Addendum to: An Upper Bound on the Error of Alignment-Based Transfer Learning Between Two Linear, Time-Invariant, Scalar Systems

Kaizad V. Raimalwala, Bruce A. Francis, and Angela P. Schoellig

Abstract—This short paper provides a derivation of the minimized $H_{\infty}$-norm of a transfer system as introduced in [1]. This system consists of two linear, time-invariant, single-input, single-output systems tasked to follow the same reference signal. The $H_{\infty}$-norm of this transfer system gives the least upper bound on the 2-norm of the transformation error, which is defined as the difference between the output of a target system and a weighted output of a source system. This paper provides the proofs of Lemma 1 and Theorem 1 in [1].

I. PROBLEM FORMULATION

Consider two first-order, linear, time-invariant (LTI), single-input, single-output (SISO) systems $S_1$ and $S_2$, whose transfer functions are given by

$$G_1(s) = \frac{k_1}{s + a_1}, \quad G_2(s) = \frac{k_2}{s + a_2},$$

where $-a_1$ and $-a_2$ are the poles, and $k_1$ and $k_2$ are the gains of $G_1$ and $G_2$ (see Fig. 1). The quantity of interest in the TL problem is the error in the estimation of $x_1(t)$ and is the output of the transfer system,

$$e_A(t) = x_1(t) - \alpha x_2(t),$$

where $\alpha$ is a constant scalar that is applied to $x_2(t)$ to estimate $x_1(t)$. The transfer function from $d(t)$ to $e_A(t)$ is

$$G_A(s) = \frac{k_1}{s + a_1} - \alpha \frac{k_2}{s + a_2}. \quad (4)$$

To assure that $G_A(s)$ is asymptotically stable, $a_1$ and $a_2$ are assumed to be positive. Furthermore, $k_2$ is assumed to be non-zero to avoid the degenerate case where $G_A = G_1$.

Design Criterion. The signal 2-norm is chosen as a measure for the signal $e_A(t)$, and is denoted by $\|:\|_2$. This measure can be determined for a specific reference signal $d(t) \in L_2[0, \infty)$, where $L_2[0, \infty)$ denotes the set of all signals that have finite energy on an infinite time interval $[0, \infty)$. However, the $H_{\infty}$-norm of $G_A$ provides the least upper bound on $\|e_A\|_2$ for all $d(t) \in D := \{d(t) : \|d\|_2 \leq 1\}$, as shown in [2]; that is,

$$\|G_A\|_\infty = \sup \{\|e_A\|_2 : d(t) \in D\}, \quad (5)$$

where the $H_{\infty}$-norm of $G_A$ is defined as

$$\|G_A\|_\infty := \sup_\omega |G_A(j\omega)|. \quad (6)$$

Definition 1. The transfer problem is formulated as minimizing $\|G_A\|_\infty^2$ with respect to $\alpha$:

$$\alpha^* := \arg\min_\alpha \|G_A\|_\infty^2. \quad (7)$$

II. AN UPPER BOUND ON THE ERROR 2-NORM

A. An Analytic Expression of the $H_{\infty}$-Norm of $G_A$

In this subsection, we derive an analytic expression for $\|G_A\|_\infty^2$ as a function of $\alpha$, $a_1$, $a_2$, $k_1$, and $k_2$, and prove the following lemma:

Lemma 1. For $G_A$ in (4), $\|G_A\|_\infty^2$ is a piecewise continuous function with respect to $\alpha$ that maximizes $|G_A(j\omega, \alpha)|^2$ with respect to $\omega$ for all $a_1, a_2 > 0$, $k_1$, and $k_2 \neq 0$. It is given by

$$\gamma_A^2(\alpha) := \|G_A\|_\infty^2 = \begin{cases} \phi(\alpha) & \text{if } \alpha_2 < \alpha < \alpha_1 \\ \psi(\alpha) & \text{otherwise} \end{cases}. \quad (8)$$

Proof. The squared magnitude of $G_A(j\omega, \alpha)$ is

$$|G_A(j\omega, \alpha)|^2 = \frac{\lambda_1(\alpha)\omega^2 + \lambda_2(\alpha)}{\omega^4 + \lambda_1\omega^2 + \lambda_5}, \quad (9)$$

where

$$\lambda_1(\alpha) = (k_1 - k_2 \alpha)^2, \quad (10)$$
$$\lambda_2(\alpha) = (k_1 a_2 - k_2 a_1 \alpha)^2. \quad (11)$$

Fig. 1. In this control block diagram, $x_2(t)$ is multiplied by a scalar $\alpha$ to match $x_1(t)$. While $x_1(t)$ and $x_2(t)$ are outputs of sub-systems $S_1$ and $S_2$, the output of the overall system is $e_A(t)$.
\[ \lambda_4 = a_1^2 + a_2, \]
\[ \lambda_5 = a_1^2 a_2. \]

Note that all four \( \lambda \) parameters are non-negative, and that \( \lambda_1 \) and \( \lambda_2 \) are functions of \( \alpha \). To reduce clutter, \( \lambda_1(\alpha) \) and \( \lambda_2(\alpha) \) are denoted by \( \lambda_1 \) and \( \lambda_2 \) in the remainder of the paper. If \( a_1, a_2 > 0 \), then for all \( \alpha \in \mathbb{R} \),
\[
\lim_{\omega \to \pm \infty} |G_A(j \omega, \alpha)|^2 = 0. \tag{14}
\]

Let the frequency that maximizes the squared magnitude for a given value of \( \alpha \) be
\[
\omega^*(\alpha) = \arg \max_{\omega} |G_A(j \omega, \alpha)|^2. \tag{15}
\]

The maximum of \( |G_A(j \omega, \alpha)|^2 \) can be obtained by finding the roots of the derivative of \( |G_A(j \omega, \alpha)|^2 \) with respect to \( \omega \),
\[
\frac{\partial |G_A(j \omega, \alpha)|^2}{\partial \omega} = 0 \tag{16}
\]

\[
\Leftrightarrow -2 \omega \left( \lambda_1 \omega^4 + 2 \lambda_2 \omega^2 + \lambda_2 \lambda_4 - \lambda_1 \lambda_5 \right) = 0 \tag{17}
\]

\[
\Leftrightarrow \omega \left( \lambda_1 \omega^4 + 2 \lambda_2 \omega^2 + \lambda_2 \lambda_4 - \lambda_1 \lambda_5 \right) = 0 \tag{18}
\]

\[
\Rightarrow \omega \left( \frac{1}{4} \omega^4 + p(\alpha) \omega^2 + q(\alpha) \right) = 0, \tag{19}
\]

where
\[
p(\alpha) = \frac{\lambda_2}{2 \lambda_1}, \tag{20}
\]
\[
q(\alpha) = \frac{\lambda_2 \lambda_4 - \lambda_1 \lambda_5}{4 \lambda_1}. \tag{21}
\]

In (18), the equation is divided by \( 4 \lambda_1 \) to obtain a standard form of the quartic term in (19), whose roots are known functions of \( p(\alpha) \) and \( q(\alpha) \) (see the Appendix). Note that \( p(\alpha) \) is always non-negative, whereas \( q(\alpha) \) can be negative. Case A2 in the Appendix is not possible since if \( p(\alpha) = 0 \), then \( q(\alpha) = -\lambda_5/4 < 0 \), contradicting case A2.

To find \( \max_{\omega} |G_A(j \omega, \alpha)|^2 \), the real roots of the polynomial in (19) need to be found. Therefore, there are two cases of interest:

**Case 1:** This case corresponds to cases A1 and A3 from the Appendix, and considers \( q(\alpha) \geq 0 \). In this case, the only real root of (19) is 0. After verifying that \( \omega = 0 \) is a local maximum of \( |G_A(j \omega, \alpha)|^2 \) with the second derivative, we obtain \( \omega^* = 0 \) for all \( \alpha \) due to (14). Therefore,
\[
|G_A|_\infty^2 = |G_A(0j, \alpha)|^2 = \frac{\lambda_2}{\lambda_5} \tag{22}
\]
\[
= \frac{(k_1 a_2 - k_2 a_1 \alpha)^2}{a_1^2 a_2^2} := \psi(\alpha). \tag{23}
\]

As a result, \( |G_A|_\infty^2 \) is a quadratic function of \( \alpha \).

**Case 2:** This case corresponds to case A4 from the Appendix, and considers \( q(\alpha) < 0 \). In this case, there are three real roots. In addition to the real root \( \omega_1 = 0 \), the quartic term has two real roots \( \pm \omega_2 \). Since there are three real roots with non-zeros second derivative and because of (14), the vertices at \( \pm \omega_2 \) must be maxima and the vertex at \( \omega_1 = 0 \) must be a minimum. Therefore,
\[
(\omega^*(\alpha))^2 = \frac{\omega_2^2}{2} \tag{24}
\]
\[
= 2 \sqrt{p^2(\alpha) - q(\alpha) - 2p(\alpha)} \tag{25}
\]
\[
= \sqrt{\lambda_2^2 - \lambda_1 \lambda_2 \lambda_4 + \lambda_2^2 \lambda_5 - \lambda_2} \tag{26}
\]

Evaluating (9) at \( \omega^2 = \omega_2^2 \) results in
\[
|G_A|_\infty^2 = |G_A(\pm \omega_2, \alpha)|^2 \tag{27}
\]
\[
= \frac{\lambda_2^2}{g(\alpha) + 2 \sqrt{f(\alpha)}} := \phi(\alpha), \tag{28}
\]

where
\[
f(\alpha) := \lambda_5 \lambda_1^2 - \lambda_4 \lambda_1 \lambda_2 + \lambda_2^2, \tag{29}
\]
\[
g(\alpha) := \lambda_5 \lambda_1 - 2 \lambda_2. \tag{30}
\]

In this case, \( |G_A|_\infty^2 \) is a nonlinear function of \( \alpha \).

The last step in proving Lemma 1 is to re-work the conditions in Case 1 and Case 2, which are expressed in terms of \( q(\alpha) \) and not \( \alpha \).

We first consider the special case where \( a_2 = a_1 = a \). Then \( q(\alpha) = a^4/4 \). In this case, \( q(\alpha) > 0 \) for all \( \alpha \) and according to (24),
\[
|G_A|_\infty^2 = \frac{(k_1 - \alpha k_2)^2}{a^2}. \tag{31}
\]

Since (32) is quadratic in \( \alpha \), a unique minimizing \( \alpha \) exists:
\[
\alpha = \frac{k_1}{k_2}, \tag{33}
\]

which is the ratio of the system gains.

When \( a_2 \neq a_1 \), \( q(\alpha) \) can be negative. To obtain a solution, we find the roots of \( q(\alpha) \) by solving
\[
0 = q(\alpha) \tag{34}
\]

\[
\Leftrightarrow 0 = \lambda_2 \lambda_4 - \lambda_1 \lambda_5 \tag{35}
\]

\[
\Leftrightarrow 0 = \left( a_1^2 k_2^2 \right) \alpha^2 - (2 a_1 a_2 k_1 k_2 (a_1^2 - a_1 a_2 + a_2^2)) \alpha + (a_1^4 k_1^2). \tag{36}
\]

We obtain two real roots:
\[
\alpha_1 = \frac{k_1 a_2}{k_2 a_1^2} (a_1^2 + a_2^2 - a_1 a_2 + \eta), \tag{37}
\]
\[
\alpha_2 = \frac{k_1 a_2}{k_2 a_1^2} (a_1^2 + a_2^2 - a_1 a_2 - \eta), \tag{38}
\]

with
\[
\eta = \sqrt{(a_1 - a_2)^2 (a_1^2 + a_2^2)} \tag{39}
\]

It is clear that \( \alpha_2 < \alpha_1 \) for all \( a_1, a_2 > 0 \), \( k_1 \), and \( k_2 \neq 0 \). When \( \alpha \leq \alpha_2 \) or \( \alpha \geq \alpha_1 \), \( q(\alpha) \geq 0 \) and \( |G_A|_\infty^2 = \psi(\alpha) \).

When \( \alpha_2 < \alpha < \alpha_1 \), \( q(\alpha) < 0 \) and \( |G_A|_\infty^2 = \phi(\alpha) \). To summarize, \( |G_A|_\infty^2 \) and \( \omega^*(\alpha) \) are piecewise functions of \( \alpha \).
and are given by
\[ \gamma_A^2(\alpha) := \|G_A\|^2_\infty = \begin{cases} 
\phi(\alpha) & \text{if } \alpha_2 < \alpha < \alpha_1, \\
\psi(\alpha) & \text{otherwise,}
\end{cases} \] (40)
\[
\omega^*(\alpha) = \begin{cases} 
\pm\omega_2 & \text{if } \alpha_2 < \alpha < \alpha_1, \\
0 & \text{otherwise,}
\end{cases}
\] (41)
with
\[
\omega_2 := \sqrt{2\sqrt{p^2(\alpha) - q(\alpha)} - 2p(\alpha)}.
\]

To prove continuity of \(\gamma_A(\alpha)\), it must be shown that
(i) \(\psi(\alpha)\) is continuous in the intervals \((-\infty, \alpha_2]\) and \([\alpha_1, \infty)\),
(ii) \(\phi(\alpha)\) is continuous in the open interval \((\alpha_2, \alpha_1)\),
(iii) and that \(\phi(\alpha_1) = \psi(\alpha_1)\) and \(\phi(\alpha_2) = \psi(\alpha_2)\)
for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\).

The first condition is true because \(\psi(\alpha)\) is a polynomial and is thus continuous over its domain.

Secondly, the function \(\phi(\alpha)\) is continuous under the following two conditions:

1) We first make sure that the square-root term in (29) is well-defined:
\[
0 \leq f(\alpha) = k_1k_2(a_1 - a_2)^2 \left[ (2a_1k_2^2(a_1 + a_2))\right] \alpha^2 \\
+ (-k_1k_2(a_1^2 + 6a_1a_2 + a_2^2))\alpha \\
+ 2a_1k_2^2(a_1 + a_2) \alpha.
\] (42)
The inequality is true when \(\alpha \geq \alpha_{f,1}\) and \(0 \leq \alpha \leq \alpha_{f,2}\), where
\[
\alpha_{f,1} = \frac{k_1}{k_2} \frac{(a_1 + a_2)}{2a_1},
\] (44)
\[
\alpha_{f,2} = \frac{k_1}{k_2} \frac{2a_1}{a_1 + a_2}.
\] (45)
It can be shown that \(\alpha_{f,1} > \alpha_{f,2}\) for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\). To show that the inequality in (43) is true over the interval \((\alpha_2, \alpha_1)\), it can be shown that either \(\alpha_{f,2} > \alpha_1\) or that \(\alpha_{f,1} < \alpha_2\); that is, \(\alpha_{f,2}, \alpha_{f,1} \notin (\alpha_2, \alpha_1)\).

2) The second condition is that
\[
g(\alpha) + 2\sqrt{f(\alpha)} \neq 0.
\] (46)
To satisfy (46), \(\alpha \neq \alpha_0\) with \(\alpha_0 := k_1k_2^{-1}\) because
\[
0 = f(\alpha) - \frac{g^2(\alpha)}{4} = -\frac{(a_1^2 - a_2^2)^2(k_1 - \alpha k_2)^4}{4}.
\] (47)

It can be shown that either \(\alpha_0 < \alpha_2\), or \(\alpha_0 > \alpha_1\), for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\). Therefore, \(\phi(\alpha)\) is continuous in the open interval \((\alpha_2, \alpha_1)\).

The third condition can be shown to be true by evaluating \(\phi(\alpha)\) and \(\psi(\alpha)\) at \(\alpha_1\) and \(\alpha_2\). We used analytic simplification techniques in MATLAB to determine that for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\), \(\phi(\alpha_1) = \psi(\alpha_1)\) and \(\phi(\alpha_2) = \psi(\alpha_2)\).

\[ \square \]

B. Minimizing the \(H_\infty\)-Norm of \(G_A\) with Respect to the Transformation Parameter

In this subsection, we minimize \(\gamma_A^2(\alpha)\) with respect to \(\alpha\) to prove the following theorem:

**Theorem 1.** For \(G_A\) in (4), the parameter \(\alpha\) that minimizes \(\gamma_A^2(\alpha)\) for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\) is given by
\[
\alpha^* = \frac{k_1}{k_2} \frac{2\alpha_2}{(4a_1 + a_2 - \sqrt{8a_1^2 + a_2^2})}.
\] (49)

**Proof.** To find the minimum of \(\gamma_A^2(\alpha)\), one possibility is to find the minimum of \(\phi(\alpha)\) for all \(\alpha \in (\alpha_2, \alpha_1)\) and the minimum of \(\psi(\alpha)\) for all \(\alpha \notin (\alpha_2, \alpha_1)\), and then compare the two. However, it can be shown that \(\alpha_2 < \alpha^* < \alpha_1\) for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\), thereby limiting the search for the minimum to \(\phi(\alpha)\). To prove that \(\alpha_2 < \alpha^* < \alpha_1\), one can use Bolzano’s Theorem [3], which considers the derivative of \(\phi(\alpha)\) with respect to \(\alpha\) denoted by \(\phi'(\alpha)\) and states: if \(\phi'(\alpha)|_{\alpha=\alpha_2} < 0\) and \(\phi'(\alpha)|_{\alpha=\alpha_1} > 0\), then \(\phi(\alpha)\) has a minimum in the interval \((\alpha_2, \alpha_1)\). MATLAB can be used to check the aforementioned conditions on \(\phi'(\alpha)\) at the points \(\alpha_1\) and \(\alpha_2\). However, it is easier to validate that \(\phi'(\alpha) = \psi'(\alpha)\) at \(\alpha_1\) and \(\alpha_2\), and use the fact that \(\psi(\alpha)\) satisfies the above conditions. The function \(\psi(\alpha)\) is a convex parabola with a single minimum at
\[
\alpha_{\psi} = \frac{k_1}{k_2} \frac{a_2}{a_1}.
\] (50)
It is sufficient to show that \(\alpha_2 < \alpha_{\psi} < \alpha_1\) for all \(\alpha_1, \alpha_2 > 0\), \(k_1\), and \(k_2 \neq 0\).

It now remains to find \(\alpha^*\) by solving \(\phi'(\alpha) = 0\). The derivative \(\phi'(\alpha)\) is
\[
\phi'(\alpha) = -\frac{2k_2}{(2\sqrt{f(\alpha) - \lambda_2} + \lambda_4 \lambda_3)^2} z_1(\alpha),
\] (51)
with
\[
z_1(\alpha) = h^3(\alpha) \left[ 4\sqrt{f(\alpha)} + 2g(\alpha) + h(\alpha)z_2(\alpha) \right],
\] (52)
\[
z_2(\alpha) = m(\alpha) + \frac{n(\alpha)}{\sqrt{f(\alpha)}},
\] (53)
where
\[
h(\alpha) = \sqrt{\lambda_1},
\] (54)
\[
m(\alpha) = \frac{(a_2 - a_1)}{(k_2(a_1 + a_2) - k_1(a_2 - a_1))},
\] (55)
\[
n(\alpha) = k_1(a_2 - a_1)^2 ((a_1 + a_2)(3a_1k_2^2a_2^2 + a_2k_2^2) - k_1k_2(a_1^2 + 6a_1a_2 + a_2^2) a).\]

Besides the three roots at \(\alpha_0 = k_1k_2^{-1}\) found from the term \(h^3(\alpha)\) in (51), the other roots of \(\phi'(\alpha)\) are found by solving
\[
0 = 4\sqrt{f(\alpha)} + 2g(\alpha) + h(\alpha)z_2(\alpha).
\] (56)
As it was previously shown that \( f(\alpha) \geq 0 \) in the interval \((\alpha_2, \alpha_1)\), further modifications of (56) result in a new equation to solve for \( \alpha \):

\[
f(\alpha) = \left( \frac{4f(\alpha) + n(\alpha)h(\alpha)}{2g(\alpha) + m(\alpha)h(\alpha)} \right)^2 
\]

\[
\Leftrightarrow 0 = f(\alpha) (2g(\alpha) + m(\alpha)h(\alpha))^2 - (4f(\alpha) + n(\alpha)h(\alpha))^2
\]

\[
\Leftrightarrow 0 = (k_1 - k_2\alpha)^5 (2k_2^2a_1(a_1 + a_2)\alpha^2 - a_2k_1k_2(4a_1 + a_2)\alpha + a_1^2k_1^2)
\]

In addition to the several roots at \( \alpha_0 \), the quadratic expression in (59) yields two real roots:

\[
\alpha_{\phi,1} = \frac{k_1}{k_2} \frac{2a_2}{4a_1 + a_2 - \sqrt{8a_1^2 + a_2^2}}
\]

\[
\alpha_{\phi,2} = \frac{k_1}{k_2} \frac{2a_2}{4a_1 + a_2 + \sqrt{8a_1^2 + a_2^2}}
\]

One can show that \( \alpha_2 < \alpha_{\phi,1} < \alpha_1 \) and \( \phi'(\alpha)|_{\alpha=\alpha_{\phi,1}} = 0 \) for all \( a_1, a_2 > 0, k_1, \) and \( k_2 \neq 0 \). The same is not true for \( \alpha_{\phi,2} \), which represents a degenerate case obtained by squaring (56). Therefore, \( \alpha^* = \alpha_{\phi,1} \).

**APPENDIX**

Consider the general biquadratic equation,

\[
\frac{1}{4}x^4 + px^2 + q = 0.
\]

The discriminant of this equation is given by

\[
D = q(p^2 - q).
\]

The four roots of the biquadratic equation are given in terms of \( p \) and \( q \) for four cases:

**Case A1**: For \( q > p^2 \), there are four complex roots,

\[
x_{1,2,3,4} = \pm \sqrt{q - p} \pm j\sqrt{q + p}.
\]

**Case A2**: For \( p \leq 0 \leq q \leq p^2 \), there are four real roots,

\[
x_{1,2,3,4} = \pm \sqrt{q - p} \pm \sqrt{-q - p}.
\]

**Case A3**: For \( p > 0 \) and \( 0 \leq q \leq p^2 \), there are four roots with zero real part,

\[
x_{1,2,3,4} = j(\pm \sqrt{p^2 + q} \pm \sqrt{p - q}).
\]

**Case A4**: For \( q < 0 \), there are two real roots and two roots with zero real part,

\[
x_{1,2} = \pm \sqrt{2p^2 - q - 2p},
\]

\[
x_{3,4} = \pm j\sqrt{2p^2 - q + 2p}.
\]

**REFERENCES**

