Robust algebraic state estimation of chaotic systems

Johann Reger[†⋆], Philipp Mai[†], and Hebertt Sira-Ramírez[‡]

***Abstract*— In this article, we propose an improvement of a recently introduced algebraic approach for the non-asymptotic state and parameter estimation of nonlinear systems. In particular, we increase the robustness of the estimation method with respect to zero mean, high frequency, measurement noises by introducing a so-called *invariant filtering* technique. In order to reduce an already fast transient to the convergence, when subject to measurement noise, we devise an estimation policy consisting of two overlapping estimators with appropriate switchings between their results. These are two identical time-shifted estimators running in parallel with an overlapping estimation period. The benefits of our method are demonstrated on the state observation of a chaotic system of the Rössler type.**

I. INTRODUCTION

Synchronization of chaotic systems, in particular *chaotic based communication*, is one of the most active and promising fields of research in communication. A coarse picture of the method is as follows: a chaotic system, the so-called master system, and an associated subsystem or copy of the master system, called slave system, share a set of common states by means of which the master system is to drive, some, or all, of the states of the slave system towards synchronization. For a wide class of chaotic systems, this synchronization turns out to be robust even in noisy environments [12], [2]. A feature of chaotic systems is that the estimation of the long-term state evolution, out of an observed segment, is quite ill-conditioned with noise-like behaviors. Nevertheless, as the states constitute a homogenous response of a nonlinear dynamic system, they bear more structure than stochastic systems. In this respect, the spectra of chaotic systems are flat and widespread, implying a large bandwidth, so it may be expected a low probability of detection and low degree of interception with transmitters on the same band. On the other hand, in the immediate, near-term, state evolution chaotic systems may be easily controlled. All the afore-mentioned features add up to the complexity-free signal generation in simple circuits and, in fact, these features are recently exploited in novel real-world applications in communication [1] where a secure data transmission scheme from the transmitter (master) to the receiver (slave) is established.

In the following, we assume that the slave system has no influence on the master system and, as a consequence, the slave system may be considered as an observer for the states of the master system [11]. Additionally, with the system parameters known, eavesdropping of at least a part of the non encrypted driving states may open up the reconstruction of all other master system states. In other words, the transmission becomes insecure as long as the state reconstruction process is fast enough. For rapidness, instead of referring to standard observers that exhibit an exponential error dynamics convergence, we propose to employ a faster, generally non-asymptotic, algebraic approach for the open-loop state estimation of nonlinear systems [6], [7], [14], [13]; for applications in identification see [5]. This method renders possible the determination of derivative estimates with respect to an arbitrary time signal, e. g. the measured driver states, under considerable noise levels. Together with the fast convergence of the method, the robustness against noise may enable to find the super-key (the set of parameters of the chaotic system) even in more realistic setups; for the noise-free case a method was already presented in [16],[10].

The main results of this contribution are two amendments of the algebraic estimation scheme which improve the performance under zero mean high frequency noise. In the first instance, we derive a formula for the derivative estimates at a fixed instant of time, all based on the inversion of a formula presented in [13]. By means of this formula, an invariant filtering technique is applied which amounts to low pass filter independently and uniformly the numerator and denominator expressions in the estimator formula. The objective is that of suppressing the corresponding noises, leaving the estimation velocity unaffected. In the second instance, we adapt our method to improve the transient behavior exhibited when the resetting of the estimators are needed in significant noise level situations. The solution we propose is to run two overlapping estimators in parallel and to take the estimate of that one which already has converged out of the fast transient.

This contribution is organized as follows: Section II briefly introduces the notion of algebraic derivative estimation. Resorting to results from [13], a formula is derived for the filter-based estimation of time signal derivatives at some time instant. Section III recalls the concept of observability in the context of nonlinear systems. In Section IV, the observability of the chaotic Rössler system is examined. For improving the convergence of the derivative estimators under noise, the notion of overlapping estimators is introduced in Section V. For the general attenuation of high frequency zero mean noise, an invariant filtering technique is motivated in Section VI before conclusions are drawn in Section VII. The proofs associated with Section II may be found in a complete form in the Appendix section.

⋆ Author for correspondence

† J. Reger and P. Mai are with the Institute of Automation and Control (EIT 8.1), Universität der Bundeswehr München, Werner-Heisenberg-Weg 39, 85579 Neubiberg, Germany, email: reger@ieee.org, pmai@ieee.org

‡ Sección de Mecatrónica, Centro de Investagación y Estudios Avanzados del IPN (CINVESTAV-IPN), 07300 México D.F., México, email: hsira@cinvestav.mx

II. Algebraic Time Derivative Estimation

For simplicity, we recall some steps of the derivation presented in [13]. Consider a real-valued polynomial function

$$x_N(t) = \sum_{i=0}^{N} \frac{x_N^{(i)}(t_\mathrm{r})}{i!}\,(t-t_\mathrm{r})^i \in \mathbb{R}[t]$$

of degree N, $t \geq t_\mathrm{r}$, which may be considered as an N-th order Taylor series approximation of a time signal $x(t)$.

We may derive $x_N(t)$ at least $(N+1)$-times with respect to time, so as to obtain an expression identical to zero. In the operator domain, this reads

$$s^{N+1}X_N(s) - \sum_{i=0}^{N} s^{N-i}x_N^{(i)}(t_r) = 0$$

which involves standard rules of operational calculus, only. In order to eliminate the initial conditions, $x_N^{(i)}(t_\mathrm{r})$, differentiate $N+1$ times wrt. the complex operator s and obtain:

$$\frac{d^{N+1}}{ds^{N+1}}\left(s^{N+1}X_N(s)\right) = 0\,.$$

Since the highest polynomial degree in this expression is $N+1$, it easy to see that we may express the $(N+1)$-st time derivative of $x_N(t)$ in terms of lower order derivatives. Therefore, pre-multiplication by an iterated integration operator results in a recursive system for determining all required time derivatives of $x_N(t)$, i. e. we may solve in a recursive fashion the linear system

$$\frac{1}{s^\nu}\frac{d^{N+1}}{ds^{N+1}}\left(s^{N+1}X_N(s)\right) = 0, \quad \nu = 0,1,\ldots,N$$

for the transformed time derivatives $s^iX_N(s)$. We refer to [13] for the result

$$x_N^{(k)}(t) = \frac{(N+k)!}{(N-k)!\,k!}\,\frac{x_N(t)}{(t-t_\mathrm{r})^k} + \sum_{j=1}^{k}\binom{N+k-j}{k-j}\frac{(N-j)!}{(N-k)!}\,\frac{z_j(N,t)}{(t-t_\mathrm{r})^{N+1+k-j}} \quad (1)$$

where the filter states $z_j(N,t)$, $j=1,\ldots,N$, obey

$$\dot z_j(N,t) = \binom{N+1}{j+1}^2 (j+1)!\,(-1)^j\;(t-t_\mathrm{r})^{N-j}x_N(t) + z_{j+1}(N,t)$$
$$\dot z_N(N,t) = (N+1)!\,(-1)^N\;x_N(t) \quad (2)$$

which is an N-th order time-varying, linear filter with homogeneous initial conditions $z_j(N,t_\mathrm{r}) = 0$.

In the Appendix, we use formulae (1) and (2) so as to derive the expression for the initial conditions $x_N^{(i)}(t_\mathrm{r})$. To this end, we solve for the coefficients, $x_N^{(i)}(t_\mathrm{r})$, from

$$x_N^{(k)}(t) = \sum_{i=k}^{N}\frac{t^{i-k}}{(i-k)!}\,x_N^{(i)}(t_\mathrm{r})$$

viewed as a linear system. From the Appendix, we recall

$$x_N^{(i)}(t_\mathrm{r}) = \frac{x_N(t)}{(t-t_\mathrm{r})^i}\,\frac{(N+i)!}{(N-i)!\,i!}(-1)^{N+i} + \sum_{j=1}^{N}\frac{(N+i-j)!\,(-1)^{N+i}}{(N-i)!\,i!}\,\frac{z_j(N,t)}{(t-t_\mathrm{r})^{N+1+i-j}} \quad (3)$$

for the the i-th ($i=0,\ldots,N$) derivative estimate at time $t=t_\mathrm{r}$. Equation (3) is written in terms of the linear filter states introduced in (2).

III. Observability of Autonomous Nonlinear Systems

We center our attention on the observability of autonomous nonlinear SISO systems, i.e., on systems characterized by

$$\dot x = f(x)\,,$$
$$y = h(x)\,, \quad (4)$$

where $x \in \mathbb{R}^n$ is the system state, and $y \in \mathbb{R}$ is the output of the system.[1] We assume that $g(\cdot)$ and $h(\cdot)$ are sufficiently smooth. The system is said to be *locally observable* from the output $y = h(x)$ if the map

$$\begin{bmatrix} y \\ \dot y \\ \vdots \\ y^{(n-1)} \end{bmatrix}\begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1}h(x) \end{bmatrix} \quad (5)$$

is locally full rank n, see [8]. In this map, the expression $L_f^k h(x)$ is the Lie-derivative of the smooth function $h(x)$ with respect to the vector field f, which is defined as

$$L_f^k h(x) = \frac{\partial L_f^{k-1}h(x)}{\partial x}f(x)\,, \quad L_f^0 h(x) = h(x)\,,$$

recursively. It is a well known result, that if the above map is locally full rank n, then the state vector, x, of the system can be locally expressed as a smooth *differential function* of $y(t)$, i. e. a smooth function of y and a finite number (in fact $n-1$) of its time derivatives [3]. We also address this type of function as a *differential parametrization* of the state x in terms of the observable output y, that is

$$x = \Phi(y, \dot y, \ddot y, \cdots, y^{(n-1)}) \quad (6)$$

for some smooth function Φ. Hence, the state x is uniformly observable from the output y. In this framework, the observation issue reduces to properly estimate the derivatives of some measured signal, as already pointed out in [4].

IV. Rössler System

In the remainder of this paper, we will discuss algebraic state estimation (observation) on the basis of a well-established chaotic system, the Rössler system given by

$$\dot x_1 = -x_2 - x_3\,,$$
$$\dot x_2 = x_1 + a x_2\,, \quad (7)$$
$$\dot x_3 = b + x_1x_3 - c x_3\,.$$

We specify the system parameters $a = b = 0.2$, $c = 5$ and set the initial condition to $x_1(0) = 8$, $x_2(0) = -5$ and

[1] For the purposes of Section IV, it would have been sufficient to introduce the concept of linear observability. However, as the estimation methods developed apply in the same manner to the case of nonlinear systems, this was reason enough to account for nonlinear observability here.

$x_3(0) = 1$. It is easy to verify that $y = x_2$ is an appropriate output for representing the remaining states, x_1 and x_3. This results in the differential parameterizations

$$x_1 = \dot{y} - a\,y\,, \tag{8}$$

$$x_3 = -y + a\,\dot{y} - \ddot{y}\,. \tag{9}$$

In the following sections, the algebraic derivative estimation method, explained in Section II of the paper, is applied to the Rössler system (7) using the noisy output $y = x_2(t) + \xi(t)$ as measured signal. Here $\xi(t)$ denotes zero mean Gaussian noise with standard deviation of $\sigma = 0.001$. The system (7) as well as the estimator equations were integrated using MATLAB's standard solver ode45.

V. OVERLAPPING DERIVATIVE ESTIMATORS TECHNIQUE

From equation (3) it is clear that for numerical reasons, a small interval of time ϵ has to be elapsed before the results of the estimators become valid (note the singularity at $t = t_{\mathrm{r}}$). Also, depending on the amount of noise $\xi(t)$ associated with the measured signal $y = x_2(t) + \xi(t)$, a certain period of integration time is also needed in the filters for neutralizing the noise effects. On the other hand, due to the nature of a Taylor series approximation of finite length, our estimation will start to diverge from the true signal $x_2(t)$ when the time difference $t - t_{\mathrm{r}}$ becomes significantly large. Figure 1 illustrates these effects considering the estimation of $\dot{y}$ of the Rössler system (7) using an approximation order of $N = 5$ (without noise).

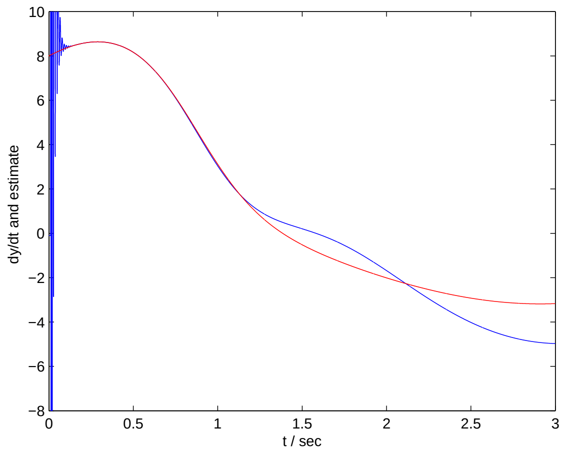

Fig. 1. Output derivative $\dot{y} = \dot{x}_2$ (red line) of the Rössler system and estimate (blue line) — no resetting, no overlapping estimation windows, noiseless

For a proper update of the derivative estimation, a reset of the filter states to zero may be necessary at some instant, implying a new ϵ-period of integration time before we again achieve accurate estimation results. This fact is illustrated in Figure 2 where we reconstruct x_1 by means of (3) and by the respective differential parametrization (8). We used an approximation order of $N = 5$ and a reset interval of length $h = 0.35$sec., which means that the filter is reset to zero whenever $t = k\,h$, $k \in \mathbb{N}$.[2]

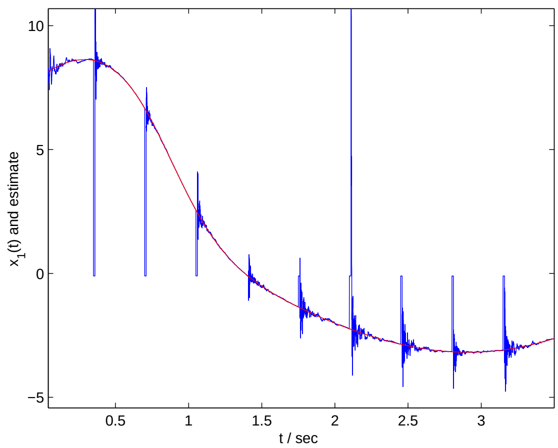

Fig. 2. State x_1 (red line) of the Rössler system and its estimate (blue line) referring to an order $N = 5$ estimator — under resetting, with noise

In order to enlarge the intervals of validity of the estimation scheme proposed, we introduce the following estimation policy: we run two estimators simultaneously but in an overlapping fashion so as to get valid results at all times except from the very first ϵ-interval. Let $h > \epsilon$ be the duration in which the estimation results are considered valid (before divergence). Then we will obtain valid estimates at any time when alternatingly reading out the estimates from the two estimators that are reset every $2h$ seconds each, but shifted in time by an interval of length h. More precisely, we switch to the estimates of, say, the first estimator whenever $t \in (2kh, (2k+1)h]$, and adopt those of the second estimator whenever $t \in ((2k-1)h, 2kh]$. We shall call this policy *a switched overlapping estimators technique*. Fig. 3 shows the advantages of this technique applied to the estimation of the state x_3 (most sensitive since $\ddot{y}$ is involved) of the Rössler system, using equation (9). Under the given noise level, we obtain an estimation accuracy for x_3 which would have been impossible to achieve using only one estimator.

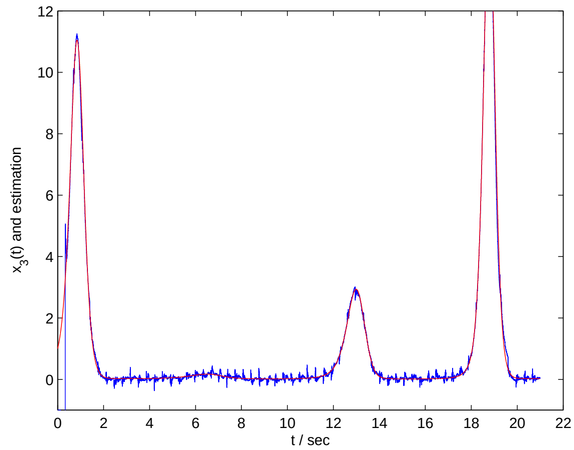

Fig. 3. State x_3 (red line) of the Rössler system and its estimate (blue line) calculated from order $N = 5$ *overlapping estimators* — with noise

[2] Of course, there are more refined reset policies. For details on absolute error bound and integral error bound related reset policies, respectively, see [14] and [13].

VI. INVARIANT FILTERING

In addition to the above-presented technique, we may further reduce the effect of measurement noise on $x_2(t)$ by invoking a particular low pass filtering technique. Splitting the time varying expressions within the estimator equation (3) into numerator $n(t)$ and denominator $d(t)$ we may then write, in a slight abuse of denotation,

$$x_N^{(i)}(0) = \frac{n(t)}{d(t)} = \text{const.} \;\rightarrow\; x_{N,\mathrm{f}}^{(i)}(0) = \frac{n_\mathrm{f}(t)}{d_\mathrm{f}(t)} = \frac{F(s)\,n(t)}{F(s)\,d(t)}$$

Hence, we *invariantly filter* $n(t)$ and $d(t)$ employing, for example, the following second order filter

$$\ddot{n}_\mathrm{f}(t) + 2\xi\omega_0\,\dot{n}_\mathrm{f}(t) + \omega_0^2\,n_\mathrm{f}(t) = \omega_0^2\,n(t)\,, \qquad (10)$$

$$\ddot{d}_\mathrm{f}(t) + 2\xi\omega_0\,\dot{d}_\mathrm{f}(t) + \omega_0^2\,d_\mathrm{f}(t) = \omega_0^2\,d(t)\,. \qquad (11)$$

In this manner, we may attenuate high frequency measurement noise in an appropriate bandwidth by specifying a cut-off frequency ω_0 and a damping factor $\xi \geq \sqrt{2}/2$. Figure 4 highlights the improvement compared to Fig. 3. The effect of the additive noise is substantially reduced. The devised low pass filter was a second order filter with cut-off frequency $\omega = 0.1$ and damping $\xi = 1$.

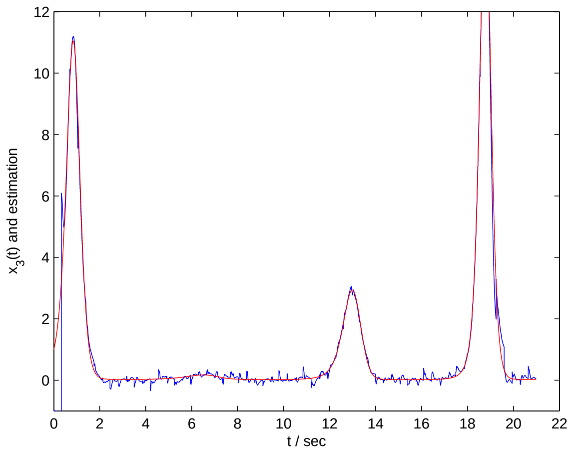

Fig. 4. State x_3 (red line) of the Rössler system and its estimate (blue line) calculated from order $N = 5$ *overlapping estimators* with *invariant filtering* — with noise

Figure 5 shows the accuracy of the time derivative estimates substituted in the differential parametrization (9).

VII. CONCLUSION

In this article, two amendments for improving the robustness of algebraic state estimators, which are subject to substantial noise, are proposed and illustrated considering the state estimation of the chaotic Rössler system. Since the time-varying filters used in the algebraic state estimator realizations may require some time to yield valid results under noise, a switched overlapping estimators policy is employed: From two shifted estimators the estimate value achieved by the convergent estimator is used while the other is starting to diverge and is properly reset. The outcome of this technique is a considerable acceleration of the convergence blocking off the computation transients occurring just after the estimator

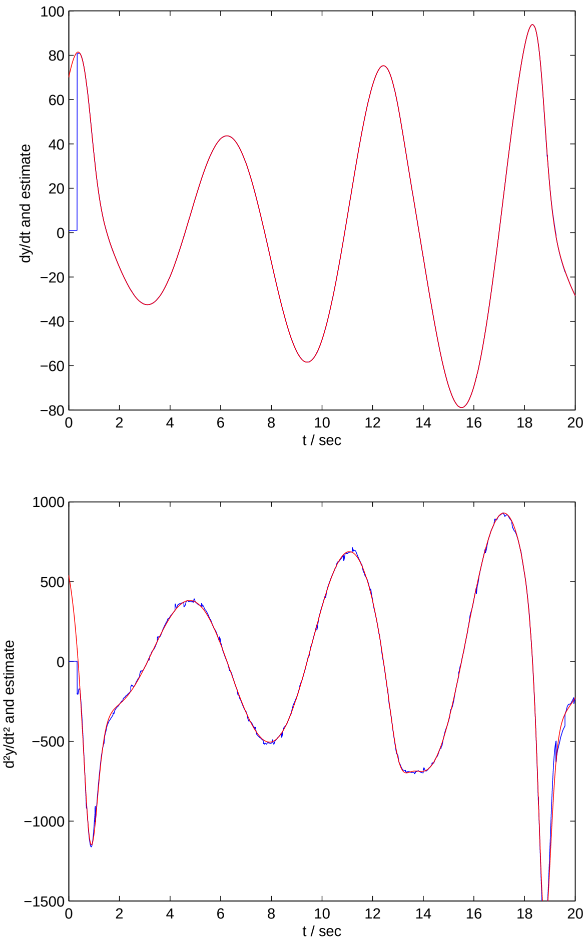

Fig. 5. Output derivatives $\dot{y}$ and $\ddot{y}$ (red line) of the Rössler system and its state estimate (blue line) calculated from order $N = 5$ *overlapping estimators* with *invariant filtering* — with noise

resettings. A second enhancement is the *invariant filtering* technique proposed here for the overall attenuation of zero mean high frequency noise effects. This is achieved by introducing low pass filtering, without detrimental impacts on the velocity of convergence of the derivatives estimator. Both improvements add up to a significant reduction of the estimator sensitivity wrt. noise and may render possible the application of algebraic derivative estimation in the observation of a wide class of nonlinear systems.

In a forthcoming paper, the method will be extended to the simultaneous estimation of system parameters and states of chaotic systems. In particular, we will use the afore-presented improvements of the algebraic estimation scheme for a "fast super-key" decoding device in chaotic communication systems.

REFERENCES

[1] A. Argyris et al., "Chaos-based communications at high bit rates using commercial fibre-optic links", *Nature*, vol. 438, November, 2005, pp. 343–346.

[2] K. M. Cuomo, A. V. Oppenheim and S. H. Strogatz, "Synchronization of Lorenz-Based Chaotic Circuits with Applications to Communications", *IEEE Trans. Circuits Syst-II: Analog and Digital Signal Processing*, vol. 40, 1993, pp. 626–633.

[3] S. Diop and M. Fliess, "Nonlinear observability, identifiability and persistent trajectories", *Proc. 36th IEEE Conference on Decision and Control*, Brighton, 1991, pp. 714–719.

[4] S. Diop, J. W. Grizzle and F. Chaplais, "On numerical differentiation algorithms for nonlinear estimation", *Proc. 39th IEEE Conference on Decision and Control (CDC'00)*, Australia, 2000.

[5] M. Fliess and H. Sira-Ramírez, "An algebraic framework for linear identification", *ESAIM, Control, Optimization and Calculus of Variations*, vol. 9, 2003, pp. 151–168.

[6] M. Fliess and H. Sira-Ramírez, "Reconstructeurs d'Etat" *C.R. Academie des Sciences de Paris*, t. 332, Série I, vol. 338(1), 2004, pp. 91–96.

[7] M. Fliess and H. Sira-Ramírez, "Control via State Estimations of Flat Systems", *Proc. 6th IFAC Symposium on Nonlinear Control Systems*, Stuttgart, Germany, 2004.

[8] R. Hermann and A. Krener, "Nonlinear Controllability and Observability", *IEEE Transactions on Automatic Control*, vol. 22(5), 1977, pp. 728–740.

[9] H. J. Huijberts, H. Nijmeijer and R. M. Willems, "A control perspective on communications using chaotic systems", *Proc. 37th IEEE Conf. on Decision and Control*, Tampa, Florida, 1998, pp. 1957-1962.

[10] L. Liu, X. Wu and H. Hu, "Estimating system parameters of Chua's circuit from synchronizing signal", *Physics Letters*, vol. A 324, 2004, pp. 36–41.

[11] H. Nijmeijer and M. Y. Mareels, "An Observer Looks at Synchronization", *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 44(10), 1997, pp. 882–890.

[12] L. Pecora and T. Carroll, "Driving systems with chaotic signals", *Phys. Rev.*, vol. A44(4), 1991, pp. 2374–2383.

[13] J. Reger, H. Sira-Ramírez and M. Fliess, "On non-asymptotic state estimation of nonlinear systems", *Proc. 44th IEEE Conference on Decision and Control (CDC'05)*, Seville, 2005.

[14] H. Sira-Ramírez and M. Fliess, "On the output feedback control of a synchronous generator", *Proc. 43rd IEEE Conference on Decision and Control (CDC'04)*, Atlantis, Bahamas Islands, 2004.

[15] H. Sira-Ramírez and M. Fliess, "An algebraic state estimation approach for the recovery of chaotically encrypted messages" *International J. of Bifurcation and Chaos*, Vol. 16, No. 2, pp. 1-15, 2006.

[16] P. G. Vaidya and S. Angadi, "Decoding chaotic cryptography without access to the superkey", *Chaos, Solitions and Fractals*, vol. 17, 2003, pp. 379–386.

VIII. Appendix

Here we prove the result in equation (3). From [13] we know that for $t > t_\mathrm{r} > 0$

$$x_N^{(k)}(t) = \frac{(N+k)!}{(N-k)!\,k!}\ \frac{1}{(t-t_\mathrm{r})^k}\,x_N(t) + \sum_{j=1}^{k}\binom{N+k-j}{k-j}\frac{(N-j)!}{(N-k)!}\ \frac{1}{(t-t_\mathrm{r})^{N+1+k-j}}\,z_j(N,t)$$

is valid. Here, $z_j(N,t)$, are the filter states from (2). Since $x_N(t)$ is the Taylor series expansion of $x(t)$, we also have,

$$x_N^{(k)}(t) = \sum_{\nu=k}^{N}\frac{(t-t_\mathrm{r})^{\nu-k}}{(\nu-k)!}\,x_N^{(\nu)}(t_\mathrm{r}) \qquad (12)$$

for $k = 0,\dots,N$. This is a linear system of equations in $x_N^{(\nu)}(t_\mathrm{r})$. For convenience in the following steps of derivation, we may write t instead of $t - t_\mathrm{r}$ and re-substitute at the end.

The right hand side of (12) contains the matrix elements

$$[M_{k\nu}] = \begin{cases} \dfrac{t^{\nu-k}}{(\nu-k)!} & , k \le \nu \le N \\ 0 & , 0 \le \nu < k \end{cases}$$

The elements of the corresponding inverse matrix are

$$[M_{ik}]^{-1} = \begin{cases} \dfrac{(-t)^{k-i}}{(k-i)!} & , i \le k \le N \\ 0 & , 0 \le k < i \end{cases}$$

which may be seen from

$$\sum_{k=0}^{N}[M_{ik}]^{-1}[M_{k\nu}] = \sum_{k=i}^{\nu}\frac{(-t)^{k-i}t^{\nu-k}}{(k-i)!(\nu-k)!} =$$

$$= t^{\nu-i}\sum_{k=i}^{\nu}\frac{(-1)^{k-i}}{(k-i)!(\nu-k)!} = t^{\nu-i}\sum_{\bar{k}=0}^{\nu-i}\frac{(-1)^{\bar{k}}}{\bar{k}!\,(\nu-i-\bar{k})!} =$$

$$= t^{\nu-i}\sum_{\bar{k}=0}^{\bar{\nu}}\frac{(-1)^{\bar{k}}}{\bar{k}!\,(\bar{\nu}-\bar{k})!} = t^{\nu-i}\frac{1}{(\nu-i)!}\sum_{\bar{k}=0}^{\bar{\nu}}\binom{\bar{\nu}}{\bar{k}}(-1)^{\bar{k}} =$$

$$= \left\{\begin{matrix} \dfrac{t^{\nu-i}}{(\nu-i)!}(1-1)^{\nu-i}\,, & \nu \ge i \\ 0 & , \text{ otherwise} \end{matrix}\right\} = \delta_{i\nu}\,.$$

With this inverse matrix we obtain

$$\sum_{\nu=0}^{N}\delta_{i\nu}\,x_N^{(\nu)}(0) = x_N^{(i)}(0) = \sum_{k=0}^{N}\frac{(-t)^{k-i}}{(k-i)!}x_N^{(k)}(t) =$$

$$= \frac{x_N(t)}{t^i}\underbrace{\left(\sum_{k=i}^{N}\frac{(N+k)!\,(-1)^{k-i}}{(k-i)!\,k!\,(N-k)!}\right)}_{=:\,a} +$$

$$\underbrace{\sum_{k=i}^{N}\sum_{j=1}^{k}\binom{N+k-j}{k-j}\frac{(N-j)!\,(-1)^{k-i}}{(k-i)!\,(N-k)!}\ \frac{z_j(N,t)}{t^{N+1+i-j}}}_{=:\,b}\,. \qquad (13)$$

We recall that

$$a = (-1)^i\sum_{k=i}^{N}\frac{(N+k)!\,(-1)^k}{(k-i)!\,k!\,(N-k)!} = \frac{(N+i)!\,(-1)^{N+i}}{(N-i)!\,i!}\,.$$

For evaluating b, note that

$$\sum_{k=i}^{N}\sum_{j=1}^{k}\dots = \sum_{j=1}^{i}\sum_{k=i}^{N}\dots + \sum_{j=i+1}^{N}\sum_{k=j}^{N}\dots$$

Hence, we may write

$$b = \sum_{j=1}^{i}\frac{z_j(N,t)}{t^{N+1+i-j}}(N-j)!\underbrace{\sum_{k=i}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!}}_{=:\,c} + \sum_{j=i+1}^{N}\frac{z_j(N,t)}{t^{N+1+i-j}}(N-j)!\underbrace{\sum_{k=j}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!}}_{=:\,d}$$

and, again, recall that

$$c = d = \frac{(N+i-j)!\,(-1)^{N+i}}{(N-j)!(N-i)!\,i!}\,.$$

It is now obvious that with c and d it follows

$$b = \sum_{j=1}^{N}\frac{(N+i-j)!\,(-1)^{N+i}}{(N-i)!\,i!}\ \frac{z_j(N,t)}{t^{N+1+i-j}}\,.$$

Finally, substituting the expressions for a and b in (13) and allowing for the time shift to t_r results in

$$x_N^{(i)}(t_\mathrm{r}) = \frac{x_N(t)}{(t-t_\mathrm{r})^i}\ \frac{(N+i)!}{(N-i)!\,i!}(-1)^{N+i} +$$

$$+\sum_{j=1}^{N}\frac{(N+i-j)!\,(-1)^{N+i}}{(N-i)!\,i!}\;\frac{z_j(N,t)}{(t-t_{\mathrm{r}})^{N+1+i-j}}$$

and the result follows.

It remains to prove some equivalences:

1) We show: $\sum_{k=i}^{N}\frac{(N+k)!\,(-1)^k}{(k-i)!\,k!\,(N-k)!}=\frac{(N+i)!\,(-1)^N}{(N-i)!\,i!}$.

Simple manipulations add up to the equivalent statement

$$\sum_{k=i}^{N}\binom{N-i}{k-i}\frac{(N+k)!}{k!}(-1)^k=\frac{(N+i)!}{i!}(-1)^N .$$

Hence, we define

$$f(x)=\sum_{k=i}^{N}\binom{N-i}{k-i}\frac{(N+k)!}{k!}\,x^k$$

and show the validity of

$$f(-1)=\frac{(N+i)!}{i!}(-1)^N .$$

Starting with the definition above, we derive

$$\begin{aligned}
f(x)&=\sum_{k=i}^{N}\binom{N-i}{k-i}(N+k)\cdots(1+k)\,x^k=\\
&=\sum_{k=i}^{N}\binom{N-i}{k-i}D_N\,x^{N+k}=D_N\sum_{k=i}^{N}\binom{N-i}{k-i}\,x^{N+k}=\\
&=D_N\sum_{\bar k=0}^{\bar N}\binom{\bar N}{\bar k}\,x^{\bar N+\bar k+2i}=D_N x^{N+i}\sum_{\bar k=0}^{\bar N}\binom{\bar N}{\bar k}\,x^{\bar k}=\\
&=D_N\left(x^{N+i}(1+x)^{N-i}\right)=\\
&=\sum_{\kappa=0}^{N}\binom{N}{\kappa}\left(D_{N-\kappa}x^{N+i}\right)\left(D_\kappa(1+x)^{N-i}\right)=\\
&=\sum_{\kappa=0}^{N-i}\binom{N}{\kappa}\frac{(N+i)!}{(\kappa+i)!}x^{\kappa+i}\frac{(N-i)!}{(N-i-\kappa)!}(1+x)^{N-i-\kappa}
\end{aligned}$$

and conclude that

$$\begin{aligned}
&f(-1)=\\
&=\sum_{\kappa=0}^{N-i}\binom{N}{\kappa}\frac{(N+i)!}{(\kappa+i)!}(-1)^{\kappa+i}\frac{(N-i)!}{(N-i-\kappa)!}0^{N-i-\kappa}=\\
&=\binom{N}{N-i}\frac{(N+i)!}{N!}(-1)^N(N-i)!=\frac{(N+i)!}{i!}(-1)^N
\end{aligned}$$

which is the desired result.

2) We show: $\sum_{k=j}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!}=$

$$=\sum_{k=i}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!}=\frac{(N+i-j)!\,(-1)^{N+i}}{(N-j)!(N-i)!\,i!} .$$

Obviously, it is equivalent to show that

$$\sum_{k=j}^{N}\binom{N-i}{k-i}\frac{(N+k-j)!}{(k-j)!}(-1)^{k-j}=\frac{N!(N+i-j)!\,(-1)^{N-j}}{(N-j)!\,i!} .$$

Similar to the steps in proof 1 from above, we define

$$f(x)=\sum_{k=j}^{N}\binom{N-i}{k-i}\frac{(N+k-j)!}{(k-j)!}\,x^{k-j}$$

and show that

$$f(-1)=\frac{N!(N+i-j)!\,(-1)^{N-j}}{(N-j)!\,i!} .$$

From the definition we may readily obtain

$$\begin{aligned}
f(x)&=\sum_{k=j}^{N}\binom{N-i}{k-i}(N+k-j)\cdots(1+k-j)\,x^{k-j}=\\
&=\sum_{k=j}^{N}\binom{N-i}{k-i}D_N\,x^{N+k-j}=D_N\sum_{k=j}^{N}\binom{N-i}{k-i}x^{N+k-j}=\\
&=D_N\sum_{\bar k=j-i}^{\bar N}\binom{\bar N}{\bar k}\,x^{\bar N+\bar k+2i-j}=\\
&=D_N\left(\sum_{\bar k=0}^{\bar N}\binom{\bar N}{\bar k}x^{\bar N+\bar k+2i-j}-\sum_{\bar k=0}^{j-i-1}\binom{\bar N}{\bar k}x^{\bar N+\bar k+2i-j}\right)=\\
&=D_N\left(x^{N+i-j}\sum_{\bar k=0}^{\bar N}\binom{\bar N}{\bar k}\,x^{\bar k}\right)-\\
&\underbrace{\sum_{\bar k=0}^{j-i-1}\binom{\bar N}{\bar k}D_N x^{\bar k+N+i-j}}_{=0}=D_N\left(x^{N+i-j}(1+x)^{N-i}\right)
\end{aligned}$$

in which the under-braced expression in the last line is zero because $D_N x^{\bar k|_{\max}+N+i-j}=D_N x^{N-1}=0$.

A direct consequence is that

$$\sum_{k=j}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!}=\sum_{k=i}^{N}\binom{N+k-j}{k-j}\frac{(-1)^{k-i}}{(k-i)!(N-k)!} .$$

Furthermore, we obtain

$$\begin{aligned}
&f(x)=D_N\left(x^{N+i-j}(1+x)^{N-i}\right)=\\
&=\sum_{\kappa=0}^{N}\binom{N}{\kappa}\left(D_{N-\kappa}x^{N+i-j}\right)\left(D_\kappa(1+x)^{N-i}\right)=\\
&=\sum_{\kappa=0}^{N-i}\binom{N}{\kappa}\frac{(N+i-j)!}{(\kappa+i-j)!}x^{\kappa+i-j}\frac{(N-i)!}{(N-i-\kappa)!}(1+x)^{N-i-\kappa}
\end{aligned}$$

and conclude

$$\begin{aligned}
&f(-1)=\\
&=\sum_{\kappa=0}^{N-i}\binom{N}{\kappa}\frac{(N+i-j)!}{(\kappa+i-j)!}(-1)^{\kappa+i-j}\frac{(N-i)!}{(N-i-\kappa)!}0^{N-i-\kappa}=\\
&=\binom{N}{N-i}\frac{(N+i-j)!}{(N-j)!}(-1)^{N-j}(N-i)!=\\
&=\frac{N!(N+i-j)!\,(-1)^{N-j}}{(N-j)!\,i!} .
\end{aligned}$$