High Precision Feedback Control Design for Dual-Actuator Systems

Tielong Shen and Minyue Fu

Abstract—Dual-actuator control systems are designed by piggybacking a small secondary actuator onto a primary actuator. Yet the control design for such systems is challenging due to the limited dynamic range of the secondary actuator and its coupling with the primary actuator. In this paper, we propose a new control design approach for dual-actuator systems. We first show that, through a simple coordinate transformation, a dual-actuator system can be separated into two subsystems, one of them controlled only by the primary actuator, which can be designed using conventional methods, and the other one involving tracking control using the secondary actuator. We then devise a switching control strategy for the secondary actuator to achieve high precision control. Two dual-actuator systems are used to demonstrate the use of dual actuators and effectiveness of the new design approach.

I. INTRODUCTION

The structural characteristic of a dual-actuator system is that there are two cascaded actuators operating along a common axis. The primary actuator, called primary actuator, is of long range but with poor accuracy and slow response time. The minor actuator, called secondary actuator, is typically piggybacked onto the primary actuator and delivers much higher precision and faster response but has a limited dynamic range. A good example of dual-actuator systems is the so-called dual-stage hard disk drive servo system which consists of a voice coil motor as the primary actuator and a piezo-electric transducer as the secondary actuator [1]-[2]. Other examples include VLSI chip manufacturing machines [3], [4] and high precision machining tools [5], [6].

Although the mechanical design of dual-actuator systems appears to be simple and the functionalities of each actuator appear to be qualitatively clear, it is a challenging task to coordinate the two actuators to yield an optimal performance. The main obstacle is that secondary actuator typically has a very limited motion range. A naive design approach is to tune the primary actuator to achieve a rough positioning or tracking and then to activate the secondary actuator when the primary actuator reaches its steady state. This approach appears to make sense because the secondary actuator does not have much influence when the tracking error is large; and conversely, the secondary actuator should take over when the primary actuator has exhausted its ability to reduce the tracking error. However, such an approach is inadequate for at least two reasons: 1) The overall response time is typically excessively long; 2) In many applications the system is subject to disturbances or the tracking reference signal is time-varying, which implies that the primary actuator can never be in a steady state. Many other design approaches also exist. For example, a common design approach for dual-stage hard disk drive systems is to design two separate control loops, one for each actuators by assuming that the interaction between the two control loops is negligible. The two loops can operate in either series [7] or parallel [8]. In such a design, the primary and secondary actuators are assigned to control low and frequency responses, respectively. However, it is well known that this control approach typically results in poor stability performance due to the coupling between the two actuators. Indeed, the most difficult aspect of the dual-actuator control design is to figure out how to coordinate the two actuators in a region where both actuators have some but limited control authorities.

In this paper, we propose a new control design approach for dual-actuator systems. Our research starts with an interesting observation that many dual-actuator systems can be effectively decoupled through a simple coordinate transformation. The result of it is that we have two subsystems, a primary system and a secondary system behaving as follows. The primary system is approximately independent of the secondary system and is purely driven by the primary actuator. The controller of the primary system can be designed using standard means, and thus is not of primary interest. On the other hand, the secondary system is controlled by the secondary actuator and the coupling from the primary system is only through the reference signal for positioning or tracking. Motivated by the observation above, we then focus on the control design for the secondary actuator. Modeling the secondary actuator as a saturated controller, this problem can be cast as a tracking control problem with an input saturation. Existing design approaches range from anti-windup compensators [9]-[10] to Riccati equation or linear matrix inequality based approaches [11]-[13]. In our design approach, we choose to optimize a quadratic performance function. To get around the difficulty with the input saturation, we exploit a switching control strategy which applies different control gains when the state of the system lie in different regions. These regions are characterized by nested ellipsoids with outer ellipsoids corresponding to looser performance bounds and the inner ones to tighter bounds. As the tracking error converges, the controller gradually switches from low performance gains to high performance ones. Two design examples are used to demonstrate our design approach. The first one is a simple two-mass system exemplifying a single-axis linear-motion dual-actuator system. The second one is a dual-stage synchronization system commonly used in the micro-manufacturing industry.
Next, we show that the model (1) can be decoupled into two subsystems via a simple coordinate transformation. This is motivated by the fact that the primary system typically has a much larger mass than the secondary system. Let

\[ z_1 = \frac{m_x x_1 + m_p z_1}{m_x + m_p}, \quad z_2 = \xi_2 = \xi_1 \]  

That is, we choose the mass center \( \xi_1 \) and its derivative as a new coordinate for the primary system. It follows that

\[ \dot{\xi}_2 = b u_p(t) - a_2 \xi_2(t) - a_1 \xi_1(t) + \epsilon(t) \]  

where \( b = (m_x + m_p)^{-1}, a_1 = k_p m_x m_p^{-1}, a_2 = c_p m_x m_p^{-1} \) and \( \epsilon(t) = m_s k_s (x_1 + c_s x_2)/(m_p^2 + m_x m_p) \) describes the coupling force from the motion of the secondary system. Due to the fact that \( m_p \gg m_x \), we can simply ignore the \( \epsilon(t) \) term. It follows that the dynamics of the primary system is approximately decoupled from that of the secondary system and the design of \( u_p(t) \) can be done using any conventional method.

On the other hand, if we define \( e_1 = x_1 - y_r \) and \( e_2 = \dot{e}_1 = x_2 - \dot{y}_r \), then

\[ \dot{e}_2 = g[\sigma[u_s(t)] - d(t)] - h_2 e_2(t) - h_1 e_1(t) \]  

where \( g = m_x^{-1}, h_1 = k_s m_x^{-1}, h_2 = c_p m_x^{-1} \) and

\[ d(t) = c_s (\dot{y}_r - z_2) + k_s [r - (z_1 + l_0)] + m_s \ddot{y}_r. \]

The model (5) now becomes the model for the secondary system. We observe that the coupling signal \( d(t) \) is dominated by the error between the state of the primary actuator \( (z_1, z_2) \) and the desired state \( (y_r - l_0, \dot{y}_r) \) and the acceleration command \( m_s \ddot{y}_r. \) If the error or the acceleration command is large, it is not possible to completely compensate \( d(t) \) by using \( u_s(t). \) It is only when \( d(t) \) is within a small region (in comparison with a) that the secondary actuator has some meaningful control effect.

The observation above means the following. We can use the input of the primary actuator to drive the mass center such that the new reference signal \( d(t) \) converges to a small region as quickly as possible. Prior to such convergence of \( d(t) \), the secondary actuator should be turned off because it does not have much effect. The control gain of \( u_s(t) \) should be gradually increased when \( d(t) \) becomes small. This motivates the idea of a switching control strategy for the secondary actuator, which we discuss in the next two sections.

III. CONTROL DESIGN FOR SECONDARY ACTUATOR

In this section, we focus on the design problem for the secondary system in a dual-actuator system. In view of the discussions in Section II, we consider the following model:

\[ \dot{x} = Ax + b[\sigma(u) - d(t)], \quad x(0) = x_0 \]  

where \( x \in \mathbb{R}^n \) is the state with initial condition \( x_0, u \in \mathbb{R} \) is the control input, \( d(t) \in \mathbb{R} \) is a given reference signal, and \( A \in \mathbb{R}^{n \times n} \) and \( b \in \mathbb{R}^{n \times 1} \) are given. Without loss of generality, we assume that the saturation level is \( a = 1. \)

Roughly speaking, the design goal is to find a state feedback control law such that the state \( x \) is driven from its initial
state $x(0)$ to the origin with a pre-specified performance cost or less and that the region of such initial state is maximal.

It is obvious that if $|d(t)| > 1$ holds persistently, it is not possible to achieve the design goal. Therefore, it is necessary for the primary actuator the bring the steady state value of $d(t)$ to be within $[-1, 1]$. In view of this, we assume that the control action of the primary actuator has been applied prior to $t = 0$ such that $|d(t)| \leq d_0$, $\forall t \geq 0$ for some $d_0 \in (0, 1)$.

Under the assumption, we can compensate the disturbance by introducing a feed-forward term in $u(t)$ as follows:

$$u(t) = v(t) + d(t) \tag{8}$$

Then we can rewrite (7) as

$$\dot{x} = Ax + b\sigma_d[v], \quad x(0) = x_0 \tag{9}$$

where $\sigma_d[\cdot]$ is defined as in (2) with $a = 1 - d(t) > 0$.

We now consider the following quadratic cost function

$$J(x_0, v) = \int_0^\infty \{x^T Q x + r\sigma_d^2[v]\} dt \tag{10}$$

for some $Q = Q^T \geq 0$ and $r > 0$ with $(A, Q^{1/2})$ being detectable, where $\sigma_d[v] = \sigma[u] - d(t)$. We seek to minimize $J(x_0, v)$ by using a linear state feedback $v = Kx$. If there is no saturation, it is well-known that the optimal solution for $v$ is given by

$$v^* = -r^{-1}b^TP_0x \tag{11}$$

and the minimum quadratic cost is $x_0^TP_0x_0$, where $P_0 = P_0^T > 0$ is the solution to the following Riccati equation:

$$A^TP_0 + P_0A + Q - r^{-1}P_0bb^TP_0 = 0 \tag{12}$$

In the presence of saturation, the optimal $v$ is difficult to compute. To get around this difficulty, we follow the approach in [12] by parameterizing the controller by using a bound on $v$. More precisely, given $\rho \geq 0$, we restrict the $v$ to be such that

$$|v| \leq \rho + (1 - d_0) \tag{13}$$

That is, $\rho$ acts as an over-saturation bound. It is easy to verify that for any $v$ constrained by (13), $\sigma_d[v]$ lies in the following sector bound as shown in Fig. 3

$$\delta(v) = \sigma_d[v] - \rho_1 v; \quad |\delta(v)| \leq \rho_2 |v| \tag{14}$$

where

$$\rho_1 = \frac{2(1 - d_0) + \rho}{2(1 - d_0 + \rho)} \quad \rho_2 = \frac{\rho}{2(1 - d_0 + \rho)} \tag{15}$$

Now, for a given $\rho > 0$, consider a Lyapunov function candidate of the form $V(x) = x^TP_\rho x$ for some $P_\rho = P_\rho^T$ to be determined and define

$$\Omega_\rho = A^TP_\rho + P_\rho A - r^{-1}P_\rho bb^TP_\rho + Q \tag{16}$$

It is straightforward to verify that

$$J(x_0, v) = \lim_{T \to -\infty} V(x_0) - V(x(T)) + \int_0^\infty (V + x^TQ x + r\sigma_d^2[v])dt \leq V(x_0) + \int_0^\infty \varphi(x, v, \delta(v))dt \tag{17}$$

where

$$\varphi(x, v, \delta(v)) = x^T(A^TP_\rho + P_\rho A + Q)x + \rho\sigma_d^2[v] + 2x^TP_\rho b\sigma_d[v] = x^T\Omega_\rho x + r(\rho_1v + \delta(v) - v^*)^2 \tag{18}$$

This implies that if $\varphi(x, v, \delta(v)) \leq 0$ along the trajectory of the system (9) with the feedback control, we have

$$J(x_0, v) \leq V(x_0) \tag{19}$$

Using the analysis above, we formulate the following (relaxed) auxiliary optimal control problem: For a given $\rho \geq 0$, design $P_\rho$ and $v$ to minimize $V(x_0)$ subject to $\varphi(x, v, \delta(v)) \leq 0$ for all $x \in \mathbb{R}^n$ and all $\delta(\cdot)$ satisfying the following sector bound:

$$|\delta(v)| \leq \rho_2 |v| \tag{20}$$

In addition, determine the largest invariant set of the form

$$X_\rho = \{x : x^TP_\rho x \leq \mu_\rho^2\} \tag{21}$$

such that for any $x_0 \in X_\rho$, $x(t) \in X_\rho$ and $|v(t)| \leq 1 - d_0 + \rho$ for all $t \geq 0$ and $J(x_0, v) \leq V(x_0)$.

**Theorem 1:** Consider the system in (9), the quadratic cost function in (10) and a given bound $\rho \geq 0$ on the level of over-saturation. Define $\rho_0 = \rho_2/\rho_1$. Suppose the equation

$$A^TP_\rho + P_\rho A - r^{-1}(1 - \rho_0^2)P_\rho bb^TP_\rho + Q = 0 \tag{22}$$

has a solution $P_\rho > 0$. Then the optimal $K_\rho$ and the associated $X_\rho$ for the auxiliary optimal control problem are given by

$$K_\rho = -\rho_1^{-1}r^{-1}b^TP_\rho \tag{23}$$

$$\mu_\rho = \frac{2(1 - d_0) + \rho}{2\sqrt{b^TP_\rho b}} \tag{24}$$
We note that the upper bound of the Riccati equation, we know that given the optimal value of \( \rho \), we have the following result:

\[
\max_{x, \rho} |Kx| = |\rho^{-1}r - B^T P_{\rho} x| = 1 - d_0 + \rho
\]

where \( \rho \) is given by

\[
\rho = \frac{\mu_0}{\|P^{1/2}b\|}
\]

Substituting \( x = P_{\rho}^{-1/2} x_0 \) into (24), we obtain (23).

Remark 1: If \( \rho = 0 \), the Riccati equation (21) and the control gain (22) recover the results in (12) and (11) for no saturation. The associated invariant set is given by

\[
X_0 := \{x : x^T P_0 x \leq \mu_0\}, \quad \mu_0 = \frac{2(1 - d_0)r}{2\sqrt{b^T P_0 b}}
\]

IV. Nested Switching Control

Roughly speaking, a practical way to avoid or reduce input saturation is to make the controller gain small when the state is large and to increase the gain when the state is small. Switching control with a sequence of nested control laws is motivated by this basic idea. To do this, a sequence of control laws need to be designed in such a way that the invariant sets are nested. More precisely, we need to find the controllers \( K_i, i = 0, 1, \ldots, N \) such that the corresponding invariant sets \( X_i = \{x : x^T P_i x \leq \mu_i^2\}, i = 0, 1, \ldots, N \), satisfy \( X_0 \subset X_1 \subset \cdots \subset X_N \). Furthermore, we also demand \( P_0 < P_1 < \cdots < P_N \) so that the quadratic performance gets improved gradually when the control gain switch from \( K_N \) to \( K_0 \).

The desired nesting properties above can indeed be achieved by using the design method in the previous section with different values of \( \rho \). This is due to the monotonicity of the Riccati equation (21). Indeed, we can rewrite (21) as

\[
A^T S_{\rho} + S_{\rho} A - r^{-1}(1 + \rho_0)S_{\rho} B B^T S_{\rho} + Q_{\rho} = 0
\]

where \( S_{\rho} = (1 - \rho_0)P_{\rho} \) and \( Q_{\rho} = (1 - \rho_0)Q \). Using \( S_{\rho} \) and \( Q_{\rho} \), the invariant set (20) can be expressed as

\[
X_{\rho} = \left\{ x : x^T P_{\rho} x \leq \frac{(1 - d_0)^2 r^2}{(1 - \rho_0)b^T S_{\rho} B} \right\}
\]

or equivalently,

\[
X_{\rho} = \left\{ x : x^T S_{\rho} x \leq \frac{(1 - d_0)^2 r^2}{b^T S_{\rho} B} \right\}
\]

Theorem 2: The solution \( S_{\rho} \) to equation (28) is monotonically decreasing in \( \rho > 0 \), i.e., \( S_{\rho + \varepsilon} < S_{\rho} \), if \( 0 \leq \rho < \rho + \varepsilon \). Consequently, \( X_{\rho} \) has the following nesting property:

\[
X_{\rho} \subset X_{\rho + \varepsilon}, \quad 0 \leq \rho < \rho + \varepsilon
\]

Similarly, we have

\[
P_\rho < P_{\rho + \varepsilon}, \quad 0 \leq \rho < \rho + \varepsilon
\]

Proof: It suffices to establish the result above for sufficiently small \( \varepsilon > 0 \). It is easy to verify that \( \rho_0 = \rho / 2(1 - \rho_0) \).
\(d_0 + \rho\) is monotonically increasing in \(\rho\). Thus, for any \(\varepsilon > 0\), we can write
\[
\frac{\rho + \varepsilon}{2(1 - d_0) + \rho + \varepsilon} = \rho_0 + \varepsilon_1
\]
with some \(\varepsilon_1 > 0\). It follows that the Riccati equation (28) for \(\rho + \varepsilon\) can be represented as
\[
A^T S_{\rho + \varepsilon} + S_{\rho + \varepsilon} A - r^{-1}(1 + \rho_0 + \varepsilon_1) S_{\rho + \varepsilon} B B^T S_{\rho + \varepsilon} + (1 - \rho_0 - \varepsilon_1) Q = 0
\]
(32)
Let \(E = S_{\rho} - S_{\rho + \varepsilon}\). From (28) and (32), we have
\[
A^T E + E A + r^{-1}(1 + \rho_0 + \varepsilon_1) E B B^T E + \varepsilon_1 Q + \varepsilon_1 S_{\rho} B B^T S_{\rho} = 0
\]
(33)
where \(A = r^{-1}(1 + \rho_0 + \varepsilon_1) B B^T S_{\rho}\). From (28), we know that \(A - r^{-1}(1 + \rho_0) B B^T S_{\rho}\) is Hurwitz. Therefore, \(A - r^{-1}(1 + \rho_0 + \varepsilon_1) B B^T S_{\rho}\) is also Hurwitz when \(\varepsilon_1\) (or equivalently, \(\varepsilon\)) is sufficiently small. It follows from (33) and the detectability of \((A, Q^{1/2})\) that \(E > 0\), i.e. \(S_{\rho} > S_{\rho + \varepsilon}\). The nesting property of \(X_\rho\) then follows from (29).

The monotonicity of \(P_i\) is proved similarly.

Based-on the nesting property of the invariant set \(X_\rho\), we can apply Theorem 1 to design a sequence of control gains \(K_i(= K_{\rho_i}), i = 0, 1, \cdots, N\), for a sequence of over-saturation bounds \(0 = \rho_0 < \rho_1 < \rho_2 < \cdots < \rho_N\), and switch the gain to \(K_i\) whenever \(x \in X_{\rho_i}(= X_{\rho_{i-1}})\) but is outside of \(X_{\rho_{i-1}}(= X_{\rho_{i-2}})\). The next result shows the performance improvement by using this switching control.

**Theorem 3:** Suppose the switching controller above is applied to the system in (9) with \(x_0 \in X_N\). Let \(\tau_i\) be the time instance \(K_i\) is switched on, \(i = 0, 1, \cdots, N\). In particular, \(\tau_N = 0\). Then,
\[
J(x_0, v) = x_0^T P_N x_0 - \sum_{i=0}^{N-1} x^T (\tau_i) \Delta P_i x(\tau_i)
\]
(34)
where \(\Delta P_i = P_{i+1} - P_i > 0, i = 0, 1, \cdots, N - 1\).

**Proof:** Following the proof of Theorem 1, we have
\[
x^T Q x + r \sigma^2_a[k x] \leq \frac{d}{dt} \{x^T P_i x\}
\]
(35)
along the trajectory of \(x(t), t \in [\tau_i+1, \tau_i]\). Integrating the inequality above yields (34).

**V. DESIGN EXAMPLES**

**A. Single-axis Dual-Stage Positioning control**

Recall the dual-stage positioning system in Fig. 2. The control problem is to drive the output \(y\) to a given command \(y_r\). The physical parameters of the system are shown in Table I. The saturation level for the secondary stage is \(a = 0.02\) and the offset between the two stages is \(l_0 = 0.05\). Rescaling the control input and the input matrix as \(0.02 u \rightarrow u\) and \(0.02 B \rightarrow B\), the system can be rewritten as model (1) with the saturation level \(a = 1\). Choose \(Q\) and \(r\) as
\[
Q = \begin{bmatrix} 50 & 0 \\ 0 & 10 \end{bmatrix}, \quad r = 0.1
\]
(36)

In order to show the effectiveness of the nested switching control strategy, we are mainly interested in the control design for the secondary stage. The primary actuator is simply controlled by a PID controller with the gains \(k_p = 50,0, \quad k_f = 5, \quad k_d = 0\).

To design the secondary stage, we take \(d_0 = 0.01\) and choose the over-saturation levels to be \(\rho_4 = 50, \rho_3 = 10, \rho_2 = 5, \rho_1 = 1\) and \(\rho_0 = 0\). This leads to the switching feedback control
\[
K_4 = [-76.0594 - 57.8713]
\]
\[
K_3 = [-50.1572 - 30.7876]
\]
\[
K_2 = [-40.1313 - 23.2691]
\]
\[
K_1 = [-25.0582 - 13.6835]
\]
\[
K_0 = [-17.9129 - 9.6805]
\]
The corresponding domains of attraction have \(\mu_4 = 1.145, \mu_3 = 0.1069, \mu_2 = 0.0449, \mu_1 = 0.0108\) and \(\mu_0 = 0.0051\). The corresponding Lyapunov matrices for \(\mu_4\) and \(\mu_0\), for example, are given by, respectively,
\[
P_4 = \begin{bmatrix} 28.1408 & 1.9384 \\ 1.9384 & 1.4749 \end{bmatrix}, P_0 = \begin{bmatrix} 24.4199 & 0.8956 \\ 0.8956 & 0.4840 \end{bmatrix}
\]

Simulation results are shown in Figs. 4-5 (the left plots) which demonstrate the step response to the set-point command \(y_r = 0.5\). The domains of attraction are also shown in Fig. 5. To compare the performance of the proposed switching controller, we also show the response (the right plots in Figs. 4-5) when the secondary stage is controlled by a P controller with “optimized” gains \(k_p = 0.45, \quad k_i = 0, \quad k_d = 0\). It is clear from the simulations that the proposed switching controller can drive the output \(y\) to \(y_r\) much faster than a PID controller.

**B. Synchronizing control of multi-stages**

Synchronization of two or more mechanical machines is very important when the machines must cooperate. Stepping- and-laser scanning systems are typical machines of this kind widely used in the semiconductor manufacturing industry.
Fig. 6 shows a typical setup of such a system. In this setup, a wafer is placed on the master stage (called wafer stage) and a reticle are placed on a slave stage (called reticle stage). The two stages are required to move in synchrony with high precision along the same axis so that the content in the reticle can be exposed to the wafer. In order to achieve the high precision in synchronization, a dual-stage is usually used for the reticle stage. The control of the secondary stage is targeted at tracking the motion of the wafer-stage.

The parameters of the system used in the simulation are shown in Table II. For simplicity, the parameters of the dual-stage system is chosen to be the same as in Example 1. Consequently, the same switching controller for Example 1 is used here. However, the PID controller for the primary stage is adjusted to have \( k_p = 100, k_i = 15, k_d = 0 \) to improve the synchronizing performance. Two different command signals of \( y_r \), ramp and sine wave, are used in the simulation. The responses of the master stage are shown in Fig. 7 (the plots in the top), and the corresponding synchronization errors are shown in Fig. 7 (the plots in the middle).

To compare the effectiveness of the proposed synchronization approach where the reticle stage is actuated by the dual-stage with the switching controller, we also show in Fig. 7 (the last two plots) the synchronization error when the reticle stage is driven by a single actuator, i.e. the secondary stage is fixed on the primary stage. It is clear from Fig. 7 that the proposed switching controller performs much better.

VI. CONCLUSIONS

For dual-actuator systems, an effective way to achieve fine control quality is to design the control strategy with delicate sensitivity, due to the primary and the secondary actuator have different range and different response properties. Particularly, the secondary actuator is usually saturated and of small range. In this paper, we have shown that a feasible method to this end is to focus the design arm on the secondary actuator, since the control can be decoupled by a simple coordinate change. We proposed a nested switching control approach for the secondary actuator which takes the saturation and the interaction from the primary actuator into account.

REFERENCES