak{I}) + \sum_j b_{\bar{s}^2 j} \psi_{\bar{s}^2 j_2}(\mathfrak{I}) \\ &= \sum_j b_{s^2 j} (\psi_{s^2 j_1} + \psi_{s^2 j_2})(\mathfrak{I}). \end{aligned}$$

Hence the order condition associated to this linear combination is

$$\sum_j b_{s^2 j} (\psi_{s^2 j_1} + \psi_{s^2 j_2})(\mathfrak{I}) = \frac{1}{6}.$$

An approximation of  $S(x)$  satisfying all the order conditions in Tables 5 and 6 is given by

$$\begin{aligned}
 \bar{S}(x, \cdot, h) = & \frac{1}{2}(S(z_{11} + z_{12}) + S(z_{11} - z_{12})) \\
 & + \frac{1}{12}h(S(z_6)D^2H(z_2)S(x) - S(x)D^2H(z_2)S(z_6)) \\
 & + \frac{3}{7}h(S(z_3)Q(x, z_5)S(z_4) + S(z_4)Q(x, z_5)S(x, z_3)) \\
 & + \frac{8}{105}hS(x)Q(x, z_7)S(x) + \frac{1}{15}hS(x)Q(x, x)S(x) \\
 & + h^2S(z_2)Q(x, z_5)S(z_8)Q(x, z_5)S(z_2) \\
 & - \frac{1}{12}h^2S(z_2)D^2H(z_2)S(z_2)D^2H(z_2)S(z_2) \\
 & + \frac{1}{6}h^2(S(z_2) - S(x))D^2H(x)S(x)Q(x, x)S(x) \\
 & - \frac{1}{6}h^2S(x)Q(x, x)S(x)D^2H(x)(S(z_2) - S(x)) \\
 & + h^3S(x)Q(x, x)S(x)Q(x, x)S(x)Q(x, x) \\
 & - \frac{1}{12}h^3S(x)D^2H(x)S(x)D^2H(x)S(x)Q(x, x)S(x) \\
 & - \frac{1}{12}h^3S(x)Q(x, x)S(x)D^2H(x)S(x)D^2H(x)S(x),
 \end{aligned} \tag{89}$$

with

$$\begin{aligned}
 z_1 &= x + \frac{1}{3}hf(x), & z_2 &= x + \frac{1}{2}hf(x), & z_3 &= x + \frac{7 - \sqrt{7}}{12}hf(z_1), \\
 z_4 &= x + \frac{7 + \sqrt{7}}{12}hf(z_1), & z_5 &= x + \frac{2}{3}hf(z_2), & z_6 &= x + hf(z_2), \\
 z_7 &= x + \frac{5}{4}hf(z_2), & z_8 &= x + \frac{4}{3}hf(z_2), & z_9 &= x + hf(z_6), & z_{10} &= x + hf(z_9), \\
 z_{11} &= \frac{1}{3}(x + z_2 + z_6) + \frac{1}{12}(-z_9 + z_{10}), & z_{12} &= \frac{\sqrt{3}}{36}(7x - 2z_2 - 4z_6 + z_9 - 2z_{10}),
 \end{aligned}$$

and hence a discrete gradient scheme with this  $\bar{S}(x, \cdot, h)$  and any  $\bar{\nabla} \in C^3(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  will be of fourth order.

One advantage of choosing the AVF discrete gradient is that the resulting scheme generally requires fewer computations at each time step. This is clearly evident in the above example: if  $\bar{\nabla} = \bar{\nabla}_{\text{AVF}}$ , then (89) collapses to (67). However, if the AVF discrete gradient is difficult to calculate, there can also be much to gain in computational cost by choosing a symmetric discrete gradient, like the symmetrized Itoh–Abe discrete gradient (23) or the Furihata discrete gradient (25). Then one can ignore the order condition for any combination (88) for which  $s_j = \blacktriangle$  and  $\mu_j = \emptyset$  for some  $j \in [1, n]$ ,





since this corresponds to elementary differentials involving  $Q(x, x)$ , which we recall is zero when the discrete gradient is symmetric. If we consider the conditions for fourth order presented in Tables 5 and 6, this eliminates 17 of the 22 conditions for trees with  $\blacktriangle$  in the stem. By considering the remaining order conditions we get that, if  $\bar{\nabla}H \in C^3(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R}^d)$  and  $\bar{\nabla}H(x, y) = \bar{\nabla}H(y, x)$ , the discrete gradient scheme (9) is of fourth order if

$$\begin{aligned} \bar{S}(x, \cdot, h) = & \frac{1}{2}(S(z_5 + z_6) + S(z_5 - z_6)) \\ & + \frac{1}{12}h(S(z_2)D^2H(z_1)S(x) - S(x)D^2H(z_1)S(z_2)) \\ & + \frac{8}{9}hS(z_1)Q(x, z_7)S(z_1) - \frac{1}{12}h^2S(z_1)D^2H(z_1)S(z_1)D^2H(z_1)S(z_1), \end{aligned} \quad (90)$$

with

$$\begin{aligned} z_1 &= x + \frac{1}{2}hf(x), & z_2 &= x + hf(z_1), & z_3 &= x + hf(z_2), \\ z_4 &= x + hf(z_3), & z_5 &= \frac{1}{3}(x + z_1 + z_2) + \frac{1}{12}(-z_3 + z_4), \\ z_6 &= \frac{\sqrt{3}}{36}(7x - 2z_1 - 4z_2 + z_3 - 2z_4), & z_7 &= x + \frac{3}{4}hf(z_1). \end{aligned}$$

If  $S$  is constant, (90) simplifies to

$$\bar{S}(x, \cdot, h) = S + \frac{8}{9}hSQ(x, z_7)S - \frac{1}{12}h^2SD^2H(z_1)SD^2H(z_1)S. \quad (91)$$

## 6 Numerical experiments

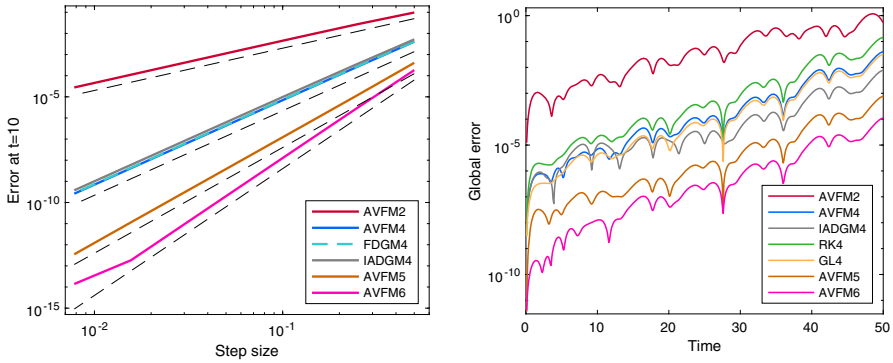
For all the examples considered in the following, the resulting discrete gradient schemes are nonlinearly implicit systems, which we here solve using Newton's method. Note that whenever  $\bar{S}(x, \hat{x}, h)$  is independent of  $\hat{x}$ , the Jacobian of

$$F(\hat{x}) = \hat{x} - x - h\bar{S}(x, \hat{x}, h)\bar{\nabla}H(x, \hat{x})$$

is given by

$$J(\hat{x}) = I - h\bar{S}(x, \hat{x}, h)D_2\bar{\nabla}H(x, \hat{x}), \quad (92)$$

where  $D_2\bar{\nabla}H(x, \hat{x})$  may also be used for the calculation of  $Q(x, \hat{x})$ . The extra computational cost of using a higher-order scheme with an explicit  $\bar{S}$  thus only lies in the computation of this  $\bar{S}$  once each time step. A scheme with an implicitly defined  $\bar{S}(x, \hat{x}, h)$  can give a much more complicated Jacobian, but a quasi-Newton method using the approximate Jacobian given by (92) may still be very efficient.



**Fig. 1** Error plots for the Hénon–Heiles system (93) solved by various discrete gradient methods: AVFM2/4/5/6 are AVF discrete gradient methods of different orders; FDGM4 is a fourth order Furihata discrete gradient method; IADGM4 is a fourth order Itoh–Abe discrete gradient method. RK4 is the classic Runge–Kutta method and GL4 is the fourth order Gauss–Legendre method, included for comparison. The black dashed lines in the order plot are reference lines of order two, four, five and six. The step size in the right plot is  $h = 0.1$

### 6.1 Hénon–Heiles system

The Hénon–Heiles system can be written in the form (6) with

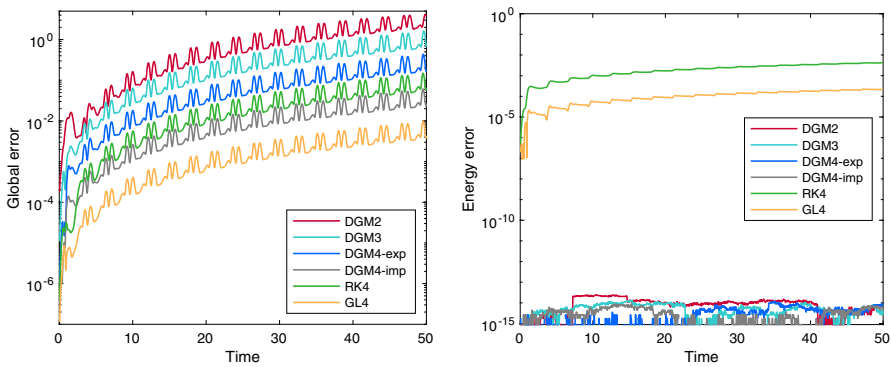
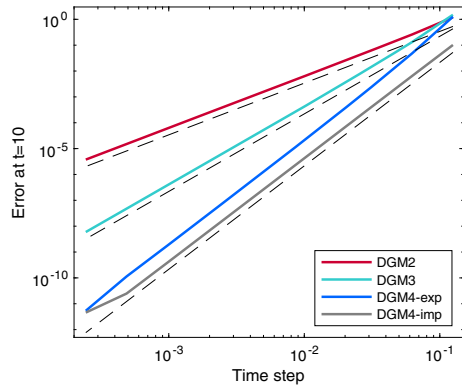
$$S = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \quad H(q, p) = \frac{1}{2}(q_1^2 + q_2^2 + p_1^2 + p_2^2) + q_1^2 q_2 - \frac{1}{3} q_2^3, \quad (93)$$

where  $I$  is the  $2 \times 2$  identity matrix. We use here the same initial conditions used in [37]:  $q_1 = \frac{1}{10}$ ,  $q_2 = -\frac{1}{2}$ ,  $p_1 = p_2 = 0$ . The order of some of the energy-preserving methods proposed in this paper are confirmed by the left plot in Fig. 1. We compare the performance of the fourth order discrete gradient methods obtained by using the  $\bar{S}$  given by (79) coupled with three different discrete gradients: the Itoh–Abe discrete gradient (22), the Furihata discrete gradient (25), and the AVF discrete gradient (10) (right plot of Fig. 1). The symmetrized Itoh–Abe discrete gradient (23) is for this  $H$  identical to the Furihata discrete gradient. The AVF and Furihata discrete gradient methods perform in this case very similarly, and thus the error from the Furihata discrete gradient method is excluded from the right plot in Fig. 1. We observe that, although it initially performs on par with the AVF method, the Itoh–Abe discrete gradient method gives a lower global error than the other fourth order methods as time goes on. Note however that this method requires the most computations at every time step.

### 6.2 Lotka–Volterra

The methods should also be tested on a skew-gradient system with non-constant  $S$ . We choose the Lotka–Volterra system also used for numerical experiments in [10]. It is given by

**Fig. 2** Order of discrete gradient methods applied to the Lotka–Volterra system (94), with different  $\bar{S}$ :  $\bar{S}(x, \hat{x}) = S(\frac{x+\hat{x}}{2})$  for DGM2, (65) for DGM3, (67) for DGM4-exp, (66) for DGM4-imp. The dashed lines are reference lines of order two, three and four



**Fig. 3** Error in the solution and in the energy for discrete gradient methods with  $\bar{S}$  given by  $\bar{S}(x, \hat{x}) = S(\frac{x+\hat{x}}{2})$  for DGM2, (65) for DGM3, (67) for DGM4-exp and (66) for DGM4-imp, applied to the Lotka–Volterra system, with step size  $h = 0.05$ . For comparison, errors from using the standard fourth order Runge–Kutta (RK4) and Gauss–Legendre (GL4) methods are also included

$$S = \frac{1}{2} \begin{pmatrix} 0 & -x_1x_2 & x_1x_3 \\ x_1x_2 & 0 & -2x_2x_3 \\ -x_1x_3 & 2x_2x_3 & 0 \end{pmatrix}, \quad H(x) = 2x_1 + x_2 + 2x_3 + \ln(x_2) - 2 \ln(x_3), \tag{94}$$

and initial conditions  $x_1 = 1, x_2 = \frac{19}{10}, x_3 = \frac{1}{2}$ . For this  $H$ , the Itoh–Abe, Furihata and AVF discrete gradients are all equivalent. Order plots are given in Fig. 2. We consider fourth order discrete gradient methods where  $\nabla S$  is given either dependent on or independent of  $\hat{x}$ ; that is, (66) or (67). The implicitly given (67) yields a significantly lower error in the solution of the corresponding discrete gradient method, as can be witnessed from the left plot in Fig. 3. In contrast to what we observed for the canonical Hamiltonian system studied above, none of the discrete gradient methods give a global error lower than that of the fourth order Gauss–Legendre method.

### 6.3 Hamiltonian neural networks

We present here results on the pendulum problem also considered in [19, 27]. Training data is first generated by adding noise to solutions of the system (6) with

$$S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad H(q, p) = 2mgl(1 - \cos q) + \frac{l^2}{2m}p^2,$$

with  $l = m = 1$  and  $g = 3$ , for various times  $t \in [0, 20]$  and 50 randomly sampled initial coordinates. Then a Hamiltonian neural network (HNN) is used to learn the Hamiltonian. If the automatic discrete differentiation algorithm is used for the training, the corresponding discrete gradient is learned simultaneously. To compute  $Q(x, y)$ , we require that the network also learns the Jacobian of the discrete gradient with respect to the second argument. This necessitates a modification of the network developed by Matsubara et al [27], which we have done using the autograd class of PyTorch. The resulting Jacobian is in any case very useful for integrating the system by any discrete gradient method, since we then can employ a root-finding algorithm that use (92), e.g. Newton's method.

We have tested second, third and fourth order discrete gradient methods both for training the network and for integrating the obtained dynamical system. For the fourth order scheme, we have used the  $\bar{S}$  given by (91) and for the third order scheme we have used

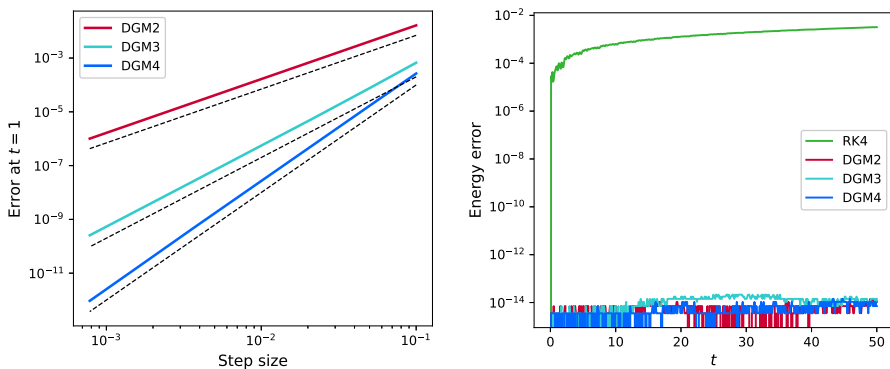
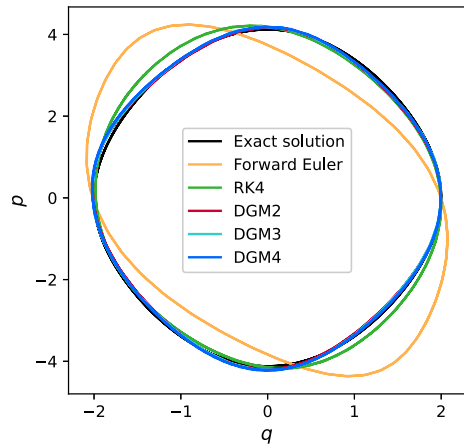
$$\bar{S}(x, \cdot, h) = S + hSQ(x, z)S - \frac{1}{12}h^2SD^2H(x)SD^2H(x)S, \quad (95)$$

where  $z = x + \frac{2}{3}hS\nabla H(x)$ . Note that  $\nabla H$  and  $D^2H$  can be obtained from the network through the relations  $\nabla H(x) = \bar{\nabla}H(x, x)$  and  $D^2H(x) = 2D_2\bar{\nabla}H(x, x)$ .

As is also noted in [27], energy preservation seems to be more important than high order for the method used to train the system. In Fig. 4 we see an example of this: the second order discrete gradient method outperforms the fourth order Runge–Kutta method. Whether or not higher-order discrete gradient methods give an improved performance that justifies the extra computational cost is not clear on this example. A thorough investigation of this for more complex problems is a matter for future studies.

For energy-preserving integration of the learned system, much can be gained in accuracy by using a higher-order discrete gradient method, at the expense of not much additional computational cost. The order plots in Fig. 5 are obtained by comparing to a solution found by the fourth order discrete gradient method with a smaller step size. Python code for the neural networks used to produce the plots in Figs. 4 and 5, and additional results on a mass-spring system, is available at <https://github.com/solveeid/dgnet4>.

**Fig. 4** The trajectories that goes through  $(q, p) = (2, 0)$  for the dynamical systems learned by an HNN with the Euler method, the classic Runge–Kutta method and discrete gradient methods with  $\bar{S}$  given by  $S$  (DGM2), (95) (DGM3) and (91) (DGM4). The plots for DGM2 and DGM3 are almost entirely hidden behind the plot for DGM4. The step size used in the training is  $h = 0.5$



**Fig. 5** Comparison of different integrators for the pendulum system found using an HNN with the second order discrete gradient method. *Left:* Order plots for discrete gradient methods with  $\bar{S}$  given by  $S$  (DGM2), (95) (DGM3) and (91) (DGM4). The dashed lines are reference lines of order two, three and four. *Right:* Energy errors for the same methods, and the fourth-order Runge–Kutta method. Step size  $h = 0.1$

## 7 Conclusions and future work

The main purpose of this paper has been to develop order theory for discrete gradient methods. This was achieved through the introduction of the function  $Q$ , Lemma 2 and a generalization of B- and P-series results. Propositions 3, 5, 6 and 7 present results which simplifies the derivation of the conditions from the general theory. The techniques introduced in this paper for building on B- and P-series methods can possibly be used on more methods than the discrete gradient methods. Future research may utilize this.

We have proposed some higher order schemes satisfying the derived order conditions. The development of specific schemes has however not been the main focus of this paper; analysis to find more optimal schemes is something we leave for the future. After such an analysis is performed, the methods could be tested on more advanced problems than those considered above, e.g. for the temporal discretization of Hamil-

tonian partial differential equations, and their performance as measured by accuracy relative to computational cost could be compared to existing methods.

The use of neural networks to train dynamical systems with preservation properties is an emerging field of study where new developments are coming with a high frequency. To our knowledge, the use of higher-order one-step energy-preserving integrators in this setting has not previously been studied, and may not even have been known to be possible until the novel results of this paper. The potential utility of the methods remain largely unexplored and will be considered in future work. This will include using methods of order higher than four and analysis of the methods compared to and in connection with other recent developments, both in the training of the network and for the integration of the resulting system. It would also be interesting to investigate how the methods perform on more advanced examples, including Hamiltonian PDEs, and for Lagrangian neural networks [1, 11].

The order theory presented in this paper can also be developed further in several different directions. The schemes given with  $\bar{S}$  independent of  $\hat{x}$  are linearly implicit when  $H$  is quadratic; if the order theory is extended to the polarized discrete gradient methods of [12, 13, 28], we could get higher order linearly implicit multi-step schemes for systems with polynomial first integrals of any degree. Another avenue could be to consider order conditions for the discrete Riemannian gradient methods presented in [4]. Then the results in Sect. 5 are especially interesting, since the integral in the AVF discrete Riemannian gradient can be challenging to compute analytically. Lastly, building on [2], order conditions for the exponential AVF method is given in [33]. This could be extended to exponential integrators using any discrete gradient by the theory presented in this paper.

**Acknowledgements** The author is grateful to Brynjulf Owren for the invaluable suggestions he has provided during the writing of this paper. The author would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme *Geometry, compatibility and structure preservation in computational differential equations*, where work on this paper was undertaken. This work was supported by EPSRC Grant No. EP/R014604/1.

**Funding** Open access funding provided by SINTEF AS

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Aoshima, T., Matsubara, T., Yaguchi, T.: Deep discrete-time Lagrangian mechanics. ICLR2021 Workshop on Deep Learning for Simulation (SimDL) (2021)
2. Berland, H., Owren, B., Skaflestad, B.:  $B$ -series and order conditions for exponential integrators. SIAM J. Numer. Anal. **43**(4), 1715–1727 (2005). <https://doi.org/10.1137/040612683>

3. Cai, J., Hong, J., Wang, Y., Gong, Y.: Two energy-conserved splitting methods for three-dimensional time-domain Maxwell's equations and the convergence analysis. *SIAM J. Numer. Anal.* **53**(4), 1918–1940 (2015). <https://doi.org/10.1137/140971609>
4. Celledoni, E., Eidnes, S., Owren, B., Ringholm, T.: Energy-preserving methods on Riemannian manifolds. *Math. Comp.* **89**(322), 699–716 (2020). <https://doi.org/10.1090/mcom/3470>
5. Celledoni, E., Grimm, V., McLachlan, R.I., McLaren, D.I., O'Neale, D., Owren, B., Quispel, G.R.W.: Preserving energy resp. dissipation in numerical PDEs using the "Average Vector Field" method. *J. Comput. Phys.* **231**(20), 6770–6789 (2012). <https://doi.org/10.1016/j.jcp.2012.06.022>
6. Celledoni, E., McLachlan, R.I., McLaren, D.I., Owren, B., Quispel, G.R.W., Wright, W.M.: Energy-preserving Runge–Kutta methods. *M2AN Math. Model. Numer. Anal.* **43**(4), 645–649 (2009). <https://doi.org/10.1051/m2an/2009020>
7. Chartier, P., Faou, E., Murua, A.: An algebraic approach to invariant preserving integrators: the case of quadratic and Hamiltonian invariants. *Numer. Math.* **103**(4), 575–590 (2006). <https://doi.org/10.1007/s00211-006-0003-8>
8. Chen, Z., Zhang, J., Arjovsky, M., Bottou, L.: Symplectic recurrent neural networks. arXiv preprint, [arXiv:1909.13334](https://arxiv.org/abs/1909.13334) (2019)
9. Chorin, A.J., McCracken, M.F., Hughes, T.J.R., Marsden, J.E.: Product formulas and numerical algorithms. *Comm. Pure Appl. Math.* **31**(2), 205–256 (1978). <https://doi.org/10.1002/cpa.3160310205>
10. Cohen, D., Hairer, E.: Linear energy-preserving integrators for Poisson systems. *BIT* **51**(1), 91–101 (2011). <https://doi.org/10.1007/s10543-011-0310-z>
11. Cranmer, M., Greydanus, S., Hoyer, S., Battaglia, P., Spergel, D., Ho, S.: Lagrangian neural networks. arXiv preprint, [arXiv:2003.04630](https://arxiv.org/abs/2003.04630) (2020)
12. Dahlby, M., Owren, B.: A general framework for deriving integral preserving numerical methods for PDEs. *SIAM J. Sci. Comput.* **33**(5), 2318–2340 (2011). <https://doi.org/10.1137/100810174>
13. Eidnes, S., Li, L., Sato, S.: Linearly implicit structure-preserving schemes for Hamiltonian systems. *J. Comput. Appl. Math.* **387**, 112489 (2021). <https://doi.org/10.1016/j.cam.2019.112489>
14. Eidnes, S., Owren, B., Ringholm, T.: Adaptive energy preserving methods for partial differential equations. *Adv. Comput. Math.* **44**(3), 815–839 (2018). <https://doi.org/10.1007/s10444-017-9562-8>
15. Faou, E., Hairer, E., Pham, T.L.: Energy conservation with non-symplectic methods: examples and counter-examples. *BIT* **44**(4), 699–709 (2004). <https://doi.org/10.1007/s10543-004-5240-6>
16. Furihata, D.: Finite difference schemes for  $\partial u / \partial t = (\partial / \partial x)^\alpha \delta G / \delta u$  that inherit energy conservation or dissipation property. *J. Comput. Phys.* **156**(1), 181–205 (1999). <https://doi.org/10.1006/jcph.1999.6377>
17. Furihata, D., Matsuo, T.: *DiscretE Variational Derivative Method*. Chapman & Hall/CRC Numerical Analysis and Scientific Computing. CRC Press, Boca Raton (2011). *A Structure-Preserving Numerical Method for Partial Differential Equations*
18. Gonzalez, O.: Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.* **6**(5), 449–467 (1996). <https://doi.org/10.1007/s003329900018>
19. Greydanus, S., Dzamba, M., Yosinski, J.: Hamiltonian neural networks. *Advances in Neural Information Processing Systems* **32**, 15379–15389 (2019)
20. Hairer, E.: Energy-preserving variant of collocation methods. *JNAIAM. J. Numer. Anal. Ind. Appl. Math.* **5**(1–2), 73–84 (2010)
21. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration*, Springer Series in Computational Mathematics, vol. 31, 2nd edn. Springer, Berlin (2006). *Structure-Preserving Algorithms for Ordinary Differential Equations*
22. Harten, A., Lax, P.D., van Leer, B.: On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.* **25**(1), 35–61 (1983). <https://doi.org/10.1137/1025002>
23. Itoh, T., Abe, K.: Hamiltonian-conserving discrete canonical equations based on variational difference quotients. *J. Comput. Phys.* **76**(1), 85–102 (1988). [https://doi.org/10.1016/0021-9991\(88\)90132-5](https://doi.org/10.1016/0021-9991(88)90132-5)
24. Jiang, C., Sun, J., Li, H., Wang, Y.: A fourth-order AVF method for the numerical integration of sine-Gordon equation. *Appl. Math. Comput.* **313**, 144–158 (2017). <https://doi.org/10.1016/j.amc.2017.05.055>
25. Kriksin, Y.A.: A conservative difference scheme for a system of Hamiltonian equations with external action. *Zh. Vychisl. Mat. i Mat. Fiz.* **33**(2), 206–218 (1993)
26. Marsden, J.E.: *Lectures on Mechanics*, London Mathematical Society Lecture Note Series, vol. 174. Cambridge University Press, Cambridge (1992)



27. Matsubara, T., Ishikawa, A., Yaguchi, T.: Deep energy-based modeling of discrete-time physics. *Adv. Neural Inf. Process. Syst.* **33**, 13100–13111 (2020)
28. Matsuo, T., Furihata, D.: Dissipative or conservative finite-difference schemes for complex-valued nonlinear partial differential equations. *J. Comput. Phys.* **171**(2), 425–447 (2001). <https://doi.org/10.1006/jcph.2001.6775>
29. Matsuo, T., Sugihara, M., Furihata, D., Mori, M.: Spatially accurate dissipative or conservative finite difference schemes derived by the discrete variational method. *Jpn. J. Ind. Appl. Math.* **19**(3), 311–330 (2002). <https://doi.org/10.1007/BF03167482>
30. McLachlan, R.I., Quispel, G.R.W., Robidoux, N.: Geometric integration using discrete gradients. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**(1754), 1021–1045 (1999). <https://doi.org/10.1098/rsta.1999.0363>
31. McLaren, D.I., Quispel, G.R.W.: Integral-preserving integrators. *J. Phys. A* **37**(39), L489–L495 (2004). <https://doi.org/10.1088/0305-4470/37/39/L01>
32. McLaren, D.I., Quispel, G.R.W.: Bootstrapping discrete-gradient integral-preserving integrators to fourth order. In: Daniel, M., Rajasekar, S. (Eds.) *Nonlinear Dynamics*, pp. 157–171. Narosa Publishing House (2009)
33. Mei, L., Huang, L., Wu, X.: Energy-preserving exponential integrators of arbitrarily high order for conservative or dissipative systems with highly oscillatory solutions. *J. Comput. Phys.* **442**, Paper No. 110429, 22 (2021). <https://doi.org/10.1016/j.jcp.2021.110429>
34. Nørsett, S.P., Wolfbrandt, A.: Order conditions for Rosenbrock type methods. *Numer. Math.* **32**(1), 1–15 (1979). <https://doi.org/10.1007/BF01397646>
35. Norton, R.A., McLaren, D.I., Quispel, G.R.W., Stern, A., Zanna, A.: Projection methods and discrete gradient methods for preserving first integrals of ODEs. *Discrete Contin. Dyn. Syst.* **35**(5), 2079–2098 (2015). <https://doi.org/10.3934/dcds.2015.35.2079>
36. Norton, R.A., Quispel, G.R.W.: Discrete gradient methods for preserving a first integral of an ordinary differential equation. *Discrete Contin. Dyn. Syst.* **34**(3), 1147–1170 (2014). <https://doi.org/10.3934/dcds.2014.34.1147>
37. Quispel, G.R.W., McLaren, D.I.: A new class of energy-preserving numerical integration methods. *J. Phys. A* **41**(4), 045206, 7 (2008). <https://doi.org/10.1088/1751-8113/41/4/045206>
38. Yaguchi, T., Matsuo, T., Sugihara, M.: The discrete variational derivative method based on discrete differential forms. *J. Comput. Phys.* **231**(10), 3963–3986 (2012). <https://doi.org/10.1016/j.jcp.2012.01.035>
39. Zhong, Y.D., Dey, B., Chakraborty, A.: Symplectic ODE-net: learning Hamiltonian dynamics with control. In: *International Conference on Learning Representations* (2019)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.