Transient Error Reduction in High Gain Observers via Direct State Reconstruction

Florian Meiners[1] and Jürgen Adamy[1]

***Abstract*— As a consequence of unknown initial states or unmodeled disturbances, state observers suffer from a transient phase, during which their estimate converges to the true system state. We propose an approach to mitigating this effect in high gain observers, via an algebraic reconstruction of the system state calculated from the derivatives of the input and output. Although this direct estimate is unreliable due to noise, we show that it significantly reduces the transient observation error. We prove that the proposed modified estimators can be tuned to yield asymptotically stable error dynamics in a stochastic sense. The results are illustrated in simulation.**

I. Introduction

Using output error injection to stabilize observer error dynamics is the de-facto standard approach to state estimation [1]–[6]. Since the invention of the Luenberger observer and Kalman filter, the methods have evolved substantially, and led to the extension of the theory to nonlinear systems. High gain observers, which have been applied in nonlinear state estimation for well over 30 years, represent one of the most prominent types of such estimators. Their success in a number of applications (see [4], [7] and [8], for examples) and the persistent interest in them in control theory research (see [3] and [9] for an overview and references) led to a number of improvements over the original ideas in [5], [6]. For example, [10] suggests a dynamic-gain design. It is a surrogate for the extension of high gain observers to a wider variety of systems by loosening the Lipschitz requirement (see, for instance [11]–[13]). Most notably, modified high gain observers seek to mitigate the so-called peaking phenomenon [14]. In [15] and [16], augmented observers are designed, which are shown to reduce the required gain. Similarly, in [17], a lowering of the gain is achieved by combining high gain design methods with an LMI approach.

Like all observers stabilized by output error feedback, however, high gain designs exhibit a settling time, during which the estimation error is usually prohibitively large. To the best of the authors' knowledge, actively fighting this transient error is a novel idea. For linear observers, combining an algebraic reconstruction of the state with conventional output error injection has already been investigated [18], [19]. Interestingly, the method proposed in [18] lends itself to an application to large classes of nonlinear systems as well. This is because most triangular canonical forms (cf. [6], [9]), for which high gain observers can be designed, are *required* to yield an injective relationship between the state and successive time deriviatives of the output. Locally, this is made precise by the notion of differential observability (see [9]). We leverage this fact to directly calculate the system state from the input, output, and their respective derivatives. Such an idea has already been proposed in the form of so-called Newton observers [20], [21]. Naturally, algebraic reconstructions of the state, by themselves, are unsatisfactory due to the inevitable error in the derivative estimates. In tandem with a conventional observer, however, they can be shown to strongly improve the transient behavior.

We propose simple extensions to two existing high gain observers. The idea is to smoothly fade between a direct reconstruction of the state and the estimate of a classical observer, pairing the immediacy of the direct calculation with the stability guarantees and precision of the state observer. As soon as a rough estimate of the state brings the observation error close enough to zero, the classical observer takes over to complete the convergence. Stability of the error dynamics is shown in the stochastic framework. Crucially, a direct reconstruction of the state from the input and output derivatives is possible under relatively mild assumptions regarding observability of the considered canonical forms. Hence, the ideas presented here are applicable to other high gain observer structures. This holds, in particular, for modified designs (see [15]–[17]) that reduce impulsive error peaking. The simplicity of the proposed estimators facilitates tuning.

II. Preliminaries

Throughout, v_i denotes the i^{th} element of a vector $\mathbf{v}$ and $M_{i,j}$ denotes the entry in the i^{th} row and j^{th} column of a matrix $\mathbf{M}$. For a vector $\mathbf{x} \in \mathbb{R}^n$, we define $\mathbf{x}_i \triangleq [x_1, \ldots, x_i]^\top$. The Euclidean norm of a vector $\mathbf{v} \in \mathbb{R}^n$ and the spectral norm of a matrix $\mathbf{M} \in \mathbb{R}^{n\times p}$ are denoted by $\|\mathbf{v}\|$ and $\|\mathbf{M}\|$, respectively.

A. A High Gain Observer for the Lower Triangular Form

Let $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n$ and $u \in \mathcal{U} \subseteq \mathbb{R}$ hold for the state and the control variable, respectively. Consider systems $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u)$ of the form

$$\begin{aligned} \dot{x}_i &= f_i(\mathbf{x}_{i+1}, u),\ i = 1, \ldots, n-1, \\ \dot{x}_n &= f_n(\mathbf{x}, u), \\ y &= h(\mathbf{x}_1, u). \end{aligned} \tag{1}$$

[1]The authors are with the Control Methods and Intelligent Systems Lab of the Institute of Automation, Electrical Engineering and Information Technology, TU Darmstadt, Landgraf-Georg Str. 4, 64283 Darmstadt, Germany {florian.meiners, juergen.adamy}@tu-darmstadt.de

We assume that the functions f_i and h are continuously differentiable w.r.t. the state sufficiently many times, and that

$$0 < a \leq \left|\frac{\partial h}{\partial x_1}\right| \leq b,\ a \leq \left|\frac{\partial f_i}{\partial x_{i+1}}\right| \leq b,\ \ i = 1, \dots, n-1, \tag{2}$$

holds for all $\mathbf{x} \in \mathcal{X}$ and all $u \in \mathcal{U}$. All f_i are assumed to be Lipschitz with constant $\frac{L}{\sqrt{n}}$. According to [6],

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, u) + \mathbf{\Theta}\mathbf{l}(y - \hat{y}) \tag{3}$$

is an observer for systems in the lower triangular form of (1), where $\hat{y} = h(\hat{\mathbf{x}}, u)$, $\mathbf{l} \in \mathbb{R}^n$ and $\mathbf{\Theta} = \mathrm{diag}(\theta, \theta^2, \dots, \theta^n)$. The parameter $\theta \in \mathbb{R}^+$ is chosen large enough to guarantee stability of the estimation error. Tuning of such an observer is discussed in Lemma 2.1 and Theorem 2.2 of [6]. Central to the functioning of the observers proposed in Section III is the possibility of directly calculating the state from the output and its successive derivatives via a map $\mathbf{\Phi} : \mathbb{R}^n \to \mathbb{R}^n$, which is parameterized by the input and its derivatives. Applying the chain rule to the time derivatives of the output yields

$$\begin{aligned}
y &= h(x_1, u),\\
\dot{y} &= \frac{\partial h}{\partial x_1} f_1 + \frac{\partial h}{\partial u}\dot{u},\\
\ddot{y} &= \frac{\partial^2 h}{\partial x_1^2} f_1^2 + 2\frac{\partial^2 h}{\partial u \partial x_1} f_1 \dot{u} + \frac{\partial h}{\partial x_1}\left(\frac{\partial f_1}{\partial x_1} f_1 + \frac{\partial f_1}{\partial u}\dot{u}\right)\\
&\quad \dots + \frac{\partial^2 h}{\partial u^2}\dot{u}^2 + \frac{\partial h}{\partial x_1}\frac{\partial f_1}{\partial x_2} f_2, \quad \text{etc.}
\end{aligned} \tag{4}$$

Note that the assumption of bounded derivatives in (2) guarantees that (4) can be solved for x_i iteratively, and that the resulting relation is unambiguous. For the system (1), we therefore obtain a map $\mathbf{\Phi} : (y, \dots, y^{(n-1)}, u, \dots, u^{(n-1)}) \mapsto \mathbf{x}$. This is a stronger property than differential observability (see [9]). Though always possible under the assumptions above, the left-inversion of (4) can be cumbersome for systems of higher order. Note, however, that large classes of systems of practical interest allow for a simplification of the computation of $\mathbf{\Phi}$. For example, all expressions containing $u, \dot{u}, \dots$ vanish if the system has full relative degree. The existence of a map $\mathbf{\Phi}$ is also guaranteed if y is a differentially flat output.

B. A Dynamic High Gain Observer for the State-Affine Form

Again, let $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^n$ and $u \in \mathcal{U} \subseteq \mathbb{R}$, and consider systems $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, u) + \mathbf{A}\mathbf{x}$ of the form

$$\begin{aligned}
\dot{x}_i &= g_i(\mathbf{x}_i, u) + x_{i+1},\ i = 1, \dots, n-1,\\
\dot{x}_n &= g_n(\mathbf{x}, u),\\
y &= x_1 = \mathbf{c}^\top \mathbf{x},
\end{aligned} \tag{5}$$

as a special case of the lower triangular form (1). Note that systems of this type encompass systems in nonlinear controller canonical form given in [22]. To loosen the requirement of Lipschitzness of the system dynamics, [10] proposes the assumption that there is a $\mathcal{C}^{n-1}$ function $\gamma : \mathbb{R} \to \mathbb{R}^+$, such that

$$\begin{aligned}
&|g_i(x_1, x_2 + \xi_2, \dots, x_i + \xi_i, u) - g_i(x_1, x_2, \dots, x_i, u)|\\
&\quad \leq \gamma(y)\left(|\xi_2| + \dots + |\xi_i|\right)
\end{aligned} \tag{6}$$

holds for all i, $\mathbf{x} \in \mathcal{X}$ and $u \in \mathcal{U}$. As shown in [10],

$$\dot{\hat{\mathbf{x}}} = \mathbf{g}(\hat{\mathbf{x}}, u) + \mathbf{A}\hat{\mathbf{x}} + \mathbf{\Theta}(\theta)\mathbf{l}(y - \hat{x}_1),\ \ \dot{\theta} = l(\theta, y), \tag{7}$$

is an observer for (5), where $\mathbf{l} \in \mathbb{R}^n$ is chosen to stabilize $\mathbf{A} - \mathbf{l}\mathbf{c}^\top$ and $\mathbf{\Theta}(\theta)$ is defined as in the previous section. The dynamic gain scaling variable θ is subject to

$$\dot{\theta} = l(\theta, y) = -\frac{1}{\tilde{b}}\theta\left(\frac{\tilde{a}}{3}(\theta - 1) - \frac{2(n-1)}{\sqrt{\tilde{q}}}\gamma(y)\right). \tag{8}$$

Therein, $\tilde{a}, \tilde{b}, \tilde{q} \in \mathbb{R}^+$ are chosen such that they satisfy

$$\begin{aligned}
&\mathbf{Q}(\mathbf{A} - \mathbf{l}\mathbf{c}^\top) + (\mathbf{A} - \mathbf{l}\mathbf{c}^\top)^\top\mathbf{Q} \preceq -\tilde{a}\mathbf{Q}, \quad \tilde{q}\mathbf{I} \preceq \mathbf{Q} \preceq \mathbf{I},\\
&\text{and} \quad -\tilde{b}\mathbf{Q} \preceq \mathbf{Q}\mathbf{\Delta} + \mathbf{\Delta}\mathbf{Q} \preceq \tilde{b}\mathbf{Q},
\end{aligned} \tag{9}$$

with $\mathbf{\Delta} = \mathrm{diag}(0, \dots, n-1)$. As stated in [10], this is always possible. Since $\mathbf{A}$ is in companion form, determining the map $\mathbf{\Phi} : (y, \dots, y^{(n-1)}, u, \dots, u^{(n-1)}) \mapsto \mathbf{x}$ is trivially possible. One only needs to require that the functions g_i be continuously differentiable sufficiently many times (cf. (4)).

C. Stochastic Systems

While determining a map $\mathbf{\Phi}$ is always possible for systems (1) or (5), we cannot assume that the required derivatives $\dot{y}, \dots, y^{(n-1)}$ and $\dot{u}, \dots, u^{(n-1)}$ are known without error. Hence why we respect an additive error around the true value of the state $\mathbf{x}$ whenever a direct reconstruction $\mathbf{x}^{\mathrm{d}}$ is calculated from estimates $\tilde{y}^{(i)}$ and $\tilde{u}^{(i)}$ of $y^{(i)}$ and $u^{(i)}$, respectively. As in the linear case (see [18]), we assume that

$$\mathbf{x}^{\mathrm{d}} = \mathbf{\Phi}\left(y, \dot{\tilde{y}}, \dots, \tilde{y}^{(n-1)}, u, \dot{\tilde{u}}, \dots, \tilde{u}^{(n-1)}\right) = \mathbf{x} + \mathbf{n} \tag{10}$$

holds for the direct reconstruction of the state. Therein, the entries of $\mathbf{n} = [n_1, \dots, n_n]^\top$ are modeled as random walks with independent identically distributed increments of zero mean and finite variance. This random walk approaches a Brownian motion $\mathbf{B}$ as the time between increments approaches zero by Skorokhod's embedding thereom (see [23]). Thus follows the form $\mathbf{x}^{\mathrm{d}} = \mathbf{x} + \mathbf{W}\mathbf{B}$ for the direct estimate, where $\mathbf{W} \in \mathbb{R}^{n \times n}$ is a scaling matrix. Since all subsequent error equations are in the form of stochastic systems, we revise the required stability properties. Consider the following system in differential form:

$$\mathrm{d}\mathbf{x} = \mathbf{f}(\mathbf{x}, t)\mathrm{d}t + \mathbf{G}(\mathbf{x}, t)\mathrm{d}\mathbf{B}, \quad \mathbf{G} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n \times p}. \tag{11}$$

It is disturbed by a p-dim. Brownian motion $\mathbf{B}$, as defined in [24]. The equation (11) is the standard form of a system with random input (see [25], [26]). Details can be found in [27]. For the stability analysis of (11), the differential operator

$$\mathcal{L} = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbf{G}\mathbf{G}^\top\right)_{i,j}\frac{\partial^2}{\partial x_i \partial x_j} \tag{12}$$

is applied to a Lyapunov function candidate [26] to assess the following stability property [28]:

Definition 1 *The zero solution* $\mathbf{x}(t)=\mathbf{0},\ t>0$ *of Eq.* (11), *is said to be stable with probability* $\mathrm{P}(\cdot)$ *one if, for any* $\rho\in(0,1]$ *and any* $\varepsilon>0$, *there is a* $\delta(\rho,\epsilon)$ *such that*

$$\mathrm{P}\left(\sup_{t>0}|\mathbf{x}(t)|\geq\varepsilon\right)\leq\rho$$

for any initial state $|\mathbf{x}_0|\leq\delta$. *If, in addition to that,*

$$\mathrm{P}\left(\lim_{t\to\infty}\mathbf{x}(t)=\mathbf{0}\right)=1$$

holds for any $\mathbf{x}_0\in\mathbb{R}^n$, *the zero solution is called globally asymptotically stable with probability one.*

III. Using Direct State Reconstruction in Observers

The method we propose to reduce the transient error in high gain observers consists of a continuous blending of the nominal observer and the direct reconstruction $\mathbf{x}^{\mathrm{d}}$ from (10). For (1), we propose the following modification of (3):

$$\begin{aligned}&\dot{\hat{\mathbf{x}}}=\mathbf{f}(\hat{\mathbf{x}},u)+\alpha\mathbf{\Theta}\mathbf{l}(y-\hat{y})+(1-\alpha)\mathbf{\Gamma}(\mathbf{x}^{\mathrm{d}}-\hat{\mathbf{x}}),\\&\text{where}\quad\alpha(y-\hat{y})=\exp\left(-\mu\left(y-\hat{y}\right)^2\right),\end{aligned}\tag{13}$$

with a feedback matrix $\mathbf{\Gamma}\in\mathbb{R}^{n\times n}$ and a tuning parameter $\mu\in\mathbb{R}^+$. For brevity, and to keep the presentation clean, the argument of α is suppressed throughout. We obtain

Theorem 1 *Let* $\mathbf{l}$ *and* θ *be chosen such that* (3) *is an observer for* (1) *with asymptotically stable error dynamics. Then, there is a matrix* $\mathbf{\Gamma}$ *and a* μ *that make* (13) *an observer for* (1) *whose error dynamics are asymptotically stable in the sense of Definition 1.*

Proof: Given the error $\mathbf{e}=\mathbf{x}-\hat{\mathbf{x}}$ and recalling $\mathbf{x}^{\mathrm{d}}=\mathbf{x}+\mathbf{W}\mathbf{B}$, the error dynamics for the observer (13) can be written as the stochastic differential equation (see (11))

$$\begin{aligned}\mathrm{d}\mathbf{e}=&\left[\mathbf{f}(\mathbf{x},u)-\mathbf{f}(\hat{\mathbf{x}},u)-\alpha\mathbf{\Theta}\mathbf{l}(y-\hat{y})-(1-\alpha)\mathbf{\Gamma}\mathbf{e}\right]\mathrm{d}t\\&-(1-\alpha)\mathbf{\Gamma}\mathbf{W}\mathrm{d}\mathbf{B}.\end{aligned}\tag{14}$$

Recall the definition of $\mathbf{x}_i$ in Section II. Using the mean value theorem, the nonlinear term can be rewritten as

$$\begin{aligned}&f_i(\mathbf{x},u)-f_i(\hat{\mathbf{x}},u)=f_i(\mathbf{x}_i,x_{i+1},u)-f_i(\hat{\mathbf{x}}_i,\hat{x}_{i+1},u)\\&=f_i(\mathbf{x}_i,x_{i+1},u)-f_i(\mathbf{x}_i,\hat{x}_{i+1},u)+f_i(\mathbf{x}_i,\hat{x}_{i+1},u)\\&-f_i(\hat{\mathbf{x}}_i,\hat{x}_{i+1},u)\\&=\underbrace{f_i(\mathbf{x}_i,\hat{x}_{i+1},u)-f_i(\hat{\mathbf{x}}_i,\hat{x}_{i+1},u)}_{\triangleq\hat{f}_i}+\underbrace{\frac{\partial f_i(\mathbf{x}_i,\delta_i,u)}{\partial x_{i+1}}}_{\triangleq\tilde{f}_{i+1}}e_{i+1}\end{aligned}$$

for some δ_i and $i=1,\dots,n-1$. This is the key to Lemma 2.1 and Theorem 2.2 in [6]. Taking $\hat{f}_n=f_n(\mathbf{x},u)-f_n(\hat{\mathbf{x}},u)$ and $\tilde{f}_1=\frac{\partial h}{\partial x_1}(x_1,\delta_0,u)$, we can define constant matrices $\tilde{\mathbf{c}}^\top\triangleq\left[\tilde{f}_1,0,\dots,0\right]$,

$$\tilde{\mathbf{A}}\triangleq\begin{bmatrix}0&\tilde{f}_2&\dots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\dots&\tilde{f}_n\\0&0&\dots&0\end{bmatrix},\quad\text{and}\quad\tilde{\mathbf{F}}\triangleq\begin{bmatrix}\hat{f}_1\\\vdots\\\hat{f}_n\end{bmatrix}.$$

Thereby, the error equation can be reformulated as

$$\begin{aligned}\mathrm{d}\mathbf{e}=&\left(\left[\tilde{\mathbf{A}}-\alpha\mathbf{\Theta}\mathbf{l}\tilde{\mathbf{c}}^\top-(1-\alpha)\mathbf{\Gamma}\right]\mathbf{e}+\tilde{\mathbf{F}}\right)\mathrm{d}t\\&+(1-\alpha)\mathbf{\Gamma}\mathbf{W}\mathrm{d}\mathbf{B}.\end{aligned}$$

Using $\boldsymbol{\varepsilon}\triangleq\mathbf{\Theta}^{-1}\mathbf{e}$ to scale the error as in [6] we obtain

$$\begin{aligned}\mathrm{d}\boldsymbol{\varepsilon}=&\left(\left[\theta\left(\tilde{\mathbf{A}}-\alpha\mathbf{l}\tilde{\mathbf{c}}^\top\right)-(1-\alpha)\mathbf{\Theta}^{-1}\mathbf{\Gamma}\mathbf{\Theta}\right]\boldsymbol{\varepsilon}+\mathbf{\Theta}^{-1}\tilde{\mathbf{F}}\right)\mathrm{d}t\\&+(1-\alpha)^2\mathbf{\Theta}^{-1}\mathbf{\Gamma}\mathbf{W}\mathrm{d}\mathbf{B}.\end{aligned}$$

Observe that, by the assumption of stable tuning of the nominal observer in [6], $\mathbf{l}$ is chosen to fulfill

$$\left(\tilde{\mathbf{A}}-\mathbf{l}\tilde{\mathbf{c}}^\top\right)^\top\mathbf{S}+\mathbf{S}\left(\tilde{\mathbf{A}}-\mathbf{l}\tilde{\mathbf{c}}^\top\right)\preceq-\lambda\mathbf{I}\tag{15}$$

for some $\mathbf{S}\succ0$ and some $\lambda>0$. Consider the Lyapunov function $V(\boldsymbol{\varepsilon})=\boldsymbol{\varepsilon}^\top\mathbf{S}\boldsymbol{\varepsilon}$. Applying (12) to V yields

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon})=&\,2\alpha\theta\boldsymbol{\varepsilon}^\top\mathbf{S}\left(\tilde{\mathbf{A}}-\mathbf{l}\tilde{\mathbf{c}}^\top\right)\boldsymbol{\varepsilon}+2\boldsymbol{\varepsilon}^\top\mathbf{S}\mathbf{\Theta}^{-1}\tilde{\mathbf{F}}\\&+2(1-\alpha)\theta\boldsymbol{\varepsilon}^\top\mathbf{S}\left(\tilde{\mathbf{A}}-\frac{\mathbf{\Theta}^{-1}\mathbf{\Gamma}\mathbf{\Theta}}{\theta}\right)\boldsymbol{\varepsilon}\\&+(1-\alpha)^2\underbrace{\operatorname{trace}\left(\mathbf{S}\mathbf{\Theta}^{-1}\mathbf{\Gamma}\mathbf{W}\mathbf{W}^\top\mathbf{\Gamma}^\top\mathbf{\Theta}^{-1}\right)}_{\triangleq c}.\end{aligned}\tag{16}$$

Since all f_i are Lipschitz with constant $\frac{L}{\sqrt{n}}$, we get

$$\begin{aligned}&\|\mathbf{\Theta}^{-1}\tilde{\mathbf{F}}\|=\sqrt{(\theta^{-1}\hat{f}_i)^2+\dots+(\theta^{-n}\hat{f}_n)^2}\\&\leq\sqrt{\left(\theta^{-1}\frac{L}{\sqrt{n}}|e_1|\right)^2+\dots+\left(\theta^{-n}\frac{L}{\sqrt{n}}\|\mathbf{e}\|\right)^2}\\&\leq\sqrt{\left(\frac{L}{\sqrt{n}}|\varepsilon_1|\right)^2+\dots+\left(\frac{L}{\sqrt{n}}\|\boldsymbol{\varepsilon}\|\right)^2}\leq\sqrt{n\left(\frac{L}{\sqrt{n}}\|\boldsymbol{\varepsilon}\|\right)^2}\\&\leq L\|\boldsymbol{\varepsilon}\|\end{aligned}$$

for $\theta>1$. We therefore obtain

$$\begin{aligned}\mathcal{L}V\leq&\left(2\|\mathbf{S}\|L-\alpha\theta\lambda\right)\|\boldsymbol{\varepsilon}\|^2\\&+2(1-\alpha)\theta\boldsymbol{\varepsilon}^\top\mathbf{S}\left(\tilde{\mathbf{A}}-\frac{\mathbf{\Theta}^{-1}\mathbf{\Gamma}\mathbf{\Theta}}{\theta}\right)\boldsymbol{\varepsilon}+(1-\alpha)^2c.\end{aligned}$$

By choosing $\mathbf{\Gamma}=\mathbf{\Theta}\mathbf{l}\left[\frac{\kappa}{2},\dots,0\right]+\frac{\phi}{2}\mathbf{I}$ with $a\leq\kappa\leq b$ and a positive tuning parameter ϕ, we can conclude that

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon})\leq&\left(2\|\mathbf{S}\|L-\alpha\theta\lambda-(1-\alpha)\theta\lambda-(1-\alpha)\underline{\sigma}(\mathbf{S})\phi\right)\|\boldsymbol{\varepsilon}\|^2\\&+(1-\alpha)^2c=-k\|\boldsymbol{\varepsilon}\|^2+(1-\alpha)^2c\end{aligned}\tag{17}$$

for some $k>0$, where $\underline{\sigma}(\mathbf{S})$ is the smallest singular value of $\mathbf{S}$. This is because $\tilde{\mathbf{A}}-\mathbf{l}\left[\kappa,\dots,0\right]$ is Hurwitz for $a\leq\kappa\leq b$ by assumption. According to Theorem 3 in [29], the zero solution of (14) is asymptotically stable in the sense of Definition 1, if there is a positive definite function $V(\boldsymbol{\varepsilon})$ with $V(\mathbf{0})=0$ such that $\mathcal{L}V(\boldsymbol{\varepsilon})\leq-\tilde{\kappa}(\boldsymbol{\varepsilon})\leq0$ for some function $\tilde{\kappa}:\mathbb{R}^n\to\mathbb{R}_0^+$ with zero set $\mathbf{0}\in\mathbb{R}^n$. From the bounds on $\frac{\partial h}{\partial x_1}$ it can be concluded that h is Lipschitz with some constant L_h. Note that α decreases with increasing difference ε_1 because of the monotonicity of h. Critical values of the error occur for $\boldsymbol{\varepsilon}=\tau\left[1,0,\dots,0\right]^\top$, $\tau\in\mathbb{R}$, because errors with entries $\varepsilon_i \neq 0, i = 2, \ldots, n$ lower the quadratic term without affecting α. We get the requirement that

$$(1-\alpha(h(x_1,u)-h(\hat{x}_1,u)))^2 c \leq \left(1-\exp(-\mu L_h^2 |e_1|^2)\right)^2 c$$
$$= \left(1-\exp(-\mu L_h^2 |\theta\varepsilon_1|^2)\right)^2 c$$
$$= \left(1-\exp(-\mu L_h^2 \theta\tau^2)\right)^2 c < k\tau^2.$$

This inequality is fulfilled, in particular, if

$$\frac{\partial}{\partial\tau}\left(k\tau^2 - \left(1-\exp\left(-\mu L_h^2\theta\tau^2\right)\right)^2\right) \neq 0 \Leftrightarrow$$
$$\frac{k}{2L_h^2\theta c} \neq \mu\left(1-\exp\left(-\mu L_h^2\theta\tau^2\right)\right)\exp\left(-\mu L_h^2\theta\tau^2\right) \leq \frac{\mu}{4}.$$

Thus, if $\mu < 2k/(L_h^2\theta c)$, the quadratic term in (17) indeed outweighs $(1-\alpha)^2 c$. We therefore know that $\mathcal{L}V(\boldsymbol{\varepsilon}) < -\tilde{k}\|\boldsymbol{\varepsilon}\|^2$ for some $\tilde{k} > 0$. Thereby, Theorem 3 in [29] is applicable. ■

For the state-affine form in (5), consider the modification

$$\begin{aligned}\dot{\hat{\mathbf{x}}} &= \mathbf{g}(\hat{\mathbf{x}},u) + \mathbf{A}\hat{\mathbf{x}} + \alpha\boldsymbol{\Theta}(\theta)\mathbf{l}(y-\hat{x}_1)\\ &\quad + (1-\alpha)\boldsymbol{\Gamma}(\theta)(\mathbf{x}^{\mathrm{d}}-\hat{\mathbf{x}}),\ \dot{\theta} = \tilde{l}(\theta,y),\end{aligned} \tag{18}$$

of (7) with α as defined in (13). Likewise, we obtain

Theorem 2 *Let the gain* $\mathbf{l}$*, as well as the constants* $\tilde{a}, \tilde{b}, \tilde{q}$ *in* (9)*, be chosen such that* (7) *is an observer for* (5) *with asymptotically stable error dynamics. Then, there is a matrix* $\boldsymbol{\Gamma}(\theta)$ *and functions* $\tilde{l}(\theta,y), \mu(\theta)$ *that make* (18) *an observer for* (5) *whose error dynamics are asymptotically stable in the sense of Definition 1.*

Proof: We transform the error $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ via $\varepsilon_i = e_i\theta^{1-i-\tilde{b}}$ as in [10]. Applying the chain rule yields

$$\dot{\boldsymbol{\varepsilon}} = \frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\dot{\mathbf{e}} - \frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}}}\left(\boldsymbol{\Delta} + \tilde{b}\mathbf{I}\right)\dot{\theta}\mathbf{e},$$

with $\boldsymbol{\Delta}$ as defined in (9). The dependence of $\boldsymbol{\Theta}$ on θ is suppressed for brevity. Recall $\mathbf{x}^{\mathrm{d}} = \mathbf{x} + \mathbf{W}\mathbf{B}$. As in the proof of Theorem 1, this leads to the stochastic error dynamics

$$\begin{aligned}\mathrm{d}\boldsymbol{\varepsilon} = \Big[&\alpha\theta\left(\mathbf{A}-\mathbf{l}\mathbf{c}^\top\right)\boldsymbol{\varepsilon} + (1-\alpha)\theta\left(\mathbf{A} - \frac{\boldsymbol{\Theta}^{-1}\boldsymbol{\Gamma}(\theta)\boldsymbol{\Theta}}{\theta}\right)\boldsymbol{\varepsilon}\\ &+ \frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\left(\mathbf{g}(\mathbf{x},u) - \mathbf{g}(x_1,\ldots,x_n - \theta^{n-1+\tilde{b}}\varepsilon_n,u)\right)\\ &- \left(\boldsymbol{\Delta}+\tilde{b}\mathbf{I}\right)\frac{\dot{\theta}}{\theta}\boldsymbol{\varepsilon}\Big]\mathrm{d}t + (1-\alpha)\,\frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\boldsymbol{\Gamma}(\theta)\mathbf{W}\mathrm{d}\mathbf{B}.\end{aligned}$$

Consider $V(\boldsymbol{\varepsilon}) = \boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon}$. Applying (12), we obtain

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon}) &= 2\alpha\theta\boldsymbol{\varepsilon}^\top\mathbf{Q}\left(\mathbf{A}-\mathbf{l}\mathbf{c}^\top\right)\boldsymbol{\varepsilon} - 2\boldsymbol{\varepsilon}^\top\mathbf{Q}\left(\boldsymbol{\Delta}+\tilde{b}\mathbf{I}\right)\frac{\dot{\theta}}{\theta}\boldsymbol{\varepsilon}\\ &+ 2(1-\alpha)\theta\boldsymbol{\varepsilon}^\top\mathbf{Q}\left(\mathbf{A} - \frac{\boldsymbol{\Theta}^{-1}\boldsymbol{\Gamma}(\theta)\boldsymbol{\Theta}}{\theta}\right)\boldsymbol{\varepsilon}\\ &+ 2\boldsymbol{\varepsilon}^\top\mathbf{Q}\frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\left(\mathbf{g}(\mathbf{x},u) - \mathbf{g}(x_1,\ldots,x_n-\theta^{n-1+\tilde{b}}\varepsilon_n,u)\right)\\ &+ (1-\alpha)^2\underbrace{\mathrm{trace}\left(\mathbf{Q}\frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\boldsymbol{\Gamma}(\theta)\mathbf{W}\mathbf{W}^\top\boldsymbol{\Gamma}^\top(\theta)\frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\right)}_{\triangleq c(\theta)}.\end{aligned}$$

Observe that the nonlinearity is bounded as

$$\begin{aligned}&\left|\boldsymbol{\varepsilon}^\top\mathbf{Q}\frac{\boldsymbol{\Theta}^{-1}}{\theta^{\tilde{b}-1}}\left(\mathbf{g}(\mathbf{x},u) - \mathbf{g}(x_1,\ldots,x_n-\theta^{n-1+\tilde{b}}\varepsilon_n,u)\right)\right|\\ &\leq \sum_{i=1}^n \left|\left(\boldsymbol{\varepsilon}^\top\mathbf{Q}\right)_i\right|\left|\frac{g_i(\mathbf{x},u) - g_i(x_1,\ldots,x_n-\theta^{n-1+\tilde{b}}\varepsilon_n,u)}{\theta^{i-1+\tilde{b}}}\right|\\ &\leq \sum_{i=1}^n \left|\left(\boldsymbol{\varepsilon}^\top\mathbf{Q}\right)_i\right|\frac{\gamma(y)(\theta^{1+\tilde{b}}|\varepsilon_2| + \cdots + \theta^{i-1+\tilde{b}}|\varepsilon_i|)}{\theta^{i-1+\tilde{b}}}\\ &\leq \gamma(y)\sum_{i=1}^n \left|\left(\boldsymbol{\varepsilon}^\top\mathbf{Q}\right)_i\right|(|\varepsilon_2| + \cdots + |\varepsilon_i|)\\ &\leq \gamma(y)(n-1)\left(\left|\boldsymbol{\varepsilon}^\top\mathbf{Q}\right||\boldsymbol{\varepsilon}|\right) \leq \frac{n-1}{\sqrt{\tilde{q}}}\gamma(y)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon}\end{aligned}$$

for $\theta > 1$ due to (6). With $\boldsymbol{\Gamma}(\theta) = \mathbf{A} + \frac{\theta\phi\mathbf{I}}{2}$, $\phi > \tilde{a}$, we get

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon}) &\leq \left(\frac{2(n-1)}{\sqrt{\tilde{q}}}\gamma(y) + \tilde{b}\frac{|\dot{\theta}|}{\theta} - \alpha\theta\tilde{a} - 2\tilde{b}\frac{\dot{\theta}}{\theta}\right)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon}\\ &+ 2(1-\alpha)\theta\boldsymbol{\varepsilon}^\top\mathbf{Q}\left(\mathbf{A} - \frac{\boldsymbol{\Theta}^{-1}\boldsymbol{\Gamma}(\theta)\boldsymbol{\Theta}}{\theta}\right)\boldsymbol{\varepsilon} + (1-\alpha)^2 c(\theta)\\ &\leq \left(\frac{2(n-1)}{\sqrt{\tilde{q}}}\gamma(y) + \tilde{b}\frac{|\dot{\theta}|}{\theta} - \alpha\theta\tilde{a} - 2\tilde{b}\frac{\dot{\theta}}{\theta} - (1-\alpha)\theta\phi\right)\\ &\quad\cdot\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon} + (1-\alpha)^2 c(\theta),\end{aligned}$$

where we have used the inequalities in (9) and, in particular, $-2\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\Delta}\boldsymbol{\varepsilon}\dot{\theta} \leq \tilde{b}\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon}|\dot{\theta}|$. We make a slight adjustment to the dynamics of θ proposed in (8), namely

$$\dot{\theta} = -\frac{\theta}{\tilde{b}}\left(\frac{\alpha\tilde{a} + (1-\alpha)\phi}{3}(\theta-1) - \frac{2(n-1)}{\sqrt{\tilde{q}}}\gamma(y)\right). \tag{19}$$

It is shown in [10] that, if initialized larger than one, θ stays above one. Invoking Theorem 3 in [29] again, we therefore only have to distinguish between positive and negative values of $\dot{\theta}$ to show that $\mathcal{L}V \leq -\tilde{\kappa}(\boldsymbol{\varepsilon}) \leq 0$ for some $\tilde{\kappa} : \mathbb{R}^n \to \mathbb{R}_0^+$ with zero set $\mathbf{0} \in \mathbb{R}^n$. For $\dot{\theta} \geq 0$, we obtain

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon}) &\leq -\frac{2\theta+1}{3}\left(\alpha\tilde{a} + (1-\alpha)\phi\right)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon} + (1-\alpha)^2 c(\theta)\\ &\leq -\left(\alpha\tilde{a} + (1-\alpha)\phi\right)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon} + (1-\alpha)^2 c(\theta).\end{aligned}$$

Likewise, for $\dot{\theta} \leq 0$, we use $\gamma(y) \geq 0$ to arrive at

$$\begin{aligned}\mathcal{L}V(\boldsymbol{\varepsilon}) &\leq -\left(\frac{4(n-1)}{\sqrt{\tilde{q}}}\gamma(y) + (\alpha\tilde{a} + (1-\alpha)\phi)\right)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon}\\ &+ (1-\alpha)^2 c(\theta) \leq -\left(\alpha\tilde{a} + (1-\alpha)\phi\right)\boldsymbol{\varepsilon}^\top\mathbf{Q}\boldsymbol{\varepsilon} + (1-\alpha)^2 c(\theta).\end{aligned}$$

Hence why, for all $\boldsymbol{\varepsilon} \neq \mathbf{0}$ and all $\dot{\theta}$, we have

$$\mathcal{L}V(\boldsymbol{\varepsilon}) < -\tilde{k}\|\boldsymbol{\varepsilon}\|^2 + (1-\alpha)^2 c(\theta)$$

for some constant $\tilde{k} > 0$. As in the proof of Theorem 1, we argue that the critical error vectors occur for $\boldsymbol{\varepsilon} = \tau\,[1,0,\ldots,0]^\top$ and show that, if $\mu(\theta)$ is chosen accordingly,

$$\tilde{k}\|\boldsymbol{\varepsilon}\|^2 > (1-\alpha)^2 c(\theta). \tag{20}$$

Plugging in $\boldsymbol{\varepsilon} = \tau\,[1,0,\ldots,0]^\top$ results in

$$\left(1-\exp\left(\mu(\theta)\theta^{2\tilde{b}}\tau^2\right)\right)^2 c(\theta) < \tilde{k}\tau^2.$$

By the same reasoning as in the proof above, we obtain

$$\frac{\partial}{\partial\tau}\left(\tilde{k}\tau^2 - \left(1-\exp\left(\mu(\theta)\theta^{2\tilde{b}}\tau^2\right)\right)^2 c(\theta)\right) \neq 0$$
$$\Leftrightarrow \frac{\tilde{k}}{2\theta^{2\tilde{b}}c(\theta)} \neq \mu(\theta)\left(1-\exp\left(\mu(\theta)\theta^{2\tilde{b}}\tau^2\right)\right)$$
$$\cdot\exp\left(\mu(\theta)\theta^{2\tilde{b}}\tau^2\right) \leq \frac{\mu(\theta)}{4}.$$

The desired inequality (20) is fulfilled by choosing

$$\mu(\theta) = 2\hat{\mu}\left(\theta^{2\tilde{b}}c(\theta)\right)^{-1} \quad \text{with} \quad \hat{\mu} < \tilde{k},$$

which again shows the applicability of Theorem 3 in [29]. ■

Remark 1 *The bounds on μ and $\mu(\theta)$ found in the proofs of Theorems 1 and 2 directly depend on the values of c and $c(\theta)$, respectively. Recalling the definitions of c and $c(\theta)$, it is obvious that the bound lowers with increasing gain in the feedback of $\mathbf{x}^{\mathrm{d}}$. This highlights the trade-off between speed of convergence due to the direct reconstruction and suppression of noise by to the nominal high gain observer.*

IV. Simulation and Comparison

To illustrate the transient error reduction, we compare the proposed observers to the original high gain designs in the simulation of a bioreactor model. It can be found in a number of publications (e.g., [5], [12]). With the microorganism concentrations η_1 and η_2, the system equations are

$$\dot{\eta}_1 = \frac{3\eta_1\eta_2}{\eta_1+\eta_2} - u\eta_1, \quad \dot{\eta}_2 = -\frac{3\eta_1\eta_2}{\eta_1+\eta_2} + (0.1-\eta_2)u,$$
$$y = \eta_1. \tag{21}$$

As pointed out in [5], a forward invariant set is given by

$$\mathcal{X}(\eta_2, u_{\max}) = \{\eta : \eta_1 > \epsilon_1, \eta_2 > \epsilon_2, 0.1 > \eta_1+\eta_2\},$$
$$\text{with} \quad \epsilon_1 = \frac{(3-u_{\max})\epsilon_2}{u_{\max}} \quad \text{and} \quad u_{\min} = \frac{3\epsilon_2}{0.1-\epsilon_2},$$

while a diffeomorphism of $\mathcal{X}$ onto its image is given by

$$\mathbf{T}(\boldsymbol{\eta}) : \begin{bmatrix}\eta_1\\ \eta_2\end{bmatrix} \mapsto \begin{bmatrix}\eta_1\\ \frac{3\eta_1\eta_2}{\eta_1+\eta_2}\end{bmatrix} \triangleq \begin{bmatrix}x_1\\ x_2\end{bmatrix}.$$

We use this map to transform the system into

$$\dot{x}_1 = x_2 - ux_1 = f_1(x_1, x_2, u),$$
$$\dot{x}_2 = \left(\frac{0.1u}{3x_1^2} + \frac{2}{x_1}\right)x_2^2 - \left(3+u+\frac{0.2u}{x_1}\right)x_2 + 0.3u$$
$$= f_2(x_1, x_2, u), \tag{22}$$

i.e. the triangular form (1). For (22) in the transformed coordinates, the high gain observer

$$\dot{\hat{x}}_1 = f_1(\hat{x}_1, \hat{x}_2, u) + L_1\theta(x_1-\hat{x}_1),$$
$$\dot{\hat{x}}_2 = f_2(\hat{x}_1, \hat{x}_2, u) + L_2\theta^2(x_1-\hat{x}_1), \tag{23}$$

can be designed as described in [6]. The observer gains can be chosen as $L_1 = 1$ and $L_2 = 0.5$, while a scaling with

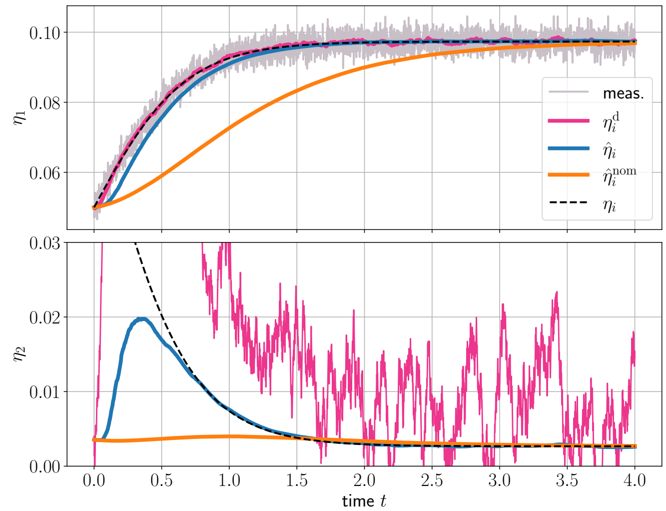

Fig. 1. Simulation of the high gain observer in (3) (orange) versus the observer (13) using the direct reconstruction (blue). The dashed line shows the true state, while the direct reconstruction is given in pink. The measurement is plotted in grey.

$\theta = 1$ is sufficient to achieve stability. For the observer (13) in $\mathbf{x}$-coordinates, we choose the same gains L_1, L_2, and the matrix $\mathbf{\Gamma} = \mathbf{L}\tilde{\mathbf{c}}^\top + 2000\mathbf{I}$ for the feedback of the direct state reconstruction. This choice quickly forces the observer estimate towards the direct estimate during the transient phase. The tuning parameter is set to $\mu = 50$ for a fast transition to the nominal high gain behavior. To caluculate the direct state reconstruction, the output $y = x_1$ and its first derivative $\dot{y} = x_2 - ux_1$ are solved for the state. This results in $x_1 = y$ and $x_2 = \dot{y} + uy$. A sliding mode differentiator (see [30]) is used to obtain estimates $\tilde{y}$ and $\dot{\tilde{y}}$ of y and $\dot{y}$ from a noisy output measurement $y^{\text{meas}} = x_1 + n$. The differentiator

$$\dot{z}_0 = -1.5G^{\frac{1}{2}}\left(|z_0 - y^{\text{meas}}|\right)^{\frac{1}{2}}\operatorname{sign}\left(z_0 - y^{\text{meas}}\right) + z_1 \triangleq v_0$$
$$\dot{z}_1 = -1.1G\operatorname{sign}\left(z_1 - v_0\right),$$

uses the tuning parameter $G = 1$ and is initialized with $z_0(0) = 0.05$, $z_1(0) = 0$. Setting $\tilde{y} = z_0$ and $\dot{\tilde{y}} = z_1$ yields

$$\mathbf{x}^{\mathrm{d}} = \begin{bmatrix}z_0\\ z_1 + uz_0\end{bmatrix}.$$

We simulate the step response of the reactor model for $u = 0.08$. As the initial values, we take $\eta_1(0) = \hat{x}_1(0) = 0.05$, $\eta_2(0) = 0.05$ and $\hat{x}_2(0) = 0.01$ for both observers alike. Since both observers operate in $\mathbf{x}$-coordinates, we apply

$$\hat{\boldsymbol{\eta}} = \mathbf{T}^{-1}(\hat{\mathbf{x}}) = \begin{bmatrix}\hat{x}_1\\ \frac{\hat{x}_1\hat{x}_2}{3\hat{x}_1-\hat{x}_2}\end{bmatrix}$$

to calculate estimates in the original coordinates. Additive measurement noise with an SNR of 30 dB at 1 kHz is assumed. The results of the simulation are shown in Fig. 1. Note that the nominal observer $\hat{\boldsymbol{\eta}}^{\text{nom}}$ follows the true state progression only after a noticeable settling time during which the observation error is large. Note that, despite the large error in the direct reconstruction (pink curve), the proposed observer with state $\hat{\boldsymbol{\eta}}$ exhibits a much smaller transient error.

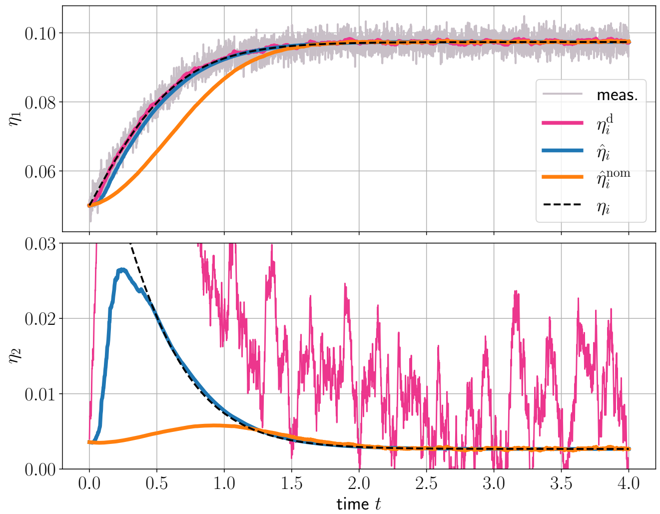

Fig. 2. Simulation of the high gain observer in (7) (orange) versus the observer (18) using the direct reconstruction (blue). The dashed line shows the true state, while the direct reconstruction is given in pink. The measurement is plotted in grey.

As stated in Section II, the state-affine form (5) is a special case of the lower triangular form (1). From the state taking values in $\mathcal{X}$ we can conclude that the nonlinear functions f_i in (22) can be bounded as in (6), say by $\gamma(y) = 4y$. Hence why (21) also allows for the design of the dynamic gain observer in (7). For the observer gain $\mathbf{l} = [2, 1]^\top$, the constants $\tilde{a} = 1.17, \tilde{b} = 2.3, \tilde{q} = 0.15$ fulfill the inequalities (9). The gain for the feedback of $\mathbf{x}^{\mathrm{d}}$ is set to $\boldsymbol{\Gamma} = \tilde{\mathbf{A}} + 2000\theta\mathbf{I}$. This time, however, we choose the dynamically adjusted $\mu(\theta) = 500/\theta^{2\tilde{b}}$. Mind that, for the modified observer, θ obeys (19). The setup is otherwise identical to that of the first simulation. Figure 2 shows the results. Indeed, the error of the nominal observer $\hat{\boldsymbol{\eta}}^{\mathrm{nom}}$ is reduced due to the dynamics of the gain scaling variable θ. The proposed observer $\hat{\boldsymbol{\eta}}$, however, further reduces the transient error compared to both the nominal observers and the design in (13). Moreover, it performs even better in terms of steady-state noise suppression due to the dynamic lowering of the scaling variable θ.

V. Summary

It was shown that the use of a direct state reconstruction calculated from the input and output of a system can be used to reduce the transient error of high gain observers. Two well-known observers were modified so as to leverage a direct reconstruction of the system state calculated from the derivatives of the input and output. Proofs of stability of the proposed observer structures were provided alongside simulations demonstrating the benefits of the modifications.

References

[1] R. Kalman and R. Bucy, “New Results in Linear Filtering and Prediction Theory,” *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.
[2] D. Luenberger, “Observing the state of a linear system,” *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74–80, 1964.
[3] H. K. Khalil and L. Praly, “High-gain observers in nonlinear feedback control,” *International Journal of Robust and Nonlinear Control*, vol. 24, no. 6, pp. 993–1015, 2014.
[4] H. K. Khalil, “High-gain observers in feedback control: Application to permanent magnet synchronous motors,” *IEEE Control Systems Magazine*, vol. 37, no. 3, pp. 25–41, 2017.
[5] J.-P. Gauthier, H. Hammouri, and S. Othman, “A simple observer for nonlinear systems applications to bioreactors,” *IEEE Transactions on automatic control*, vol. 37, no. 6, pp. 875–880, 1992.
[6] J.-P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.
[7] M. Ghanes, J. De Leon, and A. Glumineau, “Cascade and high-gain observers comparison for sensorless closed-loop induction motor control,” *IET Control Theory & Applications*, vol. 2, no. 2, pp. 133–150, 2008.
[8] R. Escobar, C. Astorga-Zaragoza, A. Téllez-Anguiano, D. Juárez-Romero, J. Hernández, and G. Guerrero-Ramírez, “Sensor fault detection and isolation via high-gain observers: Application to a double-pipe heat exchanger,” *ISA transactions*, vol. 50, no. 3, pp. 480–486, 2011.
[9] P. Bernard, *Observer Design for Nonlinear Systems*. Springer, 2019.
[10] L. Praly, “Asymptotic stabilization via output feedback for lower triangular systems with output dependent incremental rate,” *IEEE Transactions on Automatic Control*, vol. 48, no. 6, pp. 1103–1108, 2003.
[11] V. Andrieu, L. Praly, and A. Astolfi, “Homogeneous approximation, recursive observer design, and output feedback,” *SIAM Journal on Control and Optimization*, vol. 47, no. 4, pp. 1814–1850, 2008.
[12] ——, “High gain observers with updated gain and homogeneous correction terms,” *Automatica*, vol. 45, no. 2, pp. 422–428, 2009.
[13] P. Bernard, L. Praly, and V. Andrieu, “Observers for a non-lipschitz triangular form,” *Automatica*, vol. 82, pp. 301–313, 2017.
[14] F. Esfandiari and H. K. Khalil, “Output feedback stabilization of fully linearizable systems,” *International Journal of Control*, vol. 56, no. 5, pp. 1007–1037, 1992.
[15] D. Astolfi and L. Marconi, “A high-gain nonlinear observer with limited gain power,” *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3059–3064, 2015.
[16] L. Wang, D. Astolfi, L. Marconi, and H. Su, “High-gain observers with limited gain power for systems with observability canonical form,” *Automatica*, vol. 75, pp. 16–23, 2017.
[17] A. Zemouche, F. Zhang, F. Mazenc, and R. Rajamani, “High-gain nonlinear observer with lower tuning parameter,” *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3194–3209, 2018.
[18] F. Meiners and J. Adamy, “Observer error reduction via direct state reconstruction,” *IEEE Transactions on Automatic Control*, 2024.
[19] ——, “Improving discrete observers via unreliable direct state reconstruction,” *Systems & Control Letters*, vol. 196, p. 105991, 2025.
[20] J. Grizzle and P. Moraal, “Newton, observers and nonlinear discrete-time control,” in *29th IEEE Conference on Decision and Control*. IEEE, 1990, pp. 760–767.
[21] P. Moraal and J. Grizzle, “Observer design for nonlinear systems with discrete-time measurements,” *IEEE Transactions on automatic control*, vol. 40, no. 3, pp. 395–404, 1995.
[22] M. Zeitz, “Canonical forms for nonlinear systems,” in *Nonlinear Control Systems Design*. Elsevier, 1990, pp. 33–38.
[23] P. Billingsley, *Probability and measure*. John Wiley & Sons, 2017.
[24] B. Oksendal, *Stochastic Differential Equations: an Introduction with Applications*. Springer Science & Business Media, 2013.
[25] H. Deng, M. Krstic, and R. Williams, “Stabilization of stochastic nonlinear systems driven by noise of unknown covariance,” *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1237–1253, 2001.
[26] M. Zakai, “On the ultimate boundedness of moments associated with solutions of stochastic differential equations,” *SIAM Journal on Control*, vol. 5, no. 4, pp. 588–593, 1967.
[27] E. Dynkin, *Markov Processes*. Springer, 1965.
[28] H. Kushner, “On the stability of stochastic dynamical systems,” *Proc. of the National Academy of Sciences*, vol. 53, no. 1, pp. 8–12, 1965.
[29] ——, “Stochastic stability,” in *Stability of Stochastic Dynamical Systems: Proc. of the International Symposium, University of Warwick, 1972*. Springer, 1972, pp. 97–124.
[30] A. Levant, “Higher-order sliding modes, differentiation and output-feedback control,” *International journal of Control*, vol. 76, no. 9-10, pp. 924–941, 2003.