SIAM J. NUMER. ANAL.
Vol. 56, No. 2, pp. 1022–1047

FINDING THE NEAREST POSITIVE-REAL SYSTEM*

NICOLAS GILLIS† AND PUNIT SHARMA†

Abstract. The notion of positive realness for linear time-invariant (LTI) dynamical systems, equivalent to passivity, is one of the oldest in system and control theory. In this paper, we consider the problem of finding the nearest positive real (PR) system to a non-PR system: given an LTI control system defined by $E\dot{x} = Ax + Bu$ and $y = Cx + Du$, minimize the Frobenius norm of $(\Delta_E, \Delta_A, \Delta_B, \Delta_C, \Delta_D)$ such that $(E + \Delta_E, A + \Delta_A, B + \Delta_B, C + \Delta_C, D + \Delta_D)$ is a PR system. We first show that a system is extended strictly PR if and only if it can be written as a strict port-Hamiltonian system. This allows us to reformulate the nearest PR system problem into an optimization problem with a simple convex feasible set. We then use a fast gradient method to obtain a nearby PR system to a given non-PR system and illustrate the behavior of our algorithm with several examples. This is, to the best of our knowledge, the first algorithm that computes a nearby PR system to a given non-PR system that (i) is not based on the spectral properties of related Hamiltonian matrices or pencils, (ii) allows one to perturb all matrices (E, A, B, C, D) describing the system, and (iii) does not make any assumption on the original given system.

Key words. distance to positive realness, passivity, port-Hamiltonian system, fast gradient method

AMS subject classifications. 93B40, 93E24, 34A30, 65K05, 93D99

DOI. 10.1137/17M1137176

1. Introduction.

In this paper, we consider m-input m-output linear time-invariant (LTI) control systems of the form

$$\begin{aligned} E\dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{1.1}$$

on the unbounded interval $t \in [t_0, \infty)$. Here, $A, E \in \mathbb{R}^{n,n}$, $B \in \mathbb{R}^{n,m}$, $C \in \mathbb{R}^{m,n}$, and $D \in \mathbb{R}^{m,m}$ are given matrices, $x(t)$ is the vector of state variables, $u(t)$ is the vector of inputs, and $y(t)$ is the vector of outputs. The linear system is called a *standard system* when $E = I_n$, where I_n is the identity matrix of size $n \times n$, and a *descriptor system* when E is not invertible. We use the matrix quintuple (E, A, B, C, D) to refer to a system in the form (1.1).

As mentioned in [13], the restriction to systems (1.1) with the same number of inputs and outputs is necessary to have positive real (PR) systems, which are the focus of this paper. Indeed, positive realness of an LTI dynamical system is equivalent to passivity, which means that the system does not generate energy: the system (1.1) is called *passive* if there exists a nonnegative scalar valued function $\mathcal{V}$, called the storage function, such that $\mathcal{V}(0) = 0$ and the dissipation inequality

$$\mathcal{V}(x(t_1)) - \mathcal{V}(x(t_0)) \leq \int_{t_0}^{t_1} y(t)^T u(t) dt \tag{1.2}$$

*Received by the editors July 5, 2017; accepted for publication (in revised form) February 13, 2018; published electronically April 17, 2018.

http://www.siam.org/journals/sinum/56-2/M113717.html

Funding: The authors were supported by the ERC (starting grant 679515). The first author also was supported by the F.R.S.-FNRS (incentive grant for scientific research F.4501.16).

†Department of Mathematics and Operational Research, Faculté Polytechnique, Université de Mons, 7000 Mons, Belgium (nicolas.gillis@umons.ac.be, punit1718@gmail.com).

holds for all admissible u, t_0, and $t_1 \geq t_0$; see, for example, [2, 28]. The energy is defined via the inner product of the input and output vectors $u(t)$ and $y(t)$ of the system hence these vectors need to be of the same length.

Given (E, A, B, C, D), the goal of this paper is to find $(\Delta_E, \Delta_A, \Delta_B, \Delta_C, \Delta_D)$ with minimum (weighted) Frobenius norm[1] such that $(E+\Delta_E, A+\Delta_A, B+\Delta_B, C+\Delta_C, D+\Delta_D)$ is a PR system. We will refer to this problem as the *nearest PR system* problem. We will also consider the case of nearest standard PR systems when $E = I_n$ imposing $\Delta_E = 0$. Since passivity and positive realness are equivalent for LTI systems, the distance to positive realness has direct applications in passive model approximations. For example, when a real-world problem is approximated by a model (1.1), the passivity of the physical system may not be preserved, that is, the approximation process (for example, finite element or finite difference models, linearization, or model order reduction) makes the passive system nonpassive. The nonpassive system has to be approximated by a nearby passive system by perturbing E, A, B, C, and D.

Several algorithms tackle this problem using the spectral properties of the related Hamiltonian/skew-Hamiltonian matrices or pencils for the input systems that are asymptotically stable, controllable, observable, and almost passive; see [19, 38, 51, 49] and the references therein. The algorithms in [19] and [38] impose additional assumptions on the input system, namely, $E = I_n$ and $D + D^T$ nonsingular, and are restricted to perturbations of the matrix C only. These algorithms are based on the displacement of the imaginary eigenvalues of the related Hamiltonian matrix. The methods in [51] and [49] can deal with general systems (i.e., when E is not identity) by using the spectral properties of (skew-)Hamiltonian pencils, but they only allow perturbations in either B or C. In [9], authors allow perturbations in all matrices but E and assume that the system is almost passive. Their approach is based on the perturbation of a general nondissipative system to enforce dissipativity using first-order spectral perturbation results for para-Hermitian pencils. As far as we know, no algorithm exists that does not make any assumption on the input system and that allows perturbations of all matrices (E, A, B, C, D) describing the system.

The nearest PR system problem is complementary with the distance to nonpassivity for control systems; see [36] for complex standard systems. These problems are closely related to the Hamiltonian matrix nearness problems [1, 20]. For example, it is well known [8, 3] that an asymptotically stable standard system (1.1) (i.e., with $E = I_n$) with positive definite $D + D^T$ is PR if and only if the Hamiltonian matrix

$$\mathcal{H} = \begin{bmatrix} A - B(D+D^T)^{-1}C & -B(D+D^T)^{-1}B^T \\ C^T(D+D^T)^{-1}C & -(A - B(D+D^T)^{-1}C)^T \end{bmatrix} \tag{1.3}$$

has no purely imaginary eigenvalues. Therefore one can use the minimal perturbations from [1] and [20] that moves all eigenvalues of $\mathcal{H}$ away from the imaginary axis and find (if they exist) the corresponding perturbations $(\Delta_A, \Delta_B, \Delta_C, \Delta_D)$ making the system passive. The latter step, however, is in general not possible as it involves dealing with the additional block structure in the Hamiltonian matrix hence solving highly nonlinear matrix equations [20]. It is an open problem to extract system

[1]The choice for the (weighted) Frobenius norm is twofold: (i) it is arguably one of the most popular norms used to measure distances, and (ii) it is smooth and hence will make the optimization problem easier to tackle. However, our algorithm can easily be extended to any other smooth objective function, e.g., any (weighted) ℓ_p norm with $1 < p < +\infty$ (only the computation of the gradient of the corresponding objective function will change).

matrices (A, B, C, D) from a given Hamiltonian matrix that has no purely imaginary eigenvalues, that is, to express the Hamiltonian matrix as a matrix of the form (1.3).

In this paper, we compute a nearby PR system to a given non-PR system using the set of linear port-Hamiltonian (PH) systems. Our algorithm is based on the generalization of the results from [17] and [16], where authors used the structure of PH systems to find a nearby stable standard system and a nearby stable descriptor system to an unstable one, respectively. As opposed to the previously proposed methods, our algorithm is not based on the spectral properties of Hamiltonian matrices or pencils and can be applied to any given LTI dynamical system.

The paper is organized as follows. In section 2, we introduce the notation and definitions that will be used throughout the paper. In section 3, we show that a system is extended strictly PR if and only if it can be written as a strict PH system (Theorem 3.9). This allows us to reformulate, in section 4, the nearest PR system problem into an optimization problem with a simple convex feasible set. In section 5, we use a fast gradient method to tackle our reformulation and obtain a nearby PR system to a given non-PR system for both standard and general systems. We also propose several initialization strategies. The behavior of the algorithm is analyzed on several examples in section 6.

2. Notation, preliminaries, and problem definition. In the following, we denote by $\|\cdot\|_F$ the Frobenius norm and by * the complex conjugate transpose. We write $A \succ 0$ (resp., $A \succeq 0$) if A is symmetric positive definite (resp., semidefinite). The real part of $s \in \mathbb{C}$ is denoted by $\operatorname{Re} s$ and j stands for the imaginary number.

In the next two subsections, we define admissible and PR systems (sections 2.1 and 2.2). This allows us to give a formal definition of the nearest PR system problem in section 2.3. In section 2.4, we briefly describe PH systems that will be our main tool to reformulate the nearest PR system problem.

2.1. Admissible systems. The system (1.1) is called *regular* if the matrix pair (E, A) is regular, that is, if $\det(\lambda E - A) \neq 0$ for some $\lambda \in \mathbb{C}$; otherwise it is called *singular*. For a regular matrix pair (E, A), the roots of the polynomial $\det(zE - A)$ are called *finite eigenvalues* of the pencil $zE - A$ or of the pair (E, A). A regular pencil $zE - A$ has ∞ *as an eigenvalue* if E is singular.

A regular real matrix pair (E, A) (with $E, A \in \mathbb{R}^{n,n}$) can be transformed to *Weierstraß canonical form* [14], that is, there exist nonsingular matrices $W, T \in \mathbb{C}^{n,n}$ such that

$$E = W \begin{bmatrix} I_q & 0 \\ 0 & N \end{bmatrix} T \quad \text{and} \quad A = W \begin{bmatrix} J & 0 \\ 0 & I_{n-q} \end{bmatrix} T,$$

where $J \in \mathbb{C}^{q,q}$ is a matrix in *Jordan canonical form* associated with the q finite eigenvalues of the pencil $zE - A$ and $N \in \mathbb{C}^{n-q,n-q}$ is a nilpotent matrix in Jordan canonical form corresponding to $n - q$ times the eigenvalue ∞. If $q < n$ and N has degree of nilpotency $\nu \in \{1, 2, \ldots\}$, that is, $N^\nu = 0$ and $N^i \neq 0$ for $i = 1, \ldots, \nu - 1$, then ν is called the *index of the pair* (E, A). If E is nonsingular, then by convention the index is $\nu = 0$; see, for example, [32, 48]. The index ν does not depend on the transformation to canonical form [27, Lemma 2.10].

The matrix pair $(E, A) \in (\mathbb{R}^{n,n})^2$ is said to be *stable* (resp., *asymptotically stable*) if all the finite eigenvalues of $zE - A$ are in the closed (resp., open) left half of the complex plane and those on the imaginary axis are semisimple. The matrix pair (E, A) is said to be *admissible* if it is regular, of index at most one, and asymptotically stable. A dynamical system (E, A, B, C, D) in the form (1.1) is called (asymptotically) stable

if the matrix pair (E, A) is (asymptotically) stable. Similarly, it is called admissible if the matrix pair (E, A) is admissible.

2.2. Positive real systems. To define PR systems, throughout this section we assume that the system (1.1) is regular. The system (1.1) can be described by its *transfer function* $G(s) : \mathbb{C} \to (\mathbb{C} \cup \{\infty\})^{m,m}$, defined by

$$G(s) := C(sE - A)^{-1}B + D, \quad s \in \mathbb{C}. \tag{2.1}$$

Conversely, given a rational function $G(s) : \mathbb{C} \to (\mathbb{C} \cup \{\infty\})^{m,m}$, any representation of $G(s)$ in the form (2.1) is called a realization of $G(s)$. A realization is called minimal if the matrices A and E are of the smallest possible dimension. In this case the poles of the transfer function $G(s)$ are exactly the eigenvalues of the pencil $zE - A$.

Positive realness is a well-known concept in system, circuit, and control theory. In control theory, the PR systems have a significant role in stability analysis [37, 2]; see also [26] and the references therein for applications. The PR systems have been defined in several different ways in the literature; see [2, 52, 29, 39, 25, 21, 23] for standard linear systems, [50, 13, 28] for continuous-time descriptor systems, and [54] for continuous- and discrete-time descriptor systems. We follow [39] and define the positive realness in the frequency domain as follows.

DEFINITION 2.1. *The system* (1.1) *is said to be*

(1) PR *if its transfer function* $G(s)$ *satisfies*
   (a) $G(s)$ *has no pole in* $\operatorname{Re} s > 0$ *and*
   (b) $G(s) + G(s)^* \succeq 0$ *for all* s *such that* $\operatorname{Re} s > 0$*;*
(2) strictly positive real (SPR) *if its transfer function* $G(s)$ *satisfies*
   (a) $G(s)$ *has no pole in* $\operatorname{Re} s \geq 0$ *and*
   (b) $G(jw) + G(jw)^* \succ 0$ *for* $w \in [0, \infty)$*;*
(3) extended strictly positive real (ESPR) *if it is SPR and* $G(j\infty) + G(j\infty)^* \succ 0$.

Note that condition (a) in the definition of SPR is equivalent to the system being asymptotically stable. An asymptotically stable system (1.1) with a minimal realization is passive (resp., strictly input passive) if and only if it is PR (resp., ESPR). For more details of these facts, we refer to [2] and [12, pp. 174–175]. Furthermore, ESPR implies SPR, which further implies PR.

Note also that $G(s) = C(sE - A)^{-1}B + D$ is a rational function and has a power series expansion about $s = \infty$ of the form

$$G(s) = C(sE - A)^{-1}B + D = \sum_{i=-p}^{\infty} \frac{H_i}{s^i}, \tag{2.2}$$

where H_i are real matrices of size m. If $s = \infty$ is not a pole of $G(s)$ (i.e., when E is invertible), then $p = 0$ and $G(\infty) = D = H_0$. This implies that for a standard system (I_n, A, B, C, D) with $D + D^T \succ 0$, the notions of SPR and ESPR are the same, because $G(j\infty) + G(j\infty)^* \succ 0$ if and only if $D + D^T \succ 0$. In the descriptor case (i.e., when E is not invertible), then the order of the pole at $s = \infty$ is greater than or equal to one (i.e., $p \geq 1$ in (2.2)). In this case, $G(\infty)$ (if it exists) is not necessarily equal to D. To illustrate this consider the system

(2.3)
$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ \alpha \end{bmatrix}, C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^T, D = \frac{1}{2},$$

where α is a constant. This is an admissible system with the transfer function

$$G(s) = \frac{1}{s+1} - \alpha - \frac{1}{2},$$

so that

$$G(jw) + G(jw)^* = \frac{2}{w^2+1} - 2\alpha - 1.$$

Therefore, for $\alpha = -\frac{1}{2}$, the system is SPR but not ESPR despite $D \succ 0$.

Remark 1. We have chosen to define PR and ESPR systems in the frequency domain for regular descriptor systems and then used their LMI characterization to compute a nearby regular passive system to a given nonpassive one; see sections 3 and 4. However, it is possible to define passivity in the time domain for a general system with no assumptions such as regularity, controllability, and minimality [24]. This could be a direction for further research to use the results in [24] to define passivity in the time domain and use their equivalent LMI characterization to find a nearby passive system under fewer assumptions. This could also be a way to avoid the regularity assumption on PR systems.

2.3. Nearest positive-real system problems. We can now define the nearest system problems that will be studied in the following sections. Let us formulate the problem in a rather generic way.

PROBLEM. *For a given system (E, A, B, C, D) and a given set $\mathcal{D}$, find the nearest system $(\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}) \in \mathcal{D}$ to (E, A, B, C, D), that is, solve*

$$\inf_{(\tilde{E},\tilde{A},\tilde{B},\tilde{C},\tilde{D})\in\mathcal{D}} \mathcal{F}(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{E}),$$

where

(2.4)
$$\mathcal{F}(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}, \tilde{E}) = \|A - \tilde{A}\|_F^2 + \|B - \tilde{B}\|_F^2 + \|C - \tilde{C}\|_F^2 + \|D - \tilde{D}\|_F^2 + \|E - \tilde{E}\|_F^2.$$

We will consider the following two variants of this problem:

(1) Nearest PR system $(\mathcal{P})$: $\mathcal{D} = \mathbb{S}$, where $\mathbb{S}$ is the set of all PR systems $(\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$.
(2) Nearest ESPR system $(\mathcal{P}_e)$: $\mathcal{D} = \mathbb{S}_e$, where $\mathbb{S}_e$ is the set of all admissible ESPR systems $(\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ with $\tilde{D} + \tilde{D}^T \succ 0$.

We will also consider the variants of $(\mathcal{P})$ and $(\mathcal{P}_e)$ for standard systems with the additional constraints that $\tilde{E} = E = I_n$.

These problems are challenging because the feasible sets $\mathbb{S}$ and $\mathbb{S}_e$ are unbounded, highly nonconvex, and neither open nor closed. To see this, let us consider the system from (2.3). This system is ESPR with $\alpha = -1$, and thus $(E, A, B, C, D) \in \mathbb{S}_e \subseteq \mathbb{S}$. For

$$\Delta_E = \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ and } \Delta_A = \begin{bmatrix} \delta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

the transfer function of the perturbed system $(E + \Delta_E, A + \Delta_A, B, C, D)$ is given by

$$G(s) = \frac{1}{s(1+\epsilon_1) + (1-\delta)} + \frac{1}{s\epsilon_2 - 1} - \alpha + \frac{1}{2},$$

hence

$$G(jw)+G(jw)^*=\frac{2(1-\delta)}{(1-\delta)^2+w^2(1+\epsilon_1)^2}-\frac{2}{1+w^2\epsilon_2^2}-2\alpha+1.$$

For $\delta=\epsilon_1=0$ and $\epsilon_2>0$, the perturbed system is regular, of index one, but has two finite eigenvalues $\lambda_1=-1$ and $\lambda_2=\frac{1}{\epsilon_2}>0$. This implies that the system is not stable, hence not PR. This shows that $\mathbb{S}$ and $\mathbb{S}_e$ are not open. For $\epsilon_1=-\delta$ with $0\le\delta<1$ and $\epsilon_2=0$, the perturbed system is ESPR. The perturbed matrix pair $(E+\Delta_E, A+\Delta_A)$ becomes singular as $\delta\to 1$ so that the perturbed system becomes non-PR as $\delta\to 1$. This shows that $\mathbb{S}$ and $\mathbb{S}_e$ are not closed. Further the sets $\mathbb{S}$ and $\mathbb{S}_e$ are nonconvex: considering the systems

$$\Sigma_1=\left(I_2,\underbrace{\begin{bmatrix}-0.3 & 10\\ 0 & -0.3\end{bmatrix}}_{A_1},\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix}^T,\frac{1}{2}\right)\quad\text{and}$$

$$\Sigma_2=\left(I_2,\underbrace{\begin{bmatrix}-0.3 & 0\\ 10 & -0.3\end{bmatrix}}_{A_2},\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}^T,\frac{1}{2}\right),$$

it is easy to check that $\Sigma_1,\Sigma_2\in\mathbb{S}_e\subset\mathbb{S}$ but $\gamma\Sigma_1+(1-\gamma)\Sigma_2\notin\mathbb{S}$ for $\gamma=\frac{1}{2}$ since $\frac{1}{2}A_1+\frac{1}{2}A_2$ has an eigenvalue $\lambda=4.7$ in the right half complex plane.

To address the challenging problems defined above, we aim to reformulate them so that it is easier to derive optimization algorithms. An important property of such a reformulation is that the projection onto the feasible set can be performed efficiently. Such a reformulation exists and can be obtained by extending the results from [17] (resp., [16]) that used PH systems for computing the nearest stable matrix (resp., matrix pair).

2.4. Port-Hamiltonian systems. An LTI input-state-output PH system can be written as

$$\begin{aligned}E\dot{x}(t)&=(J-R)Qx(t)+(F-P)u(t),\\ y(t)&=(F+P)^TQx(t)+(S+N)u(t),\end{aligned}\tag{2.5}$$

where the following conditions must be satisfied:

- The matrix $Q\in\mathbb{R}^{n,n}$ is invertible, $E\in\mathbb{R}^{n,n}$, and $Q^TE=E^TQ\succeq 0$. The function $x\to\frac{1}{2}x^TQ^TEx$ is the *Hamiltonian* and describes the energy of the system.
- The matrix $J^T=-J\in\mathbb{R}^{n,n}$ is the structure matrix that describes flux among energy storage elements.
- The matrix $R\in\mathbb{R}^{n,n}$ with $R\succeq 0$ is the dissipation matrix and describes the energy dissipation/loss in the system.
- The matrices $F\pm P\in\mathbb{R}^{n,m}$ are the port matrices describing the manner in which energy enters and exits the system.
- The matrix $S+N$, with $0\preceq S\in\mathbb{R}^{m,m}$ and $N^T=-N\in\mathbb{R}^{m,m}$, describes the direct feed-through from input to output.
- The matrices R, P, and S satisfy

$$K=\begin{bmatrix}R & P\\ P^T & S\end{bmatrix}\succeq 0.$$

In the following, we will refer to K as the *cost matrix* of the PH system, because it corresponds to the cost matrix of an infinite horizon linear quadratic optimal control problem. For $K \succ 0$, we refer to (2.5) as a *strict PH system*. We note that this definition of PH system is slightly more restrictive than that of PH systems in [5], where it is not required for the matrix Q to be invertible.

PH systems generalize the classical Hamiltonian systems and recently have received a lot attention in energy-based modeling; see [42, 43, 45, 46, 47, 44] for some major references. The Hamiltonian $\mathcal{H}(x) = \frac{1}{2}x^T Q^T E x$ defines a storage function associated with the supply rate $y(t)^T u(t)$ and satisfies

$$\mathcal{H}(x(t_1)) - \mathcal{H}(x(t_0)) \leq \int_{t_0}^{t_1} y(t)^T u(t) dt, \tag{2.6}$$

which guarantees the passivity of the system; see (1.2).

We note that regular PH systems are always stable (see [30, Lemma 3.1] for standard PH systems and [16, Lemma 2] for descriptor PH systems): the matrix pair $(E, (J-R)Q)$ is a so-called dissipative Hamiltonian matrix pair. In particular, if $R \succ 0$, then $(E, (J-R)Q)$ is admissible; see [16] for more details and [30] and [31] for various structured distances to asymptotic stability for complex PH systems and real PH systems, respectively.

3. Key results for positive real systems. In this section, we study the link between PR systems and PH systems. The main result of this section, which is the main theoretical result of this paper, is to prove that a system is ESPR with $D + D^T \succ 0$ if and only if it can be written as a strict PH system; see Theorem 3.9 at the end of the section.

The positive realness of a system (1.1) can be characterized in terms of solutions X to the following linear matrix inequalities (LMIs):

$$\begin{bmatrix} A^T X + X^T A & X^T B - C^T \\ B^T X - C & -D - D^T \end{bmatrix} \preceq 0 \quad \text{and} \quad E^T X = X^T E \succeq 0. \tag{3.1}$$

THEOREM 3.1 (see [13, Theorem 3.1]). *Consider a regular system (E, A, B, C, D) in the form* (1.1)*. If the LMIs* (3.1) *have a solution $X \in \mathbb{R}^{n,n}$, then (E, A, B, C, D) is PR.*

The converse of Theorem 3.1 is true with some additional assumptions. In fact the PR lemma for standard systems [2] proves that if a system is PR and minimal, then the existence of a solution to the LMIs (3.1) is also necessary. Similarly, with an additional condition, the PR lemma for descriptor systems [13] proves that the existence of a solution to the LMIs (3.1) is also necessary for positive realness.

Theorem 3.1 gives an alternative way (compared to the one described in (2.6)) to show that every PH system is PR by providing an explicit solution to (3.1), as shown in the following theorem, which is a generalization of [44, Theorem 7.1], where only standard systems are considered.

THEOREM 3.2. *Every regular PH system in the form* (2.5) *is PR.*

Proof. Let (E, A, B, C, D) be a regular PH system with $A = (J-R)Q$, $B = F-P$, $C = (F+P)^T Q$, and $D = S + N$, where $J^T = -J$, $N^T = -N$, $\left[\begin{smallmatrix} R & P \\ P^T & S \end{smallmatrix}\right] \succeq 0$, Q is invertible, and $E^T Q = Q^T E \succeq 0$. By Theorem 3.1, to prove that this system is PR, it suffices to prove the existence of a solution X to the LMIs (3.1). It turns out that

$X = Q$ is one. In fact, we have

$$\begin{aligned}
&\begin{bmatrix} A^TQ + Q^TA & Q^TB - C^T \\ B^TQ - C & -D - D^T \end{bmatrix} \\
&= \begin{bmatrix} ((J-R)Q)^TQ + Q^T(J-R)Q & Q^T(F-P) - ((F+P)^TQ)^T \\ (F-P)^TQ - (F+P)^TQ & -(S+N) - (S+N)^T \end{bmatrix} \\
&= -2\begin{bmatrix} Q^TRQ & Q^TP \\ P^TQ & S \end{bmatrix} = -2\begin{bmatrix} Q^T & 0 \\ 0 & I_m \end{bmatrix}\begin{bmatrix} R & P \\ P^T & S \end{bmatrix}\begin{bmatrix} Q & 0 \\ 0 & I_m \end{bmatrix} \preceq 0
\end{aligned}$$

since $\left[\begin{smallmatrix} R & P \\ P^T & S \end{smallmatrix}\right] \succeq 0$. ∎

Note that the proof of Theorem 3.2 does not require Q to be invertible. In the following, a necessary and sufficient condition for a system in the form (1.1) to be ESPR is obtained in terms of the existence of a solution of the LMIs (3.1). We will use this result to characterize the set of all admissible ESPR systems in terms of PH systems.

THEOREM 3.3 (see [54, Theorem 2]). *Let (E, A, B, C, D) define a system (1.1). Then it is admissible, ESPR, and satisfying $D + D^T \succ 0$ if and only if there exists a solution X to the LMIs*

$$\begin{bmatrix} A^TX + X^TA & X^TB - C^T \\ B^TX - C & -D - D^T \end{bmatrix} \prec 0 \quad \text{and} \quad E^TX = X^TE \succeq 0. \tag{3.2}$$

Using Theorems 3.2 and 3.3, we prove the following result.

THEOREM 3.4. *Every PH system in the form (2.5) with a positive definite cost matrix is admissible and ESPR.*

Proof. Let $\Sigma = (E, (J-R)Q, (F-P), (F+P)^TQ, S+N)$ be a PH system in the form (2.5) with the cost matrix $K = \left[\begin{smallmatrix} R & P \\ P^T & S \end{smallmatrix}\right] \succ 0$. Following the same steps as in Theorem 3.2, one can show that $X = Q$ satisfies the LMIs in (3.2) because $K \succ 0$ and Q is invertible. Hence, by Theorem 3.3, Σ is admissible and ESPR. ∎

We note that the admissibility of the PH system Σ in Theorem 3.4 also follows by the fact that $(E, (J-R)Q)$ is a dissipative Hamiltonian (DH) matrix pair with $R \succ 0$ [16, Theorem 4]. We recall from [16, Definition 2] that (E, A) is called a DH matrix pair if there exists an invertible matrix Q, $J^T = -J$ and $R \succeq 0$ such that $Q^TE \succeq 0$ and $A = (J-R)Q$. In order to show that the converse of Theorem 3.4 is also true, we define the *PH-form* for a system (1.1).

DEFINITION 3.5. *A system (E, A, B, C, D) is said to admit a PH-form if there exists a PH system as defined in (2.5) such that*

$$A = (J-R)Q, \quad B = F - P, \quad C = (F+P)^TQ, \quad \text{and} \quad D = S + N.$$

In view of Theorem 3.4, if (E, A, B, C, D) admits a PH-form with positive definite cost matrix, then it is admissible and ESPR. Similarly, by Theorem 3.2, it follows that every regular system (E, A, B, C, D) that admits a PH-form, is PR. However, the converse, that is, every PR system admits a PH-form, is not true, as there exist PR systems with $D + D^T \prec 0$, for instance, (2.3) with $\alpha = -\frac{3}{2}$ and replacing $D = -1/2$.

We now show that whenever the LMIs (3.1) have an invertible solution X, the system (E, A, B, C, D) admits a PH-form. The main part of its proof is identical to [44, Theorem 7.1], which showed a similar result for standard systems. Unlike [44], we assume X is invertible, which allows us to give an explicit PH-form.

THEOREM 3.6. *Let $\Sigma = (E, A, B, C, D)$ be a system in the form (1.1). If the LMIs (3.1) have an invertible solution $X \in \mathbb{R}^{n,n}$, then Σ admits a PH-form.*

Proof. Let X be an invertible solution of the LMIs (3.1). Define

$$J := \frac{AX^{-1} - (AX^{-1})^T}{2}, \quad R := -\frac{AX^{-1} + (AX^{-1})^T}{2}, \quad Q := X, \quad S := \frac{1}{2}(D+D^T),$$

$$N := \frac{1}{2}(D - D^T), \quad F := \frac{1}{2}(B + X^{-1}C^T), \quad \text{and} \quad P := \frac{1}{2}(-B + X^{-1}C^T). \tag{3.3}$$

Let us show that the matrices J, R, Q, F, P, N, and S provide a PH-form for Σ. We have

$$(J - R)Q = A, \quad F - P = B, \quad (F + P)^T Q = C, \quad \text{and} \quad S + N = D.$$

Further, we have that $E^T Q \succeq 0$ (using the second LMI in (1.1)), $J^T = -J$, $N^T = -N$, and

$$\begin{aligned} K = \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} &= -\frac{1}{2}\begin{bmatrix} AX^{-1} + X^{-1}A^T & -B + X^{-1}C^T \\ -B^T + CX^{-1} & -D - D^T \end{bmatrix} \\ &= -\frac{1}{2}\begin{bmatrix} -X^{-1} & 0 \\ 0 & I_m \end{bmatrix}\begin{bmatrix} A^T X + XA & XB - C^T \\ B^T X - C & -D - D^T \end{bmatrix}\begin{bmatrix} -X^{-1} & 0 \\ 0 & I_m \end{bmatrix} \succeq 0, \end{aligned}$$

which follows from the first LMI in (3.1). ∎

For standard systems, [4, Corollary 2] shows that every minimal PR system (I_n, A, B, C, D) is equivalent to a system in PH-form, that is, there exist invertible matrices $T \in \mathbb{R}^{n,n}$ and $V \in \mathbb{R}^{m,m}$ such that the transformed system

$$(I_n, T^{-1}AT, T^{-1}BV, V^T CT, V^T DV)$$

admits a PH-form. Theorem 3.6 implies a stronger result: a minimal PR standard system itself admits a PH-form (no transformation is necessary).

COROLLARY 3.7. *If the system (I_n, A, B, C, D) is minimal and PR, then it admits a PH-form.*

Proof. This follows from the PR lemma for minimal PR standard systems [53, p. 363] (which guarantees the existence of an invertible solution X of the LMIs (3.1)) and Theorem 3.6. ∎

We note that as opposed to a standard PR system [53, p. 363], minimality of a PR system does not guarantee the solvability of the LMIs (3.1) in the descriptor case. For this to hold, an additional condition that $D + D^T \succeq \lim_{s\to\infty} \left(G(s) + G(s)^T\right)$ is also needed [13, Theorem 3.2].

COROLLARY 3.8. *Every admissible and ESPR dynamical system (E, A, B, C, D) with $D + D^T \succ 0$ admits a PH-form with positive definite cost matrix.*

Proof. By Theorem 3.3, there exists a solution X to the LMIs

$$\begin{bmatrix} A^T X + X^T A & X^T B - C^T \\ B^T X - C & -D - D^T \end{bmatrix} \prec 0 \quad \text{and} \quad E^T X = X^T E \succeq 0. \tag{3.4}$$

This implies that $A^T X + X^T A \prec 0$, and therefore X is invertible. The remainder of the proof follows using the same arguments as that of Theorem 3.6 with the solution X of the LMIs (3.4). ∎

In the following, we summarize several equivalent characterizations of a system to be admissible and ESPR.

THEOREM 3.9. *Let $\Sigma = (E, A, B, C, D)$ be a system in the form* (1.1)*. Then the following are equivalent:*

(1) *Σ is admissible and ESPR with $D + D^T \succ 0$.*

(2) *There exists a solution X to the LMIs*

$$\begin{bmatrix} A^T X + X^T A & X^T B - C^T \\ B^T X - C & -D - D^T \end{bmatrix} \prec 0 \quad \textit{and} \quad E^T X = X^T E \succeq 0.$$

(3) *Σ admits a PH-form with positive definite cost matrix.*

Proof. This follows from Theorems 3.3 and 3.4 and Corollary 3.8. ∎

In the next section, we reformulate the nearest ESPR system problem $(\mathcal{P}_e)$ using the PH-form for an admissible ESPR system (E, A, B, C, D) with $D + D^T \succ 0$. As mentioned in section 2.2, for a standard ESPR system (I_n, A, B, C, D) we have that $D + D^T \succ 0$; thus the condition $D + D^T \succ 0$ for standard systems is redundant. However, the PH-form characterization of an admissible ESPR descriptor system depends on the existence of a solution of the LMIs (3.2) when $D + D^T \succ 0$. This justifies the restriction $D + D^T \succ 0$ on defining the set $\mathbb{S}_e$ for the nearest ESPR system problem in section 2.3.

4. Reformulation of the nearest PR system problems. In this section, we exploit the results obtained in the previous section and present a new framework based on PH systems to attack $(\mathcal{P})$ and $(\mathcal{P}_e)$ defined in section 2.3, as well as their variants for standard systems.

Let us define the following two sets:

- The set $\mathbb{S}_{PH}$ containing all systems (E, A, B, C, D) in PH-form, that is,

$$\begin{aligned} \mathbb{S}_{PH} &:= \{(E, A, B, C, D) \mid (E, A, B, C, D) \text{ admits a PH-form}\} \\ &= \left\{ (E, (J - R)Q, F - P, (F + P)^T Q, S + N) \;\middle|\; J^T = -J, N^T = -N, \right. \\ &\qquad \left. E^T Q \succeq 0, Q \text{ invertible}, K = \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0 \right\}. \end{aligned}$$

- The set $S_{PH}^{\succ 0} \subset \mathbb{S}_{PH}$ containing all systems (E, A, B, C, D) in strict PH-form, that is,

$$\mathbb{S}_{PH}^{\succ 0} := \left\{ (E, (J - R)Q, F - P, (F + P)^T Q, S + N) \in \mathbb{S}_{PH} \;\middle|\; K \succ 0 \right\}.$$

By Theorem 3.9, $\mathbb{S}_e = \mathbb{S}_{PH}^{\succ 0}$.

We have discussed in section 3 that the set $\mathbb{S}_e$ of all ESPR systems is neither open nor closed and clearly the PH characterization of $\mathbb{S}_e$ does not change this. In fact, the sets $\mathbb{S}_{PH}$ and $\mathbb{S}_{PH}^{\succ 0}$ are neither closed (due to the constraint that Q is invertible) nor open (due to the constraint $E^T Q \succeq 0$).

Since we want to work with a set onto which it is easy (and possible) to project, we consider the closure $\overline{\mathbb{S}_{PH}}$ of $\mathbb{S}_{PH}$ which is equal to the set $\mathbb{S}_{PH}$ except that Q can be singular. Moreover, we have that $\overline{\mathbb{S}_{PH}} = \overline{\mathbb{S}_{PH}^{\succ 0}}$. Therefore the values of the infimum over the sets $\mathbb{S}_{PH}$, $\mathbb{S}_{PH}^{\succ 0}$, and $\overline{\mathbb{S}_{PH}}$ are the same. We have the following result.

THEOREM 4.1. *Let* (E, A, B, C, D) *be a system in the form* (1.1) *and* $\mathcal{F}$ *be defined as in* (2.4). *Then*

$$\inf_{(M,(J-R)Q,F-P,(F+P)^TQ,S+N)\in\overline{\mathbb{S}_{PH}}} \mathcal{F}((J-R)Q, F-P, (F+P)^TQ, S+N, M) \tag{4.1}$$

coincides with the infimum of $(\mathcal{P}_e)$ *while it is an upper bound for the infimum of* $(\mathcal{P})$.

Proof. This follows directly from the fact that $\mathbb{S}_e = \mathbb{S}_{PH}^{\succ 0}$ and $\mathbb{S}_e \subseteq \mathbb{S}$. ∎

We will refer to (4.1) as the nearest PH system problem. The same result holds for the variants of $(\mathcal{P})$ and $(\mathcal{P}_e)$ for standard systems since the only difference is that M is imposed to be equal to $E = I_n$.

Remark 2. We note that for standard systems we have $\overline{\mathbb{S}_{PH}} \subseteq \mathbb{S}$; this is due to Theorem 3.2 as its proof does not require Q to be invertible. In the descriptor case, $\overline{\mathbb{S}_{PH}}$ could contain systems which are not regular. This shows that in this case a feasible solution of (4.1) may not be a PR system. To rule out such situations, one can impose the matrix R to satisfy $R \succeq \delta I_n$ for some fixed small $\delta > 0$, because in this case $(E, (J-R)Q)$ is a DH matrix pair with positive definite R and therefore the system is guaranteed to be regular by [16, Corollary 1]. This does not make the problem more complicated as the projection is still straightforward, but it gives a nearby regular descriptor PH system (hence a PR system; see Theorem 3.2) to a given system.

Remark 3. Although the value of the infimum in (4.1) coincides with the infimum of $(\mathcal{P}_e)$, the solution of (4.1) may not solve problem $(\mathcal{P}_e)$, as the solution found may not even be PR; see Remark 2. However, it is possible to obtain a nearby strict PH system (hence admissible and ESPR system with $D + D^T \succ 0$; see Theorem 3.9) by rewriting (4.1) using lower bounds on the eigenvalues of matrix E^TQ and K; see also Remark 6.

5. Algorithmic solution to the nearest PH system problem. In this section, we propose an algorithm to tackle (4.1). We analyze separately standard systems when $E = I_n$ and E is not subject to perturbation and general systems when E is subject to perturbation.

5.1. Standard systems. For standard systems, $M = E = I_n$ and (4.1) can be simplified as follows:

$$\inf_{J,R,Q,F,P,S,N} \|A-(J-R)Q\|_F^2 + \|B-(F-P)\|_F^2 + \|C-(F+P)^TQ\|_F^2 + \|D-(S+N)\|_F^2$$
$$\text{such that} \quad J^T = -J, Q \succeq 0, N^T = -N \text{ and } \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0. \tag{5.1}$$

For any given square matrix Z and any skew-symmetric matrix X, we have

$$\|Z - X\|_F^2 = \left\| \frac{(Z+Z^T)}{2} + \frac{(Z-Z^T)}{2} - X \right\|_F^2 = \left\| \frac{(Z+Z^T)}{2} \right\|_F^2 + \left\| \frac{(Z-Z^T)}{2} - X \right\|_F^2 \tag{5.2}$$

since symmetric and skew-symmetric matrices are orthogonal (their inner product is zero). Therefore the infimum in (5.2) over all skew-symmetric matrices X is attained when $X = \frac{Z-Z^T}{2}$, that is,

$$\min_{X^T=-X} \|Z - X\|_F = \left\| Z - \frac{(Z-Z^T)}{2} \right\|_F. \tag{5.3}$$

This implies that the optimal N in (5.1) is given by $\frac{D-D^T}{2}$ since S is symmetric, so that (5.1) can be simplified to

(5.4)
$$\inf_{J,R,Q,F,P,S} \mathcal{G}(J,R,Q,F,P,S) \quad \text{such that} \quad J^T = -J, Q \succeq 0 \text{ and } \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0,$$

where

$$\mathcal{G}(J,R,Q,F,P,S) = \|A-(J-R)Q\|_F^2 + \|B-(F-P)\|_F^2 + \|C-(F+P)^TQ\|_F^2 + \left\| \frac{D+D^T}{2} - S \right\|_F^2.$$

Similarly as was done in [17] to find the nearest stable matrix to an unstable one, we develop a fast projected gradient method (FGM) to solve (5.4). FGM has the advantage of being in general much faster than the standard projected gradient method [33], [34, p. 90] (see section 6 for some examples), even in the nonconvex case [35], while being relatively simple to implement as long as one can do the following:

- Compute the gradient: all the terms in the objective function are of the form $f(X) = \|AX-B\|_F^2$ whose gradient is $\nabla_X f(X) = 2A^T(AX-B)$.
- Project onto the feasible set: the projection onto the set $\{X : X = -X^T\}$ is given in (5.3), while projection onto the set of positive semidefinite (PSD) matrices can be computed using an eigenvalue decomposition [22].

For the sake of completeness, we describe in Algorithm 1 the variant of FGM we use. Note that we have included a restarting procedure which is necessary in our case since the objective function is not convex and hence FGM without restart is not guaranteed to converge [15]. We have observed that, in most cases, Algorithm 1 does not need to restart very often (in the numerical examples presented in section 6, it restarts on average less than once every 1000 iterations). In our implementation, we have also added the possibility of giving different importance to each term in the objective function using nonnegative weights $w_i \geq 0$ $(1 \leq i \leq 4)$ and minimizing

(5.5)
$$w_1\|A-(J-R)Q\|_F^2 + w_2\|B-(F-P)\|_F^2 + w_3\|C-(F+P)^TQ\|_F^2 + w_4\left\| \frac{D+D^T}{2} - S \right\|_F^2.$$

Parameter settings. For the initial step length, we use $\gamma = 1/L$, where $L = \|Q\|_2^2$ is the Lipschitz constant of the gradients of $\mathcal{G}$ with respect to J (and R). The reason for this choice is that this step length would guarantee the decrease of the objective function if we would only update J (or R) since the subproblem in J (and R) is convex. Note that the Lipschitz constant of the gradients of $\mathcal{G}$ with respect to F (resp., S) is $L+1$ (resp., 1). Hence, except maybe for the variable Q, this value of γ has a good order of magnitude while being simple to compute. In fact, computing the Lipschitz constant of the full gradient of $\mathcal{G}$ is nontrivial and computationally more expensive, while the Hessian of $\mathcal{G}$ is mostly block diagonal (only the variable Q appears with other variables). We choose $\alpha_1 = 0.5$, which seems to work well in most cases, although FGM can be quite sensitive to this parameter and it is sometimes rewarding to try different values. In fact, even in the convex case, there is, as far as we know, no theoretical recommendation on how to choose this value; it is problem dependent. (We also tried $\alpha_1 = 0.1, 0.9$, which performed on average slightly worse than $\alpha_1 = 0.5$.)

Algorithm 1. FGM with restart.

Input: The (nonconvex) function $f(x)$, the feasible set $\mathcal{X}$, an initial guess $x \in \mathcal{X}$, a parameter $\alpha_1 \in (0,1)$, lower bound for the step length $\underline{\gamma}$.

Output: An approximate solution x to the problem $\operatorname{argmin}_{z\in\mathcal{X}} f(z)$.

$y = x$; initial step length $\gamma > \underline{\gamma}$.
for $k = 1, 2, \ldots$ **do**
  % Keep the previous iterate in memory.
  $\hat{x} = x$.
  % Project the gradient step onto $\mathcal{X}$, where $\mathcal{P}_{\mathcal{X}}(z) = \operatorname{argmin}_{\tilde{z}\in\mathcal{X}} \|z - \tilde{z}\|$.
  $x = \mathcal{P}_{\mathcal{X}}\big(y - \gamma\nabla f(y)\big)$.
  % Check if the objective function f has decreased, otherwise decrease the step length.
  while $f(x) > f(\hat{x})$ and $\gamma \geq \underline{\gamma}$ **do**
    $\gamma = \frac{2}{3}\gamma$.
    $x = \mathcal{P}_{\mathcal{X}}\big(y - \gamma\nabla f(y)\big)$.
  end while
  % If the step length has reached the lower bound (f could not be decreased from y), reinitialize y (the next step will be a regular gradient descent step).
  if $\gamma < \underline{\gamma}$ **then**
    Restart fast gradient: $y = x$; $\alpha_k = \alpha_1$.
    Reinitialize γ at the last value for which it allowed decrease of f.
  else
    $\alpha_{k+1} = \frac{1}{2}\left(\sqrt{\alpha_k^4 + 4\alpha_k^2} - \alpha_k^2\right)$, $\beta_k = \frac{\alpha_k(1-\alpha_k)}{\alpha_k^2+\alpha_{k+1}}$.
    $y = x + \beta_k\,(x - \hat{x})$.
  end if
  $\gamma = 2\gamma$.
end for

Remark 4 (closed form for F). The optimal solution for the variable F in (5.4) can be written in closed form,

$$\begin{aligned}\hat{F} &= \operatorname{argmin}_F \|B - (F - P)\|_F^2 + \|C^T - Q^T(F + P)\|_F^2 \\ &= (I_n + QQ^T)^{-1}\left(P + B + QC^T + QQ^TP\right),\end{aligned} \tag{5.6}$$

since

$$\begin{aligned}&\frac{1}{2}\nabla_F(\|F - (P + B)\|_F^2 + \|Q^TF - (C^T + Q^TP)\|_F^2) \\ &\quad = (F - (P + B)) + Q(Q^TF - (C^T + Q^TP)).\end{aligned}$$

However, we did not inject $\hat{F}$ in (5.4) as it makes the objective function very complicated, in particular because of the term $\|C^T - Q^T(\hat{F} + P)\|_F^2$.

5.2. General systems. Similarly as for standard systems in (5.4), (4.1) can be simplified to

$$\begin{aligned}\inf_{J,R,Q,M,F,P,S} &\mathcal{G}(J, R, Q, F, P, S) + \|E - M\|_F^2 \\ &\text{such that } J^T = -J, M^TQ \succeq 0 \text{ and } \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0.\end{aligned} \tag{5.7}$$

As opposed to (5.4), it is difficult to project on the feasible domain of (5.7) because of the coupling constraint $M^TQ \succeq 0$. Moreover, this constraint was observed to get standard optimization schemes stuck in suboptimal solutions; see [16, Example 3] for an example. Following the strategy used in [16], we introduce the variable $Z = M^TQ$ so that $M^T = ZQ^{-1}$. This allows us to reformulate (5.7) into an equivalent optimization problem with a simpler feasible set,

$$\begin{aligned} \inf_{J,R,Q,Z,F,P,S} & \mathcal{G}(J,R,Q,F,P,S) + \|E^T - ZQ^{-1}\|_F^2 \\ \text{such that } & J^T = -J, Z \succeq 0 \text{ and } \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0, \end{aligned} \tag{5.8}$$

for which we have implemented Algorithm 1 (the gradient of $\|E^T - ZQ^{-1}\|_F^2$ with respect to Q is given in [16, Appendix A]). As for standard systems, we have also added the possibility of using weights for the different terms in the objective function as in (5.5) adding the term $w_5\|E^T - ZQ^{-1}\|_F^2$ with $w_5 \geq 0$. For the initial step length, we use $\gamma = 1/L$, where $L = \max(\|Q\|_2^2, \|Q^{-1}\|_2^2)$ is the maximum between the Lipschitz constants of the gradients of $\mathcal{G}$ with respect to J, R, and Z.

5.3. Initializations. Since we are dealing with nonconvex optimization problems, it is expected that choosing good initial points will be crucial to obtain good solutions. This will be confirmed in the numerical experiments. In the next two subsections, we propose several initialization strategies. We believe that designing other initialization strategies is an important direction for further research, in particular, taking advantage of the particular structure of the problem at hand.

5.3.1. Standard initialization. The first initialization, which we refer to as "standard," uses $Q = I_n$ and $P = 0$. For these values of Q and P, the optimal solutions for the other variables can be computed explicitly:

$$J = \big(A - A^T\big)/2, R = \mathcal{P}_{\succeq}\big((-A - A^T)/2\big), S = \mathcal{P}_{\succeq}\big((D^T + D)/2\big), F = \big(B + C^T\big)/2,$$

and $Z = \mathcal{P}_{\succeq}(E^T)$ for general systems. The notation $\mathcal{P}_{\succeq}(X)$ stands for the projection of a matrix X on the cone of PSD matrices. This initialization has the advantage of being very simple to compute while working reasonably well in many cases; see section 6 for numerical experiments.

5.3.2. LMI-based initializations. Given a system that does not admit a PH-form, the LMIs (3.1) will not have a solution. However, since we are looking for a nearby system that will admit a solution to these LMIs, it makes sense to try to find a solution X to nearby LMIs. We propose the following to relax the LMIs (3.1):

$$\begin{aligned} \min_{\delta,X} \quad & \delta^2 \\ \text{such that} \quad & \begin{bmatrix} -A^TX - X^TA & C^T - X^TB \\ C - B^TX & D + D^T \end{bmatrix} + \delta I_{n+m} \succeq 0, \\ & E^TX + \delta I_n \succeq 0. \end{aligned} \tag{5.9}$$

Let us denote (δ^*, X^*) an optimal solution of (5.9). By Theorem 3.6, if $\delta^* = 0$ and X^* is invertible, then the system (E, A, B, C, D) admits a PH-form that can be constructed explicitly; see (3.3). Moreover, as long as X^* is invertible, the matrices (J, R, Q, S, N, P, Z) can be constructed using (3.3) and projected onto the feasible set

$\overline{\mathbb{S}_{PH}}$ to obtain an initial system in PH-form. We will refer to this initialization as "LMIs + formula." If one wants to obtain a better initial point, given $Q = X^*$, it is possible to compute the matrices (J, R, S, N, P) by solving a semidefinite program (SDP):

$$\min_{J,R,S,N,P} \mathcal{G}(J, R, Q, F, P, S) \quad \text{such that} \quad J^T = -J \text{ and } \begin{bmatrix} R & P \\ P^T & S \end{bmatrix} \succeq 0, \tag{5.10}$$

while taking $Z = \mathcal{P}_{\succeq}(E^T Q)$ (as $Q = X^*$ can be ill-conditioned). We will refer to this initialization as "LMIs + solve." By construction, it provides an initial point with smaller objective function value than LMIs + formula (at the cost of solving another SDP).

We will see that the LMI-based initializations work well when the initial system is close to being passive (that is, when δ^* is small); otherwise, it may provide rather bad initial points (see section 6 for some examples). However, in most applications, the systems of interest are usually close to being passive (cf. The introduction); hence we believe these initializations will be particularly useful in practice.

Remark 5 (finding the nearest stable matrix (pair)). Our proposed algorithm to find the nearest PH system is a generalization of the algorithm of [17] (resp., [16]) to find the nearest stable matrix (resp., matrix pair). In fact, it can be used on the system $(I, A, [\,], [\,], [\,])$ (resp., $(E, A, [\,], [\,], [\,])$), where $[\,]$ is the empty matrix, to recover a stable approximation of A (resp., (E, A)), not allowing (resp., allowing) perturbation in E. However, for the nearest stable matrix problem, the algorithm of this paper does not perform as well because authors in [17] used an additional heuristic, namely, a scaling of the iterates (J, R, Q) to reduce the Lipschitz constant of the objective function. Improving the performance of our algorithm using heuristics is a topic for further research.

Note that the LMI-based initializations were not introduced in [17, 16] (only the standard initialization was) and could be particularly useful to obtain better solutions, especially when the input matrix (pair) is close to being stable.

6. Numerical experiments. Our code is available from https://sites.google.com/site/nicolasgillis/ and the numerical examples presented below can be directly run from this online code. All tests are preformed using MATLAB R2015a on a laptop Intel Core i7-7500U CPU @2.9GHz, 24GB RAM. The algorithm runs in $O(n^3)$ operations per iteration (assuming $m \leq n$), including projections on the set of PSD matrices, inversion of the matrix Q (for general systems) and all necessary matrix-matrix products. Hence FGM can be applied on a standard laptop with n and m up to a thousand. To solve the SDPs (5.9) and (5.10), we used the interior point method SDPT3 (version 4.0) [40, 41] and use CVX as a modeling system [11, 18]. On a standard laptop, it can be solved for n up to about a hundred (it took about 4 minutes to solve problem (5.9) with $n = 100$ and $m = 10$). For larger problems, it would be interesting to use first-order methods. This is a direction for further research. Therefore, as of now, for large systems ($n \gg 100$), FGM can be used only in combination with the standard initialization scheme.

In the first experiment (section 6.1), we will compare our FGM with the standard projected gradient method (this is FGM restarted at each iteration) to show that FGM converges significantly faster. In the second experiment (section 6.2), we apply FGM on the small-scale problem from [51]. In the third experiment (section 6.3), we use larger mass-spring-damper (MSD) systems. In all cases, we compare the performance of the different initializations strategies from section 5.3. In the fourth

experiment (section 6.4), we compare the deterministic initializations schemes with random initializations, while in the last experiment (section 6.5) we compare the initialization schemes on randomly generated systems.

6.1. Standard system from [7]. Consider the following standard LTI system (E, A, B, C, D) from [7, section 6], where $E = I_4$:

$$A = \begin{bmatrix} -0.08 & 0.83 & 0 & 0 \\ -0.83 & -0.08 & 0 & 0 \\ 0 & 0 & -0.7 & 9 \\ 0 & 0 & -9 & -0.7 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix}^T,$$

$$C = \begin{bmatrix} 0.4 & 0 & 0.4 & 0 \\ 0.6 & 0 & 1 & 0 \end{bmatrix}, \text{ and } D = \begin{bmatrix} 0.3 & 0 \\ 0 & -0.15 \end{bmatrix}. \tag{6.1}$$

This system is asymptotically stable but not PR because the transfer function $G(s)$ does not satisfy the second condition in the Definition 2.1 of PR, e.g., for $s = 1 + 2j$.

Applying FGM on (5.4) with the standard initialization, we obtain (up to four digits of accuracy)

$$\hat{A} = \begin{bmatrix} -0.0810 & 0.8300 & -0.4 & -0.0104 \\ -0.8301 & -0.0799 & 0.0012 & -0.2 \\ -0.0021 & 0.0013 & -0.8521 & 9.1 \\ -0.0146 & -0.9 & -8.9861 & -0.8512 \end{bmatrix}, \hat{B} = \begin{bmatrix} 0.9994 & 1.8 \\ 0.1 & -0.9 \\ 0.9851 & -0.8691 \\ -0.0070 & 0.2 \end{bmatrix},$$

$$\hat{C} = \begin{bmatrix} 0.4010 & -0.1 & 0.4185 & 0.0073 \\ 0.5990 & 0.0017 & 0.8281 & 0.0158 \end{bmatrix}, \hat{D} = \begin{bmatrix} 0.3089 & -0.0647 \\ -0.0647 & 0.4318 \end{bmatrix}.$$

Figure 6.1 (right) displays the evolution of the objective function: FGM converges in about half a second, while the gradient method requires about 5 seconds. However, both methods converge to the same solution. This gives a nearby standard PR system with error

$$\|A - \hat{A}\|_F^2 + \|B - \hat{B}\|_F^2 + \|C - \hat{C}\|_F^2 + \|D - \hat{D}\|_F^2 = 0.4411.$$

In terms of relative error for each matrix, we have

$$\frac{\|A-\hat{A}\|_F}{\|A\|_F} = 1.68\%, \ \frac{\|B-\hat{B}\|_F}{\|B\|_F} = 6.60\%, \ \frac{\|C-\hat{C}\|_F}{\|C\|_F} = 13.41\%, \ \frac{\|D-\hat{D}\|_F}{\|D\|_F} = 175.62\%.$$

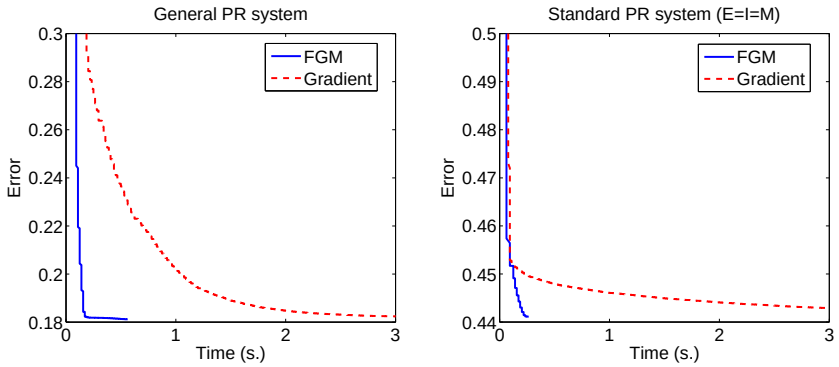

FIG. 6.1. *Evolution of the objective function for FGM and GM for the system in section* 6.1, *using the standard initialization.*

Note that the approximation error for D is rather large. The reason is twofold: (i) D is indefinite and the symmetric part of D has to be approximated by a PSD matrix (namely, S), hence the relative error is at least $\frac{0.15}{\sqrt{0.3^2+0.15^2}} = 44.7\%$ (this error can be obtained by increasing the weight for D in the objective function), and (ii) the norm of D compared to the other matrices (in particular A) is smaller, hence it implicitly has less importance in the objective function.

To give more importance to A and B, we can choose, for example, the weights $w_1 = 7/4$, $w_2 = 7/4$, $w_3 = 1/4$ and $w_4 = 1/4$ in the objective function (5.5). Doing so, we get another nearby standard PR system with objective function 0.13, and with the following relative error for each matrix:

$$\frac{\|A-\tilde{A}\|_F}{\|A\|_F} = 0.36\%,\ \frac{\|B-\tilde{B}\|_F}{\|B\|_F} = 1.08\%,\ \frac{\|C-\tilde{C}\|_F}{\|C\|_F} = 19.79\%,\ \frac{\|D-\tilde{D}\|_F}{\|D\|_F} = 197.13\%.$$

Allowing perturbations in E, we obtain with FGM a nearby PR system with the approximation error 0.1812 where weights are all equal to one (note that, as expected, it is smaller than for the more constrained standard case with error 0.4411). Figure 6.1 (left) displays the evolution of the objective function; FGM and the gradient method have behavior similar to the standard case. The relative errors are

$$\frac{\|E-\hat{E}\|_F}{\|E\|_F} = 11.49\%,\ \frac{\|A-\hat{A}\|_F}{\|A\|_F} = 0.19\%,\ \frac{\|B-\hat{B}\|_F}{\|B\|_F} = 1.81\%,$$
$$\frac{\|C-\hat{C}\|_F}{\|C\|_F} = 3.32\%,\ \frac{\|D-\hat{D}\|_F}{\|D\|_F} = 105.24\%.$$

Similarly, choosing the weights $w_1 = 7/4,\ w_2 = 7/4,\ w_3 = 1/4,\ w_4 = 1/4$ and $w_5 = 1$, we obtain an objective function value of 0.07 and the relative errors are the following:

$$\frac{\|E-\tilde{E}\|_F}{\|E\|_F} = 6.81\%,\ \frac{\|A-\tilde{A}\|_F}{\|A\|_F} = 0.06\%,\ \frac{\|B-\tilde{B}\|_F}{\|B\|_F} = 0.43\%,$$
$$\frac{\|C-\tilde{C}\|_F}{\|C\|_F} = 6.05\%,\ \frac{\|D-\tilde{D}\|_F}{\|D\|_F} = 131.64\%.$$

For this example, the LMI-based initializations from section 5.3.2 perform worse and lead to solutions with larger error (see Table 6.2 later, which summarizes all the results). The reason is that the original system is far from being in PH-form since $\lambda_{\min}(D) = -0.15$; see the experiments in sections 6.3 and 6.5 for more insight on these initializations.

6.2. Descriptor system from [51]. Consider the LTI system (E, A, B, C, D) from [51] (see also [9]) where

$$E = \begin{bmatrix} 16 & 12 & -4 & 14 \\ 14 & 8 & 4 & -14 \\ -14 & 8 & -4 & 34 \\ 6 & -4 & 0 & -10 \end{bmatrix}, A = \begin{bmatrix} 6 & -19 & 7 & -9 \\ 11 & 3 & -21 & 18 \\ 25 & -9 & 35 & -16 \\ -27 & 6 & -16 & 38 \end{bmatrix}, \tag{6.2}$$
$$B = \begin{bmatrix} -0.6 & 1.0 & 0.2 & -0.3 \end{bmatrix}^T, C = \begin{bmatrix} 3.2 & 1.4 & 2.6 & 1.4 \end{bmatrix},\ D = 0.105.$$

The matrix pair (E, A) is of index two with two finite eigenvalues $-0.5 \pm \sqrt{2}j$ hence it is not admissible. This system is stable and remains stable if E and A are not perturbed. However, it is highly sensitive to small perturbation in E and A because

the matrix pair (E, A) has Jordan block at ∞; see, e.g., [10]. For example, replacing E with $E + 10^{-6} I_n$ makes the pencil (E, A) unstable with an eigenvalue at 8002. In [51], C is perturbed to $\hat{C} = [3.0876 \;\; 1.4736 \;\; 2.6 \;\; 1.4]^T$ with $\|C - \hat{C}\|_F^2 = 0.018$ to make the system passive.

FGM with the standard initialization obtains a nearby PR system $(\hat{E}, \hat{A}, \hat{B}, \hat{C}, \hat{D})$ with approximation error 1.28 in 2 seconds, where the relative errors in the different matrices are

$$\frac{\|E - \hat{E}\|_F}{\|E\|_F} = 1.26\%, \frac{\|A - \hat{A}\|_F}{\|A\|_F} = 0.53\%, \frac{\|B - \hat{B}\|_F}{\|B\|_F} = 24.59\%,$$
$$\frac{\|C - \hat{C}\|_F}{\|C\|_F} = 5.04\%, \frac{\|D - \hat{D}\|_F}{\|D\|_F} = 703.69\%.$$

This error is not comparable to the one obtained by [51] because FGM provides a PH system for which $(\hat{E}, \hat{A})$ is admissible with four finite eigenvalues (namely, $-1.98 \pm 9.06j$, $-0.50 \pm 1.42j$).

The initialization "LMIs + solve" provides a worse but reasonable solution with error 12.47 (note that $\|E\|_F^2 + \|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2 + \|D\|_F^2 = 8764$), while FGM with initialization "LMIs + formula" performs very badly with error larger than 10^6. The reason is that the optimal solution X^* of (5.9), which is the initial value for Q, is ill-conditioned (condition number of $3.4\ 10^5$). Note that the original system is far from being in PH form since $\delta^* = 2.78$ in (5.9).

Let us replace E with I_4 for which the system is not stable because we have $\max_i \mathrm{Re}\lambda_i(E, A) = 67.6$. This will illustrate the fact that the different initializations may perform rather differently compared to the previous example. The LMI-based initializations provide a solution with error 2.05, while the standard initialization provides a solution with error 263.79 (note that $\|E\|_F^2 + \|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2 + \|D\|_F^2 = 6102$). In this case, although the initial system is far from being stable, the LMI-based initializations perform very well (note that $\delta^* = 0.4705$ is smaller than in the previous case).

6.3. Mass-spring-damper system. Let us consider the following system: (E, A) is generated as in [16, section 5.3], that is,

(6.3)
$$E = \begin{bmatrix} V & 0 \\ 0 & I_p \end{bmatrix}, A = (J - R)Q, J = \begin{bmatrix} 0 & I_p \\ -I_p & 0 \end{bmatrix}, R = \begin{bmatrix} W & 0 \\ 0 & 0 \end{bmatrix}, Q = \begin{bmatrix} I_p & 0 \\ 0 & H \end{bmatrix},$$

where $V \succ 0$, $W \succ 0$, and $H \succ 0$ are respectively mass, damping, and stiffness matrices of an MSD system. The entries of $B \in \mathbb{R}^{2p,m}$ are generated using the uniform distribution in the interval $[0, 1]$. Generating each entry of $L \in \mathbb{R}^{m,m/2}$ using the normal distribution (mean 0, standard deviation 1), we set $D = LL^T \succeq 0 \in \mathbb{R}^{m,m}$, which is rank deficient, and $C = B^T Q$. This system clearly admits a PH-form and therefore is PR (Theorem 3.2). To make this system non-PR, we perturb R by $\tilde{R} = R + \Delta_R$ as in [16] with

$$\Delta_R = \begin{bmatrix} 0 & 0 \\ 0 & -\epsilon I_p \end{bmatrix}$$

for some $\epsilon > 0$. For the numerical experiment, we take such systems of size $n = 2p = 20$ and $m = 4$, and $n = 2p = 40$ and $m = 6$. We use $\epsilon = 2/(k\,n)$ for $n = 20$ and $\epsilon = 1/(k\,n)$ for $n = 40$ with $k = 1, 2, 3, 4$.

As shown in Table 6.1, the corresponding perturbed systems do not admit a solution to the LMIs (3.1) or, equivalently, $\delta^* > 0$ in (5.9). However, as ϵ decreases,

TABLE 6.1
Optimal value δ^* of (5.9), and largest real part of the eigenvalues of the pair (E, A) for the different perturbed MSD systems.

$n = 20$, $m = 4$	$\epsilon = 2/n$	$\epsilon = 1/n$	$\epsilon = 3/(2n)$	$\epsilon = 1/(2n)$
δ^* of (5.9)	5.4269	2.3647	1.7452	0.1295
$\max_i \mathrm{Re}\lambda_i(E, A)$	2.5749	0.7514	0.0760	−0.0031
$n = 40$, $m = 6$	$\epsilon = 1/n$	$\epsilon = 1/(2n)$	$\epsilon = 1/(3n)$	$\epsilon = 1/(4n)$
δ^* of (5.9)	7.7355	0.6577	0.2488	0.1304
$\max_i \mathrm{Re}\lambda_i(E, A)$	0.9004	-0.0007	-0.0007	-0.0007

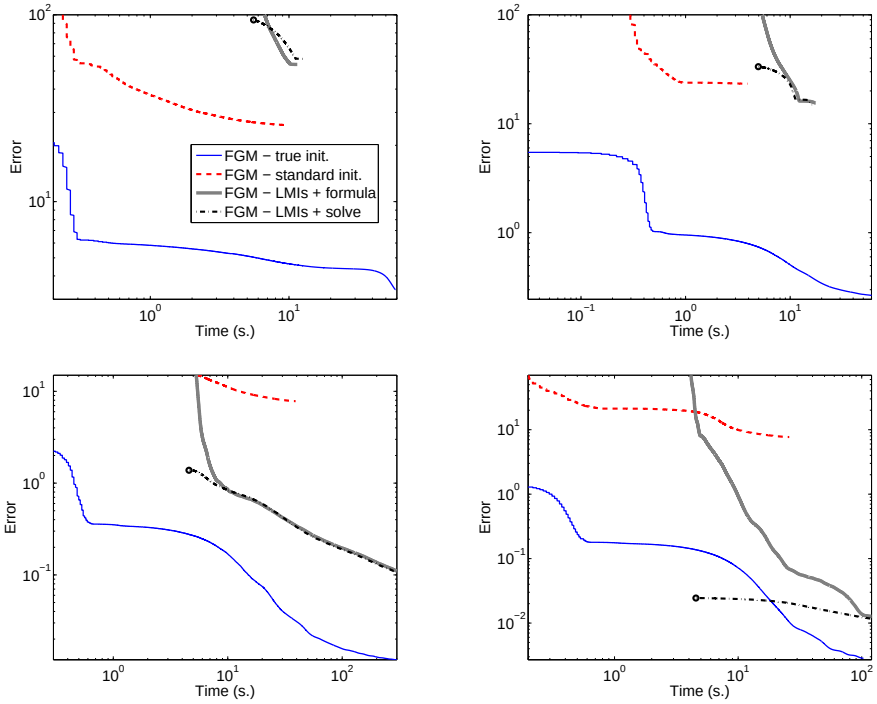

FIG. 6.2. *Evolution of the objective function for FGM with the different initializations for the perturbed MSD system with $n = 20$, $m = 4$, and $\epsilon = 2/(nk)$, with $k = 1$ (top left), $k = 2$ (top right), $k = 3$ (bottom left), and $k = 4$ (bottom right).*

the system gets closer to a system admitting a PH-form in the sense that δ^* decreases. Moreover, for smaller values of ϵ, the pair (E, A) is asymptotically stable; see Table 6.1.

We compare four different initializations: the standard initialization (section 5.3.1), LMIs + formula, and LMIs + solve (section 5.3.2), and the initialization using the unperturbed system, that is, taking (J, R, Q) as in (6.3), $F = B$, $P = 0$, $S = D$, and $N = 0$. We will refer to this last initialization as the "true" initialization as it corresponds to the groundtruth unperturbed PH system.

Figure 6.2 (resp., Figure 6.3) displays the evolution of the objective function values using these different initializations for $n = 20$ and $m = 4$ (resp., $n = 40$ and $m = 6$) for the different values of ϵ. The weights in the objective function are all equal to one.

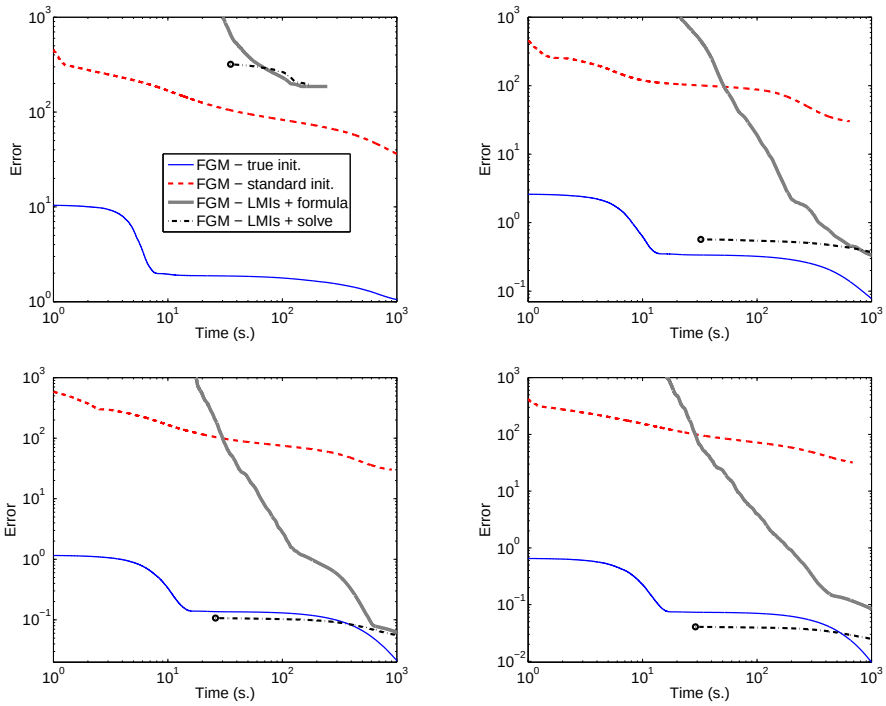

FIG. 6.3. *Evolution of the objective function for FGM with the different initializations for the perturbed MSD system with* $n = 40$, $m = 6$, *and* $\epsilon = 1/(nk)$, *with* $k = 1$ *(top left),* $k = 2$ *(top right),* $k = 3$ *(bottom left), and* $k = 4$ *(bottom right).*

Table 6.2 gives the final error obtained by the different algorithms with a maximum time limit of 300 seconds for $n = 20$ and 1000 seconds for $n = 40$. Before we comment on these results, it is important to put these numbers into perspective: we have $\|E\|_F^2 + \|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2 + \|D\|_F^2 = 5456$ (resp., $= 43078$) for $n = 20$ (resp., $n = 40$). Hence, for example, the largest error of 58.00 (resp., 196.11) of LMIs + solve for $n = 20$ (resp., $n = 40$) and $k = 1$ corresponds to a reasonable approximation, although it is much larger than for some other approaches.

For both dimensions, we observe similar behavior by FGM for the different initializations:

- FGM converges in most cases at a sublinear rate; see Figures 6.2 and 6.3, where the objective function values decrease roughly linearly in a logarithmic time scale.
- For the true initialization, FGM recovers a solution with the smallest error after sufficiently many iterations. This is rather natural since the initialization corresponds to the original unperturbed PH system.
- For the standard initialization, FGM converges to systems with error larger than with the true initialization and gets stuck is some local minima. This illustrates the importance of choosing good initial points, although, as mentioned above, in terms of relative error, these solutions still provide good approximations. For high perturbations ($k = 1$), it provides significantly better solutions than the LMI-based initializations.

TABLE 6.2

Comparison of the error obtained by the different initialization schemes on the systems from sections 6.1, 6.2, *and* 6.3. *For the* 100 *random initializations, we also report the mean and standard deviation in parentheses (for the error obtained within* 10 *seconds). Bold indicates the lowest error among all initializations.*

	Random	Standard	LMI+form.	LMI+sol.	True
(6.1) $\hat{E} = I$	**0.4411** (0.52 ± 0.35)	**0.4411**	4.07	3.26	/
(6.1)	**0.1812** (3.38 ± 11.9)	**0.1812**	2.27	2.22	/
(6.2)	**0.62** (172 ± 164)	1.28	$> 10^6$	12.47	/
(6.2) $E = I_4$	**1.39** (63.43 ± 148)	263.79	2.05	2.05	/
n=20, k=1	**1.90** (37.57 ± 22.20)	26.69	54.22	58.00	3.38
n=20, k=2	1.48 (34.52 ± 25.56)	23.31	15.50	14.97	**0.25**
n=20, k=3	**0.01** (32.54 ± 30.38)	7.82	0.11	0.11	**0.01**
n=20, k=4	0.92 (33.91 ± 24.48)	7.68	9.21 10^{-3}	8.96 10^{-3}	**2.85** 10^{-3}
n=40, k=1	66.19 (1708 ± 613)	36.26	185.93	196.11	**1.05**
n=40, k=2	13.38 (1815 ± 673)	30.42	0.33	0.37	**0.08**
n=40, k=3	299.65 (1578 ± 653)	30.31	0.06	0.06	**0.02**
n=40, k=4	18.32 (1704 ± 603)	32.05	0.09	0.03	**0.02**

- For LMIs + formula, the initial error is rather high (even for a small perturbation ϵ), because the formula (3.3) does not provide a good estimate of the PH-form when the system is not PH.
  For large ϵ ($k = 1$), it is not able to recover a solution close to the one obtained with the true initialization.
  For ϵ sufficiently small, and after sufficiently many iterations, it is able to recover a solution with error similar to that of FGM initialized with LMIs + solve and close to that obtained with the true initialization. LMIs + formula and LMIs + solve often converge to similar solutions, which can be explained by the fact that both initializations use the same initial Q and Z.
- For LMIs + solve, FGM is able to recover better and better solutions as ϵ decreases. For the largest ϵ ($k = 1$), it performs worse than the standard initialization. For the smallest ϵ ($k = 4$), the initial point obtained with LMIs + solve has smaller error than the true initialization. Since the initialization LMIs + solve computes the optimal values for J, R, P, S, N, and F for fixed Q (at a larger initial computational cost), it is not surprising that it has a much lower initial error than LMIs + formula.

Remark 6 (nearest strict PH system). The PH system obtained with FGM is not necessarily strict since the cost matrix K can be rank deficient. For example, with the true initialization for $n = 20$ and $k = 1$, it has 11 eigenvalues with modulus smaller than 10^{-12}. It is possible to impose that the system be strict (hence admissible, and

ESPR if $D + D^T \succ 0$; see Theorem 3.9) using a lower bound on the eigenvalues of K, which does not make the projection step more complicated. Note that it is also possible to use a lower bound ν for the eigenvalues of Z to have Q invertible (as long as it is initialized with an invertible matrix). In fact, the objective function is guaranteed to decrease under the updates of FGM; hence the term $\|E^T - ZQ^{-1}\|_F^2$ remains bounded, which guarantees Q will be invertible since we would have $Z \succeq \nu I_n$. We have included this option in our code.

6.4. Random initializations. So far, we have only used deterministic initializations. As expected, in some cases, they do not lead to good solutions (see, for example, Table 6.2). Therefore, an important direction for further research is to design new initialization schemes, possibly depending on the problem at hand. For example, we have seen in the previous sections that if the perturbed system is close to being passive, then the LMI-based initializations perform well. We will confirm this behavior on randomly generated systems in the next section. A simple initialization scheme is to use random matrices. In this section, we perform some numerical experiments to get some insight on whether this allows us to recover good solutions for the systems presented in the previous sections.

Defining a Gaussian matrix as a matrix whose entries are generated randomly using the normal distribution $N(0, 1)$ (we used the function `randn` in MATLAB), we initialize the variables as follows: Q is the product of a Gaussian matrix with its transpose so that it is full rank,[2] J, Z, and $\left[\begin{smallmatrix} R & S \\ S^T & P \end{smallmatrix}\right]$ are Gaussian matrices projected onto the feasible set, and F is chosen optimal using (5.6).

Table 6.2 summarizes the results using 100 random initializations with a time limit of 10 seconds for Algorithm 1. However, the comparison is not fair for the MSD systems of section 6.3 that were run 300 and 1000 seconds, respectively, for $n = 20$ and $n = 40$, but this would take a long time to run 100 times. Hence, for these systems, Algorithm 1 was run for 300 and 1000 additional seconds only on the best solution obtained after 10 seconds among the 100 random initial points.

In many cases, random initialization identifies good solutions; in fact, it achieves the lowest error compared to the three deterministic approaches for all problems of size $n \leq 20$, except for the MSD system with $n = 20$ and $k = 4$. It even competes similarly as the "true" initialisation for the MSD systems (providing a lower error for $n = 2$ and $k = 1$, and the same error for $n = 2$ and $k = 3$). However, for the larger MSD systems with $n = 40$, it is not able to compete with the LMI-based initializations that perform well in this situation (for $k \geq 2$). The reason is that the number of local minima increases and the algorithm converges in the basin of attraction of worse local minima: not surprisingly, the standard deviation of the errors obtained with random initializations increases with the dimension of the problem. For example, for $n = 40$ and $k = 3$, the error obtained is rather high, namely, 299.65. However, running the algorithm with 100 other random initializations, we obtained an error of 28.07. For larger problems, a direction for further research is therefore to design more sophisticated heuristics to identify better solutions and avoid the basin of attraction of bad local minimizers.

6.5. Randomly generated systems. In the previous sections, we observed that if the perturbed system is close to being passive, then the LMI-based initializa-

[2]For n large, the conditioning of Q can be bad ($\gg 10^4$). In that case, we compute the SVD of Q and set the smallest singular values of Q to $\sigma_{\max}(Q)/\kappa$ so that the conditioning of Q is a most κ (we used $\kappa = 10^4$).

tions are able to recover a passive system that is closer to the perturbed one than the initial true passive system. This means that the LMI-based initializations provide an initial point that is in the basin of attraction of a very good local minimum (possibly the global minimum, although this is difficult to verify). Intuitively, the reason is that the LMIs (5.10) are only slightly perturbed, hence the solution will be close to the solution of the system for the unperturbed passive system. This is closely related to perturbation analysis of optimization problems [6]. Although we are not able to prove this important observation (which would be a very interesting direction of further research), we perform in this section additional numerical experiments to support it.

To generate systems randomly, we use the same strategy as in the previous section to obtain a PH system and set $N = 0$. Then, given the parameter δ, we perturb R and S in the same way as follows: given a matrix X,

- compute its singular values decomposition $[U, \Sigma, V]$ with $X = U\Sigma V^T$,
- set the smallest singular value in Σ to zero (note that, in many cases, R and S already have a singular value equal to zero since they are the projection of Gaussian matrices onto the PSD cone),
- compute $\Sigma' = \Sigma - \delta\sigma_{\max}$, where $\sigma_{\max}$ is the largest entry of Σ, and replace X with $\tilde{X} = U\Sigma' V$.

With this procedure, the perturbed R and S do not belong to the PSD cone, for any $\delta > 0$. Since S is the symmetric part of D, this will generate a perturbed system that does not admit a PH form. The parameter δ is chosen such that a certain relative distance ϵ is achieved between the randomly generated PH system (E, A, B, C, D) and its perturbation $(\tilde{E}, \tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$:

$$\frac{\|(A - \tilde{A}, E - \tilde{E}, B - \tilde{B}, C - \tilde{C}, D - \tilde{D})\|_F}{\|(\tilde{A}, \tilde{E}, \tilde{B}, \tilde{C}, \tilde{D})\|_F} = \epsilon.$$

(We used a simple bisection scheme to find δ, given ϵ.) We will compare the three deterministic initialization schemes as in the previous sections (standard, LMI+formula, LMI+solve) along with the "true" initialization, which is the original unperturbed randomly generated PH system. Hence, the "true" initialization is guaranteed to achieve a relative error smaller ϵ, since Algorithm 1 is guaranteed to decrease the objective function at each step.

We generate systems with $n = 20$ and $m = 5$ as described above, and Table 6.3 summarizes the average relative error in percent among 10 such randomly generated systems, with a time limit of 100 seconds, for different values of ϵ.

The standard initialization scheme consistently performs worse than the other approaches. LMIs + formula performs rather well, similarly to the "true" initialization, while LMIs + solve surprisingly performs best in all scenarios, being able to identify better solutions than the "true" initialization.

TABLE 6.3

Average relative error in percent for the different initialization schemes for randomly generated systems of size $n = 20$ and $m = 5$.

ϵ	Standard init.	LMIs + formula	LMIs + solve	True init.
50%	17.95	11.21	2.26	13.38
10%	12.37	0.89	0.24	0.87
1%	12.60	0.52	0.0071	0.21
0.1%	15.89	0.19	0.0028	0.035

This experiment allows us to confirm our previous observation: as the perturbed system gets closer to a PH system, LMI-based initialization is able to recover better and better solutions.

7. Conclusion and further research. In this paper, we have proposed the first algorithm to tackle the nearest positive-real (PR) system problem that allows the perturbation of all matrices (E, A, B, C, D) describing an LTI system. Our approach combines a reformulation of PR systems as port-Hamiltonian (PH) systems and a fast gradient method (FGM). We have illustrated the effectiveness of our approach with several examples. In particular, we observed that if the initial system is close to being PH, then the proposed LMI-based initializations allow us to recover nearby PH systems. An interesting direction for further research would be to characterize this rigorously, e.g., providing error bounds for the LMI-based initializations. Another observation is that FGM is sensitive to initialization and does not always converge fast (often at a sublinear rate); hence further research includes, for example, the design of (i) new initialization schemes, (ii) more efficient algorithms (e.g., using second-order information), and (iii) globalization approaches to escape local minima.

Acknowledgment. The authors thank the anonymous reviewers for their insightful comments which helped improve the paper.

REFERENCES

[1] R. Alam, S. Bora, M. Karow, V. Mehrmann, and J. Moro, *Perturbation theory for Hamiltonian matrices and the distance to bounded-realness*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 484–514.

[2] B. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis*, Prentice-Hall, Englewood Cliffs, NJ, 1973.

[3] A. C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, Adv. Des. Control 6, SIAM, Philadelphia, 2005.

[4] C. Beattie, V. Mehrmann, and H. Xu, *Port-Hamiltonian Realizations of Linear Time Invariant Systems*, Preprint 23-2015, Institut für Mathematik, TU Berlin, 2015.

[5] C. Beattie, V. Mehrmann, H. Xu, and H. Zwart, *Port-Hamiltonian Descriptor Systems*, arXiv:1705.09081, 2017.

[6] J. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2013.

[7] S. Boyd, V. Balakrishnan, and P. Kabamba, *A bisection method for computing the H_∞ norm of a transfer matrix and related problems*, Math. Control Signals Systems, 2 (1989), pp. 207–219.

[8] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Stud. Appl. Numer. Math. 15, SIAM, Philadelphia, 1994.

[9] T. Brull and C. Schröder, *Dissipativity enforcement via perturbation of para-Hermitian pencils*, IEEE Trans. Circuits Systems, 60 (2013), pp. 164–177.

[10] R. Byers and N. Nichols, *On the stability radius of a generalized state-space system*, Linear Algebra Appl., 188 (1993), pp. 113–134.

[11] *CVX: MATLAB Software for Disciplined Convex Programming, Version* 2.0, http://cvxr.com/cvx (2012).

[12] C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, New York, 1975.

[13] R. Freund and F. Jarre, *An extension of the positive real lemma to descriptor systems*, Optim. Methods Softw., 19 (2004), pp. 69–87.

[14] F. Gantmacher, *The Theory of Matrices* I, Chelsea Publishing, New York, 1959.

[15] S. Ghadimi and G. Lan, *Accelerated gradient methods for nonconvex nonlinear and stochastic programming*, Math. Program., 156 (2016), pp. 59–99.

[16] N. Gillis, V. Mehrmann, and P. Sharma, *Computing the nearest stable matrix pairs*, Numer. Linear Algebra Appl., (2018), e2153.

[17] N. Gillis and P. Sharma, *On computing the distance to stability for matrices using linear dissipative Hamiltonian systems*, Automatica, 85 (2017), pp. 113–121.

[18] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, V. Blondel, S. Boyd, and H. Kimura, eds., Lecture Notes in Control and Inform. Sci. 371, Springer-Verlag, Berlin, 2008, pp. 95–110.
[19] S. Grivet-Talocia, *Passivity enforcement via perturbation of Hamiltonian matrices*, IEEE Trans. Circuits Systems, 51 (2004), pp. 1755–1769.
[20] N. Guglielmi, D. Kressner, and C. Lubich, *Low rank differential equations for Hamiltonian matrix nearness problems*, Numer. Math., 129 (2015), pp. 279–319.
[21] M. Haddad and D. Bernstein, *Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle and Popov theorems and their application to robust stability*, in Proceedings of the 30th IEEE Conference on Decision and Control, Vol. 3, 1991, pp. 2618–2623.
[22] N. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl., 103 (1988), pp. 103–118.
[23] C.-H. Huang, P. Ioannou, J. Maroulas, and M. Safonov, *Design of strictly positive real systems using constant output feedback*, IEEE Trans. Automat. Control, 44 (1999), pp. 569–573.
[24] T. H. Hughes, *A theory of passive linear systems with no assumptions*, Automatica, 86 (2017), pp. 87–97.
[25] P. Ioannou and G. Tao, *Frequency domain conditions for strictly positive real functions*, IEEE Trans. Automat. Control, 32 (1987), pp. 53–54.
[26] S. Joshi, *Control of Large Flexible Space Structures*, Lecture Notes in Control and Inform. Sci., 131, Springer-Verlag, Berlin, 1989.
[27] P. Kunkel and V. Mehrmann, *Differential-Algebraic Equations: Analysis and Numerical Solution*, EMS Textbk. Math., European Mathematical Society, Zürich, 2006.
[28] R. Lozano, B. Brogliato, O. Egeland, and B. Maschke, *Dissipative Systems Analysis and Control: Theory and Applications*, Springer, New York, 2013.
[29] R. Lozano-Leal and S. Joshi, *Strictly positive real transfer functions revisited*, IEEE Trans. Automat. Control, 35 (1990), pp. 1243–1245.
[30] C. Mehl, V. Mehrmann, and P. Sharma, *Stability radii for linear Hamiltonian systems with dissipation under structure-preserving perturbations*, SIAM J. Matrix Anal. Appl., 37 (2016), pp. 1625–1654.
[31] C. Mehl, V. Mehrmann, and P. Sharma, *Stability radii for real linear Hamiltonian systems with perturbed dissipation*, BIT, 57 (2017), pp. 811–843, doi:10.1007/s10543-017-0654-0.
[32] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Solution*, Lecture Notes in Control and Inform. Sci. 163, Springer-Verlag, Berlin, 1991.
[33] Y. Nesterov, *A method of solving a convex programming problem with convergence rate $o(1/k2)$*, Soviet Math. Dokl., 27 (1983), pp. 372–376.
[34] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, Appl. Optim. 87, Springer, New York, 2004.
[35] M. O'Neill and S. Wright, *Behavior of Accelerated Gradient Methods Near Critical Points of Nonconvex Problems*, arXiv:1706.07993, 2017.
[36] M. Overton and P. Van Dooren, *On computing the complex passivity radius*, in Proceedings of the 44th IEEE Conference on Decision and Control, 2005, pp. 7960–7964.
[37] V. Popov, *Hyperstability of Control Systems*, Springer, New York, 1973.
[38] C. Schröder and T. Stykel, *Passivation of LTI Systems*, Preprint 368, MATHEON, Berlin, 2007.
[39] W. Sun, P. Khargonekar, and D. Shim, *Solution to the positive real control problem for linear time-invariant systems*, IEEE Trans. Automat. Control, 39 (1994), pp. 2034–2046.
[40] K.-C. Toh, M. Todd, and R. Tütüncü, *SDPT3—a MATLAB software package for semidefinite programming, version* 1.3, Optim. Methods Softw., 11 (1999), pp. 545–581.
[41] R. Tütüncü, K. Toh, and M. Todd, *Solving semidefinite-quadratic-linear programs using SDPT3*, Math. Program., 95 (2003), pp. 189–217.
[42] A. van der Schaft, *Port-Hamiltonian systems: An introductory survey*, in Proceedings of the International Congress of Mathematicians, Vol. III, Invited Lectures, J. V. M. Sanz-Sole and J. Verdura, eds., Madrid, Spain, 2006, pp. 1339–1365.
[43] A. van der Schaft, *Port-Hamiltonian Differential-Algebraic Systems*, in Surveys in Differential-Algebraic Equations, Springer, New York, 2013, pp. 173–226.
[44] A. van der Schaft and D. Jeltsema, *Port-Hamiltonian systems theory: An introductory overview*, Found. Trends Systems Control, 1 (2014), pp. 173–378.
[45] A. van der Schaft and B. Maschke, *The Hamiltonian formulation of energy conserving physical systems with external ports*, Arch. Elektron. Übertragungstech., 45 (1995), pp. 362–371.
[46] A. van der Schaft and B. Maschke, *Hamiltonian formulation of distributed-parameter systems with boundary energy flow*, J. Geom. Phys., 42 (2002), pp. 166–194.

[47] A. van der Schaft and B. Maschke, *Port-Hamiltonian systems on graphs*, SIAM J. Control Optim., 51 (2013), pp. 906–937.

[48] A. Varga, *On stabilization methods of descriptor systems*, Systems Control Lett., 24 (1995), pp. 133–138.

[49] M. Voigt and P. Benner, *Passivity enforcement of descriptor systems via structured perturbation of Hamiltonian matrix pencils*, pressented at Meeting of the GAMM Activity Group Dynamics and Control Theory, Linz, 2011.

[50] H.-S. Wang and F.-R. Chang, *The generalized state-space description of positive realness and bounded realness*, in Proceedings of the 39th Midwest Symposium on Circuits and Systems, Vol. 2, 1996, pp. 893–896.

[51] Y. Wang, Z. Zhang, C. Koh, G. Pang, and N. Wong, *PEDS: Passivity enforcement for descriptor systems via Hamiltonian-symplectic matrix pencil perturbation*, in proceedings of the IEEE/ACM International Conference on Computer-Aided Design, 2010, pp. 800–807.

[52] J. Wen, *Time domain and frequency domain conditions for strict positive realness*, IEEE Trans. Automat. Control, 33 (1988), pp. 988–992.

[53] J. Willems, *Dissipative dynamical systems—part* II*: Linear systems with quadratic supply rates*, Arch. Ration. Mech. Anal., 45 (1972), pp. 352–393.

[54] L. Zhang, J. Lam, and S. Xu, *On positive realness of descriptor systems*, IEEE Trans. Circuits Systems, 49 (2002), pp. 401–407.