

Trajectory convergence from coordinate-wise decrease of general convex energy functions

Julien M. Hendrickx^a, Balázs Gerencsér^b

^a*ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium*

^b*Alfréd Rényi Institute of Mathematics, Budapest, Hungary,
Eötvös Loránd University, Department of Probability and Statistics, Budapest, Hungary*

Abstract

We consider arbitrary trajectories subject to a coordinate-wise energy decrease: the sign of the derivative of each entry is never the same as that of the corresponding entry of the gradient of some convex energy function. We show that this simple condition guarantees convergence to a point, to the minimum of the energy functions, or to a set where its Hessian has very specific properties. This extends and strengthens recent results that were restricted to quadratic energy functions.

Key words: Asymptotic Stabilization, Lyapunov methods, Convex energy function, Single Trajectory Asymptotics, Networked Control Systems

1 Introduction

We consider the convergence properties of a trajectory $y : \mathbb{R}^+ \rightarrow \mathbb{R}^n$, $t \mapsto y(t)$ whose evolution is constrained by a convex energy function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, $x \mapsto V(x)$ via the set of inequalities

$$\dot{y}_i \frac{\partial V}{\partial x_i} \Big|_{y(t)} \leq 0, \quad \forall t \geq 0, i = 1, \dots, n, \quad (1)$$

i.e. the derivative of a coordinate y_i of y and the corresponding coordinate of the gradient of V at y always have opposite sign if they are both nonzero. We note that we will always use the letter y when referring to a trajectory or its accumulation points, and the letter x for the points in the ambient space \mathbb{R}^n .

For energy functions of the form $V = x^T Q x$, condition (1) was shown in our recent work [8] to be often sufficient for convergence of y . This was motivated by a platoon cooperative control application involving dead-zone control and bounded arbitrary disturbance. It allowed in particular solving a conjecture on the convergence of such systems [4], and a related problem of consensus under bounded disturbance [2, 5]. In this work, we extend

these results to general convex functions V , and characterize more precisely the alternative long-term behavior of the trajectory when convergence is not guaranteed. We also explore the tightness of our conditions.

We stress that the trajectory y is not assumed to be generated by a vector field or a system of differential equations. It can be completely arbitrary provided it satisfies the constraints (1). By contrast, a large proportion of the convergence results based on decrease of energy functions rely on variations or extensions of Lyapunov-Krasovskii-LaSalle Theorems [9], and typically assume that trajectories follow some ordinary differential equation such as $\dot{y}(t) = f(y(t), t)$ or $\dot{y}(t) = f(y(t))$ for an f satisfying some (uniform) continuity conditions [1, 3]. For example, LaSalle theorem guarantees (under some conditions) the convergence of $\dot{y} = f(y)$ to an invariant set, but not to a single point, provided $f(x)^T \nabla V(x) \leq 0$ everywhere [10]. Convergence to 0 can then be guaranteed under the additional assumption that $\frac{d}{dt} V(y(t)) = f(y(t))^T \nabla V(y(t))$ is not uniformly zero *along any trajectory* other than that staying at 0 [13]. For more detail on various cases of unforced systems, we refer the reader to [11] as a starting point.

Vector fields $f(x, t)$ over the state space may not be naturally available in systems whose evolution is driven by external elements. Think of discrete communications in cyber-physical applications, systems designed to be

Email addresses: julien.hendrickx@uclouvain.be (Julien M. Hendrickx), gerencser.balazs@renyi.hu (Balázs Gerencsér).

robust to adversarial input signals that could depend on the trajectory and its history, or systems involving some random decisions (though more complex descriptions may be available, see [6,14]). Similarly, many modern control laws are not easily described by a continuous field f , think, e.g., of event-triggered or self-triggered mechanisms [7,12]. Hence it is desirable to have results guaranteeing the convergence of a *single trajectory based on properties satisfied along that specific trajectory* without assuming or constructing a corresponding vector field, nor speculating about the properties of potential other trajectories. Currently available results for single trajectories require a sufficiently negative decrease, e.g., $\frac{d}{dt}V(y(t)) \leq -\lambda V(y(t))$ for some positive λ , which allows guaranteeing convergence to the minimum of V at a certain rate, see again [9]. This precludes their use when no such uniform condition can be guaranteed, or for situations where the rate of convergence cannot be known, which could happen for example if parts of the system can occasionally pause. On the other hand, simply requiring $\frac{d}{dt}V(y(t)) < 0$ does not imply convergence, as can be verified on the simple two-dimensional example $y(t) = (1 + e^{-t})(\cos t, \sin t)$ and $V(x) = \|x\|_2$. As a source of intuition at a very informal level, one could say that the condition $\frac{d}{dt}V(y(t)) < 0$ implies the decrease and convergence of the energy V along the trajectory, but allows for persistent significant energy transfer between the different coordinates. By contrast, our condition (1) forces the decrease of energy on every coordinate. This remains at the level of intuition though, as the energy V in general cannot be separated along the different coordinates.

Our paper is organized as follows: We state our main convergence result in Section 2, together with convenient Corollaries specializing it. We study its tightness in Section 3, with examples showing that our conditions cannot simply be removed. The main proof is presented in Section 4, together with the intuition on how the elements are built together. We draw conclusions and discuss potential continuations and open problems in Section 5.

2 Main results

We first present our most general result with minimal assumptions, thus allowing for most possibilities for the asymptotic behavior.

Theorem 1 *Let $V : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex twice differentiable function with a locally Lipschitz Hessian and $y : \mathbb{R}^+ \rightarrow \mathbb{R}^n$ a trajectory that is absolutely continuous, also implying that $\dot{y}(t)$ exists almost everywhere. Suppose that where it exists,*

$$\dot{y}_i(t) \frac{\partial V}{\partial x_i} \Big|_{y(t)} \leq 0 \quad \forall t \geq 0, i = 1, \dots, n. \quad (2)$$

Then, at least one of the following three conditions holds

- (a) *y converges;*
- (b) *for every accumulation point \bar{y} , there holds $\nabla V(\bar{y}) = 0$ and hence $\bar{y} \in \arg \min_x V(x)$;*
- (c) *for every accumulation point \bar{y} , the kernel of $\nabla^2 V(\bar{y})$ contains a nonzero vector with a zero coordinate.*

We note that this theorem allows the possibility of $y(t)$ having no accumulation point, in which case (b) and (c) are trivially satisfied. This is for example the case for $y(t) = -t$ with $V(x) = \exp(x)$ on \mathbb{R} .

The proof of Theorem 1 is presented in Section 4. We now deduce useful special cases by strengthening some assumptions, first in view of positive semidefinite Hessians, and then for positive definite ones.

Corollary 2 *Under the conditions of Theorem 1, assume that the trajectory y is bounded and the kernel of the Hessian $\nabla^2 V(x)$ does not have any nonzero vector with zero component for any x . Then either y converges to a point y^* , or it converges to the set $\arg \min_x V(x)$.*

PROOF. Since y is bounded, it must have at least one accumulation point. Thus the claim of Theorem 1 is not empty. The current conditions explicitly exclude case (c), the remaining conditions (a) and (b) correspond to the statement of the Corollary. \square

Corollary 2 can for example be applied with any function of the form $V(x) = \tilde{V}(\Pi_v x)$ where Π_v is the orthogonal projection onto a space orthogonal to a vector v with $v_i \neq 0$ for every i , and \tilde{V} is strongly convex.

Corollary 3 *Under the conditions of Theorem 1,*

- (i) *if V is strongly convex, then y converges;*
- (ii) *if $\nabla^2 V(x) \succ 0$ for every x and V admits a minimum x^* , then y converges.*

PROOF. We prove (ii), of which (i) is a particular case. We suppose without loss of generality that $x^* = 0$ and $V(x^*) = 0$.

We first show that $\bar{y}(t)$ is bounded. Since $\nabla^2 V(0) \succ 0$ and $\nabla^2 V(x)$ is continuous, there exists $\epsilon, \lambda > 0$ such that $\nabla(x) \succ \lambda I$ for all $x \in B(0, \epsilon)$. By convexity, $v^T \nabla V(sv) \geq 0$ for all $s \geq 0$. Moreover, if $s \geq \epsilon$ we can

develop the following bound

$$\begin{aligned} v^T \nabla V(sv) &= \int_0^s v^T \nabla^2 V(lv) v dl \\ &\geq \int_0^\epsilon v^T \nabla^2 V(lv) v dl \\ &\geq \int_0^\epsilon \lambda \|v\|^2 dl = \epsilon \lambda. \end{aligned}$$

Consequently, for any $q \geq \epsilon$ there holds

$$V(qv) = \int_0^q v^T \nabla V(vs) ds \geq \int_0^\epsilon 0 ds + \int_\epsilon^q \epsilon \lambda ds = (q - \epsilon) \epsilon \lambda.$$

In particular, for any $x \notin B(0, \epsilon)$, we have $V(x) \geq (\|x\| - \epsilon) \epsilon \lambda$. Since $V(y(t))$ is non-increasing as $\dot{y}^T \nabla V(y(t)) \leq 0$ follows from (2), this implies $(\|y(t)\| - \epsilon) \epsilon \lambda \leq V(y(t)) \leq V(y(0))$, and hence the boundedness of y .

This allows applying Corollary 2. Since the set of minimizers of V is shown by standard argument to consist of a single point x^* in the current case, both possible conclusions of this Corollary imply the convergence of y . \square

Note that the assumption of the existence of a minimum x^* in condition (ii) of Corollary 3 is needed, as the other part of the assumption does not necessarily imply the existence of a minimum nor the boundedness of the level sets, as it is the case with $V(x) = \exp(x)$ on \mathbb{R} mentioned above.

3 Tightness

We now show that assumptions of Theorem 1 allow for situations where only condition (b) or (c) holds demonstrating thus that these situations cannot be excluded without additional assumptions.

Example 1 (Condition (b)) Let $V(x) : \mathbb{R}^2 \rightarrow \mathbb{R}$, $V(x) = (x_2 - x_1)^2$, and consider the trajectory $y(t) = (\sin(t), \sin(t))$.

We have $\nabla V(x) = 2(x_1 - x_2, x_2 - x_1)^\top$ and thus $\nabla V(y(t)) = 0$, so that our assumption (2) is trivially satisfied and Theorem 1 applies. The trajectory y does not converge, so condition (a) does not hold. Now, the set of accumulation points of y is $\{(a, a) : a \in [-1, 1]\}$. On these points (as everywhere in \mathbb{R}^2), the kernel of $\nabla^2 V$ is always $\text{span}\{(1, 1)\}$, which contains thus no nonzero vector with a zero coordinate, so condition (c) does not hold either. Hence only condition (b) applies, and indeed $\nabla V(\bar{y}) = 0$ holds for all accumulation points \bar{y} .

Example 2 (Condition (c)) Let $C = [-1, 1]^2 \subset \mathbb{R}^2$, and $V(x) = d^4(x, C)$, i.e. the fourth power of the Euclidean distance to C . Consider the trajectory $y(t) = (2 + e^{-t}, \sin(t))$.

V is convex, and one can verify that its Hessian is locally Lipschitz. Besides, y remains in $[1, \infty) \times [-1, 1]$, on which $d(x, C) = (x_1 - 1)$ and hence $V(x) = (x_1 - 1)^4$. Furthermore the differentials in this region are

$$\nabla V(x) = \begin{pmatrix} 4(x_1 - 1)^3 \\ 0 \end{pmatrix} \quad \nabla^2 V(x) = \begin{pmatrix} 12(x_1 - 1)^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

In particular $\nabla V(y(t)) = (4(e^{-t} + 1)^3, 0)^\top$, which together with $\dot{y}(t) = (-e^{-t}, \cos(t))$ implies that our assumption (2) is satisfied, and Theorem 1 applies. Again, the trajectory does not converge, so condition (a) does not hold. Its set of accumulation points \bar{y} is $\{(2, a) : a \in [-1, 1]\}$, and on such points $\nabla V(\bar{y}) = (4, 0)^\top \neq 0$, so conditions (b) does not hold. Hence solely condition (c) holds, and one can indeed verify that the kernel of the Hessian at those points contains the vector $(0, 1)^\top$.

4 Proof of Theorem 1

4.1 Introduction and proof structure

For the ease of reading, we will slightly abuse notations and use ∇V_i to denote $\frac{\partial V}{\partial x_i}$ and $\nabla V_i(z)$ to denote $\frac{\partial V}{\partial x_i}|_z$.

We first observe that, although we did not assume V to be bounded from below, this assumption is automatically satisfied along the trajectory if there exists an accumulation point.

Lemma 4 Under the assumptions of Theorem 1, if y admits an accumulation point \bar{y} , then $V(y(t)) \geq V(\bar{y})$ for all t , and $\lim_{t \rightarrow \infty} V(y(t)) = V(\bar{y})$.

PROOF. It follows from assumption (2) that

$$\frac{d}{dt} V(y(t)) = \sum_{i=1}^n \nabla V_i(y(t)) \dot{y}_i(t) \leq 0,$$

implying that $V(y(t))$ is non-increasing. Since $y(t)$ gets arbitrary close to \bar{y} for arbitrarily large times, the continuity of V implies $\limsup_{t \rightarrow \infty} V(y(t)) \geq V(\bar{y})$, and thanks to the monotonicity of $V(\bar{y}(t))$ we have then

$$\inf_t V(y(t)) = \lim_{t \rightarrow \infty} V(y(t)) = \limsup_{t \rightarrow \infty} V(y(t)) \geq V(\bar{y}). \quad \square$$

We now define

$$K_i := \{x : \nabla V_i(x) = 0\},$$

the set on which the i^{th} coordinate of the gradient of V cancels. These sets are closed by continuity of ∇V . We say that an accumulation point \bar{y} is *locally K -minimal* if there is a non-trivial ball centered on \bar{y} containing no accumulation point that belongs to a smaller number of K_i than \bar{y} . We first prove in Section 4.2 the result for locally K -minimal accumulation points, and will extend it to the general case in Section 4.3 using topological arguments.

The intuition behind our proof is the following. In the non-trivial case of the theorem where the trajectory y does not converge but admits a (K -minimal) accumulation point \bar{y} , this trajectory must repeatedly approach \bar{y} and then leave it at a non-vanishing distance. We will exploit this to define a “direction” v that is (asymptotically) followed infinitely often when the trajectory leaves \bar{y} . We will argue that for those i for which $\bar{y} \in K_i$, there must hold $(Hv)_i = 0$ with H the Hessian of V at \bar{y} , because otherwise there would be an accumulation point of the form $\bar{y} + \delta v$ at which $\nabla V_i \neq 0$, i.e. that does not belong to K_i , contradicting the local K -minimality of \bar{y} . We will also argue that for those i for which $\bar{y} \notin K_i$ i.e. $\nabla V_i \neq 0$, there must hold $v_i = 0$, for otherwise, following v would result in an impossible repeated decrease of energy $\nabla V(\bar{y})v \leq \nabla V_i(\bar{y})v_i < 0$ (where we use our assumption (2)). Hence there we will have $v^T H v = \sum_i v_i (Hv)_i = 0$, i.e. the direction v is in the kernel of H . The analysis of the structure of this v will then give cases (b) and (c).

4.2 K -minimal accumulation points

Proposition 5 *Under the assumptions of Theorem 1, if y does not converge, every locally K -minimal accumulation point \bar{y} satisfies (at least) one of the following conditions:*

- (b') $\nabla V(\bar{y}) = 0$ and $\bar{y} \in \arg \min_x V(x)$.
- (c') the kernel of $\nabla^2 V(\bar{y})$ contains a nonzero vector with a zero coordinate.

We suppose y does not converge and fix a locally K -minimal accumulation point \bar{y} (in the absence of such point, the claim trivially holds). We may re-index the coordinates without loss of generality in such a way that \bar{y} belongs to K_1, \dots, K_k and not to the $n - k$ other K_i , with k potentially equal to 0. This choice and the local K -minimality of \bar{y} imply that the two following conditions are satisfied for any sufficiently small ϵ , and hence we assume them to be satisfied in the sequel for the values of ϵ considered.

- (i) $\bar{B}(\bar{y}, 3\epsilon) \cap K_i = \emptyset$ for $i > k$, where \bar{B} denotes the closed ball.
- (ii) there are no accumulation points on less than k sets K_i within $\bar{B}(\bar{y}, 3\epsilon)$.

We first show that locally, the trajectory will be asymptotically constrained towards the k kernel spaces

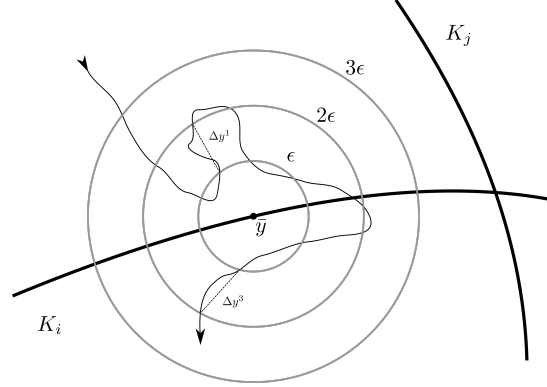


Figure 1. Representation of different constructions in the proof of Proposition 5 for a given accumulation point \bar{y} : (i) the sets K_i , with \bar{y} belonging to one of them, K_i , so that $k = 1$; (ii) the three balls of radius ϵ , 2ϵ and 3ϵ centered on \bar{y} , which do not intersect K_j ; and (iii) the trajectory y and some vectors Δy^m connecting each time a point of the trajectory at distance ϵ from \bar{y} to a subsequent point at distance 2ϵ from \bar{y} (these points are respectively $y(t_1^m)$ and $y(t_2^m)$, and are not represented here)

K_1, \dots, K_k prescribed by \bar{y} . In the representation of Figure 1, this means that y will approach closely K_i in the long run for $i = 1, \dots, k$.

Since $y(t)$ does not converge to \bar{y} , then for ϵ small enough, we can find arbitrary large t for which $y(t) \notin B(\bar{y}, 2\epsilon)$. On the other hand, since \bar{y} is an accumulation point, we can also find arbitrary large t for which $y(t) \in B(\bar{y}, \epsilon)$. Hence, there exists a sequence of disjoint intervals $[t_1^m, t_2^m]$ with $t_1^m, t_2^m \rightarrow \infty$ such that $y(t_1^m) \in \partial B(\bar{y}, \epsilon)$, $y(t_2^m) \in \partial B(\bar{y}, 2\epsilon)$ and $y([t_1^m, t_2^m]) \in \bar{B}(\bar{y}, 2\epsilon)$. We let then $\Delta y^m = y(t_2^m) - y(t_1^m)$, be the vector linking the beginning and end of these pieces of trajectories, and already observe that $\|\Delta y^m\| \in [2\epsilon, 4\epsilon]$. See again Figure 1 for a representation.

Since these vectors live in a compact set, they have an accumulation point Δy^ϵ , still with $\|\Delta y^\epsilon\| \in [2\epsilon, 4\epsilon]$. This vector can be interpreted as a direction in which the trajectory repeatedly leaves the small neighborhood of \bar{y} . We will show that $(\Delta y^\epsilon)^T H \Delta y^\epsilon = O(\epsilon^3)$ for H the Hessian of V at \bar{y} . For this purpose we show that $|(H \Delta y^\epsilon)_i| = O(\epsilon^2)$ for $i \leq k$ and $\Delta y_i^\epsilon = 0$ for $i > k$.

Claim 1: For $i \leq k$, for any increasing diverging sequence of times t^m such that $y(t^m) \in \bar{B}(\bar{y}, 3\epsilon)$, there holds

$$\lim_{m \rightarrow \infty} d(y(t^m), K_i) = 0 \text{ and } \lim_{m \rightarrow \infty} \nabla V_i(y(t^m)) = 0$$

In particular $\nabla V_i(y(t_2^m)) \rightarrow 0$ and $\nabla V_i(y(t_1^m)) \rightarrow 0$.

PROOF. If $\lim_{m \rightarrow \infty} d(y(t^m), K_i) = 0$ did not hold, there would be an infinite subsequence $\bar{y}(t^{m'})$ at a dis-

tance larger than $\delta > 0$ from K_i , which would admit an accumulation point $y^\delta \in \bar{B}(\bar{y}, 3\epsilon)$ since $y(t^m)$ remains in that compact set. Moreover, y^δ would not belong to any K_j for $j > k$ since $\bar{B}(\bar{y}, 3\epsilon)$ does not intersect with any such K_j . Hence we would have an accumulation point of y belonging to less than k sets K_i in contradiction with \bar{y} being K -minimal. The second part of the claim follows then by continuity of the gradient.¹ \square

We now show how Claim 1 implies the direction Δy^ϵ is “not too far” from the kernel of the first k rows of the Hessian of V .

Claim 2: Let $H = \nabla^2 V(\bar{y})$. For any $i \leq k$ there holds $|(H\Delta y^\epsilon)_i| \leq C\epsilon^2$, for some C possibly depending on \bar{y} but not on ϵ .

PROOF. We prove

$$\limsup_{m \rightarrow \infty} |(H\Delta y^m)_i| \leq C\epsilon^2, \quad (3)$$

which implies the result by definition of Δy^ϵ as an accumulation point of Δy^m . For this purpose we first show that the difference of the gradient $\nabla V(y(t_2^m)) - \nabla V(\bar{y})$ can be approximated by $H\Delta y^m$ up to $O(\epsilon^2)$. Indeed, we can write $\nabla V(y(t_2^m)) - \nabla V(\bar{y})$ as the following integral

$$\begin{aligned} & \int_{s=0}^1 \nabla^2 V(\bar{y} + (y(t_2^m) - \bar{y})s) (y(t_2^m) - \bar{y}) ds \\ &= \int_{s=0}^1 H(y(t_2^m) - \bar{y}) ds \\ &+ \int_{s=0}^1 (\nabla^2 V(\bar{y} + (y(t_2^m) - \bar{y})s) - H) (y(t_2^m) - \bar{y}) ds \end{aligned}$$

Since the Hessian is assumed to be locally Lipschitz continuous, we have, for a Lipschitz constant $L(\bar{y})$,

$$\begin{aligned} \|\nabla^2 V(\bar{y} + (y(t_2^m) - \bar{y})s) - H\| &\leq L(\bar{y})s \|y(t_2^m) - \bar{y}\| \\ &= O(\epsilon), \end{aligned}$$

with the implicit constant only depending on \bar{y} . Hence, slightly abusing the $O(\epsilon)$ notation for the sake of conciseness, there holds

$$\begin{aligned} \nabla V(y(t_2^m)) - \nabla V(\bar{y}) &= H(y(t_2^m) - \bar{y}) + \int_{s=0}^1 O(\epsilon)(y(t_2^m) - \bar{y}) ds \\ &= H(y(t_2^m) - \bar{y}) + O(\epsilon^2), \end{aligned}$$

¹ We need to work on the closed ball. Because it is compact and ∇V is continuous, Heine-Cantor theorem implies uniform continuity, hence approaching the set K_i on which ∇V_i is 0 implies that ∇V_i goes to 0.

where we have used $\|y(t_2^m) - \bar{y}\| = 2\epsilon$. Similarly $\nabla V(y(t_1^m)) - \nabla V(\bar{y}) = H(y(t_1^m) - \bar{y}) + O(\epsilon^2)$. Hence

$$\nabla V(y(t_2^m)) - \nabla V(y(t_1^m)) = H\Delta y^m + O(\epsilon^2). \quad (4)$$

By Claim 1, we know that $\nabla V_i(y(t_2^m)) \rightarrow 0$ and $\nabla V_i(y(t_1^m)) \rightarrow 0$. Therefore, it follows from (4), applied to each component $i = 1, \dots, k$, that

$$\nabla V_i(y(t_2^m)) - \nabla V_i(y(t_1^m)) = H_i \Delta y^m + O(\epsilon^2) \rightarrow 0,$$

which implies (3) and thus the claim. \square

As a next step, we show that $\Delta y_i^\epsilon = 0$ for $i > k$. The idea of the proof is that every Δy_i^m , of which Δy_i^ϵ is an accumulation point, results in a proportional decrease of energy “along the i coordinate” that cannot be compensated by the other coordinates due to our elementwise decrease condition (2).

Claim 3: $\Delta y_i^\epsilon = 0$ for $i > k$.

PROOF. We show that $\lim_{m \rightarrow \infty} \Delta y_i^m = 0$ for $i > k$, which implies the claim as Δy^ϵ is an accumulation point of Δy_i^m .

Since $\nabla V_i(\bar{y}) \neq 0$ for $i > k$ by definition of k , the continuity of ∇V implies that if ϵ is sufficiently small, we have $|\nabla V_i(x)| > c > 0$ for some $c > 0$ for all $x \in \bar{B}(\bar{y}, 2\epsilon)$. Therefore, since $y(t) \in \bar{B}(\bar{y}, 2\epsilon)$ for $t \in [t_1^m, t_2^m]$, there holds

$$|\Delta y_i^m| \leq \int_{t_1^m}^{t_2^m} |\dot{y}_i| dt \leq \frac{1}{c} \int_{t_1^m}^{t_2^m} |\nabla V_i(y)| |\dot{y}_i| dt.$$

Our main assumption on coordinate-wise decrease (2) implies that $|\nabla V_i(y)| |\dot{y}_i| = -\nabla V_i(y) \dot{y}_i$, and generally that $-\nabla V_j(y) \dot{y}_j \geq 0$ for every j . Hence,

$$\begin{aligned} |\Delta y_i^m| &\leq -\frac{1}{c} \int_{t_1^m}^{t_2^m} \left(\nabla V_i(y) \dot{y}_i + \sum_{i \neq j} \nabla V_j(y) \dot{y}_j \right) dt \\ &= \frac{1}{c} (V(y(t_1^m)) - V(y(t_2^m))). \end{aligned}$$

This last inequality holds for every m , so that

$$\sum_m |\Delta y_i^m| \leq \frac{1}{c} \sum_m (V(y(t_1^m)) - V(y(t_2^m))) < \infty,$$

as $V(y(t))$ is non-increasing and the overall decrease of $V(y(t))$ is finite by Lemma 4. Therefore, there holds $|\Delta y_i^m| \rightarrow 0$ as $m \rightarrow \infty$, which implies $\Delta y_i^\epsilon = 0$. \square

Claim 4: Let $H = \nabla^2 V(\bar{y})$. If ϵ is small enough, there holds

$$(\Delta y^\epsilon)^T H (\Delta y^\epsilon) \leq C' \epsilon^3 \quad (5)$$

for some constant C' depending only on \bar{y} .

PROOF. For $i \leq k$, it follows from Claim 2 that

$$|\Delta y_i^\epsilon (H \Delta y^\epsilon)_i| \leq \|\Delta y^\epsilon\| |(H \Delta y^\epsilon)_i| \leq 4\epsilon C \epsilon^2 =: C' \epsilon^3.$$

For $i > k$, we have $\Delta y_i^\epsilon (H \Delta y^\epsilon)_i = 0$ because $\Delta y_i^\epsilon = 0$ by Claim 3, so there holds

$$(\Delta y^\epsilon)^T H (\Delta y^\epsilon) \leq C' \epsilon^3. \quad \square$$

We are now ready to prove Proposition 5.

PROOF. First, remember that $\|\Delta y^\epsilon\| \in [2\epsilon, 4\epsilon]$, hence any sequence of $\frac{\Delta y^\epsilon}{\epsilon}$ admits an accumulation point. In particular, among the small enough ϵ there exists a sequence of ϵ_ℓ converging to 0 and a vector v with $\|v\| \in [2, 4]$ such that $v = \lim_{\ell \rightarrow \infty} \frac{1}{\epsilon_\ell} \Delta y^{\epsilon_\ell}$.

For $i > k$, Claim 3 implies that $\frac{1}{\epsilon_\ell} \Delta y_i^{\epsilon_\ell} = 0$ and thus $v_i = 0$. Besides, from Claim 4, we have

$$\begin{aligned} v^T H v &= \lim_{\ell \rightarrow \infty} \left(\frac{\Delta y^{\epsilon_\ell}}{\epsilon_\ell} \right)^T H \left(\frac{\Delta y^{\epsilon_\ell}}{\epsilon_\ell} \right) \\ &= \lim_{\ell \rightarrow \infty} \frac{1}{\epsilon_\ell^2} (\Delta y^{\epsilon_\ell})^T H (\Delta y^{\epsilon_\ell}) \\ &\leq \lim_{\ell \rightarrow \infty} \frac{1}{\epsilon_\ell^2} C' \epsilon_\ell^3 = C' \epsilon_\ell = 0. \end{aligned}$$

So we have found a nonzero v in the kernel of H such that $v_i = 0$ for all $i > k$. If $k < n$ then the kernel contains a nonzero vector with a zero entry, i.e. case (c') of Proposition 5 holds. On the other hand, if $k = n$, then the definition of k implies $\bar{y} \in K_i$ for every i , that is, $\nabla V_i(\bar{y}) = 0$ for every i so that case (b') holds: $\nabla V(\bar{y}) = 0$ and $\bar{y} \in \arg \min_x V(x)$ by the convexity of V . Since these conclusions hold for any locally K -minimal accumulation points \bar{y} , we have established Proposition 5. \square

4.3 Generalization to all accumulation points

We now prove Theorem 1 by extending the result of Proposition 5 to all accumulation points, whether K -minimal or not.

Lemma 6 *Under the assumptions of Theorem 1, the set of $x \in \mathbb{R}^n$ satisfying conclusion (b') or (c') of Proposition 5 is closed.*

PROOF. By continuity of $\nabla V(x)$, the set of $x \in \mathbb{R}^n$ such that $\nabla V(x) = 0$ is closed, which proves the claim for (b'). For (c') we let T_i be the set of points such that the kernel of $\nabla^2 V(x)$ has a nonzero vector whose i^{th} component is zero, and show that each of the T_i is closed, which implies the result since the target set is the union of the T_i . We assume without loss of generality that $i = n$. A point x belongs to T_n if and only if there is a non-trivial $w \in \mathbb{R}^{n-1}$ such that

$$0 = H \begin{pmatrix} w \\ 0 \end{pmatrix} = \begin{pmatrix} H_{1:n-1}(x) & h_n(x) \end{pmatrix} \begin{pmatrix} w \\ 0 \end{pmatrix} = H_{1:n-1}(x)w,$$

i.e. if and only if $H_{1:n-1}(x)$ is rank deficient. Since the rank of a matrix is the size of its largest nonsingular square submatrix, being rank deficient can be checked by checking that all the n submatrices of size $(n-1) \times (n-1)$ of $H_{1:n-1}(x)$ have a zero determinant, i.e. that n continuous functions of x are zero. The set T_n is thus an intersection of zero sets of continuous functions and is therefore closed. \square

The next lemma will allow us to deduce that the set of K -minimal accumulation points is dense within the set of accumulation points, i.e. its closure contains all accumulation points. We state it for functions defined on subsets of \mathbb{R}^n , but it actually directly extends to general topological spaces, with the same proof.

Lemma 7 *Let $S \subset \mathbb{R}^n$ and $g : S \rightarrow \mathbb{N}$. The set of local minima of g is dense in S .*

PROOF. Let $M_k \subset S$ be the set of locally minimal points with value k . By definition,

$$M_k = g^{-1}(k) \setminus \overline{\bigcup_{i=0}^{k-1} g^{-1}(i)}. \quad (6)$$

We show by induction that $\overline{\bigcup_{i=0}^k M_i} = \overline{\bigcup_{i=0}^k g^{-1}(i)}$. For $k = 0$ this is immediate because (6) becomes $M_0 = g^{-1}(0)$. Let us now assume the relation holds for $k-1$. Using (6), we may write

$$\begin{aligned} \overline{M_k \cup \bigcup_{i=0}^{k-1} M_i} &= \overline{\left(g^{-1}(k) \setminus \bigcup_{i=0}^{k-1} g^{-1}(i) \right) \cup \bigcup_{i=0}^{k-1} g^{-1}(i)} \\ &= \overline{\left(g^{-1}(k) \cup \bigcup_{i=0}^{k-1} g^{-1}(i) \right)} = \overline{\bigcup_{i=0}^k g^{-1}(i)}, \end{aligned}$$

which confirms the induction step. Since S is the domain

of g , we have then

$$\begin{aligned} S = g^{-1}(\mathbb{N}) &= \bigcup_{i=0}^{\infty} g^{-1}(i) \subseteq \bigcup_{k=0}^{\infty} \overline{\bigcup_{i=0}^k g^{-1}(i)} \\ &= \bigcup_{k=0}^{\infty} \overline{\bigcup_{i=0}^k M_i} \subseteq \overline{\bigcup_{i=0}^{\infty} M_i}, \end{aligned}$$

i.e. the closure of local minima covers S . \square

To complete the proof of the main theorem, we let S be the set of accumulation points of y , and define on this set the function g assigning to each point the number of set K_i to which it belongs. Observe that the set S_{\min} of K -minimal accumulation points is exactly the set of local minima of g . Hence it follows from Lemma 7 that S_{\min} is dense in S , and thus that $S \subseteq \bar{S}_{\min}$. Now Proposition 5 states that, in the absence of convergence of y , every point of S_{\min} satisfies condition (b') or (c'), and we have seen in Lemma 6 that the set of points satisfying either of these conditions are closed. Hence every point of $S \subseteq \bar{S}_{\min}$ also satisfies (b') or (c').

Finally, we observe that if one accumulation point \bar{y} satisfied condition (b'), i.e. $\nabla V(\bar{y}) = 0$, the convexity of V implies that $V(\bar{y}) = \min_x V(x)$, and Lemma 4 implies then that $\lim_{t \rightarrow \infty} V(y(t)) = \min_x V(x)$ and thus that all accumulation points of y must similarly satisfy condition (b'). Consequently, if y does not converge, either one accumulation point satisfies (b') and then all of them do, leading to condition (b) of the theorem, or no accumulation point does, and then they all have to satisfy (c'), implying condition (c) of the theorem.

5 Conclusions and Open Research Directions

We have extended the results of [8] to general convex energy functions as opposed to simply quadratic ones, and clarified the possible impact of zero component in vectors of the Hessian kernel at some positions, significantly increasing their applicability. Our result allows establishing the convergence of trajectories under very simple and easily verifiable assumptions, and guarantees in other situation simple and strong properties for the accumulation points of the trajectory.

We hope our results will serve as a useful tool for the analysis of the evolution of various systems, to take a shortcut in confirming convergence when otherwise there is a high complexity in the description of the dynamics. One may think of multi-agent interactions with communication issues, cooperation or race conditions, measurement errors and quantizations, exogenous randomness and more.

Note that currently our results per se do not provide information on the convergence speed, but this is a consequence of an approach applicable to trajectories with potentially arbitrarily slow convergence. There remains, however, several open questions.

Guaranteeing convergence: We have seen in Examples 1 and 2 that the alternative (b) and (c) to the convergence of the trajectory cannot simply be discarded from the possible conclusions of Theorem 1. However, it might be possible to strengthen Theorem 1 by modifying our coordinate-wise decrease assumption (2). Observe indeed that both Examples 1 and 2 involve the trajectory freely moving along coordinates for which the corresponding gradient is zero, and hence would not satisfy the following stronger variation of assumption (2):

$$\dot{y}_i(t) \neq 0 \Rightarrow \dot{y}_i(t) \frac{\partial V}{\partial x_i} |_{y(t)} < 0 \quad \forall t \geq 0, i = 1, \dots, n, \quad (7)$$

i.e. the derivative of the i^{th} coordinate of y should have opposite sign as the corresponding coordinate of ∇V , and must be zero if the latter is zero². Whether or not (7) is a sufficient condition for convergence is an open question. An even stronger assumption would also force $\dot{y}_i(t)$ to be nonzero when $\frac{\partial V}{\partial x_i} |_{y(t)} \neq 0$, but this would significantly decrease the applicability of the result, as it would forbid coordinates from fully stopping in most situations.

Extensions: The extension of Theorem 1 to functions V that are not necessarily convex is an open question. We observe that our proof would immediately apply to purely concave functions, though this case appears less relevant. Another extension would consist in relaxing the local Lipschitz continuity requirement for $\nabla^2 V$, which could further ease the practical design of energy functions, allowing e.g. for easily connecting linear and non-linear parts of V .

Coordinate-free formulation: Finally, the application of our results strongly depend on the choice of coordinates. They can be extended by embedding a change of coordinates, but formulating a truly coordinate-independent version of Theorem 1 remains future work.

Acknowledgements

J. Hendrickx was supported by the ‘‘RevealFlight’’ Concerted Research Action (ARC) of the F ed eration Wallonie-Bruxelles, and by the by F.R.S.-FNRS via the

² We note that this variation of the assumption was actually independently needed in Theorem 4 of the earlier work [8] for guaranteeing the monotonicity of $\min x_i(t)$ and $\max x_i(t)$, and was indeed satisfied when Theorem 4 was used in Theorem 5 of the same paper.

Incentive Grant for Scientific Research (MIS) “Learning from Pairwise Data” and the “KORNET” project. B. Gerencsér was supported by NRD (National Research, Development and Innovation Office) grant KKP 137490 and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences.

References

- [1] Radu Balan. An extension of Barbashin-Krasovski-LaSalle theorem to a class of nonautonomous systems. *arXiv preprint math/0506459*, 2005.
- [2] Dario Bauso, Laura Giarré, and Raffaele Pesenti. Consensus for networks with unknown but bounded disturbances. *SIAM Journal on Control and Optimization*, 48(3):1756–1770, 2009.
- [3] Hanen Damak, Mohamed A. Hammami, and Boris Kalitine. On the global uniform asymptotic stability of time-varying systems. *Differential Equations and Dynamical Systems*, 22(2):113–124, 2014.
- [4] Baris Fidan and Brian D.O. Anderson. Switching control for robust autonomous robot and vehicle platoon formation maintenance. In *Proceedings of 15th Mediterranean Conference on Control and Automation*, pages 1–6, 2007.
- [5] Laura Giarré and Raffaele Pesenti. Erratum: Consensus for networks with unknown but bounded disturbances. *SIAM Journal on Control and Optimization*, 59(2):1757–1757, 2021.
- [6] Rafal Goebel, Ricardo G Sanfelice, and Andrew R Teel. *Hybrid dynamical systems*. Princeton University Press, 2012.
- [7] Wilhelmus PMH Heemels, Karl Henrik Johansson, and Paulo Tabuada. An introduction to event-triggered and self-triggered control. In *2012 IEEE 51st IEEE conference on decision and control (cdc)*, pages 3270–3285. IEEE, 2012.
- [8] Julien M Hendrickx, Balázs Gerencsér, and Baris Fidan. Trajectory convergence from coordinate-wise decrease of quadratic energy functions, and applications to platoons. *IEEE Control Systems Letters*, 4(1):151–156, 2019.
- [9] Hassan K. Khalil and Jessy W. Grizzle. *Nonlinear systems*. Prentice Hall, Upper Saddle River, NJ, 2002.
- [10] Joseph LaSalle. Some extensions of Liapunov’s second method. *IRE Transactions on Circuit Theory*, 7(4):520–527, 1960.
- [11] Rodolphe Sepulchre, Mrdjan Jankovic, and Petar V Kokotovic. *Constructive nonlinear control*. Springer, 1997.
- [12] Mingming Shi, Claudio De Persis, and Pietro Tesi. Self-triggered network coordination over noisy communication channels. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 3942–3947. IEEE, 2017.
- [13] N. Sreedhar. Concerning Liapunov functions for linear systems - I. *International Journal of Control*, 11(1):165–171, 1970.
- [14] Arjan J Van Der Schaft and Johannes Maria Schumacher. *An introduction to hybrid dynamical systems*, volume 251. Springer London, 2000.