A. Discontinuous Systems and Differential Inclusions

Recall that for an initial value problem (IVP)

$$\dot{x}(t) = F(x(t)) \tag{37a}$$

$$x(0) = x_0 \tag{37b}$$

with $F : \mathbb{R}^n \to \mathbb{R}^n$, the typical way to check for existence of solutions is by establishing continuity of F. Likewise, to establish unicity of solution, we typically seek Lipschitz continuity. When F is discontinuous, we may understand (37a) as the Filippov differential inclusion

$$\dot{x}(t) \in \mathcal{K}[F](x(t)), \tag{38}$$

where $\mathcal{K}[F] : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ denotes the Filippov set-valued map given by

$$\mathcal{K}[F](x) \triangleq \bigcap_{\delta>0} \bigcap_{\mu(S)=0} \overline{\text{co}}\, F(B_\delta(x) \setminus S), \tag{39}$$

where μ denotes the usual Lebesgue measure and $\overline{\text{co}}$ the convex closure, *i.e.* closure of the convex hull co. For more details, see (Paden & Sastry, 1987). We can generalize (38) to the differential inclusion (Bacciotti & Ceragioli, 1999)

$$\dot{x}(t) \in \mathcal{F}(x(t)), \tag{40}$$

where $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is some set-valued map.

Definition 1 (Carathéodory/Filippov solutions)**.** We say that $x : [0, \tau) \to \mathbb{R}^n$ with $0 < \tau \leq \infty$ is a *Carathéodory solution* to (40) if $x(\cdot)$ is absolutely continuous and (40) is satisfied a.e. in every compact subset of $[0, \tau)$. Furthermore, we say that $x(\cdot)$ is a *maximal* Carathéodory solution if no other Carathéodory solution $x'(\cdot)$ exists with $x = x'|_{[0,\tau)}$. If $\mathcal{F} = \mathcal{K}[F]$, then Carathéodory solutions are referred to as *Filippov solutions*.

For a comprehensive overview of discontinuous systems, including sufficient conditions for existence (Proposition 3) and uniqueness (Propositions 4 and 5) of Filippov solutions, see the work of Cortés (2008). In particular, it can be established that Filippov solutions to (37) exist, provided that the following assumption (Assumption 1) holds.

Assumption 1 (Existence of Filippov solutions)**.** $F : \mathbb{R}^n \to \mathbb{R}^n$ is defined almost everywhere (a.e.) and is Lebesgue-measurable in a non-empty open neighborhood $U \subset \mathbb{R}^n$ of $x_0 \in \mathbb{R}^n$. Further, F is locally essentially bounded in U, *i.e.*, for every point $x \in U$, F is bounded a.e. in some bounded neighborhood of x.

More generally, Carathéodory solutions to (40) exist (now with arbitrary $x_0 \in \mathbb{R}^n$), provided that the following assumption (Assumption 2) holds.

Assumption 2 (Existence of Carathéodory solutions)**.** $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ has nonempty, compact, and convex values, and is *upper semi-continuous*.

Filippov & Arscott (1988) proved that, for the Filippov set-valued map $\mathcal{F} = \mathcal{K}[F]$, Assumptions 1 and 2 are equivalent (with arbitrary $x_0 \in \mathbb{R}^n$ in Assumption 1).

Unicity of solution requires further assumptions. Nevertheless, we can characterize the Filippov set-valued map in a similar manner to Clarke's generalized gradient, as seen in the following proposition.

Proposition 1 (Theorem 1 of Paden & Sastry (1987))**.** *Under Assumption 1, we have*

$$\mathcal{K}[F](x) = \left\{ \lim_{k\to\infty} F(x_k) : x_k \in \mathbb{R}^n \setminus (\mathcal{N}_F \cup S) \text{ s.t. } x_k \to x \right\} \tag{41}$$

for some (Lebesgue) zero-measure set $\mathcal{N}_F \subset \mathbb{R}^n$ and any other zero-measure set $S \subset \mathbb{R}^n$. In particular, if F is continuous at a fixed x, then $\mathcal{K}[F](x) = \{F(x)\}$.

For instance, for the GF (1), we have $\mathcal{K}[-\nabla f](x) = \{-\nabla f(x)\}$ for every $x \in \mathbb{R}^n$, provided that f is continuously differentiable. Furthermore, if f is only locally Lipschitz continuous and regular (see Definition 3 of Appendix B), then $\mathcal{K}[-\nabla f](x) = -\partial f(x)$, where

$$\partial f(x) \triangleq \left\{ \lim_{k\to\infty} \nabla f(x_k) : x_k \in \mathbb{R}^n \setminus \mathcal{N}_f \text{ s.t. } x_k \to x \right\} \tag{42}$$

denotes Clarke's generalized gradient (Clarke, 1981) of f, with $\mathcal{N}_f$ denoting the zero-measure set over which f is not differentiable (Rademacher's theorem). It can be established that ∂f coincides with the subgradient of f, provided that f is convex. Therefore, the GF (1) interpreted as Filippov differential inclusion may also be seen as a continuous-time variant of subgradient descent methods.

B. Finite-Time Stability of Differential Inclusions

We are now ready to focus on extending some notions from traditional Lipschitz continuous systems to differential inclusions.

Definition 2. We say that $x^\star \in \mathbb{R}^n$ is an *equilibrium* of (40) if $x(t) \equiv x^\star$ on some small enough non-degenerate interval is a Carathéodory solution to (40). In other words, if and only if $0 \in \mathcal{F}(x^\star)$. We say that (40) is *(Lyapunov) stable* at $x^\star \in \mathbb{R}^n$ if, for every $\varepsilon > 0$, there exists some $\delta > 0$ such that, for every maximal Carathéodory solution $x(\cdot)$ of (40), we have $\|x_0 - x^\star\| < \delta \implies \|x(t) - x^\star\| < \varepsilon$ for every $t \geq 0$ in the interval where $x(\cdot)$ is defined. Note that, under Assumption 2, if (40) is stable at $x^\star$, then $x^\star$ is an equilibrium of (40) (Bacciotti & Ceragioli, 1999). Furthermore, we say that (40) is *(locally and strongly) asymptotically stable* at $x^\star \in \mathbb{R}^n$ if is stable at $x^\star$ and there exists some $\delta > 0$ such that, for every maximal Carathéodory solution $x : [0, \tau) \to \mathbb{R}^n$ of (40), if $\|x_0 - x^\star\| < \delta$ then $x(t) \to x^\star$ as $t \to \tau$. Finally, (40) is *(locally and strongly) finite-time stable* at $x^\star$ if it is asymptotically stable and there exists some $\delta > 0$ and $T : B_\delta(x^\star) \to [0, \infty)$ such that, for every maximal Carathéodory solution $x(\cdot)$ of (40) with $x_0 \in B_\delta(x^\star)$, we have $\lim_{t \to T(x_0)} x(t) = x^\star$.

We will now construct a Lyapunov-based criterion adapted from the literature of finite-time stability of Lipschitz continuous systems.

Lemma 1. *Let $\mathcal{E}(\cdot)$ be an absolutely continuous function satisfying the differential inequality*

$$\dot{\mathcal{E}}(t) \leq -c\,\mathcal{E}(t)^\alpha \tag{43}$$

a.e. in $t \geq 0$, with $c, \mathcal{E}(0) > 0$ and $\alpha < 1$. Then, there exists some $t^\star > 0$ such that $\mathcal{E}(t) > 0$ for $t \in [0, t^\star)$ and $\mathcal{E}(t^\star) = 0$. Furthermore, $t^\star > 0$ can be bounded by

$$t^\star \leq \frac{\mathcal{E}(0)^{1-\alpha}}{c(1-\alpha)}, \tag{44}$$

with this bound tight whenever (43) holds with equality. In that case, but now with $\alpha \geq 1$, then $\mathcal{E}(t) > 0$ for every $t \geq 0$, with $\lim_{t\to\infty} \mathcal{E}(t) = 0$. This will be represented by $t^\star = \infty$, with $\mathcal{E}(\infty) \triangleq \lim_{t\to\infty} \mathcal{E}(t)$.

Proof. Suppose that $\mathcal{E}(t) > 0$ for every $t \in [0, T]$ with $T > 0$. Let $t^\star$ be the supremum of all such T's, thus satisfying $\mathcal{E}(t) > 0$ for every $t \in [0, t^\star)$. We will now investigate $\mathcal{E}(t^\star)$. First, by continuity of $\mathcal{E}$, it follows that $\mathcal{E}(t^\star) \geq 0$. Now, by rewriting

$$\dot{\mathcal{E}}(t) \leq -c\,\mathcal{E}(t)^\alpha \iff \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\mathcal{E}(t)^{1-\alpha}}{1-\alpha}\right] \leq -c, \tag{45}$$

a.e. in $t \in [0, t^\star)$, we can thus integrate to obtain

$$\frac{\mathcal{E}(t)^{1-\alpha}}{1-\alpha} - \frac{\mathcal{E}(0)^{1-\alpha}}{1-\alpha} \leq -c\,t, \tag{46}$$

everywhere in $t \in [0, t^\star)$, which in turn turn leads to

$$\mathcal{E}(t) \leq [\mathcal{E}(0)^{1-\alpha} - c(1-\alpha)t]^{1-\alpha} \tag{47}$$

and

$$t \leq \frac{\mathcal{E}(0)^{1-\alpha} - \mathcal{E}(t)^{1-\alpha}}{c(1-\alpha)} \leq \frac{\mathcal{E}(0)^{1-\alpha}}{c(1-\alpha)}, \tag{48}$$

where the last inequality follows from $\mathcal{E}(t) > 0$ for every $t \in [0, t^\star)$. Taking the supremum in (48) then leads to the upper bound (44). Finally, we conclude that $\mathcal{E}(t^\star) = 0$, since $\mathcal{E}(t^\star) > 0$ is impossible given that it would mean, due to continuity of $\mathcal{E}$, that there exists some $T > t^\star$ such that $\mathcal{E}(t) > 0$ for every $t \in [0, T]$, thus contradicting the construction of $t^\star$.

Finally, notice that if $\mathcal{E}$ is such that (43) holds with equality, then (47) and the first inequality in (48) hold with equality as well. The tightness of the bound (44) thus follows immediately. Furthermore, notice that if $\alpha \geq 1$, and $\mathcal{E}$ is a tight solution to the differential inequality (43), *i.e.* $\mathcal{E}(t) = [\mathcal{E}(0)^{1-\alpha} - c(1-\alpha)t]^{1-\alpha}$, then clearly $\mathcal{E}(t) > 0$ for every $t \geq 0$ and $\mathcal{E}(t) \to 0$ as $t \to \infty$. ∎

Cortés & Bullo (2005) proposed (Proposition 2.8) a Lyapunov-based criterion to establish finite-time stability of discontinuous systems, which fundamentally coincides with our Lemma 1 for the particular choice of exponent $\alpha = 0$. Their proposition was, however, directly based on Theorem 2 of Paden & Sastry (1987). Later, Cortés (2006) proposed a second-order Lyapunov criterion, which, on the other hand, fundamentally translates to $\mathcal{E}(t) \triangleq V(x(t))$ being strongly convex. Finally, Hui et al. (2009) generalized Proposition 2.8 of Cortés & Bullo (2005) in their Corollary 3.1, to establish semistability. Indeed, these two results coincide for isolated equilibria.

We now present a novel result that generalizes the aforementioned first-order Lyapunov-based results, by exploiting our Lemma 1. More precisely, given a Laypunov candidate function $V(\cdot)$, the objective is to set $\mathcal{E}(t) \triangleq V(x(t))$, and we aim to check that the conditions of Lemma 1 hold. To do this, and assuming V to be locally Lipschitz continuous, we first borrow and adapt from Bacciotti & Ceragioli (1999) the definition of *set-valued time derivative* of $V : \mathcal{D} \to \mathbb{R}$ w.r.t. the differential inclusion (40), given by

$$\dot{V}(x) \triangleq \{a \in \mathbb{R} : \exists v \in \mathcal{F}(x) \text{ s.t. } a = p \cdot v, \forall p \in \partial V(x)\}, \tag{49}$$

for each $x \in \mathcal{D}$. Notice that, under Assumption 2 for Filippov differential inclusions $\mathcal{F} = \mathcal{K}[F]$, the set-valued time derivative of V thus coincides with with the set-valued Lie derivative $\mathcal{L}_F V(\cdot)$. Indeed, more generally $\dot{V}$ could be seen as a set-valued Lie derivative $\mathcal{L}_{\mathcal{F}} V$ w.r.t. the set-valued map $\mathcal{F}$.

Definition 3. $V(\cdot)$ is said to be *regular* if every directional derivative, given by

$$V'(x; v) \triangleq \lim_{h \to 0} \frac{V(x + h\, v) - V(x)}{h}, \tag{50}$$

exists and is equal to

$$V^{\circ}(x; v) \triangleq \limsup_{x' \to x\; h \to 0^+} \frac{V(x' + h\, v) - V(x')}{h}, \tag{51}$$

known as *Clarke's upper generalized derivative* (Clarke, 1981).

In practice, regularity is a fairly mild and easy to guarantee condition. For instance, it would suffice that V is convex or continuously differentiable to ensure that it is Lipschitz and regular.

Assumption 3. $V : \mathcal{D} \to \mathbb{R}$ is locally Lipscthiz continuous and regular, with $\mathcal{D} \subseteq \mathbb{R}^n$ open.

Under Assumption 3, Clarke's generalized gradient

$$\partial V(x) \triangleq \{p \in \mathbb{R}^n : V^{\circ}(x; v) \geq p \cdot v, \forall v \in \mathbb{R}^n\} \tag{52}$$

is non-empty for every $x \in \mathcal{D}$, and is also given by

$$\partial V(x) = \left\{ \lim_{k \to \infty} \nabla V(x_k) : x_k \in \mathbb{R}^n \setminus \mathcal{N}_V \text{ s.t. } x_k \to x \right\}, \tag{53}$$

where $\mathcal{N}_V$ denotes the set of points in $\mathcal{D} \subseteq \mathbb{R}^n$ where V is not differentiable (Rademachers theorem) (Clarke, 1981).

Through the following lemma (Lemma 2), we can formally establish the correspondence between the set-valued time-derivative of V and the derivative of the energy function $\mathcal{E}(t) \triangleq V(x(t))$ associated with an arbitrary Carathéodory solution $x(\cdot)$ to the differential inclusion (40).

Lemma 2 (Lemma 1 of Bacciotti & Ceragioli (1999))**.** *Under Assumption 3, given any Carathéodory solution $x : [0, \tau) \to \mathbb{R}^n$ to (40), then $\mathcal{E}(t) \triangleq V(x(t))$ is absolutely continuous and $\dot{\mathcal{E}}(t) = \frac{\mathrm{d}}{\mathrm{d}t} V(x(t)) \in \dot{V}(x(t))$ a.e. in $t \in [0, \tau)$.*

We are now ready to state and prove our Lyapunov-based sufficient condition for finite-time stability of differential inclusions.

Theorem 3. *Suppose that Assumptions 2 and 3 hold for some set-valued map $\mathcal{F} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and some function $V : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D} \subseteq \mathbb{R}^n$ is an open and positively invariant neighborhood of a point $x^\star \in \mathbb{R}^n$. Suppose that V is positive definite w.r.t. $x^\star$ and that there exist constants $c > 0$ and $\alpha < 1$ such that*

$$\sup \dot{V}(x) \leq -c\, V(x)^{\alpha} \tag{54}$$

a.e. in $x \in \mathcal{D}$. Then, (40) is finite-time stable at $x^\star$, with settling time upper bounded by

$$t^\star \leq \frac{V(x_0)^{1-\alpha}}{c(1-\alpha)}, \tag{55}$$

where $x(0) = x_0$. In particular, any Carathéodory solution $x(\cdot)$ with $x(0) = x_0 \in \mathcal{D}$ will converge in finite time to $x^\star$ under the upper bound (55). Furthermore, if $\mathcal{D} = \mathbb{R}^n$, then (40) is globally finite-time stable. Finally, if $\dot{V}(x)$ is a singleton a.e. in $x \in \mathcal{D}$ and (54) holds with equality, then the bound (55) is tight.

Proof. Note that, by Proposition 1 of (Bacciotti & Ceragioli, 1999), we know that (40) is Lyapunov stable at $x^\star$. All that remains to be shown is local convergence towards $x^\star$ (which must be an equilibrium) in finite time. Indeed, given any maximal solution $x : [0, t^\star) \to \mathbb{R}^n$ to (40) with $x(0) = x_0 \neq x^\star$, we know by Lemma 2, that $\mathcal{E}(t) = V(x(t))$ is absolutely continuous with $\dot{\mathcal{E}}(t) \in \dot{V}(x(t))$ a.e. in $t \in [0, t^\star]$. Therefore, we have

$$\dot{\mathcal{E}}(t) \leq \sup \dot{V}(x(t)) \leq -c\, V(x(t))^\alpha = -c\, \mathcal{E}(t)^\alpha \tag{56}$$

a.e. in $t \in [0, t^\star]$. Since $\mathcal{E}(0) = V(x_0) > 0$, given that $x_0 \neq x^\star$, the result then follows by invoking Lemma 1 and noting that $\mathcal{E}(t^\star) = 0 \iff V(t^\star, x(t^\star)) = 0 \iff x(t^\star) = x^\star$. ∎

Finite-time stability still follows without Assumption 2, provided that $x^\star$ is an equilibrium of (40). In practical terms, this means that trajectories starting arbitrarily close to $x^\star$ may not actually exist, but nevertheless there exists a neighborhood $\mathcal{D}$ of $x^\star$ over which, any trajectory $x(\cdot)$ that indeed exists and starts at $x(0) = x_0 \in \mathcal{D}$ must converge in finite time to $x^\star$, with settling time upper bounded by $T(x_0)$ (the bound still tight in the case that (54) holds with equality).

C. Proof of Theorem 1

Let us focus first on the q-RGF (19) with the candidate Lyapunov function $V \triangleq f - f^\star$. Clearly, V is Lipschitz continuous and regular (given that it is continuously differentiable). Furthermore, V is positive definite w.r.t. $x^\star$.

Notice that, due to the dominated gradient assumption, $x^\star$ must be an isolated stationary point of f. To see this, notice that, if $x^\star$ were not an isolated stationary point, then there would have to exist some $\tilde{x}^\star$ sufficiently near $x^\star$ such that $\tilde{x}^\star$ is both a stationary point of f, and satisfies $f(\tilde{x}^\star) > f^\star$, since $x^\star$ is a strict local minimizer of f. But then, we would have

$$0 = \frac{p-1}{p} \|\nabla f(\tilde{x}^\star)\|^{\frac{p}{p-1}} \geq \mu^{\frac{1}{p-1}} (f(\tilde{x}^\star) - f^\star) > 0, \tag{57}$$

and subsequently $0 > 0$, which is absurd.

Therefore, $F(x) \triangleq -c\nabla f(x)/\|\nabla f(x)\|^{\frac{q-2}{q-1}}$ is continuous for every $x \in \mathcal{D} \setminus \{0\}$, for some small enough open neighborhood $\mathcal{D}$ of $x^\star$. Let us assume that $\mathcal{D}$ is positively invariant w.r.t. (19), which can be achieved, for instance, by replacing $\mathcal{D}$ with its intersection with some small enough strict sublevel set of f. Notice that $\|F(x)\| = c\|\nabla f(x)\|^{\frac{1}{q-1}}$ with $q \in (p, \infty] \subset (1, \infty]$, *i.e.*, $\frac{1}{q-1} \in [0, \infty)$. If $q = \infty$, which results in the normalized gradient flow $\dot{x} = -\frac{\nabla f(x)}{\|\nabla f(x)\|}$ proposed by Cortés (2006), then $\|F(x)\| = c > 0$. We can thus show that $F(x)$ is discontinuous at $x = 0$ for $q = \infty$. On the other hand, if $q \in (p, \infty) \subset (1, \infty)$, then we have $\|F(x)\| \to 0$ as $x \to x^\star$, and thus $F(x)$ is continuous (but not Lipschitz) at $x = x^\star$. Regardless, we may freely focus exclusively on $\mathcal{D} \setminus \{x^\star\}$ since $\{x^\star\}$ is obviously a zero-measure set.

Let $\mathcal{F} \triangleq \mathcal{K}[F]$. We thus have, for each $x \in \mathcal{D} \setminus \{x^\star\}$,

$$\begin{aligned}
\sup \dot{V}(x) &= \sup \{a \in \mathbb{R} : \exists v \in \mathcal{F}(x) \text{ s.t. } a = p \cdot v, \forall p \in \partial V(x)\} && \text{(58a)}\\
&= \sup \{\nabla V(x) \cdot v : v \in \mathcal{F}(x)\} && \text{(58b)}\\
&= \nabla V(x) \cdot F(x) && \text{(58c)}\\
&= -c\|\nabla f(x)\|^{2-\frac{q-2}{q-1}} && \text{(58d)}\\
&= -c\|\nabla f(x)\|^{\frac{1}{\theta'}} && \text{(58e)}\\
&\leq -c[C(f(x) - f^\star)^\theta]^{\frac{1}{\theta'}} && \text{(58f)}\\
&= -cC^{\frac{1}{\theta'}} V(x)^{\frac{\theta}{\theta'}}. && \text{(58g)}
\end{aligned}$$

Since $\frac{\theta}{\theta'} < 1$, given that $s > 1 \mapsto \frac{s-1}{s}$ is strictly increasing, then the conditions of Theorem 3 are satisfied. In particular, we have finite-time stability at $x^\star$ with a settling time $t^\star$ upper bounded by

$$t^\star \leq \frac{(f(x_0) - f^\star)^{1-\frac{\theta}{\theta'}}}{c\,C^{\frac{1}{\theta'}}\left(1 - \frac{\theta}{\theta'}\right)} \leq \frac{(\|\nabla f(x_0)\|/C)^{\frac{1}{\theta}\left(1-\frac{\theta}{\theta'}\right)}}{c\,C^{\frac{1}{\theta'}}\left(1 - \frac{\theta}{\theta'}\right)} = \frac{\|\nabla f(x_0)\|^{\frac{1}{\theta}-\frac{1}{\theta'}}}{cC^{\frac{1}{\theta}}\left(1 - \frac{\theta}{\theta'}\right)} \tag{59}$$

for each $x_0 \in \mathcal{D}$, which completes the proof.

Remark 1. Regarding the flow given by equation (15), as we mentioned in the main paper, it does require knowledge of f^*, which is generally unknown. However, if we replace $f^\star$ by some estimate $\tilde{f}^\star$ of it, then we still obtain a (locally) finite-time convergent flow, only that it will converge to either an invariant subset of $\{x \in \mathbb{R}^n : f(x) = \tilde{f}^\star\}$ if $\tilde{f}^\star \geq 0$, or it will converge to $x^\star$ if $\tilde{f}^\star < f^\star$. For instance, when a non-negative energy function is being minimized, then we can use $\tilde{f}^\star = 0$. However, in most applications, this flow appears to be less useful, which motivated us to look at the other first-order flows, as presented in Section 3.

D. Proof of Theorem 2

Let us focus first on (27), since the proof for (28) follows similar steps. The idea is to show that it is finite-time stable at $x^\star$, with the inequality in Theorem 3 holding exactly for $V(x) = \|\nabla f(x)\|^2$. First, notice that $F(x) \triangleq -c\|\nabla f(x)\|^{2\alpha} \frac{[\nabla^2 f(x)]^r \nabla f(x)}{\nabla f(x)^\top [\nabla^2 f(x)]^{r+1} \nabla f(x)}$ is continuous near (but not at) $x = x^\star$, and undefined at $x = x^\star$ itself. Furthermore, we have

$$\|F(x)\| = c\|\nabla f(x)\|^{2\alpha} \frac{\|[\nabla^2 f(x)]^r \nabla f(x)\|}{\nabla f(x)^\top [\nabla^2 f(x)]^{r+1} \nabla f(x)} \tag{60a}$$

$$\leq c\|\nabla f(x)\|^{2\alpha} \frac{\lambda_{\max}(\nabla^2 f(x))^r \|\nabla f(x)\|}{\lambda_{\min}(\nabla^2 f(x))^{r+1} \|\nabla f(x)\|^2} \tag{60b}$$

$$\leq c \frac{\lambda_{\max}(\nabla^2 f(x))^r}{\lambda_{\min}(\nabla^2 f(x))^{r+1}} \|\nabla f(x)\|^{2\alpha - 1}, \tag{60c}$$

with $2\alpha - 1 \geq 0$ and $\lambda_{\min}[\nabla^2 f(x)] \geq m > 0$ everywhere near $x = x^\star$ for some $m > 0$ (m-strong convexity). Therefore, F is Lebesgue integrable (and thus measurable) and locally essentially bounded, which means that Assumption 1 is satisfied.

Set $V(x) \triangleq \|\nabla f(x)\|^2$, defined over $\mathcal{D}$. If $\mathcal{D}$ is not positively invariant w.r.t. (27), we can always replace it by a smaller open subset that is, *e.g.* a sufficiently small strict sublevel set contained within $\mathcal{D}$. Clearly, V is continuously differentiable, thus satisfying Assumption 3. Furthermore, it is positive definite w.r.t. $x^\star$ and, given $x \in \mathcal{D} \setminus \{x^\star\}$, we have

$$\sup \dot{V}(x) = \sup\{a \in \mathbb{R} : \exists v \in \mathcal{K}[F](x) \text{ s.t. } a = p \cdot v, \forall p \in \partial V(x)\} \tag{61a}$$

$$= \sup \left\{ \left(2\nabla^2 f(x) \nabla f(x)\right) \cdot \left(-c\|\nabla f(x)\|^{2\alpha} \frac{[\nabla^2 f(x)]^r \nabla f(x)}{\nabla f(x)^\top [\nabla^2 f(x)]^{r+1} \nabla f(x)} \right) \right\} \tag{61b}$$

$$= -2c\|\nabla f(x)\|^{2\alpha} \tag{61c}$$

$$= -(2c)\, V(x)^\alpha \tag{61d}$$

with $\alpha < 1$. Furthermore, $\dot{V}(x^\star) = \{\nabla V(x^\star) \cdot v : v \in \mathcal{K}[F](x^\star)\} = \{0\}$ since $\nabla V(x^\star) = 2\nabla^2 f(x^\star) \nabla f(x^\star) = 0$. The result thus follows by invoking Theorem 3.

We now proceed to establish finite-time stability of (28) at $x^\star$. Like before, we notice that $F(x) = -c\|\nabla f(x)\|_1^{2\alpha - 1} x \frac{[\nabla^2 f(x)]^r \operatorname{sign}(\nabla f(x))}{\operatorname{sign}(\nabla f(x))^\top [\nabla^2 f(x)]^{r+1} \operatorname{sign}(\nabla f(x))}$ is continuous near, but not at, $x = x^\star$. Furthermore, notice that

$$\|F(x)\| = c\|\nabla f(x)\|_1^{2\alpha - 1} \frac{\|[\nabla^2 f(x)]^r \operatorname{sign}(\nabla f(x))\|}{\operatorname{sign}(\nabla f(x))^\top [\nabla^2 f(x)]^{r+1} \operatorname{sign}(\nabla f(x))} \tag{62a}$$

$$\leq c\|\nabla f(x)\|_1^{2\alpha - 1} \frac{\lambda_{\max}(\nabla^2 f(x))^r \|\operatorname{sign}(\nabla f(x))\|}{\lambda_{\min}(\nabla^2 f(x))^{r+1} \|\operatorname{sign}(\nabla f(x))\|^2} \tag{62b}$$

$$\leq c \frac{\lambda_{\max}(\nabla^2 f(x))^r}{\lambda_{\min}(\nabla^2 f(x))^{r+1}} \frac{\|\nabla f(x)\|_1^{2\alpha - 1}}{\|\operatorname{sign}(\nabla f(x))\|} \tag{62c}$$

$$\leq c \frac{\lambda_{\max}(\nabla^2 f(x))^r}{\lambda_{\min}(\nabla^2 f(x))^{r+1}} \frac{\|\nabla f(x)\|_1^{2\alpha - 1}}{\|\operatorname{sign}(\nabla f(x))\|_1 / \sqrt{n}} \tag{62d}$$

$$\leq \frac{c}{\sqrt{n}} \frac{\lambda_{\max}(\nabla^2 f(x))^r}{\lambda_{\min}(\nabla^2 f(x))^{r+1}} \|\nabla f(x)\|_1^{2\alpha - 1} \tag{62e}$$

(62f)

for every $x \in \mathcal{D} \setminus \mathcal{N}$, with $\mathcal{N} = \bigcup_{i=1}^{n} \left\{x \in \mathcal{D} : \frac{\partial f}{\partial x_i}(x) = 0\right\}$. Since $\mathcal{N}$ is a zero-measure set due to being a finite union of hypersurfaces in $\mathbb{R}^n$ ($n \geq 1$), and also recalling that f is strongly convex near $x^\star$ and $p - 1 \geq 0$, it follows that F is Lebesgue integrable and locally essentially bounded. Therefore, Assumption 1 is once again satisfied.

Now consider the candidate Lyapunov function $V(x) = \|\nabla f(x)\|_1$, defined over $\mathcal{D}$. Clearly, V is not continuously differentiable this time. However, it is still satisfies Assumption 1 due to being a.e. differentiable. In particular, we have $\partial V(x) = \{\nabla^2 f(x) \operatorname{sign}(\nabla f(x))\}$ for every $x \in \mathcal{D} \setminus \mathcal{N}$. In other words, we have $\partial V(x) = \{\nabla^2 f(x) \operatorname{sign}(\nabla f(x))\}$ a.e. in $x \in \mathcal{D}$.

Given $x \in \mathcal{D} \setminus \mathcal{N}$, we thus have

$$
\begin{aligned}
\sup \dot{V}(x) &= \sup\{a \in \mathbb{R} : \exists v \in \mathcal{K}[F](x) \text{ s.t. } a = p \cdot v, \forall p \in \partial V(x)\} && \text{(63a)}\\
&= \Big(\nabla^2 f(x) \operatorname{sign}(\nabla f(x))\Big) \cdot \left(-\frac{c\|\nabla f(x)\|_1^{2\alpha-1} [\nabla^2 f(x)]^r \operatorname{sign}(\nabla f(x))}{\operatorname{sign}(\nabla f(x))^\top [\nabla^2 f(x)]^{r+1} \operatorname{sign}(\nabla f(x))}\right) && \text{(63b)}\\
&= -c\|\nabla f(x)\|_1^{2\alpha-1} && \text{(63c)}\\
&= -c\, V(x)^{2\alpha-1}, && \text{(63d)}
\end{aligned}
$$

with $2\alpha - 1 < 1$. The result once again follows by invoking Theorem 3.