7 Appendix

7.1 Compensated Tracking Error Dynamics

The dynamics of the compensated tracking errors are derived in three different cases.

(1) For $i = 1$

$$
\dot{x}_1 = -k_1 x_1 - \beta f_1 + u_{s1} + (f_1 - \hat{f}_1) - \beta g_1 x_2 + (g_1^0 + \hat{g}_1 + \beta g_1) x_2 + (g_1 - \hat{g}_1) x_2. 
$$

(2) For $1 < i < n$:

$$
\begin{align*}
\dot{x}_i &= -k_i x_i - \beta f_i + u_{s1} + (f_i - \hat{f}_i) \\
&\quad - (g_{i-1}^0 + \hat{g}_{i-1} + \beta g_{i-1}) x_{i-1} - \beta g_i x_{i+1} + (g_i^0 + \hat{g}_i + \beta g_i) x_{i+1} + (g_i - \hat{g}_i) x_{i+1}.
\end{align*}
$$

(3) For $i = n$ ($\ddot{x}_n = \ddot{x}$):

$$
\begin{align*}
\dot{x}_n &= -k_n x_n - (g_{n-1}^0 + \hat{g}_{n-1} + \beta g_{n-1}) x_{n-1} - \beta f_n \\
&\quad + (f_n - \hat{f}_n) - \beta g_n u_{ad} + (g_n - \hat{g}_n) u_{ad} + g_n^0 u_{sn}.
\end{align*}
$$

7.2 Command Filtering

For each $i \in [2, n]$, the signal $x_{ic}$ is required for eqn. (3) and its derivative $\dot{x}_{ic}$ is required for calculating $u_{an}$ (see eqns. (6) and (10)). They are defined by the following procedure (Farrell et al. 2004, Farrell & Polycarpou 2006).

(1) For $i = 1, \ldots, n-1$, define

$$
x_{i+i, c}^0 = \alpha_i - \tilde{\xi}_{i+i, 1}. \tag{35}
$$

The signals $x_{i+i, c}$ and $\dot{x}_{i+i, c}$ are defined by

$$
\dot{x}_{i+i, c} = -K_{i+1} (x_{i+i, c} - x_{i+i, c}^0) \tag{36}
$$

with $K_{i+1} > k_{i+1}$ being a designer specified constant and $x_{i+i, c}(0) = \alpha_i(0)$. Since the filter of (36) is being used as a means to compute $x_{i+i, c}$ and $\dot{x}_{i+i, c}$, without differentiation, the designer would typically select $K_{i+1} \gg k_{i+1}$ so that $x_{i+i, c}$ accurately tracks $x_{i+i, c}^0$ over the bandwidth of $x_{i+i, c}^0$. Since (36) is a stable linear filter, $x_{i+i, c}$ and $\dot{x}_{i+i, c}$ will be bounded if the input $x_{i+i, c}^0$ is bounded.

(2) For $i = 1, \ldots, n-1$, define

$$
\tilde{\xi}_i = -k_i \xi_i + (g_i^0 + \hat{g}_i + \beta g_i)(x_{i+i, c} - x_{i+i, c}^0) \tag{37}
$$

with $\xi_i(0) = 0$.

This is a stable low pass filter. Its input is the product of $(g_i^0 + \hat{g}_i + \beta g_i)$ which we will prove to be bounded and $(x_{i+i, c} - x_{i+i, c}^0)$ which is small. For $(x_{i+i, c} - x_{i+i, c}^0) \in D$ we always have that $|x_{i+i, c} - x_{i+i, c}^0| < 2\rho(D)$ where $\rho(D) = \max_{x_1, x_2 \in D} \|x_1 - x_2\|$ is the diameter of set $D$. For any $x$, each $\xi_i$ is bounded by $\tilde{b}_\xi$, i.e., $|\xi_i| \leq \tilde{b}_\xi$, where

$$
\tilde{b}_\xi = 2\rho(D) \max_i \left[ \sup_{\forall} (|g_i^0 + \hat{g}_i + \beta g_i|) \right] \tag{38}
$$

with $k = \min_i \{k_i\}$.

For completeness, the signal $\xi_n = 0$.

7.3 Sliding Mode

For $x(t) \notin D$, we implement sliding components within the control design to return the state $x$ to the approximation region $\bar{D}$ in finite time. The following assumption is required for the sliding mode design.

**Assumption 4** For $i = 1, \ldots, n$, there exist known upper bounds on unknown functions $|f_i(x)|$ and $|g_i(x)|$ such that $|f_i(x)| \leq \tilde{b}_f$, and $|g_i(x)| \leq \tilde{b}_g$, for any $x(t) \in \mathbb{R}^n - D$.

Note that if constants $\tilde{b}_f$, and $\tilde{b}_g$, are not known, then they could be estimated using the methods suggested in (Polycarpou 1996). We do not present such an adaptive bounding approach herein for $x \notin D$ as that portion of the state space is not the main topic of this article. Note also that the approach extends directly to more general (e.g., linear or quadratic growth) bounds than the constant bounds assumed herein.

The $u_{si}(t)$ terms in eqns. (5–6) are defined as

$$
u_{si}(t) = -r_i(t) sign(\ddot{x}_i) \tag{39}
$$

and the gain $r_i(t)$ is given by

$$
r_i(t) = \begin{cases} 
0, & \text{when } x \in D \\
\tilde{b}_f + \tilde{b}_g |x_{i+1}|, & \text{when } x \notin D
\end{cases} \tag{40}
$$

where $\tilde{b}_f$, $\tilde{b}_g$, are known bounds on $|f_i(x)|$ and $|g_i(x)|$ satisfying the Assumption 4 for $x \notin D$.

The sliding component $u_{sn}$ is defined as

$$
u_{sn} = -r_n(t) sign(\ddot{x}_n) \tag{41}
$$

$$
r_n(t) = \begin{cases} 
0, & \text{when } x \in D \\
\frac{\tilde{b}_f + \tilde{b}_g |u_{ad}|}{g_n^0 + g_n}, & \text{when } x \notin D
\end{cases} \tag{42}
$$

where $\tilde{b}_f$, and $\tilde{b}_g$, are known bounds on $|f_n(x)|$ and $|g_n(x)|$ satisfying the Assumption 4 for $x \notin D$.

The main objective of this section is to demonstrate that the definitions of $u_{si}, 1 \leq i \leq (n-1)$ in eqns. (39–40) and
we can attain methods similar to those used to derive (38). Therefore, $u$ in eqns. (32–34) and applying the sliding control of (39) since the sliding gains of (40) and (42) yield, for any \( x \) within a compact region $B$, then $|\xi| < b_\xi$ where $b_\xi = \frac{2d(s)}{k} (g_n + g_0)$. By methods similar to those used to derive (38), therefore, we can attain $\| \xi(t) \| < \frac{\gamma}{2}$ for $x \notin D$ by choosing $\frac{\gamma}{2} > b_\xi$. Therefore, for $t > T_2$, 

$$\| \dddot{x}(t) \| \leq \| \dddot{x}(t) \| + \| \dddot{\xi}(t) \| < \gamma$$

which implies that $x$ returns to within $D$ in finite time. The state in $D$ may leave that region, but will return to $D$ in finite time.

7.4 $\dddot{x}$ Modification

As we state in Section 5.1, the main drawback of the standard $\sigma$-modification is that it causes the parameter estimates to drift towards certain design vectors. This can occur for $\theta_k$, either when $x \notin S_k$ or when $x \in S_k$ and $\| \dddot{x} \|$ is small. In Section 5.1 the first issue was addressed by localization of the $\sigma$-modification term. The second issue can be addressed by a localized $\dddot{x}$-modification as presented in this section.

The localized $\dddot{x}$-modification terms are defined as,

$$Q_{f_i} = -\sigma_{f_i} \| \dddot{x} \| R_{f_i} (\theta_{f_i} - \bar{\theta}_{f_i})$$
$$Q_{g_i} = -\sigma_{g_i} \| \dddot{x} \| R_{g_i} (\theta_{g_i} - \bar{\theta}_{g_i})$$
$$Q_{\Psi_{f_i}} = -\sigma_{\Psi_{f_i}} \| \dddot{x} \| R_{\Psi_{f_i}} (\Psi_{f_i} - \bar{\Psi}_{f_i})$$
$$Q_{\Psi_{g_i}} = -\sigma_{\Psi_{g_i}} \| \dddot{x} \| R_{\Psi_{g_i}} (\Psi_{g_i} - \bar{\Psi}_{g_i})$$

for $i = 1, \ldots, n$. For consistancy, we have used the same design parameter notation as in the $\sigma$-modification approach.

Substituting the $Q$ terms defined in (47–50) into (22), we obtain the derivative of $V(t)$ as

$$\dot{V} \leq -\frac{1}{2} \| \dddot{x} \|^2 + \bar{d} + \| \dddot{x} \| \rho_1$$

where $\rho_1$ is defined in eqn. (29). In eqn. (51), using the inequality $pq \leq \alpha^2 p^2 + \frac{1}{4\alpha^2} q^2$ with $\alpha^2 = \frac{\delta}{2}$, we obtain

$$\dot{V} \leq -\frac{1}{2} \| \dddot{x} \|^2 + \bar{d} + \rho_2$$

where $\rho_2$ is a positive constant given by

$$\rho_2 = \frac{1}{2\delta} \rho_1^2$$

Therefore, we can summarize these results in the following theorem.

**Theorem 3** $[\dddot{x}$-modification] For the higher order system described by (1–2) with the adaptive feedback control law of eqns. (8), (9–10), (41–42), and the parameter adaptation laws of eqns. (13–14) and (15–16) with modification terms defined in (47–50), we have the following stability properties, for $i = 1, \ldots, n$,

1. $\dddot{x}, \dddot{\theta}_{f_i}, \dddot{\theta}_{g_i}, \dddot{\Psi}_{f_i}, \dddot{\Psi}_{g_i} \in L_\infty$;
2. $x, \theta_{f_i}, \theta_{g_i}, \Psi_{f_i}, \Psi_{g_i} \in L_\infty$;
3. $\dddot{x}, \dddot{\theta}_{f_i}, \dddot{\theta}_{g_i}, \dddot{\Psi}_{f_i}, \dddot{\Psi}_{g_i} \in L_\infty$;
4. $\dddot{x}$ is small in the mean square sense, satisfying

$$\int_0^{t+T} \| \dddot{x}(\tau) \|^2 d\tau \leq \frac{2}{\delta} V(t) + \frac{2}{\delta} (\bar{d} + \rho_2) T.$$  


Similar comments about the localized forgetting apply as were stated in the Section 5.1.

7.5 Deadzone

Another means to remove the issue of parameter drift is to include a deadzone in adaptive laws. Implementation of the deadzone requires knowledge of an assumed bound on certain terms as will be discussed below.

For the deadzone approach, the modification terms in eqns. (13–16) are defined as, for \( i = 1, \ldots, n \),

\[
Q_{f_i} = \begin{cases} 
0 & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \\
\Phi_f, \bar{x}_i, \text{ otherwise,} 
\end{cases}
\]

(55)

\[
Q_{g_i} = \begin{cases} 
0 & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \\
\Phi_g, \bar{x}_i, x_{i+1}, \text{ otherwise,} 
\end{cases}
\]

(56)

and

\[
Q_{\Psi f_i} = \begin{cases} 
-\Psi f_i, R_i (\Psi f_i - \Psi f_i^0) & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \\
-\Phi_f, \bar{x}_i, \omega (\bar{x}_i) & \text{otherwise,} 
\end{cases}
\]

(57)

\[
Q_{\Psi g_i} = \begin{cases} 
-\Psi g_i, R_i (\Psi g_i - \Psi g_i^0) & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \\
-\Phi_g, \bar{x}_i, x_{i+1}, \omega (\bar{x}_i, x_{i+1}) & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \text{ and } i < n \\
-\Phi_g, \bar{x}_n, u_{ad} \omega (\bar{x}_n, g_{ad}) & \text{if } \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \text{ and } i = n. 
\end{cases}
\]

(58)

The constant \( \rho_3 > 0 \) is a known strict upper bound on \( (\bar{d} + \rho_3) \), where \( \bar{d} \) is defined in eqn. (23) and

\[
\rho_3 = \frac{1}{2} \sum_{i=1}^{n} \left( \sigma f_i (\Psi f_i^0 - \Psi f_i) \right)^T R_i (\Psi f_i^0 - \Psi f_i) + \sigma g_i (\Psi g_i^0 - \Psi g_i) \right)^T R_i (\Psi g_i^0 - \Psi g_i). 
\]

(59)

The deadzone is in effect for \( \|\bar{x}\| \leq \sqrt{\frac{\rho_3 + \mu}{2}} \) for some positive design constant \( \mu > 0 \). For \( \|\bar{x}\| > \sqrt{\frac{\rho_3 + \mu}{2}} \), the parameter adaptation laws of \( \theta f_i \) and \( \theta g_i, i = 1, \ldots, n \) do not include any modification terms. When \( \|\bar{x}\| \leq \sqrt{\frac{\rho_3 + \mu}{2}} \), all parameter updates stop.

We are now ready to present the applicable stability theorem.

**Theorem 4** [Deadzone] Assuming the upper bound \( \rho_3 > \bar{d} + \rho_3 > 0 \) is known, for the higher order system described by (1)-(2) with the adaptive feedback control law of eqns. (8), (9–10), (41–42), and the parameter adaptation laws of eqns.(13–14) and (15–16) with modification terms defined in (55–58), we have the following stability properties, for \( i = 1, \ldots, n \),

1. \( \bar{x}_i, \hat{\theta}_{f_i}, \hat{\theta}_{g_i}, \Psi f_i, \Psi g_i \in \mathcal{L}_\infty \)
2. \( x_i, \theta_{f_i}, \theta_{g_i}, \Psi f_i, \Psi g_i \in \mathcal{L}_\infty \)
3. \( \bar{x}_i, \hat{\theta}_{f_i}, \hat{\theta}_{g_i}, \Psi f_i, \Psi g_i \in \mathcal{L}_\infty \)
4. \( \bar{x} \) is small in the mean square sense, satisfying

\[
\int_{t}^{t+T} \|\bar{x}(r)\|^2 dr \leq \frac{1}{k} V(t) + b_d^2 T 
\]

(60)

where \( b_d = \sqrt{\frac{\rho_3 + \mu}{2}} \) is the deadzone width.

(5) \( \|\bar{x}\| \) is ultimately bounded by \( b_d \) as \( t \to \infty \), i.e., the total time for \( \|\bar{x}\| > b_d \) is finite.

Assuming that the design constant \( \rho_3 > 0 \) is a strict upper bound on \( (\bar{d} + \rho_3) \), the ultimate bound in Item 5 of Theorem 4 has a useful form that allows the designer to either increase \( k \) or decrease \( \mu \) or \( \rho_3 \) to decrease the ultimate bound on the tracking error.

A disadvantage of deadzone modification is that the implementation of the deadzone requires knowledge of \( (\bar{d} + \rho_3) \) or the upper bound on it over the whole region \( \mathcal{D} \). If \( \mathcal{D} \) is relatively large, the upper bound \( \rho_3 \) can be conservative, which may result in a large deadzone.

7.6 Numerical Example

For illustrative purposes, consider the second-order system given by

\[
\dot{x}_1 = \sin(x_1 + x_2) + \left(2 + g_1(x)\right) x_2 
\]

(61)

\[
\dot{x}_2 = \sin(x_2) + \left(2 + g_2(x)\right) u. 
\]

(62)

with \( g_1(x) = g_2(x) = \frac{1}{20} \left( x_1^2 + |x_1| \right) \cos(0.01 \pi x_1) \) and \( x = [x_1, x_2]^T \). The system is designed to operate over the region \( \mathcal{D} = [-3, 3] \times [-3, 3] \). Note that eqns. (61–62) are in the form of (1–2). Assume that the known design models is

\[
\dot{x}_1 = 2 x_2 
\]

(63)

\[
\dot{x}_2 = 2 u. 
\]

(64)

where \( f_1(x) = f_2(x) = 0 \), \( g_1(x) = g_2(x) = 2 \). In this case, the unknown model errors are \( f_1(x) = \sin(x_1 + x_2), f_2(x) = \sin(x_2), g_1(x) \) and \( g_2(x) \). Each of these unknown functions will be adaptively approximated during operation.
The reference trajectory $x_d(t)$ and its derivative $\dot{x}_d(t)$ are generated as the output of a second order, unity DC gain, low-pass prefilter given by

$$\begin{align*}
\dot{z}_1 &= z_2 \\
\dot{z}_2 &= a_1 \left[ sat(a_2(sat(r) - z_1)) - z_2 \right]
\end{align*}$$

$$\begin{bmatrix}
x_d \\
\dot{x}_d
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
z_1 \\
z_2
\end{bmatrix}$$

where the two $sat(\cdot)$ functions are included to limit the input magnitude and rate such that $(x_d, \dot{x}_d) \in D$ for any $t > 0$. In our simulations, we select $a_1 = 2\zeta \omega_n$ and $a_2 = \frac{\omega_n^2}{2\zeta \omega_n}$ with $\zeta = 0.9$ and $\omega_n = 5$ such that, in the absence of magnitude and rate saturation, the transfer functions are

$$\frac{x_d^{(i)}(s)}{r(s)} = \frac{\omega_n^2 s^2}{s^2 + 2\zeta \omega_n s + \omega_n^2}, \quad i = 0, 1$$

which are Bounded-Input-Bounded-Output (BIBO) stable. As long as $r(t)$ is bounded, we obtain continuous, bounded signals $x_d$ and $\dot{x}_d$ that will be used in the computation of $x_d^{(0)}(t)$. Theoretically, the input to the prefilter $r(t)$ can be any bounded signal. For the purpose of this simulation we select $r(t) = 3\sin(0.2\pi t)$.

For approximation of $f_1, f_2, g_1$ and $g_2$, we use the same vector of basis functions $\Phi(x) = [\phi_1(x), \ldots, \phi_N(x)]^\top$ for all function approximations. The basis functions are selected as

$$\phi_k(x), \quad k = 1, \ldots, N \text{ are the normalized biquadratic kernels:}$$

$$\phi_k(x) = \begin{cases} (1 - R^2)^2, & \text{if } R < 1 \\ 0, & \text{otherwise} \end{cases} \quad (65)$$

where

$$R = \left\| \frac{|x_1 - c_{k,1}|}{\mu_{k,1}}, \frac{|x_2 - c_{k,2}|}{\mu_{k,2}} \right\|_{\infty};$$

$c_k = [c_{k,1}, c_{k,2}]^\top$ is the center location of the $k$-th basis function; and, $\mu_{k,1}$ and $\mu_{k,2}$ are the constant radii of the region of support in the $x_1$ and $x_2$ directions, respectively. For both $x_1$ and $x_2$, the centers are allocated 0.3 units apart with $\mu_{k,1} = \mu_{k,2} = 0.45$.

The simulation initial conditions for the parameter vectors are $\theta_{f_1}(0) = \theta_{g_1}(0) = \theta_{f_2}(0) = \theta_{g_2}(0) = [0, \ldots, 0]^\top$, $\Psi_{f_1}(0) = \Psi_{g_1}(0) = \Psi_{f_2}(0) = \Psi_{g_2}(0) = [0.5, \ldots, 0.5]^\top$, and $\theta_{f_1}^0 = \theta_{g_1}^0 = \theta_{f_2}^0 = \theta_{g_2}^0 = [0, \ldots, 0]^\top$, $\Psi_{f_1}^0 = \Psi_{g_1}^0 = \Psi_{f_2}^0 = \Psi_{g_2}^0 = [0.005, \ldots, 0.005]^\top$. The adaptation rate matrices in (13-16) are set to $\Gamma_{f_1} = \Gamma_{g_1} = \Gamma_{f_2} = \Gamma_{g_2} = 30I_N$, $\Gamma_{f_1} = \Gamma_{g_1} = \Gamma_{f_2} = \Gamma_{g_2} = 3I_N$ where $I_N$ is the identity matrix in $\mathbb{R}^N$. For comparison purposes, we choose the same value for the $\sigma$ parameter in simulations for each of the different modification terms. They are selected as $\sigma_{f_1} = \sigma_{g_1} = \frac{1}{300}$ and $\sigma_{f_2} = \sigma_{g_2} = \frac{1}{30}$, $i = 1, 2$.

This example compares the compensated tracking error performance when the approximator parameter estimates and bounding parameter estimates are updated based on the adaptation laws of the three different modification methods. Figs. 3, 4 and 5 show the performance, over the first four repetitions of the reference trajectory, for the adaptation algorithms using $\sigma$-modification, $\bar{x}$-modification and deadzone modification, respectively. Each figure plots compensated tracking errors $\bar{x}_1$ (top), $\bar{x}_2$ (middle) and the control input $u$ (bottom) for adaptation algorithms with localized (black solid lines) and standard (red dashed lines) $\sigma$-modification algorithms for $\theta_i$.

The control gains are selected as $k_1 = 2$ and $k_2 = 4$. In order to implement the $\omega(\cdot)$ function in the definitions of $\beta_{f_i}$ and $\beta_{g_i}$, we select $\epsilon = 0.01$.

This example compares the compensated tracking error performance when the approximator parameter estimates and bounding parameter estimates are updated based on the adaptation laws of the three different modification methods. Figs. 3, 4 and 5 show the performance, over the first four repetitions of the reference trajectory, for the adaptation algorithms using $\sigma$-modification, $\bar{x}$-modification and deadzone modification, respectively. Each figure plots compensated tracking errors $\bar{x}_1$ (top), $\bar{x}_2$ (middle) and the control input $u$ (bottom) for adaptation algorithms with localized (black solid lines) and standard (red dashed lines) $\sigma$-modification algorithms for $\theta_i$.

The control gains are selected as $k_1 = 2$ and $k_2 = 4$. In order to implement the $\omega(\cdot)$ function in the definitions of $\beta_{f_i}$ and $\beta_{g_i}$, we select $\epsilon = 0.01$.

This example compares the compensated tracking error performance when the approximator parameter estimates and bounding parameter estimates are updated based on the adaptation laws of the three different modification methods. Figs. 3, 4 and 5 show the performance, over the first four repetitions of the reference trajectory, for the adaptation algorithms using $\sigma$-modification, $\bar{x}$-modification and deadzone modification, respectively. Each figure plots compensated tracking errors $\bar{x}_1$ (top), $\bar{x}_2$ (middle) and the control input $u$ (bottom) for adaptation algorithms with localized (black solid lines) and standard (red dashed lines) $\sigma$-modification algorithms for $\theta_i$.

The control gains are selected as $k_1 = 2$ and $k_2 = 4$. In order to implement the $\omega(\cdot)$ function in the definitions of $\beta_{f_i}$ and $\beta_{g_i}$, we select $\epsilon = 0.01$.

This example compares the compensated tracking error performance when the approximator parameter estimates and bounding parameter estimates are updated based on the adaptation laws of the three different modification methods. Figs. 3, 4 and 5 show the performance, over the first four repetitions of the reference trajectory, for the adaptation algorithms using $\sigma$-modification, $\bar{x}$-modification and deadzone modification, respectively. Each figure plots compensated tracking errors $\bar{x}_1$ (top), $\bar{x}_2$ (middle) and the control input $u$ (bottom) for adaptation algorithms with localized (black solid lines) and standard (red dashed lines) $\sigma$-modification algorithms for $\theta_i$.
of $\mathcal{D}$ where it has operated in the past. After initial transients, no improvement occurs for $\bar{x}_1$ or $\bar{x}_2$ when the standard $\sigma$-modification term is used. This is because the global effect of the $\sigma$-term causes the approximated functions and bounds to lose their local approximation accuracy on any subregion $S_k$ of $\mathcal{D}$ when the trajectory leaves $S_k$. When the state returns to the same local region $S_k$ later, all parameters will need to be estimated again. The adaptation laws with the localized $\sigma$-modification fix this issue by maintaining learned knowledge for later use, since the corresponding parameters relevant to $S_k$ are left unchanged when the state is outside of $S_k$.

In Fig. 4, using the adaptation laws with the standard $\bar{x}$-modification terms (red solid) achieves improved (i.e., smaller) compensated tracking errors over each repetition of the reference trajectory. This indicates that a standard $\bar{x}$-modification term can be used to address the drifting in the parameter estimates when $x \in S_k$ and $\|\bar{x}\|$ is small. However, the parameter drifting could still occur for each parameter estimate when $x \notin S_k$, if $\|\bar{x}\|$ were not small. The localized $\bar{x}$-modification term is designed to address this issue by localizing the effect of drifting to the vicinity of the present operating point. Although we observe similar tracking performance for the standard $\bar{x}$-modification algorithm (red solid) and the localized $\bar{x}$-modification algorithm (black solid) in Fig.

4, these two algorithms have distinct learning features. When $x \notin S_k$, the parameter drifting caused by the standard $\bar{x}$-modification algorithm has a global effect, but at a much slower rate, than the standard $\sigma$-modification algorithm due to inclusion of the $\|\bar{x}\|$ term. Therefore, the local estimation accuracy could be mostly preserved in the standard $\bar{x}$-modification algorithm when the region is revisited later. The localized $\bar{x}$-modification algorithm can retain all knowledge learned from past experience for future use within $S_k$ because the parameter drifting is localized to $S_k$ and will not occur when $x \notin S_k$.

Fig. 5 has shown the tracking performance for simulations with the standard (red solid) and localized (black solid) deadzone modification algorithms. With the standard deadzone modification algorithm, we observe improvement for both $\bar{x}_1$ and $\bar{x}_2$ over each repetition of the reference trajectory. This is because the global forgetting caused by the standard deadzone modification will have no effect when $\|\bar{x}\|$ is small (i.e., within a designer specified deadzone). Therefore, local approximation accuracy on the subregion $S_k$ is partially preserved for future use in the standard deadzone modification algorithm. Compared to tracking performance for the standard $\sigma$-modification algorithm (red solid) given in Fig. 3, the standard deadzone modification algorithm is shown to be effective to eliminate the issue of parameter drift-
ing when \( x \in S_k \) and \( \| \hat{x} \| \) is small. However, the forgetting caused by the standard modification term will still have a global effect. Since the localized deadzone modification algorithm can address this issue and it will not lose any learned knowledge when \( x \notin S_k \), better tracking performance is observed for the localized deadzone modification algorithm than for the standard deadzone modification algorithm in Fig. 5.

For simulations without adaptive bounding (i.e., \( \beta \) terms are all zero), we have also compared the tracking performance for localized (black dashed lines) and standard (red dashed lines) modification algorithms in Figs. 3, 4 and 5. In this case, only the approximator parameter estimates \( \theta_i \) are updated based on either localized or standard modification algorithms, while the adaptation for the bounding parameter estimates \( \Psi_i \) is turned off. In Fig. 3, we observe the effect of global forgetting resulting from the standard \( \sigma \)-modification algorithm, while the localized \( \sigma \)-modification algorithm maintains improved (i.e., smaller) tracking error performance over each repetition of the reference trajectory. Similar comments about the \( \hat{x} \)-modification terms apply in Fig. 4 as were discussed for the case with adaptive bounding. Fig. 5 shows the same tracking error performance for localized (black dashed) and standard (red dashed) deadzone modification algorithms because the adaptation laws for \( \theta_i \) for these two cases are the same.

The goal of this paragraph is to demonstrate the issue of global forgetting caused by the standard (i.e., non-localized) modification terms. To allow illustration of the approximator performance for a 2-input function, the figures will consider the true function and the approximated function versus \( x_1 \) for a fixed value of \( x_2 \) (i.e., \( f_1(x_1, x_2) \) and \( \hat{f}_1(x_1, x_2) \)) at different times. Fig. 6 displays the \( x_1-x_2 \) trajectory from the simulation in a phase plane plot. Consider \( x_2 = 0.27 \). Fig. 6 shows that when \( x_2 \) is near 0.27 then, for the trajectory of this simulation, \( x_1 \) was either near \( x_1 = -2.8 \) or \( x_1 \in [1.5, 2.5] \). Fig. 7 plots the true \( f_1 \) function (dotted) versus \( x_1 \) for \( x_2 = 0.27 \) and the approxima-

![Fig. 6. Plot of \( x_2 \) versus \( x_1 \) for the simulation.](image)

![Fig. 7. True \( f_1 \) function (dotted) versus \( x_1 \) for \( x_2 = 0.27 \) and the approximation \( \hat{f}_1 \) at different times: \( t = 0s \) (solid), \( t = 2s \) (solid), \( t = 105s \) (dash-dot) and \( t = 150s \) (dashed) for the localized \( \sigma \)-modification (top) and standard \( \sigma \)-modification (bottom).](image)
7.7 Useful Properties

7.7.1 Discussion for Assumption 5

**Assumption 5** The scalar function $-1 \leq \omega(z) \leq 1$ satisfies

$$0 \leq |z| - z \omega \left( \frac{z}{\epsilon} \right) \leq \eta \epsilon \quad \forall z \in \mathbb{R}, \quad (66)$$

for any $\epsilon > 0$ and some constant $0 < \eta < \infty$.

The function $\omega(z) = \text{tanh}(z)$ satisfies Assumption 5 for $\eta$ satisfying $\eta = e^{-(q+1)}$, i.e. $\eta = 0.2785$ (Polycarpou 1996). The function

$$\omega(z) = \text{sat}(z) = \begin{cases} 1 \text{ for } z \geq 1 \\ z \text{ for } |z| \leq 1 \\ -1 \text{ for } z \leq -1. \end{cases}$$

satisfies Assumption 5 with $\eta = 0.25$.

The inequality (66) in Assumption 5 can be extended to another form needed in Section 5. The bound

$$0 \leq |z| - z \omega \left( \frac{z}{\epsilon} \right) \leq \left| \frac{z}{z_a} \right| \eta \epsilon \quad (67)$$

will be used to address the case of $z_a \neq z$, but $\text{sign}(z_a) = \text{sign}(z)$.

The inequality (67) can be shown by multiplying both sides of

$$0 \leq |z_a| - z_a \omega \left( \frac{z_a}{\epsilon} \right) \leq \eta \epsilon$$

by $\frac{z}{z_a}$. After algebraic manipulations, we have

$$0 \leq |z| - \text{sgn}(z_a)|z_a| \left| \frac{z}{z_a} \right| \omega \left( \frac{z}{\epsilon} \right) \leq \left| \frac{z}{z_a} \right| \eta \epsilon$$

$$0 \leq |z| - \text{sgn}(z)|z| \omega \left( \frac{z}{\epsilon} \right) \leq \left| \frac{z}{z_a} \right| \eta \epsilon$$

which completes the proof.

7.7.2 Adaptation Law with Parameter Projection

The objective of this appendix is to derive convex sets within which the parameter updates of $\theta_{g_i}$ and $\Psi_{g_i}, i = 1, \ldots, n$ can be constrained to ensure that $\dot{g}_i + \beta_{g_i}$ satisfies the controllability condition of Assumption 3.

For controllability, we must have $\dot{g}_i + \beta_{g_i} > g_i$ which is the same as

$$\theta_{g_i}^T \Phi_{g_i} + \Psi_{g_i}^T \Phi_{g_i} \omega(\cdot) > g_i, \quad i = 1, \ldots, n \quad (68)$$

where the argument dependence of the function $\omega$ is dropped for presentation simplicity. Since $-1 \leq \omega(\cdot) \leq 1$, we can show

$$-\Psi_{g_i}^T \Phi_{g_i} \leq \Psi_{g_i}^T \Phi_{g_i} \omega(\cdot) \leq \Psi_{g_i}^T \Phi_{g_i}. \quad (69)$$

Therefore, it is easy to see that inequality (68) will hold $\forall x \in D$ if and only if

$$\theta_{g_i}^T \Phi_{g_i}(x) - \Psi_{g_i}^T \Phi_{g_i}(x) > g_i, \quad \forall x \in D, \quad i = 1, \ldots, n. \quad (70)$$

If the $\Phi_{g_i}, 1 \leq i \leq n$ form a partition of unity (i.e., $\sum_i \Phi_{g_i} = 1$), we can easily show that condition (70) is satisfied if and only if, for any $x \in D$ and $i = 1, \ldots, n, j = 1, \ldots, N$,

$$\theta_{g_i,j} > g_i, \quad 0 \leq \Psi_{g_i,j} < \theta_{g_i,j} - g_i \quad (71)$$

is satisfied. The reason why we prefer to use condition (71) instead of condition (70) is that the condition (71) defines a convex set within which the projection modification can be easily applied.

Then, we use the following parameter projection to constrain the parameter updates

$$P_S \{ \theta_{g_i,j} \} = \begin{cases} \dot{\theta}_{g_i,j} & \text{if } \dot{\theta}_{g_i,j} > g_i \text{ or } \dot{\theta}_{g_i,j} > 0 \\ 0 & \text{otherwise}. \end{cases} \quad (72)$$

Similarly,

$$P_S \{ \Psi_{g_i,j} \} = \begin{cases} \dot{\Psi}_{g_i,j} & \text{if } \left( \Psi_{g_i,j} < \theta_{g_i,j} - g_i \text{ or } \dot{\Psi}_{g_i,j} < 0 \right) \\ and \quad \dot{\Psi}_{g_i,j} > 0 & \text{or } \dot{\Psi}_{g_i,j} > 0 \\ 0 & \text{otherwise}. \end{cases}$$

7.7.3 Proof of Lemma 1

**Proof:** Since the function $M$ is positive definite and satisfies the inequality

$$\frac{d}{dt} M(z, \Theta_1, \ldots, \Theta_p, t) \leq -c_1 ||z||^2 + c_2,$$

we know that $\dot{M}$ is negative definite whenever $c_1 ||z||^2 > c_2$. Assume that $||z(t)|| > \sqrt{\frac{c_2}{c_1}}$ for $t \in (\tau_1, \tau_2)$. Therefore, the function $M(z, \Theta_1, \ldots, \Theta_p, t)$ on $(\tau_1, \tau_2)$ is bounded by $M(z, \Theta_1, \ldots, \Theta_p, \tau_1)$.

Let $t \in [\tau_2, \tau_3]$ with $||z(t)|| \leq \sqrt{\frac{c_2}{c_1}}$, we will next show that each $\dot{\Theta}_i(t)$ is bounded on $[\tau_2, \tau_3]$. Since $\Theta_i$ satisfies

$$\dot{\Theta}_i = \frac{d}{dt}(\Theta_i - \Theta_i^0) = \Gamma[\Phi z_i - c_0 R(\Theta_i - \Theta_i^0)],$$

we have
Then $\Theta_i(t)$ can be explicitly solved as

$$
\Theta_i(t) = \Theta_i^0 + \left(\Theta_i(t_2) - \Theta_i^0\right)e^{-c_0\tau} \int_{t_2}^t R(\tau(\lambda))d\lambda + \int_{t_2}^t e^{-c_0\tau} \int_{\lambda}^t R(\tau(v))d\tau \Phi(\tau(\lambda))z_i(\lambda)d\lambda.
$$

Note that for $t \in [t_2, t_3]$, $|z_i(t)| \leq \sqrt{\frac{2z_i}{c_1}}$, $R(G)$ is a square diagonal matrix with nonnegative diagonal components, and $\Phi(G)$ is a vector of positive, bounded functions; therefore, $\|\Theta_i(t)\|$ is bounded on $[t_2, t_3]$ by a finite value, i.e., there exists a $\bar{\Theta}_i$ such that $\|\Theta_i(t)\| < \bar{\Theta}_i < \infty$.

Also, it is easy to show that $M$ is bounded such that $M(G, \Theta_1, \cdots, \Theta_P, t) < \varphi_2 \left(\sqrt{\frac{2z_i}{c_1}}, \bar{\Theta}_1, \cdots, \bar{\Theta}_P\right) < \infty$ for $t \in [t_2, t_3]$.

We thus conclude the boundedness of $M(G, \Theta_1, \cdots, \Theta_P)$, $|z_i(t)|$ and $\|\Theta_i(t)\|$ for any $t \in [0, t_f]$.

### 7.7.4 Proof of Theorem 1

**Proof.** When $x \notin D$, we have already shown in Section 7.3 that the sliding components $u_s, i = 1, \ldots, n$ will return the state to $D$ in finite time.

The proof for the case of $x \in D$ and $\delta_{f_i} = \delta_{g_i} = 0$ is based on the eqn. (21). The negative semi-definiteness of $\frac{dx}{dt}$ implies that the variables $\bar{x}_i, \theta_{f_i}, \theta_{g_i}, i = 1, \ldots, n$ are each bounded. Let $Z(t) = \|\bar{x}(t)\|^2$. Since $\int_0^\infty Z(\tau)d\tau \leq V(t) \leq \frac{1}{2} \int_0^\infty Z(\tau)d\tau$, $Z(t) = \sum_{i=1}^n \bar{x}_i \bar{x}_i$ is bounded, Barbalat’s lemma (p. 123 in (Slotine & Li, 1991)) applied to $Z(t)$ implies that each $\bar{x}_i$ approaches zero as $t$ approaches infinity. Finally, since

$$
\dot{V} \leq -\sum_{i=1}^n k_i \bar{x}_i^2
$$

$$
V(t) - V(0) \leq -\sum_{i=1}^n \int_0^t k_i \bar{x}_i^2(\tau)d\tau
$$

$$
V(0) \geq \sum_{i=1}^n \int_0^t k_i \bar{x}_i^2(\tau)d\tau,
$$

we show that each $\bar{x}_i$ is in $L_2$.

### 7.7.5 Proof of Theorem 2

**Proof.** For the proof herein, we only consider the case of $x \in D$. When $x \notin D$, we have already shown in Section 7.3 that sliding control terms we define will return the state to $D$.

Using Lemma 1, we have that $V(t)$ defined in (17) is bounded and then $\bar{x}_i, \hat{\theta}_f, \hat{\theta}_g, \bar{\Psi}_f, \bar{\Psi}_g \in L_\infty$. This yields directly $\bar{\theta}_f, \bar{\theta}_g, \bar{\Psi}_f, \bar{\Psi}_g \in L_\infty$. The fact that $x_1 \in L_\infty$ comes from the fact that $x_d$ and $\xi$ are bounded.

Together with the boundedness of $\Psi_f$ and $\Phi_f$, we can show directly that $u \in L_\infty$ and then $\bar{x}_i, \hat{\theta}_f, \hat{\theta}_g, \bar{\Psi}_f, \bar{\Psi}_g \in L_\infty$.

For the proof of Item 4 given by (30), we consider integrating (28) on both sides to obtain

$$
V(t + T) - V(t) \leq \int_t^{t+T} (-k\|\bar{x}(\tau)\|^2 + \bar{d} + \rho_1)d\tau
$$

$$
\frac{k}{2} \int_t^{t+T} \|\bar{x}(\tau)\|^2 d\tau \leq V(t) + \int_t^{t+T} (\bar{d} + \rho_2)d\tau
$$

which directly yields (30).

### 7.7.6 Proof of Theorem 3

**Proof.** Properties 1, 2, 3 of Theorem 3 are straightforward to show given the form of inequality (52). The proof is similar to the proof of Theorem 2 in Section 5.1 and will not be repeated here.

Property 4 can be proved by integrating on both sides of (52) as

$$
V(t + T) - V(t) \leq \int_t^{t+T} (-\frac{k}{2}\|\bar{x}(\tau)\|^2 + \bar{d} + \rho_2)d\tau
$$

$$
\frac{k}{2} \int_t^{t+T} \|\bar{x}(\tau)\|^2 d\tau \leq V(t) + \int_t^{t+T} (\bar{d} + \rho_2)d\tau
$$

which directly yields (54).

### 7.7.7 Proof of Theorem 4

**Proof.** Substituting the $Q$ terms defined in (55–58) into (22), for $x \in D$ and $\|\bar{x}\| > \sqrt{\frac{\mu + c}{k}}$, $\dot{V}$ is written as

$$
\dot{V} \leq -k\|\bar{x}\|^2 + \bar{d}
$$

$$
- \sum_{i=1}^n \left( \sigma_{\Phi_1} \bar{\Psi}_f^T R_f (\Psi_f - \Psi_f^0) + \sigma_{\Phi_2} \bar{\Psi}_g^T R_g (\Psi_g - \Psi_g^0) \right).
$$

Similarly as in Section 5.1, we obtain

$$
\dot{V} \leq -k\|\bar{x}\|^2 + \bar{d} + \rho_3 \leq -k\|\bar{x}\|^2 + \rho_3 \leq -\mu < 0.
$$

where $\rho_3$ is defined in eqn. (59). Therefore, if $\|\bar{x}\| > \sqrt{\frac{\mu}{k}}$, then $V$ is decreasing. If $\|\bar{x}\| \leq \sqrt{\frac{\mu}{k}}$ then
$\tilde{f}_i, \tilde{g}_i, \tilde{\tilde{f}}_i$ and $\tilde{\tilde{g}}_i, i = 1, \ldots, n$ are all constant and $\|\tilde{x}\|$ is bounded. Thus, $V(t)$ is bounded by the maximum of $V(0)$ or

$$\max_{\|\tilde{x}\|} V(\tilde{x}, \tilde{\tilde{f}}_i(0), \tilde{\tilde{g}}_i(0), \tilde{\tilde{f}}_i(0), \tilde{\tilde{g}}_i(0))$$

which shows that $\tilde{x}_i, \tilde{\tilde{f}}_i, \tilde{\tilde{g}}_i, \tilde{\tilde{f}}_i, \tilde{\tilde{g}}_i \in L_\infty$. Properties 2, 3 of Theorem 4 can be similarly shown.

For the proof of Item 4, we integrate (74) to obtain

$$k \int_t^{t+T} \|\tilde{x}(\tau)\|^2 d\tau \leq V(t) + \int_t^{t+T} (\tilde{d} + \rho_3) d\tau$$

$$\int_t^{t+T} \|\tilde{x}(\tau)\|^2 d\tau \leq \frac{1}{k} V(t) + \frac{1}{k} \rho_3 T$$

$$\leq \frac{1}{k} V(t) + b_4^2 T$$

which implies $\tilde{x}$ is small in the mean square sense (m.s.s.).

Next, we will show Property 5 using a method of proof similar to that in Chapter 1 of (French et al., 2003). Assume $x$ starts outside the deadzone at $t_0$, enters the deadzone at $t_{2i-1}$ and leaves it at $t_{2i}$ for $i \geq 1$. Then, at the boundary of deadzone

$$V(t_{2i-1}) = V(t_{2i})$$

in the deadzone

$$V(t) \leq V(t_{2i-1}), \forall t \in [t_{2i-1}, t_{2i}],$$

and outside the deadzone according to (75),

$$V(t_{2i+1}) - V(t_{2i}) < -\mu(t_{2i+1} - t_{2i}).$$

Therefore, the total time of $x$ staying outside the deadzone is

$$T_d = (t_1 - t_0) + \sum_{i \geq 1} (t_{2i+1} - t_{2i}),$$

and

$$T_d < \frac{1}{\mu} \left( V(t_0) - V(t_1) + \sum_{i \geq 1} (V(t_{2i}) - V(t_{2i+1})) \right)$$

$$< \frac{1}{\mu} \left( V(t_0) - V(t_1) + \sum_{i \geq 1} (V(t_{2i-1}) - V(t_{2i+1})) \right)$$

$$< \frac{V(t_0)}{\mu}$$

which is a finite value. ■