MATHEMATICS OF COMPUTATION
https://doi.org/10.1090/mcom/3470
Article electronically published on September 6, 2019

ENERGY-PRESERVING METHODS ON RIEMANNIAN MANIFOLDS

ELENA CELLEDONI, SØLVE EIDNES, BRYNJULF OWREN, AND TORBJØRN RINGHOLM

Abstract. The energy-preserving discrete gradient methods are generalized to finite-dimensional Riemannian manifolds by definition of a discrete approximation to the Riemannian gradient, a retraction, and a coordinate center function. The resulting schemes are formulated only in terms of these three objects and do not otherwise depend on a particular choice of coordinates or embedding of the manifold in a Euclidean space. Generalizations of well-known discrete gradient methods, such as the average vector field method and the Itoh–Abe method, are obtained. It is shown how methods of higher order can be constructed via a collocation-like approach. Local and global error bounds are derived in terms of the Riemannian distance function and the Levi-Civita connection. Numerical results are presented, for problems on the two-sphere, the paraboloid, and the Stiefel manifold.

1. Introduction

A first integral of an ordinary differential equation (ODE) is a scalar-valued function on the phase space of the ODE that is preserved along solutions. The potential benefit of using numerical methods that preserve one or more such invariants is well-documented, and several energy-preserving methods have been developed in recent years. Among these are the discrete gradient methods, which were introduced for use in Euclidean spaces in [10]; see also [23]. These methods are based on the idea of expressing the ODE using a skew-symmetric operator and the gradient of the first integral, and then creating a discrete counterpart to this in such a way that the numerical scheme preserves the first integral.

For manifolds in general, one can use the same schemes expressed in local coordinates. A drawback is that the numerical approximation will typically depend on the particular choice of coordinates and also on the strategy used for transition between coordinate charts. Another alternative is to use a global embedding of the manifold into a larger Euclidean space, but then it typically happens that the numerical solution deviates from the manifold. Even if the situation can be amended by using projection, it may not be desirable that the computed approximation depends on the particular embedding chosen. Crouch and Grossmann [7]

Received by the editor May 24, 2018, and, in revised form, January 17, 2019, and May 31, 2019.

2010 *Mathematics Subject Classification.* Primary 37K05; Secondary 53B99, 65L05.

Key words and phrases. Geometric integration, discrete gradients, Riemannian manifolds, numerical analysis.

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 691070.

and Munthe-Kaas [26, 27] introduced different ways of extending existing Runge–Kutta methods to a large class of differentiable manifolds. Both these approaches are generally classified as Lie group integrators; see [14] or the more recent [4] for a survey of this class of methods. They can also both be formulated abstractly by means of a post-Lie structure which consists of a Lie algebra with a flat connection of constant torsion; see, e.g., [25]. In the present paper we shall state the methods in a slightly different context, using the notion of a Riemannian manifold. It is then natural to make use of the Levi-Civita connection, which in contrast to the post-Lie setting is torsion-free, and which in general has a non-zero curvature. For our purposes it is also an advantage that the Riemannian metric provides an intrinsic definition of the gradient. Taking an approach more in line with this, Leimkuhler and Patrick [19] considered mechanical systems on the cotangent bundle of a Riemannian manifold and succeeded in generalizing the classical leap-frog scheme to a symplectic integrator on Riemannian manifolds.

Some classical numerical methods in Euclidean spaces preserve certain classes of invariants; for instance, symplectic Runge–Kutta methods preserve all quadratic invariants. This can be useful when there is a natural way of embedding a manifold into a linear space by using constraints that are expressed by means of such invariants. An example is the two-sphere which can be embedded in $\mathbb{R}^3$ by adding the constraint that these vectors should have unit length. The classical midpoint rule will automatically ensure that the numerical approximations remain on the sphere as it preserves all quadratic invariants. In general, however, the invariants preserved by these methods are expressed in terms of coordinates. Hence the preservation property of the method may be lost under coordinate changes if the invariant is no longer quadratic. In [5], a generalization of the discrete gradient method to differential equations on Lie groups and a broad class of manifolds was presented. Here we develop this further by introducing a Riemannian structure that can be used to provide an intrinsic definition of the gradient and a means to measure numerical errors.

The structure of this paper is as follows: In section 2, we formulate the problem to be solved and introduce discrete Riemannian gradient methods, as well as presenting some particular examples with special attention to a generalization of the Itoh–Abe discrete gradient. We also briefly discuss the Euclidean setting as a special choice of manifold and show how the standard discrete gradient methods are recovered in this case. In the third section, we consider higher order energy-preserving methods based on generalization of a collocation strategy introduced by Hairer [12] to Riemannian manifolds. We present some error analysis in section 4, and show numerical results in section 5, where the methods are applied to models of a body moving on the two-sphere and on the paraboloid, and on a system on the Stiefel manifold.

2. Energy preservation on Riemannian manifolds

Consider an initial value problem on the finite-dimensional Riemannian manifold (M, g),

$$\dot{u} = F(u), \quad u(0) = u^0 \in M, \tag{2.1}$$

where F is a smooth vector field, $u^0 \in M$ is the initial value, and g is a Riemannian metric. We denote by $\mathcal{F}(M)$ the space of smooth functions on M. The set of smooth vector fields and differential one-forms are denoted $\Gamma(TM)$ and $\Gamma(T^*M)$, respectively, and for the duality pairing between these two spaces we use the angle brackets $\langle\cdot,\cdot\rangle$.

A first integral associated with a vector field $F \in \Gamma(TM)$ is a function $H \in \mathcal{F}(M)$ such that $\langle \mathrm{d}H, F\rangle$ vanishes identically on M. First integrals are preserved along solutions of (2.1), since

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u(t)) = \langle \mathrm{d}H(u(t)), \dot{u}(t)\rangle = \langle \mathrm{d}H(u(t)), F(u(t))\rangle = 0.$$

2.1. **Preliminaries.** The fact that a vector field F has a first integral H is closely related to the existence of a tensor field $\Omega \in \Gamma(TM \otimes T^*M) =: \Gamma(\mathcal{T}_1^1 M)$, skew-symmetric with respect to the metric g, such that

$$F(u) = \Omega(u)\ \mathrm{grad}\, H(u), \tag{2.2}$$

where $\mathrm{grad}\, H \in \Gamma(TM)$ is the Riemannian gradient, the unique vector field satisfying $\langle \mathrm{d}H, \cdot\rangle = g(\mathrm{grad}\, H, \cdot)$. Any ODE (2.1) where F is of this form preserves H, since

$$\frac{\mathrm{d}}{\mathrm{d}t}H(u) = \langle \mathrm{d}H(u), \dot{u}\rangle = \langle \mathrm{d}H(u), \Omega\ \mathrm{grad}\, H(u)\rangle = g(\mathrm{grad}\, H(u), \Omega\ \mathrm{grad}\, H(u)) = 0.$$

A converse result is detailed in the following proposition.

Proposition 1. *Any system* (2.1) *with a first integral H can be written with an F of the form* (2.2)*. The skew tensor field Ω can be chosen so as to be bounded near every non-degenerate critical point of H.*

Proof. Similar to the proof of Proposition 2.1 in [23], we can write an explicit expression for a possible choice of Ω,

$$\Omega y = \frac{g(\mathrm{grad}\, H, y)\, F - g(F, y)\ \mathrm{grad}\, H}{g(\mathrm{grad}\, H, \mathrm{grad}\, H)}. \tag{2.3}$$

Clearly, $g(y, \Omega y) = 0$ for all y. Since H is a first integral, $g(F, \mathrm{grad}\, H) = \langle \mathrm{d}H, F\rangle = 0$, so $\Omega\ \mathrm{grad}\, H = F$. For a proof that Ω is bounded near non-degenerate critical points, see [23]. $\square$

In fact, such a tensor field Ω often arises naturally from a two-form ω through the assignment $\Omega y = \omega(\cdot, y)^\sharp$. A well-known example is when ω is a symplectic two-form. Note that Ω is not necessarily unique.

Retractions, viewed as maps from TM to M, will play an important role in the methods we discuss here. Their formal definition can be found, e.g., in [1].

Definition 1. Let ϕ be a smooth map defined on a neighborhood of M in TM and let ϕ_p denote the restriction of ϕ to the tangent space T_pM at $p \in M$, with 0_p being the zero-vector in T_pM. Then ϕ is a *retraction* if it satisfies the conditions

(1) ϕ_p is defined in an open ball $B_{r_p}(0_p) \subset T_pM$ of radius r_p about 0_p,
(2) $\phi_p(x) = p$ if and only if $x = 0_p$,
(3) $D\phi_p\big|_{0_p} = \mathrm{Id}_{T_pM}$.

A canonical example of a retraction on (M, g) is obtained via the Riemannian exponential, setting $\phi_p(x) = \exp_p(x)$, i.e., following along the geodesic emanating from p in the direction x. The Riemannian exponential may be more computationally expensive to evaluate than other retractions, but its geometric position in the Riemannian framework could provide for an informative error analysis.

2.2. **The discrete Riemannian gradient method.** We adapt the discrete gradients in Euclidean space to discrete Riemannian gradients (DRG) on (M, g) by means of a retraction map ϕ and a center point function c.

Definition 2. A discrete Riemannian gradient is a triple $(\overline{\mathrm{grad}}, \phi, c)$[1] where

(1) $c : M \times M \to M$ is a continuous map such that $c(u, u) = u$ for all $u \in M$,
(2) $\overline{\mathrm{grad}} : \mathcal{F}(M) \to \Gamma(c^*TM)$,
(3) $\phi : TM \to M$ is a retraction,

such that for all $H \in \mathcal{F}(M)$, $u \in M$, $v \in M$, $c = c(u, v) \in M$,

$$H(v) - H(u) = g(\overline{\mathrm{grad}}H(u, v), \phi_c^{-1}(v) - \phi_c^{-1}(u)), \tag{2.4}$$

$$\overline{\mathrm{grad}}H(u, u) = \mathrm{grad}\, H(u). \tag{2.5}$$

The DRG $\overline{\mathrm{grad}}H$ is a continuous section of the pullback bundle c^*TM, meaning that $\pi \circ \overline{\mathrm{grad}}H = c$, where $\pi : TM \to M$ is the natural projection. We also need to define an approximation to be used for the tensor field $\Omega \in \Gamma(\mathcal{T}_1^1 M)$. To this end we let $\overline{\Omega} \in \Gamma(c^*\mathcal{T}_1^1 M)$ be a continuous skew-symmetric tensor field such that

$$\overline{\Omega}(u, u) = \Omega(u) \quad \forall u \in M.$$

Inspired by [3, 5], we propose the scheme

$$u^{k+1} = \phi_{c^k}(W(u^k, u^{k+1})), \quad c^k = c(u^k, u^{k+1}), \tag{2.6}$$

$$W(u^k, u^{k+1}) = \phi_{c^k}^{-1}(u^k) + h\, \overline{\Omega}(u^k, u^{k+1})\, \overline{\mathrm{grad}}H(u^k, u^{k+1}), \tag{2.7}$$

where h is the step size. The scheme (2.6)–(2.7) preserves the invariant H, since

$$\begin{aligned} H(u^{k+1}) - H(u^k) &= g(\overline{\mathrm{grad}}H(u^k, u^{k+1}), \phi_{c^k}^{-1}(u^{k+1}) - \phi_{c^k}^{-1}(u^k)) \\ &= g(\overline{\mathrm{grad}}H(u^k, u^{k+1}), h\, \overline{\Omega}(u^k, u^{k+1})\, \overline{\mathrm{grad}}H(u^k, u^{k+1})) = 0. \end{aligned}$$

Here and in the following we adopt the shorthand notation $c = c(u, v)$ as long as it is obvious what the arguments of c are.

The Average Vector Field (AVF) method has been studied extensively in the literature; some early references are [13, 23, 28]. This is a discrete gradient method, and we propose a corresponding DRG satisfying (2.4)-(2.5) as follows:

$$\overline{\mathrm{grad}}_{\mathrm{AVF}}H(u, v) = \int_0^1 (D_{\gamma_\xi}\phi_c)^{\mathrm{T}}\ \mathrm{grad}\, H(\phi_c(\gamma_\xi))\, \mathrm{d}\xi, \quad \gamma_\xi = (1-\xi)\phi_c^{-1}(u) + \xi\phi_c^{-1}(v), \tag{2.8}$$

where $(D_x\phi_c)^{\mathrm{T}} : T_{\phi_c(x)}M \to T_cM$ is the unique operator satisfying

$$g((D_x\phi_c)^{\mathrm{T}} a, b) = g(a, D_x\phi_c\, b) \quad \forall x, b \in T_cM, \quad a \in T_{\phi_c(x)}M.$$

Furthermore, we have the generalization of Gonzalez's midpoint discrete gradient [10],

$$\overline{\mathrm{grad}}_{\mathrm{MP}}H(u, v) = \mathrm{grad}\, H(c(u, v)) + \frac{H(v) - H(u) - g(\mathrm{grad}\, H(c(u, v)), \eta)}{g(\eta, \eta)}\eta \tag{2.9}$$

where $\eta = \phi_c^{-1}(v) - \phi_c^{-1}(u)$.

Note that both these DRGs involve the gradient of the first integral. This may be a disadvantage if H is non-smooth or if its gradient is expensive to compute. Also, the implicit nature of the schemes requires the solution of an n-dimensional

[1] To avoid cluttered notation we will just write $\overline{\mathrm{grad}}$ for the triple $(\overline{\mathrm{grad}}, \phi, c)$ in what follows.

non-linear system of equations at each time step. An alternative is to consider the Itoh–Abe discrete gradient [15], also called the coordinate increment discrete gradient [23], which in certain cases requires only the solution of n decoupled scalar equations. We now present a generalization of the Itoh–Abe discrete gradient to finite-dimensional Riemannian manifolds.

2.3. **Itoh–Abe discrete Riemannian gradient.**

Definition 3. For any tangent space T_cM one can choose a basis $\{E_1, ..., E_n\}$ composed of tangent vectors E_i, $i = 1, ..., n$, orthonormal with respect to the Riemannian metric g. Then, given $u, v \in M$, there exists a unique $\{\alpha_i\}_{i=1}^n$ so that

$$\phi_c^{-1}(v) - \phi_c^{-1}(u) = \sum_{i=1}^n \alpha_i E_i.$$

The *Itoh–Abe DRG* of the first integral H is then given by

$$\overline{\mathrm{grad}}_{\mathrm{IA}} H(u,v) = \sum_{j=1}^n a_j E_j, \tag{2.10}$$

where

$$a_j = \begin{cases} \dfrac{H(w_j) - H(w_{j-1})}{\alpha_j} & \text{if } \alpha_j \neq 0, \\ g(\mathrm{grad}\, H(w_{j-1}), D\phi_c(\eta_{j-1})E_j) & \text{if } \alpha_j = 0, \end{cases}$$

$$w_j = \phi_c(\eta_j), \quad \eta_j = \phi_c^{-1}(u) + \sum_{i=1}^j \alpha_i E_i.$$

We refer to [3] for proof that this is indeed a DRG satisfying (2.4)-(2.5).

2.4. **Euclidean setting.** Let $M = V$ be an $\mathbb{R}$-linear space, and let g be the Euclidean inner product, $g(x,y) = x^{\mathrm{T}}y$. The operator Ω is a solution dependent skew-symmetric $n \times n$ matrix $\Omega(u)$. For any $u \in V$, we have $T_uV \equiv V$. The retraction $\phi : V \to V$ is defined as $\phi_p(x) = p + x$, the Riemannian exponential on V, so that $\phi_c^{-1}(v) - \phi_c^{-1}(u) = v - u$. The gradient $\mathrm{grad}\, H$ is an n-vector whose ith component is $\frac{\partial H}{\partial u_i}$, and the definition of the discrete Riemannian gradient coincides with the standard discrete gradient, since (2.4) now reads

$$H(v) - H(u) = \overline{\mathrm{grad}}(u,v)^{\mathrm{T}}(v-u).$$

Furthermore, (2.6)-(2.7) simply becomes the discrete gradient method introduced in [10], given by the scheme

$$u^{k+1} - u^k = h\overline{\Omega}(u^k, u^{k+1})\, \overline{\mathrm{grad}} H(u^k, u^{k+1}), \tag{2.11}$$

where $\overline{\Omega}$ is a skew-symmetric matrix approximating Ω. Two typical choices are $\overline{\Omega}(u^k, u^{k+1}) = \Omega(u^k)$, or $\overline{\Omega}(u^k, u^{k+1}) = \Omega((u^{k+1} + u^k)/2)$ if one seeks a symmetric method.

The DRGs (2.8) and (2.9) become the standard AVF and midpoint discrete gradients in this case. For the Itoh–Abe DRG, the practical choice for the orthogonal basis would be the set of unit vectors, $\{e_1, ..., e_n\}$, so that $\alpha_i = v_i - u_i$, and we get (2.10) with

$$a_j = \begin{cases} \dfrac{H(w_j) - H(w_{j-1})}{v_j - u_j} & \text{if } u_j \neq v_j, \\ \frac{\partial H}{\partial u_j}(w_{j-1}) & \text{if } u_j = v_j, \end{cases}$$
$$w_j = \sum_{i=1}^{j} v_i e_i + \sum_{i=j+1}^{n} u_i e_i,$$

which is a reformulation of the Itoh–Abe discrete gradient as it is given in [15], [23] and the literature otherwise.

2.5. **Lie group setting.** Consider the case where M is a Lie group, $M = G$, equipped with a right-invariant Riemannian metric g. The methods described in reference [5] can be seen as a special case of the methods presented in the current paper, with the retraction map chosen to be the Lie group exponential; see [5] for details. In the special case when the Riemannian metric is bi-invariant (and the exponential map of the Lie group setting coincides with the Riemannian exponential [24]) the methods of [5] are an example of the methods presented here, implemented using normal coordinates (see [17, p. 76]).

3. Methods of higher order

In the Euclidean setting, a strategy to obtain energy-preserving methods of higher order was presented in [2] and later in [12]; see also [6]. This technique is generalized to a Lie group setting in [5]. Here we will formulate these methods in the context of Riemannian manifolds.

3.1. **Energy-preserving collocation-like methods on Riemannian manifolds.** Let $\mathsf{c}_1, ..., \mathsf{c}_s$ be distinct real numbers, where s is the order of the collocation polynomial specified below. Consider the Lagrange basis polynomials,

$$l_i(\xi) = \prod_{j=1, j\neq i}^{s} \frac{\xi - \mathsf{c}_j}{\mathsf{c}_i - \mathsf{c}_j}, \quad \text{and let} \quad b_i := \int_0^1 l_i(\xi)\, d\xi. \tag{3.1}$$

We assume that $\mathsf{c}_1, \ldots, \mathsf{c}_s$ are such that $b_i \neq 0$ for all i. A step of the energy-preserving collocation-like method, starting at $u^0 \in M$, is defined via a polynomial $\sigma : \mathbb{R} \to T_cM$ of degree s satisfying

$$\sigma(0) = \phi_c^{-1}(u^0), \tag{3.2}$$

$$\frac{d}{d\xi}\sigma(\xi h)\Big|_{\xi=\mathsf{c}_j} = D_{U_j}\phi_c^{-1}\left(\Omega_j \operatorname{grad}_j H\right), \quad U_j := \phi_c\left(\sigma(\mathsf{c}_j h)\right), \tag{3.3}$$

$$u^1 := \phi_c\left(\sigma(h)\right), \tag{3.4}$$

where

$$\operatorname{grad}_j H := \int_0^1 \frac{l_j(\xi)}{b_j}\left(D_{U_j}\phi_c^{-1}\right)^{\mathrm{T}} \left(D_{\sigma(\xi h)}\phi_c\right)^{\mathrm{T}} \operatorname{grad} H\left(\phi_c(\sigma(\xi h))\right) \mathrm{d}\xi,$$

and $\Omega_j := \Omega(U_j)$. Notice that with $s = 1$ and independently on the choice of c_1, we reproduce the DRG method (2.6)-(2.7) with the AVF DRG (2.8).

Using Lagrange interpolation and (3.3), the derivative of $\sigma(\xi h)$ at every point ξh is

$$\frac{d}{d\xi}\sigma(\xi h) = \sum_{j=1}^{s} l_j(\xi) D_{U_j}\phi_c^{-1}\left(\Omega_j \operatorname{grad}_j H\right), \tag{3.5}$$

from which by integrating we get

$$\sigma(\tau h) = \phi_c^{-1}(u_0) + h\sum_{j=1}^{s}\int_0^{\tau} l_j(\xi)\,\mathrm{d}\xi\, D_{U_j}\phi_c^{-1}\left(\Omega_j \operatorname{grad}_j H\right).$$

The defined method is energy preserving, which we see by using

$$\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\phi_c(\sigma(\xi h))\right) = D_{\sigma(\xi h)}\phi_c\left(\frac{\mathrm{d}}{\mathrm{d}\xi}\sigma(\xi h)\right),$$

and (3.5) to get

$$\begin{aligned}
&H(u^1) - H(u^0)\\
&\quad= \int_0^1 g\left(\operatorname{grad} H\left(\phi_c(\sigma(\xi h))\right), \frac{\mathrm{d}}{\mathrm{d}\xi}\phi_c(\sigma(\xi h))\right)\mathrm{d}\xi\\
&\quad= \int_0^1 g\left(\operatorname{grad} H\left(\phi_c(\sigma(\xi h))\right), D_{\sigma(\xi h)}\phi_c\left(\sum_{j=1}^{s} l_j(\xi) D_{U_j}\phi_c^{-1}\left(\Omega_j \operatorname{grad}_j H\right)\right)\right)\mathrm{d}\xi\\
&\quad= \int_0^1 g\left(\left(D_{\sigma(\xi h)}\phi_c\right)^{\mathrm{T}}\operatorname{grad} H\left(\phi_c(\sigma(\xi h))\right), \sum_{j=1}^{s} l_j(\xi) D_{U_j}\phi_c^{-1}\left(\Omega_j \operatorname{grad}_j H\right)\right)\mathrm{d}\xi\\
&\quad= \sum_{j=1}^{s} b_j g\left(\int_0^1 \frac{l_j(\xi)}{b_j}\left(D_{U_j}\phi_c^{-1}\right)^{\mathrm{T}}\left(D_{\sigma(\xi h)}\phi_c\right)^{\mathrm{T}}\operatorname{grad} H\left(\phi_c(\sigma(\xi h))\right)\mathrm{d}\xi, \Omega_j \operatorname{grad}_j H\right)\\
&\quad= \sum_{j=1}^{s} b_j g\left(\operatorname{grad}_j H, \Omega_j \operatorname{grad}_j H\right) = 0,
\end{aligned}$$

and hence repeated use of (3.2)-(3.4) ensures $H(u^k) = H(u^0)$ for all $k \in \mathbb{N}$.

3.2. **Higher order extensions of the Itoh–Abe DRG method.** From the Itoh–Abe DRG one can get a new DRG, also satisfying (2.4), by

$$\overline{\operatorname{grad}}_{\mathrm{SIA}} H(u,v) = \frac{1}{2}\left(\overline{\operatorname{grad}}_{\mathrm{IA}} H(u,v) + \overline{\operatorname{grad}}_{\mathrm{IA}} H(v,u)\right). \tag{3.6}$$

We call this the *symmetrized Itoh–Abe DRG*. Note that we need the base point c to be the same in the evaluation of $\overline{\operatorname{grad}}_{\mathrm{IA}} H(u,v)$ and $\overline{\operatorname{grad}}_{\mathrm{IA}} H(v,u)$. When $c(u,v) = c(v,u)$ and $\overline{\Omega}_{(u,v)} = \overline{\Omega}_{(v,u)}$, we get a symmetric DRG method (2.6)-(2.7), which is therefore of second order.

Alternatively, one can get a symmetric two-stage method by a composition of the Itoh–Abe DRG method and its adjoint. Furthermore, one can get energy-preserving methods of any order using a composition strategy. To ensure symmetry of an s-stage composition method, one needs $c_i(u,v) = c_{s+1-i}(v,u)$ for different center points c_i belonging to each stage and, similarly, $\overline{\Omega}_i(u,v) = \overline{\Omega}_{s+1-i}(v,u)$.

4. Error analysis

4.1. **Local error.** In this section, $\varphi_t(u)$ is the t-flow of the ODE vector field F. The most standard discrete gradient methods have a low or moderate order of convergence, and this is also the case for the DRG methods unless special care is taken in designing $\overline{\Omega}$ and $\overline{\mathrm{grad}}H$. We shall not pursue this approach here, but refer to the collocation-like methods if high order of accuracy is required. We shall see, however, that the methods designed here are consistent and can be made symmetric. Analysis of the local error can be done in local coordinates, assuming that the step size is always chosen sufficiently small, so that within a fixed step, $u^k, u^{k+1}, c(u^k, u^{k+1})$ and the exact local solution $u(t_{k+1})$ all belong to the same given coordinate chart. From the definition (2.6)-(2.7) it follows immediately that the representation of $u^{k+1}(h)$ satisfies $u^{k+1}(0) = u^k$ and $\frac{d}{dh}u^{k+1}(0) = F(u^k)$. Then, by equivalence of local coordinate norms and the Riemannian distance, we may conclude that the local error in DRG methods satisfies

$$d(u^{k+1}, \varphi_h(u^k)) \leq Ch^2.$$

Similar to what was also observed in [5], the DRG methods (2.6)-(2.7) are symmetric whenever $\overline{\mathrm{grad}}H(u,v) = \overline{\mathrm{grad}}H(v,u)$, $\overline{\Omega}(u,v) = \overline{\Omega}(v,u)$, and $c(u,v) = c(v,u)$ for all $u, v \in M$. In that case we obtain an error bound for the local error of the form $d(u^{k+1}, \varphi_h(u^k)) \leq Ch^3$.

The collocation-like methods of section 3 have associated nodes $\{\mathsf{c}_i\}_{i=1}^s$ and weights $\{b_i\}_{i=1}^s$ defined by (3.1). The order of the local error depends on the accuracy of the underlying quadrature formula given by these nodes and weights. The following result is a simple consequence of Theorem 4.3 in [6].[2]

Theorem 1. *Let ψ_h be the method defined by* (3.2)-(3.4)*. The order of the local error is at least*

$$p = \min(r, 2r - 2s + 2),$$

where r is the largest integer such that $\sum_{i=1}^s b_i c_i^{q-1} = \frac{1}{q}$ for all $1 \leq q \leq r$. This means that there are positive constants C and h_0 such that

$$d(\psi_h(u), \varphi_h(u)) \leq C\, h^{p+1} \quad \textit{for } h < h_0,\ u \in M.$$

Proof. Choose h small enough such that the solution can be represented in the form $u(h\xi) = \phi_c(\gamma(\xi h)), \xi \in [0,1]$, and consider the corresponding differential equation for γ in T_cM:

$$\frac{d}{dt}\gamma(t) = (\phi_c^* F)(\gamma(t)) = \left(T_{\gamma(t)}\phi_c\right)^{-1} \Omega \ \mathrm{grad}\, H\left(\phi_c(\gamma(t))\right). \tag{4.1}$$

Notice that $(T_\gamma \phi_c)^{-1} = T_U \phi_c^{-1}$ where $U = \phi_c \circ \gamma$ and $T_{U(t)}\phi_c^{-1} : T_{U(t)}M \to T_cM$ for every t. We obtain

$$\frac{d}{dt}\gamma(t) = T_{U(t)}\phi_c^{-1}\Omega\ \left(T_{U(t)}\phi_c^{-1}\right)^{\mathrm{T}} \left(T_{\gamma(t)}\phi_c\right)^{\mathrm{T}} \mathrm{grad}\, H\left(\phi_c(\gamma(t))\right). \tag{4.2}$$

Considering the Hamiltonian $\widetilde{H} : T_cM \to \mathbb{R}$, $\widetilde{H}(\gamma) := \phi_c^* H(\gamma) = H \circ \phi_c(\gamma)$, we can then rewrite (4.1) in the form

$$\frac{d}{dt}\gamma(t) = \widetilde{\Omega}(\gamma)\ \mathrm{grad}\, \widetilde{H}(\gamma), \quad \widetilde{\Omega}(\gamma) := T_{U(t)}\phi_c^{-1}\Omega\ \left(T_{U(t)}\phi_c^{-1}\right)^{\mathrm{T}}, \tag{4.3}$$

[2] The local error results of this section are valid for general retractions. For the special choice $\phi_c = \exp_c$, an analysis in a purely Riemannian setting could provide sharper geometric insight into the properties of the error.

where we have used that $\operatorname{grad} \widetilde{H} = T_{\gamma(t)}\phi_c^{\mathrm{T}} \operatorname{grad} H\left(\phi_c(\gamma(t))\right)$, which is now a gradient on the linear space T_cM with respect to the metric inherited from M, $g|_c$. Locally in a neighborhood of c, (3.2)-(3.4) applied to (4.3) coincides with the methods of Cohen and Hairer, and therefore the order result [6, Thm 4.3] can be applied. Since the Riemannian distance $d(\cdot,\cdot)$ and any norm in local coordinates are equivalent, the result follows. □

4.2. **Global error.** We prove the following result for the global error of DRG methods.

Theorem 2. *Let $u(t)$ be the exact solution to* (2.1) *where F is a complete vector field on a connected Riemannian manifold (M, g) with flow $u(t) = \varphi_t(u^0)$. Let ψ_h represent a numerical method $u^{k+1} = \psi_h(u^k)$ whose local error can be bounded for some $p \in \mathbb{N}$ as*

$$d(\psi_h(u), \varphi_h(u)) \le Ch^{p+1} \quad \text{for all } u \in M.$$

Suppose there is a constant L such that

$$\|\nabla F\|_g \le L,$$

where ∇ is the Levi-Civita connection and $\|\cdot\|_g$ is the operator norm with respect to the metric g. Then the global error is bounded as

$$d\left(u(kh), u^k\right) \le \frac{C}{L}(\mathrm{e}^{khL} - 1)h^p \quad \text{for all } k > 0.$$

Proof. Denoting the global error as $e^k := d(u(kh), u^k)$, the triangle inequality yields

$$e^{k+1} \le d\left(\varphi_h(u(kh)), \varphi_h(u^k)\right) + d\left(\varphi_h(u^k), \psi_h(u^k)\right).$$

The first term is the error at nh propagated over one step, the second term is the local error. For the first term, we find via a Grönwall-type inequality of [16],

$$d\left(\varphi_h(u(kh)), \varphi_h(u^k)\right) \le \mathrm{e}^{hL} d\left(u(kh), u^k\right) = \mathrm{e}^{hL} e^k.$$

Using the local error estimate for the second term, we get the recursion

$$e^{k+1} \le \mathrm{e}^{hL} e^k + Ch^{p+1},$$

which yields

$$e^k \le C \frac{\mathrm{e}^{khL} - 1}{\mathrm{e}^{hL} - 1} h^{p+1} \le \frac{C}{L}(\mathrm{e}^{khL} - 1)h^p.$$

□

Remark. Following Theorem 1.4 in [16], the condition that F is complete can be relaxed if $\varphi_t(u^0)$ and $\{u^k\}_{k\in\mathbb{N}}$ lie in a relatively compact submanifold N of M containing all the geodesics from u^k to $\varphi_{kh}(u^0)$. This is the case if, for instance, H has compact, geodesically convex sublevel sets, since both $\varphi_t(u^0)$ and $\{u^k\}_{k\in\mathbb{N}}$ are restricted to the level set $M_{H(u^0)} = \{p \in M \mid H(p) = H(u^0)\}$ and hence lie in the sublevel set $N_{H(u^0)} = \{p \in M \mid H(p) \le H(u^0)\}$.

5. Examples and numerical results

To demonstrate how to construct schemes of the type presented, we consider first an example on the two-sphere. The AVF DRG and Midpoint DRG schemes presented in this example could also be obtained by the discrete differential methods on homogeneous manifolds presented in [5]. The novel schemes here are the Itoh–Abe DRG scheme and its symmetrized variant, and the higher order methods obtained by composition or collocation techniques. Then, to demonstrate the usefulness of our methods for problems on more challenging manifolds, we consider first the motion of a particle under gravity on a paraboloid, and then a conservative system on the Stiefel manifold.

5.1. **Example 1: Perturbed spinning top.** We consider a non-linear perturbation of a spinning top; see [22]. This is a body whose orientation is represented by a vector s of unit length in $\mathbb{R}^3$, so that s lies on the manifold $M = S^2 = \left\{s \in \mathbb{R}^3 : \|s\| = 1\right\}$. Here and in what follows, $\|\cdot\|$ denotes the two-norm. The ODE system can be written in the form

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \Omega(s)\ \mathrm{grad}\, H(s), \quad s \in S^2, \quad H \in \mathcal{F}\left(S^2\right), \tag{5.1}$$

where $\Omega(s)y = s \times y$. Given the inertia tensor $\mathbb{I} = \mathrm{diag}(\mathbb{I}_1, \mathbb{I}_2, \mathbb{I}_3)$, and denoting by s^2 the component-wise square of s, we consider the Hamiltonian

$$H(s) = \frac{1}{2}(\mathbb{I}^{-1}s)^{\mathrm{T}}(s + \frac{2}{3}s^2).$$

Geometric integrators for spin systems are discussed widely in the literature; see, e.g., [9, 20–22] and the references therein. The two-sphere has a simple geometry which makes it attractive for illustrating our new schemes while at the same time being different from Euclidean space, where the standard discrete gradient schemes can be used.

The Riemannian metric g on S^2 restricts to the so-called round metric, coinciding with the Euclidean inner product on the tangent plane of the sphere. Our choice of retraction ϕ is as in [5], given by its restriction to p,

$$\phi_p(x) = \frac{p + x}{\|p + x\|}, \tag{5.2}$$

with the inverse

$$\phi_p^{-1}(u) = \frac{u}{p^{\mathrm{T}}u} - p$$

defined when $p^{\mathrm{T}}u > 0$. We note that $p^{\mathrm{T}}x = 0$ for all $x \in T_pS^2$. The derivative of the retraction and its inverse are given by

$$D_x\phi_p = \frac{1}{\|p + x\|}\left(I - \frac{(p + x) \otimes (p + x)}{\|p + x\|^2}\right), \qquad D_u\phi_p^{-1} = \frac{1}{p^{\mathrm{T}}u}\left(I - \frac{u \otimes p}{p^{\mathrm{T}}u}\right), \tag{5.3}$$

where $\otimes$ denotes the outer product[3] of the vectors.

We approximate the system (5.1) numerically, testing the scheme (2.6)-(2.7) with different discrete Riemannian gradients: the AVF (2.8), the midpoint (2.9), the Itoh–Abe (2.10), and its symmetrized version (3.6). For the three symmetric

[3]If x and y are in $\mathbb{R}^3$, $x \otimes y$ is the matrix-matrix product of x taken as a 3×1 matrix and y taken as a 1×3 matrix.

methods, we have chosen $c(s,\widetilde{s}) = \frac{s+\widetilde{s}}{\|s+\widetilde{s}\|}$, so that $\phi_c^{-1}(\widetilde{s}) = -\phi_c^{-1}(s)$. Using that $\operatorname{grad} H(s) = \mathbb{I}^{-1}(s+s^2)$ and considering the transpose of $T_{\gamma_\xi}\phi_c$ from (5.3), the AVF DRG becomes

$$\begin{aligned}\overline{\operatorname{grad}}_{\mathrm{AVF}} H(s,\widetilde{s}) &= \int_0^1 \frac{1}{\|l_\xi\|}\left(I - \frac{l_\xi \otimes l_\xi}{\|l_\xi\|^2}\right)\mathbb{I}^{-1}(\phi_c(\gamma_\xi)+\phi_c(\gamma_\xi)^2)\,\mathrm{d}\xi\\ &= \int_0^1 \frac{1}{\|l_\xi\|}\Big(\mathbb{I}^{-1}\left(\phi_c(\gamma_\xi)+\phi_c(\gamma_\xi)^2\right)\\ &\quad -\phi_c(\gamma_\xi)^{\mathrm{T}}\mathbb{I}^{-1}\left(\phi_c(\gamma_\xi)+\phi_c(\gamma_\xi)^2\right)\phi_c(\gamma_\xi)\Big)\,\mathrm{d}\xi,\end{aligned}$$

with $\gamma_\xi = (1-\xi)\phi_c^{-1}(s) + \xi\phi_c^{-1}(\widetilde{s}) = (1-2\xi)\phi_c^{-1}(s)$ and $l_\xi = c+\gamma_\xi$. Similarly, the midpoint DRG becomes

$$\begin{aligned}\overline{\operatorname{grad}}_{\mathrm{MP}} H(s,\widetilde{s}) =& \frac{1}{\|s+\widetilde{s}\|}\Bigg(\mathbb{I}^{-1}\left(s+\widetilde{s}+\frac{2}{3}\left(s^2+s\widetilde{s}+\widetilde{s}^2\right)\right)\\ &+\frac{\frac{1}{2}\|s+\widetilde{s}\|^2-2}{\|\widetilde{s}-s\|^2}\left(H(\widetilde{s})-H(s)\right)(\widetilde{s}-s)\Bigg),\end{aligned}$$

where we have used that $g(s,s) = s^{\mathrm{T}}s = 1$ for all $s \in S^2$. To obtain the basis of T_cM for the definition of the Itoh–Abe DRG, we have used the singular-value decomposition. For the first order scheme, noting that $\phi_s^{-1}(s) = 0$, we choose $c(s,\widetilde{s}) = s$, and get $\alpha_j = \phi_s^{-1}(\widetilde{s})^{\mathrm{T}}E_j$ for $j = 1,2$. Then the DRG (2.10) can be written as

$$\begin{aligned}\overline{\operatorname{grad}}_{\mathrm{IA}} H(s,\widetilde{s}) =& \frac{H\left(\phi_s\left(\phi_s^{-1}(\widetilde{s})^{\mathrm{T}}E_1E_1\right)\right)-H(s)}{\phi_s^{-1}(\widetilde{s})^{\mathrm{T}}E_1}E_1\\ &+\frac{H(\widetilde{s})-H\left(\phi_s\left(\phi_s^{-1}(\widetilde{s})^{\mathrm{T}}E_1E_1\right)\right)}{\phi_s^{-1}(\widetilde{s})^{\mathrm{T}}E_2}E_2.\end{aligned} \tag{5.4}$$

We solve the same problem using the fourth, sixth, and eighth order variants of the collocation-like scheme (3.2)-(3.4). Choosing in the fourth order case the Gaussian nodes $\mathsf{c}_{1,2} = \frac{1}{2} \mp \frac{\sqrt{3}}{6}$ as collocation points and setting $c(s,\widetilde{s}) = s$, we get the non-linear system

$$\begin{aligned}S_1 &= h\,\phi_{s_0}\left(\frac{1}{2}T_{S_1}\phi_{s_0}^{-1}\left(\Omega_1 \operatorname{grad}_1 H\right) + \left(\frac{1}{2}-\frac{\sqrt{3}}{3}\right)T_{S_2}\phi_{s_0}^{-1}\left(\Omega_2 \operatorname{grad}_2 H\right)\right),\\ S_2 &= h\,\phi_{s_0}\left(\left(\frac{1}{2}+\frac{\sqrt{3}}{3}\right)T_{S_1}\phi_{s_0}^{-1}\left(\Omega_1 \operatorname{grad}_1 H\right) + \frac{1}{2}T_{S_2}\phi_{s_0}^{-1}\left(\Omega_2 \operatorname{grad}_2 H\right)\right),\\ s_1 &= h\,\phi_{s_0}\left(T_{S_1}\phi_{s_0}^{-1}\left(\Omega_1 \operatorname{grad}_1 H\right) + T_{S_2}\phi_{s_0}^{-1}\left(\Omega_2 \operatorname{grad}_2 H\right)\right),\end{aligned}$$

where

$$\begin{aligned}\sigma(\xi h) =& \left(\left(3+2\sqrt{3}\right)\phi_{s_0}^{-1}(S_1) + \left(3-2\sqrt{3}\right)\phi_{s_0}^{-1}(S_2)\right)\xi\\ &+\left(3\left(\sqrt{3}-1\right)\phi_{s_0}^{-1}(S_2) - 3\left(1+\sqrt{3}\right)\phi_{s_0}^{-1}(S_1)\right)\xi^2\end{aligned}$$

and we use the transposes of (5.3) and $\operatorname{grad} H(s) = \mathbb{I}^{-1}(s+s^2)$ in the evaluation of $\operatorname{grad}_1 H$ and $\operatorname{grad}_2 H$. The sixth and eighth order schemes are derived in a similar manner, using the standard Gaussian nodes.

A second order scheme is derived by composing the Itoh–Abe DRG method with its adjoint, and a fourth order scheme is obtained by composing this method

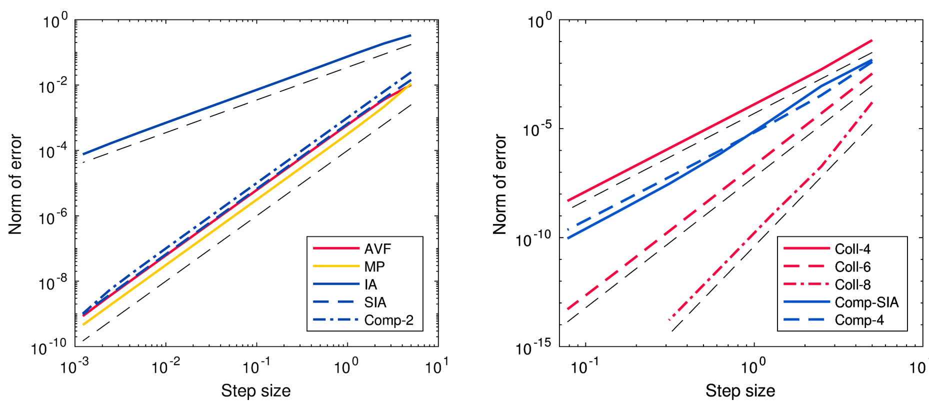

FIGURE 1. Error norm at $t = 10$ for the perturbed spinning top problem solved with different schemes, plotted with black, dashed reference lines of order 1, 2, 4, 6, and 8. Initial condition $s = (-1, -1, 1)/\sqrt{3}$ and $\mathbb{I} = \operatorname{diag}(1, 2, 4)$. *Left:* The AVF, midpoint (MP), Itoh–Abe (IA) and symmetrized Itoh–Abe (SIA) DRGs and a 3-stage composition of the IA DRG scheme (Comp-2). *Right:* Collocation-type schemes of order 4, 6, and 8, a 3-stage composition of the SIA DRG scheme (Comp-SIA), and a 6-stage composition of the IA DRG scheme (Comp-4).

again with itself, as well as one by composition of the symmetrized Itoh–Abe DRG method with itself. In all stages of these composition methods, a symmetric $c(u, v)$ is used.

Plots confirming the order of all methods can be seen in Figure 1, where solutions using the different schemes are compared to a reference solution obtained using a comparatively small step size. See the left hand panel of Figure 2 for numerical confirmation that our methods do indeed preserve the energy to machine precision, while the implicit midpoint method does not. In the right hand panel of Figure 2, the solution obtained by the Itoh–Abe DRG scheme with a step size $h = 1$ is plotted together with a solution obtained using the symmetrized Itoh–Abe DRG method with a much smaller time step. We observe, as expected for a method that conserves both the energy and the angular momentum, that the solution stays on the trajectory of the exact solution, although not necessarily at the right place on the trajectory at any given time.

5.2. **Example 2: Particle moving under gravity on a paraboloid.** We consider a particle of unit mass moving under a gravitational field on an elliptic paraboloid given by

$$P = \left\{ q \in \mathbb{R}^3 : \frac{x(q)^2}{a^2} + \frac{y(q)^2}{b^2} - 2z(q) = d \right\},$$

where x, y, z are the Cartesian coordinate functions, and $a, b, d \in \mathbb{R}^+$. See [29, pp. 106-108] and [11] for a discussion of such a system and the existence of solutions. Using, as in [18], an inertial Euclidean frame to express the position of the particle

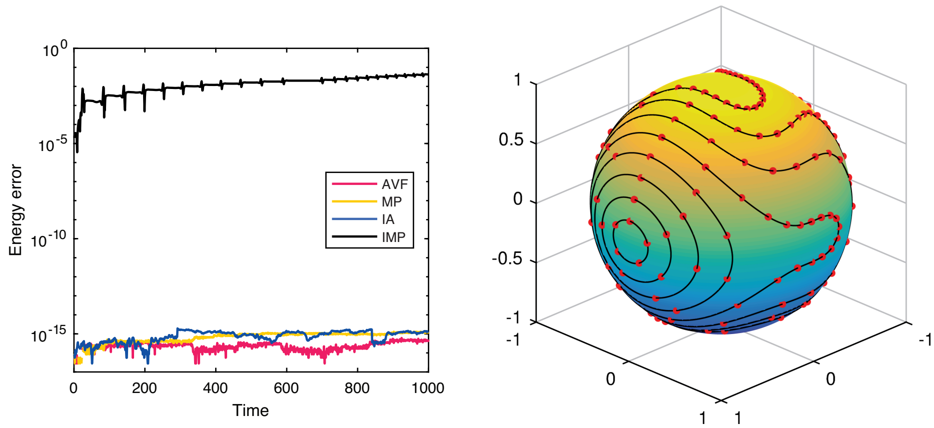

FIGURE 2. *Left:* Energy error with increasing time for the AVF, midpoint (MP) and Itoh–Abe (IA) DRG methods, as well as the implicit midpoint (IMP) method, with step size $h = 1$, initial condition $s = (-1, -1, 1)/\sqrt{3}$ and $\mathbb{I} = \mathrm{diag}(1, 2, 4)$. *Right:* Curves of constant energy on the sphere, found by our method with different starting values. The black solid line is the solution using the symmetrized Itoh–Abe DRG method with step size $h = 0.01$, while the red dots are the solutions obtained by the Itoh–Abe DRG method with step size $h = 1$.

in $\mathbb{R}^3$, we obtain the Hamiltonian

$$H(q,p) = \frac{1}{2}p^T p + g q_3,$$

where $p = \frac{\partial L}{\partial \dot{q}} = \dot{q}$ are the momentum coordinates and g is the gravitational constant. Thus the dynamics can be described on the cotangent bundle $T^*P =: M$ by the Hamiltonian equations, for $i = 1, 2, 3$,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}(q,p), \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}(q,p), \qquad q \in P, \quad p \in T_q P.$$

We define the retraction by its restriction to a center point $c = (c_P, c_T)$, $c_P \in P$, $c_T \in T_{c_P}P$, so that $\phi : TM \to M$ is given by $\phi_c(u,v) = (\phi_{P,c}(u), \phi_{T,c,u}(v))$, where $\phi_{P,c} : T_{c_P}P \to P$ and $\phi_{T,c,u} : T_{c_T}T_{c_P}P \to T_{\phi_{P,c}(u)}P$. Our choice of $\phi_{P,c}(u)$ is the projection onto the paraboloid P along the straight line in $\mathbb{R}^3$ from $c_P + u$ to the origin. The second component $\phi_{T,c}(u,v)$ is the linear projection in $\mathbb{R}^3$ of $c_T + v$ to $T_{c_P}P$. That is,

$$\phi_{P,c}(u) = \frac{d}{\alpha - c_{P,3} - u_3}(c_P + u),$$

$$\phi_{T,c,u}(v) = c_T + v - \frac{\beta(\phi_{P,c}(u))^T(c_T + v)}{\beta(\phi_{P,c}(u))^T\beta(\phi_{P,c}(u))}\beta(\phi_{P,c}(u)),$$

where

$$\alpha = \sqrt{(c_{P,3} + u_3)^2 + d\left(\frac{(c_{P,1} + u_1)^2}{a^2} + \frac{(c_{P,2} + u_2)^2}{b^2}\right)}, \quad \beta(q) = \left(\frac{q_1}{a^2}, \frac{q_2}{b^2}, -1\right)^T.$$

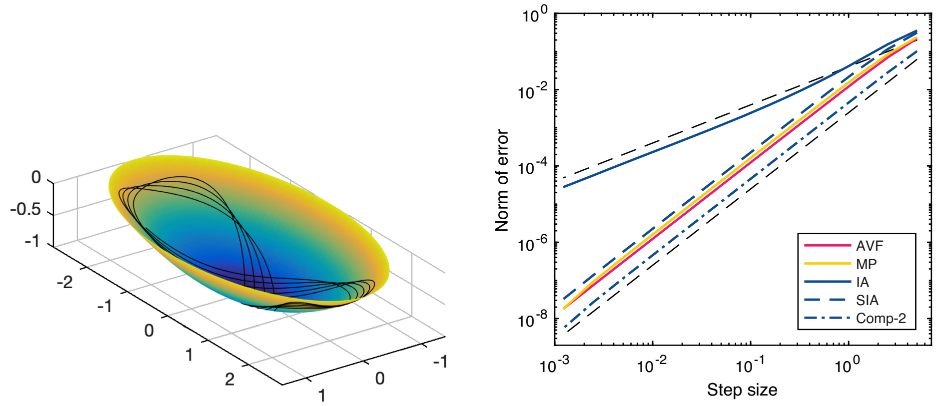

FIGURE 3. Particle moving under gravity on the paraboloid. *Left:* The system solved by the symmetrized Itoh–Abe DRG method with step size $h = 0.01$, from $t_0 = 0$ to $T = 20$. *Right:* Error norm at $t = 1$ for the problem solved with different schemes: The AVF, midpoint (MP), Itoh–Abe (IA) and symmetrized Itoh–Abe (SIA) DRG methods and a 3-stage composition of the IA DRG scheme (Comp-2), plotted against black, dashed reference lines of order 1 and 2.

This has the inverse $\phi_c^{-1}(q,p) = (\phi_{P,c}^{-1}(q), \phi_{T,c,q}^{-1}(p))$, where

$$\phi_{P,c}^{-1}(q) = \frac{c_{P,3}+d}{\beta(c_P)^T q} q - c_P, \qquad \phi_{T,c,q}^{-1}(p) = p - c_T - \frac{\beta(c_P)^T(p-c_T)}{\beta(c_P)^T\beta(q)}\beta(q).$$

We test our schemes on the problem with the paraboloid given by $a = 1$, $b = 2$, $d = 2$, and starting values $q = (1/2, 1/2, -27/32)$, $p = (3/2, 7/2, 19/16)$. We compare to the solution of standard methods in $\mathbb{R}^6$ with a comparatively small time step size to confirm numerically that the methods converge to the correct solution. The numerical results show that our schemes have the expected order; see Figure 3. Preservation of the Hamiltonian was also observed.

5.3. **Example 3: Conservative system on the Stiefel manifold.** Lastly we consider an ODE

$$\dot{Y} = S(Y)\,\mathrm{grad}\,H(Y), \quad Y \in \mathbb{V}_p(\mathbb{R}^n), \quad H \in \mathcal{F}\left(\mathbb{V}_p(\mathbb{R}^n)\right), \tag{5.5}$$

on the Stiefel manifold $M = \mathbb{V}_p(\mathbb{R}^n) = \left\{Y \in \mathbb{R}^{n\times p} : Y^TY = I_p\right\}$, i.e., the set of all $n \times p$ matrices whose p columns are orthonormal. The solution of (5.5) stays on the Stiefel manifold at all times if

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(Y^TY) &= Y^T\dot{Y} + \dot{Y}^TY \\ &= Y^TS(Y)\,\mathrm{grad}\,H(Y) + \mathrm{grad}\,H(Y)^TS(Y)^TY = 0,\end{aligned} \tag{5.6}$$

i.e., if $Y^TS(Y)\,\mathrm{grad}\,H(Y)$ is a skew-symmetric $p \times p$ matrix. We shall consider a problem on $\mathbb{V}_p(\mathbb{R}^n)$ with first integral

$$H(Y(t)) = H(Y(t_0)), \quad H(Y) = \mathrm{tr}(Y^TY^2), \tag{5.7}$$

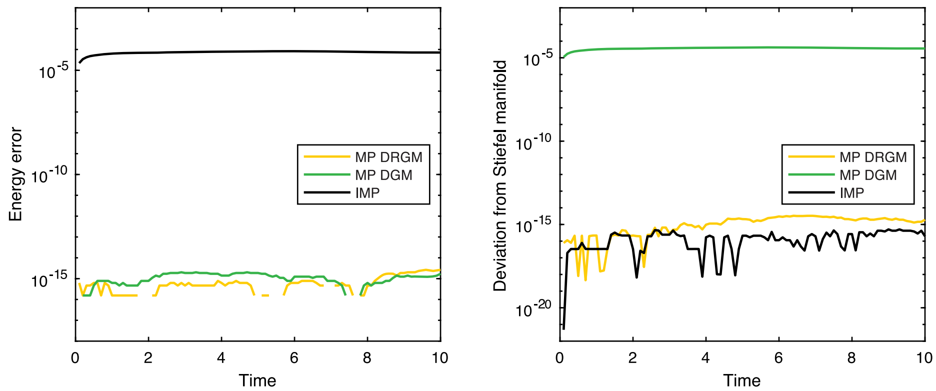

FIGURE 4. Relative energy error (left) and deviation from $\mathbb{V}_p(\mathbb{R}^n)$ (right) with increasing time for the midpoint discrete Riemannian gradient method (MP DRGM), as well as the standard midpoint discrete gradient method (MP DGM) and the implicit midpoint (IMP) method, with step size $h = 0.1$, $n = 5$, and $p = 2$. The relative energy error in step k is measured by $|(H(Y_k) - H(Y_0))/H(Y_0)|$, while the deviation from the Stiefel manifold is measured by $\|I_p - Y_k^T Y_k\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm.

where Y^2 means the component-wise square of Y. In (5.5), H is a first integral whenever $S(Y) \in \mathbb{R}^{n\times n}$ fulfills (5.6) as well as being skew-symmetric with respect to the Riemannian metric g,

$$g_Y(U, V) = \operatorname{tr}(U^T(I_n - \frac{1}{2}YY^T)V), \quad U, V \in T_Y M,$$

named the canonical metric by Edelman et al. in [8]. The Riemannian gradient of H follows from this; it is the tangent vector $\operatorname{grad} H$ satisfying $g_Y(\operatorname{grad} H, V) = \operatorname{tr}(\nabla H(Y)^T V)$ for all $V \in T_Y M$, where $\nabla H(Y)$ denotes the Euclidean gradient. That is, as stated in [8],

$$\operatorname{grad} H(Y) = \nabla H(Y) - Y\nabla H(Y)^T Y.$$

The intrinsic Riemannian structures provide the components needed in a DRG scheme. We choose the Riemannian exponential as retraction. That is, $\phi_C(V)$ is given by going the distance 1 along the geodesic path emanating from the base point $C \in M$ in the direction $V \in T_C M$. Similarly, the inverse retraction is given by the Riemannian logarithm, and the base point $C(Y, \tilde{Y})$ is given by the geodesic midpoint between Y and $\tilde{Y}$. To calculate the geodesic and the Riemannian exponential, we use the method introduced by Edelman et al. in [8, Corollary 2.2]. For the logarithm, we use the algorithm of Zimmermann [30].

Any numerical method preserving quadratic invariants, like symplectic Runge–Kutta methods, will find solutions on $\mathbb{V}_p(\mathbb{R}^n)$. However, such a method will in general not preserve the cubic invariant (5.7). A standard discrete gradient method can be implemented to preserve either (5.6) or (5.7), but not both. For our numerical experiments, we have considered (5.5) with $n = 5$ and $p = 2$, and $S(Y)$ chosen so that $H(Y)$ is conserved. As demonstrated in Figure 4, a DRG method can be used to get solutions that stay on the Stiefel manifold while preserving the first integral.

6. Conclusions and further work

We have presented a general framework for constructing energy-preserving numerical integrators on Riemannian manifolds. The main tool is to generalize the notion of discrete gradients as known from the literature. The new methods make use of an approximation to the Riemannian gradient coined the discrete Riemannian gradient, as well as a retraction map and a coordinate center function.

Particular examples of discrete Riemannian gradient methods are given as generalizations of well-known schemes, such as the average vector field method, the midpoint discrete gradient method, and the Itoh–Abe method. Extensions to higher order are proposed via a collocation-like method. We have analyzed the local and global error behavior of the methods, and they have been implemented and tested for problems on the two-sphere, the paraboloid, and the Stiefel manifold.

Possible directions for future research include a more detailed study of the stability and propagation of errors, taking into account particular features of the Riemannian manifold; for instance, it may be expected that the sectional curvature will play an important role. We believe, inspired by [3], that there is a potential for making our implementations more efficient by tailoring them to the particular manifold, as well as the ODE problem considered.

References

[1] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian Manifolds and a geometric model for the human spine*, IMA J. Numer. Anal. **22** (2002), no. 3, 359–390, DOI 10.1093/imanum/22.3.359. MR1918655

[2] L. Brugnano, F. Iavernaro, and D. Trigiante, *Hamiltonian boundary value methods (energy preserving discrete line integral methods)*, JNAIAM. J. Numer. Anal. Ind. Appl. Math. **5** (2010), no. 1-2, 17–37. MR2833606

[3] E. Celledoni, S. Eidnes, B. Owren, and T. Ringholm, *Dissipative numerical schemes on Riemannian manifolds with applications to gradient flows*, SIAM J. Sci. Comput. **40** (2018), no. 6, A3789–A3806, DOI 10.1137/18M1190628. MR3876868

[4] E. Celledoni, H. Marthinsen, and B. Owren, *An introduction to Lie group integrators—basics, new developments and applications*. part B, J. Comput. Phys. **257** (2014), no. part B, 1040–1061, DOI 10.1016/j.jcp.2012.12.031. MR3133432

[5] E. Celledoni and B. Owren, *Preserving first integrals with symmetric Lie group methods*, Discrete Contin. Dyn. Syst. **34** (2014), no. 3, 977–990, DOI 10.3934/dcds.2014.34.977. MR3094556

[6] D. Cohen and E. Hairer, *Linear energy-preserving integrators for Poisson systems*, BIT **51** (2011), no. 1, 91–101, DOI 10.1007/s10543-011-0310-z. MR2784654

[7] P. E. Crouch and R. Grossman, *Numerical integration of ordinary differential equations on manifolds*, J. Nonlinear Sci. **3** (1993), no. 1, 1–33, DOI 10.1007/BF02429858. MR1216986

[8] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl. **20** (1999), no. 2, 303–353, DOI 10.1137/S0895479895290954. MR1646856

[9] J. Frank, W. Huang, and B. Leimkuhler, *Geometric integrators for classical spin systems*, J. Comput. Phys. **133** (1997), no. 1, 160–172, DOI 10.1006/jcph.1997.5672. MR1445176

[10] O. Gonzalez, *Time integration and discrete Hamiltonian systems*, J. Nonlinear Sci. **6** (1996), no. 5, 449–467, DOI 10.1007/s003329900018. MR1411343

[11] A. Gray, A. Jones, and R. Rimmer, *Motion under gravity on a paraboloid*, J. Differential Equations **45** (1982), no. 2, 168–181, DOI 10.1016/0022-0396(82)90063-8. MR665994

[12] E. Hairer, *Energy-preserving variant of collocation methods*, JNAIAM. J. Numer. Anal. Ind. Appl. Math. **5** (2010), no. 1-2, 73–84. MR2833602

[13] A. Harten, P. D. Lax, and B. van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev. **25** (1983), no. 1, 35–61, DOI 10.1137/1025002. MR693713

[14] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta numerica, 2000, Acta Numer., vol. 9, Cambridge Univ. Press, Cambridge, 2000, pp. 215–365, DOI 10.1017/S0962492900002154. MR1883629

[15] T. Itoh and K. Abe, *Hamiltonian-conserving discrete canonical equations based on variational difference quotients*, J. Comput. Phys. **76** (1988), no. 1, 85–102, DOI 10.1016/0021-9991(88)90132-5. MR943488

[16] M. Kunzinger, H. Schichl, R. Steinbauer, and J. A. Vickers, *Global Gronwall estimates for integral curves on Riemannian manifolds*, Rev. Mat. Complut. **19** (2006), no. 1, 133–137, DOI 10.5209/rev_REMA.2006.v19.n1.16639. MR2219824

[17] J. M. Lee, *Riemannian Manifolds*, Graduate Texts in Mathematics, vol. 176, Springer-Verlag, New York, 1997. An introduction to curvature. MR1468735

[18] T. Lee, M. Leok, and N. H. McClamroch, *Global Formulations of Lagrangian and Hamiltonian Dynamics on Manifolds*, Interaction of Mechanics and Mathematics, Springer, Cham, 2018. A geometric approach to modeling and analysis. MR3699476

[19] B. Leimkuhler and G. W. Patrick, *A symplectic integrator for Riemannian manifolds*, J. Nonlinear Sci. **6** (1996), no. 4, 367–384, DOI 10.1007/s003329900015. MR1404143

[20] D. Lewis and N. Nigam, *Geometric integration on spheres and some interesting applications*, J. Comput. Appl. Math. **151** (2003), no. 1, 141–170, DOI 10.1016/S0377-0427(02)00743-4. MR1950234

[21] R. McLachlan, K. Modin, and O. Verdier, *A minimal-variable symplectic integrator on spheres*, Math. Comp. **86** (2017), no. 307, 2325–2344, DOI 10.1090/mcom/3153. MR3647960

[22] R. I. McLachlan, K. Modin, and O. Verdier, *Symplectic integrators for spin systems*, Phys. Rev. E **89** (2014), no. 6, 061301.

[23] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux, *Geometric integration using discrete gradients*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357** (1999), no. 1754, 1021–1045, DOI 10.1098/rsta.1999.0363. MR1694701

[24] J. Milnor, *Morse Theory*, Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51, Princeton University Press, Princeton, N.J., 1963. MR0163331

[25] H. Z. Munthe-Kaas and A. Lundervold, *On post-Lie algebras, Lie-Butcher series and moving frames*, Found. Comput. Math. **13** (2013), no. 4, 583–613, DOI 10.1007/s10208-013-9167-7. MR3085679

[26] H. Munthe-Kaas, *Lie-Butcher theory for Runge-Kutta methods*, BIT **35** (1995), no. 4, 572–587, DOI 10.1007/BF01739828. MR1431350

[27] H. Munthe-Kaas, *Runge-Kutta methods on Lie groups*, BIT **38** (1998), no. 1, 92–111, DOI 10.1007/BF02510919. MR1621080

[28] G. R. W. Quispel and D. I. McLaren, *A new class of energy-preserving numerical integration methods*, J. Phys. A **41** (2008), no. 4, 045206, 7, DOI 10.1088/1751-8113/41/4/045206. MR2451073

[29] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1988. With an introduction to the problem of three bodies; Reprint of the 1937 edition; With a foreword by William McCrea. MR992404

[30] R. Zimmermann, *A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric*, SIAM J. Matrix Anal. Appl. **38** (2017), no. 2, 322–342, DOI 10.1137/16M1074485. MR3639577

Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim, Norway
Email address: elena.celledoni@ntnu.no

Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim, Norway
Email address: solve.eidnes@ntnu.no

Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim, Norway
Email address: brynjulf.owren@ntnu.no

Department of Mathematical Sciences, Norwegian University of Science and Technology, N–7491 Trondheim, Norway
Email address: ringholm@gmail.com