To appear in the *International Journal of Control*
Vol. 00, No. 00, Month 20XX, 1–23

Globally convergent differentiators with variable gains https://doi.org/10.1080/00207179.2018.1448115

A. Levant[a] and M. Livne[b]

[a]*Department of Applied Mathematics, Tel Aviv University, Israel;*
[b]*Israel Aerospace Industries, Lod, Israel*

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

A new robust exact sliding-mode (SM) based differentiator is proposed which provides for the fast global finite-time convergence of its outputs to the first n exact derivatives of its input $f(t)$. The differentiator utilizes the knowledge of a function $L(t)$ providing the estimation $|f_0^{(n+1)}| \leq L(t)$, and satisfying $|\dot{L}|/L \leq M$ for a known bound M. The standard accuracy of the homogeneous SM-based differentiator is preserved in the presence of discrete sampling and noises in both f and L. The proposed discretization scheme ensures the same accuracy in computer realization.

Keywords: Differentiation, Sliding-Mode Control, Robustness, Observation.

1. Introduction

Differentiation of noisy signals is an old ill-posed core problem of observation and identification. The main approach is to approximate the sampled signal by a signal with known derivatives. Some popular tracking methods applied in that context are based on sliding-mode (SM) and high-gain control (Atassi & Khalil, 2000; Golembo, Emelyanov, Utkin, & Shubladze, 1976; Yu & Xu, 1996). Whereas SM control is used to suppress system uncertainties keeping the tracking error (sliding variable) at zero by means of high-frequency control switching, high-gain control utilizes the same idea introducing artificial singular perturbations.

Sliding mode is accurate and insensitive to disturbances (Edwards & Spurgeon, 1998; Moreno, 2014). While standard SMs nullifies sliding variables of the relative degree 1 (Isidori, 1989), higher order sliding modes (HOSMs) (Bartolini, Pisano, Punta, & Usai, 2003; Floquet, Barbot, & Perruquetti, 2003; Levant, 2005; Plestan, Moulay, Glumineau, & Cheviron, 2010; Shtessel, Taleb, & Plestan, 2012) are capable of exactly establishing constraints of higher relative degrees in finite time (FT). They are also used to alleviate the chattering effect (Fridman, 2003) caused by the high control-switching frequency (Bartolini, Ferrara, & Usai, 1998; Levant, 2003).

Roughly speaking, the order of the sliding mode $\sigma \equiv 0$ is defined as the order of the first derivative of the sliding variable σ that contains discontinuity. Let a function $f_0(t)$ be $n+1$ times differentiable, then the problem of tracking f_0 by means of a dynamic system $z_0^{(n+1)} = u$ is equivalent to establishing the $(n+1)$th-order SM ($(n+1)$-SM) $\sigma = z_0 - f_0(t) = 0$ by means of an appropriate discontinuous control u. Indeed, denoting $z_0^{(i)} = z_i$ obtain $z_i = f_0^{(i)}$, $i = 0, 1, ..., n$, in the SM, while $z_0^{(n+1)} - f_0^{(n+1)}$ contains discontinuity. Correspondingly, the HOSM theory produces robust FT-convergent exact differentiators (Angulo, Moreno, & Fridman, 2013; Bartolini, Pisano, & Usai, 2000; Bejarano & Fridman, 2010; Castillo, Fridman, & Moreno, 2017; Cruz-Zavala, Moreno, &

CONTACT A. Levant. Email: levant@post.tau.ac.il

Fridman, 2012; Efimov & Fridman, 2011; Levant, 2003, 2005, 2009; Levant & Livne, 2012; Livne & Levant, 2014; Moreno, 2012, 2014). The HOSM-based observation (Angulo, Fridman, & Levant, 2012; Bejarano & Fridman, 2010; Plestan et al., 2010) has become one of the main directions of the HOSM theory and application.

The usual requirement of HOSM-based nth-order differentiation is that the nth derivative $f_0^{(n)}$ has a known Lipschitz constant $L > 0$. Contrary to high-gain observers (Atassi & Khalil, 2000), HOSM-based differentiators do not contain large parameters and produce exact derivatives in the absence of noises. In the case of discrete measurements with the sampling interval τ and the *unknown* maximal noise magnitude $\varepsilon > 0$ the steady-state error of the ith derivative estimation is of the order $O(\max(\tau^{n-i+1}, \varepsilon^{(n-i+1)/(n+1)}))$. That accuracy is shown to be asymptotically optimal for continuous measurements, i.e., for $\tau = 0$ (Levant, 2003; Levant, Livne, & Yu, 2017). The features of the SM differentiators are derived from the homogeneity of their error dynamics (Levant, 2005).

After the differentiator coefficients are fixed, the Lipschitz constant L actually remains the only adjustable parameter of the standard HOSM-based differentiator (Levant, 2003). That differentiator is also known to keep its FT stability for variable parameter L, $|f_0^{(n+1)}(t)| \leq L(t)$, provided its logarithmic derivative $\dot{L}/L$ is bounded (Levant & Livne, 2012).

Unfortunately, such differentiator with variable parameter L features two major disadvantages. First, the convergence is in general only local (global convergence is preserved for monotonously growing differentiable L (Moreno, 2017)). Though it is known that the region of attraction to zero in the space of errors is proportional to L, the actual region size depends on the differentiator gains and can only be found by simulation. This means that the differentiator can even lose its stability for sufficiently large noises. The second drawback is inherited from the original homogeneous differentiator (Levant, 2003): its convergence rate is low, if the initial errors are large.

Both issues are settled in this paper. New quasi-linear terms are for this end added to the *recursive* form of the differentiator producing a *hybrid* differentiator combining the features of the homogeneous diffirentiator (Levant, 2003) and the linear filter (Atassi & Khalil, 2000). The new differentiator converges in FT for any initial values of differentiation errors. It generalizes results obtained for constant or growing L (Levant, 2009; Moreno, 2017) or only for the first-order differentiators (Castillo et al., 2017; Moreno, 2012). The convergence rate is exponential outside of a vicinity of the origin in the error space, and is easily regulated. The accuracy and robustness features of the standard differentiator (Levant, 2003) are preserved.

This paper extends and corrects similar results by Levant (2014). The upper bound M, $|\dot{L}/L| \leq M$, has become one of the differentiator parameters. Moreover, the additional input $L(t)$ can itself be sampled with noises. Exactly like the constant parameter L was the only actual parameter of the homogeneous differentiator, the new parameter M is the only actual parameter of the hybrid differentiator, whereas other coefficients can be fixed in advance forever. We list valid parameters for $n \leq 7$, which makes the application of the hybrid differentiator very simple.

The standard accuracy presented above is obtained under the assumption that, while the measurements are discrete, the differentiator is still described by differential equations, i.e. the integration is performed in continuous time. In practice the differentiator is a computer-based real-time system. The natural one-step Euler discretization destroys the theoretical asymptotics, and the discrete scheme requires modification (Livne & Levant, 2014). The proposed discrete hybrid differentiator restores the standard accuracy asymptotics.

The simulation for the first time demonstrates the explosive divergence of the standard differentiator in spite of the boundedness of $\dot{L}/L$, whereas the proposed hybrid differentiator features fast FT exact convergence. The observed accuracy asymptotics fits the theory.

2. Structure of differentiators

A signal $\Xi(t)$ is called available in *real time* if at each time t the values of $\Xi(\hat{t})$ are available for any sampling instant $\hat{t} \in [0, t]$. By *real-time estimations* of a signal we understand estimations based on the information available over the time interval $[0, t)$ only. That exclusion of the values at t does not influence real-time estimations in the case of continuous-time sampling due to the continuity of the differentiator outputs, but it is significant for the causality of discrete schemes.

2.1 *Standard robust exact differentiator*

Let the input signal $f(t)$ be available in real-time for $t \in [0, \infty)$ and let it consist of a bounded Lebesgue-measurable noise $\eta(t)$ with unknown features and an unknown basic signal $f_0(t)$ whose nth derivative has a known Lipschitz constant $L > 0$. The problem of finding real-time robust estimations of $f_0(t), \dot{f}_0(t), \ddot{f}_0(t), ..., f_0^{(n)}(t)$, being exact in the absence of measurement noises is solved by the differentiator (Levant, 2003)

$$\begin{array}{rcl} \dot{z}_0 & = & v_0 = \lambda_n L^{\frac{1}{n+1}} |z_0 - f(t)|^{\frac{n}{n+1}} \operatorname{sign}(z_0 - f(t)) + z_1, \\ \dot{z}_1 & = & v_1 = \lambda_{n-1} L^{\frac{1}{n}} |z_1 - v_0|^{\frac{n-1}{n}} \operatorname{sign}(z_1 - \dot{z}_0) + z_2, \\ & ... & \\ \dot{z}_{n-1} & = & v_{n-1} = \lambda_1 L^{\frac{1}{2}} |z_{n-1} - v_{n-2}|^{\frac{1}{2}} \operatorname{sign}(z_{n-1} - \dot{z}_{n-2}) + z_n, \\ \dot{z}_n & = & \lambda_0 L \operatorname{sign}(z_n - v_{n-1}). \end{array} \qquad (1)$$

Here and further all differential equations are understood in the Filippov sense (Filippov, 1988). An infinite positive parametric sequence $\lambda_0, \lambda_1, ..., \lambda_n, ...$, is proved to exist, which provides for the FT convergence of the nth-order differentiator for any n and L. In particular, the choice $\lambda_0 = 1.1, \lambda_1 = 1.5, \lambda_2 = 2, \lambda_3 = 3, \lambda_4 = 5, \lambda_5 = 8$ is sufficient for $n \leq 5$ (Levant, 2005).

The following equalities are true in the absence of input noises after a FT transient process:

$$z_0 = f_0(t); \qquad z_i = v_{i-1} = f_0^{(i)}(t), \ \ i = 1, ..., n. \qquad (2)$$

Note that differentiator (1) has a recursive structure: once the parameters $\lambda_0, \lambda_1, ..., \lambda_{n-1}$ are chosen properly for the $(n-1)$th order differentiator with the Lipschitz constant $L = 1$, only one parameter λ_n is needed to be chosen for the nth order differentiator. It simply is taken sufficiently large to provide for the differentiator convergence. Any $\lambda_0 > 1$ can be used to start this process at $n = 0$. Such differentiator can be used in any feedback, trivially providing for the separation principle (Atassi & Khalil, 2000). Unfortunately, due to the slow growth of the fractional-power functions in (1) the differentiator convergence is slow for large initial errors.

2.2 *The problem statement and the new hybrid differentiator structure*

Assume that the input function $f(t)$ consists of an *unknown* basic function $f_0(t)$ and a Lebesgue-measurable noise $\eta(t)$, $f(t) = f_0(t) + \eta(t)$. The basic function $f_0(t)$ is known to have n derivatives. The nth derivative $f_0^{(n)}(t)$ is assumed to be a locally Lipschitz function, for almost any t satisfying the inequality $|f_0^{(n+1)}(t)| \leq L_0(t)$ for an *unknown* time-variable positive absolutely-continuous function $L_0(t) > 0$. Also L_0 is sampled in real time as $L(t) = L_0(t) + \eta_L(t)$, where $\eta_L(t)$ is some small Lebesgue-measurable noise, $|\eta_L(t)|/L_0(t) < 1$. The only available numeric information consists of n, the sampled noisy signals $f(t)$, $L(t)$ and a constant $M \geq 0$ introduced below.

The logarithmic derivative of L_0 is assumed to be bounded, $|\dot{L}_0/L_0| \leq M$ for some known M. It is also assumed that the noises $\eta(t)$, $\eta_L(t)$ are not large compared with $L_0(t)$, $|\eta(t)| \leq \varepsilon L_0(t)$,

$|\eta_L(t)| \le \varepsilon_L L_0(t)$, $\varepsilon_L < 1$, where the constant parameters $\varepsilon, \varepsilon_L$ are considered unknown.

The goal is to accurately restore the functions $f_0(t), \dot{f}_0(t), ..., f_0^{(n)}(t)$.

Note that the assumptions on the noise η are not restrictive, since the function $L(t)$ can be changed by the "user". Any bounded noise η is allowed if L_0 is separated from 0. For example, one can take $\tilde{L} = L+1$ and $\tilde{L}_0 = L_0 + 1$. At the same time one cannot easily remove the noise $\eta_L(t)$ by redefining L. It is not simple to increase L_0 to $\tilde{L}_0(t)$, $\tilde{L}_0(t) \ge L(t)$, while preserving the condition $|\dot{\tilde{L}}_0/\tilde{L}_0| \le M$.

Presence of the noise $\eta_L(t)$ is essential for the generality of the considered problem. If $f(t)$ is an output of an uncertain process, the upper bound of $|f_0^{(n+1)}|$ is also not exact. For example, if $f(t)$ is a coordinate of an aircraft, like pitch or roll, then the bounds of its high derivatives depend on other sensor-measured coordinates and the aero-dynamic model parameters. The linear aero-dynamic model of the nonlinear system is determined by wind-tunnel experiments at a number of key points defined by their altitude and velocity. The closest current point is chosen at each moment. Thus $L(t)$ is neither exact nor even continuous. An attempt to produce a smooth reliable bound $L(t)$ only becomes a source of additional errors.

The proposed *hybrid* differentiator is of the form

$$\begin{array}{rcl} \dot{z}_0 &=& v_0 = -\varphi_0(t, z_0 - f(t)) + z_1, \\ \dot{z}_1 &=& v_1 = -\varphi_1(t, z_1 - v_0) + z_2, \\ & ... & \\ \dot{z}_n &=& -\varphi_n(t, z_n - v_{n-1}), \end{array} \tag{3}$$

where φ_i are the functions

$$\varphi_i(t,s) = \lambda_{n-i} L(t)^{\frac{1}{n-i+1}} |s|^{\frac{n-i}{n-i+1}} \operatorname{sign} s + \mu_{n-i} M s, \tag{4}$$

and $\lambda_{n-i}, \mu_{n-i} > 0$, $i = 0, 1, ..., n$.

Note that the recursive form (3) is equivalent to the standard dynamic-system form

$$\begin{array}{rcl} \dot{z}_0 &=& -\varphi_0(t, z_0 - f(t)) + z_1, \\ \dot{z}_1 &=& -\varphi_1(t, \varphi_0(t, z_0 - f(t))) + z_2, \\ ... && \\ \dot{z}_n &=& -\varphi_n(t, \varphi_{n-1}(...(t, \varphi_0(t, z_0 - f(t)))...)). \end{array} \tag{5}$$

The hybrid differentiator (3), (4) turns into the standard differentiator (1) for $M = 0$ and into the linear high-gain observer (HGO) (Atassi & Khalil, 2000) for $L = 0$, $M >> 1$. The coefficients of the resulting HGO are $\mu_n, \mu_n\mu_{n-1}, ..., \mu_n\mu_{n-1}\cdots\mu_0$ from the top down. Similar differentiators were first considered in (Levant, 2009) for the constant parameter L and in (Levant, 2014) for variable L, $\eta_L = 0$. In both cases the parameter M did not appear (i.e. $M = 1$ was taken in (4)). Parameters μ_i were assumed to be separately adjusted for each upper bound of $|\dot{L}|/L$ (Levant, 2014).

Contrary to the standard differentiator (1), the non-recursive form (5) of the differentiator (5) cannot be simplified. In particular, the first-order hybrid differentiator (5) gets the form

$$\begin{array}{l} \dot{z}_0 = -\lambda_1 L^{\frac{1}{2}} |z_0 - f(t)|^{\frac{1}{2}} \operatorname{sign}(z_0 - f(t)) - \mu_1 M (z_0 - f(t)) + z_1, \\ \dot{z}_1 = -\lambda_0 L \operatorname{sign}(z_0 - f(t)) - \mu_0 \lambda_1 L^{\frac{1}{2}} M |z_0 - f(t)|^{\frac{1}{2}} \operatorname{sign}(z_0 - f(t)) - \mu_0 \mu_1 M^2 (z_0 - f(t)), \end{array} \tag{6}$$

which was also independently proposed by Moreno (2009, 2012), where it was suggested to provide for faster convergence in the case of constant L. It differs from the simpler standard form (Moreno, 2014) by the additional nonlinear term in the second equation. The FT convergence of the hybrid differentiator (6) for $|\dot{L}|/L \leq M$ has been recently proved by the Lyapunov method (Castillo et al., 2017).

2.3 *Discretization*

In practice the input signal is sampled at discrete time instants, and the differentiator itself is to be a discrete computer-based system. It requires some real-time numerical integration of differential equations (3) over each sampling interval. Unfortunately, the accuracy of differentiator (3) is destroyed by the most natural one-Euler-step integration (Livne & Levant, 2014).

Let t_k be the sampling instants, $\tau_k = t_{k+1} - t_k > 0$. The proposed discrete differentiator

$$\begin{aligned}
z_0(t_{k+1}) &= z_0(t_k) - \varphi_0(t_k, z_0(t_k) - f(t_k))\tau_k \\
&\qquad + \textstyle\sum_{j=1}^{n} \frac{z_j}{j!}\tau_k^j, \\
z_1(t_{k+1}) &= z_1(t_k) - \varphi_1(t_k, z_1(t_k) - v_0(z(t_k), t_k))\tau_k \\
&\qquad + \textstyle\sum_{j=2}^{n} \frac{z_j}{(j-1)!}\tau_k^{j-1}, \\
&\dots \\
z_i(t_{k+1}) &= z_i(t_k) - \varphi_i(t_k, z_i(t_k) - v_{i-1}(z(t_k), t_k))\tau_k \\
&\qquad + \textstyle\sum_{j=i+1}^{n} \frac{z_j}{(j-i)!}\tau_k^{j-i}, \\
&\dots \\
z_n(t_{k+1}) &= z_n(t_k) - \varphi_n(t_k, z_n(t_k) - v_{n-1}(z(t_k), t_k))\tau_k.
\end{aligned} \tag{7}$$

has additional Taylor-like terms. Functions v_i, φ_i are defined in (3), (4). Unlike (Livne & Levant, 2013, 2014), in (7) the additional terms are added directly to the recursive form, but the result is algebraically identical.

Note that the one-step Euler discretization is obtained from (7) by cancelling all higher-order terms that contain τ_k^l with $l > 1$. In particular, the first-order differentiator (6) does not get such terms in (7). I.e. (7) corresponds to the usual Euler discretization for $n = 0, 1$.

3. Main results

The outputs (solutions) $z(t)$ of the hybrid differentiator (3), (4) are indefinitely extendable in time for any choice of the parameters λ_i, μ_i, M and locally bounded Lebesgue-measurable inputs $L(t)$, $f(t)$. Indeed, it easily follows from (5) that $||\dot{z}(t)|| \leq \theta_1(L(t))||z|| + \theta_2(L(t))|f(t)| + \theta_3(L(t))$, where θ_j are some positive continuous functions.

Introduce the normalized errors $\sigma_i = (z_i - f^{(i)}(t))/L_0(t)$, $i = 0, 1, ..., n$. Let $\varepsilon = \varepsilon_L = 0$. We call the FT convergence of errors to zero *uniform* for a fixed parameter $M \geq 0$, if for any $R > 0$ there exists a constant $T(R) > 0$ such that for any inputs f_0, L_0 satisfying the above assumptions and any $t_0 \geq 0$ the inequality $||\sigma(t_0)|| \leq R$ implies $\sigma(t) = 0$ for any $t \geq t_0 + T(R)$. In other words there exists a transient-time estimation which does not depend on the inputs $f_0(\cdot), L_0(\cdot)$.

Theorem 1: *For any $\lambda_0 > 1$, $\mu_0 > 1$ there exists a positive double sequence $\{\lambda_i, \mu_i\}$, $i = 0, 1, ...,$ such that under the above conditions for any n and any sufficiently small ε_L*

1. *in the absence of the noise η, i.e. for $\varepsilon = 0$, the differentiator (3), (4) uniformly converges in FT for any $M \geq 0$ and equalities $z_i(t) \equiv f_0^{(i)}(t)$, $i = 0, ..., n$, are established;*
2. *there exist constants $c_{1,n}, c_{2,n} > 0$ such that for any $M > 0$ and $t_0 \geq 0$ the convergence time is estimated as $T_{conv} \leq M^{-1}[c_{1,n} \ln \max_i(|\sigma_i(t_0)|M^{n+1-i}, 1) + c_{2,n}]$ for $\varepsilon = 0$;*

(3) normalized errors σ_i remain bounded for any finite magnitude ε of the normalized noise η/L_0.

Here and further all proofs are placed in the appendix.

Remark 1: The above convergence time estimation shows that the convergence is accelerated for larger M, but that estimation is very crude for small M. Parameters $\{\lambda_i, \mu_i\}$ are chosen recursively, pair-by-pair, sufficiently large in the list order. The proved way (Lemma 2, Appendix C) is for any $n > 0$ to assign any $\bar{\mu}_n > 0$ and to choose sufficiently large λ_n and $\mu_n = \bar{\mu}_n \lambda_n^{(n+1)/n}$. It is also required that $\mu_n > 1$. Experimentally checked parameters valid for $n \leq 7$ are listed in Section 4.

Theorem 2: *Let the infinite positive sequences $\{\lambda_i\}$, $\{\mu_i\}$ provide for the uniform FT convergence of the differentiator* (3), (4) *in the absence of noises for any n and any input $f(t)$, $L(t) > 0$, $|f^{(n+1)}(t)| \leq L(t)$, $M \geq 0$, $|\dot{L}/L| \leq M$. The following are some properties of these parameters.*

(1) The sequence $\{\lambda_i\}$ is a valid parametric sequence for the standard differentiator (1) *with constant L.*
(2) Let $\alpha, \beta \geq 1$. Then sequences $\{\alpha^{1/(i+1)}\lambda_i\}$, $\{\beta\mu_i\}$ are also valid.
(3) Fix some n_0 and let $\beta \geq 1$, $\alpha = \beta^{n_0}$. Then sequences $\{\beta^{n_0/(i+1)}\lambda_i\}$, $\{\beta\mu_i\}$ provide for the β-times-faster convergence of the n_0th-order differentiator (3), (4). *I.e. for any $R \geq 0$ the maximal convergence time from the error region $\max_{i\in[0,n_0]} |z_i(0) - f_0^{(i)}(0)| \leq R$ to zero is reduced by at least β times. That estimation does not hold for differentiators of orders different from n_0.*

The proof of Theorem 2 in appendix C utilizes that any signal $f(t)$ characterized by some $L(t)$ and M at the same time is described by $\alpha L(t)$ and βM, $\alpha, \beta \geq 1$. Such transformation is not allowed for $\alpha < 1$ or $\beta < 1$, since it reduces the class of allowed inputs.

Remark 2: The proof of Theorem 1 (Appendix C) implies that the parameters are chosen so that for any n the polynomial $x^{n+1} + \mu_n x^n + \mu_n\mu_{n-1}x^{n-1} + ... + \mu_n\mu_{n-1}\cdots\mu_0$ is Hurwitz. One can prove that the same polynomial is neutrally stable, i.e. has no roots with positive real parts and no multiple imaginary roots under the conditions of Theorem 2.

Consider now the case of noisy inputs and discrete sampling, when $f(t_k), L(t_k)$ are substituted for $f(t), L(t)$ in (3) and (5), $t_k \leq t < t_{k+1}$, $\tau_k = t_{k+1} - t_k$.

Theorem 3: *Let the positive sequences $\{\lambda_i\}$, $\{\mu_i\}$ be chosen as in Theorem 1, and the sampling intervals τ_k not exceed some $\tau > 0$. Then the normalized errors σ_i remain bounded for any ε and sufficiently small τ, ε_L. Moreover, let $\delta = \max(\tau, \varepsilon^{1/(n+1)})$, where the case of the continuous time sampling is formally included for $\tau = 0$. Then there exist such positive constants γ_i that for any sufficiently small δ the differentiator* (3), (4) *in FT provides the accuracy*

$$|z_i - f_0^{(i)}(t)| \leq \gamma_i L(t)\delta^{n-i+1}, i = 0, ..., n. \tag{8}$$

Note that accuracy (8) coincides with the asymptotic accuracy of the standard differentiator (1), in which case it is kept for *any* $\delta \geq 0$ (Levant, 2003; Levant & Livne, 2012). In the case of continuous time sampling this asymptotic accuracy cannot be improved (Levant, 2003; Levant et al., 2017). Consider now the discrete version (7) of the differentiator.

Theorem 4: *Under the conditions of Theorem 3 discrete differentiator* (7) *provides for the normalized errors σ_i bounded for any ε and sufficiently small τ, ε_L. Moreover, for any sufficiently small δ the accuracy*

$$|z_i(t_k) - f_0^{(i)}(t_k)| \leq \gamma_i L(t_k)\delta^{n-i+1},\ i = 0, ..., n,\ \delta = \max(\tau, \varepsilon^{1/(n+1)}), \tag{9}$$

is maintained with possibly different γ_i.

Remark 3: The simplest one-step Euler scheme is also always applicable, but the resulting asymptotic accuracy with respect to τ is worse (Levant, Livne, & Lunz, 2016; Livne & Levant, 2014). More important is that the same scheme (7) is valid for any virtual measurement at $t = \hat{t}_{k+1} \in [t_k, t_{k+1}]$. Indeed, the right-hand side of (7) does not depend on the measurements at t_{k+1}. Therefore, the scheme produces estimations at any moment t between the real measurements and even for **the prediction** up to $t = t_k + \tau$ while preserving the same accuracy.

4. Simulation

Consider the input function

$$f(t) = f_0(t) + \eta(t),\ f_0(t) = 2\sin\left(\tfrac{1}{2}t^2\right),\ |\eta(t)| \le \varepsilon, \tag{10}$$

where η is the noise. It is easy to check that $|f_0^{(4)}| \le L(t)$,

$$L(t) = 2t^4 + 12t^2 + 6,\ \tfrac{|\dot{L}|}{L} \le 1.6,\ M = 2. \tag{11}$$

The sequence $\{(\lambda_i, \mu_i)\}_{i=0}^n = (1.1, 2), (1.5, 3), (2, 4), (3, 7), (5, 9), (7, 13), (10, 19), (12, 23), ...$ has been experimentally validated for $n \le 7$ and can be extended up to $n = \infty$. In particular, the third order discrete differentiator (7) takes on the form

$$\begin{aligned}
&z_0(t_{k+1}) = z_0(t_k) + v_0\tau_k + \tfrac{1}{2}z_2(t_k)\tau_k^2 + \tfrac{1}{6}z_3(t_k)\tau_k^3,\\
&v_0 = -3L(t_k)^{1/4}|z_0(t_k) - f(t_k)|^{3/4}\,\mathrm{sign}(z_0(t_k) - f(t_k))\\
&\qquad -7M(z_0(t_k) - f(t_k)) + z_1(t_k),\\[2ex]
&z_1(t_{k+1}) = z_1(t_k) + v_1\tau_k + \tfrac{1}{2}z_3(t_k)\tau_k^2,\\
&v_1 = -2L(t_k)^{1/3}|z_1(t_k) - v_0(t_k)|^{2/3}\,\mathrm{sign}(z_1(t_k) - v_0(t_k))\\
&\qquad -4M(z_1(t_k) - v_0(t_k)) + z_2(t_k),\\[2ex]
&z_2(t_{k+1}) = z_2(t_k) + v_2\tau_k,\\
&v_2 = -1.5L(t_k)^{1/2}|z_2(t_k) - v_1(t_k)|^{1/2}\,\mathrm{sign}(z_2(t_k) - v_1(t_k))\\
&\qquad -3M(z_2(t_k) - v_1(t_k)) + z_3(t_k),\\[2ex]
&z_3(t_{k+1}) = z_3(t_k) - 1.1L(t_k)\,\mathrm{sign}(z_3(t_k) - v_2(t_k))\tau_k\\
&\qquad -2M(z_3(t_k) - v_2(t_k))\tau_k.
\end{aligned} \tag{12}$$

Let $z(0) = (10, 50, -70, 800)$. The convergence of the outputs z_i to the exact derivatives $f_0^{(i)}$ in the absence of noise ($\varepsilon = 0$) is demonstrated in Fig. 1. The accuracies $|z_0 - f_0| \le 6.9 \cdot 10^{-12}$, $|z_1 - \dot{f}_0| \le 7.5 \cdot 10^{-8}$, $|z_2 - \ddot{f}_0| \le 5.9 \cdot 10^{-4}$, $|z_3 - \dddot{f}_0| \le 2.8$ are kept for $\tau = 10^{-4}$ over the time interval $[5, 7]$. Correspondingly, the normalized accuracies $|\sigma_0| \le 1.4 \cdot 10^{-15}$, $|\sigma_1| \le 1.5 \cdot 10^{-11}$, $|\sigma_2| \le 1.2 \cdot 10^{-7}$, $|\sigma_3| \le 5.4 \cdot 10^{-4}$ are obtained, $\sigma_i = (z_i - f^{(i)})/L$ (Fig. 3(a)). Note that for the specified initial values the standard differentiator does not converge till 9.

The same test is now carried out for $\varepsilon = 1 \cdot 10^{-3}, 2 \cdot 10^{-3}, ..., 1 \cdot 10^{-2}$. The noise η is uniformly distributed in $[-\varepsilon, \varepsilon]$ at each integration/sampling step. As before, the accuracies are calculated over the time interval $[5, 7]$. The decimal logarithmic plots of the normalized differentiator errors and the corresponding linear regressions are shown in Fig. 2. The slopes $(0.89, 0.63, 0.43, 0.21)$ are obtained as compared to the theoretical counterparts $(1, 0.75, 0.5, 0.25)$ due to (9).

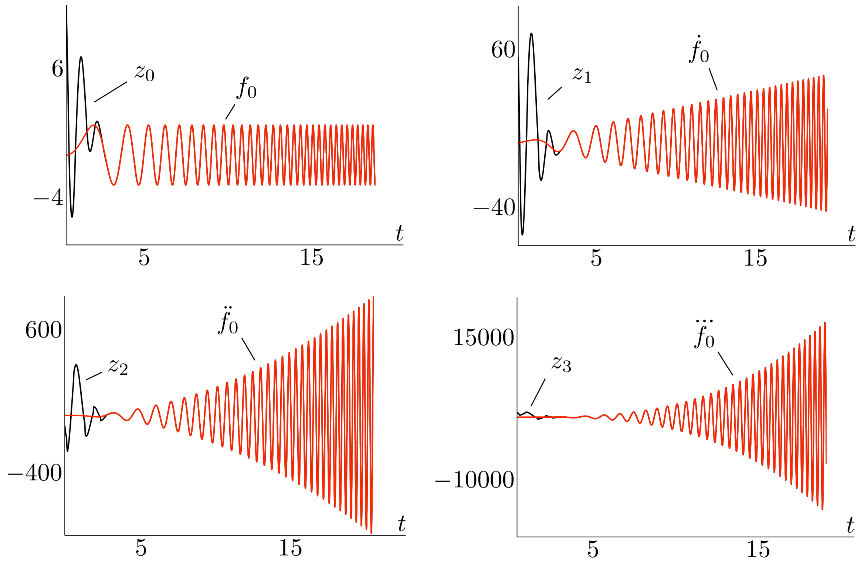


Figure 1. Convergence of the hybrid differentiator (12) for the inputs (10), (11).

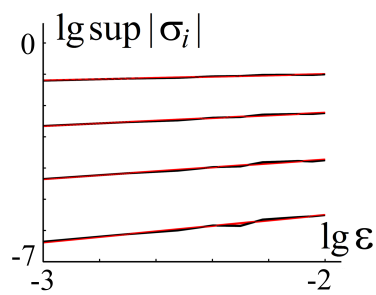
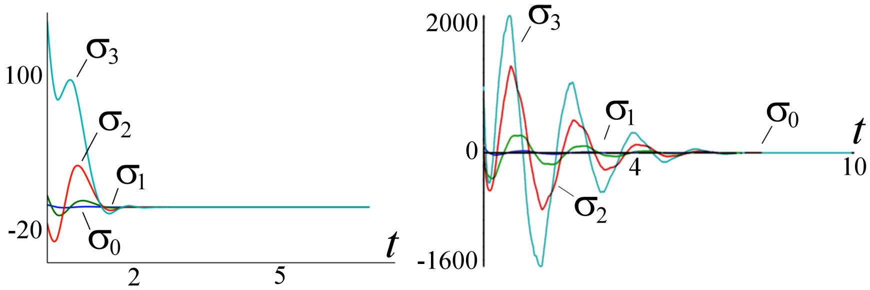


Figure 2. Accuracy asymptotics of the hybrid differentiator (12) for the inputs (10), (11). Graphs correspond to the derivative orders 0, 1, 2, 3 from the bottom to the top.



(a) Polynomial Lipschitz gain $L(t)$, $M = 2$ (b) Comparison test, $M = 1$

Figure 3. Normalized differentiation errors' convergence of the hybrid differentiator (12).

4.1 *Comparison of the hybrid and the standard differentiators*

It is shown in (Levant & Livne, 2012) that the standard differentiator (1) is FT stable (FTS) for any sufficiently small initial errors, provided $L(t)$ has the bounded logarithmic derivative $\dot{L}/L$. In practice it turns out that the standard differentiator almost always converges (Moreno, 2017), though the convergence is very slow for large initial errors. It seems that in order to demonstrate the differentiator divergence one needs some *feedback* regulation of $L(t)$.

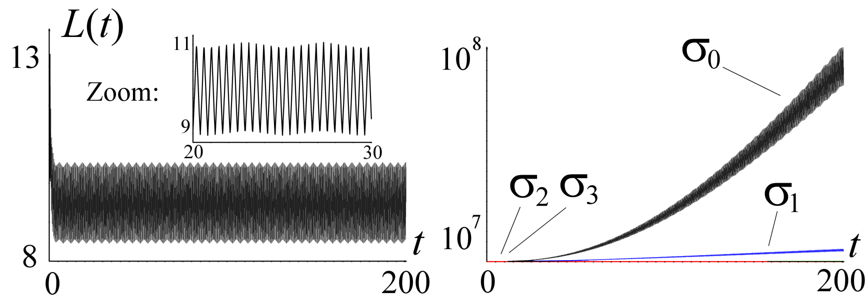


Figure 4. Comparison test (13). Standard differentiator (12) divergence, $M = 0$.

Below we probably for the first time demonstrate the divergence of the standard differentiator (1) in spite of the valid function $L(t)$ with a bounded logarithmic derivative. We also demonstrate that the new differentiator rapidly converges also under these severe conditions.

The following input f_0 and the function $L(t)$ are specially constructed during simulation tests to destroy the convergence of differentiator (1) in the experiment. Let $\varepsilon = 0$, $z_0(0) = 1000$, $z_1(0) = -1000$, $z_2(0) = 1200$, $z_3(0) = 10000$,

$$\begin{array}{l} f(t) = f_0(t) = \cos t - \sin(0.5t),\ L(0) = 10,\ M = 1, \\ \phi(t) = 2 + \cos(15t), \dot{L} = \begin{cases} -L\ \mathrm{sign}(L - \phi(t))\ \mathrm{if}(z_0 - f_0)(z_1 - \dot{f}_0) \geq 0\, \mathrm{and}\, L \geq 5\phi(t), \\ -L\ \mathrm{sign}(L - 10\phi(t))\ \mathrm{if}(z_0 - f_0)(z_1 - \dot{f}_0) < 0\, \mathrm{or}\, L < 5\phi(t), \end{cases} \end{array} \tag{13}$$

Obviously, $|\dot{L}/L| \leq M = 1$.

Hybrid discrete differentiator (12) with $\tau_k = \tau = 10^{-4}$ is applied. The convergence takes about 8 time units and its normalized errors are shown in Fig. 3(b). The accuracies $|z_0 - f_0| \leq 1.7 \cdot 10^{-14}$, $|z_1 - \dot{f}_0| \leq 2.9 \cdot 10^{-10}$, $|z_2 - \ddot{f}_0| \leq 1.8 \cdot 10^{-6}$, $|z_3 - \dddot{f}_0| \leq 0.0044$ are kept for $\tau = 10^{-4}$. Correspondingly, the normalized accuracies $|\sigma_0| \leq 1.4 \cdot 10^{-15}$, $|\sigma_1| \leq 2.4 \cdot 10^{-11}$, $|\sigma_2| \leq 1.4 \cdot 10^{-7}$, $|\sigma_3| \leq 3.5 \cdot 10^{-4}$ are obtained. Note that the normalized accuracies actually remain the same for very different inputs f_0 and L.

The same simulation is now run for the standard differentiator (1), i.e. for $M = 0$. The corresponding function $L(t)$ and the explosive divergence of the normalized errors are demonstrated in Fig. 4. The simulation is performed for $t \in [0, 200]$ in order to exclude the possibility of very long, but still practically feasible transient. Hence the new differentiator converges whereas the standard differentiator demonstrates practical divergence for the same initial errors.

5. Conclusion

For the first time a globally convergent FTS arbitrary-order robust exact differentiator is demonstrated which differentiates the signal $f(t)$ with a variable Lipschitz parameter $L(t)$, $|f^{n+1}(t)| \leq L(t)$, having a known bound $M \geq 0$, $|\dot{L}|/L \leq M$. Its proposed discretization preserves the theoretical accuracy of the standard HOSM differentiator. Both $f(t), L(t)$ may be sampled with noises.

Such differentiators are especially useful for the global output-feedback control of systems with growth not faster than linear, and in various observation problems which require parameter adaptation. They also practically remove the problem of choosing the initial values.

References

Andrieu, V., Praly, L., & Astolfi, A. (2008). Homogeneous approximation, recursive observer design, and output feedback. *SIAM Journal on Control and Optimization*, *47*(4), 1814–1850.

Angulo, M., Fridman, L., & Levant, A. (2012). Output-feedback finite-time stabilization of disturbed {LTI} systems. *Automatica*, *48*(4), 606 - 611.

Angulo, M., Moreno, J., & Fridman, L. (2013). Robust exact uniformly convergent arbitrary order differentiator. *Automatica*, *49*(8), 2489–2495.

Atassi, A., & Khalil, H. (2000). Separation results for the stabilization of nonlinear systems using different high-gain observer designs. *Systems & Control Letters*, *39*(3), 183-191.

Bartolini, G., Ferrara, A., & Usai, E. (1998). Chattering avoidance by second-order sliding mode control. *IEEE Transactions on Automatic Control*, *43*(2), 241-246.

Bartolini, G., Pisano, A., Punta, E., & Usai, E. (2003). A survey of applications of second-order sliding mode control to mechanical systems. *International Journal of Control*, *76*(9/10), 875-892.

Bartolini, G., Pisano, A., & Usai, E. (2000). First and second derivative estimation by sliding mode technique. *Journal of Signal Processing*, *4*(2), 167-176.

Bejarano, F., & Fridman, L. (2010). High order sliding mode observer for linear systems with unbounded unknown inputs. *International Journal of Control*, *83*(9), 1920-1929.

Castillo, I., Fridman, L., & Moreno, J. (2017). Super-twisting algorithm in presence of time and state dependent perturbations. *International Journal of Control*(published online 16 Jan 2017).

Cruz-Zavala, E., Moreno, J., & Fridman, L. (2012). Uniform robust exact differentiator. *IEEE Transactions on Automatic Control*, *56*(11), 2727-2733.

Edwards, C., & Spurgeon, S. (1998). *Sliding mode control: Theory and applications.* London: Taylor & Francis.

Efimov, D., & Fridman, L. (2011). A hybrid robust non-homogeneous finite-time differentiator. *IEEE Transactions on Automatic Control*, *56*(5), 1213-1219.

Filippov, A. (1962). On some problems of optimal control theory, Vestnik Moskov. Univ. Mat. 1958 (2), 25-32. *English transl.: SIAM J. Control*, *1*, 76–84.

Filippov, A. (1988). *Differential equations with discontinuous right-hand sides.* Dordrecht: Kluwer Academic Publishers.

Floquet, T., Barbot, J., & Perruquetti, W. (2003). Higher-order sliding mode stabilization for a class of nonholonomic perturbed systems. *Automatica*, *39*(6), 1077-1083.

Fridman, L. (2003). Chattering analysis in sliding mode systems with inertial sensors. *International Journal of Control*, *76*(9/10), 906-912.

Golembo, B., Emelyanov, S., Utkin, V., & Shubladze, A. (1976). Application of piecewise-continuous dynamic-systems to filtering problems. *Automation and remote control*, *37*(3), 369–377.

Isidori, A. (1989). *Nonlinear control systems, second edition.* New York: Springer Verlag.

Levant, A. (2003). Higher order sliding modes, differentiation and output-feedback control. *International J. Control*, *76*(9/10), 924-941.

Levant, A. (2005). Homogeneity approach to high-order sliding mode design. *Automatica*, *41*(5), 823-830.

Levant, A. (2009). Non-homogeneous finite-time-convergent differentiator. In *Proceedings of the 48th IEEE Conference on Decision and Control, 2009, held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009* (p. 8399-8404).

Levant, A. (2014). Globally convergent fast exact differentiator with variable gains. In *Proc. of the European Conference on Control, Strasbourg, France, June 24-27, 2014.*

Levant, A., & Livne, M. (2012). Exact differentiation of signals with unbounded higher derivatives. *IEEE Transactions on Automatic Control*, *57*(4), 1076-1080.

Levant, A., & Livne, M. (2016). Weighted homogeneity and robustness of sliding mode control. *Automatica*, *72*(10), 186–193.

Levant, A., Livne, M., & Lunz, D. (2016). On discretization of high order sliding modes. In *L. Fridman, J.P. Barbot, F. Plestan (eds). Recent trends in sliding mode control* (pp. 177–202). IET.

Levant, A., Livne, M., & Yu, X. (2017). Sliding-mode-based differentiation and its application. In *Proc. of the 20th IFAC World Congress, Toulouse, July 9-14, France, 2017.*

Livne, M., & Levant, A. (2013). Homogeneous discrete differentiation of functions with unbounded higher derivatives. In *IEEE 52nd Annual Conference on Decision and Control (CDC 2013)* (p. 1762-1767).

Livne, M., & Levant, A. (2014). Proper discretization of homogeneous differentiators. *Automatica*, *50*, 2007-2014.

Moreno, J. (2009). A linear framework for the robust stability analysis of a generalized super-twisting

algorithm. In *Proc. of the 6th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE'2009, 10-13 Jan. 2009, Toluca, Mexico* (pp. 1–6).
Moreno, J. (2012). Lyapunov approach for analysis and design of second order sliding mode algorithms. In *Sliding modes after the first decade of the 21st century* (pp. 113–149). Springer.
Moreno, J. (2014). On strict Lyapunov functions for some non-homogeneous super-twisting algorithms. *Journal of the Franklin Institute*, *351*(4), 1902-1919.
Moreno, J. (2017). Levant's arbitrary order differentiator with varying gain. In *Proc. of the 20th IFAC World Congress, Toulouse, July 9-14, France, 2017.*
Plestan, F., Moulay, E., Glumineau, A., & Cheviron, T. (2010). Robust output feedback sampling control based on second order sliding mode. *Automatica*, *46*(6), 1096-1100.
Shtessel, Y., Taleb, M., & Plestan, F. (2012). A novel adaptive-gain supertwisting sliding mode controller: methodology and application. *Automatica*, *48*(5), 759–769.
Yu, X., & Xu, J. (1996). Nonlinear derivative estimator. *Electronic Letters*, *32*(16), 1445-1447.

Appendix A. Some notation and notions

For any symbol α denote $\vec{\alpha}_k = (\alpha_0, \alpha_1, ..., \alpha_k)$, $\overleftarrow{\alpha}_i = (\alpha_i, \alpha_{i+1}, ..., \alpha_k)$ with k known from the context; $\|\cdot\|$ is the Euclidian norm, whereas $\|x\|_\infty = \max_i |x_i|$ for any vector $x \in \mathbb{R}^k$, $\|x\|_{h\infty} = \max_i |x_i|^{\frac{k+1-i}{k+1}}$ is the homogeneous norm of $x = (x_0, ..., x_k) \in \mathbb{R}^{k+1}$. Also denote $\lfloor s \rceil^p = |s|^p \operatorname{sign} s$, $\lfloor s \rceil^0 = \operatorname{sign} s$. We denote differentiator (5) by $\dot{z} = D_{n,\lambda,\mu}(z, f, L, M)$.

Solutions of differential inclusions (DIs) are defined as any locally absolutely continuous functions satisfying DIs almost everywhere. Whenever the right-hand-side of a DI contains a discontinuous function it is replaced with the set obtained by the Filippov procedure (Filippov, 1988) producing a convex compact upper-semicontinuous vector-set field.

For any function $\Theta(\cdot)$ and a set X define $\Theta(X) = \{\Theta(x) \mid x \in X\}$. For any binary operation $\otimes$ define $X \otimes Y = \{x \otimes y \mid x \in X, y \in Y\}$. One-element sets are not distinguished from their only element in that context.

Appendix B. Normal forms. Reduction to the case $M = 1$

Lemma 1: *Differentiator* (3), (4) *converges for any input under conditions of Theorem 1 iff it converges for any input with* $M = 1$. *For* $M > 0$ *the transformation to the form* (3), (4) *with* $M = 1$ *is the time-coordinate transformation* $t = \hat{t}/M$, $\hat{z}_i(\hat{t}) = z_i(\hat{t}/M)M^{-i}$, $\hat{f}_0(\hat{t}) = f_0(\hat{t}/M)$, $\hat{L}_0(\hat{t}) = L_0(\hat{t}/M)M^{-(n+1)}$ *and the corresponding transformation of noises.*

Proof of Lemma 1. Convergence for $M = 1$ implies convergence for $M = 0$ (the proof is exactly part (1) of the proof of Theorem 2, since it is performed for $M = 1$). Thus assume $M > 0$. The proof is performed by induction. The 0th-order differentiator (3), (4) gets the form

$$\dot{z}_0 = -\lambda_0 L(t) \operatorname{sign}(z_0 - f(t)) + \mu_0 M(z_0 - f(t)), \tag{B1}$$

where $|\dot{f}(t)| \leq L_0(t)$, $|\dot{L}_0(t)| \leq M L_0(t)$, $L(t) = L_0(t) + \eta_L(t)$.

Introduce $t = \hat{t}/M$, $\hat{z}_0(\hat{t}) = z_0(\hat{t}/M)$, $\hat{f}(\hat{t}) = f(\hat{t}/M)$, $\hat{L}(\hat{t}) = L(\hat{t}/M)/M$, $\hat{L}_0(\hat{t}) = L_0(\hat{t}/M)/M$. Denoting $(\cdot)' = \frac{d}{d\hat{t}}(\cdot)$ obtain

$$\hat{z}_0' = -\lambda_0 \hat{L}(t) \operatorname{sign}(\hat{z}_0 - \hat{f}(\hat{t})) + \mu_0(\hat{z}_0 - \hat{f}(\hat{t})), \tag{B2}$$

where $|\hat{f}'| \leq \hat{L}_0$ and $|\hat{L}_0'| \leq \hat{L}_0$.

Now suppose that the substitution $t = \hat{t}/M$, $\hat{z}_i(\hat{t}) = z_i(\hat{t}/M)M^{-i}$, $\hat{f}(\hat{t}) = f(\hat{t}/M)$, $\hat{L}(\hat{t}) = L(\hat{t}/M)M^{-k-1}$, $\hat{L}_0(\hat{t}) = L_0(\hat{t}/M)M^{-k-1}$ reduces differentiators of the orders $k = 0, 1, ..., n-1$ to the form with $M = 1$. The nth-order differentiator (3), (4) gets the form

$$\begin{aligned} &\dot{z}_0 = v_0 = -\lambda_n L(t)^{\frac{1}{n+1}} \lfloor (z_0 - f(t)) \rceil^{\frac{n}{n+1}} + \mu_n M(z_0 - f(t)) + z_1, \\ &\tfrac{d}{dt}\overleftarrow{z}_1 = D_{n-1,\vec{\lambda}_{n-1},\vec{\mu}_{n-1}}(\overleftarrow{z}_1, v_0, L(t), M). \end{aligned} \tag{B3}$$

According to the induction assumption equation (B3) can be rewritten as

$$\begin{aligned} &\hat{z}_0' = \hat{v}_0 = -\lambda_n \hat{L}(t)^{\frac{1}{n+1}} \left\lfloor (\hat{z}_0 - \hat{f}(\hat{t})) \right\rceil^{\frac{n}{n+1}} + \mu_n(\hat{z}_0 - \hat{f}(\hat{t})) + \hat{z}_1, \\ &\tfrac{d}{d\hat{t}}\overleftarrow{\tilde{z}}_1 = D_{n-1,\vec{\lambda}_{n-1},\vec{\mu}_{n-1}}(\overleftarrow{\tilde{z}}_1, v_0(\hat{t}/M), L(\hat{t}/M)M^{-n}, 1). \end{aligned}$$

where $t = \hat{t}/M$, $\hat{z}_0(\hat{t}) = z_0(\hat{t}/M)$, $\hat{f}(\hat{t}) = f(\hat{t}/M)$, $\hat{z}_1 = z_1(\hat{t}/M)/M$, $\hat{v}_0 = v_0(\hat{t}/M)/M$, $\tilde{z}_i(\hat{t}) = z_i(\hat{t}/M)M^{-i+1}$, $i = 1, \ldots, n$.

Dividing now both sides of the second equation by M and denoting $\hat{z}_i = \tilde{z}_i(\hat{t}/M)/M$ obtain the required normal form. □

From now on we assume $M = 1$, $|\frac{\dot{L}_0}{L_0}| \leq 1$.

Subtracting $f_0^{(i+1)}(t)$ from the both sides of the equation for z_i of (3), dividing by $L_0(t)$ and denoting $\sigma_i = (z_i - f_0^{(i)})/L_0(t)$, $\xi = \eta(t)/L_0(t) \in \varepsilon[-1, 1]$, get $\dot{\sigma}_i = \sigma_{i+1} - (\dot{L}_0/L_0)\sigma_i$, $i = 0, ..., n-1$. Correspondingly obtain

$$\begin{aligned} &\dot{\sigma}_0 = -\lambda_n(1+\eta_L)^{\frac{1}{n+1}} \lfloor \sigma_0 + \xi \rceil^{\frac{n}{n+1}} - \mu_n(\sigma_0 + \xi) - \tfrac{\dot{L}_0}{L_0}\sigma_0 + \sigma_1, \ |\tfrac{\dot{L}_0}{L_0}| \leq 1, \\ &\dot{\sigma}_1 = -\lambda_{n-1}(1+\eta_L)^{\frac{1}{n}} \left\lfloor \sigma_1 - \dot{\sigma}_0 - \tfrac{\dot{L}_0}{L_0}\sigma_0 \right\rceil^{\frac{n-1}{n}} - \mu_{n-1}(\sigma_1 - \dot{\sigma}_0 - \tfrac{\dot{L}_0}{L_0}\sigma_0) - \tfrac{\dot{L}_0}{L_0}\sigma_1 + \sigma_2, \\ &\quad \ldots \\ &\dot{\sigma}_n = -\lambda_0(1+\eta_L) \left\lfloor \sigma_n - \dot{\sigma}_{n-1} - \tfrac{\dot{L}_0}{L_0}\sigma_{n-1} \right\rceil^0 - \mu_0(\sigma_n - \dot{\sigma}_{n-1} - \tfrac{\dot{L}_0}{L_0}\sigma_{n-1}) - \tfrac{\dot{L}_0}{L_0}\sigma_n - \tfrac{1}{L_0} f_0^{(n+1)}. \end{aligned} \tag{B4}$$

In the particular case $\dot{L}_0 = 0$, $\eta_L = \xi = 0$, after substituting $|f_0^{(n+1)}| \in [-L_0, L_0]$ system (B4) becomes a DI autonomous and homogeneous in bilimit (Andrieu, Praly, & Astolfi, 2008) with a negative approximating homogeneity degree at zero and the zero homogeneity degree at infinity. Correspondingly, it should feature exponential convergence from infinity and FT convergence in a vicinity of zero, if it is globally asymptotically stable.

The following normal form is suitable for large coordinate values and describes the transient from the ball $\|\vec{\sigma}_n\|_\infty \leq R$, $R \geq R_n$ for some R_n. Denote $\zeta_i = \sigma_i/R$. Dividing (B4) by R obtain

$$\begin{aligned} &\dot{\zeta}_0 = -\lambda_n R^{-\frac{1}{n+1}}(1+\eta_L)^{\frac{1}{n+1}} \left\lfloor \zeta_0 + \tfrac{\xi}{R} \right\rceil^{\frac{n}{n+1}} - \mu_n(\zeta_0 + \tfrac{\xi}{R}) - \tfrac{\dot{L}_0}{L_0}\zeta_0 + \zeta_1, \ |\tfrac{\dot{L}_0}{L_0}| \leq 1, \\ &\dot{\zeta}_1 = -\lambda_{n-1} R^{-\frac{1}{n}}(1+\eta_L)^{\frac{1}{n}} \left\lfloor \zeta_1 - \dot{\zeta}_0 - \tfrac{\dot{L}_0}{L_0}\zeta_0 \right\rceil^{\frac{n-1}{n}} - \mu_{n-1}(\zeta_1 - \dot{\zeta}_0 - \tfrac{\dot{L}_0}{L_0}\zeta_0) - \tfrac{\dot{L}_0}{L_0}\zeta_1 + \zeta_2, \\ &\quad \ldots \\ &\dot{\zeta}_n = -\lambda_0 R^{-1}(1+\eta_L) \left\lfloor \zeta_n - \dot{\zeta}_{n-1} - \tfrac{\dot{L}_0}{L_0}\zeta_{n-1} \right\rceil^0 - \mu_0(\zeta_n - \dot{\zeta}_{n-1} - \tfrac{\dot{L}_0}{L_0}\zeta_{n-1}) - \tfrac{\dot{L}_0}{L_0}\zeta_n \\ &\qquad\qquad - R^{-1}\tfrac{1}{L_0} f_0^{(n+1)}. \end{aligned} \tag{B5}$$

In the vicinity of zero the following form is suitable, which describes the transient from the homogeneous ball $\|\vec{\sigma}_n\|_{h\infty} \leq R$, $0 < R \leq \hat{R}_n$ for some $\hat{R}_n$. Denote $\hat{\zeta}_i = \sigma_i R^{-\frac{n+1-i}{n+1}}$, $\hat{t} = tR^{-\frac{1}{n+1}}$.

Dividing the equation on σ_i of (B4) by $R^{\frac{n-i}{n+1}}$ obtain

$$
\begin{aligned}
\hat\zeta_0' &= -\lambda_n(1+\eta_L)^{\frac{1}{n+1}}\left\lfloor \hat\zeta_0 + \tfrac{\xi}{R}\right\rceil^{\frac{n}{n+1}} - R^{\frac{1}{n+1}}\mu_n(\hat\zeta_0 + \tfrac{\xi}{R}) - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_0 + \hat\zeta_1,\ |\tfrac{\dot L_0}{L_0}| \le 1,\\
\hat\zeta_1' &= -\lambda_{n-1}(1+\eta_L)^{\frac{1}{n}}\left\lfloor \hat\zeta_1 - \hat\zeta_0' - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_0\right\rceil^{\frac{n-1}{n}} - R^{\frac{1}{n+1}}\mu_{n-1}(\hat\zeta_1 - \hat\zeta_0' - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_0)\\
&\qquad - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_1 + \hat\zeta_2,\\
&\dots\\
\hat\zeta_n' &= -\lambda_0(1+\eta_L)\left\lfloor \hat\zeta_n - \hat\zeta_{n-1}' - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_{n-1}\right\rceil^{0} - R^{\frac{1}{n+1}}\mu_0(\hat\zeta_n - \hat\zeta_{n-1}' - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_{n-1})\\
&\qquad - R^{\frac{1}{n+1}}\tfrac{\dot L_0}{L_0}\hat\zeta_n - R\tfrac{1}{L_0}f_{0\hat t}^{(n+1)},
\end{aligned}
\tag{B6}
$$

where $\hat\zeta_i' = \dot{\hat\zeta}_i R^{\frac{1}{n+1}}$ is the derivative with respect to $\hat t$, $R^{\frac{1}{n+1}}\frac{\dot L_0}{L_0} = \frac{L_0'}{L_0}$, $(\frac{d}{dt})^{n+1}f_0 = R(\frac{d}{d\hat t})^{n+1}f_0 = Rf_{0\hat t}^{(n+1)} \in [-L_0, L_0]$.

Note that systems (B4), (B5), (B6) are in the recursive form, i.e. derivatives of the state variables appear on the both sides. Nevertheless, similarly to (5) they have equivalent standard non-recursive form, which validates the application of the theory by Filippov (1988). In particular, their solutions continuously depend on the noise magnitudes $\varepsilon, \varepsilon_L$.

Appendix C. Proof of the Theorems

The proof of the theorems is based on Lemma 2 below which analyses the dynamics of the auxiliary systems (B4), (B5), (B6). Lemma 2 also describes the assignment of parameters λ_n, μ_n, $n \ge 1$, after the parameters $\{\lambda_i, \mu_i\}$, $i = 0, ..., n-1$, have been already set for the differentiation orders $0, 1, ..., n-1$.

The choice of $\lambda_0, \mu_0 > 1$ for $n = 0$ is arbitrary. In order to simplify the choice of parameters for $n \ge 1$ introduce **the artificial linkage** $\mu_i = \bar\mu_i\lambda_i^{(i+1)/i}$, where $\bar\mu_i > 0$, $i = 1, 2, ...$, are *any* fixed coefficients. Thus increasing/decreasing λ_i simultaneously increases/decreases μ_i.

Lemma 2 in its turn is based on Lemma 3 to be formulated later in Appendix D. Lemmas 2, 3 and Theorem 1 are proved together in parallel by induction in n. It means that the proofs of Lemmas use that the statements of Lemmas 2, 3 and Theorem 1 hold for the parameters $\{\lambda_i, \mu_i\}$, $i = 0, 1, ..., n-1$ and differentiation orders $0, 1, ..., n-1$.

Lemma 2: *Let the positive finite sequences $\{\lambda_i, \mu_i\}$, $i = 0, ..., n-1$, be chosen so that this lemma and Theorem 1 hold up to the differentiation order $n-1$, $n \ge 1$. There are no conditions for $n = 0$ except $\lambda_0, \mu_0 > 1$. Fix **any** $\bar\mu_n, R_n, \hat R_n > 0$, $R_n \ge R_{n-1}$, $\hat R_n \ge \hat R_{n-1}$, and let $\mu_n = \lambda_n^{(n+1)/n}\bar\mu_n > 1$. Then the following holds for $M = 1$, sufficiently large positive λ_n, ΔT_n and sufficiently small ε_L.*

1. *There exist such $q_n, Q_n > 0$ that for any $R \ge R_n$, $|\xi| \le q_n R$ all trajectories of* (B4) *starting within the ball $\|\vec\sigma_n\|_\infty \le R$, in the time ΔT_n concentrate within the ball $\|\vec\sigma_n\|_\infty \le R/2$ without leaving the ball $\|\vec\sigma_n\|_\infty \le Q_n R$ on the way. Then the trajectories stay in the ball $\|\vec\sigma_n\|_\infty \le R/2$ forever.*
2. *There exist such $\hat q_n, \hat Q_n > 0$ that for any R, $0 < R \le \hat R_n$, and $|\xi| \le \hat q_n R$ all trajectories of* (B4) *starting within the ball $\|\vec\sigma_n\|_{h\infty} \le R$, in the time $R^{\frac{1}{n+1}}\Delta T_n$ concentrate within the homogeneous ball $\|\vec\sigma_n\|_{h\infty} \le R/2$ without leaving the homogeneous ball $\|\vec\sigma_n\|_{h\infty} \le \hat Q R$ on the way. Then the trajectories stay in the ball $\|\vec\sigma_n\|_{h\infty} \le R/2$ forever.*

The lemma is proved in section D. Since $\hat R_n, R_n$ are chosen voluntarily, the regions of the approximate homogeneity are further made *overlapping.*

Proof of Theorem 1. Build a sequence of recursive differentiators (3), (4) for $n = 0, 1, \ldots$. According to Lemma 2 choose two monotonously growing sequences $R_n, \hat{R}_n > 0$, such that each ball (actually cube) $||\vec{\sigma}_n||_\infty \leq R_n$ lies entirely inside the homogeneous ball (parallelepiped) $||\vec{\sigma}_n||_{h\infty} \leq \hat{R}_n$. Fix some arbitrary $\lambda_0, \mu_0 > 1$, and an arbitrary sequence $\bar{\mu}_n > 0$ for $n > 0$. Now recursively choosing the sequences λ_n, $\mu_n = \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n$, $n = 1, 2, \ldots$, sufficiently large according to Lemma 2, obtain the infinite sequences λ_n, μ_n providing for the convergence of the differentiators (3), (4) for any $n \geq 0$.

Let $\varepsilon = 0$, $M = 1$. The FT stability of the error dynamics (B4) directly follows from Lemma 2. Indeed, let the initial moment be t_0. Then the convergence time is estimated as

$$T_{conv} \leq \Delta T_n \left(\log_2 \max[1, \tfrac{1}{R_n} \|\vec{\sigma}_n(t_0)\|_\infty] + \hat{R}_n^{\frac{1}{n+1}} \sum_{k=0}^{\infty} 2^{-k\frac{1}{n+1}} + 1 \right).$$

Taking into account the transformation described in Lemma 1 obtain the general estimation for $M > 0$.

Statement (1) of Lemma 2 implies the practical stability of error dynamics (B4). Indeed, for any ε all trajectories in FT stabilize in $||\vec{\sigma}_n||_\infty \leq R_\varepsilon$ for $R_\varepsilon = \max[R_n, \varepsilon/q_n]$. □

Proof of Theorem 2.

(1) Consider system (B6) with $\xi = 0, \eta_L = 0$, i.e. in the absence of noises. According to the theorem conditions, normalized errors uniformly converge to zero. Let R be close enough to zero. Using $\dot{L}_0/L_0 \in [-1, 1]$ obtain that solutions of (B6) approximate solutions of the system

$$\begin{aligned}
\hat{\zeta}_0' &= -\lambda_n \left\lfloor \hat{\zeta}_0 \right\rceil^{\frac{n}{n+1}} + \hat{\zeta}_1, \\
\hat{\zeta}_1' &= -\lambda_{n-1} \left\lfloor \hat{\zeta}_1 - \hat{\zeta}_0' \right\rceil^{\frac{n-1}{n}} + \hat{\zeta}_2, \\
&\ldots \\
\hat{\zeta}_n' &= -\lambda_0 \left\lfloor \hat{\zeta}_n - \hat{\zeta}_{n-1}' \right\rceil^{0} + u(\hat{t}),\ u(\hat{t}) = R(\tfrac{d}{d\hat{t}})^{n+1} f(\hat{t}) \in [-1, 1].
\end{aligned} \tag{C1}$$

Solutions of this system are exactly the solutions of the DI for u replaced with $[-1, 1]$ (Filippov, 1962), which describes the errors of the standard differentiator (1). It is homogeneous of the homogeneity degree -1 with the weights $\deg \hat{\zeta}_i = n + 1 - i$. Since solutions of (B6) uniformly converge to 0, DI (C1) features the contractivity property (Levant, 2005): its trajectories starting in the homogeneous parallelepiped $\|\vec{\hat{\zeta}}_n\|_{h\infty} \leq 1$ in FT gather in the homogeneous parallelepiped $\|\vec{\hat{\zeta}}_n\|_{h\infty} \leq 0.5$. Hence, DI (C1)is FT stable (FTS) (Levant, 2005).

(2) If differentiator converges for some $L_0(t)$, $|\dot{L}_0|/L_0 \leq M$, it also should converge for the same signals $f_0(t)$, but for $\tilde{L}_0, \tilde{M}$, where $\tilde{L}_0(t) = \alpha L_0(t)$, $|\dot{\tilde{L}}_0|/\tilde{L}_0 \leq M \leq \tilde{M} = \beta M$.

(3) The new sequences $\{\beta^{n_0/(i+1)}\lambda_i\}$, $\{\beta\mu_i\}$, $i = 0, 1, \ldots, n_0$, correspond to the time transformation $\tilde{t} = \beta^{-1}t$ that is equivalent to taking $\tilde{L}_0(t) = \beta^{n_0} L_0(t)$, $\tilde{M} = \beta M$.

□

Proof of Theorem 3. First let $R \geq R_n$. Consider the cases $R \in [R_n, R_M]$ and $R > R_M$ for some large R_M. For $R \in [R_n, R_M]$ the compactness reasoning works. For R sufficiently close to arbitrary $R_* \in [R_n, R_M]$ and for sufficiently small $\varepsilon, \varepsilon_L$ the right-hand side of (B4) divided by R and modified for solutions with discrete measurements *as graphs* (Filippov, 1988) uniformly approximates the graph of the right-hand side of (B5) for $R = R_*$. In the case $R > R_M$ all nonlinear terms in (B5) are negligible (see a similar detailed reasoning in the proof of Lemma 2). Thus, in both cases contraction Lemma 2 implies convergence to the region $||\sigma||_\infty \leq R_n$.

Let $R \le \hat{R}_n$. Also here for sufficiently small $\varepsilon, \varepsilon_L$ solutions with discrete measurements uniformly approximate the solutions (B6) for $R \in [R_*, \hat{R}_n]$ for a small enough $R_* > 0$. Due to continuous dependence of the solutions on the right-hand side (Filippov, 1988) and the compactness of the segment $[R_*, \hat{R}_n]$ contraction Lemma 2 remains true, which implies the convergence into a small vicinity of zero.

Thus the solutions gather in a small vicinity of 0. Fix some value $L_* = L_0(t_*)$, and let $\varpi_i = \frac{1}{L_*}[z_i - f_0^{(i)}] = \frac{L_0}{L_*}\sigma_i$, then solutions of system (B4) satisfy

$$\begin{array}{l} \dot{\varpi}_0 = -\hat{\varphi}_0(\varpi_0 + \eta/L_*) + \varpi_1, \\ \dot{\varpi}_1 = -\hat{\varphi}_1(\hat{\varphi}_0(\varpi_0 + \eta/L_*)) + \varpi_2, \\ \quad \ldots \\ \dot{\varpi}_n \in -\hat{\varphi}_n(t, \hat{\varphi}_{n-1}(t, \ldots\hat{\varphi}_0(\varpi_0 + \eta/L_*))\ldots)) - \frac{L_0}{L_*}\frac{1}{L_0} f_0^{(n+1)}, \\ \text{where } \hat{\varphi}_i(t,s) = \lambda_{n-i}(1 + \frac{\eta_L(t)}{L_0(t)})^{\frac{1}{n-i+1}}[\frac{L_0(t)}{L_*}]^{\frac{1}{n-i+1}}\lfloor s\rceil^{\frac{n-i}{n-i+1}} + \mu_{n-i}s. \end{array} \tag{C2}$$

Hence for some small $\delta > 0$ during the time interval $[t_*, t_* + \delta]$ get

$$|\tfrac{\dot{L}_0}{L_0}| \le 1,\ \tfrac{L_0}{L_*} \in [e^{-\delta}, e^{\delta}],\ \dot{\sigma}_i + \tfrac{\dot{L}_0}{L_0}\sigma_i = \tfrac{1}{L_0}\tfrac{d}{dt}(z_i - f_0^{(i)}) \in [e^{-\delta}, e^{\delta}]\dot{\varpi}_i \text{ for } t \in [t_*, t_* + \delta],\ i = 0, \ldots, n. \tag{C3}$$

Within a sufficiently small parallelepiped $B(\rho)$ defined by the inequality $\|\vec{\varpi}\|_{h\infty} \le e^{\delta}\rho$ for sufficiently small noises η, η_L the linear terms of $\hat{\varphi}_i$ are small compared with the nonlinear ones. Thus, due to (C3) for sufficiently small δ solutions of (C2) satisfy the DI

$$\begin{array}{l} \dot{\varpi}_0 \in -\tilde{\lambda}_n[1-\Delta, 1+\Delta]\lfloor \varpi_0 + \varepsilon[1-\Delta, 1+\Delta]\rceil^{\frac{n}{n+1}} + \varpi_1, \\ \dot{\varpi}_1 \in -\tilde{\lambda}_{n-1}[1-\Delta, 1+\Delta]\lfloor \varpi_0 + \varepsilon[1-\Delta, 1+\Delta]\rceil^{\frac{n-1}{n+1}} + \varpi_2, \\ \quad \ldots \\ \dot{\varpi}_n \in -\tilde{\lambda}_0[1-\Delta, 1+\Delta]\lfloor \varpi_0 + \varepsilon[1-\Delta, 1+\Delta]\rceil^{0} + (1+\Delta)[-1, 1], \end{array} \tag{C4}$$

where $\Delta > 0$ is some small constant determined by the size ρ of the vicinity $B(\rho)$, $||\varpi||_\infty \le \rho$. It is a disturbed homogeneous FTS system (C1) of errors of the standard differentiator (1), $\deg \varpi_i = n + 1 - i$, of the homogeneity degree -1.

According to Theorem 2 the standard differentiator dynamics with the parameters λ_i is FTS. Thus for sufficiently small $\Delta > 0$ and $\varepsilon = 0$ a FTS inclusion is obtained (Levant, 2005). Fix such value $\rho_0 > 0$. Errors of (5) corresponding to the discrete noisy sampling can be considered as solutions of DI (C2) corresponding to the noisy discrete sampling of ϖ_0 with the noise magnitude ε.

In the absence of noises and sampling delays the trajectories starting within $B(\rho)$ at the time 0 finish at $\varpi = 0$ in some finite time $T > 0$ (Levant, 2005; Levant & Livne, 2016). In the presence of small noises, i.e. for small ε, the trajectories will finish in some vicinity of 0 *inside* $B(\rho)$ (Filippov, 1988). The points of these trajectories over the time interval $[0, T]$ constitute the set $\Omega(\rho, \varepsilon)$ invariant with respect to (C4). The region $\Omega(\rho, \varepsilon)$ contracts to 0 as $\rho, \varepsilon, \tau \to 0$. Reduce and fix the values of ρ, $\varepsilon \le q_n R_n$, τ for which $\Omega(\rho, \varepsilon)$ lies in the interior of $B(\hat{R}_n)$. Hence, for such noises and discrete measurements all trajectories converge into $B(\hat{R}_n/2)$ to stay there.

Theorem 3 is now obtained by standard homogeneous technique (Levant & Livne, 2016). □

Proof of Theorem 4. Similarly to the proof of Theorem 3, due to Lemma 2 all solutions converge into a small homogeneous ball $\|\vec{\sigma}\| \le R_*$, $\hat{Q}_n R_* < \hat{R}_n$. Using the Taylor expansion

$$f_0^{(i)}(t_{k+1}) = f_0^{(i)}(t_k) + \textstyle\sum_{j=i+1}^{n} \frac{f_0^{(j)}(t_k)}{(j-1)!}\tau_k^{j-i} + \frac{\theta_i}{(n-i+1)!}\tau_k^{n-i+1},\ |\theta_i| \le L_0(t_k)e^{\tau_k},$$

subtracting $f_0^{(i+1)}(t_{k+1})$ from the both sides of the equation for z_i of (7) and dividing by $L_* = L(t_*)$, get

$$\begin{aligned}
\varpi_0(t_{k+1}) &= \varpi_0(t_k) - \hat\varphi_0(t_k, \varpi_0(t_k) + \eta(t_k)/L_*)\tau_k + \sum_{j=1}^{n} \tfrac{\varpi_j(t_k)}{j!}\tau_k^j - \tfrac{\theta_0}{(n+1)!}\tau_k^{n+1},\\
\varpi_1(t_{k+1}) &= \varpi_1(t_k) - \hat\varphi_1(t_k, \hat\varphi_0(t_k, \varpi_0(t_k) + \eta(t_k)/L_*))\tau_k + \sum_{j=2}^{n} \tfrac{\varpi_j(t_k)}{(j-1)!}\tau_k^{j-1} - \tfrac{\theta_1}{n!}\tau_k^n,\\
&\ldots\\
\varpi_n(t_{k+1}) &= \varpi_n(t_k) - \hat\varphi_n(t_k, \hat\varphi_{n-1}(t_k, ...(t_k, \varpi_0(t_k) + \eta(t_k)/L_*))...))\tau_k - \theta_n\tau_k.
\end{aligned}$$

Within the ball $\|\vec\varpi\| \le R_*$ for sufficiently small ΔT any solution of the above discrete system can be presented as a sampled solution of the differential inclusion

$$\begin{aligned}
\dot\varpi_0 &\in -\tilde\lambda_n[1-\Delta, 1+\Delta] \times \lfloor \varpi_0(t-\tau[0,1]) + \varepsilon[-1-\Delta, 1+\Delta]\rceil^{\frac{n}{n+1}}\\
&\qquad + \sum_{j=1}^{n} \tfrac{\varpi_j(t-\tau[0,1])}{j!}\tau^{j-1}[-1,1] - \tfrac{[-1-\Delta,1+\Delta]}{(n+1)!}\tau^n,\\
\dot\varpi_1 &\in -\tilde\lambda_{n-1}[1-\Delta, 1+\Delta] \times \lfloor \varpi_0(t-\tau[0,1]) + \varepsilon[-1-\Delta, 1+\Delta]\rceil^{(n-1)/(n+1)}\\
&\qquad + \sum_{j=2}^{n} \tfrac{\varpi_j(t-\tau[0,1])}{j!}\tau^{j-1}[-1,1] - \tfrac{[-1-\Delta,1+\Delta]}{n!}\tau^{n-1},\\
&\ldots\\
\dot\varpi_n &\in -\tilde\lambda_0[1-\Delta, 1+\Delta] \times \lfloor \varpi_0(t-\tau[0,1]) + \varepsilon[-1-\Delta, 1+\Delta]\rceil^0 + (1+\Delta)[-1,1],
\end{aligned} \tag{C5}$$

where $\Delta > 0$ is small. Obtained DI (C5) is homogeneous with respect to the transformation $t \mapsto \kappa t$, $\tau \mapsto \kappa\tau$, $\varepsilon \mapsto \kappa^{n+1}\varepsilon$, $\varpi^{(i)} \mapsto \kappa^{n-i+1}\varpi^{(i)}$, $i = 0, ..., n$, which transfers solutions into solutions, while changing the system parameters ε and τ. The rest of the proof is based on (Levant & Livne, 2016) and is similar to (Livne & Levant, 2014).

The value of R_* is chosen so that for $\varepsilon = \varepsilon_L = 0$ the above homogeneous inclusion converges to zero during the time interval of the length δ, whereas the noise magnitudes $\varepsilon, \varepsilon_L$ are to be so small as to provide for the invariance of $\|\vec\varpi\| \le R_*$ with respect to DI (C5) (Levant & Livne, 2016). □

Appendix D. Proof of Lemma 2

Remark 4: We often apply the following simple logic. Let $A(p)$, $B(p)$ be Boolean functions depending on the parameter $p \in P$, and the goal be to prove that $\exists p: A(p)$, i.e. that $A(p)$ holds for some p. *Assume* that $B(p)$ holds. Basing on that assumption find a set P_B, such that $A(p)$ is true for $p \in P_B$. Now choose a subset $P_A \subset P_B$ such that $\forall p \in P_A: B(p)$ holds. The set P_A is a set we need to be not empty. Indeed $\forall p \in P_A: B(p)$, $\forall p \in P_A: B(p) \to A(p)$ is the formal proof. A variant of the same technique is $\forall\alpha\exists p: A(\alpha, p)$ proved as

$$\forall\alpha\exists P_A \subset P: [P_A \ne \emptyset, \quad \forall p \in P_A: (B(\alpha,p) \to A(\alpha,p))\ \&\ B(\alpha,p)].$$

The proof of Lemma 2 exploits the next technical lemma to be proved in parallel by induction.

Lemma 3: *Let the finite sequences $\{\lambda_i, \mu_i\}$, $i = 0, ..., n-1$, $\lambda_0, \mu_0 > 1$, be chosen so that this lemma, Lemma 2 and Theorem 1 hold up to the differentiation order $n-1$, $n = 0, 1, ...$. There are no conditions for $n = 0$. Let $\bar\mu_n > 0$, $\mu_n = \lambda_n^{(n+1)/n}\bar\mu_n > 1$, $R_n, \hat R_n > 0$. Then the following holds for $M = 1$, sufficiently large λ_n and sufficiently small ε_L.*

(1) Let $||\vec\sigma_n(t_0)||_\infty \le R$, $R \ge R_n$, and $\xi = \xi_\omega(t)$ satisfy the condition that the integral $\int_{I_\omega} |\xi_\omega(t)/R|dt$ over the time interval $I_\omega = [t_0, t_0+\omega]$, $\omega > 0$, is less than some fixed $K > 0$.

Then there are such positive $W(K) > 1$ and $W_\xi(K)$ that for any sufficiently small ω any trajectory of (B5) *starting within the ball $\|\vec{\zeta}_n\|_\infty \le 1$ does not leave the ball $\|\vec{\zeta}_n\|_\infty \le W$ during the time interval I_ω. Moreover, keeping $|\xi| \le W_\xi R$ for $t \ge t_0 + \omega$ ensures keeping $\|\vec{\zeta}_n\|_\infty \le W$ forever.*

(2) *Let $||\vec{\sigma}_n(t_0)||_{h\infty} \le R$, $R \le \hat{R}_n$, and $\xi = \xi_\omega(t)$ satisfy the condition that the integral $\int_{I_\omega} |\xi_\omega(t)/R|dt$ over the time interval $I_\omega = [t_0, t_0 + R^{\frac{1}{n+1}}\omega]$, $\omega > 0$, is less than some fixed $K > 0$. Then there is such positive $\hat{W}_\xi(K) > 1$ and $\hat{W}(K)$ that for any sufficiently small ω any trajectory of* (B6) *starting within the ball $\|\vec{\hat{\zeta}}_n\|_{h\infty} \le 1$ does not leave the ball $\|\vec{\hat{\zeta}}_n\|_{h\infty} \le \hat{W}$ during the time interval I_ω. Moreover, keeping $|\xi| \le \hat{W}_\xi R$ for $t \ge t_0 + \omega$ ensures keeping $\|\vec{\hat{\zeta}}_n\|_{h\infty} \le \hat{W}$ forever.*

Proof of Lemma 2. Lemma 2 is obviously true for $n = 0$, since then for $|\xi| \le q_0 R$ obtain

$$\begin{array}{l} \dot{\zeta}_0 \in -\frac{\lambda_0}{R}[1-\varepsilon_L, 1+\varepsilon_L]\,\mathrm{sign}(\zeta_0 + q_0[-1,1]) - \mu_0(\zeta_0 + q_0[-1,1]) + [-1,1]\zeta_0 + \frac{1}{R}[-1,1], \\ \hat{\zeta}_0' \in -\lambda_0[1-\varepsilon_L, 1+\varepsilon_L]\,\mathrm{sign}(\hat{\zeta}_0 + q_0[-1,1]) - R\mu_0(\hat{\zeta}_0 + q_0[-1,1]) + R[-1,1]\hat{\zeta}_0 + [-1,1]. \end{array} \tag{D1}$$

Note that ζ_0, $\hat{\zeta}_0$ in FT converge to zero for any $R > 0$, provided $\varepsilon_L < 1 - \lambda_0^{-1}$, $q_0 = 0$, $\lambda_0, \mu_0 > 1$. The statements hold for any $R_0 > 0$, $R \ge R_0$, $q_0 \le 0.5R_0$ for ζ_0, and for any $\hat{R}_0 > 0$, $R \le \hat{R}_0$, $\hat{q}_0 \le 0.5R$ for $\hat{\zeta}_0$.

Induction step. Let now Lemmas 2, 3 and Theorem 1 be true for the differentiation orders $0, 1, ..., n-1$. The plan of the induction step proof is as follows.

Consider system (B5) for $R \ge R_n$ (system (B6) for $R \le \hat{R}_n$). At the initial time t_0 get $||\zeta||_\infty \le 1$ (respectively $||\hat{\zeta}||_{h\infty} \le 1$).

Introduce parameter $p = (\lambda_{n0}, \varepsilon_{L0}, q_n(\cdot))$, where $q_n(\lambda_n) > 0$ is some function of $\lambda_n \ge \lambda_{n0}$. Define event A as the ability of making the variables $\zeta_0, \dot{\zeta}_0$ (respectively $\hat{\zeta}_0, \hat{\zeta}_0'$) arbitrarily small in arbitrarily short time and keeping them small forever. Formally it means that for any predefined $d_0, d_1, d_2 > 0$ there exists such p that for any $\lambda_n \ge \lambda_{n0}$, $\varepsilon_L \le \varepsilon_{L0} < 1, q_n \le q_{n0}(\lambda_n)$ the inequalities $|\zeta_0| \le d_0$, $|\dot{\zeta}_0| \le d_1$ (respectively $|\hat{\zeta}_0| \le d_0$, $|\hat{\zeta}_0'| \le d_1$) are kept for any $t \ge t_0 + d_2$ over any solution.

The first equation of (B5) (of (B6)) produces the input $\theta = \tilde{\xi}/R = \dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0$ (respectively $\theta = \hat{\zeta}_0' + R^{\frac{1}{n+1}}\frac{\dot{L}_0}{L_0}\hat{\zeta}_0$) for the lower $(n-1)$th-order error system. We assume that $||\overleftarrow{\zeta}_1||_\infty$ for $R \ge R_n$ ($||\overleftarrow{\hat{\zeta}}_1||_{h\infty}$ for $R \le \hat{R}_n$) keeps the known bound $W > 1$ due to Lemma 3 applied to the lower subsystem with already fixed parameters. Lemma 3 imposes its conditions (event B of Remark 4) on θ: (i) $\int_{I_\omega} |\theta|dt \le K = 2$ for $\omega \le \omega_0$, where ω is the corresponding transient time; (ii) the inequality $|\theta| \le W_\xi$ holds outside of I_ω. The values $\omega_0, W, W_\xi > 0$ are produced by Lemma 3 in correspondence to $K = 2$, ω_0 can be arbitrarily reduced without changing W, W_ξ.

Part 1. "High-gain" reasoning provides for the statement A provided B holds. *Part 2.* Further increasing λ_{n0} and decreasing ε_{L0}, q_{n0} fulfil the conditions of Lemma 3 (event B). Reversing the calculations prove that conditions of Lemma 3 (event B) hold for the resulting p. Thus, due to Remark 4 event A holds. *Part 3.* Since ζ_0 and $\hat{\zeta}_0$ are already kept small, Lemma 2 for the order n follows now from Lemma 2 applied to the lower subsystem.

The case of approximate linearity, $R \ge R_n$. Consider (B5) with $||\zeta(t_0)|| \le 1$ at the initial moment t_0.

Part 1. According to the listed plan let W be an upper bound of $\|\overleftarrow{\zeta}_1\|$. Then $|\zeta_1| \le W$. Note that

$\|\overleftarrow{\zeta}_1(t_0)\| \leq \|\zeta(t_0)\| \leq 1 < W$. It follows from (B5) that

$$\dot{\zeta}_0 \in -\lambda_n[R(1+[-\varepsilon_L,\varepsilon_L])]^{-\frac{1}{n+1}}\lfloor\zeta_0+[-q_n,q_n]\rceil^{\frac{n}{n+1}}+\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n(\zeta_0+[-q_n,q_n])-\tfrac{\dot{L}_0}{L_0}\zeta_0+\zeta_1. \quad \text{(D2)}$$

Recall that $\mu_n = \lambda_n^{(n+1)/n}\bar{\mu}_n > 1$, $|\frac{\dot{L}_0}{L_0}| \leq 1$. It is easy to see that for sufficiently small ε_L, q_n, and sufficiently large λ_n, i.e. $\lambda_n \geq \lambda_{n0}$, $\varepsilon_L \leq \varepsilon_{L0}$, $|q_n| \leq q_{n0}$, get

$$\begin{array}{ll} \zeta_0\dot{\zeta}_0 < 0, \quad |\dot{\zeta}_0| \geq 2W & \text{for } |\zeta_0| \geq \beta_n(\lambda_n);\ \beta_n = 4W\bar{\mu}_n^{-1}\lambda_n^{-\frac{n+1}{n}}, \\ |\dot{\zeta}_0| \leq W_* = 5W + 3\bar{\mu}_n^{-\frac{n+1}{n}}W^{\frac{n+1}{n}} & \text{for } |\zeta_0| \leq \beta_n(\lambda_n), \end{array} \quad \text{(D3)}$$

Moreover, $\beta_n(\lambda_n)$ becomes arbitrarily small if λ_{n0} is further increased. Note the redundancy added to the lower bound of (D3) for the robustness. If $|\zeta_0(t_0)| \leq \beta_n$ it is kept further together with $|\dot{\zeta}_0| \leq W_*$.

The upper estimation of the time ΔT_{n0} needed for the entrance of ζ_0 into the segment $[-\beta_n, \beta_n]$ can be done as small as needed independently of $R \geq R_n$ and small noises, $\varepsilon_L \leq \varepsilon_{L0}$, $|q_n| \leq q_{n\beta}\beta_n$, by enlarging λ_n. Indeed, (D2) is FTS for $\zeta_1 \equiv 0$, $q_n = 0$.

Taking into account that during that transient neither ζ_0 nor $\dot{\zeta}_0$ change their sign, and ζ_0 changes by not more than $1-\beta_n < 1$, from (D3) obtain that $\int|\dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0|dt \leq 1 - \beta_n + W_*\Delta T_{n0} \leq 1$ for sufficiently large λ_{n0}. Recall that the constant W is defined by Lemma 3 for $K = 2$. Require $\beta_n < W_\xi/2$. It also may require additional increase of λ_{n0}. Hence due to Lemma 3 the upper bounds W and W_ξ depend only on $\vec{\lambda}_{n-1}$, $\vec{\mu}_{n-1}$, i.e. are independent of λ_n.

Part 2. Prove now that after the establishment of $|\zeta_0| \leq \beta_n$ at some moment t_* also the input $\dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0$ to the next equation of (B5) becomes small in arbitrarily short time ΔT_{n1}.

The further proof idea is as follows. Due to our construction ζ_0 enters the segment $[-\beta_n, \beta_n]$ with relatively high velocity $\dot{\zeta}_0$ and stays there. While remaining in this small segment ζ_0 slows down, so that also $\dot{\zeta}_0$ enters its own small interval on the axis $\dot{\zeta}_0$ to stay there. During the first part ΔT_{n1} of this braking process $\dot{\zeta}_0$ cannot change its sign. Thus $\int|\dot{\zeta}_0|dt \leq 2\beta_n$ over the time ΔT_{n1} and, provided $\beta_n \leq 1/8$, Lemma 3 is applicable to the whole time interval $\omega = \Delta T_{n0} + \Delta T_{n1}$ for the integral value $\int|\dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0|dt \leq 1 + 2\beta_n + \beta_n\Delta T_{n1} \leq 1.5 \leq K = 2$ for sufficiently large λ_n.

Fix some $\delta > 0$ and let $L_* = L_0(t_*)$, $\varpi_i = \frac{1}{L_*R}[z_i - f_0^{(i)}] = \frac{L_0}{L_*}\zeta_i$. Introduce the virtual noise $\tilde{\eta}_L(t)$, $|\tilde{\eta}_L(t)| \leq \tilde{\varepsilon}_L$. Then during the time interval $[t_*, t_* + \delta]$ get

$$\tfrac{L_0}{L_*}, \tfrac{L_*}{L_0} \in [e^{-\delta}, e^{\delta}],\ \dot{\zeta}_i + \tfrac{\dot{L}_0}{L_0}\zeta_i = \tfrac{L_0}{L_*}\dot{\varpi}_i \text{ for } i \geq 0,\ |\varpi_0| \leq e^{\delta}\beta_n,\ |\varpi_i| \leq e^{\delta}W \text{ for } i \geq 1, \quad \text{(D4)}$$

and, taking $\tilde{\eta}_L(t) = \eta_L(t)$, correspondingly get

$$\begin{aligned} \dot{\varpi}_0 &= -\lambda_n(\tfrac{L_0}{L_*R}(1+\tilde{\eta}_L))^{\frac{1}{n+1}}\left\lfloor\varpi_0 + \tfrac{L_0}{L_*}\tfrac{\xi}{R}\right\rceil^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}}\bar{\mu}_n(\varpi_0 + \tfrac{L_0}{L_*}\tfrac{\xi}{R}) + \varpi_1, \\ \dot{\varpi}_1 &= -\lambda_{n-1}(\tfrac{L_0}{L_*R}(1+\eta_L))^{\frac{1}{n}}\lfloor\varpi_1 - \dot{\varpi}_0\rceil^{\frac{n-1}{n}} - \mu_{n-1}(\varpi_1 - \dot{\varpi}_0) + \varpi_2, \\ &\dots \\ \dot{\varpi}_n &\in -\lambda_0\tfrac{L_0}{L_*R}(1+\eta_L)\lfloor\varpi_n - \dot{\varpi}_{n-1}\rceil^0 - \mu_0(\varpi_n - \dot{\varpi}_{n-1}) + e^{\delta}[-1,1]/R. \end{aligned} \quad \text{(D5)}$$

We need to prove that $\dot{\varpi}_0$ rapidly (in arbitrarily short time) becomes small in order to justify the application of Lemma 3 (event B). First consider the case $\xi \equiv 0$, $q_n = 0$, $\tilde{\eta}_L \equiv 0$. Differentiating the first equation of (D5) obtain

$$\ddot{\varpi}_0 = -\lambda_n\tfrac{n}{n+1}\left(\tfrac{L_0}{L_*R}\right)^{\frac{1}{n+1}}|\varpi_0|^{-\frac{1}{n+1}}\dot{\varpi}_0 - \tfrac{\lambda_n}{n+1}\left(\tfrac{L_0}{L_*R}\right)^{\frac{1}{n+1}}\tfrac{\dot{L}_0}{L_0}\lfloor\varpi_0\rceil^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}}\bar{\mu}_n\dot{\varpi}_0 + \dot{\varpi}_1. \quad \text{(D6)}$$

Therefore, due to $|\dot{L}_0| \leq L_0$, $L_0/L_* \in [e^{-\delta}, e^{\delta}]$, get from (D6) that

$$\begin{aligned}\ddot{\varpi}_0 \in & -\lambda_n \tfrac{n}{n+1} R^{-\frac{1}{n+1}} \left[e^{-\frac{\delta}{(n+1)}}, e^{\frac{\delta}{(n+1)}}\right] |\varpi_0|^{-\frac{1}{n+1}} \dot{\varpi}_0 \\ & + \tfrac{\lambda_n}{n+1} R^{-\frac{1}{n+1}} \left[-e^{\delta}, e^{\delta}\right] \lfloor \varpi_0 \rceil^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n \dot{\varpi}_0 + \dot{\varpi}_1.\end{aligned} \tag{D7}$$

It follows from (D3), (D4), (D5) that $|\dot{\varpi}_1| \leq e^{\delta}[(\lambda_{n-1} R_n^{-1/n} + \mu_{n-1})(W + W_* + \beta_n) + W]$. On the other hand provided $e^{\delta} < n e^{-\delta/(n+1)}$ and $|\varpi_0| \leq e^{\delta}\beta_n$ get that the first term on the right of (D7) is larger in its absolute value than the second for $|\dot{\varpi}_0| \geq e^{\delta}\beta_n$. Thus, for $\delta < \ln n$ and sufficiently large λ_n the derivative $\dot{\varpi}_0$ enters the segment $|\dot{\varpi}_0| \leq e^{\delta}\beta_n$ to stay there in some arbitrarily short time ΔT_1. That satisfies the conditions of Lemma 3 for sufficiently large λ_{n0} and $q_n = \tilde{\varepsilon}_L = 0$ and $\omega = \Delta T_0 + \Delta T_1$.

Let now $q_n, \tilde{\varepsilon}_L$ be small positive numbers. In a vicinity of each concrete R the required conditions are obtained for sufficiently small $q_n = q_n(R)$ and $\tilde{\varepsilon}_L$ due to the continuous dependence of the solutions on the right-hand side over the time interval $\omega + 1$, since they were kept with some redundancy. Then the conditions are trivially extended to the next interval of the length 1, etc. Unfortunately, the parametric set $R \geq R_n$ is not compact. Introduce an artificial upper bound $R \leq R_M$, and consider the cases $R \in [R_n, R_M]$ and $R > R_M$ separately. The value $R_M > 1$ is assigned below and is to be sufficiently large.

First let $R > R_M$. Denote

$$\Xi(R_M, \beta_n, q_n) = \lambda_n((1 + \tilde{\varepsilon}_L)\tfrac{e^{\delta}}{R_M})^{\frac{1}{n+1}} |\beta_n + (1 + \tilde{\varepsilon}_L) e^{\delta} \tfrac{q_n}{R_M}|^{\frac{n}{n+1}} [-1, 1].$$

The set Ξ contracts to $\{0\}$ as $R_M \to \infty$, $q_n \to 0$. Correspondingly obtain

$$\begin{aligned} \dot{\varpi}_0 &= u - \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n (\varpi_0 + (1 + \tilde{\eta}_L) \tfrac{L_0}{L_*} \tfrac{\xi}{R}) + \varpi_1, \ u(t) \in \Xi(R_M, \beta_n, q_n) \\ \dot{\varpi}_1 &= -\lambda_{n-1} (\tfrac{L_0}{L_* R}(1 + \eta_L))^{\frac{1}{n}} \lfloor \varpi_1 - \dot{\varpi}_0 \rceil^{\frac{n-1}{n}} - \mu_{n-1}(\varpi_1 - \dot{\varpi}_0) + \varpi_2, \\ & \dots \\ \dot{\varpi}_n &\in -\lambda_0 \tfrac{L_0}{L_* R}(1 + \eta_L) \lfloor \varpi_n - \dot{\varpi}_{n-1} \rceil^0 - \mu_0(\varpi_n - \dot{\varpi}_{n-1}) + e^{\delta}[-1, 1]/R, \end{aligned} \tag{D8}$$

Consider the case when not only the noises $\xi, \tilde{\eta}_l$ are absent, but also the generalized noise Ξ is removed from (D8), $R > R_M$. Obviously, due to the induction assumption the statement of Lemma 2 is right without that "noise" for any $R > R_M$, for the same value β_n. Adding the real and the generalized noise for sufficiently small $q_n, \tilde{\varepsilon}_L$ and sufficiently large R_M and taking $|\eta_L| \leq \min(\tilde{\varepsilon}_L, \varepsilon_{L0})$ obtain the conditions of Lemma 3 for $R > R_M$ due to the continuous dependence of solutions on the right-hand side of the DI obtained for $\xi \in q_n R[-1, 1], \eta_l \in \varepsilon_L[-1, 1], \tilde{\eta}_l \leq \tilde{\varepsilon}_L[-1, 1]$.

Now let $R \leq R_M$. For each $R \in [R_n, R_M]$ there is some small vicinity of R such that the lemma statement remains true for sufficiently small q_n, $\tilde{\varepsilon}_L$ and $|\xi| \leq q_n R$ due to the continuous dependence of solutions on the right-hand side (Filippov, 1988). By the compactness of $[R_n, R_M]$ Lemma 2 is then true for any $R \in [R_n, R_M]$ for sufficiently small q_n, $\tilde{\varepsilon}_L$. Choosing the least values of q_n and $\tilde{\varepsilon}_L, \varepsilon_{L0}$ from the both cases $R \in [R_n, R_M]$ and $R > R_M$ obtain the lemma conditions in the presence of noises.

Finishing Remark-4 logic. Show that the parameter choice $p = (\lambda_{n0}, \varepsilon_{L0}, q_n(\lambda_n))$ provides for the conditions of Lemma 3 for $K = 2$ applied to the lower subsystem (event B of Remark 4). Recall that, as the result, Lemma 3 produces parameters $W, W_\xi, \omega_0 > 0$, $W > 1$. Choose d_1, d_2 (see formulation of A) and the corresponding p so that $|\theta| \leq W_\xi/2$ holds outside of I_ω.

Event B has two conditions (i), (ii) to be true for $t \geq t_0$. Recall that the transient time ω introduced above is a function of p, W, ω_0 and satisfies $0 < \omega \leq \omega_0$. Let B_t be the condition that (i), (ii) hold over the time segment $[t_0, t]$. Naturally, condition (i) is redefined as $\int_{t_0}^{t} |\theta| dt \leq 2$ for $t_0 \leq t \leq t_0 + \omega$. Condition (ii) still means that $|\theta(t)| \leq W_\xi$ for $t - t_0 \geq \omega$. Then $B = (\forall t \geq t_0 : \ B_t)$.

Obviously B_{t_0} always holds. Suppose B_{t_1} is wrong at some time $t_1 > t_0$, then conditions are violated on some solution of (B5). At the last moment t_*, $t_0 \le t_* \le t_1$, when B_{t_*} is still right, $t_* = \sup\{t \in [t_0, t_1]\ |\forall \tau \in [t_0, t] :\ B_\tau\}$, get that either $\int_{t_0}^{t_*} |\theta| dt = 2$, $t_* - t_0 \le \omega \le \omega_0$, or $0 \le \omega < t_* - t_0$, $\int_{I_\omega} |\theta| dt \le 2$ and $|\theta| \ge W_\xi$.

In both cases B_t holds over $[t_0, t_*]$, and all calculations are valid over this interval. In the first case, due to the choice of p the equality $\int_{t_0}^{t_*} |\theta| dt = 2$ is not possible (λ_n is chosen so as to provide for $\int_{I_\omega} |\theta| dt \le 1.5$). In the second case another contradiction is got, since p is chosen so that $|\theta(t)| \le W_\xi/2$ whenever $t - t_0 \ge \omega$. Thus B holds, and therefore also A holds.

Part 3. Meantime we got that $||\overleftarrow{\zeta_1}||_\infty \le W$ is indefinitely kept. But we need the contraction $||\overleftarrow{\zeta_1}||_\infty \le 1/2$. Require $\beta_n < q_{n-1}$ which may impose additional increase on $\lambda_{n,0}$. Now due to the validity of Lemma 2 for lower orders, the subsystem $\overleftarrow{\sigma}_1$ converges in finite time $\Delta T_{n2} > 0$ from $||\overleftarrow{\sigma}_1||_\infty \le WR$ into $||\overleftarrow{\sigma}_1||_\infty \le R/2$. It corresponds to the convergence of $\overleftarrow{\zeta_1}$ into $||\overleftarrow{\zeta_1}||_\infty \le 1/2$.

Choose any ΔT_n, $\Delta T_n > 2(\Delta T_{n0} + \Delta T_{n1} + \Delta T_{n2})$, and any Q_n, $Q_n > W$.

Since $\delta < \ln n$ and β_n can be done arbitrarily small obtain the statement of the lemma for any $R > R_n$, sufficiently large λ_{n0}, and sufficiently small q_n, ε_{L0} over the time interval ΔT_n. At the middle of that interval apply the proved result for the next time interval ΔT_n, thus prolonging it by $0.5\Delta T_n$, etc.

The case of the negative approximate homogeneity degree, $R \le \hat{R}_n$. The proof has the same logic and actually extends the method from (Levant, 2003). We only provide the main points.

Part 1. System (B6) implies that

$$\hat{\zeta}_0' \in -\lambda_n (1 + \varepsilon_L[-1,1])^{\frac{1}{n+1}} \left\lfloor \hat{\zeta}_0 + \hat{q}_n[-1,1] \right\rceil^{\frac{n}{n+1}} - R^{\frac{1}{n+1}} \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n (\hat{\zeta}_0 + \hat{q}_n[-1,1]) - R^{\frac{1}{n+1}} \frac{\dot{L}_0}{L_0} \hat{\zeta}_0 + \hat{\zeta}_1, \quad \text{(D9)}$$

Once more let W, W_ξ correspond to Lemma 3 for the $(n-1)$th order sub-system and $K = 2$. Let W be an upper bound of $\|\overleftarrow{\hat{\zeta}}_1\| \ge |\hat{\zeta}_1|$. In the presence of noises for sufficiently small $\varepsilon_L, \hat{q}_n$ get

$$\begin{array}{ll} \hat{\zeta}_0 \hat{\zeta}_0' < 0, \quad |\hat{\zeta}_0'| \ge 2W & \text{for } |\hat{\zeta}_0| \ge \hat{\beta}_n(\lambda_n);\ \hat{\beta}_n = (4W)^{\frac{n+1}{n}} \lambda_n^{-\frac{n+1}{n}}, \\ |\hat{\zeta}_0'| \le W_* = 6W + 16\hat{R}_n^{\frac{1}{n+1}} \bar{\mu}_n W^{\frac{n+1}{n}} & \text{for } |\hat{\zeta}_0| \le \hat{\beta}_n(\lambda_n), \end{array} \quad \text{(D10)}$$

Moreover, $\hat{\beta}_n(\lambda_n)$ becomes arbitrarily small if λ_{n0} is further increased. Note the redundancy added to the lower bound of (D10) for the robustness. Obviously $|\hat{\zeta}_0| \le \hat{\beta}_n$ is fulfilled after a short transient at some moment $\hat{t}_*$. If $|\zeta_0(t_0)| \le \hat{\beta}_n$ it is kept further together with $|\hat{\zeta}_0'| \le W_*$.

Part 2. In order to prove that also $\hat{\zeta}_0'$ rapidly becomes small consider a time interval $\hat{t} \in [\hat{t}_*, \hat{t}_* + \delta]$, $L_* = L_0(t_*)$. It corresponds to the original-time interval $t \in [t_*, t_* + R^{\frac{1}{n+1}}\delta]$. Then $\varpi_i = L_*^{-1} R^{-\frac{n+1-i}{n+1}} (z_i - f_0^{(i)}) = \frac{L_0}{L_*} \hat{\zeta}_i$, $\varpi_i' = \frac{L_0}{L_*} \hat{\zeta}_i' + R^{\frac{1}{n+1}} \frac{\dot{L}_0}{L_0} \frac{L_0}{L_*} \hat{\zeta}_i$, $L_0 \in [e^{-\hat{\delta}}, e^{\hat{\delta}}] L_*$, $\hat{\delta} = \hat{R}_n^{\frac{1}{n+1}} \delta$. The diapason of the L_0 variance is deliberately increased here.

Once more introduce the virtual noise $\tilde{\eta}_L(t)$, $|\tilde{\eta}_L(t)| \le \tilde{\varepsilon}_L$. Recall that in the reality $\tilde{\eta}_L(t) \equiv \eta_L(t)$. Correspondingly obtain

$$\begin{array}{l} \varpi_0' \in -\lambda_n (\frac{L_0}{L_*}(1 + \tilde{\eta}_L))^{\frac{1}{n+1}} \left\lfloor \varpi_0 + [e^{-\hat{\delta}}, e^{\hat{\delta}}] \frac{\xi}{R} \right\rceil^{\frac{n}{n+1}} \\ \qquad\qquad - \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n R^{\frac{1}{n+1}} (\varpi_0 + [e^{-\hat{\delta}}, e^{\hat{\delta}}] \frac{\xi}{R}) + \varpi_1, \\ \varpi_1' \in -\lambda_{n-1} (\frac{L_0}{L_*}(1 + \eta_L))^{\frac{1}{n}} \lfloor \varpi_1 - \varpi_0' \rceil^{\frac{n-1}{n}} - \mu_{n-1} R^{\frac{1}{n+1}} (\varpi_1 - \varpi_0') + \varpi_2, \\ \quad \dots \\ \varpi_n' \in -\lambda_0 \frac{L_0}{L_*}(1 + \eta_L) \lfloor \varpi_n - \varpi_{n-1}' \rceil^0 - \mu_0 R^{\frac{1}{n+1}} (\varpi_n - \varpi_{n-1}') + e^{\hat{\delta}}[-1,1], \end{array} \quad \text{(D11)}$$

Let $\xi = 0, \tilde{\eta}_L = 0$. The first line of (D11) becomes an equation. Differentiating it obtain

$$\varpi_0'' = -\lambda_n \frac{n}{n+1}\left(\frac{L_0}{L_*}\right)^{\frac{1}{n+1}} |\varpi_0|^{-\frac{1}{n+1}}\varpi_0' - \frac{\lambda_n}{n+1}\left(\frac{L_0}{L_*}\right)^{\frac{1}{n+1}} R^{\frac{1}{n+1}} \frac{\dot{L}_0}{L_0}\lfloor\varpi_0\rceil^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}}\bar{\mu}_n R^{\frac{1}{n+1}}\varpi_0' + \varpi_1'. \tag{D12}$$

Therefore, due to $|\dot{L}_0| \le L_0$, $L_0/L_* \in [e^{-\hat{\delta}}, e^{\hat{\delta}}]$, $R \in [0, \hat{R}_n]$, get from (D12) that

$$\begin{aligned} \varpi_0'' \in & -\lambda_n \frac{n}{n+1}\left[e^{-\frac{\hat{\delta}}{(n+1)}}, e^{\frac{\hat{\delta}}{(n+1)}}\right] |\varpi_0|^{-\frac{1}{n+1}}\varpi_0' \\ & + \frac{\lambda_n}{n+1}\hat{R}_n^{\frac{1}{n+1}}\left[-e^{\hat{\delta}}, e^{\hat{\delta}}\right]|\varpi_0|^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}}\bar{\mu}_n R^{\frac{1}{n+1}}\varpi_0' + \varpi_1'. \end{aligned} \tag{D13}$$

Here $|\varpi_0| \le e^{\hat{\delta}}\hat{\beta}_n$, $|\varpi_0'| \le e^{\hat{\delta}}(W_* + \hat{R}_n^{\frac{1}{n+1}}\hat{\beta}_n)$ and $|\varpi_1| \le e^{\hat{\delta}}W$, which due to (D11) implies $|\varpi_1'| \le e^{\hat{\delta}}[\lambda_{n-1}(W_* + \hat{R}_n^{\frac{1}{n+1}}\hat{\beta}_n + W)^{\frac{n-1}{n}} + \mu_{n-1}\hat{R}_n^{\frac{1}{n+1}}(W_* + \hat{R}_n^{\frac{1}{n+1}}\hat{\beta}_n + W) + W]$.

Choose such sufficiently small $\hat{\delta}$ and $\hat{\beta}_n$ that $e^{\hat{\delta}}\hat{R}_n^{\frac{1}{n+1}}[e^{\hat{\delta}}\hat{\beta}_n]^{\frac{n}{n+1}} < ne^{-\frac{\hat{\delta}}{(n+1)}}$ and $e^{\hat{\delta}}\hat{\beta}_n < 1$ hold, and fix them. It may require additional increase of λ_n. Then for sufficiently large λ_n the value ϖ_0' during the time $\hat{t} \in [t_*, t_* + \hat{\delta}]$ converges into the segment $\hat{c}\hat{\beta}_n[-1, 1]$, where the constant $\hat{c} > 0$ is easily calculated from (D13). For sufficiently small $\hat{\beta}_n$, i.e. sufficiently large λ_n, obtain the contraction property.

Part 3. Applying now Lemma 3 for $K = 2$, assigning $\hat{Q}_n = W$ and using continuous dependence on $\varepsilon_L, \hat{q}_n$ finish the lemma proof. □

Appendix E. Proof of Lemma 3

Proof of Lemma 3. Consider the case $n = 0$. Then $\lambda_0, \mu_0 > 1$, whereas $R_0, \hat{R}_0 > 0$ are any numbers. Check that the Lemma is true for the inclusions

$$\begin{aligned} \dot{\zeta}_0 &\in -\lambda_0(1 + \varepsilon_L[-1, 1])\tfrac{1}{R}\,\mathrm{sign}(\zeta_0 + \tfrac{\xi}{R}) - \mu_0(\zeta_0 + \tfrac{\xi}{R}) + [-1, 1]\zeta_0 + [-1, 1]\tfrac{1}{R},\ R \ge R_0 \\ \hat{\zeta}_0' &\in -\lambda_0(1 + \varepsilon_L[-1, 1])\,\mathrm{sign}(\hat{\zeta}_0 + \tfrac{\xi}{R}) - R\mu_0(\hat{\zeta}_0 + \tfrac{\xi}{R}) + R[-1, 1]\hat{\zeta}_0 + [-1, 1],\ R \le \hat{R}_0, \end{aligned} \tag{E1}$$

which are obtained from (B5) and (B6), $|\zeta_0(t_0)|, |\hat{\zeta}_0(\hat{t}_0)| \le 1$.

Let $R \ge R_0$. First let $\xi = 0$, $\varepsilon_L = 0$. Then $|\dot{\zeta}_0| \le (\lambda_0 + 1)\frac{1}{R_0} + (\mu_0 + 1)|\zeta_0| + \mu_0\frac{\xi}{R_0}$ implies that following the time interval I_ω of the length ω

$$|\zeta_0(t_0 + \omega)| \le 1 + (\lambda_0 + 1)\tfrac{1}{R_0}\omega + (\mu_0 + 1)\int_{I_\omega} |\zeta_0| dt + \mu_0\tfrac{1}{R_0}K.$$

Assuming $|\zeta_0| \le W$, get $1 + ((\lambda_0 + 1)\frac{1}{R_0} + (\mu_0 + 1)W)\omega + \mu_0\frac{1}{R_0}K \le W$. Hence, any $W \ge \mu_0\frac{2}{R_0}K + 4$ can be taken, provided $(\mu_0 + 1)\omega \le 0.5$, and $(\lambda_0 + 1)\omega \le R_0$.

Furthermore, for any $W_\xi < (\mu_0 - 1)R_0$ keeping $\xi \le W_\xi$ implies that $|\zeta_0| \le W$ is kept afterwards, since $\lambda_0, \mu_0 > 1$.

Similarly the case of $R \le \hat{R}_0$ is checked. Continuous dependence on ε_L finishes the proof for $n = 0$.

Induction step. Let now Lemmas 2, 3 and Theorem 1 be true for $0, ..., n-1$. Note that Lemma 2 has been already proved for the order n and can be used as well. Let λ_n be chosen sufficiently large providing for the statement of Lemma 2.

The case of approximate linearity, $R \geq R_n$**.** System (B5) implies

$$\begin{array}{l} \dot{\zeta}_0 = -\lambda_n R^{-\frac{1}{n+1}}(1+\eta_L)^{\frac{1}{n+1}} \left\lfloor \zeta_0 + \frac{\xi}{R} \right\rceil^{\frac{n}{n+1}} - \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n(\zeta_0 + \frac{\xi}{R}) - \frac{\dot{L}_0}{L_0}\zeta_0 + \zeta_1; \\ |\dot{\zeta}_0| \leq \lambda_n R^{-\frac{1}{n+1}}(1+\varepsilon_L) \left[|\zeta_0|^{\frac{n}{n+1}} + R^{-\frac{n}{n+1}}|\xi|^{\frac{n}{n+1}} \right] + (\lambda_n \bar{\mu}_n + 1)|\zeta_0| + \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n \frac{1}{R}|\xi| + |\zeta_1|. \end{array} \tag{E2}$$

Similarly to the proof of Lemma 2 the following proof utilizes Remark 4. Let event A be the statement that $|\zeta_0| \leq W_0$ holds over the time interval I_ω for any $0 \leq \omega \leq \omega_0$. Suppose that the inequalities $||\overleftarrow{\zeta}_1||_\infty \leq W_1$, $|\zeta_0| \leq W_0$ hold during the time interval I_{ω_0} (event B). Parameter p consists of ω_0, W_0, W_1. Let $W_0, W_1 > 1$.

Note that at the initial moment $||\zeta||_\infty \leq 1$, in particular $|\zeta_0| \leq 1$. Thus due to the Hölder inequality integrating obtain that

$$\begin{array}{l} \int_{I_\omega} 1 \cdot |\xi|^{\frac{n}{n+1}} dt \leq (\int_{I_\omega} dt)^{\frac{1}{n+1}} (\int_{I_\omega} |\xi| dt)^{\frac{n}{n+1}} \leq \omega^{\frac{1}{n+1}} (\int |\xi(t)| dt)^{\frac{n}{n+1}}; \\ |\zeta_0| \leq 1 + \int_{I_\omega} \dot{\zeta}_0 dt \leq 1 + \lambda_n(1+\varepsilon_L) R_n^{-\frac{1}{n+1}} \left[W_0\omega + R_n^{-\frac{n}{n+1}} \omega^{\frac{1}{n+1}} \left(\int |\xi(t)| dt \right)^{\frac{n}{n+1}} \right] \\ \qquad + (\lambda_n^{\frac{n+1}{n}} \bar{\mu}_n + 1) W_0 \omega + \lambda_n^{\frac{n+1}{n}} \bar{\mu}_n R_n^{-1} K + \omega W_1. \end{array}$$

Correspondingly, for sufficiently small ω_0, $\omega \leq \omega_0$, get $|\zeta_0| \leq 2 + 2\lambda_n^{\frac{n+1}{n}} \bar{\mu}_n R_n^{-1} K \leq W_0$. Writing down that condition and further increasing W_0 obtain that

$$\begin{array}{l} \zeta_0 \dot{\zeta}_0 < 0, \ |\dot{\zeta}_0| \geq 1 \text{ for } |\zeta_0| \geq W_0, \ |\zeta_1| \leq W_1, \ \xi \equiv 0, \\ 2 + 2\lambda_n^{\frac{n+1}{n}} \bar{\mu}_n R_n^{-1} K \leq W_0 \end{array} \tag{E3}$$

hold due to (E2) for $\xi = 0$ since $\lambda_n^{\frac{n+1}{n}} \bar{\mu}_n > 1$. For any fixed W_1 conditions (E3) hold for any sufficiently large W_0.

Find now the appropriate value of W_1. The input to the $\overleftarrow{\zeta}_1$-subsystem of (B5) is $\dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0$. Since $|\zeta_0| \leq W_0$ and due to (E2) the inequality $\int_{I_\omega} |\dot{\zeta}_0 + \frac{\dot{L}_0}{L_0}\zeta_0| dt \leq 2\lambda_n^{(n+1)/n} \bar{\mu}_n R_n^{-1} K + 2$ holds for ω small enough. Thus, applying Lemma 3 (the induction assumption), get that also $\overleftarrow{\zeta}_1$ remains bounded by some $\tilde{W}_1 > 0$ over I_ω. Note that W_1 does not depend on W_0. Fix $W_1 = 2\tilde{W}_1$, and choose any W_0 satisfying (E3) for sufficiently small ω_0, i.e. for any $\omega \leq \omega_0$. Fix some sufficiently small value $\omega_0 > 0$ satisfying all the restrictions imposed above.

Finishing Remark-4 logic. The choice of the parameter $p = (\omega_0, W_0, W_1)$ implies B and therefore A due to Remark 4. Indeed, let the event B_ω, $0 \leq \omega \leq \omega_0$, mean that inequalities $||\overleftarrow{\zeta}_1||_\infty \leq W_1$, $|\zeta_0| \leq W_0$ hold during the time interval $[t_0, t_0 + \omega]$. Then $B = (\forall \omega \in [0, \omega_0] : \ B_\omega)$.

At the initial moment t_0 event B_ω is true for $\omega = 0$. Suppose that B_ω is violated for some $\omega_* \leq \omega_0$. Take the corresponding solution of system (B5). Then along this solution the condition B_{t-t_0} is for the first time violated at some time $t_1 = t_0 + \omega_1$, $t_1 = \sup\{t \geq t_0 \,|\forall \tau \in [t_0, t] : \ B_{\tau - t_0}\}$, i.e. $||\overleftarrow{\zeta}_1(t_1)||_\infty = W_1$ or $|\zeta_0(t_1)| = W_0$, and the condition B_{ω_1} is true. It means that all the above calculations hold and none of these two equalities is feasible for the chosen ω_0, W_0, W_1, i.e. we come to the contradiction.

Thus, $|\zeta_0| \leq W_0$, $||\overleftarrow{\zeta}_1||_\infty \leq W_1$ hold over the time interval I_ω for $\omega \leq \omega_0$. Take any $W \geq \max[W_0, W_1]$, then $||\zeta||_\infty \leq W$ is kept over I_ω. Note that W can always be increased.

It remains to prove that for sufficiently small W_ξ keeping $|\xi| \leq RW_\xi$ provides for indefinitely keeping $|\zeta| \leq W$ also after the time interval I_ω. The transient dynamics during the time period I_ω results in a shift to some point $\vec{\sigma}_n(t_0 + \omega)$, $||\vec{\sigma}_n(t_0 + \omega)||_\infty \leq WR$. Apply Lemma 2 which has been already proved above for the order n using Lemma 3 for the lower subsystem only. Hence, $\vec{\sigma}_n$ converges into $||\vec{\sigma}_n||_\infty \le WR/2$ during the time ΔT_n without leaving $||\vec{\sigma}_n||_\infty \le Q_n WR$ on the way, provided starting from the time $t_0+\omega$ the inequality $|\xi| \le W_\xi R$ is kept for $W_\xi = q_n W$. Redefine $W = \tilde{W}$ taking $\tilde{W} \ge \max[W, Q_n W]$. It finishes the case $R \ge R_n$.

The case of the negative approximate homogeneity degree, $R \le \hat{R}_n$**.** The proof logic is the same, and only the main points are presented. Dynamics (B6) imply

$$\hat{\zeta}_0' = -\lambda_n(1+\eta_L)^{\frac{1}{n+1}}\left\lfloor \hat{\zeta}_0 + \tfrac{\xi}{R}\right\rceil^{\frac{n}{n+1}} - R^{\frac{1}{n+1}}\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n(\hat{\zeta}_0 + \tfrac{\xi}{R}) - R^{\frac{1}{n+1}}\tfrac{\dot{L}_0}{L_0}\hat{\zeta}_0 + \hat{\zeta}_1;$$
$$|\hat{\zeta}_0'| \le \lambda_n(1+\varepsilon_L)\left[|\hat{\zeta}_0|^{\frac{n}{n+1}} + |\tfrac{\xi}{R}|^{\frac{n}{n+1}}\right] + R^{\frac{1}{n+1}}(\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n + 1)|\hat{\zeta}_0| + R^{\frac{1}{n+1}}\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n|\tfrac{\xi}{R}| + |\hat{\zeta}_1|.$$

Suppose that the inequalities $||\overleftarrow{\hat{\zeta}}_1||_{h\infty} \le W_1$, $|\hat{\zeta}_0| \le W_0$ hold during the time interval I_ω. Also note that $|\hat{\zeta}_0| \le 1$, $||\overleftarrow{\hat{\zeta}}_1||_{h\infty} \le 1$ at the initial moment t_0.

Note that the integration over the time interval I_ω corresponds to the $\hat{t}$-interval of the length ω. Hence, integrating, obtain according to the Hölder inequality that

$$|\hat{\zeta}_0| \le 1 + \lambda_n(1+\varepsilon_L)\left[W_0\omega + \omega^{\frac{1}{n+1}}K^{\frac{n}{n+1}}\right] + \hat{R}_n^{\frac{1}{n+1}}(\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n + 1)W_0\omega + \lambda_n^{\frac{n+1}{n}}\bar{\mu}_n\hat{R}_n^{\frac{1}{n+1}}K + \omega W_1$$

holds over the interval. Thus for sufficiently small ω get $|\hat{\zeta}_0| \le 2 + 2\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n\hat{R}_n^{\frac{1}{n+1}}K \le W_0$. That estimation of W_0 does not depend on W_1. Additionally increase W_0 to ensure $\hat{\zeta}_0\dot{\hat{\zeta}}_0 < 0$ for $|\hat{\zeta}_0| \ge W_0$, $|\hat{\zeta}_1| \le W_1$ and $\xi = 0$.

The input to the $\overleftarrow{\zeta}_1$-subsystem of (B6) is $\hat{\zeta}_0' + R^{\frac{1}{n+1}}\frac{\dot{L}_0}{L_0}\hat{\zeta}_0$. Simple calculations show that the inequality $\int_{I_\omega}|\hat{\zeta}_0' + R^{\frac{1}{n+1}}\frac{\dot{L}_0}{L_0}\hat{\zeta}_0|dt \le 2\lambda_n^{\frac{n+1}{n}}\bar{\mu}_n\hat{R}_n^{\frac{1}{n+1}}K + 2$ holds for small ω. Thus, due to the induction assumption, also $\overleftarrow{\hat{\zeta}}_1$ remains bounded by some $\tilde{W}_1$ over I_ω. Fix the corresponding value of $W_1 = 2\tilde{W}_1$. The previous conditions allow to choose W_0. Define $W = \max[W_0, W_1]$.

Similarly to the previous case also here Lemma 2 finishes the proof. □