Approximation by Simple Poles—Part II:
System Level Synthesis Beyond Finite Impulse Response

Michael W. Fisher, Gabriela Hug, and Florian Dörfler

Abstract—In Part I, a novel Galerkin-type method for finite dimensional approximations of transfer functions in Hardy space was developed based on approximation by simple poles. In Part II, this approximation is applied to system level synthesis, a recent approach based on a clever reparameterization, to develop a new technique for optimal control design. To solve system level synthesis problems, prior work relies on finite impulse response approximations that lead to deadbeat control, and that can experience infeasibility and increased suboptimality, especially in systems with large separation of time scales. The new design method does not result in deadbeat control, is convex and tractable, always feasible, can incorporate prior knowledge, and works well for systems with large separation of time scales. Suboptimality bounds with convergence rate depending on the geometry of the pole selection are provided. An example demonstrates superior performance of the method.

Index Terms—$H_{\infty}$ control, optimal control, optimization, system level synthesis (SLS).

I. INTRODUCTION

In Part I [1], the approximation by a finite collection of transfer functions with simple poles was studied as a Galerkin-type method for approximating transfer functions in Hardy space. This article applies this simple pole approximation (SPA) to optimal design of linear feedback controllers. A powerful approach for solving optimal control problems involves not directly optimizing over the controller, but rather over an affine function of a closed-loop transfer function (which depends implicitly on the controller), and then recovering the optimal controller that realizes this closed-loop behavior afterward. Examples include sensitivity minimization [2], the Youla parameterization [3], Q-parameterization [4], input–output parameterization (IOP) [5], and system level synthesis (SLS) [6], [7]. For this article, we focus on the closed-loop system responses for state feedback controllers, and so restrict our attention to SLS rather than sensitivity minimization (which minimizes a different closed-loop transfer function), Youla or Q-parameterization (which do not directly parameterize using the closed-loop transfer function), or IOP (which focuses on output feedback).

Mixed $H_2/H_\infty$ control synthesis is valuable for applications and has a long history (see, e.g., [8]). However, it remains challenging to solve efficiently as methods for $H_2$ and $H_\infty$ synthesis alone do not readily yield optimal solutions to mixed $H_2/H_\infty$ synthesis. The SLS reparameterization for mixed $H_2/H_\infty$ synthesis results in a convex but infinite dimensional optimization problem. In order to solve it, prior work [9] has approximated that the closed-loop responses are finite impulse responses (FIR) in order to arrive at a tractable finite dimensional optimization problem. However, this results in deadlock control (DBC), which often experiences poorly damped oscillations between discrete sampling times that can even persist in steady state, as well as lack of robustness to model uncertainty and parameter variations because of the high control gains required to reach the origin in finite time [10]. We denote SLS with the FIR approximation by DBC for the rest of this article.

With DBC, the number of poles in the closed-loop transfer functions is equal to the length of the FIR, potentially resulting in large numbers of poles that can lead to high computational complexity for the control design, lack of robustness in the resulting controller, and implementation challenges in practice [11, Ch. 19]. This is especially problematic when the optimal solution has a long settling time, such as in systems with large separation of time scales, where short sampling times are needed to capture the fast dynamics, which are also coupled with much slower dynamics. This leads to closed-loop impulse responses settling only after a large number of time steps. In addition, FIR closed-loop responses have all poles at the origin, which results in infeasibility in case of stable but uncontrollable poles in the plant. To resolve this, DBC introduces a slack variable enabling constraint violation, which leads to additional suboptimality [9]. Furthermore, in this case, DBC leads to a quasi-convex problem, requiring an iterative approach, such as golden section search, to solve rather than a single convex optimization [9]. This approach then requires inversion of a transfer function with an order equal to the length of the FIR, which is potentially large and may be numerically unstable, to recover the optimal closed-loop transfer functions and resulting controller [9].

This article combines SLS with simple pole approximation (SPA) [1] to develop a new control method, which addresses these limitations. This approach is not FIR, so it does not suffer from the drawbacks of DBC. Moreover, the number of poles is independent of the settling time of the optimal closed-loop

Received 1 October 2023; revised 22 May 2024; accepted 29 August 2024. Date of publication 17 September 2024; date of current version 28 February 2025. This work was supported in part by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research under Grant OSR-2019-COE-NEOM-4178.11 and in part by the European Union’s Horizon 2020 research and innovation program under Grant 883985. Recommended by Associate Editor Zhan Shu.

(Corresponding author: Michael W. Fisher.)

Michael W. Fisher is with the University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: mfisher@ethz.ch).
Gabriela Hug and Florian Dörfler are with the ETH Zürich, 8092 Zürich, Switzerland (e-mail: hug@eeh.ee.ethz.ch; dorfler@ethz.ch).

Digital Object Identifier 10.1109/TAC.2024.3462517

© 2024 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.
responses, and therefore SPA even works well for systems with large separation of timescales. It results in a convex and tractable optimization for the design, avoiding the need for iterative methods, does not need to invert transfer functions to recover the optimal closed-loop solution, requires only a small number of poles, guarantees feasibility for stabilizable systems without introducing slack variables, and additional suboptimality resulting from these can be avoided. Finally, if prior information is known about the optimal solution, such as the locations of some of the optimal poles (e.g., for model matching [12], model reference control [13], design based on the internal model principle [14], expensive control [15, Th. 3.12(b)], etc.), then these can be incorporated directly into the design for improved performance.

A suboptimality certificate is provided, which shows the convergence rate of SPA to the ground-truth optimal solution based on the geometry of the pole selection. Unlike a similar certificate for DBC, this does not require a long enough time horizon for the optimal impulse response to decay to be valid, and its convergence rate does not depend on this decay rate. This certificate is then specialized to a particular pole selection based on an Archimedes spiral as in [1, Th. 4]. An example shows superior performance of SPA over DBC, and is fully reproducible with all code publicly available [16].

The rest of this article is organized as follows. Section II provides preliminaries and problem setup, Section III provides the SPA method and suboptimality certificates, Section IV shows an illustrative example, and Section V gives the proofs. Finally, Section VI concludes this article.

II. Preliminaries

We use the same notation as in Part I [1], and refer the reader to the preliminaries and main results sections there for further details. Recall also Assumptions A1–A5 from Part I.

Consider the following LTI system in discrete time:

\[ x(k+1) = Ax(k) + Bu(k) + Bw(k) \]
\[ y(k) = Cx(k) + Du(k) \]

(1)

where \( x(k) \in \mathbb{R}^n \), \( u(k) \in \mathbb{R}^p \), \( w(k) \in \mathbb{R}^q \), and \( y(k) \in \mathbb{R}^m \) are the state, controller input, disturbance input, and performance output vectors at time step \( k \), respectively. Let \( \sigma \) be the stable plant poles (i.e., the stable eigenvalues of \( A \)). It will be useful to introduce the following related system:

\[ x(k+1) = Ax(k) + Bu(k) + v(k) \]
\[ y(k) = Cz(k) + Du(k) \]

where \( v(k) \in \mathbb{R}^n \) and the other signals are defined analogously to (1). Consider a linear (possibly dynamic) state feedback control law of the form \( u(z) = K(z)x(z) \), where \( K \in \mathbb{R}^{k \times n} \), and let \( T_{\text{desired}}(z) \) be some desired closed-loop transfer functions for model reference or model matching control (note that we can set \( T_{\text{desired}}(z) = 0 \), if desired). For any signals \( a(z) \) and \( b(z) \), let \( T_{a \rightarrow b}(z) \) denote the closed-loop transfer function from \( a(z) \) to \( b(z) \). For any transfer function \( F(z) \), let \( \Im(F(z)) \) and \( \Re(F(z)) \) denote its impulse response and convolution (i.e., causal Toeplitz) operators, respectively (see [1] for more details). The goal is to choose a controller \( K(z) \) that is a solution to the mixed \( H_2/H_{\infty} \) [11], [17] optimal control problem given by

\[
\min_{K(z)} ||T_{w \rightarrow y}(z) - T_{\text{desired}}(z)||_{H_2} + \lambda ||T_{w \rightarrow y}(z) - T_{\text{desired}}(z)||_{H_{\infty}} \\
\text{s.t. } T_{v \rightarrow z}(z), T_{v \rightarrow u}(z) \in \mathbb{R}^{k \times n}_{\infty} \tag{2}
\]

where \( \lambda \in [0, \infty) \) is constant. As \( T_{w \rightarrow y}(z) \) is nonconvex in \( K(z) \), (2) is known to be a challenging problem. We make the following feasibility assumption.

(A6): A solution to (2) exists, i.e., \((A,B)\) is stabilizable, and the optimal closed-loop transfer functions are rational (hence they have finitely many poles).

While one can construct pathological examples where this assumption does not hold (e.g., a controllable single input single output (SISO) system with \( y = x \) and \( T_{\text{desired}}(z) = z^{-1} \)), in the standard mixed \( H_2/H_{\infty} \) setting, Assumption A6 is satisfied automatically [18].

By Assumption A6, there exists an optimal solution \( T^*_{v \rightarrow z}(z) \) to (2). As \( T^*_{v \rightarrow x}, T^*_{v \rightarrow u} \in \mathbb{R}^{k \times n}_{\infty} \), we can write their partial fraction decompositions as

\[
T^*_{v \rightarrow z}(z) = \sum_{q=0}^{m_q} \sum_{j=1}^{\hat{m}_q} \frac{H^*_q(i,j)}{(z - q)} \\
T^*_{v \rightarrow x}(z) = \sum_{q=0}^{m_q} \sum_{j=1}^{\hat{m}_q} G^*_q(i,j) \frac{1}{(z - q)} \tag{3}
\]

where \( \Omega \) and \( \hat{\Omega} \) are finite sets of stable poles closed under complex conjugation, \( H^*_q(i,j) \) and \( G^*_q(i,j) \) are coefficient matrices, and \( m_q \) and \( \hat{m}_q \) are the multiplicities of the pole \( q \) in \( T^*_{v \rightarrow z} \) and \( T^*_{v \rightarrow x} \), respectively. It will be shown (in the proof of Lemma 2) that the following relationship between the poles \( \Omega, \hat{\Omega} \) holds: \( \hat{\Omega} \subset \Omega \cup \sigma \). Thus, each pole of \( T^*_{v \rightarrow z} \) must be a pole of at least one of \( T^*_{v \rightarrow u} \), and the plant.

A recent approach was proposed to solve problem (2) for the special case, where \( y = (Qx)^T \) \( (Ru)^T \) for constant matrices \( Q \) and \( R \), \( T_{\text{desired}}(z) = 0 \), and \( B = I \). This approach is known as SLS [9], and the key idea is to reparameterize the control design in terms of the closed-loop transfer functions \( \Phi_x(z) = T_{v \rightarrow z}(z) \) and \( \Phi_u(z) = T_{v \rightarrow u}(z) \). This transforms (2) into an infinite dimensional convex optimization problem at the price of the additional affine constraint

\[
(zI - A)\Phi_x - B\Phi_u = I. \tag{4}
\]

After solving (2) subject to (4), a controller that yields the optimal closed-loop responses can be recovered via \( K(z) = \Phi_u(z)\Phi_x^{-1}(z) \), and realizations of \( K(z) \) exist, which do not require transfer function inversion.

A. FIR Approximation

To obtain a tractable optimization problem, in [9], the FIR approximation is made for the closed-loop transfer functions \( \Phi_x \) and \( \Phi_u \), i.e., \( \Phi_x(z) = \sum_{i=1}^{T} G_i z^{-i} \) and \( \Phi_u(z) = \sum_{i=1}^{T} H_i z^{-i} \), where \( G_i \) and \( H_i \) are coefficient matrices and \( T \) is the length of the FIR, resulting in DBC. For an uncontrollable plant, it is not feasible to achieve FIR closed-loop transfer functions, so to maintain feasibility, DBC introduces a slack variable \( V \) that allows (4) to be violated. The objective then becomes nonconvex, so DBC uses a quasi-convex upper bound of the objective [9]. The true (i.e., realized) closed-loop responses are then given by \( T_{v \rightarrow z}(k) = \Phi_u(z)(I + \frac{V}{\sigma^T})^{-1} \) and \( T_{v \rightarrow u}(k) = \Phi_u(z)(I + \frac{V}{\sigma^T})^{-1} \). Let \( J^* \) be the optimal cost of (2), \( J(T) \) the optimal cost of DBC with an FIR of length \( T \), and \( C_\sigma, \rho > 0 \) such that \( ||J(T_{v \rightarrow z}(k))||_2 \leq C_\sigma p_k^T \) for all \( k \geq 0 \). Then, for \( T \) sufficiently large such that \( C_\sigma p_k^T < 1 \), DBC is feasible and satisfies the following suboptimality bound [9, Th. 4.7] for some
c > 0:
\[
\frac{J(T) - J^*}{J^*} \leq \frac{C_c \rho^T}{1 - C_c \rho^T} \left( 1 + \frac{\lambda c}{1 - \rho^T} \right).
\]  

When \( \rho \) is small (i.e., the optimal closed-loop response is slow), \( C_c \rho^T < 1 \) may require large \( T \), the convergence rate of \( C_c \rho^T \) in (5) is slow, and the term \( \frac{1}{1 - \rho^T} \) (which arises from the slack variable \( V \)) will further slow convergence.

### III. MAIN RESULTS

#### A. SPA Control Design

To introduce our new method, we begin by reformulating (2) using the SLS reparameterization, which results in the following convex but infinite dimensional optimization problem, which is a strict generalization of the formulation in [9].

Recall that \( \sigma \) are the stable poles of the plant, where each \( q \in \sigma \) has multiplicity \( m_q \), and \( \mathcal{P} \) represents a selection of poles within the unit disk [1]. To obtain a tractable optimization problem, we approximate \( \Phi_\sigma \) and \( \Phi_u \) using \( \mathcal{P} \) and \( \sigma \) by

\[
\Phi_\sigma(z) = \sum_{p \in \mathcal{P}} H_p \frac{1}{z - p} \quad \text{and} \quad \Phi_u(z) = \sum_{p \in \mathcal{P}} G_p \frac{1}{z - p} + \sum_{q \in \sigma} \sum_{i = 1}^{m_q + 1} G_{q,i} \frac{1}{z - q},
\]

where \( H_p, G_p \), and \( G_{q,i} \) are coefficient matrices. We refer to this as the SPA since all of the poles of \( \Phi_\sigma \) are simple. Lemma 1 shows that SPA is always feasible for \((A, B)\) stabilizable, and its proof explains the asymmetry in the poles of \( \Phi_x \) and \( \Phi_u \) due to the plant poles.

**Lemma 1:** Under Assumption A6, the SPA of (6) yields a feasible solution for (2) with the SLS constraint (4).

Although it is possible to select any poles \( \mathcal{P} \subset \mathbb{D} \) for the SPA method, we provide several recommendations that often lead to improved performance. First, we suggest to include the stable poles of the plant \( \sigma \) in \( \mathcal{P} \) to allow the design to cancel out any controllable modes of the plant for which it is advantageous to do so. In addition, for any poles of the optimal solution that are known a priori (see Section I), including these in \( \mathcal{P} \) can lead to a dramatic improvement in performance. For the remaining poles, the Archimedean spiral is a natural choice as it provides an approximately even pole selection over \( \mathbb{D} \) and converges at the rate \( (|\mathcal{P}| + 2)^{-1/2} \) [1].

For any \( q \in \sigma \), let \( \tilde{m}_q = 1 \) if \( q \in \mathcal{P} \) and \( \tilde{m}_q = 0 \), otherwise. Then, the SPA of (6) applied to (2) subject to the SLS constraint (4) results in the following optimal control design problem, consisting of the objective:

\[
\min_{H_p, G_p, G_{q,i}} \left\| J(C\Phi_x B) + J(D\Phi_u B) - J(T_{\text{desired}}) \right\|_F + \lambda \left\| C(C\Phi_x B) + C(D\Phi_u B) - C(T_{\text{desired}}) \right\|_2
\]

subject to the following SLS constraints (whose form given below is derived in the proof of Lemma 2):

\[
G_{q,2} + (qI - A)G_{q,1} - BH_q = 0 \quad \forall q \in \sigma \cap \mathcal{P}
\]

\[
(pI - A)G_p - BH_p = 0 \quad \forall p \in \mathcal{P} \cap \sigma
\]

\[
G_{q,i+1} + (qI - A)G_{q,i} = 0 \quad \forall q \in \sigma
\]

\[
i \in \{1, \tilde{m}_q, \ldots, m_q\}
\]

### B. Suboptimality Bounds

Recall that \( d(z, \mathcal{P}) \) is the distance from \( z \) to \( \mathcal{P} \), where \( \max_{z \in \mathbb{D}} d(z, \mathcal{P}) \) measures the geometric approximation error between approximating poles \( \mathcal{P} \) and optimal poles \( \Omega \), and \( D(\mathcal{P}) = \max_{z \in \mathbb{D}} d(z, \mathcal{P}) \) measures the worst-case geometric approximation error (for unknown \( \Omega \)). In addition, \( r \in (0, 1) \) is such that \( \mathcal{P} \subset B_r \), and \( \delta \) is a measure of the minimum distance between each approximating pole in \( \mathcal{P} \) and \( \sigma \) (see [1] for further details). Also, recall Assumptions A1–A5 from Part I [1]. Our main theoretical result shows that the relative error of the SPA method decays at least linearly with \( D(\mathcal{P}) \).

**Theorem 1 (General Suboptimality Bound):** Let \( J^* \) denote the optimal cost of (2), and let \( J(\mathcal{P}) \) denote the optimal cost of (7)-(9) for any choice of \( \mathcal{P} \). Suppose that Assumption A6 is
met, and $\mathcal{P}$ satisfies Assumptions A1–A5. Then, there exists a constant $\hat{K} = \hat{K}(\Omega, G^*_p(q,j), H^*_p(q,j), r, \delta) > 0$ such that

$$\frac{J(\mathcal{P}) - J^*}{J^*} \leq \hat{K} \mathcal{D}(\mathcal{P}).$$ \hspace{1cm} (10)

While the DBC suboptimality bound in (5) only holds for $T$ sufficiently large such that $||T_{\tau_{min}}(T)||_2 \leq C_\rho T < 1$, the SPA bound in (10) does not have this requirement. The constant term in the DBC bound is expressed in terms of the $\mathcal{H}_\infty$ norm of the optimal controller, whereas $\hat{K}$ from Theorem 1 depends on the partial fraction decomposition of the optimal closed-loop responses. Furthermore, the DBC includes a multiplicative term $\frac{1}{\sqrt{C_\rho T}}$ resulting from the slack variable, whereas the SPA bound has no such term because it does not need a slack variable. Finally, the convergence for the DBC bound depends on the rate of decay of the optimal closed-loop impulse response, whereas the SPA bound depends on the distance between $\mathcal{P}$ and the optimal closed-loop poles. Therefore, SPA is preferable when the optimal impulse response takes long to decay, such as in stabilizable systems with large separation of time scales. However, DBC may be preferable when the optimal responses decay fast, or when many poles are desired, since its convergence rate approaches exponential as the number of poles approaches infinity. In addition, if some optimal poles can be included in $\mathcal{P}$ due to prior knowledge (see Section III-A), this will typically have the effect of decreasing both $D(\mathcal{P})$ and $\hat{K}$ in (10), significantly reducing the relative error of SPA. In contrast, it is not clear how such prior knowledge could be included with DBC to reduce its relative error.

Corollary 1 shows that for the Archimedes spiral pole selection in [1], the relative error of SPA converges to zero at the rate $(|\mathcal{P}| + 2)^{-1/2}$ since $|\mathcal{P}| = 2n - 2$ for each $n > 0$.

**Corollary 1 (Spiral Suboptimality Bound):** For each even integer $n > 0$, let $\mathcal{P}_n$ denote the selection of $(2n - 2)$ poles given by $p_k$ for $k \in \{-n, ..., n - 1\}$, where

$$\theta_k = 2\sqrt{n}k, \quad r_k = \sqrt{n}, \quad p_k = (r_k, \theta_k), \quad p_{-k} = (r_k, -\theta_k).$$

Then, there exists a constant $\hat{K} = \hat{K}(\Omega, G^*_p(q,j), H^*_p(q,j)) > 0$ and $N > 0$ such that $n \geq N$ implies

$$\frac{J(\mathcal{P}_n) - J^*}{J^*} \leq \frac{\hat{K}}{\sqrt{n}}.$$ \hspace{1cm} (11)

Note that in Corollary 1 only needs to be chosen to ensure that $D(\mathcal{P}_N) < 1$, $|\mathcal{P}_N| \geq n_{\text{max}}$, and that $\delta > 0$ for $\mathcal{P}_N$, and so is typically satisfied in practice with small $N$ (see the remark following [1, Th. 3]).

**IV. NUMERICAL EXAMPLE**

To compare DBC and SPA, we consider the example of using a power converter to provide frequency and voltage control services to the power grid, which arises naturally as a result of interfacing renewable generation to the grid [19]. This example served as the motivation to develop the SPA method, because of the inadequate performance of DBC resulting from the large separation of time scales in power systems containing power converter interfaced devices [20]. Let $w$ represent the frequency and voltage magnitude at the connection point, and let $y$ represent the power output of the converter. Then, this can be formulated in the form of (2) with matrices given by

$$A = \begin{bmatrix} 0.088 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0.995 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.005 & 0 \\ 0 & 0.1 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 0 \ 0 & 0 \ 0.01 & 0 \ 0 & 0.01 \end{bmatrix}, \quad B = \begin{bmatrix} -0.0001 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0.829 & -0.428 & 1.02 & 0 \\ 0 & 0.428 & 0.829 & 0 & -1.02 \end{bmatrix},$$

$$T_{\text{desired}}(z) = \begin{bmatrix} \frac{z-5.65-5}{z-0.995} & 0 \\ 0 & \frac{z-0.995}{z-0.1} \end{bmatrix}.$$
matching with the optimal transfer function than DBC, and with far fewer poles.

V. PROOFS

Proof of Lemma 1: Working in coordinates where \( A \) is in a Jordan normal form, we may have \( A = \text{diag}(A_u, A_c) \) and \( B = [B_u^T \ B_c^T]_1 \), where \( (A_c, B_c) \) is controllable and \( (A_u, B_u) \) is not. As \( (A, B) \) is stabilizable, there exists \( K_c \) such that the eigenvalues of \( A_c + B_cK_c \) lie in \( \mathcal{P} \).

Let \( K = [0 \ K_c] \). Then, \( A + BK = [A_u \ B_uK_c] \) so \( \Phi_u(z) = [(zI - A_u)^{-1}] \) and \( \Phi_u(z) = K\Phi(z) = [0 \ K_c(zI - (A_c + B_cK_c)^{-1})] \). Thus, since \((A, B)\) is stabilizable, the poles of \( \Phi_u \) lie in \( \mathcal{P} \cup \sigma \) and contain the uncontrollable plant poles, while the poles of \( \Phi_u \) lie in \( \mathcal{P} \). So, \( \Phi_u, \Phi_u \in \mathcal{Z} \mathcal{K}_\infty \) and satisfy (4).

The key technical result required to prove Theorem 1 is Lemma 2, which extends the approximation error bounds of [1, Th. 1] to bound the error between a feasible solution \( (\Phi_u, \Phi_x) \) of (7)-(9) and the optimal solution \( (\Phi_u^*, \Phi_x^*) \) of (2).

Lemma 2: Under the conditions of Theorem 1, let \( (\Phi_u^*, \Phi_x^*) \) denote the optimal solution to (2). Then, there exist \( \Phi_x, \Phi_u \in \mathcal{Z} \mathcal{K}_\infty \), which are a feasible solution to (7)-(9), and constants \( K_u^1, K_u^2, K_x^1, K_x^2 > 0 \), such that

\[
||\Phi_u - \Phi_u^*||_{\mathcal{K}_\infty} \leq K_u^1 D(\mathcal{P}), \quad ||\Phi_u - \Phi_u^*||_{\mathcal{K}_\infty} \leq K_u^2 D(\mathcal{P})
\]

\[
||\Phi_x - \Phi_x^*||_{\mathcal{K}_\infty} \leq K_x^1 D(\mathcal{P}), \quad ||\Phi_x - \Phi_x^*||_{\mathcal{K}_\infty} \leq K_x^2 D(\mathcal{P}).
\]

Before proving Lemma 2, we will require several technical results as given in the next few lemmas and corollaries. Lemma 3 bounds the distance between two simple transfer functions in terms of the distance between their poles.

Lemma 3: Let \( k \) be any integer, \( m \) a positive integer, and \( z \in \mathbb{D} \). Let \( q, p_1, \ldots, p_m \in \mathbb{D} \), and let \( \delta(q) = \max_i |p_i - q| \). If \( z \in \mathbb{D} \), let \( \delta = \delta(\mathbb{D}), \{p_i\}_{i=1}^m > 0 \) and \( \eta = d(q, \partial \mathbb{D}) > 0 \); if not, suppose that \( d(z, p_1), d(z, p_m) \geq \delta > 0 \) and \( d(z, q) \geq \eta > 0 \). Then, there exists \( K > 0 \) such that

\[
\left| \sum_{i=1}^m \frac{(z - q)^k}{(z - p_i)^k} - (z - q)^{k - m} \right| 
\leq K \delta(q).
\]

Proof of Lemma 3: Let \( \mathcal{P} = \{p_1, \ldots, p_m\} \). We compute

\[
\left| \sum_{i=1}^m \frac{(z - q)^k}{(z - p_i)^k} - (z - q)^{k - m} \right| = \left| \sum_{i=1}^m \frac{(z - q)^m - \sum_{i=1}^m (z - p_i)}{(z - q)^m - \sum_{i=1}^m (z - p_i)} \right|.
\]

Noting that the proofs of [1, Eqs. (10) and (12)] are still valid for \( z \in \mathbb{D} \) (i.e., \( |z| \leq 1 \)), applying them here we have that

\[
|z - q)^m - \sum_{i=1}^m (z - p_i) | \leq ((|q| + 2)^m - (|q| + 1)^m) \delta(q).
\]

Furthermore

\[
|z - q)^m - \sum_{i=1}^m (z - p_i) | \geq d(z, q)^{m-k} d(z, \mathcal{P})^m \geq \eta^{m-k} \delta^m.
\]

Combining these two inequalities implies that

\[
\left| \sum_{i=1}^m \frac{(z - q)^k}{(z - p_i)^k} - (z - q)^{k - m} \right| \leq K \delta(q).
\]
\[ K = \left( (|q| + 2)^m - (|q| + 1)^m \right) \eta^{m-k} g^m. \]

Lemmas 4 and 5 prove useful identities related to partial fraction decompositions of approximating transfer functions.

**Lemma 4:** Let \( k \) be a nonnegative integer, \( m \) a positive integer, and \( q, p_1, \ldots, p_m \in \mathbb{D} \). Let \( c_{p_i} = \left( \prod_{j \neq i} (p_i - p_j) \right)^{-1} \). Then, the following holds.

a) For \( k < m \)
\[
\sum_{i=1}^{m} (p_i - q)^k c_{p_i} \frac{1}{z - p_i} = \frac{(z-q)^k}{\prod_{i=1}^{m} (z - p_i)}. 
\]

b) For \( k \geq m \)
\[
\sum_{i=1}^{m} (p_i - q)^k c_{p_i} \frac{1}{z - p_i} = \frac{(z-q)^k}{\prod_{i=1}^{m} (z - p_i)} - \sum_{i=0}^{k-m} b_i (z-q)^i.
\]

**Proof of Lemma 4:** First, consider Case (a). Write the partial fraction decomposition
\[
\frac{(z-q)^k}{\prod_{i=1}^{m} (z - p_i)} = \sum_{i=1}^{m} \frac{\kappa_i}{z - p_i}.
\]
Multiplying both sides by \( \prod_{i=1}^{m} (z - p_i) \) and evaluating at \( z = p_i \) imply that
\[
\kappa_i = \frac{(p_i - q)^k}{\prod_{j \neq i} (p_i - p_j)} = (p_i - q)^k c_{p_i},
\]
which completes the proof for Case (a).

Next consider Case (b). Write the partial fraction decomposition
\[
\frac{(z-q)^k}{\prod_{i=1}^{m} (z - p_i)} = \sum_{i=1}^{m} \frac{\kappa_i}{z - p_i} + \sum_{i=0}^{k-m} b_i (z-q)^i. \tag{14}
\]
Multiplying by \( \prod_{i=1}^{m} (z - p_i) \) and evaluating at \( z = p_i \) imply
\[
\kappa_i = \frac{(p_i - q)^k}{\prod_{j \neq i} (p_i - p_j)} = (p_i - q)^k c_{p_i}.
\]
Differentiating (14) \( i \) times with respect to \( z \) and evaluating at \( z = q \) imply that
\[
0 = -i! \sum_{j=1}^{m} (p_j - q)^{k-1-i} c_{p_j} + i! b_i
\]
for \( i \in \{0, \ldots, k-m\} \), so
\[
b_i = \sum_{j=1}^{m} (p_j - q)^{k-1-i} c_{p_j}.
\]

For the rest of this section, \( \lambda \in \sigma \) will represent a stable eigenvalue of the \( A \) matrix of the planet. Let \( m \) and \( n \) be integers, and define the rising factorial \( m^{(n)} = \prod_{k=0}^{n-1} (m + k) \) and the falling factorial \( m^\underline{n} = \prod_{k=0}^{n-1} (m - k) \). For \( m \) and \( n \) nonnegatives, letting \( m! \) denote the standard factorial, we have
\[
m^{(n)} = \frac{(m+n-1)!}{(m-1)!} \text{ and } m^\underline{n} = \frac{m!}{(m-n)!}.
\]
Note, the following holds.

**Fact 3:** \( (-1)^n m^{(n)} = (-m)^n \).

**Fact 4:** \( \sum_{j=0}^{n} \binom{n}{j} m_j (m')_{n-j} = (m + m')_n \).

**Lemma 5:** Let \( k \) and \( m \) be positive integers, \( z \in \partial \mathbb{D} \), and \( q, \lambda, p_1, \ldots, p_m \in \mathbb{D} \) with \( d(\lambda, \{ p_i \}_{i=1}^{m}) \geq \delta > 0 \) and \( d(\lambda, q) \geq \eta > 0 \). Choose constants \( c_{p_i} \), as in Lemma 4. Then, the following holds.

a) There exists \( K > 0 \) such that
\[
\sum_{i=1}^{m} c_{p_i} \frac{(\lambda - p_i)^{-k}}{z - p_i} = \frac{(\lambda - z)^{-k}}{\prod_{i=1}^{m} (z - p_i)} - \frac{r(z)}{(\lambda - z)^k}
\]
and
\[
\lim_{z \to \lambda} \frac{d}{dz} \left( r(z) \prod_{i=1}^{m} (z - p_i) \right) = \begin{cases} 
1, & l = 0 \\
0, & l \in \{1, \ldots, k-1\} \\
\left( \frac{(-1)^l + (m)_l}{(\lambda - z)^{-m-k} + \epsilon} \right) \prod_{i=1}^{m} (\lambda - p_i), & l = k
\end{cases}
\]
where \( |c| \leq K \hat{d}(q) \).

b) There exist \( K'_0, \ldots, K'_{k-1} > 0 \) such that
\[
\sum_{i=1}^{m} c_{p_i} \frac{(p_i - q)^m}{(\lambda - p_i)^k} \frac{1}{z - p_i} = \frac{(z-q)^m}{(\lambda-z)^k \prod_{i=1}^{m} (z - p_i)} - \frac{1}{(\lambda-z)^{-k-m}} \sum_{n=0}^{k-1} \frac{a_n (\lambda - z)^n}{(\lambda-z)^{-k-n}}
\]
\[
|a_0 - 1| \leq K'_0 \hat{d}(q), \quad |a_n| \leq K'_n \hat{d}(q), \quad n \in \{1, \ldots, k-1\}.
\]

**Proof of Lemma 5:** For \( l \in \{0, m\} \), write the partial fraction decomposition
\[
\frac{(z-q)^l}{(\lambda-z)^k \prod_{i=1}^{m} (z - p_i)} = \sum_{i=1}^{m} \frac{\kappa_i}{z - p_i} + \frac{r(z)}{(\lambda - z)^k}
\]
and
\[
r(z) = \sum_{n=0}^{k-1} a_n (\lambda - z)^n. \tag{15}
\]
Multiplying both sides by \( (\lambda - z)^k \prod_{i=1}^{m} (z - p_i) \) yields
\[
(z-q)^l = (\lambda - z)^k \sum_{i=1}^{m} \frac{\kappa_i}{z - p_i} + r(z) \prod_{i=1}^{m} (z - p_i).
\]
Evaluating (16) at \( z = p_i \) implies that
\[
\kappa_i = c_{p_i} (p_i - q)^l (\lambda - p_i)^{-k}.
\]
For any nonnegative integer \( n \), define
\[
b_n = \left( \frac{d}{dz^n} r(z) \right)(\lambda) \equiv (-1)^n n! a_n
\]
\[
d_n = \left( \frac{d}{dz^n} \prod_{i=1}^{m} (z - p_i) \right)(\lambda) = \sum_{\nu \in \mathbb{E}^n} \prod_{v_i \neq c} \sum_{v_j \in I_{\nu_i}} \prod_{k \in I_{\nu_j}} (\lambda - p_k)
\]
\[
e_n = \left( \frac{d}{dz^n} (z - q)^l \right)(\lambda) = l_n (\lambda - q)^{l-n}. \tag{18}
\]
Note that
\[
\lim_{z \to \lambda} \frac{d}{dz^n} \left( r(z) \prod_{i=1}^{m} (z - p_i) \right) = \sum_{j=0}^{n} \binom{n}{j} d_j b_{n-j} \quad (20)
\]
for any nonnegative integer \( n \). Differentiating (16) \( n \) times with respect to \( z \), and evaluating at \( z = \lambda \) imply that
\[
e_n = \lim_{z \to \lambda} \frac{d}{dz^n} \left( r(z) \prod_{i=1}^{m} (z - p_i) \right) = \sum_{j=0}^{n} \binom{n}{j} d_j b_{n-j} \quad (21)
\]
for \( n \in \{0, \ldots, k-1\} \). Dividing by \( d_0 \) and solving for \( b_n \) imply
\[
b_n = \frac{e_n}{d_0} - \sum_{j=1}^{n} \binom{n}{j} \frac{d_j}{d_0} b_{n-j}. \quad (22)
\]
Note that
\[
d_n = \frac{d_n}{d_0} = \prod_{k \in V} \frac{1}{\lambda - p_k}.
\]
Define
\[
e'_n = \frac{d_n}{d_0} - \frac{m_n}{(\lambda - q)^n} = \sum_{\omega \in \Omega_n} \left( \prod_{k \in \omega} \frac{1}{\lambda - p_k} - \frac{1}{(\lambda - q)^n} \right)
\]
since the number of terms in the sum is \( m_n \). Thus, by Lemma 3, there exists \( k'_n > 0 \) such that
\[
\frac{d_n}{d_0} = m_n \frac{1}{(\lambda - q)^n} + e'_n, \quad |e'_n| \leq k'_n \hat{d}(q) \quad (23)
\]
Consider first Case (a): \( l = 0 \). Then, \( e_0 = 1 \) and \( e_n = 0 \) for \( n \in \{1, \ldots, k-1\} \). By (21), this implies the desired result for \( n \in \{0, \ldots, k-1\} \), so it suffices to prove the desired result for \( n = k \). By (22), \( b_0 = \frac{1}{d_0} \). We claim that there exists \( k_n > 0 \), such that
\[
-\sum_{j=1}^{n} \binom{n}{j} \frac{d_j}{d_0} b_{n-j} = \frac{(-1)^n m(n)}{(\lambda - q)^{m+n}} + e_n, \quad |e_n| \leq k_n \hat{d}(q) \quad (24)
\]
for all \( n \in \{1, \ldots, k\} \). Note that by (22), this implies that
\[
b_n = \frac{(-1)^n m(n)}{(\lambda - q)^{m+n}} + e_n, \quad |e_n| \leq k_n \hat{d}(q) \quad (25)
\]
for \( n \in \{1, \ldots, k-1\} \). We prove (24) by strong induction. For the base case, first note that
\[
b_0 = \frac{1}{d_0} = \frac{1}{(\lambda - q)^m} + \frac{1}{\prod_{i=1}^{m} (\lambda - p_i) - (\lambda - q)^m}
\]
\[
= \frac{1}{(\lambda - q)^m} + e_0, \quad |e_0| \leq k_0 \hat{d}(q) \quad (26)
\]
where such \( k_0 \) exists by Lemma 3. Then, for \( n = 1 \), we have
\[
-\frac{d_1}{d_0} b_0 = -\left( \frac{m}{\lambda - q} + e'_1 \right) \frac{1}{(\lambda - q)^m} + e_0
\]
\[
= -\frac{m}{(\lambda - q)^m} - \frac{m}{\lambda - q} e_0 - \frac{1}{(\lambda - q)^m} e'_1 - e_0 e'_1
\]
\[
= -\frac{m}{(\lambda - q)^m} + e_1, \quad |e_1| \leq k_1 \hat{d}(q)
\]
\[
k_1 = \frac{m k_0}{\lambda - q} + \frac{K'_1}{(\lambda - q)^m} + k_0 k'_1,
\]
For the induction step, assume that (24) holds for all \( j \in \{1, \ldots, n-1\} \), which, together with (26), implies that (25) holds for all \( j \in \{0, \ldots, n-1\} \). By (23) and (25), we have
\[
-\sum_{j=1}^{n} \binom{n}{j} \frac{d_j}{d_0} b_{n-j} = -\sum_{j=1}^{n} \binom{n}{j} \left( m_j \frac{1}{(\lambda - q)^j} + e'_n \right)
\]
\[
\ast \left( (-1)^{n-j} m(n-j) \frac{1}{(\lambda - q)^{m+n-j}} + e_{n-j} \right)
\]
\[
= -\sum_{j=1}^{n} \binom{n}{j} m_j (-1)^{n-j} m(n-j) \frac{1}{(\lambda - q)^{m+n}} + e_n
\]
\[
e_n = -\sum_{j=1}^{n} \binom{n}{j} \left( m_j e_{n-j} + (-1)^{n-j} m(n-j) e'_n + e_n e_{n-j} \right)
\]
\[
|e_n| \leq k_n \hat{d}(q)
\]
\[
k_n = \sum_{j=1}^{n} \binom{n}{j} \left( m_j k_{n-j} + m(n-j) k'_n + k_{n-j} k'_n \right).
\]
So, the following holds:
\[
-\sum_{j=1}^{n} \binom{n}{j} \frac{d_j}{d_0} b_{n-j} \quad \text{above identity}
\]
\[
= -\frac{1}{(\lambda - q)^{m+n}} \sum_{j=1}^{n} \binom{n}{j} m_j (-1)^{n-j} m(n-j) + e_n
\]
\[
\text{Fact 3} \quad = -\frac{1}{(\lambda - q)^{m+n}} \sum_{j=1}^{n} \binom{n}{j} m_j (-m)_{n-j} + e_n
\]
\[
\text{add} \quad = \frac{1}{(\lambda - q)^{m+n}} \left( (-m)_n - \sum_{j=0}^{n} \binom{n}{j} m_j (-m)_{n-j} \right) + e_n
\]
\[
\text{Fact 4} \quad = \frac{1}{(\lambda - q)^{m+n}} \left( (-m)_n - (m - m)_{n} + e_n \right)
\]
\[
\text{add} \quad = \frac{1}{(\lambda - q)^{m+n}} (-m)_n + e_n
\]
\[
\text{Fact 3} \quad = (-1)^n m(n) \frac{1}{(\lambda - q)^{m+n}} + e_n, \quad |e_n| \leq k_n \hat{d}(q).
\]
Thus, (24) holds. Note that \( b_k = 0 \) since \( r(z) \) is a polynomial of the order \( k-1 \). Therefore, by (20) and (24), we have that
\[
\lim_{z \to \lambda} \frac{d}{dz^k} \left( r(z) \prod_{i=1}^{m} (z - p_i) \right) = \sum_{j=0}^{k} \binom{k}{j} d_j b_{k-j}
\]
\[
= d_0 b_k + \sum_{j=1}^{k} \binom{k}{j} d_j b_{k-j} = \sum_{j=1}^{k} \binom{k}{j} d_j b_{k-j}
\]
\[
= (-d_0) \left( -\sum_{j=1}^{k} \binom{k}{j} \frac{d_j}{d_0} b_{k-j} \right)
\]
\[
= (-d_0) \left( (-1)^k m(k) \frac{1}{(\lambda - q)^{m+k}} + e_k \right), \quad |e_k| \leq k_k \hat{d}(q).
\]
which yields the result for Case (a). Next, consider Case (b): \( l = m \). Then, by (17) and (22)

\[
a_0 = b_0 = \frac{e_0}{d_0} = \frac{(\lambda - q)^m}{\prod_{i=1}^{m}(\lambda - p_i)}
\]

so

\[
|a_0 - 1| = \left| \frac{(\lambda - q)^m}{\prod_{i=1}^{m}(\lambda - p_i)} - 1 \right| \leq k_0 \hat{d}(q)
\]

where such \( k_0 > 0 \) exists by Lemma 3. We claim that there exist \( k_n > 0 \) such that

\[
|b_n| \leq k_n \hat{d}(q) \quad (27)
\]

for \( n \in \{1, \ldots, k - 1\} \). We prove (27) by strong induction. For the base case, note that by (22) and (23)

\[
b_1 = \frac{e_1}{d_0} - \frac{d_1}{d_0} b_0 = \frac{m(\lambda - q)^{m-1}}{d_0} - \left( \frac{m}{\lambda - q} + \epsilon_1 \right) \frac{e_0}{d_0}
\]

\[
= \frac{m(\lambda - q)^{m-1} - m \frac{1}{\lambda - q} (\lambda - q)^m}{d_0} - \epsilon_1 \frac{(\lambda - q)^m}{d_0}
\]

\[
= - \epsilon_1 \frac{(\lambda - q)^m}{d_0}, \quad |b_1| \leq k_1 \hat{d}(q), \quad k_1 = k_0 \frac{|\lambda - q|^m}{\prod_{i=1}^{m}(\lambda - p_i)}
\]

For the induction step, assume that (27) holds for all \( j \in \{1, \ldots, n - 1\} \). By (19), (22), (23), and the induction hypothesis, we have

\[
b_n \overset{(22)}{=} \frac{e_n}{d_0} - \sum_{j=1}^{n-1} \binom{n}{j} \frac{d_j}{d_0} b_{n-j}
\]

regrouping terms

\[
= \frac{e_n}{d_0} - \sum_{j=1}^{n-1} \binom{n}{j} \frac{d_j}{d_0} \frac{b_{n-j}}{b_{n-j}}
\]

\[
\overset{(19)}{=} \frac{m_n(\lambda - q)^{m-n}}{d_0} - \left( \frac{m_n}{\lambda - q} + \epsilon_n \right) \frac{e_0}{d_0}
\]

\[
- \sum_{j=1}^{n-1} \binom{n}{j} \left( \frac{d_j}{d_0} \left( m_j(\lambda - q)^{-j} + \epsilon_j \right) \right) b_{n-j}
\]

\[
\overset{(19)}{=} \frac{m_n(\lambda - q)^{m-n}}{d_0} - \frac{m_n(\lambda - q)^{m-n}}{d_0}
\]

\[
- \epsilon_n \frac{(\lambda - q)^m}{\prod_{i=1}^{m}(\lambda - p_i)} - \sum_{j=1}^{n-1} \binom{n}{j} \left( m_j(\lambda - q)^{-j} + \epsilon_j \right) b_{n-j}
\]

\[
= \frac{\epsilon_n (\lambda - q)^m}{\prod_{i=1}^{m}(\lambda - p_i)} - \sum_{j=1}^{n-1} \binom{n}{j} \left( m_j(\lambda - q)^{-j} + \epsilon_j \right) b_{n-j}
\]

\[
|b_n| \leq k_n \hat{d}(q)
\]

\[
k_n = k_0 \frac{|\lambda - q|^m}{\prod_{i=1}^{m}(\lambda - p_i)} + \sum_{j=1}^{n-1} \binom{n}{j} \left( m_j|\lambda - q|^{-j} + k_j \right) k_{n-j}
\]

Thus, (27) holds. Combining (27) with (15) and (17) yields the result for Case (b).

For \( z \in \mathbb{C} \), let \( J(z) \) denote an elementary Jordan block with eigenvalue \( z \). In the proof of Lemma 2 this will refer to the elementary Jordan blocks of the \( A \) matrix of the plant. Corollary 2 will help bound the error between the optimal and approximating transfer functions in terms of \( D(\phi) \).

Corollary 2: Let \( k \) be a nonnegative integer, \( m \) a positive integer, \( z \in \mathbb{D} \), and \( q, \lambda, p_1, \ldots, p_m \in \mathbb{D} \) with \( d(\lambda, \{p_i\}^m_{i=1}) \geq \delta > 0 \) and \( d(\lambda, q) \geq \eta > 0 \). Choose constants \( c_{p_i} \) as in Lemma 4.

Then, the following holds.

a) There exists \( K > 0 \) such that

\[
\left| \sum_{i=1}^{m} c_{p_i} \frac{(\lambda - q)^{m+k}}{(\lambda - p_i)^{k+1}} \right| \leq K \hat{d}(q).
\]

b) There exists \( K > 0 \) such that

\[
\left| \sum_{i=1}^{m} c_{p_i} \frac{(p_i - q)^{m-k}}{(\lambda - p_i)^{k+1}} \frac{1}{z - p_i} \right| \leq K \hat{d}(q).
\]

Proof of Corollary 2: First, we prove Case (a). By Lemma 5(a), we have

\[
\lim_{z \to \lambda} \frac{d}{dz} \left( \frac{z - p_1}{1 - r(z) \prod_{i=1}^{m} (z - p_i)} \right) = 0
\]

Furthermore, by Lemma 5(a), the numerator/denominator satisfies

\[
1 - r(z) \prod_{i=1}^{m} (z - p_i) = 0, \quad \lim_{z \to \lambda} (\lambda - z)^k \prod_{i=1}^{m} (z - p_i) = 0.
\]

As both the numerator and denominator approach zero as \( z \to \lambda \), we can evaluate the limit using L'Hopital’s rule. For any \( l \in \{1, \ldots, k - 1\} \), by Lemma 5(a), differentiating the numerator and denominator \( l \) times and taking the limit as \( z \to \lambda \) imply

\[
\lim_{z \to \lambda} \frac{d}{dz^l} \left( \frac{z - p_1}{1 - r(z) \prod_{i=1}^{m} (z - p_i)} \right) = 0
\]

Thus, (27) holds. Combining (27) with (15) and (17) yields the result for Case (b).

For Case (b), we first recall the following fact.

Fact 1: If there exist \( k_{i,j} \) and \( d \) positive such that \( |M_{i,j}| \leq k_{i,j} d \) for all \( i, j \), then there exists \( K > 0 \) such that \( |M_{i,j}| \leq K d \).

By Fact 1, it suffices to show that for each \( l \in \{0, \ldots, m_q - 1\} \), there exists \( k_{l} \) such that the \( l \)th superdiagonal of the
matrix in the desired result satisfies
\[
\left| \sum_{i=1}^{m} c_{pi}(p_i - q)^m (-1)^i (\lambda - p_i)^{-(i+1)} \frac{1}{z - p_i} \right| \leq k_1 d(q).
\]

By Lemma 5(b) and since \( z \in \partial \mathbb{D} \)
\[
\left| \sum_{i=1}^{m} c_{pi}(p_i - q)^m (-1)^i (\lambda - p_i)^{-(i+1)} \frac{1}{z - p_i} \right| \leq \frac{\left( z - q \right)^m}{(\lambda - z)^{l+1} \prod_{i=1}^{m}(z - p_i)} + \sum_{n=0}^{l} a_n (|\lambda - z|)^{n-l+1} - \sum_{n=0}^{l} \frac{|a_n|}{|\lambda - z|^{n-l+1}} = 0
\]
Lemma 3(b) triangle inequality
\[
\leq \frac{(z - q)^m}{(\lambda - z)^{l+1} \prod_{i=1}^{m}(z - p_i)} + \sum_{n=0}^{l} |a_n| (1 - |\lambda|)^{n-l+1}
\]
\[
+ \sum_{n=0}^{l} \frac{|a_n|}{|\lambda - z|^{n-l+1}} \leq k_1 d(q)
\]
\[
k_i = K'(1 - |\lambda|)^{l-1} + K_0(1 - |\lambda|)^{-l-1} + \sum_{n=0}^{l} K_n (1 - |\lambda|)^{n-l-1}
\]

where such \( K' > 0 \) exists by Lemma 3. This proves Case (b). \( \square \)

**Lemma 6:** Let \( \tilde{m} \in \{0,1\} \), \( m_q, m > 0 \) be integers, and \( q \in \mathbb{D} \). Suppose that for each \( i \in \{1,\ldots,m\} \), we have matrices \( G_i \) and \( H_i \), and for each \( j \in \{1,\ldots,l\} \), we have matrices \( H_j \) and \( G_j \), and poles \( p_j \), such that
\[
H_j = c_j H_i^*_j, \quad G_j = -J(q - p_j)^{-1} B H_j^*.
\]

Then, the following holds:
\[
\sum_{i=1}^{m} \sum_{j=1}^{m_r} G_j \frac{1}{z - p_j} = - \sum_{l=1}^{m_q} J(0)^{-1} \sum_{i=1}^{m} B H_i^* \left( \frac{1}{z - q} \right)^l + \sum_{l=2}^{m_q+1} \sum_{i=1}^{m} J(0)^{-l-2} \sum_{i=1}^{m} B H_i^* \left( \frac{1}{z - q} \right)^l
\]
\[
+ \sum_{j=1}^{m} c_j J(0)^{-1} \sum_{i=1}^{m} B H_i^* \left( \frac{1}{z - q} \right)^l + \sum_{i=1}^{m} J(0)^{-l-1} \sum_{j=1}^{m} c_j (p_j - q)^{-1} B H_i^* + J(0)^{-1} \sum_{i=1}^{m} c_j (p_j - q)^{-1} B H_i^*
\]

Furthermore, for each \( i \in \{1,\ldots,m_q\} \) and \( l \in \{0,\ldots,m_q - 1\} \), there exists \( K_{i,l} > 0 \) such that each element in the \( l \)th super-diagonal of the term multiplying \( B H_i^* \) in (29) has a difference from \( \frac{1}{(z - q)^{m_q+1}} \) bounded in absolute value by \( K_{i,l} D(3) \).

Proof of Lemma 6: We begin by proving (29). For any \( i \in \{1,\ldots,m\} \), \( j \in \{1,\ldots,l\} \), and \( k \in \{1,\ldots,m_q - 1\} \), we have that \( J(0) G_j = -c_j J(0) (q - p_j)^{-1} B H_i^* \). Writing \( J(0) = J(q - p_j) + (p_j - q) I \) implies that \( J(0) G_j = -c_j B H_i^* - c_j (p_j - q) J(q - p_j)^{-1} B H_i^* \). Iterating this process yields
\[
J(0)^{k} G_j = \sum_{t=0}^{k-1} (-c_j J(0)^{-1} (p_j - q)^{t} B H_i^* - c_j (p_j - q)^{t} J(q - p_j)^{-1} B H_i^*).
\]
Then, applying Lemma 4(a) for \( k = 1 \), setting \( z = q \), and noting that for \( m = 1 \), \( c_j \) contains a factor of \( \frac{1}{p_j - q} \) give
\[
\sum_{j=1}^{m} c_j (p_j - q)^{t} = 0
\]
for any \( i \in \{1,\ldots,m_r\} \) and \( t \in \{1,\ldots,l - 2\} \). Furthermore, applying Lemma 4(a) for \( k = 1 \) and setting \( z = q \) imply
\[
\sum_{j=1}^{m} c_j = 0
\]
for any \( i \in \{1,\ldots,m_r\} \) and \( l = 0 \). Therefore, for any \( i \in \{1,\ldots,m_r\} \) and \( k \in \{1,\ldots,m_q - 1\} \)
\[
\tilde{m} J(0)^{k-1} B H_i^* - J(0)^{k} \sum_{j=1}^{m} G_j = \sum_{t=0}^{k-1} J(0)^{k-1-t} B H_i^* \sum_{j=1}^{m} c_j (p_j - q)^{t} + \sum_{j=1}^{m} c_j (p_j - q)^{t} J(q - p_j)^{-1} B H_i^* + \tilde{m} J(0)^{k-1} c_i B H_i^* \]
\[
\sum_{t=0}^{k-1} J(0)^{k-1-t} B H_i^* \sum_{j=1}^{m} c_j (p_j - q)^{t} + \sum_{j=1}^{m} c_j (p_j - q)^{t} J(q - p_j)^{-1} B H_i^* + J(0)^{k-1} B H_i^* \sum_{j=1}^{m} c_j + \tilde{m} J(0)^{k-1} c_i B H_i^*
\]

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
We compute

\[
\Phi_x(\mathbf{z}) = \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} G_{ij} \frac{1}{z - p_{ij}}
\]

combining terms in sum

\[
+ \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} \left( \tilde{m} J(0)^{l-2} B_{H_{i}} \right) \frac{1}{(z - q)^{l+2}}
\]

Note that

\[
\sum_{i=1+\hat{m}}^{m} \sum_{j=1+\hat{m}}^{i} G_{ij} \frac{1}{z - p_{ij}}
\]

\[
- \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} (p_{ij} - q)^{k} B_{H_{i}}
\]

\[
= \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} (p_{ij} - q) G_{ij} \frac{1}{z - q}
\]

\[
- \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} \frac{1}{z - q}
\]

\[
\sum_{i=1+\hat{m}}^{m} \sum_{j=1+\hat{m}}^{i} G_{ij} \left( \frac{1}{z - p_{ij}} - \frac{1}{z - q} \right)
\]

Also, we compute

\[
P_{x}(\mathbf{z}) = \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} G_{ij} \frac{1}{z - p_{ij}}
\]

\[
+ \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} \left( \tilde{m} J(0)^{l-2} B_{H_{i}} \right) \frac{1}{(z - q)^{l+2}}
\]

\[
- \sum_{i=1}^{m} \sum_{j=1+\hat{m}}^{i} \frac{1}{z - q}
\]

\[
\sum_{i=1+\hat{m}}^{m} \sum_{j=1+\hat{m}}^{i} G_{ij} \left( \frac{1}{z - p_{ij}} - \frac{1}{z - q} \right)
\]
Substituting (35)–(37) into (34) yields
\[
\Phi_x(z) = \left( \sum_{l=0}^{m-1} J(0)^l \tilde{m} l \frac{1}{(z-q)^{l+2}} \right) BH_i^* + \sum_{i=1+\tilde{m}}^{m} \left( \sum_{j=1+\tilde{m}}^{m} J(q-p_j)^{l-1} \frac{c_j (p_j^i - q)^{m_l} - p_j^i}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} + \tilde{m} c_i J(0)^{m_l+1} \frac{1}{(z-q)^{m_l+1}} BH_i^*.
\]

This completes the proof of (29).

Next, we derive the upper bound of the difference from \( \frac{1}{(z-q)^{l+1}} \) for the elements of the superdiagonal of the term multiplying BH_i^*. Note that, by the form of (29), all elements on each such superdiagonal are identical. By (29), for \( i \in \{1+\tilde{m}, \ldots, m_l \} \) and \( l \in \{m_l-t, \ldots, m_l-2\} \), the \( t \)th superdiagonal of the term multiplying BH_i^* in \( \Phi_x(z) \) is
\[
\left( \sum_{j=1+\tilde{m}}^{m} (-1)^l (q-p_j)^{(l+1)} \frac{c_j (p_j^i - q)^{m_l} - p_j^i}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} = \left( \sum_{j=1+\tilde{m}}^{m} (-1)^l (p_j^i - q)^{(l+1)} \frac{c_j (p_j^i - q)^{m_l} - p_j^i}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} = \left( \sum_{j=1+\tilde{m}}^{m} c_j (p_j^i - q)^{m_l-1-l} \frac{1}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}}
\]
where, for Lemma 4(a), note that for \( \tilde{m} = 1, c_j^i \) contains a factor of \( \frac{1}{p_j^i - q} \). For \( i \in \{1+\tilde{m}, \ldots, m_l \} \), the \( (m_l-1) \)th superdiagonal (i.e., \( l = m_l-1 \)) of the term multiplying BH_i^* in \( \Phi_x(z) \) is
\[
\left( \sum_{j=1+\tilde{m}}^{m} (-1)^{m_l-1} (q-p_j)^{m_l-1} \frac{-c_j^i (p_j^i - q)^{m_l} - p_j^i}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} + \tilde{m} c_i \frac{1}{(z-q)^{m_l+1}} = \left( \sum_{j=1+\tilde{m}}^{m} \frac{(-1)^{m_l-1} c_j^i (p_j^i - q)^{m_l}}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} + \tilde{m} c_i \frac{1}{(z-q)^{m_l+1}} = \left( \sum_{j=1+\tilde{m}}^{m} c_j^i \frac{1}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} + \tilde{m} c_i \frac{1}{(z-q)^{m_l+1}} = \left( \sum_{j=1+\tilde{m}}^{m} c_j^i \frac{1}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} + \tilde{m} c_i \frac{1}{(z-q)^{m_l+1}} = \left( \sum_{j=1+\tilde{m}}^{m} c_j^i \frac{1}{z-p_j^i} \right) \frac{1}{(z-q)^{m_l}} = \prod_{j=1+(z-p_j^i)}^{m_l-1} \frac{1}{(z-q)^{m_l}}
\]

Thus, combining the cases above, by Lemma 3 for every \( i \in \{1, \ldots, m_l \} \) and \( l \in \{0, \ldots, m_l-1\} \), each term in the \( t \)th superdiagonal of the term multiplying BH_i^* in (29) has a difference from \( \frac{1}{(z-q)^{l+1}} \) bounded by \( K_{ij} D(P) \).

Proof of Lemma 2: The proof begins by selecting an optimal solution \( \Phi_x^*, \Phi_y^* \) to (2), and constructing \( \Phi_v(z) = \sum_{c>0} H_c^{*}z \frac{1}{(z-q)^{l+1}} \) by [1, Th. 1] to approximate \( \Phi_v^* \). By [1, Th. 1], this implies that the approximation error bounds for \( \Phi_v^* \) of (12) are satisfied. Next, \( \Phi_x \) is defined as the unique solution to the SLS constraint in (4). The remainder of the proof will show that \( \Phi_x \) is a feasible solution to (7)–(9), and that it satisfies the approximation error bounds of (13).

Toward that end, first it is shown that it suffices to work in coordinates in which \( A \) is in Jordan normal form. Next it is shown that, in these coordinates, the approximation error bounds and the SLS constraints decouple according to each elementary Jordan block in \( A \), so it suffices to prove the result for a single elementary Jordan block with eigenvalue \( \lambda \). Afterward, it is
shown that the SLS constraint uniquely determines the poles and multiplicities of $\Phi_x$ from those of $\Phi_u$ for any transfer functions $(\Phi_u, \Phi_x)$ in $1/\mathcal{RH}_\infty$ that satisfy it. From the choice of $\Phi_u$, this immediately implies that $\Phi_x$ is a feasible solution to (7)-(9).

Subsequently, for each pole $q$ in $\Phi_u^*$ that appears in $\Phi_x^*$ by [1, Th. 1], there exist poles in $\Phi_x$ for approximating the portion of $\Phi_u^*$ corresponding to pole $q$. By the relationship between $\Phi_u$ and $\Phi_x$ described above, we then consider the resulting poles that appear in $\Phi_x$ and $\Phi_u$, and denote $\lambda - \Phi P J$ for $J$ by $\Phi$ and $\tau$ $(\tau \neq 0)$. Applying this fact to (43) and the poles used to approximate it, respectively, and then applying the resulting approximation error. As $q$ was arbitrary, this then yields the desired approximation error bounds for $\Phi_x$ of (13).

First, we obtain an optimal solution to (2), and use [1, Th. 1] to find $\Phi_u$, which closely approximates $\Phi_x$. Let $(\Phi_x, \Phi_u)$ be an optimal solution to (2), which exists by Assumption A6. By [1, Th. 1], there exist coefficient matrices $\{H_p\}_{p \in \mathbb{P}}$ such that, if we define $\Phi_u(z) = \sum_{p \in \mathbb{P}} H_p \frac{1}{z - q}$, then $\Phi_u \in 1/\mathcal{RH}_\infty$, $\|\Phi_u - \Phi_u^*\|_{H_\infty} \leq K_u^* D(\mathcal{P})$, and $\|\Phi_u - \Phi_u^*\|_{H_2} \leq K_u^2 D(\mathcal{P})$. Define $\Phi_x(z) = (zI - A)^{-1}(B\Phi_u(z) + I)$, and note that this implies that $(\Phi_x, \Phi_u)$ satisfy the SLS constraint in (4) by construction.

As $\Omega$ and $\sigma$ are finite, $\eta = \min_{\eta \in \Omega, \lambda \in \sigma, \lambda \neq q} d(\lambda, \eta) > 0$ and $d(\lambda, \eta) \geq \eta$ for all such $\lambda \neq q$. By Assumption A5, for every $q \in \Omega$ and $\lambda \in \sigma$ with $\lambda \neq q$, $d(\lambda, \mathcal{P}(q)) > 0$ where $\mathcal{P}(q)$ are the $m_i$ closest poles in $\mathcal{P}$ to $q$. This implies that $\delta = \min_{q \in \Omega, \lambda \in \sigma, \lambda \neq q} d(\lambda, \mathcal{P}(q)) > 0$ and that $d(\lambda, \mathcal{P}(q)) \geq \delta$ for all such $\lambda \neq q$.

Next, we show that it suffices to work in coordinates in which $A$ is in Jordan normal form, and that in these coordinates, the SLS constraints decouple according to each elementary Jordan block. There exist matrices $J$ in Jordan normal form and $V$ invertible such that $J = VAV^{-1}$. Fix $\zeta \in \partial\Omega$ for the remainder of the proof. We will show that there exists $K > 0$ such that

$$\|\Phi_x(z) - \Phi_x(z)\|_2 \leq K D(\mathcal{P}).$$

This will imply that

$$\|\Phi_x(z) - \Phi_x(z)\|_2 \leq \sup_{z \in \partial\Omega} \|V^{-1}(\Phi_x(z) - \Phi_x(z))\|_2 \leq K^* D(\mathcal{P}),$$

$$\|\Phi_x(z) - \Phi_x(z)\|_2 \leq \sqrt{\eta}\|\Phi_x(z) - \Phi_x(z)\|_2 \leq K^2 D(\mathcal{P}),$$

$$K^* = \sup_{z \in \partial\Omega} \|V^{-1}K\|, \quad K^2 = \eta K^{*}.\quad (38)$$

So, to prove the lemma, it suffices to show that (38) holds and that $(\Phi_x, \Phi_u)$ is a feasible solution to (7)-(9). Let $J(\lambda)$ denote an elementary Jordan block with eigenvalue $\lambda$ in $J$, $M(\lambda)$ denote the restriction of the matrix $M$ to the rows corresponding to the rows of $J(\lambda)$ in $J$, and $M(\lambda)^0$ the concatenation of $M(\lambda)$ with rows of zeros. Decomposing $\Phi_x z$ and $\Phi_x^* z$ by rows gives

$$\|\Phi_x(z) - \Phi_x(z)\|_2 = \sup_{z \in \partial\Omega} \|V^{-1}(\Phi_x(z) - \Phi_x(z))\|_2 \leq \sum_{\lambda \in \sigma} \sum_{\eta \in \mathcal{P}(q)} |\Phi_x(z)|_{J(\lambda)} - |\Phi_x(z)|_{J(\lambda)}|\eta\|_2.$$
that $\Phi_x \in {\mathcal L}^2_{R(x)}$ and is a feasible solution to (7)–(9). Hence, to complete the proof, it suffices to prove (40).

In what follows, we show that to prove (13), it suffices to fix a particular pole in $\Phi_x$, and to show that a certain portion of $\Phi_x$ closely approximates the portion of $\Phi_x$, corresponding to this pole. This is done by using the construction of [1, Th. 1] to approximate $\Phi_x$ by $\Phi_u$. Let $\Omega$ denote the poles of $\Phi_x$. For each $q \in \Omega$, its contribution to the partial fraction decompositions of $\Phi_x$ and $\Phi_u$ is given, respectively, by

$$
\sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i} = c_j \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + c_j \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i},
$$

by the above fact. Since $\Phi_u$ was constructed as in [1, Th. 1], the portion of $\Phi_u$ that was chosen to approximate the pole at $q$ in $\Phi_u$ is given by

$$
\sum_{i=1}^{m_q} \sum_{j=1}^{m_q} \frac{1}{(z-q)^i} = \sum_{i=1}^{m_q} \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i}.
$$

Hence, from (45) and (47), we compute

$$
\|\Phi_x(z) - \Phi_u(z)\|_2 \\
= \left\| \sum_{q \in \Omega} \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i} \right\|_2 \\
\leq \sum_{q \in \Omega} \left\| \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} \right\|_2 + \sum_{j=1}^{m_q} \left\| \frac{1}{(z-p_j)^i} \right\|_2 \\
= \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i}.
$$

so, since $\Omega$ is finite, to prove (40), it suffices to show that there exists $K_q > 0$ such that

$$
\left\| \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i} \right\|_2 \leq K_q D(\mathcal{P})
$$

for each $q \in \Omega$. Toward that end, fix $q \in \Omega$, and for the remainder of the proof, let $\Phi_x(z)$ and $\Phi_u(z)$ denote the contributions to $\Phi_x(z)$ and $\Phi_u(z)$ given by (47) and (45), respectively, as in (48). We consider two cases.

Case 1: $q \neq \lambda$. Substituting (45)–(47) into (42)–(44) implies that

$$
-J(\lambda - q)G_m = BH_i
$$

$$
G_{i-1} = (\lambda - q)^{-1}(G_i - BH_{i-1}), \ i \in \{2, \ldots , m\}
$$

and all other coefficients in $\Phi_u$ and $\Phi_x$ are zero. The above implies that $G_m = i \sum_{q \in \Omega} J(\lambda - q)^{-i}BH_i$ for all $i \in \{1, \ldots , m\}$. Define $G_{l,i} = J(\lambda - q)^{-i}BH_i$ for all $i \in \{1, \ldots , m\}$ and $i \in \{l, \ldots , m\}$, and note that $G_i = \sum_{i=1}^{m} G_{i,i}$.

Write $G_i^* = c_j (G_{i,i}^* + \Delta G_i^*)$. Then, by (46)

$$
J(\lambda - q)G_i^*(i,i) = -BH_i^* = -\frac{1}{c_j} BH_i = -\frac{1}{c_j} J(\lambda - p_j)G_i^* = J(\lambda - p_j)G_i^* + \Delta G_i^* = (J(\lambda - q) - (q - p_j)I)(G_{i,i}^* + \Delta G_i^*)
$$

so $0 = -(p_j - q)G_i^*(i,i) + J(\lambda - p_j)\Delta G_i^*$ and $\Delta G_i^* = (p_j - q)J(\lambda - p_j)^{-1}G_i^*$. In summary

$$
\Phi_u = \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{1}{(z-q)^i} = \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{1}{(z-p_j)^i}.
$$

For any $i \in \{1, \ldots , m\}$ and $l \in \{2, \ldots , i\}$, write $J(\lambda - p_j)^{-1}G_{i,l} = G_{i,l-1} + \Delta G$. Note that for $i \in \{1, \ldots , m\}$ and $l \in \{2, \ldots , i\}$, $J(\lambda - q)G_{i,l-1} = G_{i,l}$.

Applying this equation recursively implies that for

$$
G_i^* = c_j^* G_{i,i} + c_j (p_j - q)J(\lambda - p_j)^{-1}G_i^*,
$$

so, since $\Omega$ is finite, to prove (40), it suffices to show that there exists $K_q > 0$ such that

$$
\left\| \sum_{i=1}^{m_q} \frac{1}{(z-q)^i} + \sum_{j=1}^{m_q} \frac{1}{(z-p_j)^i} \right\|_2 \leq K_q D(\mathcal{P})
$$

for each $q \in \Omega$. Toward that end, fix $q \in \Omega$, and for the remainder of the proof, let $\Phi_x(z)$ and $\Phi_u(z)$ denote the contributions to $\Phi_x(z)$ and $\Phi_u(z)$ given by (47) and (45), respectively, as in (48). We consider two cases.

Case 1: $q \neq \lambda$. Substituting (45)–(47) into (42)–(44) implies that

$$
-J(\lambda - q)G_m = BH_i
$$

$$
G_{i-1} = (\lambda - q)^{-1}(G_i - BH_{i-1}), \ i \in \{2, \ldots , m\}
$$

Therefore, the following holds:

$$
\Phi_x(z) = \sum_{i=1}^{m} \sum_{j=1}^{i} \frac{1}{(z-p_j)}
$$

above and

$$
+ \sum_{i=1}^{m} \sum_{j=1}^{i} c_j (p_j - q)J(\lambda - p_j)^{-1}G_{i,i}^* \frac{1}{(z-p_j)}
$$
reverse sum order
\begin{align*}
\sum_{i=1}^{m} \sum_{j=1}^{i} G_{i}^{*}(t, i, j) \left( \sum_{j=1}^{i} c_{j}^{*}(p_{j}^{i} - q)^{i-l} \frac{1}{z-p_{j}^{i}} \right) \\
+ \sum_{i=1}^{m} \left( \sum_{j=1}^{i} c_{j}^{*}(p_{j}^{i} - q)^{i} J(\lambda - p_{j}^{i})^{-1} \frac{1}{z-p_{j}^{i}} \right) G_{i}^{*}(t, i, j) \\
\end{align*}

Lemma 4(a)
\begin{align*}
\sum_{i=1}^{m} \sum_{j=1}^{i} G_{i}^{*}(t, i, j) (z-q)^{i-l} \prod_{j=1}^{i} (z-p_{j}^{i}) \\
+ \sum_{i=1}^{m} \left( \sum_{j=1}^{i} c_{j}^{*}(p_{j}^{i} - q)^{i} J(\lambda - p_{j}^{i})^{-1} \frac{1}{z-p_{j}^{i}} \right) G_{i}^{*}(t, i, j).
\end{align*}
Thus, the following holds:
\begin{align*}
\Phi_{x}(z) - \Phi_{x}^{*}(z) = \sum_{i=1}^{m} \sum_{j=1}^{i} G_{i}^{*}(t, i, j) \left( \frac{(z-q)^{i-l}}{\prod_{j=1}^{i} (z-p_{j}^{i})} - \frac{1}{(z-q)^{i}} \right) \\
+ \sum_{i=1}^{m} \left( \sum_{j=1}^{i} c_{j}^{*}(p_{j}^{i} - q)^{i} J(\lambda - p_{j}^{i})^{-1} \frac{1}{z-p_{j}^{i}} \right) G_{i}^{*}(t, i, j).
\end{align*}
This implies
\[||\Phi_{x}(z) - \Phi_{x}^{*}(z)||_{2} \leq \sum_{i=1}^{m} ||G_{i}^{*}(t, i, i)||_{2} \left( \sum_{j=1}^{i} c_{j}^{*}(p_{j}^{i} - q)^{i} \right) J(\lambda - p_{j}^{i})^{-1} \frac{1}{z-p_{j}^{i}} ||G_{i}^{*}(t, i, i)||_{2},\]
which proves (48) for Case 1.

It will be useful to derive an additional bound for use in the proof of Case 2. In particular, we want to show that there exist constants \( K_{i} > 0 \) for \( i \in \{1, \ldots, m\} \) such that
\[||\sum_{j=1}^{i} G_{j}^{*} - G_{i}^{*}(1, i)||_{2} \leq K_{i} D(P).\]  
(49)

Write \( G_{j}^{*} = c_{j}^{*}(G_{i}^{*}(t, i, i) + \Delta G) \). Then
\[J(\lambda - q)^{*}(G_{i}^{*}(1, i)) = -BH_{i}^{*} = \frac{1}{c_{j}^{*}} J(\lambda - p_{j}^{i})^{*}G_{j}^{i} = J(\lambda - p_{j}^{i})(G_{i}^{*}(1, i) + \Delta G),\]
so \( \Delta G = J(\lambda - p_{j}^{i})^{-1}J(\lambda - q)^{i}G_{i}^{*}(1, i) - G_{i}^{*}(1, i) \), which implies that \( G_{j}^{*} = c_{j}^{*}J(\lambda - p_{j}^{i})^{-1}J(\lambda - q)^{i}G_{i}^{*}(1, i) \). Thus, the following holds:
\[||\sum_{j=1}^{i} G_{j}^{*} - G_{i}^{*}(1, i)||_{2} = \left|\sum_{j=1}^{i} c_{j}^{*}J(\lambda - p_{j}^{i})^{-1}J(\lambda - q)^{i}G_{i}^{*}(1, i) \right|_{2} - J(\lambda - q)^{-1}J(\lambda - q)^{i}G_{i}^{*}(1, i) \right|_{2} \leq \left|\sum_{j=1}^{i} c_{j}^{*}J(\lambda - p_{j}^{i})^{-1} - J(\lambda - q)^{-1} \right|_{2} \left|J(\lambda - q)^{i}G_{i}^{*}(1, i) \right|_{2},\]

Therefore, in order to prove (49), it suffices to show that
\[\left|\sum_{j=1}^{i} c_{j}^{*}J(\lambda - p_{j}^{i})^{-1} - J(\lambda - q)^{-1} \right|_{2} \leq K_{i} D(P) \]  
(50)
for some constants \( K_{i} > 0 \). For \( l \in \{0, \ldots, m - 1\} \), the \( l \)th superdiagonal of \( \sum_{j=1}^{i} c_{j}^{*}J(\lambda - p_{j}^{i})^{-1} \) is given by
\[\sum_{j=1}^{i} c_{j}^{*}(\lambda - p_{j}^{i})^{-l} \left( \lambda - p_{j}^{i} \right)^{-l+1} = (-1)^{l} \sum_{j=1}^{i} c_{j}^{*}(\lambda - p_{j}^{i})^{-l+1},\]
thus, for each \( i \in \{1, \ldots, m\} \) and \( l \in \{0, \ldots, m - 1\} \), the difference between terms in superdiagonal \( l \) of the matrix in (50) is \( c_{i}(i, i) \), which satisfies that \( |c_{i}(i, i)| \leq K_{i} D(P) \). Therefore, by Fact 1 in the proof of Corollary 2, this implies that there exist \( K_{i} > 0 \) such that (50) holds.

Case 2: \( q = \lambda \). Let \( \bar{Q} \) denote the poles in \( \Phi_{x} \). Substituting (45)–(47) into (41)–(44) implies that
\[J(0)G_{m+q}^{*} = 0, \quad G_{i}^{*} = V - \sum_{q \neq q_{i}} G_{i}^{*}(q_{i}, q) \]
\[G_{i+1}^{*} = J(0)G_{i}^{*}, \quad i \in \{m + 1, \ldots, m + q + m - 1\} \]
\[G_{i+1}^{*} = J(0)G_{i}^{*} + BH_{i}^{*}, \quad i \in \{1, \ldots, m\} \]
\[J(0)G_{m+q}^{*} + \bar{m} = 0, \quad G_{i}^{*} = V - \sum_{q \neq q_{i}} G_{i}^{*}(q_{i}, q) \]
\[G_{i+1}^{*} = J(0)G_{i}^{*}, \quad i \in \{2, \ldots, m + \bar{m} - 1\} \]
\[G_{i+1}^{*} = J(0)G_{i}^{*} + \bar{m} \sum_{j=1}^{m} BH_{i}^{j}, \quad i \in \{1 + \bar{m} + \bar{m} - \bar{m}, \ldots, \bar{m} + \bar{m}\} \]
\[j \in \{1 + \bar{m} + \bar{m}, \ldots, \bar{m}\} \]

where \( G_{i}^{*}(q_{i}, q) \) and \( G_{i}^{*}(q_{i}, q) \) denote the coefficients of \( \frac{1}{z-q} \) in \( \Phi_{x} \) and \( \Phi_{x} \), respectively, for the pole \( \hat{q} \). For \( l \in \{1, \ldots, m\} \), define \( G_{l} = G_{l}^{*}(1 + \hat{m} + \hat{m}) \), \( G_{l} = J(0)^{l-1}G_{l} \), and \( G_{l} = J(0)^{l-1}G_{l}^{*} \). By (49), from Case 1
\[
K_1 D(\mathcal{P}), \quad K_1 = \sum_{q \in \Omega}^m \sum_{\tilde{q} \neq q}^{m_q} K_i(q, \tilde{q}).
\]

Thus, for \(l \in \{1, \ldots, m_q\}\), we have
\[
||\tilde{G}_l - \hat{G}_l||_2 \leq \|J(0)^{l-1}\|_2 K_1.\]
where \(K_1 = \|J(0)^{l-1}\|_2 K_1\). For \(l \in \{2, \ldots, m_q + \tilde{m}\}\), define
\[
\tilde{G}_l = G_l - \hat{G}_l = -\sum_{i=1}^{m_i} \sum_{j=1+\tilde{m}} G_{i,j}^l \]
\[
\tilde{G}_l = -J(0)^{l-1} \sum_{i=1}^{m_i} \sum_{j=1+\tilde{m}} G_{i,j}^l.
\]

By (52) and (55), we have
\[
\Phi^*_z = \sum_{i=1}^{m_q} \sum_{\tilde{m} = 0}^{m_q - 1} J(i)^{-1} B H_i^* \]

Therefore, for \(i \in \{1, \ldots, m_q\}\) and \(l \in \{0, \ldots, m_q - 1\}\), the \(l\)th superdiagonal of the term multiplying \(B H_i^*\) in \(\Phi^*_z\) is given by
\[
\frac{z}{(z-q)^{l+\tilde{m}}}.\]
Thus, by Lemma 6, for every \(j, j' \in \{1, \ldots, m_q\}\)
\[
\frac{(\Phi^*_x(z) - \Phi^*_x(z))_{j,j'}}{||\Phi^*_x(z) - \Phi^*_x(z)||_2} \leq K(\tilde{m}) D(\mathcal{P}) \frac{1}{(z-q)^{l+\tilde{m}}}.\]

By Fact 1 in the proof of Corollary 2, this implies that \(||\Phi^*_x(z) - \Phi^*_x(z)||_2 \leq K D(\mathcal{P})\) for some \(K > 0\), which proves (48) for Case 2.

Theorem 1 applies the approximation error bounds of Lemma 2 to the optimal solution of (2) to obtain the desired suboptimality bounds.

Proof of Theorem I: Let \((\Phi^*_x, \Phi^*_u)\) be an optimal solution to (2). By Lemma 2, there exist \(\Phi^*_x, \Phi^*_u \in \frac{1}{2} \mathbb{R}^n_{\infty}\), which are a feasible solution to (7)–(9) and satisfy the approximation error bounds (12) and (13). Letting \(J(\Phi^*_x, \Phi^*_u)\) denote the value of the objective of (2) for \((\Phi^*_x, \Phi^*_u)\), we compute
\[
J(\mathcal{P}) \leq J(\Phi^*_x, \Phi^*_u)
\]
by adding and
\[
\begin{align*}
\triangle inequality 
\leq \left\| C(\Phi^*_x(z) - \Phi^*_x(z)) B + D(\Phi^*_u(z) - \Phi^*_u(z)) \tilde{B} - T_{desired}(z) \right\|_{\infty} \\
+ \lambda \left\| C(\Phi^*_x(z) - \Phi^*_x(z)) B + D(\Phi^*_u(z) - \Phi^*_u(z)) \tilde{B} - T_{desired}(z) \right\|_{\infty} \\
+ \left\| C(\Phi^*_x(z) - \Phi^*_x(z)) B + D(\Phi^*_u(z) - \Phi^*_u(z)) \tilde{B} - T_{desired}(z) \right\|_{\infty} \\
+ \lambda \left\| C(\Phi^*_x(z) - \Phi^*_x(z)) B + D(\Phi^*_u(z) - \Phi^*_u(z)) \tilde{B} - T_{desired}(z) \right\|_{\infty}
\end{align*}
\]
This work combined SLS with SPA to develop a new control design method. Unlike DBC, SPA does not result in DBC, feasibility is automatic so it does not require slack variables, which lead to additional suboptimality, and it can be solved by a single SDP, as opposed to the iterative algorithm that DBC requires. A suboptimality certificate was provided for SPA that, unlike the DBC bound, does not require a sufficiently long time horizon that the optimal impulse response has already decayed, and does not depend on this decay rate. The bound is specialized for the Archimedes spiral pole selection [1]. An example shows that SPA achieves much better matching with the optimal solution than DBC with orders of magnitude fewer poles. Future work should address extensions to state and input constraints, application of SPA to output feedback, extensions to continuous-time, static controllers, and extensions to time-varying systems and uncertainty.

VI. CONCLUSION

This work combined SLS with SPA to develop a new control design method. Unlike DBC, SPA does not result in DBC, feasibility is automatic so it does not require slack variables, which lead to additional suboptimality, and it can be solved by a single SDP, as opposed to the iterative algorithm that DBC requires. A suboptimality certificate was provided for SPA that, unlike the DBC bound, does not require a sufficiently long time horizon that the optimal impulse response has already decayed, and does not depend on this decay rate. The bound is specialized for the Archimedes spiral pole selection [1]. An example shows that SPA achieves much better matching with the optimal solution than DBC with orders of magnitude fewer poles. Future work should address extensions to state and input constraints, application of SPA to output feedback, extensions to continuous-time, static controllers, and extensions to time-varying systems and uncertainty.

REFERENCES

[1] M. W. Fisher, G. Hug, and F. Dörfler, “Approximation by simple poles—Part I: Density and geometric convergence rate in Hardy space,” IEEE Trans. Autom. Control, vol. 69, no. 8, pp. 4904–4909, Aug. 2024.
[2] G. Zames and B. A. Francis, “Feedback, minimax sensitivity, and optimal robustness,” IEEE Trans. Autom. Control, vol. AC-28, no. 5, pp. 585–601, May 1983.
[3] D. C. Youla, H. A. Jabr, and J. J. Bongiorno, “Modern Wiener-Hopf design of optimal controllers—Part I: The multivariable case,” IEEE Trans. Autom. Control, vol. AC-21, no. 3, pp. 319–338, Jun. 1976.
[4] A. Desoer and M. J. Chen, “Design of multivariable feedback systems with stable plant,” IEEE Trans. Autom. Control, vol. AC-26, no. 2, pp. 408–415, Apr. 1981.
[5] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour, “An input-output parameterization of stabilizing controllers: Amidst Youla and system level synthesis,” IEEE Control Syst. Lett., vol. 3, no. 4, pp. 1014–1019, Oct. 2019.
[6] Y.-S. Wang, N. Matni, and J. C. Doyle, “Separable and localized system-level synthesis for large-scale systems,” IEEE Trans. Autom. Control, vol. 63, no. 12, pp. 4234–4249, Dec. 2018.
[7] Y.-S. Wang, N. Matni, and J. C. Doyle, “A system-level approach to controller synthesis,” IEEE Trans. Autom. Control, vol. 64, no. 10, pp. 4079–4093, Oct. 2019.
[8] D. S. Bernstein and W. M. Haddad, “LQG control with an $H_\infty$ performance bound: A Ricatti equation approach,” IEEE Trans. Autom. Control, vol. 34, no. 3, pp. 293–305, Mar. 1989.
[9] J. Anderson, J. C. Doyle, S. H. Low, and N. Matni, “System level synthesis,” Annu. Rev. Control, vol. 47, pp. 364–393, 2019.
[10] S. Urikura and A. Nagata, “Ripple-free deadbeat control for sampled-data systems,” IEEE Trans. Autom. Control, vol. AC-32, no. 6, pp. 474–482, Jun. 1987.
[11] K. Zhou, J. C. Doyle, and K. Glover, Robust and Optimal Control, Englewood Cliffs, NJ, USA: Pearson, 1995.
[12] S. Jiang and P. G. Voulgaris, “Performance optimization of switched systems: A model matching approach,” IEEE Trans. Autom. Control, vol. 54, no. 9, pp. 2058–2071, Sep. 2009.
[13] A. Abdullah and M. Zribi, “Model reference control of LPV systems,” J. Franklin Inst., vol. 346, pp. 854–871, 2009.
[14] B. A. Francis and W. M. Wonham, “The internal model principle of control theory,” Automatica, vol. 12, pp. 457–465, 1976.
[15] H. Kwakernaak and R. Sivan, Linear Optimal Control Systems. Hoboken, NJ, USA: Wiley, 1972.
[16] “Git repository,” 2022. [Online]. Available: https://github.com/michaelwfisher/SystemLevelSynthesis
[17] G. E. Dullerud and F. Paganini, A Course in Robust Control Theory (Texts in Applied Mathematics Series), vol. 36, New York, NY, USA: Springer, 2000.
[18] D. E. Walker and D. B. Ridgely, “Reduced order mixed $H_2/H_\infty$ optimization with a singular $H_\infty$ constraint,” in Proc. Amer. Control Conf., 1994, pp. 1128–1132.
[19] L. Huang, H. Xin, and F. Dörfler, “$H_\infty$-control of grid-connected inverters: Design, objectives and decentralized stability certificates,” IEEE Trans. Smart Grid, vol. 11, no. 5, pp. 3805–3816, Sep. 2020.
[20] I. Subotic, D. Groš, M. Colombino, and F. Dörfler, “A Lyapunov framework for nested dynamical systems on multiple time scales with application to converter-based power systems,” IEEE Trans. Autom. Control, vol. 66, no. 12, pp. 5969–5984, Dec. 2021.
[21] J. Yu, Y.-S. Wang, and J. Anderson, “Localized and distributed $H_\infty$ state feedback control,” in Proc. Amer. Control Conf., 2021, pp. 2732–2738.
[22] G. H. Golub and C. F. V. Loan, Matrix Computations, 3rd ed. Baltimore, MD, USA: The Johns Hopkins Univ. Press, 1996.

Michael W. Fisher received the M.Sc. degree in mathematics and the Ph.D. degree in electrical engineering: systems from the University of Michigan, Ann Arbor, MI, USA, in 2017 and 2020, respectively. He is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. He was a Postdoctoral Researcher with ETH Zurich. His research interests include dynamics, control, and optimization of complex systems.

Gabriela Hug received the M.Sc. degree in electrical engineering and the Ph.D. degree in information technology and electrical engineering from ETH Zurich, Zurich, Switzerland, in 2004 and 2008, respectively. After her Ph.D., she worked with Hydro One, Toronto, ON, Canada, and from 2009 to 2015, she was an Assistant Professor with Carnegie Mellon University, Pittsburgh, PA, USA. She is currently a Professor with Power Systems Laboratory, ETH Zurich. Her research focuses on the control and optimization of electric power systems.

Florian Dörfler received the diploma degree in engineering cybernetics from the University of Stuttgart, Stuttgart, Germany, in 2008, and the Ph.D. degree in mechanical engineering from University of California at Santa Barbara, Santa Barbara, CA, USA, in 2013. From 2013 to 2014, he was an Assistant Professor with University of California, Los Angeles, CA, USA. He is an Associate Professor with Automatic Control Laboratory, ETH Zurich, Zurich, Switzerland, and the Associate Head with the Department of Information Technology and Electrical Engineering. His primary research interests include control, optimization, and system theory with applications in network systems.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.