

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—Methods from machine learning have successfully been used to improve the performance of control systems in cases when accurate models of the system or the environment are not available. These methods require the use of data generated from physical trials. Transfer Learning (TL) allows for this data to come from a different, similar system. This paper studies a simplified TL scenario with the goal of understanding in which cases a simple, alignment-based transfer of data is possible and beneficial. Two linear, time-invariant (LTI), single-input, single-output systems are tasked to follow the same reference signal. A scalar, LTI transformation is applied to the output from a source system to align with the output from a target system. An upper bound on the 2-norm of the transformation error is derived for a large set of reference signals and is minimized with respect to the transformation scalar. Analysis shows that the minimized error bound is reduced for systems with poles that lie close to each other (that is, for systems with similar response times). This criterion is relaxed for systems with poles that have a larger negative real part (that is, for stable systems with fast response), meaning that poles can be further apart for the same minimized error bound. Additionally, numerical results show that using the reference signal as input to the transformation reduces the minimized bound further.

I. INTRODUCTION

Models that accurately represent dynamic systems are often difficult to derive mathematically, or are unreliable due to parameter uncertainties or unknown external disturbances affecting the system. The potential use of data from simulations or experiments to improve models of dynamic systems is a prime motivator for modern research at the intersection of machine learning and control theory. Various regression techniques have been used to learn dynamics, kinematics and disturbance models using data from physical trials or simulations of the robot [1]–[4].

Transfer Learning (TL) allows for this data to be generated by a second system. In a training phase, both systems generate data, and a transformation that aligns one dataset to the other is learned. Once this mapping is learned, the first system, also called the target system, can use data generated by the second system, the source system, in subsequent model learning (see Fig. 1). This transfer of data may be beneficial to the model learning process if the source system is less costly, difficult, or hazardous to operate than the target system.

K.V. Raimalwala and A.P. Schoellig are with the Dynamic Systems Lab (www.dynsyslab.org) at the University of Toronto Institute for Aerospace Studies (UTIAS), Canada. Email: kaizad.raimalwala@mail.utoronto.ca, schoellig@utias.utoronto.ca

B.A. Francis is with the Department of Electrical and Computer Engineering, University of Toronto, Canada. Email: bruce.francis@utoronto.ca

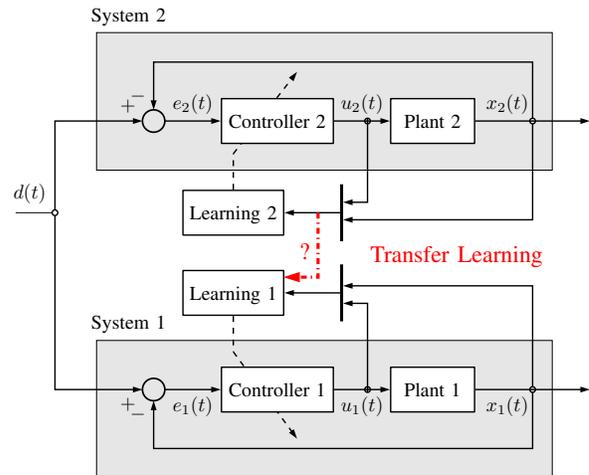


Fig. 1. Transfer Learning framework. Systems 1 and 2 learn from input-output data (dashed lines). The concept of Transfer Learning (TL) allows for System 1 to use data from System 2 for its own learning task (red dash-dotted line). In this paper, we study TL from a system-theoretic perspective and provide insight to when TL is beneficial.

In this paper, we study when such a data transfer is beneficial. We investigate TL for two simple, linear, time-invariant (LTI), single-input, single-output (SISO) systems. We use an H_∞ -norm minimization framework to define the quality of the data transfer and analyze when the transfer works well and when it does not.

In robotics, TL has often been considered in the context of speeding up a single robot’s task learning using knowledge obtained in a previous task performed by the same robot. TL has, for example, been successfully used in an Iterative Learning Control (ILC) framework to speed up learning of new tasks [5]–[7]. Research for multi-agent robotic systems is relatively sparse [8]. Most common applications aim to speed up joint or sequential learning either in an ILC framework by transferring task-dependent disturbance estimates [9], or in a Reinforcement Learning framework by transferring rules and policies of simple systems with discrete states [10]–[12]. However, TL can also be used in a model learning framework to apply a transformation on input-output data generated by one robot. This transformed data can then be input to a model learning algorithm for a second, similar robot [13]. We are interested in the latter multi-agent learning scenario.

In several applications beyond robotics, Manifold Alignment has been used to find an optimal transformation to align datasets [14], [15]. In [16], this technique is demonstrated for two simple state-space models that each need to learn

a task using Reinforcement Learning. In [13], a similar transformation technique is used on data from one robotic arm to speed up learning of a robot model for a second robotic arm.

While these works have shown that TL is feasible for some examples, they do not address the question of when the data transfer works well and when it fails. This paper provides insight on how the quality of the transformation depends on the system properties of the source and target systems. We study two first-order, LTI, SISO systems tasked to follow the same reference signal and show that the transfer quality depends on the location of the poles (i.e., the stability properties) of both systems.

Despite the simplicity of the systems we choose to study, this paper offers four novel results. Firstly, we study the feasibility of TL for control systems from a system-theoretic perspective by formulating the problem as an H_∞ -norm minimization problem. Secondly, we derive an upper bound on the 2-norm of the transformation error for a large set of reference signals and minimize this bound with respect to the transformation parameter. Thirdly, we analyze how this upper bound depends on the poles and the gains of the two systems, which provides insight into which configurations work best for TL in this simple scenario. Lastly, we analyze how these results change when the reference signal is used as an additional input to the transformation.

The remainder of the paper is structured as follows: Section II provides a background on model learning and TL for control systems. Section III presents the problem formulation considered in this paper. In Section IV, an upper bound on the 2-norm of the transformation error is derived and minimized with respect to the transformation parameters. In Section V-A, our analysis illustrates how this bound depends on the systems' poles and gains. In Section V-B, TL is demonstrated for a pair of example systems with proportional feedback control. Lastly, Section VI provides concluding remarks.

II. BACKGROUND

A. Model Learning

Dynamics and kinematics models govern a robot's behavior. While analytic models can be derived from first principles, they often do not capture the real-world dynamics accurately [2], [3]. Supervised model learning presents a solution by employing a regression tool to find a map from input data to labelled observations. Given a sequence of input-output data with input $\mathbf{x}[k] \in \mathbb{R}^n$ and output $\mathbf{y}[k] \in \mathbb{R}^p$, where $k \in \{0, 1, 2, \dots, N\}$ and N is the number of samples obtained, the problem of model learning is to find a map $\mathcal{M} : \mathbf{f}(\mathbf{x}) \rightarrow \mathbf{y}$ such that some measure of the magnitude of the error sequence,

$$\epsilon[k] = \mathbf{y}[k] - \mathbf{f}(\mathbf{x}[k]), \quad k \in \{0, 1, 2, \dots, N\}, \quad (1)$$

is minimized. For example, learning an inverse dynamics model for a robot arm can be formulated as finding a map $\mathcal{M} : \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}) \rightarrow \boldsymbol{\tau}$, where $\mathbf{q} \in \mathbb{R}^p$ is a vector of joint

angles for all p joints of the arm, $\boldsymbol{\tau} \in \mathbb{R}^p$ is a vector of applied torques to each joint and $\mathbf{x} \in \mathbb{R}^{3p}$ [17].

B. Transfer Learning

When two robots (or control systems, in general), S_1 and S_2 , execute a task, data is generated by each system. Data from S_2 can then undergo a transformation to align with data from S_1 . The problem is akin to model learning in that a map needs to be found. In [13] and [16], the authors model this map as a time-invariant, linear transformation for each data sample. We make the same assumption.

Let vectors $\mathbf{x}_1[k]$ and $\mathbf{x}_2[k]$ be sampled data from S_1 and S_2 . We thus define the problem of TL as finding a matrix \mathbf{A} such that the vector 2-norm of

$$\epsilon[k] = \mathbf{x}_1[k] - \mathbf{A}\mathbf{x}_2[k] \quad (2)$$

is minimized for all times $k \in \{0, 1, 2, \dots, N\}$. The vector \mathbf{x} can consist of system states, control inputs, or other variables that are relevant for a specific model learning algorithm. For the inverse dynamics model learning example in [13], the vector \mathbf{x} is defined as $\mathbf{x} = [\mathbf{q}^T, \dot{\mathbf{q}}^T, \ddot{\mathbf{q}}^T, \mathbf{u}^T]^T$. Once such a matrix is learned from one pair of datasets, additional training data for learning the model of S_1 can be obtained by transforming subsequent data collected from S_2 using \mathbf{A} .

To find an optimal transformation that aligns the discrete datasets, *a priori* models of each system need not be known, as the transformation only depends on data collected from physical trials of the two systems. The disadvantage of this data-alignment technique is that it is difficult to make predictions on the quality of the transformation. Furthermore, there usually is no guarantee on the performance of a given transformation on subsequent data.

Work in [13] shows that for two simulated robot arms, the data alignment worked well and sped up model learning. However, it is not obvious that the same approach works in other applications. Our work is motivated by an interest to further explore the properties of a time-invariant, linear transformation for control system data, and determine when TL is most beneficial and when it fails.

We therefore consider two first-order, LTI, SISO systems, S_1 and S_2 . We study TL in continuous time to facilitate our analysis. In this framework, $x_1(t), x_2(t) \in \mathbb{R}$ are the scalar system states, which are driven by the reference signal $d(t)$, and $\mathbf{A} = \alpha$ is a scalar that must map $x_2(t)$ to $x_1(t)$. We find the transformation α^* that optimally aligns $x_2(t)$ to $x_1(t)$ for all $t \in [0, \infty)$ in an H_∞ -norm sense. We call this approach the Output Transfer (OT) case.

In a second scenario, which we call the Input-Output Transfer (IOT) case, we study the case when the common reference signal $d(t)$ is used in the transformation; that is, $\mathbf{x}_2(t) = [x_2(t), d(t)]^T$ and $\mathbf{A} = [\beta_1, \beta_2]$. A transformation $\{\beta_1^*, \beta_2^*\}$ is numerically found to optimally align a linear combination of $x_2(t)$ and $d(t)$ to $x_1(t)$ for all $t \in [0, \infty)$ in an H_∞ -norm sense.

In both cases, we derive an upper bound on the 2-norm of the transformation error signal and minimize it with respect to the transformation parameters. We show that for

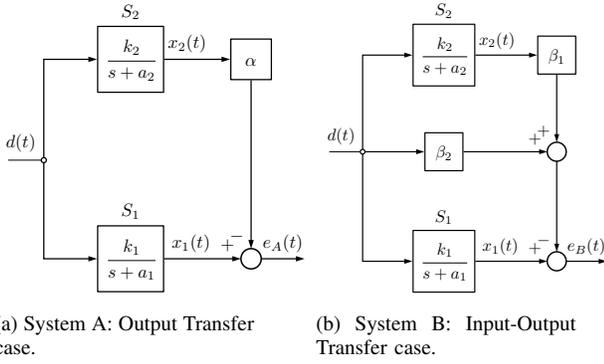


Fig. 2. On the left, $x_2(t)$ is multiplied by a scalar α to match $x_1(t)$. While $x_1(t)$ and $x_2(t)$ are outputs of sub-systems S_1 and S_2 , the output of the overall system is $e_A(t)$. On the right, a linear combination of $x_2(t)$ and $d(t)$ is used to match $x_1(t)$, and the output of the overall system is $e_B(t)$.

these simple systems, the minimized upper bound strongly depends on the poles of S_1 and S_2 , and that systems with a fast response have an advantage over systems with a slow response. We further show that the IOT case reduces the minimized error bound.

III. PROBLEM FORMULATION

In this section, we introduce a framework for analyzing TL for simple, linear control systems and define the H_∞ -norm minimization problem.

Consider two first-order, LTI, SISO systems S_1 and S_2 , whose transfer functions are given by

$$G_1(s) = \frac{k_1}{s + a_1}, \quad G_2(s) = \frac{k_2}{s + a_2}, \quad (3)$$

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. 2). 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,

$$e_A(t) = x_1(t) - \alpha x_2(t), \quad (4)$$

$$e_B(t) = x_1(t) - (\beta_1 x_2(t) + \beta_2 d(t)), \quad (5)$$

where α is a constant scalar that is applied to $x_2(t)$ to estimate $x_1(t)$ in the Output Transfer (OT) case (see Fig. 2a), and likewise $\{\beta_1, \beta_2\}$ are two constant scalars in the Input-Output Transfer (IOT) case (see Fig. 2b). The transfer functions from $d(t)$ to $e_A(t)$ and from $d(t)$ to $e_B(t)$ are

$$G_A(s) = \frac{k_1}{s + a_1} - \alpha \frac{k_2}{s + a_2}, \quad (6)$$

$$G_B(s) = \frac{k_1}{s + a_1} - \beta_1 \frac{k_2}{s + a_2} - \beta_2. \quad (7)$$

To assure that $G_A(s)$ and $G_B(s)$ are 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 signals $e_A(t)$ and $e_B(t)$, and is denoted by $\|\cdot\|_2$. This measure can be determined for a specific reference signal $d(t) \in \mathcal{L}_2[0, \infty)$, where $\mathcal{L}_2[0, \infty)$ denotes

the set of all signals that have finite energy on an infinite time interval $[0, \infty)$. However, the H_∞ -norm of G_A (and likewise for G_B) provides the least upper bound on $\|e_A\|_2$ for all $d(t) \in \mathcal{D} := \{d(t) : \|d\|_2 \leq 1\}$, as shown in [18]; that is,

$$\|G_A\|_\infty = \sup\{\|e_A\|_2 : d(t) \in \mathcal{D}\}, \quad (8)$$

where the H_∞ -norm of G_A is defined as

$$\|G_A\|_\infty := \sup_{\omega} |G_A(j\omega)|. \quad (9)$$

Definition 1. The transfer problem is formulated as minimizing $\|G_A\|_\infty^2$ with respect to α , and likewise, $\|G_B\|_\infty^2$ with respect to $\{\beta_1, \beta_2\}$:

$$\alpha^* := \arg \min_{\alpha} \|G_A\|_\infty^2, \quad (10)$$

$$\{\beta_1^*, \beta_2^*\} := \arg \min_{\{\beta_1, \beta_2\}} \|G_B\|_\infty^2. \quad (11)$$

The H_∞ -norm is useful in analyzing the properties of TL for a large set of reference signals. Assuming that any signal in \mathcal{D} is a potential reference signal, the optimal transformation α^* represents the best possible transformation that would be obtained when observing the system for an infinite amount of time under all possible reference inputs $d(t) \in \mathcal{D}$. Consequently, as long as the reference signal belongs to the set \mathcal{D} , the H_∞ -norm evaluated at α^* provides the worst possible transformation error we could get.

IV. AN UPPER BOUND ON THE ERROR 2-NORM

In this section, we derive an analytic expression for $\|G_A\|_\infty^2$ as a function of α , a_1 , a_2 , k_1 , and k_2 , and find $\min_{\alpha} \|G_A\|_\infty^2$ as a function of a_1 , a_2 and k_1 .

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

$$|G_A(j\omega, \alpha)|^2 = \frac{\lambda_1(\alpha)\omega^2 + \lambda_2(\alpha)}{\omega^4 + \lambda_4\omega^2 + \lambda_5}, \quad (12)$$

where

$$\lambda_1(\alpha) = (k_1 - k_2\alpha)^2, \quad \lambda_4 = a_1^2 + a_2^2,$$

$$\lambda_2(\alpha) = (k_1 a_2 - k_2 a_1 \alpha)^2, \quad \lambda_5 = a_1^2 a_2^2.$$

The following expressions are used in the subsequent lemma:

$$\alpha_1 = \frac{k_1 a_2}{k_2 a_1^3} (a_1^2 + a_2^2 - a_1 a_2 + \eta), \quad (13)$$

$$\alpha_2 = \frac{k_1 a_2}{k_2 a_1^3} (a_1^2 + a_2^2 - a_1 a_2 - \eta), \quad (14)$$

$$\eta = \sqrt{(a_1 - a_2)^2 (a_1^2 + a_2^2)}, \quad (15)$$

$$\phi(\alpha) = \frac{\lambda_1^2}{\lambda_4 \lambda_1 - 2\lambda_2 + 2\sqrt{\lambda_5 \lambda_1^2 - \lambda_4 \lambda_1 \lambda_2 + \lambda_2^2}}, \quad (16)$$

$$\psi(\alpha) = \frac{\lambda_2(\alpha)}{\lambda_5}. \quad (17)$$

Lemma 1. For G_A in (6), $\|G_A\|_\infty^2$ is a piecewise continuous function with respect to α that maximizes $|G_A(j\omega, \alpha)|^2$ in (12) with respect to ω for all $a_1, a_2 > 0$, k_1 , and $k_2 \neq 0$.

With the expressions defined in (13)-(17), it is given by

$$\gamma_A^2(\alpha) := \|G_A\|_\infty^2 = \begin{cases} \phi(\alpha) & \text{if } \alpha_2 < \alpha < \alpha_1 \\ \psi(\alpha) & \text{otherwise} \end{cases}. \quad (18)$$

Proof. The proof is presented in [19]. \square

The function $\gamma_A^2(\alpha)$ thus provides the least upper bound on $\|e_A\|_2^2$, subject to $d(t) \in \mathcal{D}$, as a function of the user-defined parameter α . This function can be minimized with respect to α . The result is given in Theorem 1.

Theorem 1. For G_A in (6), the parameter α that minimizes $\|G_A\|_\infty^2$ for all $a_1, a_2 > 0$, k_1 , and $k_2 \neq 0$ is given by

$$\alpha^* = \frac{k_1}{k_2} \frac{2a_2}{\left(4a_1 + a_2 - \sqrt{8a_1^2 + a_2^2}\right)}. \quad (19)$$

Proof. The proof is presented in [19]. \square

Evaluating $\|G_A\|_\infty^2$ at α^* yields

$$\begin{aligned} \gamma_A^{*2} &:= \min_\alpha \gamma_A^2(\alpha) \\ &= \frac{k_1^2 \left(4a_1^2 - a_2 \left(a_2 + \sqrt{8a_1^2 + a_2^2}\right)\right)^4}{32a_1^2(a_1 + a_2)^3 \left(4a_1a_2\sqrt{\xi_1} + (a_2 - a_1)\xi_2\right)}, \end{aligned} \quad (20)$$

where

$$\xi_1 = 2a_1^2(a_1 - a_2)^2 \left(4a_1^2 + 5a_2^2 - 3a_2\sqrt{8a_1^2 + a_2^2}\right), \quad (22)$$

$$\xi_2 = a_2^4 - 8a_1^2(a_1^2 + a_2^2) + a_2(4a_1^2 + a_2^2)\sqrt{8a_1^2 + a_2^2}. \quad (23)$$

Remark 1. If $a_1 = a_2 = a$, the optimization problem in (10) is trivial; $\alpha^* = k_1k_2^{-1}$ and $\gamma_A^* = 0$. That is, if S_1 and S_2 have identical poles (but potentially different gains), perfect transfer can be achieved with zero transfer error. In this case, α^* only needs to compensate for the difference in the gains of the two systems.

Remark 2. The minimized H_∞ -norm, γ_A^* , is independent of k_2 . This can be explained by observing that in Fig. 2a, insufficient amplification of the output $x_2(t)$, by the gain k_2 , can be compensated by the multiplier α .

Remark 3. If the poles $-a_1$ and $-a_2$ are held constant, γ_A^{*2} is proportional to k_1^2 . As a result, the transfer error decreases if $|k_1|$ is decreased with the degenerate limit case being $k_1 = 0$, $\alpha^* = 0$, $\gamma_A^* = 0$, which achieves perfect matching because the target system outputs zero.

Additionally, the result of Theorem 1 can be analyzed for varying pole combinations $(-a_1, -a_2)$. There are two different ways to interpret the result: (i) by keeping the target system gain k_1 constant, or (ii) by re-parametrizing the systems using their DC gains $g_1 = k_1a_1^{-1}$ and $g_2 = k_2a_2^{-1}$, $G_1 = g_1a_1(s + a_1)^{-1}$ and likewise for G_2 , and keeping the DC gains constant.

Corollary 1. If the target system gain k_1 is held constant, γ_A^* , defined in (20), approaches infinity as the pole of the target system S_1 approaches the imaginary axis.

Therefore, TL may not be beneficial if the target system has a slow response to reference inputs.

Proof. In (21), a_1^2 is factored out in the denominator. Therefore, $\lim_{a_1 \rightarrow 0^+} \gamma_A^* \rightarrow \infty$. \square

Corollary 1 shows that TL is problematic if the target system has a pole that is close to the imaginary axis, unless $a_2 \approx a_1$. The interpretation of this is that if $a_1 \ll a_2$, the target system response is slow and does not follow the reference signal well, while the source system does. However, keeping k_1 constant and decreasing a_1 also means that the DC gain g_1 increases, which contributes to the bad transfer quality.

Therefore, it is valuable to study the results under the assumption of constant DC gains, which is more realistic in practice. The minimized bound can then be written as

$$\gamma_A^{*2} = g_1^2 \frac{\left(4a_1^2 - a_2 \left(a_2 + \sqrt{8a_1^2 + a_2^2}\right)\right)^4}{32(a_1 + a_2)^3 \left(4a_1a_2\sqrt{\xi_1} + (a_2 - a_1)\xi_2\right)}, \quad (24)$$

where ξ_1 and ξ_2 are from (22) and (23).

Corollary 2. If g_1 is held constant, it can be shown that $\gamma_A^* \rightarrow |g_1|/4$ as $a_1 \rightarrow 0$ and $\gamma_A^* \rightarrow |g_1|$ as $a_2 \rightarrow 0$.

Proof. Corollary 2 follows directly from (22):(24) by taking the limit $a_1 \rightarrow 0$ and $a_2 \rightarrow 0$. \square

Finding equivalent analytic results for the IOT case proved to be more difficult. Results are obtained from numerically finding $\|G_B\|_\infty$ as a function $\beta = [\beta_1, \beta_2]$. Numerical methods can then be used to find an estimate for β^* , along with $\gamma_B^* := \gamma_B(\beta^*)$, where we use a similar notation as in (18): $\gamma_B(\beta) := \|G_B\|_\infty$.

V. NUMERICAL EXAMPLES

A. Influence of System Parameters on Transfer Error

In this section, contour plots illustrate how the minimized H_∞ -norms of G_A and of G_B vary as a function of the pole magnitudes a_1 and a_2 .

In Fig. 3a, we keep k_1 constant (according to Corollary 1). The minimized bound γ_A^* is normalized by $|k_1|$. To compare the minimized H_∞ -norms of the transfer functions G_A and G_B , a contour plot of the ratio γ_B^*/γ_A^* is shown in Fig. 3b. This ratio is undefined when $a_1 = a_2$ because $\gamma_A^* = \gamma_B^* = 0$. For γ_B^* , no closed-form solutions are available to suggest how γ_B^* may be normalized, if at all. Therefore, k_1 and k_2 are both set to one in Fig. 3b. To obtain β^* and γ_B^* , a basic grid-search method was employed.

In Fig. 3c, a contour plot of γ_A^* is shown, where now the DC gain g_1 is held constant (according to Corollary 2). We chose $g_1 = 1$. This is in contrast to Fig. 3a where k_1 is held constant. A figure similar to Fig. 3b can be generated for $g_1 = g_2 = 1$. Qualitatively, the plot looks very similar to Fig. 3b.

In Figs. 3a and 3c, the base-10 logarithm of the data is shown to illustrate the variation more clearly. In all the

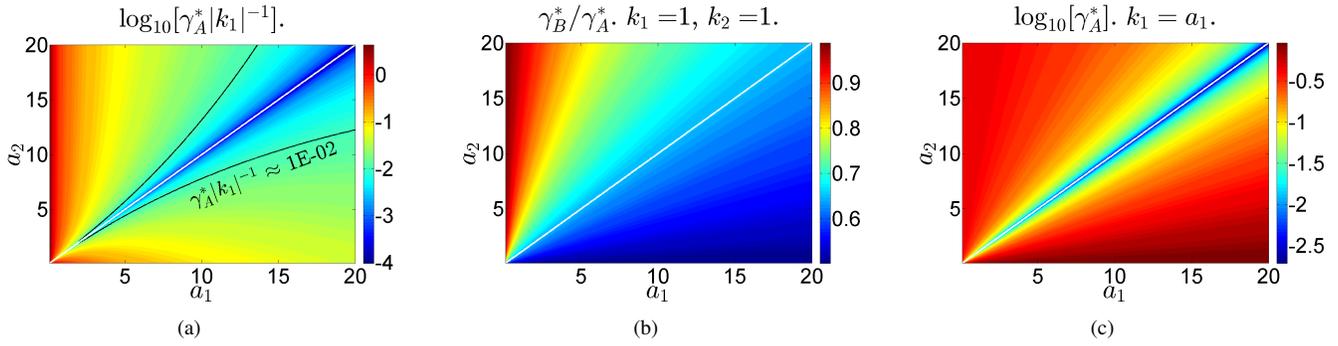


Fig. 3. On the left is a contour plot of $\log_{10}[\gamma_A^*|k_1^{-1}|]$ vs. a_1 and a_2 . When $a_1 = a_2$, $\gamma_A^*|k_1^{-1}|$ is zero. As the pole a_1 approaches zero, $\gamma_A^*|k_1^{-1}|$ grows exponentially. The black line illustrates a constant contour line where $\gamma_A^*|k_1^{-1}| = 1\text{E}-2$. In the middle, the ratio γ_B^*/γ_A^* vs. a_1 and a_2 is shown for $k_1=1$ and $k_2=1$. It can be seen that using the reference signal in the transformation always reduces the minimized H_∞ -norm of the system, and the asymmetry reveals that the benefit of doing so is most evident when $a_2 < a_1$. On the right, a contour plot of $\log_{10}[\gamma_A^*]$ vs. a_1 and a_2 is shown for $k_1 = a_1$, that is, $g_1 = 1$.

contour plots, a white line is drawn through the diagonal as $\gamma_A^* = \gamma_B^* = 0$ when $a_1 = a_2$.

Key Observation 1: In Fig. 3a, the minimized H_∞ -norm, γ_A^* , increases exponentially with diminishing a_1 , a result expected from Corollary 1 for constant k_1 . This phenomenon holds true for γ_B^* as well if k_1 and k_2 are held constant. This indicates that S_1 must not have a pole close to the imaginary axis if we want to achieve an accurate data transfer from S_2 to S_1 , unless k_1 is small as well.

Key Observation 2: For both Figs. 3a and 3c with constant k_1 and constant g_1 , respectively, the transfer quality is best if the poles of S_1 and S_2 lie close together.

Key Observation 3: For both Figs. 3a and 3c, if the poles a_1 and a_2 are of greater magnitude, they can be further apart for the minimized error bound to stay on the same contour line. For example, consider the two black contour lines for $\gamma_A^*|k_1^{-1}| = 1\text{E}-2$ in Fig. 3a. If $a_1 = 10$, then to not do worse than an error of around $1\text{E}-2$, a_2 must be approximately between 7.8 and 13.1. However, if $a_1 = 12$, then the range of allowable a_2 increases by around 47% to be approximately between 8.9 and 16.7. The same observation holds for the IOT case and for the OT analysis with constant DC gain.

Key Observation 4: According to Fig. 3b, γ_A^* is always greater than γ_B^* . This shows that using the reference signal in the transformation reduces the least upper bound on the 2-norm of the transformation error; that is, γ_B^*/γ_A^* is smaller than one. The asymmetry in Fig. 3b shows that it is especially useful to do this when $a_2 < a_1$; that is, when the target system responds faster than the source system. The same observation holds for the OT analysis with constant DC gain.

Key Observation 5: According to Fig. 3c, if the DC gain g_1 is held constant, then it is slightly more preferable to have $a_2 > a_1$, in contrast to if the gain itself, k_1 , is held constant.

B. Demonstration of Alignment-Based Transfer Learning for Two Sample Systems

In this section, we consider specific source and target systems and demonstrate TL for datasets from a specific reference signal. We consider two first-order, LTI, SISO systems with open-loop poles $-m_1$ and $-m_2$. They each have a proportional feedback controller with gains k_1 and k_2 . For S_1 , the state and input equations are

$$\dot{x}_1(t) = -m_1 x_1(t) + u_1(t), \quad (25)$$

$$u_1(t) = k_1(d(t) - x_1(t)), \quad (26)$$

and likewise for S_2 . Therefore, $a_1 = m_1 + k_1$ and likewise for a_2 . Each system is tasked to follow a sinusoidal reference signal $d(t) = 0.5 \sin(t)U(2\pi - t)$, where $U(\cdot)$ is the Heaviside step function. It can be shown that $\|d\|_2 < 1$. Each of the two systems are run in simulation for eight seconds and data is collected for $x_1(t)$, $x_2(t)$, and $d(t)$ at a sampling frequency of 100 Hz. Let \mathbf{X}_1 denote a $(1 \times N)$ matrix of samples of $x_1(t)$,

$$\mathbf{X}_1 = [x_1(0) \quad \cdots \quad x_1(k\delta) \quad \cdots \quad x_1(N\delta)]^T, \quad (27)$$

where $\delta = 0.01$ and $k \in \{0, 1, 2, \dots, N\}$. Similarly, \mathbf{X}_2 and \mathbf{D} are matrices constructed from samples of $x_2(t)$ and $d(t)$, respectively. Then, by least-squares regression, an optimal $\bar{\alpha}$ is found such that

$$\bar{\alpha} = \arg \min_{\alpha} \mathbf{E}_A^T \mathbf{E}_A, \quad (28)$$

where $\mathbf{E}_A := \mathbf{X}_1 - \alpha \mathbf{X}_2$. Similarly,

$$\{\bar{\beta}_1, \bar{\beta}_2\} = \arg \min_{\{\beta_1, \beta_2\}} \mathbf{E}_B^T \mathbf{E}_B, \quad (29)$$

where $\mathbf{E}_B := \mathbf{X}_1 - (\beta_1 \mathbf{X}_2 + \beta_2 \mathbf{D})$. Fig. 4a shows the trajectories of $x_1(t)$, $x_2(t)$ and $\alpha x_2(t)$ for $a_1 = 8$ and $a_2 = 4$, with gains $k_1 = k_2 = 5$. As expected, Fig. 4b shows that the transformation error is reduced when the reference signal is used.

To show the asymmetry of Figs. 3a and 3c, the simulation can be re-run for the case where now $a_1 = 4$ and $a_2 = 8$. By switching the poles to $a_2 > a_1$ and leaving the gains

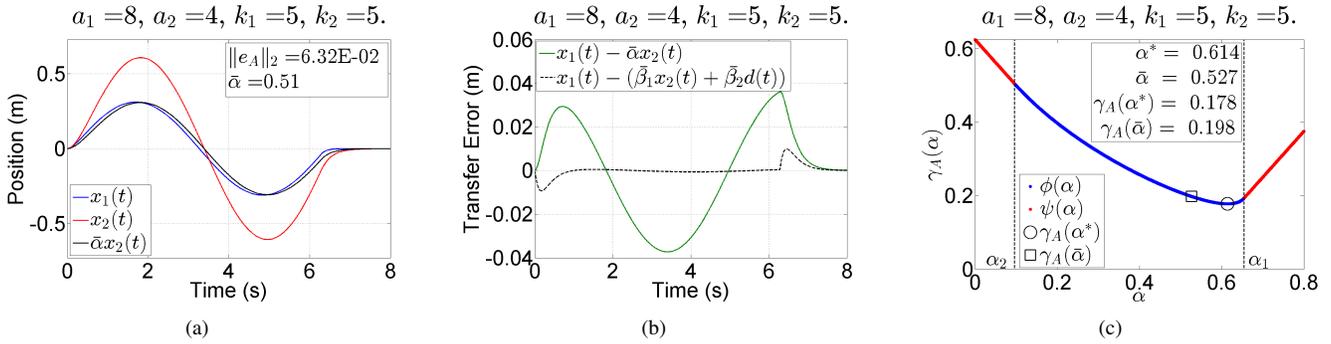


Fig. 4. On the left is an example of TL for the OT case. The objective is to align the red trajectory with the blue through $\bar{\alpha}$, see the black line. In the middle is a comparison of the transformation error of the OT case and the IOT case; the latter shows better performance. On the right, the squared H_∞ -norm of G_A from (18), $\gamma_A(\alpha)$, is illustrated as a function of α . The two dashed lines denote where the piecewise function changes from $\phi(\alpha)$ to $\psi(\alpha)$. It can be seen here that α^* , denoted by the circle, minimizes this function. The square marker denotes $\bar{\alpha}$, optimized for a specific reference signal.

constant, $\|e_A\|_2$ is increased by an order of magnitude from $6.32\text{E}-2$ to $1.24\text{E}-1$. However, if the DC gains are held constant, that is, $g_1 = .625$ and $g_2 = 1.25$, then it is slightly reduced to $6.19\text{E}-2$.

In Fig. 4c, $\bar{\alpha}$ and α^* are compared for the two sample systems, along with the values of γ_A evaluated at those optimized α values. Note that $\gamma_A(\bar{\alpha}) > \gamma_A(\alpha^*)$, indicating that although $\bar{\alpha}$ was optimally found for data generated from a specific reference signal, α^* provides a lower upper bound on $\|e_A\|_2$ for all $d(t) \in \mathcal{D}$. A similar result can be shown for the IOT case; for the same two sample systems, $\gamma_B(\beta^*) = 1.09\text{E}-2$, whereas $\gamma_B(\bar{\beta}) = 9.21\text{E}-2$.

Unlike the process to obtain solutions for α^* and γ_A^* , finding $\bar{\alpha}$ did not require any knowledge of the system models. When working with complex robotic systems, we will not be able to compute the H_∞ -norm, since TL is used in cases when no model of the systems is available in the first place. Nevertheless, our analysis provided insight and intuition on how system properties such as the pole location influences the quality of the transformation.

VI. CONCLUSIONS

This paper explored the feasibility of one robot using another robot's data for model learning by studying a simplified scenario. An LTI transformation is used to map the output from one control system to the output from another. By framing this as an H_∞ -norm minimization problem, an upper bound on the 2-norm of the transformation error is derived and minimized with respect to the transformation scalar, for a large set of reference inputs. This minimized error bound is a measure of the quality and benefit of TL, and answered the question as to when TL is beneficial. In particular, the transfer quality is good if the system poles lie close to each other, or if the system poles are further apart but both have large negative real parts. The quality is further improved if the reference signal is used in the transformation.

REFERENCES

- [1] D. Nguyen-Tuong and J. Peters, "Model learning for robot control: a survey," *Cognitive processing*, vol. 12, no. 4, pp. 319–340, 2011.
- [2] F. Berkenkamp and A. P. Schoellig, "Safe and robust learning control with Gaussian processes," in *Proc. of the European Control Conference (ECC)*, 2015, pp. 2501–2506.
- [3] C. J. Ostafew, A. P. Schoellig, and T. D. Barfoot, "Learning-based nonlinear model predictive control to improve vision-based mobile robot path-tracking in challenging outdoor environments," in *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 4029–4036.
- [4] A. P. Schoellig, F. L. Mueller, and R. D'Andrea, "Optimization-based iterative learning for precise quadcopter trajectory tracking," *Autonomous Robots*, vol. 33, no. 1–2, pp. 103–127, 2012.
- [5] M. Arif, T. Ishihara, and H. Inooka, "Incorporation of experience in iterative learning controllers using locally weighted learning," *Automatica*, vol. 37, no. 6, pp. 881–888, 2001.
- [6] P. Janssens, G. Pipeleers, and J. Swevers, "Initialization of ILC based on a previously learned trajectory," in *Proc. of the American Control Conference (ACC)*, 2012, pp. 610–614.
- [7] M. Hamer, M. Waibel, and R. D'Andrea, "Knowledge transfer for high-performance quadcopter maneuvers," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2013, pp. 1714–1719.
- [8] K. Tuyls and G. Weiss, "Multiagent learning: Basics, challenges, and prospects," *AI Magazine*, vol. 33, no. 3, p. 41, 2012.
- [9] A. P. Schoellig, J. Alonso-Mora, and R. D'Andrea, "Limited benefit of joint estimation in multi-agent iterative learning," *Asian Journal of Control*, vol. 14, no. 3, pp. 613–623, 2012.
- [10] M. E. Taylor and P. Stone, "Transfer learning for reinforcement learning domains: A survey," *Journal of Machine Learning Research*, vol. 10, pp. 1633–1685, 2009.
- [11] B. Lakshmanan and R. Balaraman, "Transfer learning across heterogeneous robots with action sequence mapping," in *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010, pp. 3251–3256.
- [12] G. Boutsioukis, I. Partalas, and I. Vlahavas, "Transfer learning in multi-agent reinforcement learning domains," in *Recent Advances in Reinforcement Learning*. Springer, 2012, pp. 249–260.
- [13] B. Bocsi, L. Csató, and J. Peters, "Alignment-based transfer learning for robot models," in *Proc. of the International Joint Conference on Neural Networks (IJCNN)*, 2013, pp. 1–7.
- [14] J. Ham, D. Lee, and L. Saul, "Semisupervised alignment of manifolds," in *Proc. of the Annual Conference on Uncertainty in Artificial Intelligence*, vol. 10, 2005, pp. 120–127.
- [15] C. Wang and S. Mahadevan, "A general framework for manifold alignment," in *AAAI Fall Symposium on Manifold Learning and Its Applications*, 2009, pp. 53–58.
- [16] C. Wang and S. Mahadevan, "Manifold alignment using procrustes analysis," in *Proc. of the International Conference on Machine Learning*, 2008, pp. 1120–1127.
- [17] D. Nguyen-Tuong, M. Seeger, and J. Peters, "Model learning with local Gaussian process regression," *Advanced Robotics*, vol. 23, no. 15, pp. 2015–2034, 2009.
- [18] J. C. Doyle, B. A. Francis, and A. Tannenbaum, *Feedback control theory*. Macmillan Publishing Company New York, 1992, vol. 1.
- [19] K. V. Raimalwala, B. A. Francis, and A. P. Schoellig, "Addendum to: An upper bound on the error of alignment-based transfer learning between two linear, time-invariant, scalar systems," University of Toronto – TSpace, Tech. Rep., 2015, available at <http://hdl.handle.net/1807/69365>.