Minimum-gain Pole Placement with Sparse Static Feedback

Vaibhav Katewa and Fabio Pasqualetti

Abstract—The minimum-gain eigenvalue assignment/pole placement problem (MGEAP) is a classical problem in LTI systems with static state feedback. In this paper, we study the MGEAP when the state feedback has arbitrary sparsity constraints. We formulate the sparse MGEAP problem as an equality-constrained optimization problem and present an analytical characterization of its solution in terms of eigenvector matrices of the closed loop system. This result is used to provide a geometric interpretation of the solution of the non-sparse MGEAP, thereby providing additional insights for this classical problem. Further, we develop an iterative projected gradient descent algorithm to solve the sparse MGEAP using a parametrization based on the Sylvester equation. We present a heuristic algorithm to compute the projections, which also provides a novel method to solve the sparse EAP. Also, a relaxed version of the sparse MGEAP is presented and an algorithm is developed to obtain approximately sparse solutions to the MGEAP. Finally, numerical studies are presented to compare the properties of the algorithms, which suggest that the proposed projection algorithm converges in almost all instances.

I. INTRODUCTION

The Eigenvalue/Pole Assignment Problem (EAP) using static state feedback is one of the central problems in the design of Linear Time Invariant (LTI) control systems. It plays a key role in system stabilization and shaping its transient behavior. Given the following LTI system

$$\mathcal{D}x(k) = Ax(k) + Bu(k), \tag{1a}$$
$$u(k) = Fx(k), \tag{1b}$$

where $x \in \mathbb{R}^n$ is the state of the LTI system, $u \in \mathbb{R}^m$ is the control input, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $\mathcal{D}$ denotes either the continuous time differential operator or the discrete-time shift operator, the EAP involves finding a real feedback matrix $F \in \mathbb{R}^{m \times n}$ such that the eigenvalues of the closed loop matrix $A_c(F) \triangleq A + BF$ coincide with a given set $\mathcal{S} = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ that is closed under complex conjugation.

It is well known that the existence of F depends on the controllability properties of the pair (A, B). Further, for single input systems ($m = 1$), the feedback vector that assigns the eigenvalues is unique and can be obtained using the Ackermann's formula [1]. On the other hand, for multi-input systems ($m > 1$), the feedback matrix is not unique and there exists a flexibility to choose the eigenvectors of the closed loop system. This flexibility can be utilized to choose a feedback matrix that satisfies some auxiliary control criteria in addition to assigning the eigenvalues. For instance, the feedback matrix can be chosen to minimize the sensitivity of the closed loop

system to perturbations in the system parameters, thereby making the system robust. This is known as Robust Eigenvalue Assignment Problem (REAP). Alternatively, one can choose the feedback matrix with minimum gain, thereby reducing the overall control effort. This is known as Minimum Gain Eigenvalue Assignment Problem (MGEAP). Both of these problems have been studied extensively [2]- [3].

Recently, considerable attention has been given to the study and design of sparse feedback control systems, where certain entries of the matrix F are required to be zero. Feedback sparsity typically arises in large scale systems in which the controllers do not have access to all the states of the system. Sparsity may also be a result of the special structure of the control system which prohibits feedback from some states to the controllers.

The feedback design problem with sparsity constraints is considerably more difficult than the unconstrained case. There have been numerous studies to determine the optimal feedback control law for H2/LQR/LQG problems with sparsity, particularly when the controllers have access to only local information (see [4]–[7] and the references therein). While the optimal H2/LQR/LQG design problems with sparsity have a rich history, studies on the REAP/MGEAP in the presence of *arbitrary* sparsity constrains are lacking. Even the problem of finding a particular (not necesary optimal) sparse feedback matrix that solves the EAP is not well studied. In this paper, we study the EAP and MGEAP with *arbitrary* sparsity constraints on the feedback matrix F. We provide analytical characterization for the solution of sparse MGEAP and provide iterative algorithms to solve the sparse EAP and MGEAP.

Related work There have been numerous studies on the optimal pole placement problem *without* sparsity constraints. For the REAP, authors have considered optimizing different metrics which capture the sensitivity of the eigenvalues, such as the condition number of the eigenvector matrix [2], [8]–[11], departure from normality [12] and others [13], [14]. Most of these methods use gradient-based iterative procedures to obtain the solutions. For surveys and performance comparisons of these REAP methods, see [8], [15], [16] and the references therein.

Early works for MGEAP, including [17], [18], presented approximate solutions using low rank feedback and successive pole placement techniques. Simultaneous robust and minimum gain pole placement were studied in [11], [19]–[21]. For a survey and performance comparison of these MGEAP studies, see [3] and the references therein. The regional pole placement problem was studied in [22], [23], where the eigenvalues were assigned inside a specified region. While all these studies have provided useful insights on REAP/MGEAP, they do not consider sparsity constraints on the feedback matrix. Further, the techniques in these studies cannot be used or simply

extended to obtain a sparse solution. In contrast, we study the sparse EAP/MGEAP by explicitly including the sparsity constraints into the problem.

There have also been numerous studies on EAP with sparse *dynamic* LTI feedback. The concept of decentralized fixed modes (DFMs) was introduced in [24] and later refined in [25]. Decentralized fixed modes are those eigenvalues of the system which cannot be shifted using a static/dynamic feedback with *fully decentralized* sparsity pattern (i.e. the case where controllers have access only to local states). The remaining eigenvalues of the system can be arbitrarily assigned. However, this cannot be achieved in general using a static decentralized controller and requires the use of *dynamic* decentralized controller [24]. Other algebraic characterizations of the DFMs were presented in [26], [27]. The notion of DFMs was generalized for an *arbitrary* sparsity pattern and the concept of structurally fixed modes (SFMs) was introduced in [28]. Graph theoretical characterizations of structurally fixed modes were provided in [29], [30]. As in the case of DFMs, assigning the non-SFMs also requires *dynamic* controllers. These studies on DFMs and SFMs present feasibility conditions and analysis methods for the EAP problem with sparse *dynamic* feedback. In contrast, we study both EAP and MGEAP with sparse *static* controllers, assuming the sparse EAP is feasible. We remark that EAP with sparsity and static feedback controller is in fact important for several network design and control problems, and easier to implement than its dynamic counterpart.

Recently, there has been a renewed interest in studying linear systems with sparsity constraints. Using a different approach than [29], the original results regarding DFMs in [24] were generalized for an arbitrary sparsity pattern by the authors in [31], [32], where they also present a sparse *dynamic* controller synthesis algorithm. Further, there have been many recent studies on minimum cost input/output and feedback sparsity pattern selection such that the system has no structurally fixed modes (see [33], [34] and the references therein). In contrast, we consider the problem of finding a *static* minimum gain feedback with a *given* sparsity pattern that solves the EAP.

Contribution The contribution of this paper is three-fold. First, we study the MGEAP with static feedback and arbitrary sparsity constraints (assuming feasibility of sparse EAP). We formulate the sparse MGEAP as an equality constrained optimization problem and present an analytical characterization of an optimal sparse solution. As a minor contribution, we use this result to provide a geometric insight for the non-sparse MGEAP solutions. Second, we develop two heuristic iterative algorithms to obtain a solution of the sparse EAP. The first algorithm is based on repeated projections on linear subspaces. The second algorithm is developed using the Sylvester equation based parametrization and it obtains a solution via projection of a non-sparse feedback matrix on the space of sparse feedback matrices that solve the EAP. Third, using the latter EAP projection algorithm, we develop a projected gradient descent method to solve the sparse MGEAP. We also formulate a relaxed version of the sparse MGEAP using penalty based optimization method and develop an algorithm to obtain approximately-sparse solutions.

Paper organization The remainder of the paper is organized as follows. In Section II we formulate the sparse MGEAP optimization problem. In Section III, we obtain the solution of the optimization problem using the Lagrangian theory of optimization. We also provide a geometric interpretation for the optimal solutions of the non-sparse MGEAP. In Section IV, we present two heuristic algorithms for solving the sparse EAP. Further, we present a projected gradient descent algorithm to solve the sparse MGEAP and also an approximately-sparse solution algorithm for a relaxed version of the sparse MGEAP. Section V contains numerical studies and comparisons of the proposed algorithms. Finally, we conclude the paper in Section VI.

II. SPARSE MGEAP FORMULATION

A. Mathematical Notation and Properties

We use the following notation throughout the paper:

$\|\cdot\|_2$	Spectral norm
$\|\cdot\|_F$	Frobenius norm
$<\cdot,\cdot>_F$	Inner (Frobenius) product
$|\cdot|$	Cardinality of a set
$\Gamma(\cdot)$	Spectrum of a matrix
$\text{tr}(\cdot)$	Trace of a matrix
$(\cdot)^+$	Moore-Penrose pseudo inverse
$(\cdot)^\mathsf{T}$	Transpose of a matrix
$\mathcal{R}(\cdot)$	Range of a matrix
$A > 0$	Positive definite matrix A
$\circ$	Hadamard (element-wise) product
$\otimes$	Kronecker product
$(\cdot)^*$	Complex conjugate
$(\cdot)^\mathsf{H}$	Conjugate transpose
$\text{supp}(\cdot)$	Support of a vector
$\text{vec}(\cdot)$	Vectorization of a matrix
$\text{diag}(a)$	$n \times n$ Diagonal matrix with diagonal elements given by $n-$dim vector a
$\text{Re}(\cdot)$	Real part of a complex variable
$\text{Im}(\cdot)$	Imaginary part of a complex variable
$1_n(0_n)$	n-dim vector of ones (zeros)
$1_{n\times m}(0_{n\times m})$	$n \times m$-dim matrix of ones (zeros)
I_n	n-dim identity matrix
e_i	i-th canonical vector
$T_{m,n}$	Permutation matrix that satisfies $\text{vec}(A^\mathsf{T}) = T_{m,n}\text{vec}(A), A \in \mathbb{R}^{m\times n}$

The Kronecker sum of two square matrices A and B with dimensions n amd m, respectively, is denoted by

$$A \oplus B = (I_m \otimes A) + B \otimes I_n.$$

Further, we use the following properties for deriving our results [35], [36]:

P.1 $\text{tr}(A) = \text{tr}(A^\mathsf{T})$ and $\text{tr}(ABC) = \text{tr}(CAB)$,

P.2 $\|A\|_F^2 = \text{tr}(A^\mathsf{T}A) = \text{vec}^\mathsf{T}(A)\text{vec}(A)$,

P.3 $\text{vec}(AB) = (I \otimes A)\text{vec}(B) = (B^\mathsf{T} \otimes I)\text{vec}(A)$,

P.4 $\text{vec}(ABC) = (C^\mathsf{T} \otimes A)\text{vec}(B)$,

P.5 $(A \otimes B)^\mathsf{T} = A^\mathsf{T} \otimes B^\mathsf{T}$ and $(A \otimes B)^\mathsf{H} = A^\mathsf{H} \otimes B^\mathsf{H}$,

P.6 $1_n^\mathsf{T}(A \circ B)1_n = \text{tr}(A^\mathsf{T}B)$,

P.7 $A \circ B = B \circ A$ and $A \circ (B \circ C) = (A \circ B) \circ C$,

P.8 $\text{vec}(A \circ B) = \text{vec}(A) \circ \text{vec}(B), (A \circ B)^\mathsf{T} = A^\mathsf{T} \circ B^\mathsf{T}$,

P.9 $\frac{d}{dX}\text{tr}(AX) = A^\mathsf{T}$, $\frac{d}{dX}\text{tr}(X^\mathsf{T}X) = 2X$, $\frac{d}{dx}(Ax) = A$,

P.10 $d(X^{-1}) = -X^{-1}dXX^{-1}$,

P.11 $d(A \otimes B) = dA \otimes B + A \otimes dB$,

P.12 Let $D_x f$ and $D_x^2 f$ be the gradient and Hessian of $f(x): \mathbb{R}^n \to \mathbb{R}$. Then, $df = (D_x f)^\mathsf{T} dx$ and $d^2 f = (dx)^\mathsf{T}(D_x^2 f)dx$,

P.13 $T_{m,n}^{-1} = T_{m,n}^\mathsf{T} = T_{n,m}$.

P.14 Projection of a vector $y \in \mathbb{R}^n$ on the null space of $A \in \mathbb{R}^{m \times n}$ is given by $y_p = [I_n - A^+ A]y$.

B. Sparse MGEAP

The sparse MGEAP involves finding a real feedback matrix $F \in \mathbb{R}^{m \times n}$ with minimum norm that assigns the closed loop eigenvalues of (1a)-(1b) at some desired locations given by set $\mathcal{S} = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$, and satisfies a given sparsity constraints. Let $\bar{\mathsf{F}} \in \{0,1\}^{m \times n}$ denote a binary matrix that specifies the sparsity structure of the feedback matrix F. If $\bar{\mathsf{F}}_{ij} = 0$ (respectively $\bar{\mathsf{F}}_{ij} = 1$), then the jth state is unavailable (respectively available) for calculating the ith input. Thus,

$$F_{ij} = \begin{cases} 0 & \text{if } \bar{\mathsf{F}}_{ij} = 0, \text{ and} \\ \star & \text{if } \bar{\mathsf{F}}_{ij} = 1, \end{cases}$$

where $\star$ denotes a real number. Let $\bar{\mathsf{F}}^c \triangleq 1_{m \times n} - \bar{\mathsf{F}}$ denote the complementary sparsity structure matrix. Further, (with a slight abuse of notation, c.f. (1a)) let $X \triangleq [x_1, x_2, \cdots, x_n] \in \mathbb{C}^{n \times n}, x_i \neq 0_n$ denote the non-singular eigenvector matrix of the closed loop matrix $A_c(F) = A + BF$.

The MGEAP can be mathematically stated as the following optimization problem

$$\min_{F,X} \quad \frac{1}{2}\|F\|_F^2 \tag{2}$$

$$\text{s.t.} \quad (A + BF)X = X\Lambda, \tag{2a}$$

$$x_i^\mathsf{H} x_i = 1 \quad \forall i = 1, 2, \cdots, n \quad \text{and} \tag{2b}$$

$$\bar{\mathsf{F}}^c \circ F = 0_{m \times n}, \tag{2c}$$

where $\Lambda = \text{diag}([\lambda_1, \lambda_2, \cdots, \lambda_n]^\mathsf{T})$ is the diagonal matrix constructed from the entries of the desired eigenvalue set $\mathcal{S}$. Equations (2a) and (2c) represent the eigenvalue assignment and sparsity constraints, respectively. Further, the normalization constraint in (2b) is included to uniquely determine the eigenvector matrix X.

The constraint (2a) is not convex in (F, X) and, therefore, the optimization problem (2) is non-convex. Consequently, multiple local minima may exist. This is a common feature in various minimum distance and eigenvalue assignment problems [37], including the non-sparse MGEAP.

Remark 1. (*Choice of norm*) *The Frobenius norm explicitly measures the element-wise gains of a matrix, which is informative in sparsity constrained problems arising, for instance, in network control problems. Further, it is also convenient for the analysis, as it allows us to analytically compute the derivatives of the cost function.* □

Definition 1. (*Fixed modes [24], [31]*) *The fixed modes of (A, B) with respect to the sparsity constraints $\bar{\mathsf{F}}$ are those*

eigenvalues of A which cannot be changed using LTI static (and also dynamic) state feedback, and are denoted by

$$\Gamma_f(A, B, \bar{\mathsf{F}}) \triangleq \bigcap_{F \,:\, F \circ \bar{\mathsf{F}}^c = 0} \Gamma(A + BF).$$

□

We make the following assumptions regarding the fixed modes and feasibility of the optimization problem (2).

A1: The fixed modes of the triplet $(A, B, \bar{\mathsf{F}})$ are included in the desired eigenvalue set $\mathcal{S}$, i.e., $\Gamma_f(A, B, \bar{\mathsf{F}}) \subseteq \mathcal{S}$.

A2: There exists at least one feedback matrix F that satisfies constraints (2a)-(2c) for the given $\mathcal{S}$.

Assumption **A1** is clearly necessary for the feasibility of the optimization problem (2). Assumption **A2** is restrictive because, in general, it is possible that a static feedback matrix with a given sparsity pattern cannot assign the closed loop eigenvalues to arbitrary locations (i.e. for an arbitrary set $\mathcal{S}$ satisfying Assumption **A1**).[1] In such cases, only a few ($< n$) eigenvalues can be assigned independently and other remaining eigenvalues are a function of them. To the best of our knowledge, there are no studies on characterizing conditions for the existence of a static feedback matrix for an arbitrary sparsity pattern $\bar{\mathsf{F}}$ and eigenvalue set $\mathcal{S}$ [38] (although such characterization is available for *dynamic* feedback laws with arbitrary sparsity pattern [28], [31], [32], and static output feedback for *decentralized* sparsity pattern [39]). Thus, for the purpose of this paper, we focus on finding the optimal feedback matrix *assuming* that at least one such feedback matrix exists.

III. Solution to the sparse MGEAP

In this section we present the solution to the optimization problem (2). We use the theory of Lagrangian multipliers for equality constrained minimization problems to derive the optimality conditions. We begin with a remark on the formalism involving complex variables.

Remark 2. (*Complex variables*) *If the desired eigenvalue set $\mathcal{S}$ contains complex entries, the corresponding eigenvectors in X will also be complex. Thus, the eigenvalue-eigenvector constraint in* (2a) *is a complex-valued constraint, and it induces the following conjugate constraint:*

$$(A + BF)X^* = X^*\Lambda^*. \tag{3}$$

In the optimization problem, we use the formalism wherein a complex number and its conjugate are treated as independent variables [40], [41] and, thus, we treat X and X^* as independent variables. □

Remark 3. (*Conjugate eigenvectors*) *We use the convention that the right eigenvectors (x_i, x_j) corresponding to two conjugate eigenvalues (λ_i, λ_j) are also conjugate. Thus, if $\lambda_i = \lambda_j^*$, then $x_i = x_j^*$. Similar convention is also used for the left eigenvectors.* □

---

[1]Although a sparse *dynamic* feedback law can assign the eigenvalues to arbitrary locations under Assumption **A1** [32].

In the theory of equality constrained optimization, the first-order optimality conditions are meaningful only when the optimal point satisfies the following regularity condition: the Jacobian of the constraints, defined by J_b, is full rank. This regularity condition is mild and usually satisfied for most classes of problems [42]. Before presenting the main result, we derive the Jacobian and state the regularity condition for the optimization problem (2).

Computation of J_b requires vectorization of the matrix constraints (2a)-(2c). For this purpose, let $x \triangleq \text{vec}(X) \in \mathbb{C}^{n^2}$, $f \triangleq \text{vec}(F) \in \mathbb{R}^{mn}$. Recalling Remark 2, let $z \triangleq [x^\mathsf{T}, x^\mathsf{H}, f^\mathsf{T}]^\mathsf{T}$ be the vector containing all the independent variables of the optimization problem. Further, let n_s denote the total number of feedback sparsity constraints (i.e. number of 1's in $\bar{\mathsf{F}}^c$):

$$n_\text{s} = |\{(i,j) \ : \ \bar{\mathsf{F}}^c = [\bar{\mathsf{f}}^c_{ij}], \ \bar{\mathsf{f}}^c_{ij} = 1\}|.$$

Finally, note that the constraint (2c) can be equivalently written as (see the proof of Lemma III.1 for a formal definition of Q)

$$Qf = 0_{n_s}, \tag{4}$$

for some binary matrix $Q \in \{0,1\}^{n_s \times mn}$.

Lemma III.1. *(Jacobian of the constraints) The Jacobian of the equality constraints (2a)-(2c) is given by*

$$J_b(z) = \begin{bmatrix} A_c(F) \oplus (-\Lambda^\mathsf{T}) & 0_{n^2 \times n^2} & X^\mathsf{T} \otimes B \\ 0_{n^2 \times n^2} & A_c(F) \oplus (-\Lambda^\mathsf{H}) & X^\mathsf{H} \otimes B \\ \bar{x} & \bar{x}^* & 0_{n \times mn} \\ 0_{n_s \times n^2} & 0_{n_s \times n^2} & Q \end{bmatrix}, \tag{5}$$

where $\bar{x} \triangleq (1_n \otimes x^\mathsf{H}) \circ (I_n \otimes 1_n^\mathsf{T})$.

Proof. We construct the Jacobian J_b by rewriting the constraints (2a)-(2c), (3) in vectorized form and taking their derivatives with respect to z. Constraint (2a) can be vectorized in the following two different ways (using P.3 and P.4):

$$[(A + BF) \oplus (-\Lambda^\mathsf{T})]x = 0_{n^2}, \tag{6a}$$
$$[A \oplus -(\Lambda^\mathsf{T})]x + (X^\mathsf{T} \otimes B)f = 0_{n^2}. \tag{6b}$$

Differentiating (6a) w.r.t. x and (6b) w.r.t f yields the first (block) row of J_b. Similar vectorizations of the conjugate eigenvalue constraint (3) and differentiation w.r.t. z yields the second (block) row of J_b. Differentiating constraint (2b) for $i = 1, 2, \cdots, n$ w.r.t. x and x^* yields

$$\frac{d}{dx} \begin{bmatrix} x_1^\mathsf{H} x_1 - 1 \\ x_2^\mathsf{H} x_2 - 1 \\ \vdots \\ x_n^\mathsf{H} x_n - 1 \end{bmatrix} = \begin{bmatrix} x_1^\mathsf{H} & 0_n^\mathsf{T} & \cdots & 0_n^\mathsf{T} \\ 0_n^\mathsf{T} & x_2^\mathsf{H} & \cdots & 0_n^\mathsf{T} \\ \vdots & \vdots & \ddots & \vdots \\ 0_n^\mathsf{T} & 0_n^\mathsf{T} & \cdots & x_n^\mathsf{H} \end{bmatrix} = \bar{x},$$

and $\bar{x}^*$, respectively. Finally, (2c) consists of n_s non-trivial sparsity constraints, which can be written as (4) where $Q = [e_{q_1} \ e_{q_2} \ \ldots \ e_{q_{n_s}}]^\mathsf{T}$ and $\{q_1, \ldots, q_{n_s}\} = \text{supp}(\text{vec}(\bar{\mathsf{F}}^c))$ is the set of indices indicating the ones in $\text{vec}(\bar{\mathsf{F}}^c)$. Differentiating (4) w.r.t. z yields the fourth (block) row of J_b, thus completing the proof. $\square$

Next, we provide the optimality conditions for the optimization problem (2).

Theorem III.2. *(Optimality conditions) Let $(\hat{X}, \hat{F})$ (equivalently $\hat{z} = [\hat{x}^\mathsf{T}, \hat{x}^\mathsf{H}, \hat{f}^\mathsf{T}]^\mathsf{T}$) satisfy the constraints (2a)-(2c). Then, $(\hat{X}, \hat{F})$ is a local minimum of the optimization problem (2) if and only if*

$$\hat{F} = -\bar{\mathsf{F}} \circ (B^\mathsf{T} \hat{L} \hat{X}^\mathsf{T}), \tag{7a}$$

where $\hat{X}$ and $\hat{L}$ are the right and left eigenvector matrices of $A_c(\hat{F})$, respectively, and satisfy

$$(A + B\hat{F})\hat{X} = \hat{X}\Lambda \tag{7b}$$
$$(A + B\hat{F})^\mathsf{T} \hat{L} = \hat{L}\Lambda^\mathsf{T} \quad and, \tag{7c}$$
$$J_b(\hat{z}) \ is \ full \ rank, \tag{7d}$$
$$P(\hat{z})\hat{D}P(\hat{z}) > 0, \tag{7e}$$

where $\hat{D}$ is the Hessian defined as

$$\hat{D} \triangleq \begin{bmatrix} 0_{n^2 \times n^2} & 0_{n^2 \times n^2} & \bar{L}^\mathsf{H} \\ 0_{n^2 \times n^2} & 0_{n^2 \times n^2} & \bar{L}^\mathsf{T} \\ \bar{L} & \bar{L}^* & 2I_{mn} \end{bmatrix}, \tag{8}$$

where $\bar{L} \triangleq T_{n,m}(B^\mathsf{T} \hat{L} \otimes I_n)$, and $P(z)$ is the projection matrix of $J_b(z)$ defined as

$$P(z) = I_{2n^2 + mn} - J_b^+(z)J_b(z). \tag{9}$$

Proof. We prove the result using the Lagrange theorem for equality constrained minimization. Let $L \in \mathbb{C}^{n \times n}, L^*, M \in \mathbb{R}^{m \times n}$ and $h \in \mathbb{R}^n$ be the Lagrange multipliers associated with constraints (2a),(3),(2b) and (2c), respectively. Further, let $H = \text{diag}(h)$. The Lagrange function for the optimization problem (2) is given by

$$\mathcal{L} \overset{P.2}{=} \frac{1}{2} \text{tr}(F^\mathsf{T} F) + \frac{1}{2} 1_n^\mathsf{T} [L \circ (A_c(F)X - X\Lambda)]1_n$$
$$+ \frac{1}{2} 1_n^\mathsf{T} [L^* \circ (A_c(F)X^* - X^*\Lambda^*)]1_n + \text{tr}((X^\mathsf{H}X - I_n)H)$$
$$+ 1_m^\mathsf{T} [M \circ (\bar{\mathsf{F}}^c \circ F)]1_n$$
$$\overset{P.6,P.7}{=} \frac{1}{2} \text{tr}(F^\mathsf{T} F) + \frac{1}{2} \text{tr}[L^\mathsf{T} (A_c(F)X - X\Lambda)]$$
$$+ \frac{1}{2} \text{tr}[L^\mathsf{H} (A_c(F)X^* - X^*\Lambda^*)] + \text{tr}((X^\mathsf{H}X - I_n)H)$$
$$+ \text{tr}[(M \circ \bar{\mathsf{F}}^c)^\mathsf{T} F].$$

We next derive the first-order necessary condition for a stationary point. Differentiating l w.r.t. X and setting to 0, we get

$$\frac{d}{dX}\mathcal{L} \overset{P.9}{=} \frac{1}{2}[A_c^\mathsf{T}(F)L - L\Lambda^\mathsf{T}] + X^* H = 0_{n \times n}. \tag{10}$$

Pre-multiplying (10) by X^T, we get

$$\frac{1}{2}[X^\mathsf{T} A_c^\mathsf{T}(F)L - X^\mathsf{T} L\Lambda^\mathsf{T}] + X^\mathsf{T} X^* H = 0_{n \times n},$$
$$\overset{(2a)}{\Rightarrow} \frac{1}{2}[\Lambda^\mathsf{T} X^\mathsf{T} L - X^\mathsf{T} L\Lambda^\mathsf{T}] + X^\mathsf{T} X^* H = 0_{n \times n}. \tag{11}$$

Since Λ is diagonal and $x_i^\mathsf{H} x_i = 1$, the vector containing the diagonal elements of the left side of (11) is h. Therefore, $h = 0_n$ (and $H = 0_{n \times n}$) and from (10), we get (7c). Equation (7b) is a restatement of (2a) for the optimal $(\hat{F}, \hat{X})$.

Differentiating $\mathcal{L}$ w.r.t. F, we get

$$\frac{d}{dF}\mathcal{L} \stackrel{P.9}{=} F + \mathrm{Re}(B^\mathsf{T} L X^\mathsf{T}) + M \circ \bar{\mathsf{F}}^c = 0_{m \times n}. \qquad (12)$$

Taking the Hadamard product of (12) with $\bar{\mathsf{F}}^c$ and using (2c), we get (since $\bar{\mathsf{F}}^c \circ \bar{\mathsf{F}}^c = \bar{\mathsf{F}}^c$)

$$\bar{\mathsf{F}}^c \circ \mathrm{Re}(B^\mathsf{T} L X^\mathsf{T}) + M \circ \bar{\mathsf{F}}^c = 0_{m \times n} \qquad (13)$$

Replacing $M \circ \bar{\mathsf{F}}^c$ from (13) into (12), we get

$$F = -\bar{\mathsf{F}} \circ \mathrm{Re}(B^\mathsf{T} L X^\mathsf{T}).$$

Since the eigenvectors in X (and $L \triangleq [l_1, l_2, \cdots .l_n]$) are closed under complex conjugation (c.f. Remark 3), the product $L X^\mathsf{T} = \sum_{i=1}^{n} l_i x_i^\mathsf{T}$ is real and (7a) follows. Equation (7d) is the necessary regularity condition and follows from Lemma III.1

Next, we derive the second-order sufficient condition for a local minimum by calculating the Hessian of $\mathcal{L}$. Taking the differential of $\mathcal{L}$ twice, we get

$$d^2\mathcal{L} = tr((dF)^\mathsf{T} dF) + tr(L^\mathsf{T} B dF dX) + tr(L^\mathsf{H} B dF dX^*)$$
$$+ 2tr((dX)^\mathsf{H} dX H)$$
$$\stackrel{(P.2, H=0)}{=} df^\mathsf{T} df + \mathrm{vec}^\mathsf{T}(dF^\mathsf{T} B^\mathsf{T} L) dx + \mathrm{vec}^\mathsf{T}(dF^\mathsf{T} B^\mathsf{T} L^*) dx^*$$
$$\stackrel{(P.3, P.5)}{=} df^\mathsf{T} df + df^\mathsf{T} \bar{L} dx + df^\mathsf{T} \bar{L}^* dx^*$$
$$= \frac{1}{2}[dx^\mathsf{H}, dx^\mathsf{T}, df^\mathsf{T}] D \begin{bmatrix} dx \\ dx^* \\ df \end{bmatrix},$$

where D is the Hessian (c.f. P.12) defined in (8). The sufficient second-order optimality condition for the optimization problem requires the Hessian to be positive definite in the kernel of the Jacobian at the optimal point [43]. That is, $y^\mathsf{T} D y > 0, \ \forall y : J_b(z) y = 0$. This condition is equivalent to $P(z) D P(z) > 0$, since $J_b(z) y = 0$ if and only if $y = P(z) s$ for a $s \in \mathbb{R}^{2n^2 + mn}$ [42]. Since the projection matrix $P(z)$ is symmetric, (7e) follows, and this concludes the proof. $\square$

Remark 4. *(Eigenvector normalization constraints) From the proof of Theorem III.2, we observe that the Lagrange multipliers associated with the eigenvector normalization constraints in (2b) are zero ($h = 0_n$). This implies that these constraints are redundant. Further, observe from (7a) that any pair of left and right eigenvector matrices $\{\hat{L} D^{-1}, \hat{X} D\}$ where D is any diagonal matrix with non-zero (and possibly complex) diagonal entries, will result in the same feedback matrix $\hat{F}$. Thus, we can always choose a suitable D to normalize $\hat{X}$. Therefore, we ignore the eigenvector normalization constraints in the remainder of the paper.* $\square$

Observe that the Hadamard product in (7a) guarantees that the feedback matrix satisfies the sparsity constraints given in (2c). However, note that the optimal sparse feedback $\hat{F}$ cannot be obtained by sparsification of the optimal non-sparse feedback. The feedback optimality condition (7a) is an implicit condition in terms of the closed loop right and left eigenvector matrices. Next, we provide an explicit optimality condition in terms of $\{\hat{L}, \hat{X}\}$.

Corollary III.3. *(Stationary point equation) $\hat{Z} \triangleq [\hat{X}^\mathsf{T}, \hat{L}^\mathsf{T}]^\mathsf{T}$ is a stationary point of the optimization problem (2) if and only if*

$$\bar{A}\hat{Z} - \hat{Z}\Lambda = \bar{B}_1[\mathsf{F} \circ (\bar{B}_1^\mathsf{T} \bar{I} \hat{Z} \hat{Z}^\mathsf{T} \bar{B}_2)]\bar{B}_2^\mathsf{T} \hat{Z}, \qquad (14)$$

where,

$$\bar{A} \triangleq \begin{bmatrix} A & 0_{n \times n} \\ 0_{n \times n} & A^\mathsf{T} \end{bmatrix}, \ \bar{B}_1 \triangleq \begin{bmatrix} B & 0_{n \times n} \\ 0_{n \times m} & I_n \end{bmatrix},$$
$$\bar{B}_2 \triangleq \begin{bmatrix} I_n & 0_{n \times m} \\ 0_{n \times n} & B \end{bmatrix}, \ \mathsf{F} \triangleq \begin{bmatrix} \bar{\mathsf{F}} & 0_{m \times m} \\ 0_{n \times n} & \bar{\mathsf{F}}^\mathsf{T} \end{bmatrix}, and$$
$$\bar{I} \triangleq \begin{bmatrix} 0_{n \times n} & I_n \\ I_n & 0_{n \times n} \end{bmatrix}.$$

Proof. Combining (7b) and (7c) and using $\Lambda^\mathsf{T} = \Lambda$, we get

$$\begin{bmatrix} A\hat{X} - \hat{X}\Lambda \\ A^\mathsf{T}\hat{L} - \hat{L}\Lambda \end{bmatrix} = - \begin{bmatrix} B\hat{F}\hat{X} \\ \hat{F}^\mathsf{T} B^\mathsf{T} \hat{L} \end{bmatrix}$$
$$\Rightarrow \bar{A}\hat{Z} - \hat{Z}\Lambda = -\bar{B}_1 \begin{bmatrix} \hat{F} & 0 \\ 0 & \hat{F}^\mathsf{T} \end{bmatrix} \bar{B}_2^\mathsf{T} \hat{Z}$$
$$= \bar{B}_1 \begin{bmatrix} \bar{\mathsf{F}} \circ (B^\mathsf{T} \hat{L} \hat{X}^\mathsf{T}) & 0 \\ 0 & \bar{\mathsf{F}}^\mathsf{T} \circ (\hat{X} \hat{L}^\mathsf{T} B) \end{bmatrix} \bar{B}_2^\mathsf{T} \hat{Z}$$
$$= \bar{B}_1 \left(\mathsf{F} \circ \left\{ \bar{B}_1^\mathsf{T} \begin{bmatrix} \hat{L} \hat{X}^\mathsf{T} & 0 \\ 0 & \hat{X} \hat{L}^\mathsf{T} \end{bmatrix} \bar{B}_2 \right\} \right) \bar{B}_2^\mathsf{T} \hat{Z}$$
$$= \bar{B}_1 \left(\mathsf{F} \circ \left\{ \bar{B}_1^\mathsf{T} \bar{I} (\hat{Z} \hat{Z}^\mathsf{T} \circ \bar{I}) \bar{B}_2 \right\} \right) \bar{B}_2^\mathsf{T} \hat{Z}$$
$$= \bar{B}_1 \{ \mathsf{F} \circ (\bar{B}_1^\mathsf{T} \bar{I} \hat{Z} \hat{Z}^\mathsf{T} \bar{B}_2) \} \bar{B}_2^\mathsf{T} \hat{Z},$$

where the above equations follow from direct manipulations using definition of Hadamard product. $\square$

Remark 5. *(Partial spectrum assignment) The results of Theorem III.2 and Corollary III.3 are also valid when only p ($p < n$) eigenvalues need to be assigned and the remaining eigenvalues are unspecified. In this case, $\Lambda \in \mathbb{C}^{p \times p}$, $\hat{X} \in \mathbb{C}^{n \times p}$ and $\hat{L} \in \mathbb{C}^{n \times p}$. Partial assignment is useful when the sparsity constraints allow only a limited number of eigenvalues to be assigned independently and the remaining eigenvalues are functionally related to them (c.f. discussion below Assumption A2). Nevertheless, in this paper we focus on assigning all the closed loop eigenvalues.* $\square$

Remark 6. *(General eigenstructure assignment) Although the optimization problem (2) is formulated by considering Λ to be diagonal, the result in Theorem III.2 is valid for any general Λ satisfying $\Gamma(\Lambda) = \mathcal{S}$. For instance, we can choose Λ in a Jordan canonical form. However, note that for a general Λ, X will cease to be an eigenvector matrix. Thus, in such cases, we need to omit the normalization constraints (2b) in the optimization problem.* $\square$

A solution of the optimization problem (2) can be obtained by numerically/iteratively solving the matrix equation (14), which resembles a Sylvester type equation with a non-linear right side, and using (7a) to compute the feedback matrix. The regularity and local minimum of the solution can be verified using (7d) and (7e), respectively. Since the optimization problem is not convex, only local minima can be obtained

via this procedure. To improve upon the local solutions, the above procedure can be repeated for several different initial conditions (for solving (14)). However, convergence to the global minimum is not guaranteed.

In this paper, instead of solving (14) directly, we use a different approach based on the gradient descent procedure to obtain a local solution. Details of this approach and corresponding algorithms are presented in Section IV.

A. Results for non-sparse MGEAP

In this subsection, we present some results specific to the case when the optimization problem (2) does not have any sparsity constraints (i.e. $\bar{F} = 1_{m \times n}$). Although the non-sparse MGEAP has been studied previously, these results are novel and further illustrate the properties of an optimal solution.

We begin by presenting a geometric interpretation of the optimality conditions provided in Theorem III.2 under the assumption that $B = I_n$, i.e. all the entries of A can be perturbed independently. In this case, the optimization problem (2) can be written as (while ignoring the eigenvector normalization constraints; c.f. Remark 4):

$$\min_X \ \frac{1}{2} \|A - X\Lambda X^{-1}\|_F^2. \tag{15}$$

Since A and $R(X) \triangleq X\Lambda X^{-1}$ are elements (or vectors) of the matrix inner product space with Frobenius norm, a solution of the optimization problem (15) is given by the projection of A on the manifold $\mathcal{M} \triangleq \{R(X) : X \text{ is non-singular}\}$. This projection can be obtained by solving the normal equation, which states that the optimal error vector $\hat{F} = A - \hat{X}\Lambda\hat{X}^{-1}$ should be orthogonal to the tangent plane of the manifold $\mathcal{M}$ at the optimal point $\hat{X}$. The next result shows that the optimality conditions derived in Theorem III.2 are in fact the normal equations for the optimization problem (15).

Lemma III.4. *(Geometric interpretation)* Let $\bar{F} = 1_{m \times n}$ and $B = I_n$. Then, Equations (7a)-(7c) are equivalent to the following normal equation:

$$< \hat{F}, \mathcal{T}_{\mathcal{M}}(\hat{X}) >_F = 0, \tag{16}$$

where $\mathcal{T}_{\mathcal{M}}(X)$ denotes the tangent space of $\mathcal{M}$ at X.

Proof. We begin by characterizing the tangent space $\mathcal{T}_{\mathcal{M}}(X)$, which is given by the first order approximation of the matrix function $R(X)$:

$$R(X + dX) = (X + dX)\Lambda(X + dX)^{-1}$$
$$\stackrel{(P.10)}{=} R(X) + dX\Lambda X^{-1} - X\Lambda X^{-1}dX X^{-1}$$
$$+ \text{ higher order terms.}$$

Thus, the tangent space is given by

$$\mathcal{T}_{\mathcal{M}}(X) = \{Y\Lambda X^{-1} - X\Lambda X^{-1}Y X^{-1} : Y \in \mathbb{C}^{n \times n}\}$$

Necessity: Using $\hat{F}$ given by (7a), we get

$$< \hat{F}, \mathcal{T}_{\mathcal{M}}(\hat{X}) > = \text{tr}(\hat{F}^{\mathsf{T}}(Y\Lambda\hat{X}^{-1} - \hat{X}\Lambda\hat{X}^{-1}Y\hat{X}^{-1}))$$
$$= -\text{tr}(\hat{X}\hat{L}^{\mathsf{T}}Y\Lambda\hat{X}^{-1}) + \text{tr}(\hat{X}\hat{L}^{\mathsf{T}}\hat{X}\Lambda\hat{X}^{-1}Y\hat{X}^{-1})$$
$$\stackrel{(P.1)}{=} -\text{tr}(\hat{L}^{\mathsf{T}}Y\Lambda) + \text{tr}(\hat{L}^{\mathsf{T}}\hat{X}\Lambda\hat{X}^{-1}Y)$$
$$\stackrel{(a)}{=} -\text{tr}(\hat{L}^{\mathsf{T}}Y\Lambda) + \text{tr}(\Lambda\hat{L}^{\mathsf{T}}\hat{X}\hat{X}^{-1}Y) \stackrel{(P.1)}{=} 0,$$

where (a) follows from the fact that Λ and $\hat{L}^{\mathsf{T}}\hat{X}$ commute. *Sufficiency:* From (16), we get

$$\text{tr}(\hat{F}^{\mathsf{T}}(Y\Lambda\hat{X}^{-1} - \hat{X}\Lambda\hat{X}^{-1}Y\hat{X}^{-1})) = 0$$
$$\stackrel{(P.1)}{\Rightarrow} \text{tr}[(\Lambda\hat{X}^{-1}\hat{F}^{\mathsf{T}} - \hat{X}^{-1}\hat{F}^{\mathsf{T}}\hat{X}\Lambda\hat{X}^{-1})Y] = 0.$$

Since the above equation is true for all $Y \in \mathbb{C}^{n \times n}$, we get

$$\Lambda\hat{X}^{-1}\hat{F}^{\mathsf{T}} - \hat{X}^{-1}\hat{F}^{\mathsf{T}}\hat{X}\Lambda\hat{X}^{-1} = 0_{n \times n}$$
$$\Rightarrow \hat{X}\Lambda\hat{X}^{-1}\hat{F}^{\mathsf{T}} = \hat{F}^{\mathsf{T}}\hat{X}\Lambda\hat{X}^{-1}$$
$$\Rightarrow A_c(\hat{F})\hat{F}^{\mathsf{T}} = \hat{F}^{\mathsf{T}}A_c(\hat{F}).$$

Thus, $A_c(\hat{F})$ and $\hat{F}^{\mathsf{T}}$ commute and therefore, have common left and right eigenspaces [44], i.e. $\hat{F}^{\mathsf{T}} = -\hat{X}G\hat{X}^{-1} = -\hat{X}\hat{L}^{\mathsf{T}}$, where G is a diagonal matrix. This completes the proof. □

Next, we show the equivalence of the non-sparse MGEAP for two orthogonally similar systems.

Lemma III.5. *(Invariance under orthogonal transformation)* Let $\bar{F} = 1_{m \times n}$ and (A_1, B_1), (A_2, B_2) be two orthogonally similar systems such that $A_2 = PA_1P^{-1}$ and $B_2 = PB_1$, with P being an orthogonal matrix. Let an optimal solutions of (2) for the two systems be denoted by $(\hat{X}_1, \hat{L}_1, \hat{F}_1)$ and $(\hat{X}_2, \hat{L}_2, \hat{F}_2)$, respectively. Then

$$\hat{X}_2 = P\hat{X}_1, \ \hat{L}_2 = P\hat{L}_1, \ \hat{F}_2 = \hat{F}_1P^{\mathsf{T}}, \ and$$
$$\|\hat{F}_1\|_F = \|\hat{F}_2\|_F. \tag{17}$$

Proof. From (7b), we have

$$(A_2 + B_2\hat{F}_2)\hat{X}_2 = \hat{X}_2\Lambda$$
$$\Rightarrow (PA_1P^{-1} + PB_1\hat{F}_1P^{\mathsf{T}})P\hat{X}_1 = P\hat{X}_1\Lambda$$
$$\Rightarrow (A_1 + B_1\hat{F}_1)\hat{X}_1 = \hat{X}_1\Lambda.$$

Similar relation can be shown between $\hat{L}_1$ and $\hat{L}_2$ using (7c). Next, from (7a), we have

$$\hat{F}_2 = -B_2^{\mathsf{T}}\hat{L}_2\hat{X}_2^{\mathsf{T}} = -B_1^{\mathsf{T}}\hat{L}_1\hat{X}_1^{\mathsf{T}}P^{\mathsf{T}} = \hat{F}_1P^{\mathsf{T}}.$$

Finally, $\|\hat{F}_1\|_F^2 = \text{tr}(\hat{F}_1^{\mathsf{T}}\hat{F}_1) \stackrel{(P.1)}{=} \text{tr}(\hat{F}_2^{\mathsf{T}}\hat{F}_2) = \|\hat{F}_2\|_F^2$, thereby completing the proof. □

Recall from Remark 5 that Theorem III.2 is also valid for MGEAP with partial spectrum assignment. Next, we consider the case when only one real eigenvalue needs to assigned for the MGEAP while the remaining eigenvalues are unspecified. In this special case, we can explicitly characterize the global minimum of (2) as shown in the next result.

Corollary III.6. *(One real eigenvalue assignment)* Let $\bar{F} = 1_{m \times n}$, $\Lambda \in \mathbb{R}$, and $B = I_n$. Then, the global minima of the optimization problem (2) is given by $\hat{F}_{gl} = -\sigma_{min}(A - \Lambda I_n)uv^{\mathsf{T}}$, where u and v are unit norm left and right singular vectors, respectively, corresponding to $\sigma_{min}(A - \Lambda I_n)$. Further, $\|\hat{F}_{gl}\|_F = \sigma_{min}(A - \Lambda I_n)$.

Proof. Since $\Lambda \in \mathbb{R}$, $\hat{X} \in \mathbb{R}^n \triangleq \hat{x}$ and $\hat{L} \in \mathbb{R}^n \triangleq \hat{l}$. Let $\hat{l} = \beta\hat{\bar{l}}$ where $\beta \triangleq \|\hat{l}\|_2 > 0$. Substituting $\hat{F} = -\hat{l}\hat{x}^\mathsf{T}$ from (7a) into (7b)-(7c), we get

$$(A - \hat{l}\hat{x}^\mathsf{T})\hat{x} = \hat{x}\Lambda \overset{(2b)}{\Rightarrow} (A - \Lambda I_n)\hat{x} = \beta\hat{\bar{l}} \quad \text{and,}$$

$$(A^\mathsf{T} - \hat{x}\hat{l}^\mathsf{T})\hat{l} = \hat{l}\Lambda \Rightarrow (A - \Lambda I_n)^\mathsf{T}\hat{\bar{l}} = \beta\hat{x}.$$

The above two equations imply that the unit norm vectors $\hat{x}$ and $\hat{\bar{l}}$ are left and right singular vectors of $A - \Lambda I_n$ associated with the singular value β. Since $\|\hat{F}\|_F^2 = \operatorname{tr}(\hat{F}^\mathsf{T}\hat{F}) = \operatorname{tr}(\hat{x}\hat{l}^\mathsf{T}\hat{l}\hat{x}^\mathsf{T}) = \beta^2$, we pick β as the minimum singular value of $A - \Lambda I_n$, and the proof is complete. $\qquad\square$

We conclude this subsection by presenting a brief comparison of the non-sparse MGEAP solution with deflation techniques for eigenvalue assignment. For $B = I_n$, an alternative method to solve the non-sparse EAP is via the Wielandt deflation technique [45]. Wielandt deflation achieves pole assignment by modifying the matrix A in n steps $A \to A_1 \to A_2 \to \cdots \to A_n$. Step i shifts one eigenvalue of A_{i-1} to a desired location λ_i, while keeping the remaining eigenvalues of A_{i-1} fixed. This is achieved by using the feedback $F_{df}^i = -(\mu_i - \lambda_i)v_i z_i^\mathsf{T}$, where μ_i and v_i are any eigenvalue and right eigenvector pair of A_{i-1}, and z is any vector such that $z_i^\mathsf{T}v_i = 1$. Thus, the overall feedback that solves the EAP is given as $F_{df} = \sum_{i=1}^n F_{df}^i$.

It is interesting to compare the optimal feedback expression in (7a), $\hat{F} = -\sum_{i=1}^n \hat{l}_i\hat{x}_i^\mathsf{T}$, with the deflation feedback. Both feedbacks are sum of n rank-1 perturbations. However, the Wielandt deflation has an inherent special structure and a restrictive property that, in each step, all except one eigenvalue remain unchanged. Furthermore, each rank-1 term in $\hat{F}$ and F_{df} involves the right/left eigenvectors of the closed and open loop matrix, respectively. Due to these reasons, it clearly follows that $\|\hat{F}\|_F \leq \|F_{df}\|_F$.

IV. SOLUTION ALGORITHMS

In this section, we present an iterative algorithm to obtain a solution to the sparse MGEAP in (2). To develop the algorithm, we first present two algorithms for computing non-sparse and approximately-sparse solutions to the MGEAP, respectively. Next, we present two heuristic algorithms to obtain a sparse solution of the EAP (i.e. any sparse solution, which is not necessarily minimum-gain). Finally, we use these algorithms to develop the algorithm for sparse MGEAP. Note that although our focus is to develop the sparse MGEAP algorithm, the other algorithms presented in this section are novel in themselves to the best of our knowledge.

We make the following assumptions:
A3: The triplet $(A, B, \bar{\mathsf{F}})$ has no fixed modes, i.e., $\Gamma_f(A, B, \bar{\mathsf{F}}) = \emptyset$.
A4: The open and closed loop eigenvalue sets are disjoint, i.e., $\Gamma(A) \cap \Gamma(\Lambda) = \emptyset$.

Assumption **A4** is not restrictive since if there are any common eigenvalues in A and Λ, we can use a preliminary sparse feedback F_p to shift the eigenvalues of A to some other locations such that $\Gamma(A + BF_p) \cap \Gamma(\Lambda) = \emptyset$. Since the system has no fixed modes (c.f. Assumption **A3**), such a

F_p always exists. Then, we can solve the modified MGEAP[2] with parameters $(A + BF_p, B, \Lambda, \bar{\mathsf{F}})$. If F is the sparse solution of this modified problem, then the solution of the original problem is $F_p + F$.

To avoid complex domain calculations in the algorithms, we use the real counterpart of (2a). For two complex conjugate eigenvalues (λ_i, λ_i^*) and corresponding eigenvectors (x_i, x_i^*), the following complex equation

$$(A + BF)\begin{bmatrix} x_i & x_i^* \end{bmatrix} = \begin{bmatrix} x_i & x_i^* \end{bmatrix}\begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_i^* \end{bmatrix}$$

is equivalent to the following real equation

$$(A + BF)\begin{bmatrix} \operatorname{Re}(x_i) & \operatorname{Im}(x_i) \end{bmatrix} = \begin{bmatrix} \operatorname{Re}(x_i) & \operatorname{Im}(x_i) \end{bmatrix}\begin{bmatrix} \operatorname{Re}(\lambda_i) & \operatorname{Im}(\lambda_i) \\ -\operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \end{bmatrix}$$

Thus, for each complex eigenvalue, the columns $\begin{bmatrix} x_i & x_i^* \end{bmatrix}$ of X are replaced by $\begin{bmatrix} \operatorname{Re}(x_i) & \operatorname{Im}(x_i) \end{bmatrix}$ to obtain a real X' and the sub-matrix $\begin{bmatrix} \lambda_i & 0 \\ 0 & \lambda_i^* \end{bmatrix}$ of Λ is replaced by $\begin{bmatrix} \operatorname{Re}(\lambda_i) & \operatorname{Im}(\lambda_i) \\ -\operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \end{bmatrix}$ to obtain a real Λ'. Clearly, X' is not the eigenvector matrix of $A + BF$ (c.f. Remark 6), and the latter can be easily obtained by combining the columns of X'. Also, the invertibility of X is equivalent to the invertibility of X'. In the remainder of this section, we assume (2a) to be in the real form.

A. Algorithms for non-sparse MGEAP

In this subsection, we present two iterative gradient based descent algorithms to obtain non-sparse and approximately-sparse solutions to the MGEAP, respectively. To develop the algorithms, we use the Sylvester equation based parametrization [11], [46]. In this parametrization, instead of defining (F, X) as free variables, we define a parameter $G \triangleq FX \in \mathbb{R}^{m \times n}$ as the free variable. With this parametrization, the non-sparse MGEAP is stated as (while ignoring the eigenvector normalization constraints; c.f. Remark 4):

$$\min_G \quad J = \frac{1}{2}\|F\|_F^2 \tag{18}$$

$$\text{s.t.} \quad AX - X\Lambda + BG = 0, \tag{18a}$$

$$F = GX^{-1}. \tag{18b}$$

Note that, for any given G, we can solve the Sylvester equation (18a) to obtain X. Assumption **A4** guarantees that (18a) has a unique solution [47]. Further, we can use (18b) to obtain a non-sparse feedback matrix F. Thus, (18) is an unconstrained optimization problem in the free parameter G.

The Sylvester equation based parametrization requires the unique solution X of (18a) to be non-singular, which holds generically if (i)$(A, -BG)$ is controllable and (ii) $(\Lambda, -BG)$ is observable [48]. Since the system has no fixed modes (c.f. Assumption **A3**), (A, B) is controllable. Therefore, the above two conditions are mild and are satisfied for almost all instances as confirmed in our simulations (see Section V).

The next result provides the gradient and Hessian of the cost J w.r.t. to the parameter $g \triangleq \operatorname{vec}(G)$.

---

[2]Although the minimization cost of the modified MGEAP is $0.5\|F_p + F\|_F^2$, it can be solved using techniques similar to solving MGEAP in (2).

Lemma IV.1. *(Gradient and Hessian of* **J***) The gradient and Hessian of the cost J in* (18) *with respect to g is given by*

$$\frac{dJ}{dg} = \underbrace{\left[(X^{-1} \otimes I_m) + (I_n \otimes B^{\mathsf{T}})\tilde{A}^{-\mathsf{T}}(X^{-1} \otimes F^{\mathsf{T}}) \right]}_{\triangleq Z(F,X)} f, \quad (19)$$

$$\frac{d^2 J}{d^2 g} \triangleq H(F,X) = Z(F,X)Z^{\mathsf{T}}(F,X)$$
$$+ Z_1(F,X)Z^{\mathsf{T}}(F,X) + Z(F,X)Z_1^{\mathsf{T}}(F,X), \quad (20)$$

where $Z_1(F,X) \triangleq (I_n \otimes B^{\mathsf{T}})\tilde{A}^{-\mathsf{T}}(X^{-1}F^{\mathsf{T}} \otimes I_n)T_{m,n}$,
and $\tilde{A} \triangleq A \oplus (-\Lambda^{\mathsf{T}})$.

Proof. Vectorizing (18a) using P.3 and taking the differential, we get

$$\tilde{A}x + (I_n \otimes B)g = 0$$
$$\Rightarrow dx = -\tilde{A}^{-1}(I_n \otimes B)dg. \quad (21)$$

Note that due to Assumption **A4**, $\tilde{A}$ is invertible. Taking the differential of (18b) and vectorizing, we get

$$dF \overset{P.10}{=} dGX^{-1} - \underbrace{GX^{-1}}_{F} dXX^{-1} \quad (22)$$

$$\overset{P.3,P.4}{\Rightarrow} df = (X^{-\mathsf{T}} \otimes I_m)dg - (X^{-\mathsf{T}} \otimes F)dx$$
$$\overset{(21)}{=} \underbrace{\left[(X^{-\mathsf{T}} \otimes I_m) + (X^{-\mathsf{T}} \otimes F)\tilde{A}^{-1}(I_n \otimes B) \right]}_{\overset{P.5}{=} Z^{\mathsf{T}}(F,X)} dg. \quad (23)$$

The differential of cost $J \overset{P.2}{=} \frac{1}{2}f^{\mathsf{T}}f$ is given by $dJ = f^{\mathsf{T}}df$. Using (23) and P.12, we get (19). To derive the Hessian, we compute the second-order differentials of the involved variables. Note that since $g(G)$ is an independent variable, $d^2 g = 0(d^2 G = 0)$ [36]. Further, the second-order differential of (18a) yields $d^2 X = 0$. Rewriting (22) as $dFX = dG - FdX$, and taking its differential and vectorization, we get

$$(d^2 F)X + (dF)(dX) = -(dF)(dX)$$
$$\Rightarrow d^2 F = -2(dF)(dX)X^{-1}$$
$$\overset{P.4}{\Rightarrow} d^2 f = -2(X^{-\mathsf{T}} \otimes dF)dx. \quad (24)$$

Taking the second-order differential of J (and omitting the parameter dependence notation), we get

$$d^2 J = (df)^{\mathsf{T}}df + f^{\mathsf{T}}(d^2 f) = (df)^{\mathsf{T}}df + (d^2 f)^{\mathsf{T}}f$$
$$\overset{(23),(24),P.5}{=} dg^{\mathsf{T}}ZZ^{\mathsf{T}}dg - 2dx^{\mathsf{T}}\underbrace{(X^{-1} \otimes (dF)^{\mathsf{T}})f}_{\overset{P.5}{=} \text{vec}((dF)^{\mathsf{T}}FX^{-\mathsf{T}})}$$
$$\overset{P.5}{=} dg^{\mathsf{T}}ZZ^{\mathsf{T}}dg - 2dx^{\mathsf{T}}(X^{-1}F^{\mathsf{T}} \otimes I_n)T_{m,n}df$$
$$\overset{(21),(23)}{=} dg^{\mathsf{T}}ZZ^{\mathsf{T}}dg$$
$$+ \underbrace{2dg^{\mathsf{T}}(I_n \otimes B^{\mathsf{T}})\tilde{A}^{-\mathsf{T}}(X^{-1}F^{\mathsf{T}} \otimes I_n)T_{m,n}Z^{\mathsf{T}}dg}_{dg^{\mathsf{T}}(Z_1 Z^{\mathsf{T}} + ZZ_1^{\mathsf{T}})dg}.$$
$$(25)$$

The Hessian in (20) follows from (25) and P.12 and the proof is complete. □

Clearly, the first-order optimality condition of the unconstrained optimization problem (18) is given by $\frac{dJ}{dg} = Z(F,X)f = 0$. The next result shows that this condition is equivalent to the first-order optimality conditions of Theorem III.2 without sparsity constraints.

Corollary IV.2. *(Equivalence of first-order optimality conditions) Let* $\bar{F} = 1_{m \times n}$. *Then, the first-order optimality condition* $Z(\hat{F},\hat{X})\hat{f} = 0$ *of* (18) *is equivalent to* (7a)-(7c), *where*

$$\hat{l} \triangleq vec(\hat{L}) = \tilde{A}^{-\mathsf{T}}(\hat{X}^{-1} \otimes \hat{F}^{\mathsf{T}})\hat{f}. \quad (26)$$

Proof. The optimality condition (7b) follows from (18a)-(18b). Equation (7a) can be rewritten as $\hat{F}\hat{X}^{-\mathsf{T}} + B^{\mathsf{T}}\hat{L} = 0$ and its vectorization using P.3 yields $Z(\hat{F},\hat{X})\hat{f} = 0$. Finally, vectorization of the left side of (7c) yields

$$\text{vec}[(A + B\hat{F})^{\mathsf{T}}\hat{L} - \hat{L}\Lambda^{\mathsf{T}}] = \text{vec}[A^{\mathsf{T}}\hat{L} - \hat{L}\Lambda^{\mathsf{T}} + (B\hat{F})^{\mathsf{T}}L]$$
$$\overset{P.3,P.5}{=} \tilde{A}^{\mathsf{T}}\hat{l} + (I_n \otimes (B\hat{F})^{\mathsf{T}})\hat{l}$$
$$\overset{(26)}{=} (\hat{X}^{-1} \otimes \hat{F}^{\mathsf{T}})\hat{f} + (I_n \otimes (B\hat{F})^{\mathsf{T}})\tilde{A}^{-\mathsf{T}}(\hat{X}^{-1} \otimes \hat{F}^{\mathsf{T}})\hat{f}$$
$$\overset{P.5}{=} (I_n \otimes \hat{F}^{\mathsf{T}})Z(\hat{F},\hat{X})\hat{f} = 0.$$

Thus, $\hat{L}$ is the right eigenvector matrix and the proof is complete. □

Using Lemma IV.1, we next present a steepest/Newton descent algorithm to solve the non-sparse MGEAP (18) [42]. In the algorithms presented in this section, we interchangeably use the the matrices (G,F,X) and their respective vectorizations (g,f,x). The conversion of a matrix to the vector (and vice-versa) is not specifically stated in the steps of the algorithms and is assumed wherever necessary.

Algorithm 1: Non-sparse solution to the MGEAP

Input: A, B, Λ, G_0.
Output: Local minimum $(\hat{F}, \hat{X})$ of (18).

Initialize: $G_0, X_0 \leftarrow$ Solution of (18a), $F_0 \leftarrow G_0 X_0^{-1}$
repeat
1    $\alpha \leftarrow$ Compute step size (see below);
2    $g \leftarrow g - \alpha Z(F,X)f$ **or;**
3    $g \leftarrow g - \alpha[H(F,X) + V(F,X)]^{-1}Z(F,X)f$;
    $X \leftarrow$ Solution of Sylvester equation (18a);
    $F \leftarrow GX^{-1}$;
until convergence;
return (F,X)

Steps 2 and 3 of Algorithm 1 represent the steepest and (damped) Newton descent steps, respectively. Since, in general, the Hessian $H(F,X)$ is not positive-definite, the Newton descent step may not result in a decrease of the cost. Therefore, we add a Hermitian matrix $V(F,X)$ to the Hessian to make it positive definite [42]. We will comment on the choice of $V(F,X)$ in Section V. In step 1, the step size α can be determined by backtracking line search or Armijo's rule [42]. For a detailed discussion of the steepest/Newton descent methods, the interested reader is referred to [42]. The computationally intensive steps in Algorithm 1 are solving the

Sylvester equation (18a) and evaluating the inverses of X and $H + V$.

Next, we present a relaxation of the optimization problem (2) and a corresponding algorithm that provides approximately-sparse solutions to the MGEAP. We remove the explicit feedback sparsity constraints (2c) and modify the cost function to penalize it when these sparsity constraints are violated. Using the Sylvester equation based parametrization, the relaxed optimization problem is stated as (while ignoring the eigenvector normalization constraints; c.f. Remark 4):

$$\min_{G} \quad J_W = \frac{1}{2} \, ||W \circ F||_F^2 \qquad (27)$$

$$\text{s.t.} \quad \text{(18a) and (18b) hold true,}$$

where $W \in \mathbb{R}^{m \times n}$ is a weighing matrix that penalizes the cost for violation of sparsity constraints, and is given by

$$W_{ij} = \begin{cases} 1 & \text{if } \bar{\mathsf{F}}_{ij} = 1, \text{ and} \\ \gg 1 & \text{if } \bar{\mathsf{F}}_{ij} = 0. \end{cases}$$

As the penalty weights of W corresponding to the sparse entries of F increase, an optimal solution of (27) becomes more sparse and approaches towards the optimal solution of (2). Note that the relaxed problem (27) corresponds closely to the non-sparse MGEAP (18). Thus, we use a similar gradient based approach to obtain its solution.

Lemma IV.3. *(Gradient and Hessian of $\mathbf{J_W}$) The gradient and Hessian of the cost J_W in (27) with respect to g is given by*

$$\frac{dJ_W}{dg} = Z(F,X)\bar{W}f, \qquad (28)$$

$$\frac{d^2 J_W}{d^2 g} \triangleq H_W(F,X) = Z(F,X)\bar{W}Z^\mathsf{T}(F,X)$$
$$+ Z_{1,W}(F,X)Z^\mathsf{T}(F,X) + Z(F,X)Z_{1,W}^\mathsf{T}(F,X), \qquad (29)$$

$$\textit{where} \quad \bar{W} \triangleq diag(vec(W \circ W)) \quad \textit{and,}$$

$$Z_{1,W}(F,X) \triangleq (I_n \otimes B^\mathsf{T})\tilde{A}^{-\mathsf{T}}(X^{-1}(W \circ W \circ F)^\mathsf{T} \otimes I_n)T_{m,n}$$

Proof. Since the constraints of optimization problems (18) and (27) are same, Equations (21)-(24) from Lemma IV.1 also hold true for problem (27). Now, $J_W \overset{P.2,P.8}{=} \frac{1}{2}(vec(W) \circ f)^\mathsf{T}(vec(W) \circ f) = \frac{1}{2}f^\mathsf{T}\bar{W}f$. Thus, $dJ_W = f^\mathsf{T}\bar{W}df$ and $d^2 J_W = (df)^\mathsf{T}\bar{W}df + f^\mathsf{T}\bar{W}d^2 f$. Using the relation $vec(W \circ W \circ F) = \bar{W}f$, the remainder of the proof is similar to proof of Lemma IV.1. $\qquad \square$

Using Lemma IV.3, we next present an algorithm to obtain an approximately-sparse solution to the MGEAP.

The step size rule and modification of the Hessian in Algorithm 2 is similar to Algorithm 1.

B. Algorithms for Sparse EAP

In this subsection, we present two heuristic algorithms to obtain a sparse solution to the EAP (i.e. any sparse solution, which is not necessarily minimum-gain). This involves finding a pair (F, X) which satisfies the eigenvalue assignment and sparsity constraints (2a), (2c). We begin with a result that combines these two constraints.

Algorithm 2: Approximately-sparse solution to the MGEAP

Input: A, B, Λ, W, G_0
Output: Local minimum $(\hat{F}, \hat{X})$ of (27).

Initialize: $G_0, X_0 \leftarrow$ Solution of (18a), $F_0 \leftarrow G_0 X_0^{-1}$
repeat
1    $\alpha \leftarrow$ Update step size;
2    $g \leftarrow g - \alpha Z(F,X)\bar{W}f$ **or;**
3    $g \leftarrow g - \alpha[H_W(F,X) + V_W(F,X)]^{-1}Z(F,X)\bar{W}f$;
   $X \leftarrow$ Solution of Sylvester equation (18a);
   $F \leftarrow GX^{-1}$;
until convergence;
return (F, X)

Lemma IV.4. *(Feasibility of $(\mathbf{F}, \mathbf{X})$) An invertible matrix $X \in \mathbb{R}^{n \times n}$ satisfies (2a) and (2c) if and only if*

$$\tilde{a}(x) \in \mathcal{R}(\tilde{B}(X)) \quad \textit{where,} \qquad (30)$$
$$\tilde{a}(x) \triangleq \tilde{A}x, \tilde{B}(X) \triangleq -(X^\mathsf{T} \otimes B)P_{\bar{\mathsf{F}}}, P_{\bar{\mathsf{F}}} \triangleq diag(vec(\bar{\mathsf{F}})).$$

Further, if (30) holds true, then the set of sparse feedback matrices that satisfy (2a) and (2c) is given by

$$\mathcal{F}_X = \{P_{\bar{\mathsf{F}}}f_{ns} : \tilde{B}(X)f_{ns} = \tilde{a}(x), f_{ns} \in \mathbb{R}^{mn}\}. \qquad (31)$$

Proof. Any feedback f which satisfies the sparsity constraint (2c) can be written as $f = P_{\bar{\mathsf{F}}}f_{ns}$ where $f_{ns} \in \mathbb{R}^{mn}$ is a non-sparse vector[3]. Vectorizing (2a) using P.3 and P.4, and substituting $f = P_{\bar{\mathsf{F}}}f_{ns}$, we get

$$\tilde{A}x = -(X^\mathsf{T} \otimes B)P_{\bar{\mathsf{F}}}f_{ns}, \qquad (32)$$

from which (30) and (31) follow. $\qquad \square$

Based on Lemma IV.4, we develop a heuristic iterative algorithm to obtain a sparse solution to the EAP. The algorithm starts with a non-sparse EAP solution (F_0, X_0) that does not satisfy (2c) and (30). Then, it takes repeated projections of $\tilde{a}(x)$ on $\mathcal{R}(\tilde{B}(X))$ to update X and F, until a sparse solution is obtained.

Algorithm 3: Sparse solution to EAP

Input: $A, B, \Lambda, \bar{\mathsf{F}}, G_0, iter_{max}$.
Output: (F, X) satisfying (2a) and (2c).

Initialize: $G_0, X_0 \leftarrow$ Solution of (18a), $F_{ns,0} \leftarrow G_0 X_0^{-1}$, $i \leftarrow 0$
repeat
1    $\tilde{a}(x) \leftarrow \tilde{B}(X)[\tilde{B}(X)]^+ \tilde{a}(x)$;
2    $x \leftarrow \tilde{A}^{-1}\tilde{a}(x)$;
3    $X \leftarrow$ Normalize X;
   $i \leftarrow i + 1$
until convergence or $i > iter_{max}$;
return $(f \in \mathcal{F}_X$ in (31)$, X)$

In step 1 of Algorithm 3, we update $\tilde{a}(x)$ by projecting it on $\mathcal{R}(\tilde{B}(X))$. Step 2 computes x from $\tilde{a}(x)$ using the fact that $\tilde{A}$

---

[3] Since f satisfies (4), it can also be characterized as $f = (I_{mn} - Q^+ Q)f_{ns}$, and thus $P_{\bar{\mathsf{F}}} = I_{mn} - Q^+ Q$.

is invertible (c.f. Assumption **A4**). Finally, the normalization in step 3 is performed to ensure invertibility of X[4].

Next, we develop a second heuristic algorithm for solving the sparse EAP problem using the non-sparse MGEAP solution in Algorithm 1. The algorithm starts with a non-sparse EAP solution (F_0, X_0). In each iteration, it sparsifies the feedback to obtain $f = P_{\bar{\mathsf{F}}} f_{ns}$ (or $F = \bar{\mathsf{F}} \circ F_{ns}$), and then solves the following non-sparse MGEAP

$$\min_{F_{ns}, X} \quad \frac{1}{2} \, ||F_{ns} - F||_F^2 \tag{33}$$

$$\text{s.t.} \quad (A + BF_{ns})X = X\Lambda, \tag{33a}$$

to update F_{ns} that is *close* to the sparse F. Thus, using the heuristics of repeated sparsification of the solution of non-sparse MGEAP in (33), the algorithm obtains a sparse solution. Note that a solution $\hat{F}_{ns}$ of the optimization problem (33) with parameters (A, B, Λ, F) satisfies $\hat{F}_{ns} = F + \hat{K}_{ns}$, where $\hat{K}_{ns}$ is a solution of the optimization problem (18) with parameters $(A + BF, B, \Lambda)$. Thus, we can use Algorithm 1 to solve (33).

Algorithm 4: Projection-based sparse solution to EAP

Input: $A, B, \Lambda, \bar{\mathsf{F}}, G_0, iter_{max}$.
Output: (F, X) satisfying (2a) and (2c).

Initialize: $G_0, X_0 \leftarrow$ Solution of (18a),
  $F_{ns,0} \leftarrow G_0 X_0^{-1}, i \leftarrow 0$
repeat
  | $F \leftarrow \bar{\mathsf{F}} \circ F_{ns}$;
1 | $(K_{ns}, X) \leftarrow$ Algorithm 1$(A + BF, B, \Lambda)$;
  | $F_{ns} \leftarrow F + K_{ns}$;
  | $i \leftarrow i + 1$;
until convergence or $i > iter_{max}$;
return (F, X)

Remark 7. (Comparison of EAP Algorithms 3 and 4)
1. Projection property: *In general, Algorithm 3 results in a sparse EAP solution F that is considerably different from the initial non-sparse $F_{ns,0}$. In contrast, Algorithm 4 provides a sparse solution F that is close to $F_{ns,0}$. This is due to the fact that Algorithm 4 updates the feedback by solving the optimization problem (33), which minimizes the deviations between successive feedback matrices. Thus, Algorithm 4 provides a good (although not necessarily orthogonal) projection of a given non-sparse EAP solution $F_{ns,0}$ on the space of sparse EAP solutions.*
2. Complexity: *The computational complexity of Algorithm 4 is considerably larger than that of Algorithm 3. This is because Algorithm 4 requires a solution of a non-sparse MGEAP problem in each iteration. In contrast, Algorithm 3 only requires projections on the range space of a matrix in each iteration. Thus, Algorithm 3 is considerably faster as compared to Algorithm 4.*
3. Convergence: *Although we do not formally prove the convergence of heuristic Algorithms 3 and 4) in this paper,*

a comprehensive simulation study in Subsection V-B suggests that Algorithm 4 converges in almost all instances. In contrast, Algorithm 3 converges in much fewer instances and its convergence deteriorates considerably as the number of sparsity constraints increase (see Subsection V-B). □

C. Algorithm for sparse MGEAP

In this subsection, we present an iterative projected gradient algorithm to compute the sparse solutions of the MGEAP in (2). The algorithm consists of two loops. The outer loop is same as the non-sparse MGEAP Algorithm 1 (using steepest descent) with an additional projection step, which constitutes the inner loop. Figure 1 represents one iteration of the algorithm. First, the gradient $Z(F_k, X_k) f_k$ is computed at a current point G_k (equivalently (F_k, X_k), where F_k is sparse). Next, the gradient is projected on the tangent plane of the sparsity constraints (4), which is given by

$$\mathcal{T}_F = \left\{ y \in \mathbb{R}^{mn} : \left[\frac{d(Qf)}{dg} \right]^\mathsf{T} y = 0 \right\}$$

$$= \left\{ y \in \mathbb{R}^{mn} : QZ^\mathsf{T}(F, X) y = 0 \right\}. \tag{34}$$

From P.14, the projection of the gradient on $\mathcal{T}_F$ is given by $P_{F_k} Z(F_k, X_k) f_k$, where $P_{F_k} = I_{mn} - [QZ^\mathsf{T}(F_k, X_k)]^+ [QZ^\mathsf{T}(F_k, X_k)]$. Next, a move is made in the direction of the projected gradient to obtain $G_{ns,k}(F_{ns,k}, X_{ns,k})$. Finally, the orthogonal projection of $G_{ns,k}$ is taken on the space of sparsity constraints to obtain $G_{k+1}(F_{k+1}, X_{k+1})$. This orthogonal projection is equivalent to solving (33) with sparsity constraints (2c), which in turn is equivalent to the original sparse MGEAP (2). Thus, the orthogonal projection step is as difficult as the original optimization problem. To address this issue, we use the heuristic Algorithm 4 to compute the projections. Although the projections obtained using Algorithm 4 are not necessarily orthogonal, they are typically good (c.f. Remark 7).
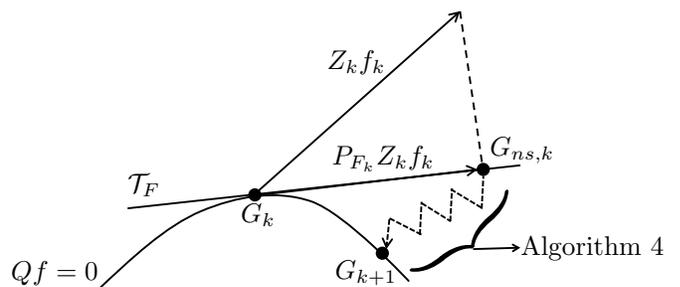

Fig. 1. A single iteration of Algorithm 5.

Algorithm 5 is computationally intensive due to the use of Algorithm 4 in step 1 to compute the projection on the space of sparse matrices. In fact, the computational complexity of Algorithm 5 is one order higher than that of non-sparse MGEAP Algorithm 1. However, a way to considerably reduce the number of iterations of Algorithm 5 is to initialize it using the approximately-sparse solution obtained by Algorithm 2. In this case, Algorithm 5 starts *near* the the local minimum and, thus, its convergence time reduces considerably.

---

[4]Since X is not an eigenvector matrix, we compute the eigenvectors from X, normalize them, and then recompute real X from the normalized eigenvectors.

Algorithm 5: Sparse solution to the MGEAP

Input: $A, B, \Lambda, \bar{\mathsf{F}}, G_0, iter_{max}$.
Output: Local minimum $(\hat{F}, \hat{X})$ of (2).

Initialize:
$(F_0, X_0) \leftarrow$ Algorithm $4(A, B, \Lambda, \bar{\mathsf{F}}, G_0, iter_{max})$,
$G_0 \leftarrow F_0 X_0, \; i \leftarrow 0$
repeat
    $\alpha \leftarrow$ Update step size;
    $g_{ns} \leftarrow g - \alpha P_F Z(F, X) f$;
    $X_{ns} \leftarrow$ Solution of Sylvester equation (18a);
    $F_{ns} \leftarrow G_{ns} X_{ns}^{-1}$;
1    $(F, X) \leftarrow$ Algorithm $4(A, B, \Lambda, \bar{\mathsf{F}}, G_{ns}, iter_{max})$;
    $G \leftarrow FX$;
    $i \leftarrow i + 1$;
until convergence or $i > iter_{max}$;
return (F, X)

V. SIMULATION STUDIES

In this section, we present the implementation details of the algorithms developed in Section IV and provide numerical simulations to illustrate their properties.

A. Implementation aspects of the algorithms

In the Newton descent step (step 3) of Algorithm 1, we need to choose (omitting the parameter dependence notation) V such that $H + V$ is positive-definite. We choose $V = \delta I_{mn} - Z_1 Z^{\mathsf{T}} - Z Z_1^{\mathsf{T}}$ where, $0 < \delta \ll 1$. Thus, from (20), we have: $H + V = Z Z^{\mathsf{T}} + \epsilon I_{mn}$. Clearly, $Z Z^{\mathsf{T}}$ is positive-semidefinite and the small additive term ϵI_{mn} ensures that $H + V$ is positive-definite. Note that other possible choices of V also exist. In step 1 of Algorithm 1, we use the Armijo rule to compute the step size α. Finally, we use $\left\| \frac{dJ}{dg} \right\|_2 < \epsilon$, $0 < \epsilon \ll 1$ as the convergence criteria of Algorithm 1. For Algorithm 2, we analogously choose $V_W = \delta I_{mn} - Z_{1,W} Z^{\mathsf{T}} - Z Z_{1,W}^{\mathsf{T}}$ and the same convergence criteria and step size rule as Algorithm 1. In both algorithms, if we encounter a scenario in which the solution X of (18a) is singular (c.f. paragraph below (18b)), we perturb G slightly such that the new solution is non-singular, and then continue the iterations. We remark that such instances occur extremely rarely in our simulations.

For the sparse EAP Algorithm 3, we use the convergence criteria $e_X = \|[I_{n^2} - \tilde{B}(X)(\tilde{B}(X))^+]\tilde{a}(x)\|_2 < \epsilon$, $0 < \epsilon \ll 1$. For Algorithm 4, we use the convergence criteria $e_F = \|F - \bar{\mathsf{F}} \circ F\|_F < \epsilon$, $0 < \epsilon \ll 1$. Thus, the iterations of these algorithms stop when x lies in a certain subspace and when the sparsity error becomes *sufficiently* small (within the specified tolerance), respectively. Further, note that Algorithm 4 uses Algorithm 1 in step 1 without specifying an initial condition G_0 for the latter. This is because in step 1, we effectively run Algorithm 1 for multiple initial conditions in order to capture its global minima. We remark that the capture of global minima by Algorithm 1 is crucial for convergence of Algorithm 4.

As the iterations of Algorithm 4 progress, the sparse matrix F achieves eigenvalue assignment with increasing accuracy. As a result, near the convergence of Algorithm 4, the eigenvalues of $A + BF$ and Λ are very close to each other. This creates numerical difficulties when Algorithm 1 is used with parameters $(A + BF, B, \Lambda)$ in step 1 (see Assumption **A4**). To avoid this issue, we run Algorithm 1 using a preliminary feedback F_p, as explained below Assumption **A4**.

Finally, for Algorithm 5, we use the following convergence criteria: $\left\| P_F \frac{dJ}{dg} \right\|_2 < \epsilon$, $0 < \epsilon \ll 1$. We choose stopping tolerance ϵ typically between 10^{-6} and 10^{-5} for all the algorithms.

B. Numerical study

We begin this subsection with the following example:

$$A = \begin{bmatrix} -3.7653 & -2.1501 & 0.3120 & -0.2484 \\ 1.6789 & 1.0374 & -0.5306 & 1.3987 \\ -2.1829 & -2.5142 & -1.2275 & 0.2833 \\ -13.6811 & -9.6804 & -0.5242 & 2.9554 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 1 & 2 & 5 \\ 1 & 3 & 4 & 2 \end{bmatrix}^{\mathsf{T}}, \qquad \bar{\mathsf{F}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix},$$

$$\mathcal{S} = \{-2, -1, -0.5 \pm j\}.$$

The eigenvalues of A are $\Gamma(A) = \{-2, -1, 1 \pm 2j\}$. Thus, the feedback F is required to move two unstable eigenvalues into the stable region while keeping the other two stable eigenvalues fixed. Table I shows the non-sparse, approximately-sparse and sparse solutions obtained by Algorithms 1, 2 and 5, respectively, and Figure 2 shows a sample iteration run of these algorithms. Since the number of iterations taken by the algorithms to converge depends on their starting points, we report the average number of iterations taken over 1000 random starting points. Further, to obtain approximately-sparse solutions, we use the weighing matrix with $W_{ij} = w$ if $\bar{\mathsf{F}}_{ij} = 1$. All the algorithms obtain three local minima, among which the first is the global minimum. The second column in $\hat{X}$ and $\hat{L}$ is conjugate of the first column (c.f. Remark 3). It can be verified that the non-sparse and sparse solutions satisfy the optimality conditions of Theorem III.2.

For the non-sparse solution, the number of iterations taken by Algorithm 1 with steepest descent are considerably larger than the Newton descent. This is because the steepest descent converges very slowly near a local minimum. Therefore, we use Newton descent steps in Algorithms 1 and 2. Next, observe that the entries at the sparsity locations of the local minimum feedbacks obtained by Algorithm 2 have small magnitude. Further, the average number of Newton descent iterations for convergence and the norm of the feedback obtained of Algorithm 2 is larger as compared to Algorithm 1. This is because the approximately-sparse optimization problem (27) is effectively more restricted than its non-sparse counterpart (18).

Finally, observe that the solutions of Algorithm 5 are sparse. Note that Algorithm 5 involves the use of projection Algorithm 4, which in turn involves running Algorithm 1 multiple times. Thus, for a balanced comparison, we present the total

TABLE I
COMPARISON OF MGEAP SOLUTIONS BY ALGORITHMS 1, 2 AND 5

Non-sparse solutions by Algorithm 1

$$\hat{X}_1 = \begin{bmatrix} -0.0831+0.3288j & -0.5053 & 0.2031 \\ 0.1919-0.4635j & 0.5612 & -0.2505 \\ 0.1603+0.5697j & 0.5617 & 0.8441 \\ 0.3546+0.3965j & -0.3379 & 0.4283 \end{bmatrix} \hat{x}_1^*$$

$$\hat{L}_1 = \begin{bmatrix} -0.5569+1.4099j & -0.9154 & 2.5568 \\ -0.2967+1.0244j & -0.5643 & 1.7468 \\ -0.0687+0.0928j & 0.0087 & 0.2722 \\ 0.2259-0.2523j & 0.0869 & -0.4692 \end{bmatrix} \hat{l}_1^*$$

$$\hat{F}_1 = \begin{bmatrix} -0.1111 & -0.1089 & -0.0312 & -0.4399 \\ -0.1774 & -0.2072 & 0.0029 & 0.1348 \end{bmatrix}, \|\hat{F}_1\|_F = 0.5580$$

$$\hat{F}_2 = \begin{bmatrix} 0.3817 & -0.3349 & 0.7280 & -0.2109 \\ -0.0873 & -0.4488 & -0.4798 & -0.0476 \end{bmatrix}, \|\hat{F}_2\|_F = 1.1286$$

$$\hat{F}_3 = \begin{bmatrix} 0.3130 & 2.0160 & 1.2547 & -0.6608 \\ 0.0683 & -0.7352 & -0.0748 & -1.0491 \end{bmatrix}, \|\hat{F}_3\|_F = 2.7972$$

Average # of Steepest/Newton descent iterations = 5402.1/15.5

Approximately-sparse solutions by Algorithm 2 with w = 30

$$\hat{F}_1 = \begin{bmatrix} 0.9652 & -1.3681 & 0.0014 & -0.0021 \\ 0.4350 & -0.0023 & -0.6746 & -0.1594 \end{bmatrix}, \|\hat{F}_1\|_F = 1.8636$$

$$\hat{F}_2 = \begin{bmatrix} -1.0599 & -1.7036 & -0.0013 & -0.0071 \\ -0.1702 & -0.0057 & 0.0263 & -0.0582 \end{bmatrix}, \|\hat{F}_2\|_F = 2.0146$$

$$\hat{F}_3 = \begin{bmatrix} 3.3768 & 2.2570 & 0.0410 & -0.0098 \\ -1.4694 & -0.0061 & 0.1242 & -3.8379 \end{bmatrix}, \|\hat{F}_3\|_F = 5.7795$$

Average # of Newton descent iterations = 19.1

Sparse solutions by Algorithm 5

$$\hat{X}_1 = \begin{bmatrix} -0.3370+0.3296i & -0.4525 & 0.5168 \\ 0.2884-0.1698i & 0.1764 & -0.3007 \\ -0.4705+0.5066i & -0.7478 & 0.7831 \\ 0.2754+0.3347i & -0.4527 & 0.1711 \end{bmatrix} \hat{x}_1^*$$

$$\hat{L}_1 = \begin{bmatrix} 25.2072+16.5227i & 44.6547 & 82.3616 \\ 11.0347+11.8982i & 12.5245 & 35.8266 \\ -12.2600-5.8549i & -24.5680 & -39.1736 \\ -0.6417-2.2060i & -0.4383 & -3.6464 \end{bmatrix} \hat{l}_1^*$$

$$\hat{F}_1 = \begin{bmatrix} 0.9627 & -1.3744 & 0.0000 & 0.0000 \\ 0.4409 & 0.0000 & -0.6774 & -0.1599 \end{bmatrix}, \|\hat{F}_1\|_F = 1.8694$$

$$\hat{F}_2 = \begin{bmatrix} -1.0797 & -1.7362 & 0.0000 & 0.0000 \\ -0.1677 & 0.0000 & 0.0264 & -0.0610 \end{bmatrix}, \|\hat{F}_2\|_F = 2.0525$$

$$\hat{F}_3 = \begin{bmatrix} 3.4465 & 2.2568 & 0.0000 & 0.0000 \\ -1.8506 & -0.0000 & 0.3207 & -4.0679 \end{bmatrix}, \|\hat{F}_3\|_F = 6.0866$$

Average # of Newton descent iterations:
1. Using random initialization = 8231
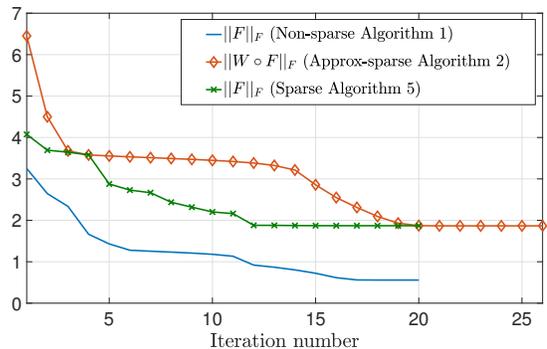2. Using initialization by Algorithm 2 = 715

Fig. 2. Optimization costs for a sample run of Algorithms 1, 2 and 5 (the algorithms converge to their global minima).
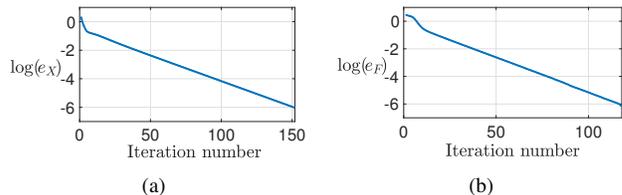

Fig. 3. Projection and sparsity errors for a sample run of (a) Algorithm 3 and (b) Algorithm 4, respectively.

projection error e_X and the sparsity error e_F decrease as the respective algorithms progress.

Next, we provide an empirical verification of the convergence of heuristic Algorithms 3 and 4. Let the sparsity ratio (SR) be defined as the ratio of the number of sparse entries to the total number of entries in F (i.e. $SR = \frac{\text{Number of } 0's \text{ in } \bar{F}}{mn}$). We perform 1000 random executions of both the algorithms. In each execution, n is randomly selected between 4 and 20 and m is randomly selected between 2 and n. Then, matrices (A, B) are randomly generated with appropriate dimensions. Next, a binary sparsity pattern matrix $\bar{F}$ is randomly generated with the number of sparsity entries given by $\lfloor SR \times mn \rfloor$, where $\lfloor \cdot \rfloor$ denotes rounding to the next lowest integer. To ensure feasibility (c.f. Assumption **A2** and discussion below), we pick the desired eigenvalue set $\mathcal{S}$ randomly as follows: we select a random F_r which satisfies the selected sparsity pattern $\bar{F}$, select $\mathcal{S} = \Gamma(A + BF_r)$. Finally, we set $iter_{max} = 1000$ and select a random starting point $G_0(F_0, X_0)$, and run both algorithms from the same starting point. Let F_{sol} denote the feedback solution obtained by Algorithms 3 and 4, respectively, and let $d_{F_{sol},F_0} \triangleq \|F_{sol} - F_0\|_F$ denote the distance between the starting point F_0 and the final solution.

number of Newton descent iterations of Algorithm 1 involved in the execution of Algorithm 5.[5] From Table I, we can observe that Algorithm 5 involves considerably more Newton descent iterations compared to Algorithms 1 and 2, since it involves computationally intensive projection calculations by Algorithm 4. One way to reduce its computation time is to initialize is *near* the local minimum using the approximately sparse solution of Algorithm 2.

Figure 3 shows a sample run of EAP Algorithms 3 and 4 for $G_0 = \begin{bmatrix} -1.0138 & 0.6851 & -0.1163 & 0.8929 \\ -1.8230 & -2.2041 & -0.1600 & 0.7293 \end{bmatrix}$. The sparse feedback obtained by Algorithms 3 and 4 are $F = \begin{bmatrix} 0.1528 & -2.6710 & 0.0000 & 0.0000 \\ -0.8382 & 0.0000 & 0.1775 & -0.1768 \end{bmatrix}$ and $F = \begin{bmatrix} 4.2595 & 4.2938 & 0.0000 & 0.0000 \\ -0.2519 & 0.0000 & -2.1258 & -1.3991 \end{bmatrix}$, respectively. Observe that the

[5]Although the number of outer iteration of Algorithm 5 are considerably less, for instance, 20 in Figure 2.

TABLE II
CONVERGENCE PROPERTIES OF EAP ALGORITHMS 3 AND 4

SR	Algorithm 3	Algorithm 4
1/4	Convergence instances = 427 Average $d_{F_{sol},F_0} = 8.47$	Convergence instances = 997 Average $d_{F_{sol},F_0} = 1.28$
1/2	Convergence instances = 220 Average $d_{F_{sol},F_0} = 4.91$	Convergence instances = 988 Average $d_{F_{sol},F_0} = 1.61$
2/3	Convergence instances = 53 Average $d_{F_{sol},F_0} = 6.12$	Convergence instances = 983 Average $d_{F_{sol},F_0} = 1.92$

Table II shows the convergence results of Algorithms 3 and 4 for three different sparsity ratios. While the convergence of Algorithm 3 deteriorates as F becomes more sparse, Algorithm 4 converges in almost all instances. This implies that in the context of Algorithm 5, Algorithm 4 provides a valid projection in step 1 in almost all instances. We remark that in the rare case that Algorithm 4 fails to converge, we can reduce the step size α to obtain a new G_{ns} and compute its projection. Further, we compute the average of distance d_{F_{sol},F_0} over all executions of the Algorithms 3 and 4 that converge. Observe that the average distance for Algorithm 4 is smaller than Algorithm 4. This shows that Algorithm 4 provides considerably better projection of F_0 in the space of sparse matrices as compared to Algorithm 3 (c.f. Remark 7).

VI. CONCLUSION

In this paper we studied the MGEAP for LTI systems with arbitrary sparsity constraints on the static feedback matrix. We presented an analytical characterization of its locally optimal solutions, thereby providing explicit relations between an optimal solution and the eigenvector matrices of the associated closed loop system. We also provided a geometric interpretation of an optimal solution of the non-sparse MGEAP. Using the Sylvester equation based parametrization, we developed a projected gradient descent algorithm to solve the MGEAP. We also presented two novel algorithms for solving the sparse EAP and an algorithm to obtain approximately sparse solution to the MGEAP. Various convergence properties of the algorithms were discussed using numerical studies.

The analysis in the paper is developed under the assumption that the sparse EAP problem with static feedback is feasible. A future direction of research include finding necessary/sufficient conditions for the feasibility of the sparse EAP, and formulating a convex relaxation of the sparse MGEAP with guaranteed distance from optimality.

REFERENCES

[1] P. J. Antsaklis and A. N. Michael. *Linear systems*. Birkhauser, 2005.
[2] J. Kautsky, N. K. Nichols, and P. Van Dooren. Pole assignment in linear state feedback. *International Journal of Control*, 44(5):1129–1155, 1985.
[3] A. Pandey, R. Schmid, and T. Nguyen. Performance survey of minimum gain exact pole placement methods. In *European Control Conference*, pages 1808–1812, Linz, Austria, 2015.
[4] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.
[5] B. Bamieh, F. Paganini, and M. A. Dahleh. Distributed control of spatially invariant systems. *IEEE Transactions on Automatic Control*, 47(7):1091–1107, 2002.
[6] F. Lin, M. Fardad, and M. R. Jovanović. Augmented Lagrangian approach to design of structured optimal state feedback gains. *IEEE Transactions on Automatic Control*, 56(12):2923–2929, 2011.
[7] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yuksel. Information structures in optimal decentralized control. In *IEEE Conf. on Decision and Control*, pages 1291 –1306, Maui, HI, USA, December 2012.
[8] R. Schmid, A. Pandey, and T. Nguyen. Robust pole placement with moores algorithm. *IEEE Transactions on Automatic Control*, 59(2):500–505, 2014.
[9] M. Ait Rami, S. E. Faiz, A. Benzaouia, and F. Tadeo. Robust exact pole placement via an LMI-based algorithm. *IEEE Transactions on Automatic Control*, 54(2):394–398, 2009.
[10] R. Byers and S. G. Nash. Approaches to robust pole assignment. *International Journal of Control*, 49(1):97–117, 1989.

[11] A. Varga. Robust pole assignment via Sylvester equation based state feedback parametrization. In *IEEE International Symposium on Computer-Aided Control System Design*, Anchorage, AK, USA, 2000.
[12] E. K. Chu. Pole assignment via the Schur form. *Systems & Control Letters*, 56(4):303–314, 2007.
[13] A. L. Tits and Y. Yang. Globally convergent algorithms for robust pole assignment by state feedback. *IEEE Transactions on Automatic Control*, 41(10):1432–1452, 1996.
[14] E. K. Chu. Optimization and pole assignment in control system design. *International Journal of Applied Mathematics and Computer Science*, 11(5):1035–1053, 2001.
[15] A. Pandey, R. Schmid, T. Nguyen, Y. Yang, V. Sima, and A. L. Tits. Performance survey of robust pole placement methods. In *IEEE Conf. on Decision and Control*, pages 3186–3191, Los Angeles, CA, USA, 2014.
[16] B. A. White. Eigenstructure assignment: a survey. *Proceedings of the Institution of Mechanical Engineers*, 209(1):1–11, 1995.
[17] L. F. Godbout and D. Jordan. Modal control with feedback-gain constraints. *Proceedings of the Institution of Electrical Engineers*, 122(4):433–436, 1975.
[18] B. Kouvaritakis and R. Cameron. Pole placement with minimised norm controllers. *IEE Proceedings D - Control Theory and Applications*, 127(1):32–36, 1980.
[19] H. K. Tam and J. Lam. Newton's approach to gain-controlled robust pole placement. *IEE Proceedings - Control Theory and Applications*, 144(5):439–446, 1997.
[20] A. Varga. A numerically reliable approach to robust pole assignment for descriptor systems. *Future Generation Computer Systems*, 19(7):1221–1230, 2003.
[21] A. Varga. Robust and minimum norm pole assignment with periodic state feedback. *IEEE Transactions on Automatic Control*, 45(5):1017–1022, 2000.
[22] S. Datta, B. Chaudhuri, and D. Chakraborty. Partial pole placement with minimum norm controller. In *IEEE Conf. on Decision and Control*, pages 5001–5006, Atlanta, GA, USA, 2010.
[23] S. Datta and D. Chakraborty. Feedback norm minimization with regional pole placement. *International Journal of Control*, 87(11):2239–2251, 2014.
[24] S. H. Wang and E. J. Davison. On the stabilization of decentralized control systems. *IEEE Transactions on Automatic Control*, 18(5):473–478, 1973.
[25] J. P. Corfmat and A. S. Morse. Decentralized control of linear multivariable systems. *Automatica*, 12(5):479–495, 1976.
[26] B. D. O. Anderson and D. J. Clements. Algebraic characterization of fixed modes in decentralized control. *Automatica*, 17(5):703–712, 1981.
[27] E. J. Davison and U. Ozguner. Characterizations of decentralized fixed modes for interconnected systems. *Automatica*, 19(2):169–182, 1983.
[28] M. E. Sezer and D. D. Siljak. Structurally fixed modes. *Systems & Control Letters*, 1(1):60–64, 1981.
[29] V. Pichai, M. E. Sezer, and D. D. Siljak. A graph-theoretic characterization of structurally fixed modes. *Automatica*, 20(2):247–250, 1984.
[30] V. Pichai, M. E. Sezer, and D. D. Siljak. A graphical test for structurally fixed modes. *Mathematical Modelling*, 4(4):339–348, 1983.
[31] A. Alavian and M. Rotkowitz. Fixed modes of decentralized systems with arbitrary information structure. In *International Symposium on Mathematical Theory of Networks and Systems*, pages 913–919, Groningen, The Netherlands, 2014.
[32] A. Alavian and M. Rotkowitz. Stabilizing decentralized systems with arbitrary information structure. In *IEEE Conf. on Decision and Control*, pages 4032–4038, Los Angeles, CA, USA, 2014.
[33] S. Pequito, S. Kar, and A. P. Aguiar. A framework for structural input/output and control configuration selection in large-scale systems. *IEEE Transactions on Automatic Control*, 61(2):303–318, 2016.
[34] S. Moothedath, P. Chaporkar, and M. N. Belur. Minimum cost feedback selection for arbitrary pole placement in structured systems. *Arxiv preprint arXiv:1706.06911*, 2017.
[35] Petersen K. B. and M. S. Pedersen. *The Matrix Cookbook*. Technical University of Denmark, 2012.
[36] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons, 1999.
[37] D. Kressner and M. Voigt. Distance problems for linear dynamical systems. In *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 559–583. Springer, 2015.
[38] J. Rosenthal and J. C. Willems. Open problems in the area of pole placement. In *Open Problems in Mathematical Systems and Control Theory*, page 181191. Springer, 1999.

[39] J. Leventides and N. Karcanias. Sufficient conditions for arbitrary pole assignment by constant decentralized output feedback. *Mathematics of Control, Signals and Systems*, 8(3):222–240, 1995.

[40] A. Hjorungnes and D. Gesbert. Complex-valued matrix differentiation: Techniques and key results. *IEEE Transactions on Signal Processing*, 55(6):2740–2746, 2007.

[41] D. H. Brandwood. A complex gradient operator and its application in adaptive array theory. *IEE Proceedings F - Communications, Radar and Signal Processing*, 130(1):11–16, 1983.

[42] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, 2 edition, 1984.

[43] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 1995.

[44] P Lancaster. *The Theory of Matrices*. American Press, 1985.

[45] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, 1988.

[46] S. P. Bhattacharyya and E. De Souza. Pole assignment via Sylvesters eqauation. *Systems & Control Letters*, 1(4):261–263, 1982.

[47] N. J. Hingham. *Accuracy and Stability of Numerical Algorithms*. SIAM, 2002.

[48] S. P. Bhattacharyya and E. De Souza. Controllability, observability and the solution of AX-XB=C. *Linear algebra and its applications*, 39(1):167–188, 1981.