Applied Mathematics & Optimization
https://doi.org/10.1007/s00245-019-09605-x

Two Structure-Preserving Time Discretizations for Gradient Flows

Ansgar Jüngel[1] · Ulisse Stefanelli[2,3] · Lara Trussardi[2]

Abstract

The equality between dissipation and energy drop is a structural property of gradient-flow dynamics. The classical implicit Euler scheme fails to reproduce this equality at the discrete level. We discuss two modifications of the Euler scheme satisfying an exact energy equality at the discrete level. Existence of discrete solutions and their convergence as the fineness of the partition goes to zero are discussed. Eventually, we address extensions to generalized gradient flows, GENERIC flows, and curves of maximal slope in metric spaces.

Keywords Gradient flow · Structure-preserving time discretization · GENERIC flows · Curves of maximal slope

Mathematics Subject Classification 35K55 · 58D25

1 Introduction

Gradient flows are the paradigm of dissipative evolution and arise ubiquitously in applications [1]. In abstract terms, they are formulated as

✉ Ulisse Stefanelli
ulisse.stefanelli@univie.ac.at
http://www.mat.univie.ac.at/∼stefanelli

Ansgar Jüngel
juengel@tuwien.ac.at
http://www.asc.tuwien.ac.at/∼juengel

Lara Trussardi
lara.trussardi@univie.ac.at
http://www.mat.univie.ac.at/∼trussardi

1 Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstraße 8-10, 1040 Wien, Austria

2 Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

3 Istituto di Matematica Applicata e Tecnologie Informatiche E. Magenes, v. Ferrata 1, 27100 Pavia, Italy

Published online: 11 September 2019

$$u'(t) + \mathrm{D}\phi(u(t)) = 0 \quad \text{for a.e. } t \in (0,T), \quad u(0) = u_0, \tag{1.1}$$

where the solution $t \mapsto u(t) \in H$ is a trajectory in a Hilbert space H, the potential ϕ is given, and u_0 is a prescribed initial datum. We assume ϕ to be smooth for the purpose of this introductory discussion, so that $\mathrm{D}\phi$ is here the Fréchet differential. Let us however anticipate that the analysis will encompass nonsmooth situations as well.

Problem (1.1) arises in a variety of applications including heat conduction, the Stefan problem, the Hele-Shaw cell, porous media, parabolic variational inequalities, some classes of ODEs with obstacles, degenerate parabolic PDEs, and the mean curvature flow for Cartesian graphs, among many others [37]. Correspondingly, (1.1) has attracted constant attention during the last half century, starting from the seminal contributions by Kōmura [28], Crandall- Pazy [14], and Brézis [9,10].

Solutions of (1.1) fulfill the *energy equality*

$$\phi(u(t)) + \int_s^t \|u'(r)\|^2 \mathrm{d}r = \phi(u(s)) \tag{1.2}$$

for all $[s,t] \subset [0,T]$. This equality expresses the fact that the energy drop $\phi(u(s)) - \phi(u(t))$ along the trajectory in the time interval $[s,t]$ *exactly* corresponds to the squared L^2 norm of the velocity u'. Here and in the following, $\|\cdot\|$ denotes the norm in H. For all u solving (1.1), the energy equality (1.2) can be rewritten as

$$\phi(u(t)) + \frac{1}{2}\int_s^t \|u'(r)\|^2 \mathrm{d}r + \frac{1}{2}\int_s^t \|\mathrm{D}\phi(u(r))\|^2 \mathrm{d}r = \phi(u(s)) \tag{1.3}$$

for all $[s,t] \subset [0,T]$. In particular, (1.2) and (1.3) are equivalent for solutions to (1.1).

If u does not solve (1.1), conditions (1.2) and (1.3) are not equivalent anymore. On the one hand, relation (1.3) and $u(0) = u_0$ imply that

$$\begin{aligned}
\frac{1}{2}\int_0^T \|u' + \mathrm{D}\phi(u)\|^2 \mathrm{d}r &= \frac{1}{2}\int_0^T \left(\|u'\|^2 + 2(\mathrm{D}\phi(u), u') + \|\mathrm{D}\phi(u)\|^2\right) \mathrm{d}r \\
&= \frac{1}{2}\int_0^T \left(\|u'\|^2 + 2(\phi \circ u)' + \|\mathrm{D}\phi(u)\|^2\right) \mathrm{d}r \\
&= \phi(u(T)) - \phi(u(0)) + \frac{1}{2}\int_0^T \|u'\|^2 \mathrm{d}r + \frac{1}{2}\int_0^T \|\mathrm{D}\phi(u)\|^2 \mathrm{d}r = 0,
\end{aligned}$$

so that u solves (1.1) and hence fulfills (1.2) as well.

On the other hand, (1.2) does not imply (1.1). In order to check this, take $\phi(u) = u_1 + u_2$ for all $u = (u_1, u_2) \in \mathbb{R}^2 = H$ and letting $u(t) = (-t, 0)$ for all $t \geq 0$, we see that u fulfills (1.2) but does not solve (1.1). All in all, we have verified that

$$(1.1) \;\Leftrightarrow\; (1.3) \;\overset{\Rightarrow}{\not\Leftarrow}\; (1.2).$$

The analysis of Problem (1.1) often relies on a time discretization. Let a partition $\{0 = t_0 < t_1 < \cdots < t_N = T\}$ be given and indicate by $\tau_i = t_i - t_{i-1}$ its time step.

A classical discretization of (1.1) is the *implicit Euler scheme*. Given u_0, this reads as

$$\frac{u_i - u_{i-1}}{\tau_i} + \mathrm{D}\phi(u_i) = 0 \quad \text{for } i = 1, \dots, N,$$

which can be equivalently reformulated in variational terms as

$$u_i \in \operatorname*{arg\,min}_{u \in H} \left(\phi(u) + \frac{\tau_i}{2} \left\| \frac{u - u_{i-1}}{\tau_i} \right\|^2 \right) \quad \text{for } i = 1, \dots, N.$$

Minimality implies stability of the scheme. Under appropriate assumptions on ϕ, convergence follows as the fineness of the partition goes to zero.

Unfortunately, the Euler scheme fails to reproduce exact dissipation dynamics at the discrete level: The dissipation $\|(u_i - u_{i-1})/\tau_i\|^2$ of the i-th time step does not correspond to the energy drop $\phi(u_{i-1}) - \phi(u_i)$ in the same step. Indeed, energy-*dissipating* schemes for gradient systems were developed quite intensively in the literature. Examples include the convex splitting method, popularized in [17], algebraically stable Runge–Kutta methods [24], and the discrete variational derivative method [18]. The exact replication of the dissipation dynamics is nevertheless a desirable feature in view of developing structure-preserving algorithms.

The aim of this note is to analyze two variants of the Euler scheme reproducing at the discrete level the energy equalities (1.2) and (1.3), respectively. The first of such modifications can be traced back at least to Gonzalez [20], see also [13,40], and will hence be termed *Gonzalez scheme* here. Given u_0, for all $i = 1, \dots, N$, the Gonzalez scheme defines u_i starting from the previous value u_{i-1} by letting $u_i = u_{i-1}$ in case $\mathrm{D}\phi(u_{i-1}) = 0$ and solving

$$\frac{u_i - u_{i-1}}{\tau_i} + \mathrm{D}\phi(u_i) + \big(\phi(u_i) - \phi(u_{i-1}) - (\mathrm{D}\phi(u_i), u_i - u_{i-1})\big) \times \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2} = 0 \tag{1.4}$$

if $\mathrm{D}\phi(u_{i-1}) \neq 0$, where $(\cdot, \cdot)$ is the scalar product in H. Indeed, by taking the scalar product of (1.4) with $u_i - u_{i-1}$, the terms containing $\mathrm{D}\phi(u_i)$ cancel out and we are left with

$$\phi(u_i) + \tau_i \left\| \frac{u_i - u_{i-1}}{\tau_i} \right\|^2 = \phi(u_{i-1}).$$

By summing up the latter over i, one obtains a discrete version of the energy balance (1.2). This implies in particular the stability of the Gonzalez scheme. We remark that the original Gonzalez scheme uses $\mathrm{D}\phi((u_i + u_{i-1})/2)$ instead of $\mathrm{D}\phi(u_i)$ in (1.4), leading in both cases to an implicit scheme. To our knowledge, no analysis for the solvability of (1.4) nor for the convergence of the scheme has been presented so far. We fill this gap in Sect. 3, where we prove that for sufficiently smooth functions ϕ and finite-dimensional spaces H, the Gonzalez scheme admits solutions (Theorem 3.1) that converge with an explicit and sharp a priori rate (Theorem 3.2).

The second modification of the Euler scheme reproduces at the discrete level relation (1.3). The scheme is inspired by the approach to steepest-descent dynamics by De Giorgi [15], and we hence refer to it as *De Giorgi scheme* in the following. Given u_0, the De Giorgi scheme reads as

$$\phi(u_i) + \frac{\tau_i}{2}\left\|\frac{u_i - u_{i-1}}{\tau_i}\right\|^2 + \frac{\tau_i}{2}\|\mathrm{D}\phi(u_i)\|^2 - \phi(u_{i-1}) = 0 \tag{1.5}$$

for $i = 1, \ldots, N$. By summing up the latter relation over i, one readily obtains a discrete version of (1.3), which directly proves stability. The De Giorgi scheme reduces to a single scalar equation and allows for an existence (Theorem 4.1) and a convergence proof (Theorem 4.2), even in nonsmooth and infinite-dimensional situations, see Sect. 4. In addition, sharp error estimates can be derived in the finite-dimensional case.

Instead of solving (1.5), one could simply minimize its residual by minimizing the functional defined by the left-hand side. As a by-product of our analysis, we show that this alternative variational approach with respect to the Euler scheme can also be considered, giving convergence and, when restricted to finite dimensions, optimal error bounds.

Remarkably, the De Giorgi scheme can serve as an a posteriori tool to check the convergence of time-discrete approximations u_i, regardless of the method used to generate them. Indeed, the analysis of the (positive parts of the) residuals to (1.5) can reveal whether the approximation converges to the unique solution of (1.1), see condition (4.6) below.

Eventually, the De Giorgi scheme can be extended to other nonlinear evolution settings. In Sect. 5 we comment on its application to the case of generalized gradient flows, GENERIC flows, and curves of maximal slope in metric spaces.

Before moving on, let us briefly put our contribution in context. Numerical schemes conserving first integrals (in particular, the energy) can be found in [20,39,41,42]. See [44] for a recent reference and [12] for a review. In particular, the Gonzalez scheme falls within the general class of *discrete-gradient methods* along with the choice

$$\nabla_d\phi(u, v) := \big(\phi(u) - \phi(v) - (\mathrm{D}\phi(u), u - v)\big)\frac{u - v}{\|u - v\|^2}$$

for the *discrete gradient*, for $u \neq v$. These methods are specifically tailored to exactly reproduce dissipation dynamics. The discrete gradient is designed to satisfy a discrete chain rule, which in turn delivers numerical integrators replicating the dissipation property of the continuous system [44]. In the specific case of the Gonzalez scheme, such discrete chain rule reads (see [20] or [25, Formula (5.10)]) $\phi(u) - \phi(v) = \nabla_d\phi(u, v) \cdot (u - v)$. A systematic study of discrete gradient methods can be found in [20,33].

Discrete-gradient methods have been applied to a large class of equations, not necessarily of gradient-flow type. Examples are schemes for conservative partial differential equations fulfilling a discrete conservation law [32] and linearly-fitted numerical schemes for conservative wave equations [30]. The latter method also preserves the dissipation structure of dissipative wave equations. Moreover, the discrete gradient method was applied to subdifferentials in [5].

As the implicit Euler method is of first order only, one may look for higher-order variational schemes. Two second-order schemes, the variational implicit midpoint and the extrapolated variational implicit Euler schemes, are proposed in [29]. A variational BDF2 (two-step backward differentiation formula) method is analyzed in [31]. While second-order convergence is expected in a smooth Hilbert space case, the convergence of rate one-half is shown in the general metric setting. All the mentioned results do not replicate the exact energy dissipation dynamics of the gradient flow problem.

One may ask whether the Gonzalez and De Giorgi scheme, analyzed in this paper, can be extended to give higher-order convergence. While higher-order consistency and convergence have been proved for discrete gradient methods for ODEs and the discrete variational derivative method for PDEs [18, Section 1.4], we are not aware of higher-order generalizations of our schemes. Moreover, as the result of [31] shows, the lack of smoothness may decrease the optimal convergence order. For quadratic potentials ϕ, the G-stability by Dahlquist allows for energy-dissipating schemes, but requires to redefine the energy as a function of (u_i, u_{i-1}) instead of u_i alone [26]. General energy-dissipative Runge–Kutta schemes are studied in [27], but again, they do not replicate the exact energy dissipation dynamics of the problem.

Our primary focus is that of approximating the limiting continuous dynamics. Still, one has to mention that the discrete schemes discussed here may bear some interest in connection with optimization as well. Indeed, the unconstrained minimization of the potential ϕ can be tackled by considering iterative approximations. In the case of a convex ϕ, the easiest example in this class is the *proximal algorithm* $u_i = \text{prox}_{\tau_i\phi}(u_{i-1})$, where $\text{prox}_{\tau_i\phi} = (\text{id} + \tau_i\partial\phi)^{-1} : H \to H$. This corresponds to iterations of the implicit Euler method, starting from an initial guess u_0. If the set of minimum points $\operatorname{argmin}\phi$ is not empty and $\sum\tau_i = \infty$, the proximal algorithm weakly converges to a minimizer [38, Props. 6.1–2, pp. 99–100]

A variety of modifications of the proximal algorithm has been advanced in order to solve optimization problems with specific structure, or with higher efficiency [6,7]. A possibility is that of augmenting the input of the proximity operator $\text{prox}_{\tau_i\phi}$ by specific corrections. This is indeed the case of the Gonzales scheme, for it corresponds to the position

$$u_i = \text{prox}_{\tau_i\phi}\left(u_{i-1} - \tau_i\nabla_d\phi(u_i, u_{i-1})\right).$$

Note, however, that the small correction $\tau_i\nabla_d\phi(u_i, u_{i-1})$ depends on u_i as well, which ultimately requires an implicit updating step. A discussion on the relation between discrete and continuous gradient-flow dynamics in presence of higher-order corrections can be found in [45].

Starting from Nesterov's result [36], acceleration methods based on inertial gradient systems of the form

$$u''(t) + \alpha(t)u'(t) + \mathrm{D}\phi(u(t)) \ni 0$$

where α is a suitably chosen damping coefficient vanishing for $t \to \infty$, are currently attracting attention, see [3,4,8,11] among many others. One could consider extensions of the Gonzales and the De Giorgi schemes to the case of inertial gradient systems as well. In order to replicate the equality between dissipation and energy drop, now including the kinetic energy, one has to introduce a higher-order correction in the time-increment term as well. We shall develop these considerations elsewhere.

In order to make the connection with optimization algorithms complete, one would have to study the convergence of the discrete sequence u_i as $i \to \infty$. This would correspond to investigate the long-time discrete dynamics and eventually connecting it with the corresponding continuous counterpart. Although some of our results hold on the whole time semiline as well, such a long time analysis is presently beyond the scope of this paper. We present some evidence of the potential of these methods in a comparative discussion in Sect. 6 below.

2 Preliminaries

Let us introduce our notation and recall some basic results. Denote by H a real separable Hilbert space with scalar product $(\cdot, \cdot)$ and associated norm $\|\cdot\|$. Let $\phi : H \to (-\infty, \infty]$ be a proper and lower semicontinuous functional, not necessarily convex, and let $u_0 \in D(\phi) = \{u \in H \ : \ \phi(u) \neq \infty\}$ (*essential domain*). The *Fréchet subdifferential* $\partial\phi(u)$ of ϕ at $u \in D(\phi)$ is the set of points $\xi \in H$ such that

$$\liminf_{v \to u} \frac{\phi(v) - \phi(u) - (\xi, v - u)}{\|v - u\|} \geq 0.$$

The latter reduces to the classical gradient $\mathrm{D}\phi(u)$ in case ϕ is Fréchet differentiable at u and to the subdifferential in the sense of convex analysis if ϕ is convex. If $\phi = \phi_1 + \phi_2$, where ϕ_1 is convex, proper, and lower semicontinuous, and $\phi_2 \in C^1(H)$, we have $\partial\phi = \partial\phi_1 + \mathrm{D}\phi_2$. We denote by $D(\partial\phi)$ the essential domain $D(\partial\phi) = \{u \in H \ : \ \partial\phi(u) \neq \emptyset\}$ and remark that $\partial\phi(u)$ is either convex or empty.

The well-posedness of Problem (1.1) for convex functions ϕ is classical [10]. Nonconvexity can also be allowed, as long as ϕ has compact sublevels, see Rossi and Savaré [43] for some comprehensive existence theory. Crucial assumptions are the conditional strong-weak closeness of $\partial\phi$ and the validity of the chain rule. These hold, for instance, if ϕ is a $C^{1,1}$ perturbation of a convex functional, which includes the case of λ-convex functionals [1]. In addition, a suitable class of dominated, not $C^{1,1}$ perturbations can be considered as well [43]. For the sake of definiteness, we focus here on the finite-time case $t \in [0, T]$. Let us, however, mention that some of our results hold for gradient flows on the semiline $t \in [0, \infty)$ as well. We will comment on this issue along the text.

Existence results for gradient flows often take the moves from time discretizations. In the following, we will consider families of partitions

$$\{0 = t_0^n < t_1^n < \cdots < t_{N^n}^n\}$$

of the interval $[0, T]$, indicate their time steps by $\tau_i^n = t_i^n - t_{i-1}^n$, and denote their fineness by $\tau^n = \max_i \tau_i^n$. Correspondingly, one is interested in finding a time-discrete solution $u_i^n \in D(\partial\phi)$ with $u_0^n = u_0$ via some time-discrete scheme.

By defining the affine functions $\ell_j^n(t) = (t - t_{j-1}^n)/\tau_j^n$, for any given vector w_i^n, $i = 0, \ldots, N^n$, we use the notation $\widehat{w}^n, \overline{w}^n : [0, \infty) \to H$ for the piecewise affine and the backward constant interpolants of the values w_i^n on the partition, namely

$$\widehat{w}^n(0) = \overline{w}^n(0) = w_0^n, \quad \widehat{w}^n(t) = \ell_j^n(t) w_j^n + (1 - \ell_j^n(t)) w_{j-1}^n,$$
$$\overline{w}^n(t) = w_j^n \quad \text{for all } t \in (t_{j-1}^n, t_j^n].$$

By letting $\tau^n \to 0$ one is then asked to show the convergence of the time-discrete trajectory $\widehat{u}^n$ to a solution to (1.1).

A *caveat* on notation: in the following, we use the same symbol C to indicate a generic positive constant, possibly depending on the data and varying from line to line. Occasionally, we may explicitly indicate dependences of a constant by subscripts.

3 The Gonzalez Scheme

As mentioned in the introduction, the Gonzalez scheme can be traced back at least to [20] and has been already applied in various thermomechanical contexts [19,21,23,40–42]. We aim here at providing some theoretical analysis by focusing on solvability (Theorem 3.1) and convergence (Theorem 3.2). Before moving on, let us equivalently rewrite relation (1.4) as

$$\phi(u_i) + \tau_i \left\| \frac{u_i - u_{i-1}}{\tau_i} \right\|^2 = \phi(u_{i-1}) \text{ and } \mathrm{D}\phi(u_i) \text{ is parallel to } u_i - u_{i-1}. \quad (3.1)$$

Indeed, a solution to (1.4) fulfills (3.1) as well: The energy equality holds and one reads off the parallelism by comparing the terms in (1.4). On the contrary, if u_i solves (3.1), we compute

$$\begin{aligned}
0 &= \left(\phi(u_i) - \phi(u_{i-1}) + \tau_i \left\| \frac{u_i - u_{i-1}}{\tau_i} \right\|^2 \right) \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2} \\
&= \left(\phi(u_i) - \phi(u_{i-1}) + \tau_i \left\| \frac{u_i - u_{i-1}}{\tau_i} \right\|^2 \right) \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2} \\
&\quad + \frac{(\mathrm{D}\phi(u_i), u_i - u_{i-1})}{\|u_i - u_{i-1}\|} \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|} - \frac{(\mathrm{D}\phi(u_i), u_i - u_{i-1})}{\|u_i - u_{i-1}\|} \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|} \\
&= \left(\phi(u_i) - \phi(u_i) + \tau_i \left\| \frac{u_i - u_{i-1}}{\tau_i} \right\|^2 \right) \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2} \\
&\quad + \mathrm{D}\phi(u_i) - \frac{(\mathrm{D}\phi(u_i), u_i - u_{i-1})}{\|u_i - u_{i-1}\|} \frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|}
\end{aligned}$$

which is exactly (1.4). We will make use of this equivalent formulation to prove the following result.

Theorem 3.1 (Existence for the Gonzalez scheme) *Let $\phi \in C^1(\mathbb{R}^d)$ be bounded from below and let $\mathrm{D}\phi(u_{i-1}) \neq 0$. Then there exists $u_i \in \mathbb{R}^d \setminus \{u_{i-1}\}$ such that (1.4) holds.*

Proof Define $K = \{u \in \mathbb{R}^d \ : \ \phi(u) + \tau_i \|(u-u_{i-1})/\tau_i\|^2 = \phi(u_{i-1})\}$. Note that K is not empty as $u_{i-1} \in K$. Denote by f the continuous function $f(u) = \|u-u_{i-1}\|^2/\tau_i$. The set K is compact, for it is the preimage of $\{\phi(u_{i-1})\}$ with respect to the continuous and coercive function $g = \phi + f$. One can hence define u_i as the maximizer of f on K. The element u_i is surely different from u_{i-1} as K does not reduce to u_{i-1} because of $\mathrm{D}\phi(u_{i-1}) \neq 0$ and the implicit function theorem.

We now check that u_i solves (3.1). The energy equality follows directly from the fact that $u_i \in K$ and one is left to verify that $\mathrm{D}\phi(u_i)$ and $u_i - u_{i-1}$ are parallel. As $\mathrm{D}\phi(u_i) = \mathrm{D}g(u_i) - \mathrm{D}f(u_i)$, such parallelism follows once we prove that $\mathrm{D}g(u_i)$ and $\mathrm{D}f(u_i)$ are parallel, as we have $\mathrm{D}f(u_i) = 2(u_i - u_{i-1})/\tau_i$.

In case $\mathrm{D}g(u_i) = 0$, the parallelism trivially holds. If $\mathrm{D}g(u_i) \neq 0$, again the implicit function theorem ensures that K is a C^1 hypersurface in a neighborhood of u_i and that $\mathrm{D}g(u_i)$ is orthogonal to K at u_i. Assume by contradiction that $\mathrm{D}f(u_i)$ is not parallel to $\mathrm{D}g(u_i)$. Then the projection ν of $\mathrm{D}f(u_i)$ onto the tangent space $\mathrm{D}g(u_i)^\perp$ is nonzero. By letting $\gamma : [-\delta, \delta] \to K$ be a C^1 curve for some small $\delta > 0$ with $\gamma(0) = u_i$ and $\gamma'(0) = \nu$, we deduce that $f \circ \gamma$ is not maximized at 0, contradicting the maximality of u_i. □

By inspecting the proof of Theorem 3.1, one realizes that the lower bound on ϕ can be replaced by the weaker quadratic bound $\phi(u) \geq -C\|u\|^2 - C$ at the price of prescribing the time step τ_i to be small enough. Indeed, what is needed in the proof is just the coercivity of the function $u \mapsto \phi(u) + \tau_i\|(u-u_{i-1})/\tau_i\|^2$, which holds for sufficiently small $\tau_i > 0$ in view of

$$\phi(u) + \tau_i \left\| \frac{u - u_{i-1}}{\tau_i} \right\|^2 \geq \left(-C + \frac{1}{2\tau_i} \right) \|u\|^2 - \left(C + \frac{1}{\tau_i} \|u_{i-1}\|^2 \right).$$

Let a sequence of partitions $\{0 = t_0^n < t_1^n < \cdots < t_{N^n}^n = T\}$ be given with $\tau^n = \max_i \tau_i^n = \max_i (t_i^n - t_{i-1}^n) \to 0$ and correspondingly, let u_i^n be solutions of the Gonzalez scheme, whose existence is guaranteed by Theorem 3.1. We have the following convergence result.

Theorem 3.2 (Convergence for the Gonzalez scheme) *Let $\phi \in C^1(\mathbb{R}^d)$ be bounded from below.*

(i) *There exists a subsequence which is not relabeled such that $\widehat{u}^n \to u$ weakly in $H^1(0, T; \mathbb{R}^d)$ as $n \to \infty$, where u solves (1.1).*
(ii) *Let $\phi \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$. Then the whole sequence $(\widehat{u}^n)$ converges strongly in $W^{1,\infty}(0, T; \mathbb{R}^d)$ and we have the error bound*

$$\|u - \widehat{u}^n\|_{W^{1,\infty}(0,T;\mathbb{R}^d)} \leq C\tau^n. \tag{3.2}$$

(iii) *Let* $\phi \in C^3(\mathbb{R}^d)$ *and assume that the condition*

$$\mathrm{D}^2\phi(v)w \ \textit{ is parallel to } \ w \ \textit{ for any } v, \ w \in \mathbb{R}^d \tag{3.3}$$

holds. Then

$$\|u - \widehat{u}^n\|_{W^{1,\infty}(0,T;\mathbb{R}^d)} \leq C(\tau^n)^2. \tag{3.4}$$

Proof *Step 1: Proof of* (i). Let us start by verifying the stability of the Gonzalez scheme. Indeed, by adding up the local energy equality, we obtain

$$\phi(u_m^n) + \sum_{i=1}^{m} \tau_i \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i} \right\|^2 = \phi(u_0) \tag{3.5}$$

for all $m \leq N^n$. This implies that $(\widehat{u}^n)$ is bounded in $H^1(0, T; \mathbb{R}^d) \hookrightarrow L^\infty(0, T; \mathbb{R}^d)$ independently of n. In particular, $\|\widehat{u}^n(t)\|$ and $\|\overline{u}^n(t)\|$ are bounded for all times.

Let us rewrite the Gonzalez scheme in the compact form

$$(\widehat{u}^n)' + \mathrm{D}\phi(\overline{u}^n) + \overline{r}^n = 0, \tag{3.6}$$

where the remainder $\overline{r}^n$ is defined by

$$r_i^n := 0 \ \text{ if } \mathrm{D}\phi(u_{i-1}^n) = 0 \ \text{ and}$$
$$r_i^n := \Big(\phi(u_i^n) - \phi(u_{i-1}^n) - \big(\mathrm{D}\phi(u_i^n), u_i^n - u_{i-1}^n\big)\Big) \frac{u_i^n - u_{i-1}^n}{\|u_i^n - u_{i-1}^n\|^2} \ \text{ if } \mathrm{D}\phi(u_{i-1}^n) \neq 0.$$

One readily checks that

$$\|r_i^n\| \leq \left| \int_0^1 \left(\mathrm{D}\phi(\xi u_i^n + (1-\xi)u_{i-1}^n) - \mathrm{D}\phi(u_i^n), \frac{u_i^n - u_{i-1}^n}{\|u_i^n - u_{i-1}^n\|} \right) \mathrm{d}\xi \right|.$$

Denote by ω a continuity modulus for $\mathrm{D}\phi$ on a ball containing $\overline{u}^n$ for all times. In particular, $\omega : [0, \infty) \to [0, \infty)$ is nondecreasing and $\lim_{t\to 0+} \omega(t) = \omega(0) = 0$. Then

$$\|r_i^n\| \leq \omega\big((1 - \xi)\|u_i - u_{i-1}\|\big) \leq \omega\big(\|u_i - u_{i-1}\|\big).$$

This implies that

$$\max_{i=1,\ldots,N^n} \|r_i^n\| \leq \max_i \ \omega\big(\|u_i - u_{i-1}\|\big) \leq \max_i \omega \left(\int_{t_{i-1}^n}^{t_i^n} \|(\widehat{u}^n)'\| \mathrm{d}r \right)$$
$$\leq \max_i \ \omega \left((t_i^n - t_{i-1}^n)^{1/2} \left(\int_{t_{i-1}^n}^{t_i^n} \|(\widehat{u}^n)'\|^2 \mathrm{d}r \right)^{1/2} \right)$$

$$\le \max_i \omega\left((t_i^n - t_{i-1}^n)^{1/2}\left(\int_0^T \|(\widehat{u}^n)'\|^2 \mathrm{d}r\right)^{1/2}\right)$$
$$\le \omega(C(\tau^n)^{1/2}).$$

In particular, $\overline{r}^n \to 0$ strongly in $L^\infty(0,T;\mathbb{R}^d)$.

Bound (3.5) ensures, upon passing to subsequences, which are not relabeled, that

$$\begin{aligned} \widehat{u}^n \rightharpoonup u \quad & \text{weakly in } \ H^1(0,T;\mathbb{R}^d), \\ \widehat{u}^n, \overline{u}^n \to u \quad & \text{strongly in } \ L^\infty(0,T;\mathbb{R}^d), \\ \mathrm{D}\phi(\overline{u}^n) \to \mathrm{D}\phi(u) \quad & \text{strongly in } \ L^q(0,T;\mathbb{R}^d) \quad \text{for all } q<\infty. \end{aligned}$$

For the last convergence, we have used that $\mathrm{D}\phi(\overline{u}^n) \to \mathrm{D}\phi(u)$ pointwise, as $\mathrm{D}\phi$ is continuous, and that $\|\mathrm{D}\phi(\overline{u}^n)\|$ is uniformly bounded. One can hence pass to the limit in (3.6) and find that u solves the gradient-flow problem (1.1).

Step 2: $\phi \in C^{1,1}_{\mathrm{loc}}(\mathbb{R}^d)$. In case $\mathrm{D}\phi$ is locally Lipschitz continuous, the solution of (1.1) is unique and we can prove an error estimate. Let us start by noting that in this case one has

$$\|r_i^n\| \le C\|u_i^n - u_{i-1}^n\| = C\tau_i \left\|\frac{u_i^n - u_{i-1}^n}{\tau_i}\right\|. \tag{3.7}$$

As $\overline{u}^n$ is bounded, (3.6) and (3.7) imply that

$$\|(\widehat{u}^n)'\| \le \|\mathrm{D}\phi(\overline{u}^n)\| + \|\overline{r}^n\| \le C + C\tau^n\|(\widehat{u}^n)'\|.$$

Consequently, for sufficiently small values of $\tau^n > 0$, we infer that $(\widehat{u}^n)'$ is bounded in $L^\infty(0,T;\mathbb{R}^d)$. Relation (3.7) then ensures that

$$\|\overline{r}^n\|_{L^\infty(0,T;\mathbb{R}^d)} \le C\tau^n\|(\widehat{u}^n)'\|_{L^\infty(0,T;\mathbb{R}^d)} \le C\tau^n. \tag{3.8}$$

Take now the difference between (3.6), written for the partition n, and the same equation for the partition m, test it against $\widehat{u}^n - \widehat{u}^m$, and integrate in time. Then, in view of the Lipschitz continuity of $\mathrm{D}\phi$,

$$\begin{aligned} &\frac{1}{2}\|(\widehat{u}^n - \widehat{u}^m)(t)\|^2 \\ &\quad = -\int_0^t (\mathrm{D}\phi(\overline{u}^n) - \mathrm{D}\phi(\overline{u}^m), \widehat{u}^n - \widehat{u}^m)\mathrm{d}r - \int_0^t (\overline{r}^n - \overline{r}^m, \widehat{u}^n - \widehat{u}^m)\mathrm{d}r \\ &\quad \le C\left(\int_0^t \|\overline{u}^n - \overline{u}^m\|\,\|\widehat{u}^n - \widehat{u}^m\|\mathrm{d}r + \int_0^t (\|\overline{r}^n\| + \|\overline{r}^m\|)\|\widehat{u}^n - \widehat{u}^m\|\mathrm{d}r\right). \end{aligned} \tag{3.9}$$

The difference of the piecewise linear and piecewise constant functions can be estimated according to

$$\int_0^T \|\widehat{u}^n - \overline{u}^n\|^2 \mathrm{d}r = \sum_{i=1}^{N^n} \int_{t_{i-1}^n}^{t_i^n} \|\widehat{u}^n - \overline{u}^n\|^2 \mathrm{d}r = \sum_{i=1}^{N^n} \|u_i^n - u_{i-1}^n\|^2 \left(\int_{t_{i-1}^n}^{t_i^n} (\ell_i^n)^2 \mathrm{d}r \right)$$
$$= \sum_{i=1}^{N^n} \frac{\tau_i^n}{3} \|u_i^n - u_{i-1}^n\|^2 \leq \frac{(\tau^n)^2}{3} \sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 \leq C(\tau^n)^2.$$

Using (3.8) and the previous estimate, relation (3.9) gives

$$\frac{1}{2}\|(\widehat{u}^n - \widehat{u}^m)(t)\|^2 \leq C \int_0^t \big(\|\overline{u}^n - \widehat{u}^n\|^2 + \|\widehat{u}^n - \widehat{u}^m\|^2 + \|\widehat{u}^m - \overline{u}^m\|^2\big)\mathrm{d}r$$
$$+ C \int_0^t \big(\|\overline{r}^n\| + \|\overline{r}^m\|\big)\|\widehat{u}^n - \widehat{u}^m\|\mathrm{d}r$$
$$\leq C \int_0^t \|\widehat{u}^n - \widehat{u}^m\|^2 \mathrm{d}r + C\big((\tau_i^n)^2 + (\tau_i^m)^2\big).$$

We deduce from the Gronwall lemma that

$$\|\widehat{u}^n - \widehat{u}^m\|_{L^\infty(0,T;\mathbb{R}^d)} \leq C(\tau^n + \tau^m). \tag{3.10}$$

Then, again by arguing on the difference of relations (3.6) and using the local Lipschitz continuity of $\mathrm{D}\phi$, we find that

$$\|(\widehat{u}^n)' - (\widehat{u}^m)'\|_{L^\infty(0,T;\mathbb{R}^d)} \leq C\|\overline{u}^n - \overline{u}^m\|_{L^\infty(0,T;\mathbb{R}^d)}$$
$$+ \|\overline{r}^n\|_{L^\infty(0,T;\mathbb{R}^d)} + \|\overline{r}^m\|_{L^\infty(0,T;\mathbb{R}^d)}.$$

Bounds (3.8) and (3.10) imply that

$$\|\widehat{u}^n - \widehat{u}^m\|_{W^{1,\infty}(0,T;\mathbb{R}^d)} \leq C(\tau^n + \tau^m).$$

By taking $m \to \infty$, we obtain the error estimate (3.2). We refer to [2, Lemma 17.2.2, p. 695] for the analogous result for the Euler method.

Step 3: $\phi \in C^3(\mathbb{R}^d)$. We now turn to the proof of second-order consistency under condition (3.3). We conclude from condition (3.3) and the estimate $\|u_i - u_{i-1}\| \leq C\tau_i^n$ that the scheme (1.4) can be written as

$$\frac{u_i - u_{i-1}}{\tau_i^n} + \mathrm{D}\phi(u_i) = -\big(\phi(u_i) - \phi(u_{i-1}) - \mathrm{D}\phi(u_i)\cdot(u_i - u_{i-1})\big)\frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2}$$
$$= \left(\frac{1}{2}(u_i - u_{i-1})\cdot \mathrm{D}^2\phi(u_i)(u_i - u_{i-1}) + \mathrm{O}(\|u_i - u_{i-1}\|^3)\right)\frac{u_i - u_{i-1}}{\|u_i - u_{i-1}\|^2}$$
$$= \frac{1}{2}\mathrm{D}^2\phi(u_i)(u_i - u_{i-1}) + \mathrm{O}((\tau_i^n)^2).$$

In particular, u_i solves

$$u_i = u_{i-1} - \tau_i^n \mathrm{D}\phi(u_i) + \frac{\tau_i^n}{2}\mathrm{D}^2\phi(u_i)(u_i - u_{i-1}) + \mathrm{O}((\tau_i^n)^3). \tag{3.11}$$

Let w be the unique solution of $w' + \mathrm{D}\phi(w) = 0$ with $w(t_{i-1}) = u_{i-1}$. Then $w'' = -\mathrm{D}^2\phi(w)w' = \mathrm{D}^2\phi(w)\mathrm{D}\phi(w) \in C^1(\mathbb{R}^d)$. A Taylor expansion for some $\eta \in [t_{i-1}, t_i]$ leads to

$$\begin{aligned} w(t_i) &= w(t_{i-1}) + \tau_i^n w'(t_i) - \frac{(\tau_i^n)^2}{2} w''(t_i) + \frac{(\tau_i^n)^3}{6} w'''(\eta) \\ &= u_{i-1} - \tau_i^n \mathrm{D}\phi(w(t_i)) - \frac{(\tau_i^n)^2}{2}\mathrm{D}^2\phi(w(t_i))\mathrm{D}\phi(w(t_i)) + \frac{(\tau_i^n)^3}{6} w'''(\eta) \\ &= u_{i-1} - \tau_i^n \mathrm{D}\phi(w(t_i)) + \frac{\tau_i^n}{2}\mathrm{D}^2\phi(w(t_i))(w(t_i) - u_{i-1}) + \mathrm{O}((\tau_i^n)^3), \end{aligned} \tag{3.12}$$

where in the last step we used the fact that $w(t_i) - u_{i-1} = w(t_i) - w(t_{i-1}) = -\tau_i^n \mathrm{D}\phi(w(t_i)) + \mathrm{O}((\tau_i^n)^2)$. Take now the difference of relations (3.12) and (3.11), multiply it by $w(t_i) - u_i$ and use the local Lipschitz continuity of the functions $\mathrm{D}\phi$ and $\xi \mapsto \mathrm{D}^2\phi(\xi)(\xi - u_{i-1})$ in order to obtain

$$\|w(t_i) - u_i\|^2 \leq \tau_i^n C \|w(t_i) - u_i\|^2 + C(\tau_i^n)^3 \|w(t_i) - u_i\|.$$

Hence, for sufficiently small $\tau^n > 0$, we infer that

$$\|w(t_i) - u_i\| \leq C(\tau_i^n)^3. \tag{3.13}$$

The error control (3.4) follows from the stability of Problem (1.1). Indeed, let u_1, u_2 be two solutions of the system $u' + \mathrm{D}\phi(u) = 0$, it follows for all $0 \leq s \leq t$ that

$$\|(u_1 - u_2)(t)\| \leq \mathrm{e}^{L(t-s)} \|(u_1 - u_2)(s)\|, \tag{3.14}$$

where $L = \max_{t,i} \|\mathrm{D}^2\phi(u_i(t))\|$. We use (3.13) and (3.14) to find that

$$\begin{aligned} &\|u(t_i) - u_i\| \\ &\quad \leq \|u(t_i) - w(t_i)\| + \|w(t_i) - u_i\| \leq \mathrm{e}^{L\tau_i} \|u(t_{i-1}) - u_{i-1}\| + \|w(t_i) - u_i\| \\ &\quad \leq \mathrm{e}^{L\tau_i} \|u(t_{i-1}) - u_{i-1}\| + C(\tau_i^n)^3 \leq \cdots \leq Ci \max\{1, \mathrm{e}^{LT}\}(\tau_i^n)^3 \leq C(\tau^n)^2 \end{aligned} \tag{3.15}$$

whence the bound $\|u - \widehat{u}^n\|_{L^\infty(0,T;\mathbb{R}^d)} \leq C(\tau^n)^2$ holds. Estimate (3.4) follows from the local Lipschitz continuity of $\mathrm{D}\phi$. □

Some of the arguments of the proof of Theorem 3.2 can be adapted to the case of a gradient flow on an unbounded time domain $t \in [0, \infty)$. In particular, by possibly asking ϕ to be coercive, one can reproduce the convergence proof of point (i). On the

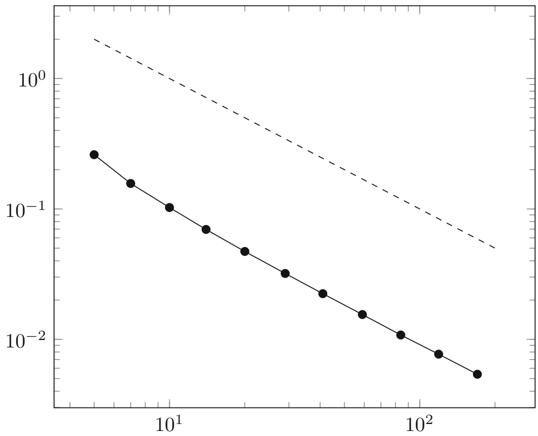

Fig. 1 L^∞ error of the difference of the solution of the Gonzalez scheme and the exact solution for the choice (3.16) as a function of $1/\tau$ in log-log scale. The dashed line corresponds to order τ. The order of convergence in (3.2) is sharp

other hand, the error bounds at points (ii)–(iii) call for some Gronwall argument, see (3.10) and (3.15), which necessarily calls for the finiteness of the time interval $[0, T]$.

Note that condition (3.3) ensuring the second-order convergence (3.4) is always fulfilled in one dimension ($d = 1$) and is sharp, as the example in Fig. 2 $u' + u = 0$, $u(0) = 1$ shows.

In several dimensions, condition (3.3) holds for radial functions ϕ. On the other hand, by testing on a nonradial potential, one can check that the first-order convergence rate in (3.2) is sharp, as the choice

$$\phi(u) = u_1^2 + \tfrac{1}{4}u_2^2 \quad \text{for } u = (u_1, u_2) \in \mathbb{R}^2,\ u_0 = (1, 1), \tag{3.16}$$

with exact solution $u(t) = (\exp(-2t), \exp(-t/2))$ shows, see Fig. 1. Moreover, condition (3.3) holds for quadratic potentials.

Note that in Theorem 3.2 one could approximate the initial data $u_0^n \to u_0$ as well. This would still imply the convergence by including an additional term in the error estimate, taking into account the initial error $\|u_0 - u_0^n\|$.

Before closing this section, let us comment on some shortcomings of the Gonzalez scheme. First of all, the analysis of the Gonzalez scheme is at the moment restricted to C^1 energies with compact sublevels, which in turn enforces a finite-dimensional setting.

In some nonsmooth cases, the existence of an update could be conditional to the smallness of the time step. An example in this direction is given by the scalar energy $\phi(u) = -u + I_{[\sqrt{2},\infty)}(u)$, where $I_{[\sqrt{2},\infty)}$ stands for the indicator function of the half-line $[\sqrt{2}, \infty)$. For $u_0 = 2$, the unique strong solution to (1.1) is $u(t) = \max\{2-t, \sqrt{2}\}$. For all given *rational* time steps $\tau > 0$, one can however identify a number $i \in \mathbb{N}$ such that $2 - i\tau < \sqrt{2} < 2 - (i-1)\tau$ and check that the Gonzalez scheme cannot be solved at step i.

Eventually, the Gonzalez scheme seems not to be related to a variational principle. The actual computation of the update from the previous step involves the solution of the scalar energy equality as well as the discussion on the alignment condition, which makes the incremental problem a nonlinear system.

4 The De Giorgi Scheme

Let now the Hilbert space H be general, possibly infinite-dimensional. We shall consider energies of the form $\phi = \phi_1 + \phi_2$, where $\phi_1 : H \to (-\infty, \infty]$ is convex, proper, lower semicontinuous and $\phi_2 \in C^{1,\alpha}_{\mathrm{loc}}(H)$ for some $\alpha \in (0, 1]$. For the sake of notational simplicity, we assume $\partial\phi_1$ to be single-valued and remark that $u \in D(\partial\phi) \equiv D(\partial\phi_1)$. The arguments easily extend to not single-valued operators $\partial\phi_1$ with a bit notational intricacy.

In case ϕ is nonconvex, one can easily find situations for which the De Giorgi scheme (1.5) has no solution. An example in this direction is the scalar energy $\phi(u) = u(1-u)$ from $u_{i-1} = 0$. Indeed, in this case (1.5) reads as

$$u(1-u) + \frac{u^2}{2\tau} + \frac{\tau}{2}(1-2u)^2 = 0,$$

which admits no real solution, independently of the choice of $\tau > 0$. We are hence forced to *generalize* the De Giorgi scheme by allowing the possibility of solving (1.5) with some tolerance. Define the functionals $G_i : D(\partial\phi) \times D(\partial\phi) \to (-\infty, \infty]$ as

$$G_i(u, v) = \phi(u) + \frac{\tau_i}{2}\left\|\frac{u-v}{\tau_i}\right\|^2 + \frac{\tau_i}{2}\|\partial\phi(u)\|^2 - \phi(v).$$

Given $u_{i-1} \in D(\partial\phi)$, one can find $u_i \in D(\partial\phi)$ such that $(G_i(u_i, u_{i-1}))^+$ is arbitrarily small. In particular, we look for $u_i \in D(\partial\phi)$ such that

$$\phi(u_i) + \frac{\tau_i}{2}\left\|\frac{u_i - u_{i-1}}{\tau_i}\right\|^2 + \frac{\tau_i}{2}\|\partial\phi(u_i)\|^2 = \phi(u_{i-1}) + \rho_i \tag{4.1}$$

where the *residuals* ρ_i are such that ρ_i^+ is small enough, see Theorem 4.2 below.

The following existence result holds.

Theorem 4.1 (Existence for the De Giorgi scheme) *Let $\phi = \phi_1 + \phi_2$ have compact sublevels, $\phi_1 : H \to (-\infty, \infty]$ be convex, proper, lower semicontinuous with $\partial\phi_1$ being single-valued, and $\phi_2 \in C^{1,\alpha}_{\mathrm{loc}}(H)$ for $\alpha \in (0, 1]$. Furthermore, let $u_{i-1} \in D(\partial\phi)$ be given with $\partial\phi(u_{i-1}) \neq 0$. Then there exists $u_i \in D(\partial\phi)$ with $u_i \neq u_{i-1}$ and*

$$G_i(u_i, u_{i-1}) \le \frac{L}{1+\alpha}\|u_i - u_{i-1}\|^{1+\alpha}, \tag{4.2}$$

where L is the Hölder constant of $\mathrm{D}\phi_2$. In particular, (4.1) can be solved with $\rho_i \le L\|u_i - u_{i-1}\|^{1+\alpha}/(1+\alpha)$. In case ϕ is convex, namely $\phi_2 = 0$, and $\|\partial\phi\|$ is strongly continuous along segments in $D(\partial\phi)$, one can find u_i such that $G_i(u_i, u_{i-1}) = 0$.

In the convex case $\phi_2 = 0$, the functional $u \mapsto \|\partial\phi(u)\|$ is lower semicontinuous along segments in $D(\partial\phi)$. On the other hand, the continuity of $\|\partial\phi\|$ along segments assumed in the theorem may be available even in some nonsmooth situations. As an illustration of this fact, let $H = L^2(\Omega)$ for some Lipschitz domain $\Omega \in \mathbb{R}^d$ and define ϕ to be the Dirichlet energy

$$\phi(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 \mathrm{d}x$$

for $u \in H^1(\Omega)$ and $\phi(u) = \infty$ otherwise, which is lower semicontinuous but not continuous in L^2. Then $D(\partial\phi) = H^2(\Omega)$ and $\partial\phi(u) = -\Delta u$. Hence, for all u_1, $u_2 \in D(\partial\phi)$, the mapping

$$\begin{aligned}
&\lambda \in \mathbb{R} \mapsto \|\partial\phi(\lambda u_1 + (1-\lambda)u_2)\| = \|-\Delta(\lambda u_1 + (1-\lambda)u_2)\| \\
&= \left(\lambda^2\|-\Delta u_1\|^2 + 2\lambda(1-\lambda)\int_\Omega \Delta u_1 \Delta u_2 \,\mathrm{d}x + (1-\lambda)^2\|-\Delta u_2\|^2\right)^{1/2}
\end{aligned}$$

is continuous for $\lambda \in [0, 1]$. Note that the continuity of $\|\partial\phi\|$ along segments in $D(\partial\phi)$ is weaker than the strong continuity of $\partial\phi$.

Proof of Theorem 4.1 Let us start by considering the classical Euler scheme

$$\frac{u^e - u_{i-1}}{\tau_i} + \partial\phi(u^e) = 0, \tag{4.3}$$

which can be solved by minimizing the function

$$u \mapsto \frac{1}{2\tau_i}\|u - u_{i-1}\|^2 + \phi(u).$$

Note that the latter function has compact sublevels and the direct method applies. We readily check that $u_i := u^e$ delivers $u_i \neq u_{i-1}$, for u_{i-1} is not critical. In addition, one can prove (4.2) as well. To this aim, we test relation (4.3) with $u^e - u_{i-1}$ and use the convexity of ϕ_1 and the Hölder continuity of $\mathrm{D}\phi_2$ to find that

$$\begin{aligned}
&\tau_i\left\|\frac{u^e - u_{i-1}}{\tau_i}\right\|^2 + \phi(u^e) - \phi(u_{i-1}) \\
&\quad\le \phi_2(u^e) - \phi_2(u_{i-1}) - (\mathrm{D}\phi_2(u^e), u^e - u_{i-1}) \\
&\quad= \int_0^1 \Big(\mathrm{D}\phi_2(\xi u^e + (1-\xi)u_{i-1}) - \mathrm{D}\phi_2(u^e), u^e - u_{i-1}\Big)\mathrm{d}\xi \\
&\quad\le L\|u^e - u_{i-1}\|^{1+\alpha}\int_0^1 (1-\xi)^\alpha \,\mathrm{d}\xi = \frac{L}{1+\alpha}\|u^e - u_{i-1}\|^{1+\alpha}.
\end{aligned} \tag{4.4}$$

By using again (4.3), one finds that u^e fulfills (4.2) when choosing $u_i = u^e$.

In the convex case $\phi_2 = 0$, the choice $u_i = u^e$ does not necessarily implies that $G_i(u_i, u_{i-1}) = 0$. In case $G_i(u^e, u_{i-1}) < 0$ (check $\phi(u) = u^2$ from $u_0 = 1$ for sufficiently small time steps), we can look for a point u_i along the segment $u_\lambda = \lambda u^e + (1-\lambda)u_{i-1}$ for $\lambda \in [0, 1)$ such that $G_i(u_\lambda, u_{i-1}) \equiv 0$. Note that the real function

$$\lambda \mapsto g(\lambda) := G_i(u_\lambda, u_{i-1})$$

is well defined, for $D(\phi)$ and $D(\partial\phi)$ are convex. Moreover, $\lambda \in [0, 1] \mapsto \phi(u_\lambda)$ is convex and lower semicontinuous, hence continuous on $[0, 1]$. The assumption on the continuity of $\|\partial\phi\|$ along lines implies that $\lambda \mapsto \|\partial\phi(u_\lambda)\|$ is continuous as well. Hence, g is continuous in $(0, 1)$ and we find $\lambda \in (0, 1)$ such that $G_i(u_\lambda, u_{i-1}) = g(\lambda) = 0$ as $g(0) = \tau_i\|\partial\phi(u_{i-1})\|^2/2 > 0$, for u_{i-1} is not singular, and $g(1) < 0$. □

As will be clear from the statement of Theorem 4.2 below, the smallness of the residuals ρ_i in (4.1) will be instrumental to pass to the limit in the scheme. Theorem 4.1 claims that these can be as small as $L\|u_i - u_{i-1}\|^{1+\alpha}$, which allows us to prove convergence. This suggests to consider discrete solutions u_i that *minimize* the residuals. This corresponds to formulate a variational principle of the form

$$u_i \in \arg\min_{u \in D(\partial\phi)} G_i(u, u_{i-1}). \tag{4.5}$$

Recall the classical strong/weak closure property

$$x_n \to x \text{ strongly in } H, \; y_n \rightharpoonup y \text{ weakly in } H, \; y_n \in \partial\phi_1(x_n) \;\Rightarrow\; y \in \partial\phi_1(x).$$

This minimum problem is solvable by the direct method, for the sublevels of the function $u \mapsto G_i(u, u_{i-1})$ are strongly compact. Indeed, one has $G_i(\cdot, u_{i-1}) \geq \phi(\cdot) - \phi(u_{i-1})$ and the sublevels of ϕ are compact.

In particular, let (u^n) be a strongly convergent minimizing sequence such that $u^n \to u_i$ as $n \to \infty$. Then we deduce from the strong-weak closure of the subdifferential $\partial\phi_1$ that

$$\partial\phi(u^n) = \partial\phi_1(u^n) + \mathrm{D}\phi_2(u^n) \rightharpoonup \partial\phi_1(u_i) + \mathrm{D}\phi_2(u_i) = \partial\phi(u_i) \quad \text{weakly in } H,$$

and the lower semicontinuity of $u \mapsto G_i(u, u_{i-1})$ follows. Clearly, $u_i \neq u_{i-1}$ if u_{i-1} is not critical. By inspecting the proof of Theorem 4.1, we see that the minimization in (4.5) implies that $G_i(u, u_{i-1}) \leq 0$, hence fulfilling (4.2). As such, Theorem 4.2 below will ensure the convergence of the discrete solution obtained via (4.5) as well.

Let us now introduce our convergence result. To this aim, we specify the notation with respect to a sequence of partitions $\{0 = t_0^n < t_n^1 < \cdots < t_{N^n}^n = T\}$ as

$$G_i^n(u, v) = \phi(u) + \frac{\tau_i^n}{2}\left\|\frac{u - v}{\tau_i^n}\right\|^2 + \frac{\tau_i^n}{2}\|\partial\phi(u)\|^2 - \phi(v).$$

Theorem 4.2 (Convergence for the De Giorgi scheme) *Under the assumptions of Theorem 4.1, let $u_i^n \in D(\partial\phi)$ be such that $u_0^n = u_0$ and*

$$\sum_{i=1}^{N^n} G_i^n(u_i^n, u_{i-1}^n)^+ \to 0 \;\; as \;\; n \to \infty. \tag{4.6}$$

Then $\widehat{u}^n \to u$ converges strongly in $H^1(0, T; H)$, where u solves the gradient-flow problem (1.1).

A comment on condition (4.6) is in order. For all $n \in \mathbb{N}$, let $u_i^n \in D(\partial\phi)$ be the sequence fulfilling (4.2) from Theorem 4.1. Then,

$$\begin{aligned}
&\phi(u_{N^n}^n) + \frac{1}{2}\sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 + \frac{1}{2}\sum_{i=1}^{N^n} \tau_i^n \left\| \partial\phi(u_i^n) \right\|^2 - \phi(u_0) \\
&= \sum_{i=1}^{N^n} G_i^n(u_i^n, u_{i-1}^n) \le \frac{L(\tau^n)^\alpha}{1+\alpha} \sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^{1+\alpha} \\
&\le C(\tau^n)^\alpha \left(1 + \sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 \right).
\end{aligned} \tag{4.7}$$

For sufficiently small values of $\tau^n > 0$, we can absorb the last term on the right-hand side by the corresponding term on the left-hand side, which leads to

$$\sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 \le C.$$

Inserting this estimate into (4.7) shows that

$$\sum_{i=1}^{N^n} G_i^n(u_i^n, u_{i-1}^n) \le C(\tau^n)^\alpha,$$

so that (4.6) holds. In particular, Theorem 4.2 implies the convergence for the Euler scheme as well. More generally, condition (4.6) can be seen as a criterion for checking convergence, independently of the procedure that produced the discrete solution.

Proof Arguing as in (4.7), we have

$$\begin{aligned}
&\phi(u_m^n) + \frac{1}{2}\sum_{i=1}^{m} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 + \frac{1}{2}\sum_{i=1}^{m} \tau_i^n \left\| \partial\phi(u_i^n) \right\|^2 - \phi(u_0) \\
&\le \sum_{i=1}^{N^n} G_i^n(u_i^n, u_{i-1}^n)^+ \quad \text{for all } m \le N^n.
\end{aligned} \tag{4.8}$$

As the right-hand side converges for $n \to \infty$, by assumption, one deduces the bounds

$$\sup_m \phi(u_m^n) + \frac{1}{2}\sum_{i=1}^{N^n} \tau_i^n \left\| \frac{u_i^n - u_{i-1}^n}{\tau_i^n} \right\|^2 + \frac{1}{2}\sum_{i=1}^{N^n} \tau_i^n \|\partial\phi(u_i^n)\|^2 \le C.$$

Recall that the sublevels of ϕ are assumed to be compact. We can hence apply a diagonal extraction argument (without relabeling) and obtain

$$\begin{aligned}
&\widehat{u}^n \to u \quad \text{strongly in } \ C([0,T];H) \ \text{ and weakly in } \ H^1(0,T;H),\\
&\overline{u}^n \to u \quad \text{strongly in } \ L^\infty(0,T;H),\\
&\partial\phi(\overline{u}^n) \to \partial\phi(u) \quad \text{weakly in } \ L^2(0,T;H).
\end{aligned} \tag{4.9}$$

In order to identify the limit in (4.9), we have used the strong-weak closure of the subdifferential $\partial\phi$, which follows from the very definition of the subdifferential. In particular, $\overline{u}^n(t) \to u(t)$ for all $t \in (0,T)$. Fix now $t \in (0,T)$ and, for all n, choose m such that $t_{m-1}^n < t \le t_m^n$. By passing to the limit inferior as $n \to \infty$ in estimate (4.8) and using $\phi(u_m^n) = \phi(\overline{u}^n(t))$, it follows that

$$\phi(u(t)) + \frac{1}{2}\int_0^t \|u'\|^2 \mathrm{d}r + \frac{1}{2}\int_0^t \|\partial\phi(u)\|^2 \mathrm{d}r \le \phi(u_0).$$

Therefore, we can use the chain rule [10, Lemme 3.3, p. 73] to conclude that

$$\begin{aligned}
&\frac{1}{2}\int_0^t \|u'\|^2 \mathrm{d}r + \frac{1}{2}\int_0^t \|\partial\phi(u)\|^2 \mathrm{d}r \le \phi(u_0) - \phi(u(t))\\
&\quad = -\int_0^t (\phi \circ u)' \mathrm{d}r = -\int_0^t (\partial\phi(u), u')\mathrm{d}r \le \frac{1}{2}\int_0^t \|u'\|^2 \mathrm{d}r + \frac{1}{2}\int_0^t \|\partial\phi(u)\|^2 \mathrm{d}r.
\end{aligned}$$

In particular, all these inequalities are actually equalities and

$$\frac{1}{2}\|u'\|^2 + \frac{1}{2}\|\partial\phi(u)\|^2 = -(\partial\phi(u), u') \quad \text{a.e. in } \ (0,T).$$

This implies that the equality (1.3) holds and that u is a solution of the gradient-flow problem (1.1). □

Theorem 4.2 can be formulated for gradient flows on the whole semiline $t \in [0,\infty)$ as well, by asking the convergence condition to be

$$\sum_{i=1}^{\infty} G_i^n(u_i^n, u_{i-1}^n)^+ \to 0 \ \text{ as } \ n \to \infty.$$

Indeed, estimate (4.7) holds in case $N^n = \infty$ as well. In order to pass to the limit, one uses again the analogue of (4.8). A localization and diagonal-extraction argument entails strong compactness.

We present some error estimates for the De Giorgi scheme in finite dimensions.

Proposition 4.3 (Error control for the De Giorgi scheme in $\mathbb{R}^d$) *Let $\phi \in C^2(\mathbb{R}^d)$ be bounded from below and u_i and v_i fulfill $u_0 = v_0$,*

$$G_i(u_i, u_{i-1}) = 0, \quad \textit{and} \quad v_i \in \arg\min G_i(\cdot, v_{i-1}), \tag{4.10}$$

respectively. Then for all $i = 1, \ldots, N$

$$\|u(t_i) - v_i\| \leq C\tau, \tag{4.11}$$

$$\|u(t_i) - u_i\| \leq C\tau^{1/2}, \tag{4.12}$$

where u is the unique solution of (1.1).

Proof Note first that (1.1) is uniquely solvable and that its solution u is bounded. The existence of u_i is due to Theorem 4.1 and that of v_i follows by an application of the direct method. Owing to the energy equalities (1.2) and (1.5) and the minimality of v_i, we conclude that u is bounded uniformly in time and that the values u_i and v_i are also bounded independently of i.

Step 1: proof of (4.11). Let us start by considering v_i. We show that the consistency error is of order τ^2. To this aim, let w be the unique solution to $w' + \mathrm{D}\phi(w) = 0$ with $w(t_{i-1}) = v_{i-1}$. Note that $w \in C^2$ and that w'' is bounded for all times, depending on $\phi(v_{i-1})$ only (hence, just on $\phi(u_0)$). We deduce from the minimality of v_i that

$$0 = \mathrm{D}\phi(v_i) + \frac{1}{\tau_i}(v_i - v_{i-1}) + \tau_i \mathrm{D}^2\phi(v_i)\mathrm{D}\phi(v_i),$$

which is equivalent to

$$v_i = v_{i-1} - \tau_i \mathrm{D}\phi(v_i) - \tau_i^2 \mathrm{D}^2\phi(v_i)\mathrm{D}\phi(v_i). \tag{4.13}$$

On the other hand, a Taylor expansion ensures that

$$w(t_i) = w(t_{i-1}) + \tau_i w'(t_i) - \frac{\tau_i^2}{2} w''(\eta) = v_{i-1} - \tau_i \mathrm{D}\phi(w(t_i)) - \frac{\tau_i^2}{2} w''(\eta) \tag{4.14}$$

for some $\eta \in [t_{i-1}, t_i]$. By taking the difference between (4.14) and (4.13) and multiplying it by $w(t_i) - v_i$, we find that

$$\begin{aligned}
&\|w(t_i) - v_i\|^2 \\
&\quad \leq C\tau_i \|w(t_i) - v_i\|^2 + \left(\tau_i^2 \|\mathrm{D}^2\phi(v_i)\| \|\mathrm{D}\phi(v_i)\| + \frac{\tau_i^2}{2} \|w''(\eta)\|\right) \|w(t_i) - v_i\| \\
&\quad \leq C\tau_i \|w(t_i) - v_i\|^2 + C\tau_i^2 \|w(t_i) - v_i\|,
\end{aligned}$$

which, for sufficiently small values of $\tau_i > 0$, implies that

$$\|w(t_i) - v_i\| \leq C\tau_i^2. \tag{4.15}$$

We now argue as in (3.14) and deduce from (4.15) that

$$\begin{aligned}
\|u(t_i) - v_i\| &\leq \|u(t_i) - w(t_i)\| + \|w(t_i) - v_i\| \\
&\leq \mathrm{e}^{L\tau_i}\|u(t_{i-1}) - v_{i-1})\| + \|w(t_i) - v_i\| \\
&\leq \mathrm{e}^{L\tau_i}\|u(t_{i-1}) - v_{i-1})\| + C\tau_i^2 \\
&\leq \cdots \leq Ci\max\{1, \mathrm{e}^{LT}\}\tau^2 \leq C\tau,
\end{aligned} \tag{4.16}$$

from which (4.11) follows.

Step 2: proof of (4.12). Let us address the error control for u_i. We use $G_i(u_i, u_{i-1}) = 0$ to compute

$$\begin{aligned}
\|u_i - u_{i-1} + \tau_i \mathrm{D}\phi(u_i)\|^2 &= \|u_i - u_{i-1}\|^2 + \tau_i^2\|\mathrm{D}\phi(u_i)\|^2 + 2\tau_i\mathrm{D}\phi(u_i)\cdot(u_i - u_{i-1}) \\
&= 2\tau_i(-\phi(u_i) + \phi(u_{i-1})) + 2\tau_i\mathrm{D}\phi(u_i)\cdot(u_i - u_{i-1}) \\
&= \tau_i\mathrm{D}^2\phi(\xi)(u_i - u_{i-1})\cdot(u_i - u_{i-1}) \leq C\tau_i^3,
\end{aligned} \tag{4.17}$$

where we used a Taylor expansion and ξ belongs to the segment joining u_i and u_{i-1}. For the last inequality, we also used the estimate $\|u_i - u_{i-1}\| \leq C\tau_i$.

Next, let w be the unique solution to $w' + \mathrm{D}\phi(w) = 0$ with $w(t_{i-1}) = u_{i-1}$. A Taylor expansion shows that

$$w(t_i) = u_{i-1} - \tau_i\mathrm{D}\phi(u_i) + \tau_i(\mathrm{D}\phi(u_i) - \mathrm{D}\phi(u_{i-1})) + \frac{\tau_i^2}{2}w''(\eta). \tag{4.18}$$

By (4.18) and (4.17), the estimate

$$\begin{aligned}
\|w(t_i) - u_i\| &= \left\| u_{i-1} - \tau_i\mathrm{D}\phi(u_i) + \tau_i(\mathrm{D}\phi(u_i) - \mathrm{D}\phi(u_{i-1})) + \frac{\tau_i^2}{2}w''(\eta) - u_i \right\| \\
&\leq \|u_i - u_{i-1} + \tau_i\mathrm{D}\phi(u_i)\| + C\tau_i^2 \leq C\tau_i^{3/2}
\end{aligned}$$

follows. The error control (4.12) can be proved by arguing similarly to (4.16). □

Differently from the convergence Theorem 4.2, the proof of Theorem 4.3 heavily relies on the finiteness of the time interval $[0, T]$, for it hinges on a Gronwall-like argument, see (4.16).

The convergence rates in (4.11)–(4.12) are sharp, as the one-dimensional test of Sect. 6 shows, see Fig. 2.

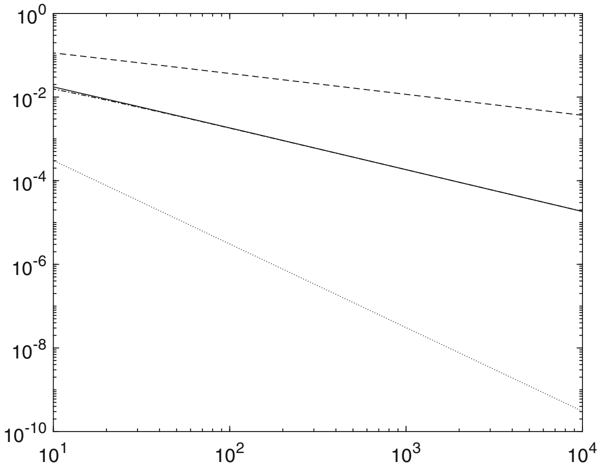

Fig. 2 L^∞ error on [0, 1] of the difference between the exact solution $u(t) = \mathrm{e}^{-t}$ of $u' + u = 0$, $u(0) = 1$ and the discrete solutions e_i (solid line), g_i (dotted line), v_i (dash-dotted line), and u_i (dashed line) as a function of $1/\tau$ in log-log scale

5 Extensions

We discuss some extensions of the De Giorgi scheme to other nonlinear evolution equations.

5.1 Generalized gradient flows

The analysis of the De Giorgi scheme can be extended to generalized gradient flows, namely

$$\partial\psi(u, u') + \partial\phi(u) \ni 0 \quad \text{for a.e. } t \in (0, T), \quad u(0) = u_0. \tag{5.1}$$

Here, $\psi : H \times H \to [0, \infty)$ and $\partial\psi(u, u')$ denotes partial subdifferentiation with respect to the second variable only and we recall that for simplicity, $\partial\phi$ is still assumed to be single-valued. More precisely, we assume that

- $$\forall u \in H : \ \psi(u, \cdot) \text{ is convex and lower semicontinuous,} \tag{5.2}$$
- the mapping $H \times H \times H \to \mathbb{R}$, $(u, v, w) \mapsto \psi(u, v) + \psi^*(u, w)$, is weakly lower semicontinuous,
- $$\exists c > 0,\ p > 1,\ \forall u,\ v,\ w \in H : \ \psi(u, v) + \psi^*(u, w) \geq c\|v\|^p + c\|w\|^{p'}, \tag{5.3}$$

where $p' = p/(p-1)$ and the Legendre–Fenchel conjugation is taken with respect to the second variable only. An example for ψ satisfying (5.2)–(5.3) is $\psi(u, u') = \beta(u)|u'|^p$, where $p > 1$ and β is sufficiently smooth, uniformly positive, and bounded. Note that, as a consequence of (5.3), for any $u \in H$ one has

$$w \in \partial\psi(u,0) \;\Leftrightarrow\; 0 = \psi(u,0) + \psi^*(u,w) \overset{(5.2)}{\geq} c|w|^q \;\Rightarrow\; w = 0. \tag{5.4}$$

The analog of equality (1.3) for generalized gradient flows reads as

$$\phi(u(t)) + \int_0^t \psi(u,u')\mathrm{d}r + \int_0^t \psi^*(u,-\partial\phi(u))\mathrm{d}r = \phi(u_0) \tag{5.5}$$

for all $t \in [0,T]$. This can be checked by equivalently rewriting (5.1) by using Fenchel's duality, $y \in \partial\psi(u,u') \Leftrightarrow \psi(u,u')+\psi^*(u,y) = (y,u')$ with $y = -\partial\phi(u)$, as

$$\psi(u,u') + \psi^*(u,-\partial\phi(u)) + (\partial\phi(u),u') = 0$$

and integrating in time. Correspondingly, one modifies the functionals G_i as follows

$$\widetilde{G}_i(u,v) := \phi(u) + \tau_i\psi\left(v, \frac{u-v}{\tau_i}\right) + \tau_i\psi^*(v,-\partial\phi(u)) - \phi(v). \tag{5.6}$$

Given the initial value $u_0 \in D(\phi)$, the De Giorgi scheme consists in finding $u_i \in D(\partial\phi)$ with $\widetilde{G}_i(u_i,u_{i-1}) = \rho_i$ and sufficiently small ρ_i^+. Note that the first occurrence in ψ and ψ^* in (5.6) is $v = u_{i-1}$, so that the scheme is implicit. Given a sequence of partitions, we use the notation

$$\widetilde{G}_i^n(u,v) := \phi(u) + \tau_i^n\psi\left(v, \frac{u-v}{\tau_i^n}\right) + \tau_i^n\psi^*(v,-\partial\phi(u)) - \phi(v).$$

Theorem 5.1 (De Giorgi scheme for generalized gradient flows) *Assume* (5.2)–(5.3) *and let $\phi = \phi_1 + \phi_2$ have compact sublevels, $\phi_1 : H \to (-\infty,\infty]$ be convex, proper, lower semicontinuous with $\partial\phi_1$ single-valued, and $\phi_2 \in C^{1,\alpha}_{\mathrm{loc}}(H)$ for $\alpha \in (0,1]$. Let $u_{i-1} \in D(\partial\phi)$ satisfy $\partial\phi(u_{i-1}) \neq 0$. Then there exists $u_i \in D(\partial\phi)$ with $u_i \neq u_{i-1}$ and*

$$\widetilde{G}_i(u_i,u_{i-1}) \leq \frac{L}{1+\alpha}\|u_i - u_{i-1}\|^{1+\alpha}, \tag{5.7}$$

where L is the Hölder constant of $\mathrm{D}\phi_2$. *In case ϕ is convex, namely $\phi_2 = 0$, and $v \mapsto \psi^*(u,-\partial\phi(v))$ is continuous (with respect to the strong topology) along segments in $D(\partial\phi)$, for all $u \in H$, one can find $u_i \neq u_{i-1}$ such that $\widetilde{G}_i(u_i,u_{i-1}) = 0$.*

Let $u_i^n \in D(\partial\phi)$ be such that $u_0^n = u_0$ and

$$\sum_{i=1}^{N^n} \widetilde{G}_i^n(u_i^n,u_{i-1}^n)^+ \to 0 \;\; as \;\; n \to \infty. \tag{5.8}$$

Then $\widehat{u}^n \to u$ converges weakly in $W^{1,p}(0,T;H)$, and u solves the generalized gradient-flow problem (5.1).

Proof By adapting the argument of Theorem 4.1, we wish to find u^e solving the Euler scheme

$$\partial\psi\left(u_{i-1}, \frac{u^e - u_{i-1}}{\tau_i}\right) + \partial\phi(u^e) \ni 0. \tag{5.9}$$

This can be done by letting

$$u^e \in \arg\min_{u\in H}\left(\psi\left(u_{i-1}, \frac{u - u_{i-1}}{\tau_i}\right) + \phi(u)\right).$$

This minimization problem is solvable as $\psi(u_{i-1}, \cdot)$ and ϕ are lower semicontinuous, ψ is nonnegative, and the sublevels of ϕ are compact. Hence the Direct Method applies. Note that the scheme in (5.9) is implicit. As $\partial\phi(u_{i-1}) \neq 0$ and $\partial\psi(u_{i-1}, 0) = 0$ from (5.4), we readily check that $u^e \neq u_{i-1}$.

We rewrite relation (5.9) equivalently as

$$\tau_i\psi\left(u_{i-1}, \frac{u^e - u_{i-1}}{\tau_i}\right) + \tau_i\psi^*\left(u_{i-1}, -\partial\phi(u^e)\right) + \left(\frac{u^e - u_{i-1}}{\tau_i}, \partial\phi(u^e)\right) = 0.$$

By arguing as in (4.4), we have that

$$\left(\frac{u^e - u_{i-1}}{\tau_i}, \partial\phi(u^e)\right) \geq \phi(u^e) - \phi(u_{i-1}) - \frac{L}{1+\alpha}\|u^e - u_{i-1}\|^{1+\alpha}.$$

In particular, it turns out that (5.7) holds for $u_i = u^e$.

In case $\phi_2 = 0$, we find that $\widetilde{G}_i(u^e, u_{i-1}) \leq 0$. Consider the map $\lambda \mapsto g(\lambda) = \widetilde{G}_i(u_\lambda, u_{i-1})$ for $u_\lambda = \lambda u^e + (1-\lambda)u_{i-1}$ and $\lambda \in [0, 1)$. It is continuous, for the functions $v \mapsto \phi(u_{i-1}, v)$, $v \mapsto \psi^*(u_{i-1}, -\partial\phi(v))$, and $v \mapsto \phi(v)$ are all continuous along the segments of $D(\partial\phi)$. Since $\partial\phi(u_{i-1}) \neq 0$, we conclude that

$$g(0) = \tau_i\psi(u_{i-1}, 0) + \tau_i\psi^*(u_{i-1}, -\partial\phi(u_{i-1})) \overset{(5.3)}{\geq} c\tau_i\|\phi(u_{i-1})\|^q > 0.$$

In case $g(1) = 0$ we have nothing to prove. If $g(1) < 0$, there exists $\lambda^* \in (0, 1)$ such that $g(\lambda^*) = \widetilde{G}_i(u_{\lambda^*}, u_{i-1}) = 0$. Now, let u_i^n fulfill (5.8). The coercivity (5.3) implies that

$$\begin{aligned}
&\phi(u_m^n) + \sum_{i=1}^m \tau_i^n\left(c\left\|\frac{u_i^n - u_{i-1}^n}{\tau_i^n}\right\|^p + c\|\partial\phi(u_i^n)\|^{p'} - \frac{1}{c}\right) - \phi(u_0)\\
&\quad\leq \phi(u_m^n) + \sum_{i=1}^m \tau_i^n\left(\psi\left(u_{i-1}^n, \frac{u_i^n - u_{i-1}^n}{\tau_i^n}\right) + \psi^*(u_{i-1}^n, -\partial\phi(u_i^n))\right) - \phi(u_0)\\
&\quad= \sum_{i=1}^{N^n} \widetilde{G}_i^n(u_i^n, u_{i-1}^n)^+ \quad \text{for all } m \leq N^n.
\end{aligned} \tag{5.10}$$

As the right-hand side converges for $n \to \infty$, we infer that $\widehat{u}^n$, $\overline{u}^n$, and $\partial\phi(\overline{u}^n)$ are bounded in $W^{1,p}(0,T;H)$, $L^\infty(0,T;H)$, and $L^{p'}(0,T;H)$, respectively, and that $\phi(\overline{u}^n)$ is bounded. Hence, we can extract subsequences (not relabeled) such that, as $n \to \infty$,

$$
\begin{aligned}
\widehat{u}^n \to u & \quad \text{strongly in } \ C([0,T];H) \ \text{ and weakly in } \ W^{1,p}(0,T;H),\\
\overline{u}^n,\ \underline{u}^n \to u & \quad \text{strongly in } \ L^\infty(0,T;H),\\
\partial\phi(\overline{u}^n) \to \partial\phi(u) & \quad \text{weakly in } \ L^{p'}(0,T;H),
\end{aligned}
$$

where $\underline{u}^n(t) = u_{i-1}$ for all $t \in (t^n_{i-1}, t^n_i]$. These convergences and (5.8) allow us to pass to the limit inferior $n \to \infty$ in the last inequality in (5.10), giving

$$
\phi(u(t)) + \int_0^t \psi(u,u')\mathrm{d}r + \int_0^t \psi^*(u,-\partial\phi(u))\mathrm{d}r - \phi(u_0) \le 0
$$

for all $t \in [0,T]$. Eventually, the chain rule $(\phi \circ u)' = (\partial\phi(u), u')$ (now in $W^{1,p}(0,T;H)$) implies the energy equality (5.5). We conclude that u solves (5.1). Indeed

$$
\begin{aligned}
&\int_0^t \psi(u,u')\mathrm{d}r + \int_0^t \psi^*(u,-\partial\phi(u))\mathrm{d}r \le \phi(u_0) - \phi(u(t)) = -\int_0^t (\phi\circ u)'\mathrm{d}r\\
&\quad = \int_0^t (-\partial\phi(u), u')\mathrm{d}r \le \int_0^t \psi(u,u')\mathrm{d}r + \int_0^t \psi^*(u,-\partial\phi(u))\mathrm{d}r,
\end{aligned}
$$

so that all inequalities are actually equalities. □

5.2 GENERIC Flows

Another extension of the De Giorgi scheme concerns GENERIC flows (General Equations for Non-Equilibrium Reversible-Irreversible Coupling),

$$
u' = L\mathrm{D}E(u) - K\partial\phi(u) \quad \text{for a.e. } t \in (0,T), \quad u(0) = u_0. \tag{5.11}
$$

Here, ϕ plays the role of an entropy (up to the sign), being nonincreasing in time; $E : H \to \mathbb{R}$ is an energy, being conserved along trajectories; $K : H \to H$ is the so called Onsager operator, being linear, symmetric, and positive definite; and $L : H \to H$ is linear, symplectic, and antiselfadjoint ($L^* = -L$). One also assumes that the following compatibility conditions hold:

$$
L^*\partial\phi(u) = K^*\mathrm{D}E(u) = 0. \tag{5.12}
$$

The GENERIC formalism [22] is a systematic approach for the variational formulation of physical models and is particularly tailored to the unified treatment of coupled conservative and dissipative dynamics. As such, GENERIC has been applied to a variety of situations ranging from complex fluids [22], to dissipative quantum mechanics [35], to thermomechanics [34], and to the Vlasov-Fokker-Planck equation [16].

Let $\psi(u) = (K^{-1}u, u)/2$. The equivalent of (1.3) for GENERIC flows is

$$\phi(u(t)) + \int_0^t \psi(u' - \mathrm{L}\mathrm{D}E(u))\mathrm{d}r + \int_0^t \psi^*(-\partial\phi(u))\mathrm{d}r = \phi(u_0) \tag{5.13}$$

for all $t \in [0, T]$. Indeed, given the compatibility (5.12), equation (5.13) and the equation in (5.11) are equivalent:

$$\begin{aligned}(5.11) \;\Leftrightarrow\; & \psi(u' - \mathrm{L}\mathrm{D}E(u)) + \psi^*(-\partial\phi(u)) - \big(u' - \mathrm{L}\mathrm{D}E(u), -\partial\phi(u)\big) = 0 \text{ a.e.}\\ \Leftrightarrow\; & \frac{\mathrm{d}}{\mathrm{d}t}\phi \circ u + \psi(u' - \mathrm{L}\mathrm{D}E(u)) + \psi^*(-\partial\phi(u)) = 0 \text{ a.e.} \;\Leftrightarrow\; (5.13).\end{aligned}$$

The De Giorgi scheme can be extended to this case as well, by considering the functional

$$\overline{G}_i(u, v) := \phi(u) + \tau_i\psi\left(\frac{u - v}{\tau_i} - \mathrm{L}\mathrm{D}E(u)\right) + \tau_i\psi^*(-\partial\phi(u)) - \phi(v).$$

The existence of suitable De Giorgi solutions, namely $u_i \in D(\partial\phi)$ such that $\overline{G}_i(u_i, u_{i-1})$ is small enough, is still open. By *assuming* that there exists a sequence u_i fulfilling condition (4.6) written for $\overline{G}$ instead of G, a necessary condition for convergence is that the real function $(u, u') \mapsto \psi(u' - \mathrm{L}\mathrm{D}E(u)) + \psi^*(-\partial\phi(u))$ is lower semicontinuous, which follows, for instance, if ψ and ϕ are convex and E is C^1.

5.3 Curves of Maximal Slope in Metric Spaces

The variational interpretation of the De Giorgi scheme from (4.5) can be extended to evolutions in metric spaces. Let (X, d) be a complete metric space and $\phi : X \to [0, \infty]$ be lower semicontinuous. A locally absolutely continuous curve $u : [0, T] \to X$ is said to be a *curve of maximal slope* for the functional ϕ if $\phi \circ u$ is nonincreasing, $u(0) = u_0$, and

$$\phi(u(t)) + \frac{1}{2}\int_0^t |u'|^2(s)\mathrm{d}s + \frac{1}{2}\int_0^t |\partial\phi|^2(u(s))\,\mathrm{d}s = \phi(u_0) \quad \text{for all } t \in [0, T]. \tag{5.14}$$

It is beyond the purpose of this note to provide a comprehensive discussion of this notion and the corresponding theory. We refer the reader to the reference monograph by Ambrosio et al. [1] for details and limit ourselves in observing that (5.14) exactly corresponds to (1.3) upon replacing the norm of the time derivative with the *metric derivative*

$$|u'|(t) := \lim_{s \to t} \frac{d(u(s), u(t))}{|s - t|} \tag{5.15}$$

and the norm of the gradient of the functional by its *local slope* at $u \in D(\phi) := \{v \in X \,:\, \phi(v) < \infty\}$ as

$$|\partial\phi|(u) = \limsup_{v \to u} \frac{(\phi(u) - \phi(v))^+}{d(u, v)}.$$

Let us recall from [1, Theorem 1.1.2] that the limit in (5.15) exists almost everywhere for trajectories in the class

$$AC^2(0, T; X) := \Big\{ u : [0, T] \to X : \exists\, m \in L^2(0, T) : \forall\, 0 \le s < t \le T : \\ d(u(s), u(t)) \le \int_s^t m(r)\, \mathrm{d}r \Big\}.$$

A continuous curve $\gamma : [0, 1] \to X$ is called a *constant-speed geodesic* [1, Definition 2.4.2] if

$$d(\gamma(t), \gamma(s)) = d(\gamma(0), \gamma(1))|t - s| \quad \text{for all } s,\ t \in [0, 1].$$

Given $\lambda \in \mathbb{R}$, we say that a functional $\eta : X \to (-\infty, \infty]$ is λ-*geodesically convex* [1, Def. 2.4.3] if for all points $u_0,\ u_1 \in D(\eta) = \{x \in X \,:\, \eta(x) < \infty\}$, there exists a constant-speed geodesic γ connecting u_0 and u_1 such that

$$\eta(\gamma(t)) \le (1 - t)\eta(u_0) + t\eta(u_1) - \frac{\lambda}{2} t(1 - t) d^2(u_0, u_1) \quad \text{for all } t \in [0, 1].$$

The De Giorgi scheme in the metric setting is defined in terms of the functionals $\widehat{G}_i : D(|\partial\phi|) \times D(|\partial\phi|) \to (-\infty, \infty]$, given by

$$\widehat{G}_i(u, v) = \phi(u) + \frac{1}{2\tau_i} d^2(u, v) + \frac{\tau_i}{2} |\partial\phi|^2(u) - \phi(v),$$

where $D(|\partial\phi|) := \{v \in D(\phi) \,:\, |\partial\phi|(v) < \infty\}$. Given a sequence of partitions, we use as before the notation

$$\widehat{G}_i^n(u, v) = \phi(u) + \frac{1}{2\tau_i^n} d^2(u, v) + \frac{\tau_i^n}{2} |\partial\phi|^2(u) - \phi(v).$$

Our existence and convergence result reads as follows.

Theorem 5.2 (De Giorgi scheme for curves of maximal slope) *Let $\tau_* > 0$ be such that for all $v \in X$ and $\tau \le \tau_*$, the mapping $u \mapsto d^2(u, v)/(2\tau) + \phi(u)$ is $(1/\tau)$-geodesically convex. Assume that ϕ has compact sublevels, $u \mapsto |\partial\phi|(u) \in [0, \infty]$ is lower semicontinuous, and the chain-rule inequality*

$$|(\phi \circ u)'| \le |\partial\phi|(u)|u'| \quad \textit{a.e. in } (0, T) \tag{5.16}$$

holds for all $u \in AC^2(0, T; X)$ *such that* $|\partial\phi|(u)|u'| \in L^1(0, T)$. *Let* $u_{i-1} \in D(|\partial\phi|)$ *be given with* $|\partial\phi|(u_{i-1}) \neq 0$. *Then, for all* $\tau_i \leq \tau_*$, *there exists* $u_i \in D(|\partial\phi|)$ *with* $u_i \neq u_{i-1}$ *and*

$$\widehat{G}_i(u_i, u_{i-1}) \leq 0. \tag{5.17}$$

Let $u_i^n \in D(|\partial\phi|)$ *be such that* $u_0^n = u_0$ *and*

$$\sum_{i=1}^{N^n} \widehat{G}_i^n(u_i^n, u_{i-1}^n)^+ \to 0 \ \text{as} \ n \to \infty. \tag{5.18}$$

Then $\overline{u}^n(t) \to u(t)$ *for all* $t \in [0, T]$ *for a subsequence* (*which is not relabeled*), *where* $u \in AC^2(0, T; X)$ *is a curve of maximal slope for the functional* ϕ.

A comment on these assumptions is in order. The geodesic convexity of the squared distance corresponds to the assumption that the metric space is *nonpositively curved*. This is the case of Euclidean and Hilbert spaces, as well as Riemannian manifolds of nonpositive sectional curvature [1]. As regards the functional ϕ, let us recall Corollary 2.4.10 in [1], which ensures that $|\partial\phi|$ is lower semicontinuous and the chain-rule inequality holds whenever ϕ is 0-geodesically convex. In particular, the $(1/\tau)$-geodesical convexity of $u \mapsto d^2(u, v)/(2\tau) + \phi(u)$ follows when ϕ is 0-geodesically convex and the metric space is nonpositively curved [1, Remark 4.0.2]. The assumptions can be weakened by requiring compactness with respect to a weaker topology, which could allow for the extension of the result to reflexive Banach spaces [1, Remark 2.0.5].

Theorem 5.2 shows that

$$u_i \in \underset{u \in D(|\partial\phi|)}{\arg\min} \left(\phi(u) + \frac{1}{2\tau_i} d^2(u, u_{i-1}) + \frac{\tau_i}{2} |\partial\phi|^2(u) - \phi(u_{i-1}) \right) \tag{5.19}$$

for $i = 1, \ldots, N$ is an alternative approximation for curves of maximal slope with respect to the classical Euler scheme

$$u_i^e \in \underset{u \in D(|\partial\phi|)}{\arg\min} \left(\phi(u) + \frac{1}{2\tau_i} d^2(u, u_{i-1}) - \phi(u_{i-1}) \right) \tag{5.20}$$

for $i = 1, \ldots, N$. Note that the compactness of the sublevels of ϕ and the lower semicontinuity of $|\partial\phi|$ imply that solutions to the minimum problem (5.19) exist.

Proof of Theorem 5.2 Let u_i^e be a solution to (5.20). We will use the *slope estimate* [1, Lemma 3.1.3]

$$|\partial\phi|(u_i^e) \leq \frac{1}{\tau_i} d(u_i^e, u_{i-1}) \tag{5.21}$$

as well as the following consequence of the $(1/\tau)$-geodesic convexity of the functional $u \mapsto d^2(u, u_{i-1})/(2\tau) + \phi(u)$ [1, Theorem 4.1.2.ii] (choose $v = u = u_{i-1}$, $u_\tau = u_i^e$, and $\lambda = 0$ in the notation of the cited theorem):

$$\phi(u_i^e) + \frac{1}{\tau_i} d^2(u_i^e, u_{i-1}) - \phi(u_{i-1}) \le 0. \tag{5.22}$$

Let u_i solve (5.19). We deduce from the minimality of u_i and estimates (5.21) and (5.22) that

$$\begin{aligned}
\widehat{G}_i(u_i, u_{i-1}) &\le \widehat{G}_i(u_i^e, u_{i-1}) \\
&= \phi(u_i^e) + \frac{1}{2\tau_i} d^2(u_i^e, u_{i-1}) + \frac{\tau_i}{2} |\partial\phi|^2(u_i^e) - \phi(u_{i-1}) \\
&\le \phi(u_i^e) + \frac{1}{2\tau_i} d^2(u_i^e, u_{i-1}) + \frac{1}{2\tau_i} d^2(u_i^e, u_{i-1}) - \phi(u_{i-1}) \\
&= \phi(u_i^e) + \frac{1}{\tau_i} d^2(u_i^e, u_{i-1}) - \phi(u_{i-1}) \le 0,
\end{aligned}$$

and (5.17) follows. If $|\partial\phi|(u_{i-1}) \neq 0$, we have $\widehat{G}_i(u_i, u_{i-1}) \le 0 < \tau_i |\partial\phi|^2(u_{i-1})/2 = \widehat{G}_i(u_{i-1}, u_{i-1})$. Hence, $u_i \neq u_{i-1}$.

Next, let u_i^n fulfill relation (5.18). We take the sum for $i = 1, \ldots, m$:

$$\phi(u_m^n) + \sum_{i=1}^m \frac{d^2(u_i^n, u_{i-1}^n)}{2\tau_i^n} + \sum_{i=1}^m \frac{\tau_i^n}{2} |\partial\phi|^2(u_i^n) - \phi(u_0) = \sum_{i=1}^m \widehat{G}_i^n(u_i^n, u_{i-1}^n)^+. \tag{5.23}$$

As the right-hand side is bounded uniformly with respect to n, one can follow the proof of [1, Theorem 2.3.3] and extract a subsequence $\overline{u}^n$ (not relabeled) such that $\overline{u}^n \to u$ pointwise in $[0, T]$, $|(\overline{u}^n)'| \to |u'|$ weakly in $L^2(0, T)$, and for all $t \in [0, T]$,

$$\phi(u(t)) \le \liminf_{n\to\infty} \phi(\overline{u}^n(t)), \quad |\partial\phi|(u(t)) \le \liminf_{n\to\infty} |\partial\phi|(\overline{u}^n(t)).$$

Passing to the limit inferior in relation (5.23) leads to

$$\begin{aligned}
&\phi(u(t)) + \frac{1}{2}\int_0^t |u'|^2(s)\, \mathrm{d}s + \frac{1}{2}\int_0^t |\partial\phi|^2(u(s))\, \mathrm{d}s - \phi(u_0) \\
&\quad \le \liminf_{n\to\infty} \sum_{i=1}^{N^n} \widehat{G}_i^n(u_i^n, u_{i-1}^n)^+ = 0
\end{aligned} \tag{5.24}$$

for all $t \in [0, T]$. Eventually, we use (5.24) and the chain-rule inequality (5.16) to obtain

$$\begin{aligned}
&\frac{1}{2}\int_0^t |u'|^2(s)\, \mathrm{d}s + \frac{1}{2}\int_0^t |\partial\phi|^2(u(s))\, \mathrm{d}s \le -\phi(u(t)) + \phi(u_0) \\
&\quad = -\int_0^t (\phi \circ u)'(s)\, \mathrm{d}s \le \int_0^t |\partial\phi|(u(s))|u'|(s)\, \mathrm{d}s \\
&\quad \le \frac{1}{2}\int_0^t |u'|^2(s)\, \mathrm{d}s + \frac{1}{2}\int_0^t |\partial\phi|^2(u(s))\, \mathrm{d}s
\end{aligned}$$

so that all inequalities are actually equalities and u solves (5.14). □

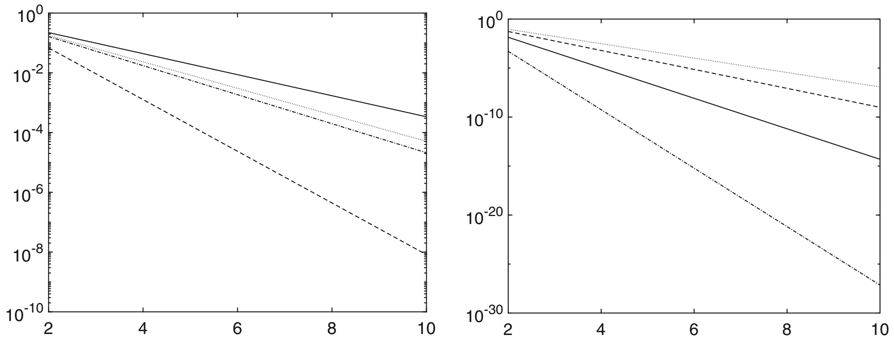

Fig. 3 Behavior of the potential $\phi(u) = \lambda u^2/2$ with $\lambda = 1$ along iterations for $\tau\lambda = 0.5$ (left) $\tau\lambda = 5$ (right) and in semilog scale; e_i (solid line), g_i (dotted line), v_i (dash-dotted line), and u_i (dashed line)

6 A Comparison Between the Gonzales and the De Giorgi Scheme

Let us close this discussion by presenting a direct comparison of the output of the Gonzales and the De Giorgi schemes in the case of the single scalar ODE problem

$$u' + \lambda u = 0, \quad u(0) = 1$$

for $\lambda > 0$ given. This corresponds to the gradient flow of the uniformly convex potential $\phi(u) = \lambda u^2/2$.

Given the time steps τ_i, one readily finds the solutions e_i of the Euler scheme, g_i of the Gonzales scheme, u_i of the De Giorgi scheme $G_i(u_i, u_{i-1}) = 0$, and v_i of the De Giorgi scheme min $G_i(\cdot, v_{i-1})$, see (4.10), as

$$e_i = \prod_{j=0}^{i} \frac{1}{1 + \tau_i\lambda}, \quad g_i = \prod_{j=0}^{i} \frac{2 - \tau_i\lambda}{2 + \tau_i\lambda},$$

$$u_i = \prod_{j=0}^{i} \frac{1 - (\tau_i\lambda)^{3/2}}{1 + \tau_i\lambda + (\tau_i\lambda)^2}, \quad v_i = \prod_{j=0}^{i} \frac{1}{1 + \tau_i\lambda + (\tau_i\lambda)^2}.$$

Observe that u_i and g_i oscillate in sign for $\tau_i\lambda > 1$ and $\tau_i\lambda > 2$, respectively, whereas e_i and v_i are always positive.

Figure 2 records the performance of the methods in terms of order of convergence for $\lambda = 1$. In accordance to Theorem 3.2(iii), the Gonzales scheme is of order τ^2, the Euler and the De Giorgi scheme v_i are of order τ, and the De Giorgi scheme u_i is of order $\tau^{1/2}$. We recall that the second-order convergence of the Gonzales scheme is limited to the scalar case only. In more than one dimension, the Gonzales scheme is only of order τ if condition (3.3) does not hold, see Fig. 1.

Let us now assess the performance of the Gonzales and the De Giorgi schemes in order to compute the minimum of ϕ. We illustrate in Fig. 3 the reduction of the potential along iterations. The performance of the Gonzales scheme and the De Giorgi

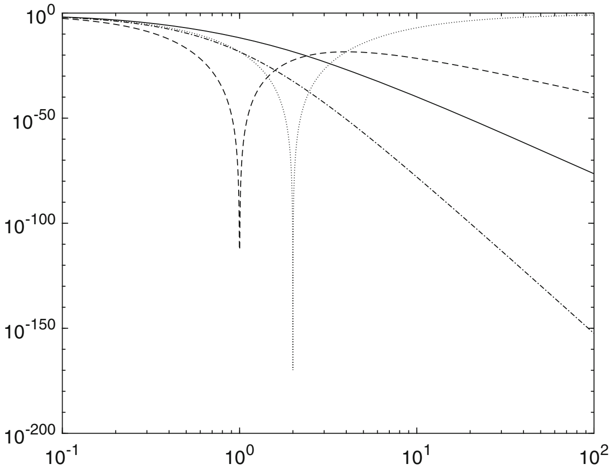

Fig. 4 Reduction of the potential $\phi(u) = \lambda u^2/2$ with $\lambda = 1$ after 20 iterations as a function of $\tau\lambda$ in log-log scale; e_i (solid line), g_i (dotted line), v_i (dash-dotted line), and u_i (dashed line)

scheme u_i depends on the size of $\tau\lambda$. The De Giorgi scheme v_i outperforms the Euler method regardless of the size of $\tau\lambda$. This fact is additionally illustrated in Fig. 4, where we record the effect of the size of $\tau\lambda$ on the effective reduction of the potential after a fixed number of iterations. One can observe how the Gonzales scheme delivers a strong energy reduction for $\tau\lambda < 2$ and virtually none for large $\tau\lambda$. Compared to the Euler method, both De Giorgi schemes are more energy-reducing for large $\tau\lambda$. While the De Giorgi scheme u_i shows some singular behavior at $\tau\lambda = 1$, the De Giorgi scheme v_i delivers an energy reduction regardless of the size of $\tau\lambda$, eventually outperforming the Euler method.

Acknowledgements This research is supported by Austrian Science Fund (FWF) project F 65. AJ is partially supported by the FWF Projects P 27352, P 30000, and W 1245. US is partially supported by the FWF Projects I 2375 and P 27052 and by the Vienna Science and Technology Fund (WWTF) through Project MA14-009. Some interesting remarks from the referees in the direction of optimization algorithms are gratefully acknowledged.

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2nd edn. Birkhäuser Verlag, Basel (2008)
2. Attouch, H., Buttazzo, G., Michaille, G.: Variational analysis in Sobolev and BV spaces. Applications to PDEs and optimization. Second edition. MOS-SIAM Series on Optimization, 17. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA (2014)
3. Attouch, H., Cabot, A.: Asymptotic stabilization of inertial gradient dynamics with time dependent viscosity. J. Differ. Equ. **263**, 5412–5458 (2017)
4. Attouch, H., Chbani, Z., Peypouquet, J., Redont, P.: Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. Math. Program. Ser. B **168**, 123–175 (2018)
5. Bagirov, A., Karasözen, B., Sezer, M.: Discrete gradient method: derivative-free method for nonsmooth optimization. J. Optim. Theory Appl. **137**, 317–334 (2008)

6. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York (2011)
7. Bertsekas, D.P.: Convex Optimization Algorithms. Athena Scientific, Belmont, MA (2015)
8. Bot, R.I., Csetnek, E.R., László, S.C.: Approaching nonsmooth nonconvex minimization through second order proximal-gradient dynamical systems. J. Evol. Equ. **18**, 1291–1318 (2018)
9. Brézis, H.: Monotonicity methods in Hilbert spaces and some application to nonlinear partial differential equations. In: Contributions to Nonlinear Functional Analysis. Proceedings of Symposia in University of Wisconsin, Madison, pp. 101–156. Academic Press, New York (1971)
10. Brézis, H.: Opérateurs Maximaux Monotones et Semi-groupes de Contractions Dans les Espaces de Hilbert. Number 5 in North Holland Math Studies. North-Holland, Amsterdam (1973)
11. Cabot, A., Engler, H., Gadat, S.: On the long time behavior of second order differential equations with asymptotically small dissipation. Trans. Am. Math. Soc. **361**, 5983–6017 (2009)
12. Celledoni, E., Grimm, V., McLachlan, R., McLaren, D., O'Neale, D., Owren, B., Quispel, G.: Preserving energy resp. dissipation in numerical PDEs using the "average vector field" method. J. Comput. Phys. **231**, 6770–6789 (2012)
13. Condette, N., Melcher, C., Süli, E.: Spectral approximation of pattern-forming nonlinear evolution equations with double-well potentials of quadratic growth. Math. Comput. **80**, 205–223 (2011)
14. Crandall, M.G., Pazy, A.: Semi-groups of nonlinear contractions and dissipative sets. J. Funct. Anal. **3**, 376–418 (1969)
15. De Giorgi, E., Marino, A., Tosques, M.: Problems of evolution in metric spaces and maximal decreasing curve. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat. **68**, 180–187 (1980)
16. Duong, M.H., Peletier, M.A., Zimmer, J.: GENERIC formalism of a Vlasov–Fokker–Planck equation and connection to large-deviation principles. Nonlinearity **26**, 2951–2971 (2013)
17. Eyre, D.: Unconditionally gradient stable time marching the Cahn–Hilliard equation. In: Computational and Mathematical Models of Microstructural Evolution (San Francisco, CA, 1998), pp. 39–46, Materials Research Society Symposium Proceedings, p. 529 (1998)
18. Furihata, D., Matsuo, T.: Discrete Variational Derivative Method. CRC Press, Boca Raton (2011)
19. García Orden, J.C., Romero, I.: Energy–Entropy–Momentum integration of discrete thermo-viscoelastic dynamics. Eur. J. Mech. Solids **32**, 76–87 (2012)
20. Gonzalez, O.: Time integration and discrete Hamiltonian systems. J. Nonlinear Sci. **6**, 449–467 (1996)
21. Gonzalez, O.: Exact energy-momentum conserving algorithms for general models in nonlinear elasticity. Comp. Methods Appl. Mech. Eng. **190**, 1763–1783 (2000)
22. Grmela, M., Öttinger, H.C.: Dynamics and thermodynamics of complex fluids. I. Development of a general formalism. Phys. Rev. E **56**, 6620–6632 (1997)
23. Groß, M., Bartelt, M., Betsch, P.: Structure-preserving time integration of non-isothermal finite viscoelastic continua related to variational formulations of continuum dynamics. Comput. Mech. **62**, 123–150 (2017)
24. Hairer, E., Lubich, C.: Energy-diminishing integration of gradient systems. IMA J. Numer. Anal. **34**, 452–461 (2014)
25. Hairer, E., Lubich, C., Wanner, G.: Geometric Numerical Integration, 2nd edn. Springer, Berlin (2006)
26. Jüngel, A., Milišić, J.-P.: Entropy dissipative one-leg multistep time approximations of nonlinear diffusive equations. Numer. Method Partial Differ. Equ. **31**, 1119–1149 (2015)
27. Jüngel, A., Schuchnigg, S.: Entropy-dissipating semi-discrete Runge–Kutta schemes for nonlinear diffusion equations. Commun. Math. Sci. **15**, 27–53 (2017)
28. Kōmura, Y.: Nonlinear semi-groups in Hilbert space. J. Math. Soc. Jpn. **19**, 493–507 (1967)
29. Legendre, G., Turinici, G.: Second-order in time schemes for gradient flows in Wasserstein and geodesic metric spaces. C. R. Math. Acad. Sci. Paris **355**, 345–353 (2017)
30. Liu, K., Wu, X., Shi, W.: A linearly-fitted conservative (dissipative) scheme for efficiently solving conservative (dissipative) nonlinear wave PDES. J. Comput. Math. **35**, 780–800 (2017)
31. Matthes, D., Plazotta, S.: A variational formulation of the BDF2 method for metric gradient flows. ESAIM Math. Model. Numer. Anal. **53**, 145–172 (2019)
32. McLachlan, R., Quispel, G.: Discrete gradient methods have an energy conservation law. Discret. Contin. Dyn. Sys. **34**, 1099–1104 (2014)
33. McLachlan, R., Quispel, G., Robidoux, N.: Geometric integration using discrete gradients. Phil. Trans. R. Soc. Lond. A **357**, 1021–1045 (1999)
34. Mielke, A.: Formulation of thermoelastic dissipative material behavior using GENERIC. Contin. Mech. Thermodyn. **23**, 233–256 (2011)

35. Mielke, A.: Dissipative quantum mechanics using GENERIC. In: Proceedings of the conference on Recent Trends in Dynamical Systems, vol. 35 of Proceedings in Mathematics & Statistics, Springer, pp. 555–585 (2013)
36. Nesterov, YuE: A method for solving the convex programming problem with convergence rate $O(1/k^2)$. Dokl. Akad. Nauk SSSR **269**, 543–547 (1983)
37. Nochetto, R., Savaré, G., Verdi, C.: A posteriori error estimates for variable time-step discretization of nonlinear evolution equations. Commun. Pure Appl. Math. **53**, 525–589 (2000)
38. Peypouquet, J.: Convex Optimization in Normed Spaces. Theory, Methods and Examples. Springer Briefs in Optimization. Springer, Cham (2015)
39. Quispel, G., McLaren, D.: A new class of energy-preserving numerical integration methods. J. Phys. A: Math. Theor. **41**, 045206 (2008)
40. Romero, I.: Thermodynamically consistent time-stepping algorithms for non-linear thermomechanical systems. Int. J. Numer. Methods Eng. **79**, 706–732 (2009)
41. Romero, I.: Algorithms for coupled problems that preserve symmetries and the laws of thermodynamics. Part I: monolithic integrators and their application to finite strain thermoelasticity. Comput. Methods Appl. Mech. Eng. **199**, 1841–1858 (2010)
42. Romero, I.: Algorithms for coupled problems that preserve symmetries and the laws of thermodynamics. Part II: fractional step methods. Comput. Methods Appl. Mech. Eng. **199**, 2235–2248 (2010)
43. Rossi, R., Savaré, G.: Gradient flows of non convex functionals in Hilbert spaces and applications. ESAIM Control Optim. Calc. Var. **12**, 564–614 (2006)
44. Sato, S., Matsuo, T., Suzuki, H., Furihata, D.: A Lyapunov-type theorem for dissipative numerical integrators with adaptive time-stepping. SIAM J. Numer. Anal. **53**, 2505–2518 (2015)
45. Shi, B., Du, S.S., Jordan, M.I., Su, W.J.: Understanding the acceleration phenomenon via high-resolution differential equations (submitted). arXiv:1810.08907 (2018)