2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Generic Interval Filter for continuous time nonlinear systems with bounded state condition

Quoc Hung Lu[1*], Soheib Fergani[1] and Carine Jauberthie[1]

***Abstract*— State estimation for continuous time nonlinear systems, in bounded error framework, is concerned in this paper. Guaranteed state estimation is obtained without assumption of differentiability on the dynamical function. The proposed filter is mainly based on interval computations, bounded state assumption and with or without Lipschitz smoothness condition on the dynamical function. This filter is applied on a case study, highlighting the potential of the proposed filtering method.**

I. Introduction

Continuous time nonlinear systems of ordinary differential equations (ODEs) represent a fundamental area of inquiry within both academic and industrial spheres, owing to their capacity to model a wide range of natural phenomena and engineering systems. However, the resolution of such systems poses a formidable challenge, prompting the development of state approximation methodologies. These encompass punctual and set-valued estimation approaches, with the former tracing its origins to the earliest investigations of ODEs, while the latter has garnered increasing attention in recent years.

Historically, the advancement of punctual estimation methodologies has occupied a central position since the nascent stages of exploration into ODEs. These methodologies, encompassing numerical integrators and deterministic approaches, aim to furnish discrete-time approximations of system states. Despite their efficacy, they may encounter challenges in capturing the nuanced behaviors inherent to nonlinear systems, particularly in contexts marked by uncertainties. Conversely, set-valued estimation methodologies have garnered prominence more recently, signaling a paradigmatic shift in system analysis and control [1].

In the fields of control theory and signal processing, filtering signals in the presence of uncertainties and nonlinearities is a significant challenge. Most of the real-world systems exhibit nonlinear dynamics with states bounded within certain intervals of uncertainties, and these systems often lack the smooth differentiability conditions required by traditional filtering approaches [2]. Recent advances in the field have sparked a paradigmatic shift towards addressing the constraints of conventional filters, leading to an upsurge in alternative methodologies. Traditional filtering techniques, such as the Kalman filter [3], rely heavily on assumptions of linearity and differentiability, which are frequently violated in practical scenarios. To overcome the heavy assumptions of linearity in the case of punctual estimation, several works around extended Kalman filter have been proposed [4], [5], [6]. On the other hand, for the nonlinear set-valued estimation, interesting approaches have been developed based on Particle Filters principle [7], [8] where efficiency and accuracy depend mostly on the number of particles used in the estimation which may require a high computation time. Other set-membership filtering techniques aiming to obtain the best guaranteed result/conservatism tradeoff have also been proposed in [9], [10]. In addition to the known advantages and drawbacks of these methods, all of them necessitate adherence to the criterion of differentiability.

Our research focuses on interval filtering as a means to overcome these limitations. Interval filtering accommodates nonlinear systems with state variables confined within predefined intervals, relaxing the necessity for strict adherence to differentiability criteria.

In this article, a generic filter for continuous time nonlinear systems with bounded state while circumventing the necessity for differentiability conditions is presented. It is organised as follows. In Section II, some important definitions, notations and basic concepts needed for the estimation strategy developments are provided. Then, Section III presents the main result of the paper which is the guaranteed state estimation obtained without assumption of differentiability on the dynamical function. Section IV shows the simulations results on two nonlinear systems: the first one is a nonlinear suspension model identified on a real vehicle and the second one is a nonlinear chemical reactor. These simulation results highlight the efficiency of the proposed estimation strategy. Finally, conclusions and future works are described in Section V.

II. Notations and basic concepts

A *real interval (matrix)* of dimension $p \times q$ is defined as $[X] \equiv [\underline{X}, \overline{X}] \triangleq \left\{X \in \mathbb{R}^{p\times q} : \underline{X} \leq X \leq \overline{X}\right\}$ where $\sup([X]) \equiv \overline{X}$, $\inf([X]) \equiv \underline{X}$, $\text{mid}([X]) = (\overline{X} + \underline{X})/2$, $\text{rad}([X]) = (\overline{X} - \underline{X})/2$, $\text{width}([X]) = \overline{X} - \underline{X}$ are called respectively the *largest*, the *smallest*, the *midpoint*, the *radius* and the *width element (or matrix)* of $[X]$. Denote also $[X] = \text{mid}([X]) \pm \text{rad}([X])$ and write $X \in [X]$ to indicate a *punctual element* X *belonging element-wise* to $[X]$. Basic interval operators $\diamond \in \{+, -, \times, \div\}$ defined in [11] are used in computation (no approximation algorithm is necessary) and more general operators are constructed by means of inclusion function $[f]$ defined in Definition 1 below. In practice, the Intlab package [12] developed for Matlab (also existing in Octave and C/C++) is used for computations.

[1] Authors are with LAAS-CNRS, Université de Toulouse, CNRS, UPS, 7 avenue du Colonel Roche, 31400 Toulouse, France. Emails: qhlu@laas.fr, sfergani@laas.fr, cjaubert@laas.fr.
[*] Corresponding author.

Definition 1: An *inclusion function* of $f : D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is a function $[f]$ that maps any interval $[x] \subset D$ to an interval $[f]([x]) \subset \mathbb{R}^n$ which contains the image set $f([x])$. The *minimal inclusion function* of f is a function $[f]^*$ such that (s.t.) for all $[x] \subset D$, $[f]^*([x])$ is the smallest interval containing $f([x])$.

A sequence of elements is noted as $w_1, ..., w_k$ or $w_1 : w_k$ or $w_{1:k}$ where the notation $p : q$ ($p \leq q$) is used for a *range* from p to q with step of 1. The function $\mathbb{I}(.)$ is defined as $\mathbb{I}(x) = 1$ if the (vector of) conditions x are true and $\mathbb{I}(x) = 0$ otherwise. Denote $\{x_t(i)\}_{i=1:n}$ for the components of an *indexed* vector $x_t \in \mathbb{R}^n$, $t \in \mathcal{I} \subset \mathbb{R}$ and $\{x_i\}_{i=1:n}$ for those of a *non indexed* vector $x \in \mathbb{R}^n$. By abuse of notation, following operators are *componentwise* : the absolute value operator, the operators $\star \in \{+, -, \times, \div, \leq, \geq, <, >, \vee, \wedge\}$ where $x \vee \delta = \max(x, \delta)$ and $x \wedge \delta = \min(x, \delta)$.

III. Main results

A. Continuous time nonlinear system

Consider the following general continuous system

$$\begin{cases} \dot{x}(t) &= f(x(t), u(t)), \\ x(0) &= x_0, \end{cases} \qquad t \geq 0 \tag{1}$$

where $x(t) \in \mathbb{R}^{n_x}$ is the system states, $u(t) \in \mathbb{R}^{n_u}$ inputs and $f : D = D_x \times D_u \to \mathbb{R}^{n_x}$ a vector-valued function s.t.

$$f(x, u) = \Big(f_1(x, u), ..., f_{n_x}(x, u)\Big)^T, x \in D_x, u \in D_u, \tag{2}$$

where $D_x \subseteq \mathbb{R}^{n_x}$, $D_u \subseteq \mathbb{R}^{n_u}$ and $\{f_i\}_{i=1:n_x}$ are real functions. Note that $\mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \equiv \mathbb{R}^{n_x+n_u}$ and x_t is used interchangeably with $x(t)$ to simplify the notation.

From the ODE's theory, a solution $x(t)$ of (1) is a *continuously differentiable function* of t on some interval $I \subseteq [0, +\infty)$ and s.t. $x(0) = x_0$. So, $x(t)$ (which is differentiable) and $\dot{x}(t)$ are continuous functions of $t \in I$ and thus, by writing

$$f(x(t), u(t)) = f \circ \vartheta(t), \quad \vartheta(t) = (x(t), u(t)), \tag{3}$$

one deduces that f, ϑ and u are necessarily continuous functions to ensure the continuity of $\dot{x}$. Furthermore, since $x(t)$ is continuous, then $\forall \tau > 0$, $\forall t \in I_\tau = [0, \tau]$:

$$\underline{x}_\tau \triangleq \inf_{t \in I_\tau} \{x(t)\} \leq x(t) \leq \sup_{t \in I_\tau} \{x(t)\} \triangleq \overline{x}_\tau \ , \tag{4}$$

or equivalently, $x(t) \in [x_\tau] = [\underline{x}_\tau, \overline{x}_\tau]$.

In the literature, a majority of researches assume the smoothness of f by its differentiability. In the present text, we consider a class of functions f possessing a weakened smoothness property but restricted by some conditions in a practical view of filtering / state estimation and controlling purposes. In that view point, there is no necessity to handle with a system in which its state magnitudes $\|x(t)\|$ tends to ∞ (as $t \to \infty$), since in that case, after a time $\tau > 0$, the states are no longer in the scope of estimation or control. Three overt ways causing this worst-case are as follows:

(W1) $\|\dot{x}(t)\| = \|f(x(t), u(t))\| \to \infty$ as $t \to \infty$.
(W2) $\exists \tau > 0$ s.t. $\forall t \geq \tau$: $\|\dot{x}(t)\| \in [\alpha]$ with $\underline{\alpha} > 0$.
(W3) $c > 0$ s.t. $\lim_{t \to \infty} \|\dot{x}(t)\| = c$.

Some researches do not consider these concerns in their developments and some other consider the problem for only a finite time horizon $[0, \tau]$. The last situation turns the problem into bounded state condition: $x(t) \in [x_\tau]$, $\forall t \in [0, \tau]$, as stated in (4).

Furthermore, practical systems, in particular engineering ones, are generally constrained by many technical limits (material, operation energy, control working conditions,...) so that their considered states are finite (e.g. rotor speed, yaw angle, suspension position,...). Thus, in the following development, we devote to a *bounded state* context as a constraint to the system (1) in which the function f is not necessarily differentiable and differs from (W1)-(W3) aforementioned. This means that for any solution $x(t)$ of (1), along the *curve* $\{\vartheta(t) = (x(t), u(t)) \in D_x \times D_u \mid t \in I\}$:

(F1) f is bounded (constant bounds containing 0),
(F2) $f = 0$ infinitely many time (or $f \to 0$) as $t \to \infty$,
(F3) f does not converge to any constant $c \neq 0$.

Besides, any valid input $u(t)$ should be finite and lies necessarily in some range. Literally, we assume the following:

(H1) There are finite intervals $[x]$ and $[u]$ so that

$$x(t) \in [x] \quad \text{and} \quad u(t) \in [u] \ , \ \forall t \in I. \tag{5}$$

Note that, in the case of finite time horizon, $[x] \equiv [x_\tau]$ for some fixed $\tau > 0$, and similarly for $[u]$.

So, any continuous function f defined on the whole space $\mathbb{R}^{n_x+n_u}$ and well describing the relation (1) can now be considered as being restricted to the domain $D = [x] \times [u]$, in the sense that f vanishes outside of D. Then the range of f is contained in an interval $[f]$ which can be obtained by means of an inclusion function. In other words, by assuming (H1), condition (F1) naturally holds for any continuous function f while (F2)-(F3) can be ignored since the system states have already assumed to be bounded.

This extension allows us to handle with various class of functions including linear functions, Lipschitz functions, L^p ($p \in [1, \infty)$) continuous functions,... defined on $\mathbb{R}^{n_x+n_u}$, provided that (H1) is satisfied. In particular, the class of Lipschitz function is the first one worthily to be considered after those of linear and continuously differentiable functions for two reasons: (a) the former effectively contains the others (with bounded domains), (b) the former property concerns the uniqueness of the solution of system (1) and provides nicely bounds for it.

The system (1) can be equivalently written as:

$$x(t) \quad = \quad x_0 \ + \ \int_0^t f(x(s), u(s)) ds \quad , \quad \forall t \in I. \tag{6}$$

Now, suppose further that:

(H2) x_0 belongs to some known interval $[x_0] \subset [x]$.

Then, the following proposition provides the largest guaranteed results that include the corresponding real states at any time instance $t \in I$. The issue that we have to resolve in the next section is how to obtain better reductions of these results at desired discrete times.

Proposition 1: Consider system (1) with assumptions (H1)-(H2). Then at any time $t \in I$:

$$\underline{x}_t \triangleq \underline{x} \vee \left(\underline{x}_0 + t.\underline{f}\right) \leq x(t) \leq \overline{x} \wedge \left(\overline{x}_0 + t.\overline{f}\right) \triangleq \overline{x}_t, \quad (7)$$

and denote $[x_t] = [\underline{x}_t, \overline{x}_t] = [x] \cap ([x_0] + t.[f])$.
The proposition is obtained by direct computation from (6) using (H1)-(H2), its proof is thus deliberately omitted here.

B. Discretized system

Let $0 = t_0 < t_1 < ...$ s.t. $T = t_k - t_{k-1}$, $k \in \mathbb{N}^+$, is small enough and assuming that $u(t)$ is constant in each interval $[t_{k-1}, t_k)$, i.e. $u(t) \equiv u_k$, $\forall t \in [t_{k-1}, t_k)$. For simplicity, let $x(t_k) = x_k$, $\forall k \in \mathbb{N}^*$, $\alpha_f = \overline{f} - \underline{f}$ and $\beta_f = |\overline{f}| \vee |\underline{f}|$. The system (1) can be discretized at $t_1, t_2, ...$ and rewritten as:

$$\begin{aligned} x_k &= x_0 + \int_0^{kT} f(x(s), u(s))ds \,, & (8)\\ &= x_{k-1} + \int_{t_{k-1}}^{t_k} f(x(s), u_k)ds \,,\ k \in \mathbb{N}^*. & (9)\end{aligned}$$

Then, consider that

$$(8) \iff x_k = x_{k-1} + T.f(x_{k-1}, u_k) + \epsilon_k \,,\ k \in \mathbb{N}^*, \quad (10)$$

with *discretization errors* ϵ_k's determined by

$$\epsilon_k = \int_{t_{k-1}}^{t_k} [f(x(s), u_k) - f(x_{k-1}, u_k)]\, ds \,. \quad (11)$$

Applying Proposition 1, one obtains

$$x_k \in [x_k] = [x(kT)] = [x] \cap ([x_0] + k.T.[f]) \,,$$

which is used as a consistency condition later on. Now, in order to refine this guaranteed interval, we propose to use further assumption as follows.

(H3) Assume that the function f is either:
componentwise Lipschitz with coefficients $L_i > 0$ and $L = (L_1, ..., L_{n_x})^T$:

$$|f(x, u(t)) - f(\tilde{x}, u(t))| \leq L.\|x - \tilde{x}\| \,, \quad (12)$$

or *totally Lipschitz* with a coefficient $L > 0$:

$$\|f(x, u(t)) - f(\tilde{x}, u(t))\| \leq L.\|x - \tilde{x}\|. \quad (13)$$

Remark 1: Noting that $|z| \leq \|z\|$, $\forall z \in \mathbb{R}^n$, so if (13) holds then (12) also holds in which L is a scalar. Reversely, if (12) holds then $\|f(x, u(t)) - f(\tilde{x}, u(t))\| \leq \|L\|.\|x - \tilde{x}\|$.

Proposition 2: Consider system (10) with assumptions (H1)-(H2), then $\forall k \in \mathbb{N}^*$:

(2.1) $|\epsilon_k| \leq T \cdot \alpha_f$ and $\|\epsilon_k\| \leq T \cdot \|\alpha_f\|$.

(2.2) Assume further that f satisfies (H3). Then the following inequalities hold true

$$\begin{aligned} |\epsilon_k| &\leq T \cdot (T \cdot \|\beta_f\| \cdot L \wedge \alpha_f) \,, & (14)\\ \|\epsilon_k\| &\leq T \cdot (T \cdot \|\beta_f\| \cdot \|L\| \wedge \|\alpha_f\|) \,. & (15)\end{aligned}$$

Remark 2: If (H3) is not verified or it is hard to find the Lipschitz coefficient L, one can consider that $L \equiv \infty$ and (14)-(15) remain valid, whence (2.1) $\equiv$ (2.2).

Proof: Let $k \in \mathbb{N}^*$ and $i = 1 : n_x$, we have:

$$\begin{aligned} |\epsilon_k(i)|^2 &= \left| \int_{t_{k-1}}^{t_k} \Big(f_i(x(s), u_k) - f_i(x_{k-1}, u_k) \Big) ds \right|^2 \\ &\leq \int_{t_{k-1}}^{t_k} 1^2 ds \int_{t_{k-1}}^{t_k} |f_i(x(s), u_k) - f_i(x_{k-1}, u_k)|^2 \, ds \\ &= T \cdot \int_{t_{k-1}}^{t_k} |f_i(x(s), u_k) - f_i(x(t_{k-1}), u_k)|^2 ds \quad (*)\end{aligned}$$

where the Holder's inequality is used to obtain the upper bound in the above evaluation. Since $f([x], [u]) \subset [f]$, then $|f_i(x(t), u_k) - f_i(x(t_{k-1}), u_k)| \leq \overline{f}_i - \underline{f}_i \equiv \text{width}([f_i])$, and also $|f_i(x(t), u_k) - f_i(x(t_{k-1}), u_k)| \leq 2 \left(|\overline{f}_i| \vee |\underline{f}_i| \right)$.

Since $\text{rad}([f_i]) \leq |\overline{f}_i| \vee |\underline{f}_i|$ then $|\epsilon_k|^2 \leq T^2 \cdot \left(\overline{f} - \underline{f}\right)^2$ and (2.1) is proved.

From (H3) and $(*)$, one obtains:

$$\begin{aligned} |\epsilon_k|^2 &\leq T \cdot \int_{t_{k-1}}^{t_k} L^2.\|x(s) - x(t_{k-1})\|^2 ds \,, \\ &= TL^2 \textstyle\sum_{i=1}^{n_x} \int_{t_{k-1}}^{t_k} \left| \int_{t_{k-1}}^{s} \dot{x}_\tau(i) d\tau \right|^2 ds \\ &\leq TL^2 \textstyle\sum_{i=1}^{n_x} \int_{t_{k-1}}^{t_k} \left(\int_{t_{k-1}}^{s} 1^2 d\tau \right) \left(\int_{t_{k-1}}^{s} |\dot{x}_\tau(i)|^2 \, d\tau \right) ds \\ &\leq T^3 L^2 \textstyle\sum_{i=1}^{n_x} \int_{t_{k-1}}^{t_k} |f_i(x(\tau), u(\tau))|^2 \, d\tau \\ &\leq T^4 L^2 \left\| \left(|\overline{f}| \vee |\underline{f}| \right) \right\|^2 \,,\end{aligned}$$

where the Holder's inequality is used in the third inequality. So (2.2) is implied using the obtained result and (2.1). ■

Corollary 1: Consider Proposition 2 and let $\delta > 0$. Then $|\epsilon_k| \leq \|\epsilon_k\| \leq \delta$, $\forall k \in \mathbb{N}^*$, if

$$T \leq \overline{T} \triangleq \left(\frac{\delta}{\|\alpha_f\|} \vee \sqrt{\frac{\delta}{\|\beta_f\|.\|L\|}} \right) .$$

Proof: Consider the right hand side of (15) and denote it by $RHS(15)$. If

$$T^2.\|\beta_f\|.\|L\| \leq \delta \quad \text{or} \quad T.\|\alpha_f\| \leq \delta \,,$$

then $\|\epsilon_k\| \leq RHS(15) \leq \delta$. So, the corollary is implied. ■

Let's use following notations in the sequel: For all $k \in \mathbb{N}^*$,

- $\hat{x}_k \triangleq \hat{x}_{k-1} + T.f(\hat{x}_{k-1}, u_k)$ be the *state estimate*,
- $e_k \triangleq x_k - \hat{x}_k$ be the *estimation error*.

Proposition 3: Consider system (10) with assumptions (H1)-(H3). Let $\delta > 0$, $T \leq \overline{T}$ according to Corollary 1. Let $[x_0] = m_0 \pm \rho_0$ with $m_0 = \text{mid}([x_0])$, $\rho_0 = \text{rad}([x_0])$ and the initial estimate be chosen as $\hat{x}_0 = m_0$. For all $k \in \mathbb{N}^*$, let $[\hat{x}_k] = \hat{x}_k \pm \rho_k$ be an *interval estimate* of x_k, where

$$\begin{aligned} \hat{x}_k &= \hat{x}_{k-1} + T.f(\hat{x}_{k-1}, u_k) \,, \\ \rho_k &= \rho_{k-1} + T.\left(\frac{\alpha_f}{\sigma} \wedge L \right).\|\rho_{k-1}\| + \delta \,, \quad (16)\end{aligned}$$

with $\sigma = \|\rho_0\| + \delta \cdot \mathbb{I}(\|\rho_0\| = 0)$. Then :

$$\begin{aligned} \rho_{k-1} &\leq \rho_k \quad \text{and} \quad \|\rho_{k-1}\| \leq \|\rho_k\| \,, & (17)\\ x_k &\in [\hat{x}_k] \,, & (18)\\ |e_k| &\leq \rho_k \,. & (19)\end{aligned}$$

Proof: Firstly, (17) is directly deduced from (16) and (18) holds if (19) is satisfied. So it remains to prove (19).

By assumptions, one gets

$$\begin{aligned} &x_0 \in [\hat{x}_0], \quad |e_0| = |x_0 - \hat{x}_0| = |x_0 - m_0| \leq \rho_0, \\ &e_1 = x_1 - \hat{x}_1 \,, \\ &\quad = x_0 - \hat{x}_0 + T.\left(f(x_0, u_1) - f(\hat{x}_0, u_1)\right) + \epsilon_1.\end{aligned}$$

Thus

$$\begin{aligned} |e_1| &\leq |x_0 - \hat{x}_0| + T.|f(x_0,u_1) - f(\hat{x}_0,u_1)| + |\epsilon_1| \\ &\leq |e_0| + T.(\alpha_f \wedge L.\|e_0\|) + \delta \\ &\leq \rho_0 + T.\left(\frac{\alpha_f}{\|\rho_0\|} \wedge L\right).\|\rho_0\| + \delta \\ &= \rho_0 + T.\left(\frac{\alpha_f}{\sigma} \wedge L\right).\|\rho_0\| + \delta \quad \equiv \quad \rho_1 \ , \end{aligned}$$

where the last equality is verified for either $\rho_0 = 0$ or not.

Assuming now that (19) is satisfied for a $k > 1$, one proves that it also holds for $k+1$. Indeed, with the similar strategy,

$$\begin{aligned} |e_{k+1}| &\leq \rho_k + T.\left(\frac{\alpha_f}{\|\rho_k\|} \wedge L\right).\|\rho_k\| + \delta \\ &\leq \rho_k + T.\left(\frac{\alpha_f}{\sigma} \wedge L\right).\|\rho_k\| + \delta \quad \equiv \quad \rho_{k+1} \ , \end{aligned}$$

regarding $\frac{1}{\|\rho_k\|} \leq \frac{1}{\|\rho_1\|} \leq \frac{1}{\sigma}$, $\forall k \geq 2$, thanks to (17). ■

C. Filter algorithms

1) Generic Guarantor algorithm: In the following, we propose a Generic Guarantor (GG) algorithm by applying Proposition 3. As this algorithm is generic, any refined consistency step could be appreciated and used in addition, e.g. using measurements with contemporary contractors.

Algorithm 1 GENERIC GUARANTOR (GG)

1: **Initialization:** L, f, $[x]$, $[u]$, $[x_0]$, δ, $\{u_k\}_{k=1:N}$.
   % Compute:
2: $[\underline{f}, \overline{f}] = [f]([x],[u])$;
3: Choose a T: $0 < T \leq \overline{T}$ (defined in Corollary 1).
4: $\hat{x}_0 = \text{mid}([x_0])$; $\rho_0 = \text{rad}([x_0])$; $[\hat{x}_0] = [x_0]$;
5: $\sigma = \|\rho_0\| + \delta \cdot \mathbb{I}(\|\rho_0\| = 0)$;
6: $T_L = T \cdot \left[(\overline{f} - \underline{f})/\sigma \wedge L\right]$;
7: **for** $k = 1, 2, 3, ...N$ **do**
   % Propagation step:
8: $\hat{x}_k = \hat{x}_{k-1} + T.f(\hat{x}_{k-1}, u_k)$;
9: $\rho_k = \rho_{k-1} + T_L.\|\rho_{k-1}\| + \delta$;
10: $[\hat{x}_k] = \hat{x}_k \pm \rho_k$;
   % Consistency step:
11: $[\hat{x}_k] \leftarrow [\hat{x}_k] \cap ([x_0] + k.T.[f]) \cap [x]$;
12: $\rho_k \leftarrow \text{rad}([\hat{x}_k])$;
13: **end for**

Remark 3: The quality of GG algorithm is affected by the Lipschitz coefficient L and the bounded range $[\underline{f}, \overline{f}]$. The smaller their sizes, the better the algorithm's performance.

2) Robust Generic Guarantor algorithm: The Robust Generic Guarantor (RGG) algorithm is proposed in order to make the GG algorithm more efficient and robust in consideration of: a) penalizing the wrapping effect of interval computations; b) tolerating the errors which might be caused by using the intersection operator inside of contractor methods.

The algorithm is based on the following two principles:

(**P1**) Penalize the ρ_k's increase by a factor γ_k. The choice of γ_k depends on f. It should be a function of k decreasing to 0 at a rate depending on the application, e.g.:

$$\gamma_k = n \cdot k^{-1} \quad \text{or} \quad \gamma_k = (\log k^n)^{-1} \quad \text{or else,}$$

with a fixed chosen number $n > 0$. Using this factor γ_k, it is still guaranteed that $\rho_k \leq \rho_{k+1}$, $\forall k \in \mathbb{N}^*$, and furthermore

$$\rho_k \leq \rho_{k+1} \leq \rho = \lim_{t \to \infty} \rho_k < \infty$$

where ρ is proportional to width($[x]$). This is coherent with the fact that: $[\hat{x}_k] \subseteq [x] \ \Rightarrow \ \rho_k \leq \text{width}([x])$, $\forall k \in \mathbb{N}^*$.

(**P2**) Apply an alternative *regularization method* vis-a-vis existing contractors to regularize the estimate consistency using linear measurements $[y_k] = \mu_k \pm \sigma_k$, where $\mu_k = \text{mid}([y_k])$, $\sigma_k = \text{rad}([y_k])$ is the measure precision and assuming that the real measurement $y_k = C.x_k \in [y_k]$. Assume further that $[y_k] \subseteq C.[\hat{x}_k]$ where $[\hat{x}_k]$ is the estimate interval provided by the propagation step of the GG algorithm. The proposed method is supported by following propositions.

Proposition 4: Let $\alpha, \beta \in \mathbb{R}$ and $[x], [y] \subset \mathbb{R}$. Then

$$\text{width}(\alpha.[x] + \beta.[y]) = |\alpha|.\text{width}([x]) + |\beta|.\text{width}([y]). \quad (20)$$

Proof: The proposition is proved by considering three cases ($\alpha, \beta \geq 0$; $\alpha, \beta \leq 0; \beta < 0 < \alpha$) using interval computations included: $[u] + [v] = [\underline{u} + \underline{v}, \overline{u} + \overline{v}]$, $-[u] = [\overline{u}, \underline{u}]$ and $\lambda.[u] = \text{sign}(\lambda)\,[|\lambda|.\underline{u}, |\lambda|.\overline{u}]$, $\forall \lambda \in \mathbb{R}$. ■

Proposition 5: Let $[x], [y] \subset \mathbb{R}$: width($[x]$) $\leq$ width($[y]$). Let $[z] = \alpha.[x] + (1-\alpha).[y]$ with $\alpha \in [0,1]$, then

a) width($[x]$) $\leq$ width($[z]$) $\leq$ width($[y]$),
b) $[x] \cap [y] \ \subseteq \ [z]$.

Proof: By applying Proposition 4, width($[z]$) is a weighted average of width($[x]$) and width($[y]$). Thus, width($[z]$) obtains an intermediate value between the other two width values. The last statement of the Proposition is proved regarding:
$[z] = \{z \ : \ \exists x \in [x], \exists y \in [y], z = \alpha x + (1-\alpha)y\}$,
$\forall u \in [x] \cap [y], \exists u \in [x], \exists u \in [y], u = \alpha u + (1-\alpha)u$. ■

The regularization method consists in replacing $[\hat{x}_k]$ in the RHS of line 11 of the GG algorithm by a *regularized term*:

$$\text{Reg}([\hat{x}_k]) = \alpha \cdot C^+ \cdot [y_k] + (1-\alpha) \cdot [\hat{x}_k] \ , \quad (21)$$

where C^+ is the Moore-Penrose pseudoinverse of C and $\alpha \in [0,1]^{n_x}$ is a chosen *regularization factor*. Knowing that $[y_k]$ contains the exact measurement $y_k = C \cdot x_k$, the set $C^+ \cdot [y_k]$ provides the best approximation of the inverse image $C^{-1}([y_k])$ regarding that:

$$\forall u \in C^+ \cdot [y_k], \ \exists y \in [y_k], \ \exists x \in [x] : \ u = C^+ \cdot y = C^+ \cdot C \cdot x,$$

and hence $C \cdot u = C \cdot x = y$ (since $C \cdot C^+ \cdot C = C$), while

$$x_k \ \in \ C^{-1}([y_k]) = \{x \in [x] : \ \exists y \in [y_k], \ C \cdot x = y\} \ .$$

If C is invertible, then $C^+ \cdot [y_k] = C^{-1} \cdot [y_k] = C^{-1}([y_k])$ but in general,

$$C^+ \cdot [y_k] \ \subseteq \ C^{-1}([y_k]) \ \subseteq \ C^+ \cdot [y_k] + \text{Ker}(C),$$

where Ker(C) denotes the kernel of C, and one desires x_k is contained in $C^+ \cdot [y_k]$ or "close" to it.

So, equation (21) provides a regularization of $[\hat{x}_k]$ by taking into account the x_k information by the use of $C^+ \cdot [y_k]$ controlled with a regularization factor α. The smaller α is used, the lesser influence of $C^+ \cdot [y_k]$ is contributed and

hence the fact that $x_k \in \text{Reg}([\hat{x}_k])$ is more conservative, and vice versa. For $\alpha = 0$ (and $\gamma_k = 1$), $\text{Reg}([\hat{x}_k]) = [\hat{x}_k]$ and the GG algorithm is recovered. In any case, by Proposition 5, $\text{Reg}([\hat{x}_k])$ contains $(C^+ \cdot [y_k]) \cap [\hat{x}_k]$ and has a reduced width compared to its former $[\hat{x}_k]$ when an appropriate α is applied. Thus, the proposed method is controllable, robust and simple to implement. It is summarized as follows:

Algorithm 2 ROBUST GENERIC GUARANTOR (RGG)

Retain the entire GG algorithm with some replacements:

i) $\rho_k = \rho_{k-1} + (T.L_a.\|\rho_{k-1}\| + \delta) * \gamma_k$ (line 9)

ii) $[\hat{x}_k] \leftarrow \text{Reg}([\hat{x}_k]) \cap ([x_0] + k.T.[f]) \cap [x]$ (line 11)

for a chosen penalization factor γ_k and $\text{Reg}([\hat{x}_k])$ is determined as (21).

IV. APPLICATIONS

A. Suspension model

In this simulation, we consider the Magneto-Rheological (MR) damper model that was used in [13] with almost all parameters and settings are kept unchanged. This model is represented by the following system:

$$\begin{cases} m_s \ddot{z}_s(t) &= -k_s z_{def}(t) - F_d(t), \\ m_{us} \ddot{z}_{us}(t) &= k_s z_{def}(t) + F_d(t) - k_t(z_{us}(t) - z_r(t)) \end{cases}, \quad (22)$$

where $z_{def}(t) = (z_s(t) - z_{us}(t))$ is the suspension deflection, z_s and z_{us} are the chassis and unsprung masses bounce, m_s and m_{us} are sprung and unsprung masses, z_r is the road disturbance and F_d is the damper force defined by

$$\begin{aligned} F_d(t) &= c_0 \dot{z}_{def}(t) + k_0 z_{def}(t) \\ &\quad + f_I(t) \tanh\left(c_1 \dot{z}_{def}(t) + k_1 z_{def}(t)\right), \end{aligned}$$

with c_0, k_0, c_1, k_1 chosen as [14]: $c_0 = 1500$ (Nsm^{-1}), $c_1 = 129$ (sm^{-1}), $k_0 = 989$ (Nm^{-1}), $k_1 = 85$ (m^{-1}), and f_I is a controllable force depending on the input current I and satisfying the dissipativity constraint $0 < f_{\min} \leq f_I \leq f_{\max}$. In this simulation, $f_{\min} = 1000$ Nm^{-1}, $f_{\max} = 1500$ Nm^{-1} and other parameter values used are presented in Table I issued from [15].

Symbol	Value	Unit	Signification
m_s	315	kg	sprung mass
m_{us}	37.5	kg	unsprung mass
k_s	29500	Nm^{-1}	suspension linearized stiffness
k_t	208000	Nm^{-1}	tire stiffness
z_{def}	$[-0.09; 0.05]$	m	suspension bound (stroke limit)

TABLE I

LINEARIZED RENAULT MÉGANE COUPÉ PARAMETERS OF THE QUARTER VERTICAL MODEL (FRONT SUSPENSION).

Putting: $x_t = [z_s(t), \dot{z}_s(t), z_{us}(t), \dot{z}_{us}(t)]^T$ as state variable under consideration whose components are $\{x_t(i)\}_{i=1:4}$, $u_t = f_I(t)$ controllable input and $w_t = z_r(t)$ disturbance, then (22) has the state-space representation

$$\dot{x}_t = f(x_t, u_t, w_t), \quad f = [f_1, ..., f_4]^T, \quad (23)$$

where $f_1(x_t, u_t, w_t) = x_t(2)$, $f_3(x_t, u_t, w_t) = x_t(4)$, $f_2(x_t, u_t, w_t) = \left(a^T x_t - u_t \tanh(b^T x_t)\right)/m_s$,

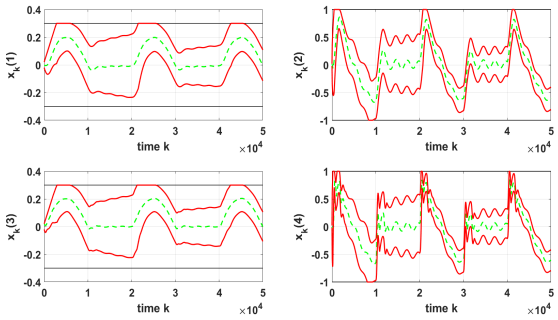

Fig. 1. MR damper model - Estimation results by RGG with $\gamma_k = k^{-1}$ and $\alpha = \mathbf{0}$.

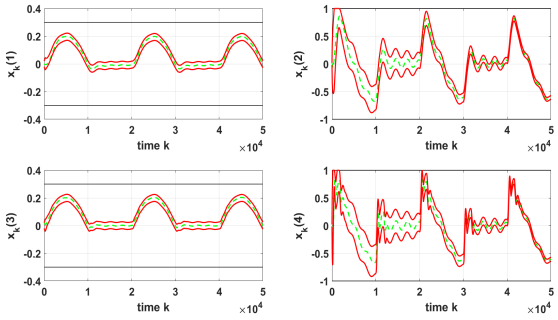

Fig. 2. MR damper model - Estimation results by RGG with $\gamma_k = k^{-1}$ and $\alpha = (2.10^{-2}, 10^{-4}, 2.10^{-2}, 10^{-4})^T$.

$f_4(x_t, u_t, w_t) = \left(c^T x_t + u_t \tanh(b^T x_t) + k_t w_t\right)/m_{us}$, $a = [-k_s - k_0, -c_0, k_s + k_0, c_0]^T$, $b = [k_1, c_1, -k_1, -c_1]^T$, $c = -a - [0, 0, k_t, 0]^T$. The functions $\{f_i\}_{i=1:4}$ are Lipschitz according to x_t with coefficients: $L_1 = L_3 = 1$, $L_2 = (\|a\| + f_{\max}.\|b\|)/m_s$, $L_4 = (\|c\| + f_{\max}.\|b\|)/m_{us}$, computed by using following properties: $\forall x, y \in \mathbb{R}$, $|\tanh(x)| \leq 1 \wedge |x|$, $|\tanh(x) - \tanh(y)| \leq |\tanh(x - y)|$.

Further numerical setting in order to apply the proposed algorithms is as follows: $[u] = [1000, 1500]$, $[w] = [0, 0.2]$, $T = 10^{-4}$, $[x] = ([-0.3, 0.3], [-1, 1], [-0.3, 0.3], [-1, 1])^T$, $\delta = 0.1$, $[x_0] = 0.02 \times ([-1, 1], [-1, 1], [-1, 1], [-1, 1])^T$.

The initial state x_0 is chosen at random in $[x_0]$. For $N = 5.10^4$, $\forall k = 1 : N$, $u_k = \text{mid}([u])$, $w_k = \sup([w]) \times \max\{0, \sin(\pi.k.T)\}$ and x_k is generated using (23). Then $y_k = z_{def} = x_k(1) - x_k(3)$ simulated by using $C = [1, 0, -1, 0]$ and $[y_k] = (C.x_k \pm 0.05) \cap [-0.09, 0.05]$ regarding the z_{def}'s constraint (see Table I).

The estimation result by applying the RGG algorithm is shown in Figures 1 and 2, in which *green dashed lines* are real states, *red lines* are $\sup([\hat{x}_k])$ and $\inf([\hat{x}_k])$. As commonly known, the GG algorithm provides more conservative estimation results but less practically efficient. In this simulation, using for instance the Forward-Backward Propagation (FBP) contractor together with the GG algorithm is also almost inefficient in terms of contraction effect. The reason is that $[x_k(1)]$ and $[x_k(3)]$ are almost equal for all $k = 1 : N$ which makes the contraction procedure insight of FBP contractor works with low success. Here, figures 1 and 2 emphasize the effect of the parameters γ_k, α on the efficiency of the proposed RGG.

B. Chemical reaction

In this part, we consider another nonlinear system describing the chemical reaction related to oil-reservoir often used in the literature [16], [17], [18]. The ODE system is as follows:

$$\dot{x}_t = f(x_t) = \left(x_t(2),\ \ x_t^2(2) - 3/[s + x_t^2(1)]\right)^T \qquad (24)$$

where $x_t = (x_t(1), x_t(2))^T$ is the system state and $s \in [10^{-4}, 10^{-1}]$ is the stiffness coefficient. The system is subject to initial value sensitivity, step-size sensitivity and stiffness dependency (Figure 3).

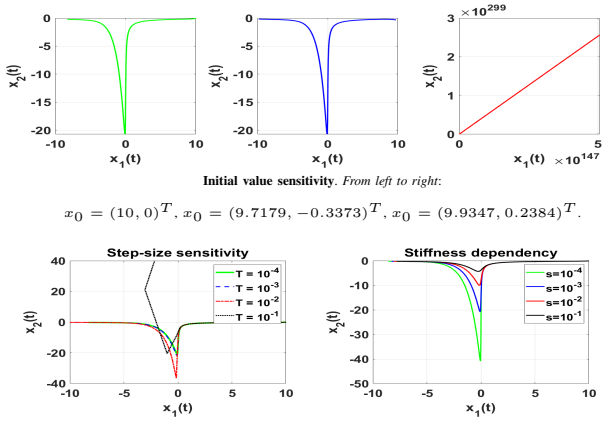

Fig. 3. Chemical reaction model sensitivity.

The RGG simulation results are shown in Figure 4, assuming the exact initial value $x_0 = (10, 0)^T$, the stiffness $s = 10^{-3}$ and time horizon $[0, 50]$(s) (similar to [16]). The step-size is chosen as $T = 10^{-4}$(s) and other parameters needed for the algorithm are $\gamma_k = k^{-2}$ and $\alpha = (0, 0)^T$ (i.e. without measurement), $\rho_0 = 0$, $\delta = 1 = \sigma$, $L = (1, 10^{10})^T$, $[x] = ([-15, 15], [-25, 5])^T$.

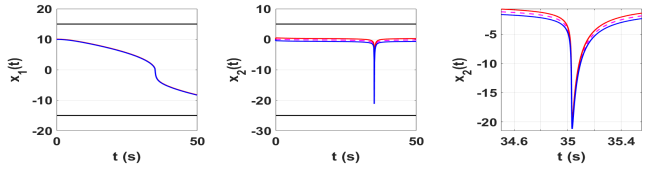

Fig. 4. Chemical reaction - Estimation results by RGG. *Dashed lines*: x_k. *Red lines*: $\sup([\hat{x}_k])$. *Blue lines*: $\inf([\hat{x}_k])$.

V. Conclusion

The generic interval filter proposed in this paper is simple, robust and controllable. It can be used in a wide range of applications, in which processes can be modeled by continuous functions restricted on bounded domains regardless of their differentiability. A crucial conception part of the method, making it flexible and efficient, is the compromise interactively of the step-size T and the control factors δ, γ_k, α. As the computation engines and sensors are more and more powerful, the use of an appropriate small step-size T is no longer an intractable issue. However, if smaller step size can not be acheived, the effect of the control factors become more crutial in the algorithm and vise versa. The proposed method might be used as a first verification of estimation results or a benchmark tool in comparison with other filters.

References

[1] L. Jaulin, "Nonlinear bounded-error state estimation of continuous-time systems," *Automatica*, vol. 38, no. 6, pp. 1079–1082, 2002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0005109801002849

[2] M. H. Davis and S. I. Marcus, "An introduction to nonlinear filtering," in *Stochastic Systems: The Mathematics of Filtering and Identification and Applications: Proceedings of the NATO Advanced Study Institute held at Les Arcs, Savoie, France, June 22–July 5, 1980*. Springer, 1981, pp. 53–75.

[3] G. Welch and G. Bishop, "An introduction to Kalman Filter," in *International Conference on Computer Graphics and Interactive Techniques*, 1995. [Online]. Available: https://api.semanticscholar.org/CorpusID:215767582

[4] S. J. Julier and J. K. Uhlmann, "New extension of the Kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. Spie, 1997, pp. 182–193.

[5] W. Bai, W. Xue, Y. Huang, and H. Fang, "On extended state based Kalman filter design for a class of nonlinear time-varying uncertain systems," *Science China Information Sciences*, vol. 61, pp. 1–16, 2018.

[6] Y. Liu, Z. Wang, X. He, and D.-H. Zhou, "Filtering and fault detection for nonlinear systems with polynomial approximation," *Automatica*, vol. 54, pp. 348–359, 2015.

[7] F. Abdallah, A. Gning, and P. Bonnifait, "Box particle filtering for nonlinear state estimation using interval analysis," *Automatica*, vol. 44, no. 3, pp. 807–815, 2008.

[8] Q. H. Lu, S. Fergani, and C. Jauberthie, "Reinforced likelihood box particle filter," *IEEE Control Systems Letters*, vol. 7, pp. 502–507, 2022.

[9] A. A. de Paula, G. V. Raffo, and B. O. Teixeira, "Zonotopic filtering for uncertain nonlinear systems: Fundamentals, implementation aspects, and extensions [applications of control]," *IEEE Control Systems Magazine*, vol. 42, no. 1, pp. 19–51, 2022.

[10] D. Bhattacharjee and K. Subbarao, "Set-membership filter for discrete-time nonlinear systems using state-dependent coefficient parameterization," *IEEE Transactions on Automatic Control*, vol. 67, no. 2, pp. 894–901, 2021.

[11] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. London: Springer-Verlag, 2001.

[12] S. Rump, "INTLAB - INTerval LABoratory," in *Developments in Reliable Computing*, T. Csendes, Ed. Dordrecht: Kluwer Academic Publishers, 1999, pp. 77–104, http://www.tuhh.de/ti3/rump/.

[13] Q. H. Lu, S. Fergani, and C. Jauberthie, "Fault detection combining adaptive degrees of freedom χ2-statistics and interval approach for nonlinear systems," in *The 22nd World Congress of the International Federation of Automatic Control (IFAC)*, Yokohama, Japan, July 2023. [Online]. Available: https://laas.hal.science/hal-04069589

[14] E. Nino-Juarez, R. Ramirez-Mendoza, R. Morales-Menendez, O. Sename, and L. Dugard, "Minimizing the frequency effect in a black box model of a magneto-rheological damper," in *Mini conference; 11th, Vehicle system dynamics, identification and anomalies*. Technical University of Budapest, 2008, pp. 733–742.

[15] S. Fergani, "Robust multivariable control for vehicle dynamics," PhD Thesis, Grenoble INP, GIPSA-lab, Control System dpt., Grenoble, France, October 2014.

[16] O. Bouissou, A. Chapoutot, and A. Djoudi, "Enclosing Temporal Evolution of Dynamical Systems Using Numerical Methods," in *5th NASA Formal Methods Symposium*, ser. LNCS. Springer Verlag, May 2013, vol. 7871, p. 108. [Online]. Available: https://hal.science/hal-00819730

[17] J. Alexandre Dit Sandretto and A. Chapoutot, "Validated Explicit and Implicit Runge-Kutta Methods," *Reliable Computing electronic edition*, vol. 22, July 2016. [Online]. Available: https://hal.science/hal-01243053

[18] J. R. Cash and A. H. Karp, "A variable order runge-kutta method for initial value problems with rapidly varying right-hand sides," *ACM Trans. Math. Softw.*, vol. 16, pp. 201–222, 1990. [Online]. Available: https://api.semanticscholar.org/CorpusID:1133671