4th Asia Pacific Conference of the Prognostics and Health Management,
Tokyo, Japan, September 11 – 14, 2023 OS07-06

# A two-step detection method for actuator and sensor failures based on self-repairing PID control

Masayoshi Hara[1], Masanori Takahashi[1]

[1] *Oita University, 700 Dannoharu, Oita, 870-1192, Japan*
{*mhara, m-takahashi* } *@oita-u.ac.jp*

## ABSTRACT

Our previous works have proposed a self-repairing PID (SRPID) control for plants with sensor failures. This method uses an I-controller (integrator) with a well-designed auxiliary signal to detect failures. However, the SRPID has not addressed actuator failures. This paper presents a two-step detection method for actuator and sensor failures based on the SRPID and confirms the effectiveness of the proposed detection method through experiments on a ball-and-beam system.

## 1. INTRODUCTION

Proportional-integral-derivative (PID) control has been used widely in various fields of industry. Generally, an assumption that sensors and actuators are healthy guarantees stability and control performance. However, sensor and actuator failures often destabilize control systems.

As a countermeasure to this problem, our previous works have proposed a self-repairing PID (SRPID) control for plants with sensor failures that exploits an I-controller (an integrator) and an auxiliary signal to detect failures exactly (Takahashi, 2016). The SRPID belongs to a class of active fault-tolerant controls (AFTC) (Jiang, Staroswiecki, & Cocquempot, 2003; Wang, Zhou, Qin, & Wang, 2008; Kawaguchi, Araki, Sato, Kuroda, & Asami, 2020). Compared to the other AFTCs, one of the advantages is that no information about the plants is necessary to detect failures. Unfortunately, issues of actuator failures have never been considered. In addition, a threshold for detection has been determined by trial and error. Hence, misdetection (false positive) sometimes occurs under control conditions, and it takes a long time to find failures.

This paper presents a two-step detection method for sensor and actuator failures based on the SRPID. Concretely, the control structure of the SRPID is modified so that thresholds for detection can be computed only by initial state values. Furthermore, by using the modified SRPID, two thresholds for the two step-detection of sensor and actuator failures are automatically generated.



This paper is organized as follows. Section 2 explains the problem of the conventional SRPID for sensor failures and also presents the modified SRPID for nonlinear double-integrator systems. Section 3 shows the two-step detection of sensor and actuator failures. Section 4 describes the effectiveness of the proposed method through experiments on a ball-and-beam system. Section 5 concludes.

## 2. SELF-REPAIRING PID CONTROL

This section shows a modified self-repairing PID control for plants with sensor failures.

### 2.1. Sensor failure detection in PID control

In this subsection, the self-repairing control in our previous work, is explained briefly. In addition, an issue on a threshold for detection is pointed out.

Consider a plant of the form:

$$\begin{aligned}\dot{x} &= Ax + bu + d\\ y &= c^{\mathsf{T}} x\end{aligned} \tag{1}$$

where $x$ is the state vector, $u$ is the control input, $y$ is the output and $d$ is a bounded disturbance. $A$, $b$ and $c$ are matrix and vectors of appropriate dimensions. Based on dynamic redundancy against failures, two sensors to measure $y$, are used; one is the primary sensor and the other is the backup. Then, the feedback signal $y_S$ is given by

$$y_S(t) = \begin{cases} y_1(t) & (t \leq t_D)\\ y_2(t) & (t > t_D)\end{cases} \tag{2}$$

where $y_1$ and $y_2$ are the outputs of the primary sensor and the backup respectively. Of course, $y_1 = y$ if the primary sensor is healthy. The backup is always maintained to be healthy, i.e., $y_2 = y$. At a detection time $t_D > 0$, the primary sensor is replaced with the backup.

For the plant (1), the PID controller is constructed as

$$u = -p\left(y_S + \frac{1}{T_I}v + T_D\dot{y}_S\right) \tag{3}$$

$$\dot{v} = y_S \tag{4}$$

where $p > 0$ is a control gain, $T_I > 0$ is an integral time, and $T_D > 0$ is a derivative time. Assume that those parameters are chosen to render the overall control system stable. Then, there exists a constant $\rho$ such that

$$\begin{aligned}&\left\|\begin{bmatrix} x(t) \\ v(t) \end{bmatrix}\right\| \\ &\leq \quad \left\|\begin{bmatrix} x(0) \\ v(0) \end{bmatrix}\right\| + \rho\left(\|d\|_\infty + \|\gamma\|_\infty\right), \ \forall t \geq 0\end{aligned} \tag{5}$$

where $\|\cdot\|_\infty = \sup_{0\leq s\leq t} \cdot(s)$.

Consider the following scenario of sensor failure.

**Scenario 1** (the failure of the primary sensor)

$$y_1(t) = \varphi, \ t \geq t_F \tag{6}$$

where $\varphi$ is an unknown value, and $t_F > 0$ is an unknown failure time.

To detect the sensor failures, an auxiliary signal is injected into the integrator (4) as follows.

$$\dot{v} = y_S + \gamma\mathrm{sgn}(y_S) \tag{7}$$

where $\gamma > 0$ is a constant. If there is no failure, then from (5), $v$ is bounded as follows.

$$|v(t)| < \Gamma_S, \ t < t_F \tag{8}$$

where

$$\Gamma_S = \left\|\begin{bmatrix} x(0) \\ v(0) \end{bmatrix}\right\| + \rho\left(\|d\|_\infty + \|\gamma\|_\infty\right) \tag{9}$$

Otherwise (if the sensor fails), $v$ obeys

$$\dot{v} = \varphi + \gamma\mathrm{sgn}(\varphi) \tag{10}$$

which implies

$$\dot{v} > \gamma, \ \varphi \geq 0 \tag{11}$$

$$\dot{v} < -\gamma, \ \varphi < 0 \tag{12}$$

This means that $v$ tends to infinity as $t \to \infty$ unless the failed sensor is replaced. By using this unstable behavior of $v$, the detection time $t_D$ is defined by

$$t_D := \inf\{t > 0 \mid |v(t)| > \Gamma_S\} \tag{13}$$

If the failure occurs, the detection time $t_D$ exists and satisfies

$$t_D - t_F < \frac{2\Gamma_S}{\gamma} \tag{14}$$

Obviously, the maximum detection time is $2\Gamma_S/\gamma$ and can be prescribed in advance. This is one of the advantages of the self-repairing PID control. However, from (9), it is shown that $\Gamma_S$ depends on $\gamma$. Hence, the maximum detection time cannot be shortened arbitrarily by $\gamma$. Furthermore, because $\Gamma_S$ also includes the term of the unknown disturbance, it should be determined by trial and error.

As a remedy, the following subsection presents a modified PID controller for nonlinear double-integrator systems, and it also shown that the threshold can be determined by only the initial values.

### 2.2. Modification of the PID control system

Consider a nonlinear double-integrator system of the form:

$$\ddot{y} = bu + d(y, \dot{y}) \tag{15}$$

where $b > 0$ is a constant, and $d$ is a nonlinear term with respect to $y$ and $\dot{y}$. Assume that $d$ satisfies

$$|d| \leq c_0 + c_1|y| + c_2|\dot{y}| \tag{16}$$

where $c_0 > 0$, $c_1 > 0$ and $c_2 > 0$ are some constants. The above model can represent the behaviors of many mechanical systems.

Now, define an augmented output by

$$z := y + \tau_D\dot{y} \tag{17}$$

where $\tau_D > 0$ is a constant. From (15) and (17), it follows that

$$\begin{aligned}\dot{z} &= -\frac{1}{\tau_D}y + \frac{1}{\tau_D}z + \tau_D bu + \tau_D\tilde{d}(z, y) \\ \dot{y} &= -\frac{1}{\tau_D}y + \frac{1}{\tau_D}z\end{aligned} \tag{18}$$

Notice that the nonlinear term $d$ is transformed to $\tilde{d}$ by $z$.

To stabilize the above system, the control input is given by

$$u = -p\left(z_S + v\right) \tag{19}$$

where

$$z_S = y_S + \tau_D\dot{y}_S \tag{20}$$

$$\dot{v} = \tau_I z_S + \gamma\mathrm{sgn}(z_S) \tag{21}$$

and $p > 0$ and $\tau_I > 0$ are constants. The block diagram of the overall control system is illustrated in Figure 1. The modified input (19) is equivalent to the PID control (3) by the

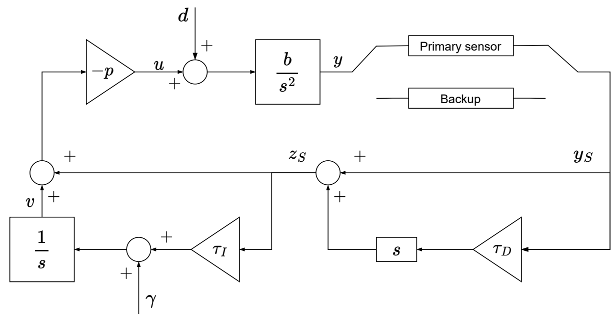



Figure 1. Block diagram of the modified PID control system.

following relationships.

$$p = \frac{p}{1+\tau_I \tau_D} \tag{22}$$

$$T_D = \frac{\tau_D}{1+\tau_I \tau_D} \tag{23}$$

$$\frac{1}{T_I} = \frac{\tau_I}{1+\tau_I \tau_D} \tag{24}$$

If there is no failure, then the following control performances are guaranteed theoretically (see Appendix). For given any $\lambda$, selecting sufficiently large $p$, $\tau_I$ and small $\tau_D$ ensures

$$\left\| \begin{bmatrix} z(t)+v(t) \\ y(t) \\ v(t) \end{bmatrix} \right\| < \max\left\{|z(0)|+|y(0)|+2|v(0)|, \lambda\right\}, \ \forall t \geq 0 \tag{25}$$

and also

$$\limsup_{t\to\infty} \left\| \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} \right\| < \lambda \tag{26}$$

From these results, clearly, all the signals are bounded. In particular, from (25), $|v|$ satisfies

$$|v(t)| < \max\left\{|z(0)|+|y(0)|+2|v(0)|, \lambda\right\}, \ \forall t \geq 0 \tag{27}$$

Then, in the detection rule (12), $v$ of (20) is monitored instead of (7), and a candidate of $\Gamma_S$ is selected by

$$\Gamma_S = \max\left\{|z_S(0)|+|y_S(0)|, \lambda\right\} \tag{28}$$

Note that $v(0)$ is set equal to zero.

Because this threshold $\Gamma_S$ is determined only by the state and is independent of $\gamma$, the maximum detection time $2\Gamma_S/\gamma$ can be shortened arbitrarily. Thus, early and exact detection can be attained for the sensor failures.

## 3. A TWO-STEP DETECTION

In the previous section, the modified PID control system and its sensor failure detection are shown. This section shows the detection of actuator failure as well as sensor failure.

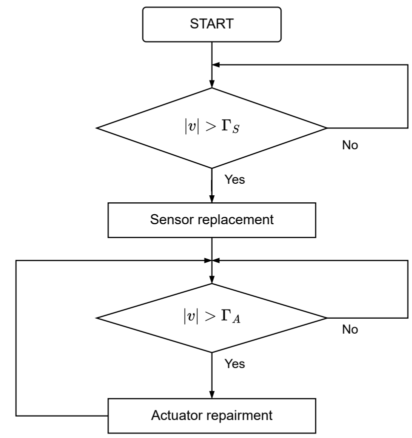



Figure 2. Flow chart of the two-step detection.

Assume that the following scenario is additionally to Scenario 1 of subsection 2.1.

**Scenario 2** (the failure of the actuator)

$$u(t) = u(t_F), \ t \geq t_F \tag{29}$$

When Scenario 2 occurs (the actuator fails), there are two cases: one is the continuation of a stable condition, and the other is the destabilization of the system. Unfortunately, the former might be extremely rare. Even if the state is moving toward an equilibrium point, from (21), $|v|$ tends to diverge.

Because it is not known in advance whether Scenario 1 or 2 will occur, the following detection procedure is employed.

**Step 1:** Suspect Scenario 1. The detection rule (13) for sensor failures is adopted, that is,

$$t_S := \inf\left\{t > 0 \mid |v(t)| > \Gamma_S\right\} \tag{30}$$

At $t = t_S$, the primary sensor is replaced with the backup, that is, the feedback signal is given by

$$y_S(t) = \begin{cases} y_1(t) & (t \leq t_S) \\ y_2(t) & (t > t_S) \end{cases} \tag{31}$$

**Step 2:** If the finite time $t_S$ exists, then the next detection rule is adopted as follows.

$$t_A := \inf\left\{t > t_S \mid |v(t)| > \Gamma_A\right\} \tag{32}$$

where $\Gamma_A$ is given by

$$\Gamma_A = \max\left\{|z_S(t_S)|+|y_S(t_S)|+|v(t_S)|, \lambda\right\} \tag{33}$$

By the above two-step procedure, actuator and sensor failures can be detected as follows.

If the sensor fails (in the case of Scenario 1), there exists $t_S$

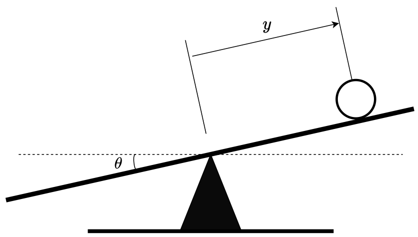



Figure 3. Mathematical model of ball-and-beam system.

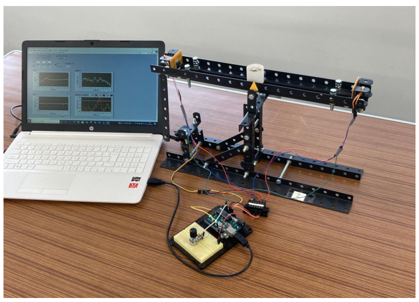


Figure 4. Experimental system.

satisfying

$$t_S - t_F \leq \frac{2\Gamma_S}{\gamma} \tag{34}$$

The control system can recover by replacing the failed sensor with the backup. Because of the boundedness (27), $|v|$ restarting from $t = t_S$ does not hit the threshold $\Gamma_A$. Hence, the finite time $t_A$ does not exists.

If the actuator fails (in the case of Scenario 2), both the finite times $t_S$ and $t_A$ exist.

Thus, the actuator and sensor failures are detected.

## 4. Application to a ball-and-beam system

The proposed detection method in the previous section is applied to a ball-and-beam system to verify its effectiveness.

Consider a ball-and-beam system (BBS) as shown in Figure 3. The mathematical model (the equation of motion) of the BBS is represented as follows (Hauser, Sastry, & Kokotovic, 1992).

$$\ddot{y} = b \sin\theta + cy(\dot{\theta})^2 \tag{35}$$

where $y$ [m] is the ball position, and $\theta$ [rad] is the beam angle (manipulated variable). The parameters of (35) are $b = 7$ and $c = 5/7$ when the gravitational acceleration is approximated as $g = 9.8$ [m/s$^2$]. A servo motor directly manipulates the beam angle $\theta$. Therefore, in this research, the control input is given by $u = \sin\theta$, and $\dot{\theta}$ is supposed to be small. Note that the ball position $y$ is measured by a sensor and is available. The control problem is to stabilize the BBS by manipulating the beam angle.

Table 1 : Physical parameters of the experimental system

| | |
|---|---|
| Ball Diameter | 30.16 [mm] |
| Mass of ball | 111.6 [g] |
| Beam length | 300.0 [mm] |
| Beam width | 40.0 [mm] |

Figure 4 shows the experimental apparatus of the BBS, which consists of the ball-and-beam experiment equipment, a microcontroller board (Arduino UNO), and a PC as the user interface. Table 1 shows the specifications of the BBS.

Through the PC, the PID parameters $p$, $\tau_I$, and $\tau_D$ are sent to the microcontroller board, and the control results are stored in the PC and indicated. The microcontroller board reads the values of the infrared sensors on both sides of the beam and calculates the ball position $y$. It also updates the control input $u$. In the experiment, the sampling period is 100 ms.

The initial values of the ball are $y(0) = -98$ [mm] and $\dot{y}(0) = 0$ [mm/s]. The parameters of the modified PID controller and the auxiliary signal are $p = 3.0$, $\tau_I = 0.2$, $\tau_D = 0.8$, $\lambda = 0.15$ and $\gamma = 0.8\tau_I$ respectively.

Suppose that the actuator fails by turning off the servo motor intentionally at $t_F = 10$ [s].

The detection procedure is as follows (also see Figure 2).

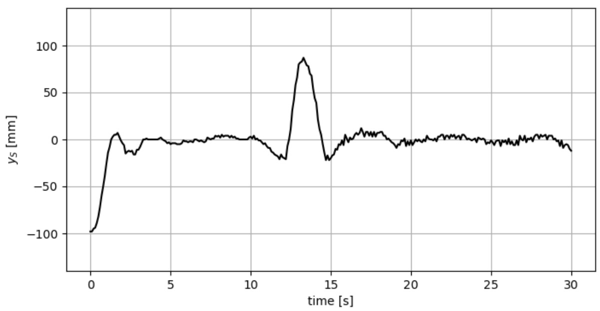
1. When $|v|$ exceeds the threshold $\Gamma_S$, the primary sensor is replaced with the backup. And the threshold $\Gamma_A$ is calculated by (33) at the same time.

2. When $|v|$ exceeds the threshold $\Gamma_A$, the actuator is repaired. Here the power of the servo motor is turned on again.

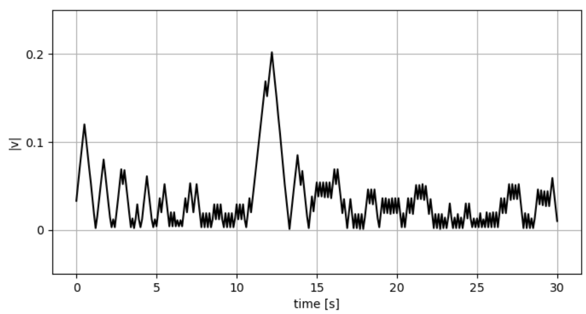
Figure 5 shows the results of the experiments. From the initial values and $\lambda = 0.15$, $\Gamma_S$ (the dashed blue line in Figure 5(c)) is given by $\Gamma_S = \lambda = 0.15$. It can be observed that $|v|$ remains lower than $\Gamma_S$ in the period of no failure. Immediately after the failure, $y_S$ is still close to zero because the beam angle $\theta$ is stuck at zero. $|v|$ increases slowly and then hits the threshold $\Gamma_S$ at $t_S = 11.7$ [s]. The primaly sensor is replaced, and $\Gamma_A$ (the dashed red line in Figure 5(c)) is calculated by $z_S(t_S)$, $y_S(t_S)$ and $v(t_S)$, $\Gamma_A = 0.153$.

After the sensor replacement, $|v|$ exceeds $\Gamma_A$ at $t_A = 11.8$ [s], and then the actuator is repaired (by turning on the servo motor again). The control system can recovery and is stabilized by the PID controller. Therefore, both $y_S$ and $v$ remains small regions.

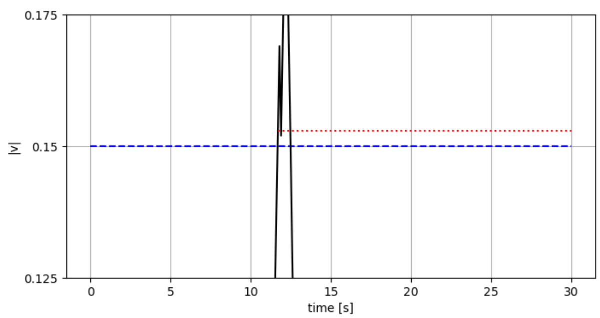
Finally, Table 2 shows the experimental results with various auxiliary signals. From these results, the actuator failures can be successfully detected and large $\gamma$ can shorten the detection times $t_S$ and $t_A$.




(a) The measured ball position, $y_S$




(b) The absolute value of $v$




(c) The absolute value of $v$ around the failure time

Figure 5. Experimantal results

## 5. Conclusion

This paper has presented the modified PID control system and the two-step detection of actuator and sensor failures. The thresholds for detection can be automatically generated by the measured states. Furthermore, experimental results have confirmed that actuator failure can be detected by the proposed method.

However, there are many issues to be considered in future works. Because the auxiliary signal degrades control performance, this signal should be redesigned. An automatic tuner for PID parameters is necessary to handle unknown systems. A class of applicable systems needs to be expanded.

Table 2 : Experimental results: the detection times $t_S$ and $t_A$

| $\gamma/\tau_I$ | $t_S$ | $\Gamma_S$ | $t_A$ | $\Gamma_A$ |
|---|---|---|---|---|
| 0.1 | 26.1 [s] | 0.15 | 26.2 [s] | 0.151 |
| 0.2 | 18.6 [s] | 0.15 | 18.7 [s] | 0.151 |
| 0.4 | 15.5 [s] | 0.15 | 15.6 [s] | 0.151 |
| 0.8 | 11.7 [s] | 0.15 | 11.8 [s] | 0.153 |

## References

Hauser, J., Sastry, S., & Kokotovic, P. (1992). Nonlinear control via approximate input - output linearization: the ball and beam example. *IEEE Trans. on Automatic Control*, *37(3)*, 392-398.

Jiang, B., Staroswiecki, M., & Cocquempot, V. (2003). Active fault tolerant control for a class of nonlinear systems. *IFAC Proceedings Volumes*, *36(5)*, 125-130.

Kawaguchi, N., Araki, N., Sato, T., Kuroda, M., & Asami, T. (2020). Active fault-diagnosis method using adaptive allocator and fault-tolerant adaptive control system design. *ROBOMECH J.*, *7(1)*.

Takahashi, M. (2016). Self-repairing pi/pid control against sensor failures. *Int. J. of Innovative Computing, Information and Control*, *12(1)*, 193-202.

Wang, Y., Zhou, D., Qin, S. J., & Wang, H. (2008). Active fault-tolerant control for a class of nonlinear systems with sensor faults. *Int. J. of Control, Automation and Systems*, *6(3)*, 339-350.

## Biographies

**Masayoshi Hara** received his B. Eng. and M. Eng. degrees from Oita University, Japan in 1991 and 1993 respectively. He is currently an Assistant Professor with the Faculty of Science and Technology, Oita University, Japan. His research interests are in the area of fault detection and actual control system.

**Masanori Takahashi** received his B. Eng., M. Eng., and D. Eng. degrees from Kumamoto University, Japan in 1992, 1994 and 1998 respectively. He is currently a Professor with the Faculty of Science and Technology, Oita University, Japan. His research interests are in the area of fault tolerant control, fault detection and adaptive switching control.

## Appendix

Consider the case when no sensor failure occurs. Then, $y_S = y$ and so $z_S = z$. Now, define a new variable by

$$e := z + v \tag{36}$$

From (17), the time derivative of $e$ is represented as

$$\begin{aligned}\dot{e} \;=\; & -\left(\tau_D bp - \frac{1}{\tau_D} - \tau_I\right) e + \tau_D \bar{d}(e, y, v) \\ & -\frac{1}{\tau_D} y - \left(\frac{1}{\tau_D} + \tau_I\right) v + \gamma \mathrm{sgn}(z) \qquad (37)\end{aligned}$$

Notice that $\tilde{d}$ is transformed to $\bar{d}$ by $e$. Moreover, the time derivatives of $y$ and $v$ are expressed as

$$\dot{y} = -\frac{1}{\tau_D} y + \frac{1}{\tau_D} e - \frac{1}{\tau_D} v \qquad (38)$$

$$\dot{v} = -\tau_I v + \tau_I e + \gamma \mathrm{sgn}(z) \qquad (39)$$

Then, consider a positive definite function as

$$V = \frac{1}{2}\left\{e^2 + y^2 + v^2\right\} \qquad (40)$$

Its time derivative is calculated as follows.

$$\begin{aligned}\dot{V} \;=\; & -\left(\tau_D bp - \frac{1}{\tau_D} - \tau_I\right) e^2 + \tau_D \bar{d}(e, y, v) e \\ & -\frac{1}{\tau_D} ye - \left(\frac{1}{\tau_D} + \tau_I\right) ve + \gamma \mathrm{sgn}(z) e \\ & -\frac{1}{\tau_D} y^2 + \frac{1}{\tau_D} ey - \frac{1}{\tau_D} vy \\ & -\tau_I v^2 + \tau_I ev + \gamma \mathrm{sgn}(z) v \qquad (41)\end{aligned}$$

From (16), the transformed nonlinear term can be evaluated by

$$|\bar{d}| \le c_0 + c_1|y| + \frac{c_2}{\tau_D}\left(|e| + |v| + |y|\right) \qquad (42)$$

Thus, $\dot{V}$ is evaluated as

$$\dot{V} \le -\frac{1}{2}\alpha_1 e^2 - \frac{1}{2}\alpha_2 y^2 - \frac{1}{2}\alpha_3 v^2 + \beta \qquad (43)$$

where

$$\begin{aligned}\alpha_1 &= 2\tau_D bp - \frac{6}{\tau_D} - 4\tau_I \\ &\qquad -\tau_D c_0 - 3\tau_D^3 c_1^2 - 3c_2 - 3c_2^2 \\ \alpha_2 &= \frac{1}{3\tau_D} \\ \alpha_3 &= \tau_I - \frac{4}{\tau_D} - c_2 \\ \beta &= \frac{\tau_D c_0}{2} + \frac{\gamma^2}{2\tau_I} \qquad (44)\end{aligned}$$

Choose sufficiently large $p$ and $\tau_I$ so that $\alpha_1 > 0$ and $\alpha_3 > 0$. Then, it can be verified that

$$\dot{V} \le -\alpha V + \beta, \;\; \alpha = \min\{\alpha_1,\ \alpha_2,\ \alpha_3\} \qquad (45)$$

Therefore, all the signals in the control system are bounded. Furthermore, it is shown that

$$\begin{aligned}V \;\le\; & V(0)e^{-\alpha t} + \frac{\beta}{\alpha}\left(1 - e^{-\alpha t}\right) \\ \le\; & \max\left\{V(0),\ \frac{\beta}{\alpha}\right\} \qquad (46)\end{aligned}$$

For given small $\lambda$, selecting large $p$, $\tau_I$ and small $\tau_D$ gives

$$\frac{\lambda^2}{2} > \frac{\beta}{\alpha} \qquad (47)$$

Thus, it can be shown that all the signals $e$, $y$ and $v$ are bounded because

$$\begin{aligned}& \left\| \begin{bmatrix} e(t) \\ y(t) \\ v(t) \end{bmatrix} \right\| \\ & \le \sqrt{2V(t)} < \max\left\{\sqrt{2V(0)}, \lambda\right\} \\ & \le \max\left\{|e(0)| + |v(0)| + |y(0)|, \lambda\right\} \\ & \le \max\left\{|z(0)| + |y(0)| + 2|v(0)|, \lambda\right\},\ \forall t \ge 0 \qquad (48)\end{aligned}$$

Furthermore, from (45) and (46), it can be seen that

$$\limsup_{t\to\infty} \left\| \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} \right\| < \lambda \qquad (49)$$