MAX-PLANCK-GESELLSCHAFT

Peter Benner Tobias Breiten*

Rational interpolation methods for symmetric Sylvester equations

MAX–PLANCK–INSTITUT
FÜR DYNAMIK KOMPLEXER
TECHNISCHER SYSTEME
MAGDEBURG

Max Planck Institute Magdeburg Preprints

MPIMD/13-23 December 10, 2013

Impressum:

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg

Publisher:
Max Planck Institute for Dynamics of Complex Technical Systems

Address:
Max Planck Institute for Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

www.mpi-magdeburg.mpg.de/preprints

MAX-PLANCK-GESELLSCHAFT

Max Planck Institute Magdeburg
Preprints

Peter Benner Tobias Breiten*

Rational interpolation methods for symmetric Sylvester equations

MAX–PLANCK–INSTITUT
FÜR DYNAMIK KOMPLEXER
TECHNISCHER SYSTEME
MAGDEBURG

*Institute for Mathematics and Scientific Computing, Heinrichstr. 36/III, University of Graz, Austria (tobias.breiten@uni-graz.at). Most of this work was completed while this author was with the Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg.

MPIMD/13-23 December 10, 2013

Abstract

We discuss low rank approximation methods for large-scale symmetric Sylvester equations. Following similar discussions for the Lyapunov case, we introduce an energy norm by the symmetric Sylvester operator. Given a rank n_r, we derive necessary conditions for an approximation being optimal with respect to this norm. We show that the norm minimization problem is related to an objective function based on the $\mathcal{H}_2$-inner product for symmetric state space systems. This objective function is shown to exhibit first-order conditions that are equivalent to the ones from the norm minimization problem. We further propose an iterative procedure and demonstrate its efficiency when used within image reconstruction problems.

1 Introduction

In this paper, we consider large-scale linear matrix equations

$$\mathbf{AXM} + \mathbf{EXH} = \mathbf{G}, \tag{1}$$

with $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}, \mathbf{M}, \mathbf{H} \in \mathbb{R}^{m\times m}$ and $\mathbf{G} \in \mathbb{R}^{n\times m}$. The sought after solution $\mathbf{X} \in \mathbb{R}^{n\times m}$ to the *Sylvester equation* (1) is of great interest within systems and control theory, see [?]. In particular, for $\mathbf{M} = \mathbf{E}^T$, $\mathbf{H} = \mathbf{A}^T$ and $\mathbf{G} = \mathbf{BB}^T$, the resulting *Lyapunov equation* characterizes stability properties of an underlying dynamical system

$$\begin{aligned} \mathbf{E}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t), \end{aligned} \tag{2}$$

where, respectively, $\mathbf{x}(t), \mathbf{u}(t)$ and $\mathbf{y}(t)$ are called *state, control* and *output* of the system. While linear matrix equations of the form (1) have been studied for several years now, the need for efficient solution methods when n, m become large has also attracted attention in the numerical linear algebra community. For a detailed introduction into linear matrix equations, we refer to the two recent survey articles [?, ?]. Since direct methods, e.g., the Bartels-Stewart algorithm ([?]) or Hammarling's method ([?]) require a cubic complexity to solve (1), they are only feasible as long as n, m are medium size. Depending on the individual computer architecture, this nowadays might cover system dimensions up to $n, m \sim 10^4$. Often dynamical systems, and thus matrix equations, however result from a spatial discretization of a partial differential equation (PDE). Here, one commonly obtains systems with millions of degrees of freedom, preventing the use of the mentioned direct methods. For the most general case where the right hand side $\mathbf{G}$ is full rank, there is still no easily applicable technique to compute $\mathbf{X}$. On the other hand, assuming that $\mathbf{G} = \mathbf{BC}^T$, where $\operatorname{rank}(\mathbf{B}), \operatorname{rank}(\mathbf{C}) \ll n, m$, the singular values of $\mathbf{X}$ often decay very fast, see [?, ?, ?, ?]. In other words, the low numerical rank of the solution allows for *low rank approximations* $\mathbf{X} \approx \mathbf{V}\mathbf{X}_r\mathbf{W}^T$, where $\mathbf{V} \in \mathbb{R}^{n\times n_r}, \mathbf{W} \in \mathbb{R}^{m\times n_r}$ and $\mathbf{X}_r \in \mathbb{R}^{n_r\times n_r}$, with $n_r \ll n, m$. This phenomenon has been used to construct several numerically efficient methods applicable for large-scale matrix equations. The most popular choices can basically be categorized into two categories: (a) methods based on alternating directions implicit (ADI) schemes; (b) methods based on projection and prolongation. Methods that specifically address the computation of low-rank approximations to the solution of Sylvester equations can be found in [?, ?, ?, ?], while the literature on low-rank solution for the Lyapunov case goes even back further and has achieved more attention; for a detailed overview on this topic, we refer to [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?]. Other techniques are based on the tensorized linear system, see [?, ?, ?], or Riemannian optimization, see [?, ?]. Especially the latter class of methods is important in the context of this article since it requires a special symmetry assumption that we comment on below.

The structure of this paper now is as follows. In Section 2, we review the use of large-scale Sylvester equations (1) for problems evolving in image restoration. For the special symmetry property arising in these applications, in Section 3 we introduce an objective function based on the energy norm of the underlying Sylvester operator. We further derive first-order necessary conditions for this objective function. In Section 4, we establish a connection between the energy norm and the $\mathcal{H}_2$-inner product of two dynamical control systems of the form (2). We show that this inner product exhibits first-order necessary conditions that are equivalent to the ones for the energy norm. Based on techniques from rational interpolation, we discuss the use of an iterative Sylvester solver applicable in large-scale settings. In Section 5, we provide numerical results from image restoration problems and demonstrate the applicability of the method. We conclude with a short summary in Section 6.

In all what follows, a matrix $\mathbf{A} \succ 0$ will be symmetric positive definite. With $\otimes$ we denote the Kronecker product of two matrices. Vectorization of a matrix $\mathbf{A}$, i.e., stacking all columns of $\mathbf{A}$ into a long vector, will be denoted by $\operatorname{vec}(\mathbf{A})$. The (matrix-valued) residue of a meromorphic matrix-valued function, $\mathbf{G}(s)$, at a point $\lambda \in \mathbb{C}$ will be denoted as $\operatorname{res}[\mathbf{G}(s), \lambda]$. All vectors and matrices are denoted by boldface letters; scalar quantities by italic letters. The Kronecker delta δ_{ij} is defined as

$$\delta_{ij} := \begin{cases} 1, & i = j, \\ 0, & \text{otherwise}. \end{cases}$$

2 Sylvester equations in image restoration

Besides their use in control theory, Sylvester equations also appear in restoration problems for degraded images. In what follows, we give a brief recapitulation of the discussions in [?, ?, ?]. Consider an image represented by a matrix $\mathbf{F} \in \mathbb{R}^{n \times m}$ with grayscale pixel values $\mathbf{F}_{ij}$ between 0 and 255. Unfortunately, often the matrix $\mathbf{F}$ is not given exactly but is perturbed by some noise or blurring process. The result is a degraded image $\mathbf{G} \in \mathbb{R}^{n \times m}$ that is obtained after an out-of-focus or atmospheric blur. One way to compute an approximately restored image $\mathbf{X} \approx \mathbf{F}$ is given by the solution to a regularized linear discrete ill-posed problem of the form

$$\min_{\mathbf{x}} \left(\|\mathbf{H}\mathbf{x} - \mathbf{g}\|_2^2 + \lambda \|\mathbf{L}\mathbf{x}\|_2^2 \right), \tag{3}$$

where $\mathbf{x} = \operatorname{vec}(\mathbf{X}), \mathbf{g} = \operatorname{vec}(\mathbf{G}), \mathbf{H}$ models the degradation process and $\mathbf{L}$ is a regularization operator with regularization parameter λ. The solution to (3) can be computed by solving the linear system

$$(\mathbf{H}^T\mathbf{H} + \lambda^2 \mathbf{L}^T\mathbf{L})\mathbf{x} = \mathbf{H}^T\mathbf{g}.$$

While the choice of an appropriate or optimal parameter λ is a nontrivial task, we rather want to focus on efficiently solving the linear system once λ has been determined. This can for example be done by using the L-curve criterion or the generalized cross validation method, see [?, ?]. Following e.g. [?], assuming certain separability properties of the blurring matrix $\mathbf{H} = \mathbf{H}_2 \otimes \mathbf{H}_1$ and the regularization operator $\mathbf{L} = \mathbf{L}_2 \otimes \mathbf{L}_1$, problem (3) has a special structure and can equivalently be solved by the Sylvester equation

$$(\mathbf{H}_1^T\mathbf{H}_1)\mathbf{X}(\mathbf{H}_2^T\mathbf{H}_2) + \lambda^2(\mathbf{L}_1^T\mathbf{L}_1)\mathbf{X}(\mathbf{L}_2^T\mathbf{L}_2) = \mathbf{G}. \tag{4}$$

In particular, we note that the matrices defining the matrix equation are symmetric positive (semi-)definite. In the next section, we see that this enables us to use special approximation techniques. Before we proceed, we mention typical structures of $\mathbf{H}$ and $\mathbf{L}$ that we pick up in the numerical examples. Again, we follow the more detailed discussions in [?, ?]. A uniform out-of-focus blur for example can be modelled by the uniform Toeplitz matrix

$$\mathbf{U}_{ij} = \begin{cases} \frac{1}{2r-1}, & |i-j| \leq r, \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Atmospheric blur can be realized by a Gaussian Toeplitz matrix

$$\mathbf{T}_{ij} = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(i-j)^2}{2\sigma^2}\right), & |i-j| \leq r, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

For the regularization operators, we use discrete first-order derivatives. Note that we do not claim that image restoration via solving Sylvester equations is an optimal choice. Nowadays, methods based on total variation and L_1-norm minimization often produce much more accurate results and might be preferable in practice. Nevertheless, this application serves here as an application providing meaningful test cases for the method to be developed in the following two sections. Note that the symmetric positive definite structure of the underlying Sylvester operator also occurs in control applications in case of symmetric state matrices with collocated actors/sensors.

3 Symmetric Sylvester equations and the energy norm

According to the previous section, we can reconstruct a blurred/noisy image via solving a symmetric Sylvester equation of the form

$$\mathbf{AXM} + \mathbf{EXH} = \mathbf{G}, \tag{7}$$

where $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n \times n} \succ 0,\ \mathbf{M}, \mathbf{H} \in \mathbb{R}^{m \times m} \succ 0$ and $\mathbf{G} \in \mathbb{R}^{n \times m}$. We already mentioned that as the size of the image increases, explicitly solving for $\mathbf{X} \in \mathbb{R}^{n \times m}$ becomes computationally infeasible. Although in the previously discussed application the matrix $\mathbf{G}$ is not necessarily of low (numerical) rank, we still might construct approximations $\mathbf{X} \approx \tilde{\mathbf{X}} := \mathbf{V}\mathbf{X}_r\mathbf{W}^T$ with $\mathbf{V} \in \mathbb{R}^{n \times n_r}, \mathbf{W} \in \mathbb{R}^{m \times n_r}$ and $\mathbf{X}_r \in \mathbb{R}^{n_r \times n_r}$. Note that we do not necessarily require $\mathbf{X}_r$ to be a square matrix but rather have the freedom to choose $\mathbf{V}$ and $\mathbf{W}$ such that they have a different number of columns. Still, using $\mathbf{X}_r \in \mathbb{R}^{n_r \times n_r}$ seems to be a natural choice and also simplifies the notation.

The most common way to evaluate the quality of an approximation is by means of the norm of the error $\|\mathbf{X} - \tilde{\mathbf{X}}\|$. As long as we use, respectively, the spectral norm or the Frobenius norm, for a given rank n_r, the best approximation is given by the singular value decomposition. This result is well-known and follows from the Eckart-Young-Mirsky theorem that can be found in standard text books such as, e.g., [?]. Unfortunately, computing an SVD-based approximant would require the full solution

$\mathbf{X}$ itself. For symmetric systems however, another natural choice for measuring errors is the energy norm. Note that if the solution to problem (3) is unique and the matrices are separable as shown above, the Sylvester operator is symmetric and positive definite. For the error $\mathbf{X}-\tilde{\mathbf{X}}$ we can thus define a norm via

$$\|\mathbf{X}-\tilde{\mathbf{X}}\|_{\mathcal{L}_S}^2 := \underbrace{\operatorname{vec}\left(\mathbf{X}-\tilde{\mathbf{X}}\right)^T}_{\mathbf{e}^T} \underbrace{(\mathbf{M}\otimes\mathbf{A}+\mathbf{H}\otimes\mathbf{E})}_{\mathcal{L}_S} \underbrace{\operatorname{vec}\left(\mathbf{X}-\tilde{\mathbf{X}}\right)}_{\mathbf{e}}. \tag{8}$$

The energy norm for matrix equations was first investigated in detail in [?, ?] and later discussed in the context of $\mathcal{H}_2$-model reduction in [?]. Note that there is also a direct connection between the Frobenius norm and the energy norm of the error $\mathbf{X}-\tilde{\mathbf{X}}$:

$$\|\mathbf{X}-\tilde{\mathbf{X}}\|_{\mathcal{L}_S}^2 = \mathbf{e}^T\mathcal{L}_S\mathbf{e} = \frac{\mathbf{e}^T\mathcal{L}_S\,\mathbf{e}}{\mathbf{e}^T\mathbf{e}}\|\mathbf{e}\|_2^2 \geq \lambda_{\min}(\mathcal{L}_S)\,\|\mathbf{X}-\tilde{\mathbf{X}}\|_F^2.$$

The last step holds due to the fact that the Rayleigh quotient $R(\mathcal{L}_S,\mathbf{e})$ is bounded below by the minimal eigenvalue of the symmetric matrix $\mathcal{L}_S$. Assume now that for a given Sylvester equation (7) and a prescribed dimension $n_r \ll n$, the goal is to find matrices $\mathbf{V}\in\mathbb{R}^{n\times n_r}, \mathbf{W}\in\mathbb{R}^{m\times n_r}$ and $\mathbf{X}_r\in\mathbb{R}^{n_r\times n_r}$ such that

$$\|\mathbf{X}-\mathbf{V}\mathbf{X}_r\mathbf{W}^T\|_{\mathcal{L}_S}^2 = \min_{\substack{\tilde{\mathbf{V}}\in\mathbb{R}^{n\times n_r}\\ \tilde{\mathbf{W}}\in\mathbb{R}^{m\times n_r}\\ \tilde{\mathbf{X}}_r\in\mathbb{R}^{n_r\times n_r}}} \|\mathbf{X}-\tilde{\mathbf{V}}\tilde{\mathbf{X}}_r\tilde{\mathbf{W}}^T\|_{\mathcal{L}_S}^2. \tag{9}$$

As a first step towards optimization, one usually considers first-order necessary conditions for $\mathbf{V},\mathbf{W}$ and $\mathbf{X}_r$. For this, we state some useful properties for computing the derivative of the trace function with respect to a matrix. According to [?], for a matrix $\mathbf{X}\in\mathbb{R}^{n\times m}$ and matrices $\mathbf{K},\mathbf{L}$ of compatible dimensions, it holds

$$\begin{aligned}\frac{\partial}{\partial\mathbf{Y}}\left[\operatorname{tr}(\mathbf{K}\mathbf{Y}\mathbf{L})\right] &= \mathbf{K}^T\mathbf{L}^T,\\ \frac{\partial}{\partial\mathbf{Y}}\left[\operatorname{tr}\left(\mathbf{K}\mathbf{Y}\mathbf{L}\mathbf{Y}^T\right)\right] &= \mathbf{K}^T\mathbf{Y}\mathbf{L}^T+\mathbf{K}\mathbf{Y}\mathbf{L}.\end{aligned} \tag{10}$$

Using these properties, we can give the folllowing generalization of similar results obtained for the special case of the Lyapunov equation in [?].

Lemma 1 *Assume that* $(\mathbf{V},\mathbf{W},\mathbf{X}_r)$ *solves* (9). *Then it holds*

$$\left(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M}+\mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H}-\mathbf{G}\right)\mathbf{W}\mathbf{X}_r^T = \mathbf{0}, \tag{11a}$$

$$\mathbf{X}_r^T\mathbf{V}^T\left(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M}+\mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H}-\mathbf{G}\right) = \mathbf{0}, \tag{11b}$$

$$\mathbf{V}^T\left(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M}+\mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H}-\mathbf{G}\right)\mathbf{W} = \mathbf{0}. \tag{11c}$$

Proof Note that by vectorization of (7), we know that it holds

$$\mathcal{L}_S\operatorname{vec}(\mathbf{X}) = \operatorname{vec}(\mathbf{G}).$$

Consequently, we obtain

$$\begin{aligned}
f(\mathbf{V}, \mathbf{W}, \mathbf{X}_r) &= \operatorname{vec}\left(\mathbf{X} - \mathbf{V}\mathbf{X}_r\mathbf{W}^T\right)^T \mathcal{L}_S \operatorname{vec}\left(\mathbf{X} - \mathbf{V}\mathbf{X}_r\mathbf{W}^T\right) \\
&= \operatorname{vec}(\mathbf{X})^T \operatorname{vec}(\mathbf{G}) - 2\operatorname{vec}\left(\mathbf{V}\mathbf{X}_r\mathbf{W}^T\right)^T \operatorname{vec}(\mathbf{G}) \\
&\quad + \operatorname{vec}\left(\mathbf{V}\mathbf{X}_r\mathbf{W}^T\right)^T (\mathbf{M} \otimes \mathbf{A} + \mathbf{H} \otimes \mathbf{E}) \operatorname{vec}\left(\mathbf{V}\mathbf{X}_r\mathbf{W}^T\right) \\
&= \operatorname{tr}\left(\mathbf{X}^T\mathbf{G}\right) - 2\operatorname{tr}\left(\mathbf{W}\mathbf{X}_r^T\mathbf{V}^T\mathbf{G}\right) \\
&\quad + \operatorname{tr}\left(\mathbf{W}\mathbf{X}_r^T\mathbf{V}^T(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M} + \mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H})\right).
\end{aligned}$$

Using that $\operatorname{tr}(\mathbf{K}) = \operatorname{tr}\left(\mathbf{K}^T\right)$ and $\operatorname{tr}(\mathbf{KL}) = \operatorname{tr}(\mathbf{LK})$ for matrices $\mathbf{K}, \mathbf{L}$ of compatible dimensions together with (10) gives

$$\begin{aligned}
\frac{\partial f}{\partial \mathbf{V}} &= 2(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M} + \mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H} - \mathbf{G})\mathbf{W}\mathbf{X}_r^T, \\
\frac{\partial f}{\partial \mathbf{W}} &= 2(\mathbf{M}\mathbf{W}\mathbf{X}_r^T\mathbf{V}^T\mathbf{A}\mathbf{V}\mathbf{X}_r + \mathbf{H}\mathbf{W}\mathbf{X}_r^T\mathbf{V}^T\mathbf{E} - \mathbf{G}^T)\mathbf{V}\mathbf{X}_r, \\
\frac{\partial f}{\partial \mathbf{X}_r} &= 2\mathbf{V}^T(\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{M} + \mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{W}^T\mathbf{H} - \mathbf{G})\mathbf{W}.
\end{aligned}$$

Since a minimizer has to fulfill first-order necessary conditions, it also holds

$$\frac{\partial f}{\partial \mathbf{V}} = \frac{\partial f}{\partial \mathbf{W}} = \frac{\partial f}{\partial \mathbf{X}_r} = \mathbf{0},$$

which shows the assertion.

Along the lines of [**?**], one might consider solving (9) by a Riemannian optimization method. While this certainly is possible, in what follows we prefer to proceed via a connection of (9) and the $\mathcal{H}_2$-inner product of two dynamical systems as this results in a conceptional simpler algorithm which is easy to implement.

4 Tangential interpolation of symmetric state space systems

In this section, it will prove beneficial to assume that the right hand side $\mathbf{G}$ is given in factored form $\mathbf{G} = \mathbf{B}\mathbf{C}^T$ with $\mathbf{B} \in \mathbb{R}^{n \times q}$ and $\mathbf{C} \in \mathbb{R}^{m \times q}$. At this point, it is not particularly important that we have $q \ll n, m$. This also means we can always ensure such a decomposition by, e.g., the SVD of $\mathbf{G}$. We now can associate the energy norm of $\mathbf{X}$ fulfilling (7) with the $\mathcal{H}_2$-inner product of two dynamical systems defined by their transfer functions. For this, recall that if a symmetric state space system is given as

$$\begin{aligned}
\mathbf{E}\dot{\mathbf{x}}(t) &= -\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t), \\
\mathbf{y}(t) &= \mathbf{B}^T\mathbf{x}(t),
\end{aligned} \tag{12}$$

with $\mathbf{x}(t) \in \mathbb{R}^{n\times n}, \mathbf{u}(t), \mathbf{y}(t) \in \mathbb{R}^q$, denoting state, control and output of the system, the *transfer function* is the rational matrix valued function

$$\mathbf{G}_1(s) = \mathbf{B}^T(s\mathbf{E} + \mathbf{A})^{-1}\mathbf{B}. \tag{13}$$

Since $\mathbf{E}, \mathbf{A} \succ 0$, system (12) is asymptotically stable and the poles of $\mathbf{G}_1(s)$ are all in the open left half of the complex plane. Hence, for $\mathbf{G}_1(s)$ and $\mathbf{G}_2(s) := \mathbf{C}^T(s\mathbf{M} + \mathbf{H})^{-1}\mathbf{C}$, the $\mathcal{H}_2$-inner product usually is defined as

$$\begin{aligned} \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{trace}\left(\overline{\mathbf{G}_1(\imath\omega)}\mathbf{G}_2(\imath\omega)^T\right) \text{d}\omega \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{trace}\left(\mathbf{G}_1(-\imath\omega)\mathbf{G}_2(\imath\omega)^T\right) \text{d}\omega. \end{aligned} \tag{14}$$

The previous expression now turns out to be exactly the square of the energy norm of $\mathbf{X}$.

Proposition 2 *Let* $\mathbf{X}$ *be the solution of* $\mathbf{AXM} + \mathbf{EXH} = \mathbf{BC}^T$. *Then it holds*

$$\|\mathbf{X}\|^2_{\mathcal{L}_S} = \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2},$$

where $\mathbf{G}_1(s) = \mathbf{B}^T(s\mathbf{E} + \mathbf{A})^{-1}\mathbf{B}$ *and* $\mathbf{G}_2(s) = \mathbf{C}^T(s\mathbf{M} + \mathbf{H})^{-1}\mathbf{C}$,

Proof First note that we have

$$\|\mathbf{X}\|^2_{\mathcal{L}_S} = \text{vec}\,(\mathbf{X})^T\,(\mathbf{M} \otimes \mathbf{A} + \mathbf{H} \otimes \mathbf{E})\,\text{vec}\,(\mathbf{X})\,.$$

Since $\mathbf{X}$ is a solution of the Sylvester equation, this implies

$$\|\mathbf{X}\|^2_{\mathcal{L}_S} = \text{vec}\,(\mathbf{X})^T\,\text{vec}\left(\mathbf{BC}^T\right).$$

Due to the properties of the trace-operator, we find

$$\|\mathbf{X}\|^2_{\mathcal{L}_S} = \text{trace}\left(\mathbf{X}^T\mathbf{BC}^T\right) = \text{trace}\left(\mathbf{B}^T\mathbf{XC}\right).$$

On the other hand, it is well-known, see, e.g., [**?**], that the solution of a Sylvester equation can be obtained by complex integration as

$$\mathbf{X} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (-\imath\omega\mathbf{E} + \mathbf{A})^{-1}\mathbf{BC}^T(\imath\omega\mathbf{M} + \mathbf{H})^{-1}\text{d}\omega.$$

Pre- and postmuliplication with, respectively, $\mathbf{B}^T$ and $\mathbf{C}$ shows the assertion.

Instead of parametrizing the minimization problem (9) via $\mathbf{V}, \mathbf{W}, \mathbf{X}_r$, the goal is to use *reduced* rational transfer functions

$$\mathbf{G}_{1,r}(s) = \mathbf{B}_r^T(s\mathbf{E}_r + \mathbf{A}_r)^{-1}\mathbf{B}_r \quad \text{and} \quad \mathbf{G}_{2,r}(s) = \mathbf{C}_r^T(s\mathbf{M}_r + \mathbf{H}_r)^{-1}\mathbf{C}_r,$$

with symmetric positive definite matrices $\mathbf{A}_r, \mathbf{E}_r, \mathbf{M}_r$ and $\mathbf{H}_r$ of dimension $n_r \times n_r$. and $\mathbf{B}_r, \mathbf{C}_r \in \mathbb{R}^{n_r \times q}$. Since using every entry of the system matrices would lead to an overparametrization, we replace $\mathbf{G}_{1,r}$ and $\mathbf{G}_{2,r}$ by their *pole-residue representations*. For this, let $\mathbf{A}_r\mathbf{Q} = \mathbf{E}_r\mathbf{Q}\mathbf{\Lambda}$ be the eigenvalue decomposition of the matrix pencil $(\mathbf{A}_r, \mathbf{E}_r)$. Since $\mathbf{A}_r$ and $\mathbf{E}_r$ are symmetric positive definite, we can choose $\mathbf{Q}^T\mathbf{E}_r\mathbf{Q} = \mathbf{I}$. Hence, we have

$$\mathbf{G}_{1,r}(s) = \mathbf{B}_r^T(s\mathbf{E}_r + \mathbf{A}_r)^{-1}\mathbf{B}_r = \mathbf{B}_r^T\mathbf{Q}(\mathbf{Q}^T(s\mathbf{E}_r + \mathbf{A}_r)\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{B}_r = \sum_{i=1}^{n_r} \frac{\mathbf{b}_i\mathbf{b}_i^T}{s+\lambda_i},$$

with $\mathbf{\Lambda} = \operatorname{diag}(\lambda_1, \ldots, \lambda_{n_r})$ and $\mathbf{B}_r^T\mathbf{Q} = [\mathbf{b}_1, \ldots, \mathbf{b}_{n_r}]$. The name of the representation is due to the fact that $\mathbf{b}_i\mathbf{b}_i^T = \operatorname{res}[\mathbf{G}_{1,r}(s), \lambda_i]$. Analogously, let $\mathbf{G}_{2,r}(s)$ be given as

$$\mathbf{G}_{2,r}(s) = \sum_{j=1}^{n_r} \frac{\mathbf{c}_j\mathbf{c}_j^T}{s+\sigma_j},$$

where the σ_j are the eigenvalues of the pencil $(\mathbf{H}_r, \mathbf{M}_r)$ and $\mathbf{c}_j\mathbf{c}_j^T = \operatorname{res}[\mathbf{G}_{2,r}(s), \sigma_j]$. Next, define an objective function via

$$\mathcal{J} := \langle \mathbf{G}_1 - \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2}. \tag{15}$$

For reduced transfer functions obtained within a projection framework, in [?], we have claimed that

$$\mathcal{J} \le \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2} - \langle \mathbf{G}_{1,r}, \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} = \|\mathbf{X} - \tilde{\mathbf{X}}\|_{\mathcal{L}_S}^2,$$

where $\tilde{\mathbf{X}}$ can be consructed by prolongation of the solution $\mathbf{X}_r$ of a reduced Sylvester equation. For the sake of completeness, we give a proof based on the following two results from [?] and [?] (stated here for multi-input multi-output systems) .

Lemma 3 ([?]) *Suppose that* $\mathbf{G}(s)$ *and* $\mathbf{H}(s) = \sum_{i=1}^m \frac{1}{s-\mu_i}\mathbf{c}_i\mathbf{b}_i^T$ *are stable and have simple poles. Then*

$$\langle \mathbf{G}, \mathbf{H} \rangle_{\mathcal{H}_2} = \sum_{i=1}^m \mathbf{c}_i^T\overline{\mathbf{G}}(-\mu_i)\mathbf{b}_i.$$

Lemma 4 ([?]) *Let* $\mathbf{H}(s) = \mathbf{B}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$ *be a symmetric state space system, and let* $\mathbf{H}_r(s) = \mathbf{B}_r^T(s\mathbf{I}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r$ *be any reduced model of* $\mathbf{H}(s)$ *constructed by a compression of* $\mathbf{H}(s)$*, i.e.,* $\mathbf{A}_r = \mathbf{V}^T\mathbf{A}\mathbf{V}, \mathbf{B}_r = \mathbf{V}^T\mathbf{B}$*. Then, for any* $s \ge 0$*,* $\mathbf{H}(s) - \mathbf{H}_r(s) \succeq 0$*.*

Lemma 5 *Let* $\mathbf{G}_1(s) = \mathbf{B}^T(s\mathbf{E} + \mathbf{A})^{-1}\mathbf{B}$ *and* $\mathbf{G}_2(s) = \mathbf{C}^T(s\mathbf{M} + \mathbf{H})^{-1}\mathbf{C}$ *be given transfer functions. Suppose that* $\mathbf{G}_{1,r}(s) = \mathbf{B}_r^T(s\mathbf{E}_r + \mathbf{A}_r)^{-1}\mathbf{B}_r = \sum_{i=1}^{n_r} \frac{\mathbf{b}_i\mathbf{b}_i^T}{s+\lambda_i}$ *and* $\mathbf{G}_{2,r}(s) = \mathbf{C}_r^T(s\mathbf{M}_r + \mathbf{H}_r)^{-1}\mathbf{C}_r = \sum_{j=1}^{n_r} \frac{\mathbf{c}_j\mathbf{c}_j^T}{s+\sigma_j}$ *have been constructed by orthogonal projections*

$$\mathbf{A}_r = \mathbf{V}^T\mathbf{A}\mathbf{V},\ \mathbf{E}_r = \mathbf{V}^T\mathbf{E}\mathbf{V},\ \mathbf{B}_r = \mathbf{V}^T\mathbf{B},$$
$$\mathbf{H}_r = \mathbf{W}^T\mathbf{H}\mathbf{W},\ \mathbf{M}_r = \mathbf{W}^T\mathbf{M}\mathbf{W},\ \mathbf{C}_r = \mathbf{W}^T\mathbf{C}.$$

Then

$$\langle \mathbf{G}_1 - \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} \le \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2} - \langle \mathbf{G}_{1,r}, \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2}.$$

Proof For the $\mathcal{H}_2$-inner product, we find

$$\begin{aligned}\langle \mathbf{G}_1 - \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} = \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2} &- \langle \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} \\ &- \langle \mathbf{G}_{2,r}, \mathbf{G}_1 - \mathbf{G}_{1,r} \rangle_{\mathcal{H}_2} - \langle \mathbf{G}_{1,r}, \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2}.\end{aligned}$$

Applying Lemma 3 to the second term gives

$$-\langle \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} = -\sum_{i=1}^{n_r} \mathbf{b}_i^T \left(\mathbf{G}_2(\lambda_i) - \mathbf{G}_{2,r}(\lambda_i) \right) \mathbf{b}_i.$$

Since $\mathbf{G}_{1,r}$ is constructed by orthogonal projection, it must have stable poles and thus $\lambda_i \ge 0$. Moreover, Lemma 4 yields $\mathbf{G}_2(s) - \mathbf{G}_{2,r}(s) \succeq 0$ which shows that

$$-\langle \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} \le 0.$$

The same argument shows $\langle \mathbf{G}_{2,r}, \mathbf{G}_1 - \mathbf{G}_{1,r} \rangle_{\mathcal{H}_2} \ge 0$ and proves the statement.

In particular, the proof indicates that equality holds for $\left(\mathbf{G}_2(\lambda_i) - \mathbf{G}_{2,r}(\lambda_i) \right) \mathbf{b}_i = \mathbf{0}$ and $\left(\mathbf{G}_1(\sigma_j) - \mathbf{G}_{1,r}(\sigma_j) \right) \mathbf{c}_j = \mathbf{0}$. Again this generalizes our SISO formulation in [**?**]. Moreover, the latter condition is directly related to the gradient of $\mathcal{J}$ with respect to the parameters $\mathbf{b}_i, \lambda_i, \mathbf{c}_i$ and σ_i.

Theorem 6 *Let $\mathbf{G}_1(s), \mathbf{G}_2(s), \mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$ be symmetric state space systems with simple poles. Suppose that $\lambda_1, \ldots, \lambda_{n_r}$ and $\sigma_1, \ldots, \sigma_{n_r}$ are the poles of the reduced transfer functions with $\mathrm{res}[\mathbf{G}_{1,r}(s), \lambda_i] = \mathbf{b}_i \mathbf{b}_i^T$ and $\mathrm{res}[\mathbf{G}_{2,r}(s), \sigma_j] = \mathbf{c}_j \mathbf{c}_j^T$, for $i, j = 1, \ldots, n_r$. The gradient of $\mathcal{J}$ with respect to the parameters listed as*

$$\{\mathbf{b}, \boldsymbol{\lambda}, \mathbf{c}, \boldsymbol{\sigma}\} = [\mathbf{b}_1^T, \lambda_1, \mathbf{c}_1^T, \sigma_1, \ldots, \mathbf{b}_{n_r}^T, \lambda_{n_r}, \mathbf{c}_{n_r}^T, \sigma_{n_r}]^T$$

is given by $\nabla_{\{\mathbf{b}, \boldsymbol{\lambda}, \mathbf{c}, \boldsymbol{\sigma}\}} \mathcal{J}$, a vector of length $2n_r \cdot (q+1)$, partitioned into n_r vectors of length $2(q+1)$ as

$$\left(\nabla_{\{\mathbf{b}, \boldsymbol{\lambda}, \mathbf{c}, \boldsymbol{\sigma}\}} \mathcal{J} \right)_k = \begin{bmatrix} 2 \left(\mathbf{G}_{2,r}(\lambda_k) - \mathbf{G}_2(\lambda_k) \right) \mathbf{b}_k \\ \mathbf{b}_k^T (\mathbf{G}_{2,r}'(\lambda_k) - \mathbf{G}_2'(\lambda_k)) \mathbf{b}_k \\ 2 \left(\mathbf{G}_{1,r}(\sigma_k) - \mathbf{G}_1(\sigma_k) \right) \mathbf{c}_k \\ \mathbf{c}_k^T (\mathbf{G}_{1,r}'(\sigma_k) - \mathbf{G}_1'(\sigma_k)) \mathbf{c}_k \end{bmatrix},$$

for $k = 1, \ldots, n_r$.

Proof Observe that for the ℓ-th entry of $\mathbf{b}_k$, we have

$$\begin{aligned}\frac{\partial \mathcal{J}}{\partial (\mathbf{b}_k)_\ell} &= \frac{\partial}{\partial (\mathbf{b}_k)_\ell} \langle \mathbf{G}_1 - \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} \\ &= -\left\langle \frac{\partial \mathbf{G}_{1,r}}{\partial (\mathbf{b}_k)_\ell}, \mathbf{G}_2 - \mathbf{G}_{2,r} \right\rangle_{\mathcal{H}_2} \\ &= -\left\langle \frac{\mathbf{e}_\ell \mathbf{b}_k^T}{s+\lambda_k}, \mathbf{G}_2 - \mathbf{G}_{2,r} \right\rangle_{\mathcal{H}_2} - \left\langle \frac{\mathbf{b}_k \mathbf{e}_\ell^T}{s+\lambda_k}, \mathbf{G}_2 - \mathbf{G}_{2,r} \right\rangle_{\mathcal{H}_2} \\ &= -\mathbf{e}_\ell^T \left(\mathbf{G}_2(\lambda_k) - \mathbf{G}_{2,r}(\lambda_k) \right) \mathbf{b}_k - \mathbf{b}_k^T \left(\mathbf{G}_2(\lambda_k) - \mathbf{G}_{2,r}(\lambda_k) \right) \mathbf{e}_\ell \\ &= -2 \cdot \mathbf{e}_\ell^T \left(\mathbf{G}_2(\lambda_k) - \mathbf{G}_{2,r}(\lambda_k) \right) \mathbf{b}_k,\end{aligned}$$

where $\mathbf{e}_\ell$ is the ℓ-th unit vector. The previous steps follow from Lemma 3 and the fact that $\mathbf{G}_2$ and $\mathbf{G}_{2,r}$ are symmetric state space systems. Similarly, for the derivative with respect to λ_k, we find

$$\begin{aligned}\frac{\partial \mathcal{J}}{\partial \lambda_k} &= \frac{\partial}{\partial \lambda_k} \langle \mathbf{G}_1 - \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \rangle_{\mathcal{H}_2} \\ &= -\left\langle \frac{\partial}{\partial \lambda_k} \mathbf{G}_{1,r}, \mathbf{G}_2 - \mathbf{G}_{2,r} \right\rangle_{\mathcal{H}_2} \\ &= \left\langle \frac{\mathbf{b}_k \mathbf{b}_k^T}{(s+\lambda_k)^2}, \mathbf{G}_2 - \mathbf{G}_{2,r} \right\rangle_{\mathcal{H}_2}.\end{aligned}$$

For the latter expression, we can use the MIMO analogue of [**?**, Lemma 2.4] and obtain

$$\frac{\partial \mathcal{J}}{\partial \lambda_k} = \mathbf{b}_k^T (\mathbf{G}_{2,r}'(\lambda_k) - \mathbf{G}_2'(\lambda_k)) \mathbf{b}_k.$$

The proofs for $\mathbf{c}_k$ and σ_k use the exact same arguments and are thus omitted here.

Remark Note the change of sign for the derivatives with respect to λ_k and σ_k compared to the special case of $\mathcal{H}_2$-optimal model reduction discussed in [**?**]. This simply follows from a different notation in this manuscript. Using $\lambda_i, \sigma_j < 0$ together with transfer function representations $\sum_{i=1}^{n_r} \mathbf{G}_{1,r}(s) = \frac{\mathbf{b}_i \mathbf{b}_i^T}{s-\lambda_i}$ and $\mathbf{G}_{r,2}(s) = \sum_{j=1}^{n_r} \frac{\mathbf{c}_j \mathbf{c}_j^T}{s-\sigma_j}$ would lead to similar expressions as in [**?**].

In [**?**], we stated the inequality from Lemma 5 and showed that equality holds if the gradient of $\mathcal{J}$ is zero. In fact, we can even show that the corresponding reduced transfer functions can be used to compute a triple $(\mathbf{V}, \mathbf{W}, \mathbf{X}_r)$ satisfying the first-order necessary conditions from Theorem 1.

Theorem 7 *Let* $\mathbf{X}$ *be the solution of the Sylvester equation*

$$\mathbf{AXM} + \mathbf{EXH} = \mathbf{BC}$$

and denote, respectively, $\mathbf{G}_1(s) = \mathbf{B}^T(s\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}$ *and* $\mathbf{G}_2(s) = \mathbf{C}^T(s\mathbf{M}+\mathbf{H})^{-1}\mathbf{C}$. *Suppose* $\mathbf{G}_{1,r}(s) = \sum_{i=1}^{n_r} \frac{\mathbf{b}_i\mathbf{b}_i^T}{s+\lambda_i}$ *and* $\mathbf{G}_{2,r}(s) = \sum_{j=1}^{n_r} \frac{\mathbf{c}_j\mathbf{c}_j^T}{s+\sigma_j}$ *satisfy*

$$\mathbf{G}_{1,r}(\sigma_k)\mathbf{c}_k = \mathbf{G}_1(\sigma_k)\mathbf{c}_k, \tag{16a}$$

$$\mathbf{c}_k^T\mathbf{G}'_{1,r}(\sigma_k)\mathbf{c}_k = \mathbf{c}_k^T\mathbf{G}_1(\sigma_k)\mathbf{c}_k, \tag{16b}$$

$$\mathbf{G}_{2,r}(\lambda_k)\mathbf{b}_k = \mathbf{G}_2(\lambda_k)\mathbf{b}_k, \tag{16c}$$

$$\mathbf{b}_k^T\mathbf{G}'_{2,r}(\lambda_k)\mathbf{b}_k = \mathbf{b}_k^T\mathbf{G}(\lambda_k)\mathbf{b}_k, \tag{16d}$$

for $k = 1,\dots,n_r$. *Define* $\mathbb{X} \in \mathbb{R}^{n_r\times n_r}, \mathbb{Y} \in \mathbb{R}^{n\times n_r}$ *and* $\mathbb{Z} \in \mathbb{R}^{m\times n_r}$ *via*

$$\mathbb{X}_{ij} = \frac{\mathbf{b}_i^T\mathbf{c}_j}{\lambda_i+\sigma_j}, \quad \mathbb{Y}_i = (\sigma_i\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}\mathbf{c}_i, \quad \mathbb{Z}_j = (\lambda_j\mathbf{M}+\mathbf{H})^{-1}\mathbf{C}\mathbf{b}_j. \tag{17}$$

Then the triple $(\mathbb{Y},\mathbb{Z},\mathbb{X}^{-1})$ *satisfies* (11).

Proof First note that (16) defines $n_r(q+1)$ constraints on, respectively, $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$. Due to the pole-residue representation, exactly the same number of parameters defines the rational matrix valued transfer functions $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$. Hence, $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}$ are uniquely determined by (16). Echoing the argumentation in [**?**, Lemma 3.11] and [**?**], w.l.o.g. we can thus assume that the reduced transfer functions are obtained by $\mathbf{E}/\mathbf{M}$-orthogonal projections via

$$\boldsymbol{\Lambda} = \mathbf{V}^T\mathbf{A}\mathbf{V}, \quad \tilde{\mathbf{B}} := [\mathbf{b}_1,\dots,\mathbf{b}_q]^T = \mathbf{V}^T\mathbf{B},$$
$$\boldsymbol{\Sigma} = \mathbf{W}^T\mathbf{H}\mathbf{W}, \quad \tilde{\mathbf{C}} := [\mathbf{c}_1,\dots,\mathbf{c}_q]^T = \mathbf{W}^T\mathbf{C},$$

where $\mathbf{V}$ and $\mathbf{W}$ are s.t.

$$\operatorname{span}\{\mathbf{V}\} \supset \operatorname*{span}_{i=1,\dots,n_r} \left\{(\sigma_i\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}\mathbf{c}_i\right\},$$
$$\operatorname{span}\{\mathbf{W}\} \supset \operatorname*{span}_{j=1,\dots,n_r} \left\{(\lambda_j\mathbf{M}+\mathbf{H})^{-1}\mathbf{C}\mathbf{b}_j\right\}.$$

Due to the definition of $\mathbb{X}_{ij}$ we further obtain:

$$\mathbb{X}_i = (\sigma_i\mathbf{I}+\boldsymbol{\Lambda})^{-1}\tilde{\mathbf{B}}\mathbf{c}_i, \quad \mathbb{X}_j^T = (\lambda_j\mathbf{I}+\boldsymbol{\Sigma})^{-1}\tilde{\mathbf{C}}\mathbf{b}_j,$$

where $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_1,\dots,\lambda_{n_r})$ and $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1,\dots,\sigma_{n_r})$. Using well-known results from projection-based rational interpolation (see [**?**]), we conclude

$$\mathbf{V}\mathbb{X}_i = \mathbb{Y}_i, \quad \mathbf{W}\mathbb{X}_j^T = \mathbb{Z}_j$$

and therefore $\mathbf{V}\mathbb{X} = \mathbb{Y}$ and $\mathbf{W}\mathbb{X}^T = \mathbb{Z}$. Keeping this in mind, for (11) we get

$$\left(\mathbf{A}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^T\mathbf{M} + \mathbf{E}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^T\mathbf{H} - \mathbf{B}\mathbf{C}^T\right)\mathbb{Z}\mathbb{X}^{-T}$$
$$= \left(\mathbf{A}\mathbb{Y}\mathbf{W}^T\mathbf{M} + \mathbf{E}\mathbb{Y}\mathbf{W}^T\mathbf{H} - \mathbf{B}\mathbf{C}^T\right)\mathbf{W}$$
$$= \mathbf{A}\mathbb{Y} + \mathbf{E}\mathbb{Y}\boldsymbol{\Sigma} - \mathbf{B}\tilde{\mathbf{C}}^T = \mathbf{0}.$$

Here, the last step follows from the definition of $\mathbb{Y}$. Similarly, it holds

$$\begin{aligned}&\mathbb{X}^{-T}\mathbb{Y}^{T}\left(\mathbf{A}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^{T}\mathbf{M}+\mathbf{E}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^{T}\mathbf{H}-\mathbf{B}\mathbf{C}^{T}\right)\\&\quad=\mathbf{V}^{T}\left(\mathbf{A}\mathbf{V}\mathbb{Z}^{T}\mathbf{M}+\mathbf{E}\mathbf{V}\mathbb{Z}^{T}\mathbf{H}-\mathbf{B}\mathbf{C}^{T}\right)\\&\quad=\mathbf{\Lambda}\mathbb{Z}^{T}\mathbf{M}+\mathbb{Z}^{T}\mathbf{H}-\tilde{\mathbf{B}}\mathbf{C}^{T}=\mathbf{0}.\end{aligned}$$

Again, the last equality is due to the definition of $\mathbb{Z}$. Finally, we have

$$\begin{aligned}&\mathbb{Y}^{T}\left(\mathbf{A}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^{T}\mathbf{M}+\mathbf{E}\mathbb{Y}\mathbb{X}^{-1}\mathbb{Z}^{T}\mathbf{H}-\mathbf{B}\mathbf{C}^{T}\right)\mathbb{Z}\\&\quad=\mathbb{X}^{T}\mathbf{V}^{T}\left(\mathbf{A}\mathbf{V}\mathbb{X}\mathbf{W}^{T}\mathbf{M}+\mathbf{E}\mathbf{V}\mathbb{X}\mathbf{W}^{T}\mathbf{H}-\mathbf{B}\mathbf{C}^{T}\right)\mathbf{W}\mathbb{X}^{T}\\&\quad=\mathbb{X}^{T}\left(\mathbf{\Lambda}\mathbb{X}+\mathbb{X}\mathbf{\Sigma}-\tilde{\mathbf{B}}\tilde{\mathbf{C}}^{T}\right)\mathbb{X}^{T}=\mathbf{0}.\end{aligned}$$

Once more, the last step is true due to the definition of $\mathbb{X}$.

Remark From the proof of Thereom 7, we find that the same approximation is obtained when $(\mathbb{Y},\mathbb{Z},\mathbb{X}^{-1})$ is replaced by $(\mathbf{V},\mathbf{W},\mathbb{X})$ where $\mathbf{V}$ and $\mathbf{W}$ are the projection matrices constructing $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$. Note further that $\mathbb{X}$ solves the projected reduced Sylvester equation. This in particular implies that the approximation $\mathbf{V}\mathbb{X}\mathbf{W}^{T}$ fulfills the common Galerkin condition on the residual (see [**?**]).

The natural question that arises is wheter triples $(\mathbf{V},\mathbf{W},\mathbf{X}_r)$ fulfilling (11) also yield reduced transfer functions $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$ with vanishing gradient $\nabla_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}$. The answer is given by the following result.

Theorem 8 *Let a triple* $(\mathbf{V},\mathbf{W},\mathbf{X}_r)$ *be given such that* (11) *holds. Suppose reduced transfer functions* $\mathbf{G}_{1,r}(s)$ *and* $\mathbf{G}_{2,r}(s)$ *are defined via*

$$\begin{aligned}\mathbf{A}_r&=\mathbf{V}^{T}\mathbf{A}\mathbf{V}, &\mathbf{E}_r&=\mathbf{V}^{T}\mathbf{E}\mathbf{V}, &\mathbf{B}_r&=\mathbf{V}^{T}\mathbf{B},\\\mathbf{H}_r&=\mathbf{W}^{T}\mathbf{H}\mathbf{W}, &\mathbf{M}_r&=\mathbf{W}^{T}\mathbf{M}\mathbf{W}, &\mathbf{C}_r&=\mathbf{W}^{T}\mathbf{C}.\end{aligned}$$

Then it holds $\nabla_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}=\mathbf{0}$.

Proof The third condition in (11) implies

$$\mathbf{A}_r\mathbf{X}_r\mathbf{M}_r+\mathbf{E}_r\mathbf{X}_r\mathbf{H}_r-\mathbf{B}_r\mathbf{C}_r^{T}=\mathbf{0}.$$

Assuming that $\mathbf{H}_r\mathbf{R}=\mathbf{M}_r\mathbf{R}\mathbf{\Sigma}$ is the eigenvalue decomposition of $(\mathbf{H}_r,\mathbf{M}_r)$, postmultiplication of the above equation with $\mathbf{r}_j:=\mathbf{R}\mathbf{e}_j$ gives

$$\mathbf{A}_r\underbrace{\mathbf{X}_r\mathbf{M}_r\mathbf{r}_j}_{\mathbf{x}_j}+\sigma_j\mathbf{E}_r\underbrace{\mathbf{X}_r\mathbf{M}_r\mathbf{r}_j}_{\mathbf{x}_j}=\mathbf{B}_r\underbrace{\mathbf{C}_r^{T}\mathbf{r}_j}_{\mathbf{c}_j}.$$

Hence, we have $\mathbf{x}_j=(\sigma_j\mathbf{E}_r+\mathbf{A}_r)^{-1}\mathbf{B}_r\mathbf{c}_j$. Also, postmultiplication of the third equation in (11) with $\mathbf{r}_j$ yields

$$\mathbf{A}\mathbf{V}\mathbf{X}_r\mathbf{M}_r\mathbf{r}_j+\sigma_j\mathbf{E}\mathbf{V}\mathbf{X}_r\mathbf{M}_r\mathbf{r}_j=\mathbf{B}\mathbf{C}_r^{T}\mathbf{r}_j.$$

In particular, we conclude $\mathbf{V}\mathbf{x}_j = (\sigma\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}\mathbf{c}_j$. This however yields

$$\mathbf{G}_{1,r}(\sigma_j)\mathbf{c}_j = \mathbf{B}^T\mathbf{V}(\sigma_j\mathbf{E}_r+\mathbf{A}_r)^{-1}\mathbf{B}\mathbf{c}_j = \mathbf{B}^T(\sigma_j\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}\mathbf{c}_j = \mathbf{G}_1(\sigma)\mathbf{c}_j$$

and

$$\begin{aligned}\mathbf{c}_j^T\mathbf{G}_{1,r}'(\sigma_j)\mathbf{c}_j &= -\mathbf{c}_j^T\mathbf{B}_r^T(\sigma_j\mathbf{E}_r+\mathbf{A}_r)^{-1}\mathbf{V}^T\mathbf{E}\mathbf{V}(\sigma_j\mathbf{E}_r+\mathbf{A}_r)^{-1}\mathbf{B}_r\mathbf{c}_j\\ &= -\mathbf{c}_j^T\mathbf{B}^T(\sigma_j\mathbf{E}+\mathbf{A})^{-1}\mathbf{E}(\sigma_j\mathbf{E}+\mathbf{A})^{-1}\mathbf{B}\mathbf{c}_j = \mathbf{c}_j^T\mathbf{G}_1'(\sigma_j)\mathbf{c}_j.\end{aligned}$$

The proof for $\mathbf{G}_{2,r}$ follows analogously.

In summary, we can state that the first-order necessary conditions for the objective functions $f(\mathbf{V},\mathbf{W},\mathbf{X}_r)$ and $\mathcal{J}(\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma})$ are equivalent to each other. For the remainder of this paper, we focus on the objective function $\mathcal{J}$. Along the lines of [?], we present the Hessian of $\mathcal{J}$ with respect to the parameters $\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}$.

Lemma 9 *The Hessian of $\mathcal{J}$ with respect to $\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}$ is given by $\nabla^2_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}$, an $(2n_r\cdot(q+1))\times(2n_r\cdot(q+1))$ matrix partitioned into n_r^2 matrices of size $2(q+1)\times 2(q+1)$ defined by*

$$\begin{aligned}\left(\nabla^2_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}\right)_{k\ell} &= \begin{bmatrix} \mathbf{0} & \mathbf{0} & 2\left(\frac{\mathbf{c}_\ell\mathbf{b}_k^T+\mathbf{c}_\ell^T\mathbf{b}_k\mathbf{I}_q}{\sigma_\ell+\lambda_k}\right) & -2\frac{\mathbf{c}_\ell\mathbf{c}_\ell^T\mathbf{b}_k}{(\sigma_\ell+\lambda_k)^2}\\ \mathbf{0} & 0 & -2\frac{\mathbf{c}_\ell^T\mathbf{b}_k\mathbf{b}_k^T}{(\lambda_k+\sigma_\ell)^2} & 2\frac{\mathbf{b}_k^T\mathbf{c}_\ell\mathbf{c}_\ell^T\mathbf{b}_k}{(\sigma_\ell+\lambda_k)^3}\\ 2\left(\frac{\mathbf{b}_\ell\mathbf{c}_k^T+\mathbf{b}_\ell^T\mathbf{c}_k\mathbf{I}_q}{\sigma_k+\lambda_\ell}\right) & -2\frac{\mathbf{b}_\ell\mathbf{b}_\ell^T\mathbf{c}_k}{(\sigma_k+\lambda_\ell)^2} & \mathbf{0} & \mathbf{0}\\ -2\frac{\mathbf{b}_\ell^T\mathbf{c}_k\mathbf{c}_k^T}{(\lambda_\ell+\sigma_k)^2} & 2\frac{\mathbf{c}_k^T\mathbf{b}_\ell\mathbf{b}_\ell^T\mathbf{c}_k}{(\lambda_\ell+\sigma_k)^3} & \mathbf{0} & 0\end{bmatrix}\\ &+\delta_{k\ell}\begin{bmatrix} 2(\mathbf{G}_{2,r}(\lambda_k)-\mathbf{G}_2(\lambda_k)) & 2(\mathbf{G}_{2,r}'(\lambda_k)-\mathbf{G}_2'(\lambda_k))\mathbf{b}_k & \mathbf{0} & \mathbf{0}\\ 2\mathbf{b}_k^T(\mathbf{G}_{2,r}'(\lambda_k)-\mathbf{G}_2'(\lambda_k)) & \mathbf{b}_k^T\left(\mathbf{G}_{2,r}''(\lambda_k)-\mathbf{G}_2''(\lambda_k)\right)\mathbf{b}_k & \mathbf{0} & 0\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & 0 & \mathbf{0} & 0\end{bmatrix}\\ &+\delta_{k\ell}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & 0 & \mathbf{0} & 0\\ \mathbf{0} & \mathbf{0} & 2(\mathbf{G}_{1,r}(\sigma_k)-\mathbf{G}_1(\sigma_k)) & 2(\mathbf{G}_{1,r}'(\sigma_k)-\mathbf{G}_1'(\sigma_k))\mathbf{c}_k\\ \mathbf{0} & 0 & 2\mathbf{c}_k^T(\mathbf{G}_{1,r}'(\sigma_k)-\mathbf{G}_1'(\sigma_k)) & \mathbf{c}_k^T\left(\mathbf{G}_{1,r}''(\sigma_k)-\mathbf{G}_1''(\sigma_k)\right)\mathbf{c}_k\end{bmatrix}.\end{aligned}$$

The proof follows by direct computation of the partial derivatives. Since a similar derivation can be found in [?] for the $\mathcal{H}_2$-optimal case, we omit the details.

Unfortunately, the objective function $\mathcal{J}$ is unbounded so that its minimization is not well defined. This can be seen by considering $n_r = 1$. In this case,

$$\mathbf{G}_{1,r}(s) = \frac{\mathbf{b}\mathbf{b}^T}{s-\lambda} \quad\text{and}\quad \mathbf{G}_{2,r}(s) = \frac{\mathbf{c}\mathbf{c}^T}{s-\mu},$$

are the reduced transfer functions. By Lemma 3, for the objective function we get

$$\mathcal{J} = \langle\mathbf{G}_1,\mathbf{G}_2\rangle_{\mathcal{H}_2} - \mathbf{b}^T\mathbf{G}_2(\lambda)\mathbf{b} - \mathbf{c}^T\mathbf{G}_1(\mu)\mathbf{c} + \frac{\mathbf{b}^T\mathbf{c}\mathbf{c}^T\mathbf{b}}{\lambda+\mu}.$$

Hence, by scaling $\alpha\mathbf{b}$ and $\frac{1}{\alpha}\mathbf{c}$, we further obtain

$$\mathcal{J} = \langle \mathbf{G}_1, \mathbf{G}_2 \rangle_{\mathcal{H}_2} - \alpha^2 \mathbf{b}^T \mathbf{G}_2(\lambda)\mathbf{b} - \frac{1}{\alpha^2}\mathbf{c}^T \mathbf{G}_1(\mu)\mathbf{c} + \frac{\mathbf{b}^T \mathbf{c}\mathbf{c}^T \mathbf{b}}{\lambda + \mu}.$$

and we can arbitrarily decrease the value of $\mathcal{J}$ by increasing α. In fact, a similar conclusion can be drawn from the Hessian in Theorem 9. Multiplication of $\left(\nabla^2_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}\right)_{11}$ with $\mathbf{z} := \begin{bmatrix}\alpha\mathbf{b}_1^T & 0 & \mathbf{c}_1^T & 0\end{bmatrix}^T$ yields

$$\begin{aligned}\mathbf{z}^T \left(\nabla^2_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}\right)_{11} \mathbf{z} &= 2\alpha^2 \mathbf{b}_1^T(\mathbf{G}_{2,r}(\lambda_1) - \mathbf{G}_2(\lambda_1))\mathbf{b}_1 + 2\mathbf{c}_1^T(\mathbf{G}_{1,r}(\sigma_1) - \mathbf{G}_2(\sigma_1))\mathbf{c}_1 \\ &\quad + 8\alpha \frac{(\mathbf{b}_1^T \mathbf{c}_1)^2}{\sigma_1 + \lambda_1}.\end{aligned}$$

For a stationary point, we thus find

$$\mathbf{z}^T \left(\nabla^2_{\{\mathbf{b},\boldsymbol{\lambda},\mathbf{c},\boldsymbol{\sigma}\}}\mathcal{J}\right)_{11} \mathbf{z} = 8\alpha \frac{(\mathbf{b}_1^T \mathbf{c}_1)^2}{\sigma_1 + \lambda_1}.$$

In other words, the Hessian is always indefinite and, consequently, all stationary points are saddle points. While this will cause problems for optimization routines, we can still extend the idea of iterative correction as in [**?**] to the MIMO Sylvester case. Algorithm 1 is a suitable generalization of a SISO version we proposed in [**?**]. Due to the iterative structure, upon convergence, the reduced transfer functions $\mathbf{G}_{1,r}(s)$ and $\mathbf{G}_{2,r}(s)$ will tangentially interpolate the original transfer function $\mathbf{G}_1(s)$ and $\mathbf{G}_2(s)$ such that the corresponding gradient in Lemma 6 vanishes. According to Theorem 7 this way we can compute stationary points of the objective function f which is obviously bounded.

Initialization

The efficiency of Algorithm 1 obviously depends on the number of iterations needed until a typical convergence tolerance is satisfied. Hence, an imporant point is the initialization of the algorithm. While in general several strategies for choosing interpolation points and tangential directions may be possible, there exists a natural choice for our applications. Recall that the blurred and noisy image is given as the right hand side $\mathbf{G} = \mathbf{B}\mathbf{C}^T$. Though $\mathbf{G}$ deviates from the original unperturbed image, it still is related to it. In other words, $\mathbf{G}$ can be seen as a (rough) approximation to the solution $\mathbf{X}$ of the underlying Sylvester equation. For this reason, if we are interested in constructing an approximation of rank n_r, we propose to use a truncated singular value decomposition of $\mathbf{G} \approx \mathbf{U}_{n_r}\mathbf{D}_{n_r}\mathbf{Z}_{n_r}^T$, with $\mathbf{U}_{n_r} \in \mathbb{R}^{n \times n_r}, \mathbf{Z} \in \mathbb{R}^{m \times n_r}$ and $\mathbf{D}_{n_r} \in \mathbb{R}^{n_r \times n_r}$. Since $\mathbf{U}_{n_r}^T \mathbf{U}_{n_r} = \mathbf{I}$ and $\mathbf{Z}_{n_r}^T \mathbf{Z}_{n_r} = \mathbf{I}$, we can construct an initial reduced model via

$$\begin{aligned}&\mathbf{A}_r = \mathbf{U}_{n_r}^T \mathbf{A}\mathbf{U}_{n_r},\ \mathbf{E}_r = \mathbf{U}_{n_r}^T \mathbf{E}\mathbf{U}_{n_r},\ \mathbf{B}_r = \mathbf{U}_{n_r}^T \mathbf{B}, \\ &\mathbf{H}_r = \mathbf{V}_{n_r}^T \mathbf{H}\mathbf{V}_{n_r},\ \mathbf{M}_r = \mathbf{V}_{n_r}^T \mathbf{M}\mathbf{V}_{n_r},\ \mathbf{C}_r = \mathbf{V}_{n_r}^T \mathbf{C}.\end{aligned}$$

Algorithm 1: MIMO $(\text{Sy})^2$IRKA

Input: Interpolation points: $\{\lambda_1, \ldots, \lambda_{n_r}\}$ and $\{\sigma_1, \ldots, \sigma_{n_r}\}$.
Tangential directions: $\tilde{\mathbf{B}} = [\mathbf{b}_1, \ldots, \mathbf{b}_{n_r}]$ and $\tilde{\mathbf{C}} = [\mathbf{c}_1, \ldots, \mathbf{c}_{n_r}]$.

Output: $\mathbf{G}_{1,r}(s), \mathbf{G}_{2,r}(s)$ satisfying (16)

1: **while** relative change in $\{\lambda_i, \sigma_i\} > \text{tol}$ **do**

2: Compute $\mathbf{V}$ and $\mathbf{W}$ from

$$\text{span}\{\mathbf{V}\} \supset \underset{i=1,\ldots,n_r}{\text{span}} \left\{(\sigma_i\mathbf{E} + \mathbf{A})^{-1}\mathbf{B}\mathbf{c}_i\right\},$$
$$\text{span}\{\mathbf{W}\} \supset \underset{j=1,\ldots,n_r}{\text{span}} \left\{(\lambda_j\mathbf{M} + \mathbf{H})^{-1}\mathbf{C}\mathbf{b}_j\right\}.$$

3: Compute $\mathbf{G}_{1,r}(s) = \mathbf{B}^T\mathbf{V}(\mathbf{V}^T(s\mathbf{E} + \mathbf{A})\mathbf{V})^{-1}\mathbf{V}^T\mathbf{B}$.

4: Compute $\mathbf{G}_{2,r}(s) = \mathbf{C}^T\mathbf{W}(\mathbf{W}^T(s\mathbf{M} + \mathbf{H})\mathbf{W})^{-1}\mathbf{W}^T\mathbf{C}$.

5: Compute $\mathbf{A}_r\mathbf{Q} = \mathbf{E}_r\mathbf{Q}\mathbf{\Lambda}$ with $\mathbf{Q}^T\mathbf{E}_r\mathbf{Q} = \mathbf{I}$.

6: Compute $\mathbf{H}_r\mathbf{R} = \mathbf{M}_r\mathbf{R}\mathbf{\Sigma}$ with $\mathbf{R}^T\mathbf{M}_r\mathbf{R} = \mathbf{I}$.

7: Update $\lambda_i = \text{diag}(\mathbf{\Lambda})$, $\tilde{\mathbf{B}} = \mathbf{B}_r^T\mathbf{Q}$, $\sigma_i = \text{diag}(\mathbf{\Sigma})$, $\tilde{\mathbf{C}} = \mathbf{C}_r^T\mathbf{R}$.

8: **end while**

Initial interpolation points and tangential directions then can be obtained by computing the pole-residue representations for $\mathbf{G}_{1,r}(s) = \mathbf{B}_r^T(s\mathbf{E}_r + \mathbf{A}_r)^{-1}\mathbf{B}_r$ and $\mathbf{G}_{2,r}(s) = \mathbf{C}_r^T(s\mathbf{M}_r + \mathbf{H}_r)^{-1}\mathbf{C}_r$. In all our numerical examples, we initialize Algorithm 1 by this procedure. Moreover, as we mentioned earlier, the right hand side $\mathbf{G}$ is not necessarily low rank and we thus have to face transfer functions with a large number of inputs and outputs. In the case of $\mathcal{H}_2$-optimal model reduction, this can slow down the convergence of iterative algorithms such as IRKA significantly, see [**?**]. For this reason, in our examples we replace $\mathbf{G}$ by its truncated singular value decomposition which is of rank n_r. While this means we are actually approximating the solution of a perturbed Sylvester equation, we will see that this does not seem to influence the quality of the restored image.

5 Numerical results

We study the performance of Algorithm 1 for two examples from image restoration. Again, we follow the setting in [**?**, **?**]. Given an original image $\mathbf{X}$, we use out-of-focus-blur (5) and atmospheric blur (6) to construct a blurred image $\hat{\mathbf{G}}$. The final degraded image $\mathbf{G}$ then is obtained by adding Gaussian white noise $\mathbf{N}$ to $\hat{\mathbf{G}}$ such that $\frac{\|\mathbf{N}\|}{\|\hat{\mathbf{G}}\|} = 10^{-2}$.

All simulations were generated on an Intel®Core™i5-3317U CPU, 3 GB RAM, Ubuntu Linux 12.10, MATLAB® Version 7.14.0.739 (R2012a) 64-bit (glnxa64).

Lothar Reichel

Due to the dedication of this special issue, the first example is an image showing Lothar Reichel.[1] The matrix $\mathbf{X} \in \mathbb{R}^{363\times 400}$ contains grayscale pixel values from the interval $[0, 255]$. The blurring matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ in (4) are Toeplitz matrices as in (5) and (6). First, we construct $\mathbf{H}_1$ with $r_1 = 5$ and $\mathbf{H}_2$ with $\sigma = 7$ and $r_2 = 2$. For the regularization operators we use discrete first-order derivatives such that

$$\mathbf{L}_1 = \begin{bmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \\ & & & 0 \end{bmatrix}, \quad \mathbf{L}_2 = \begin{bmatrix} 1 & & & \\ -1 & \ddots & & \\ & \ddots & 1 & \\ & & -1 & 0 \end{bmatrix}.$$

In Figure 1d we show the results obtained from Algorithm 1 for $n_r = 40$. We obtain a relative change less than 10^{-2} after 10 iterations. Recall that we also approximate the degraded image $\mathbf{G}$ by a low rank matrix of rank 40. We compare our result with the reconstructed image obtained by solving the Sylvester equation exactly by means of the Bartels-Stewart algorithm (1c). For both variants, the optimal value of the regularization parameter λ_{opt} is computed by minimization over a logarithmically spaced interval $[10, 10^{-3}]$ with 20 points. Figure 1 shows that the quality of the approximately reconstructed image is similar to that of the exactly reconstructed image. Actually, in terms of the relative spectral norm error, Algorithm 1 (0.0185) outperforms the full solution (0.1260).

Figure 2 shows similar results for different blurring matrices. Here, we choose $r_1 = 6, \sigma = 12, r_2 = 6$. While the quality of the reconstructed images clearly is worse than in the first setting, Algorithm 1 obviously yields far better results than we obtain by solving the Sylvester equation explicitly. Moreover, the final (energy norm optimal) iterate from Algorithm 1 is found after 20 iteration steps.

Magdeburg cathedral

The second example is an image from the cathedral in Magdeburg, Germany.[2] The matrix $\mathbf{X}$ is of size 436×556. We choose $r_1 = 4, \sigma = 7$ and $r_2 = 5$. Since the Sylvester equation is larger than in the first example, we increase the rank of the approximation to $n_r = 50$. Figure 3 shows a similar comparison as in the first example. Algorithm 1 needs 19 steps before convergence is obtained. Again, the relative spectral norm error for the approximate solution (0.018) is smaller than for the exact solution (2.890). We get similar results for the paramter values $r_1 = 5, \sigma = 7$ and $r_2 = 2$. The results are shown in Figure 4. The number of iterations needed in Algorithm 1 is 13. Once more, note that the method used for reconstruction is probably not the most sophisticated and explains the modest quality of the approximations. Still, we point out that the reconstructed images computed by an approximate solution of the Sylvester equation

[1] The photo is taken from `http://owpdb.mfo.de/detail?photo_id=3467` .

[2] The photo is taken from `http://commons.wikimedia.org/wiki/File:Magdeburger_Dom_Seitenansicht.jpg`.

(a) Original image (b) Blurred and noisy image

(c) Restored image (exact), $\lambda_{\text{opt}} = 0.079$ (d) Restored image (appr.), $\lambda_{\text{opt}} = 0.127$

Figure 1: Uniform blur ($r_1 = 5$) and atmospheric blur ($\sigma = 7, r_2 = 2$) for $n_r = 40$.

in all cases perform better than the actual exact solution. This might be due to the badly conditioned matrices which may cause numerical perturbations when one tries to compute the full solution explicitly.

6 Conclusions

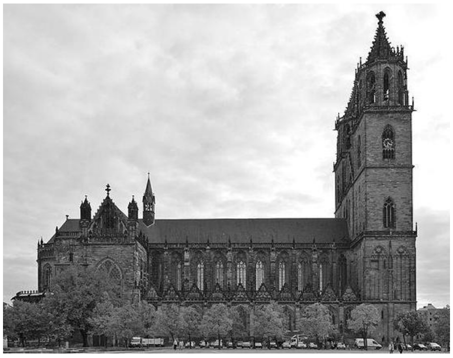
In this paper, we have studied symmetric Sylvester equations arising in image reconstruction problems. The symmetric structure of the equation allowed to measure errors of low rank approximations in terms of an energy norm induced by the Sylvester operator. For given rank n_r, we have derived first-order optimality conditions for an approximation optimal with respect to this energy norm. We have then established a connection to the $\mathcal{H}_2$-inner product of two symmetric state space systems. The

(a) Original image (b) Blurred and noisy image

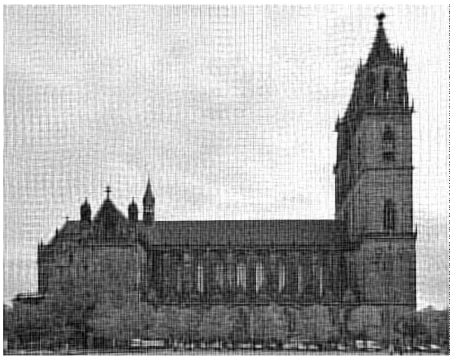
(c) Restored image (exact), $\lambda_{\text{opt}} = 1.438$ (d) Restored image (appr.), $\lambda_{\text{opt}} = 0.127$.

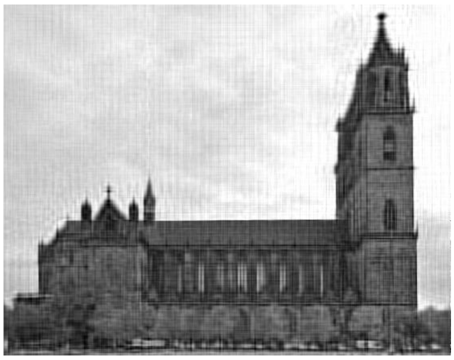
Figure 2: Uniform blur ($r_1 = 6$) and atmospheric blur ($\sigma = 12, r_2 = 6$) for $n_r = 40$.

corresponding first-order optimality conditions have been shown to be equivalent to the ones related to the energy norm minimization problem. The stationary points of the $\mathcal{H}_2$-inner product itself have been shown to be necessarily saddle points. An iterative interpolatory procedure trying to find these saddle points has been suggested. Several numerical examples demonstrated the applicability of the method for image reconstruction problems.



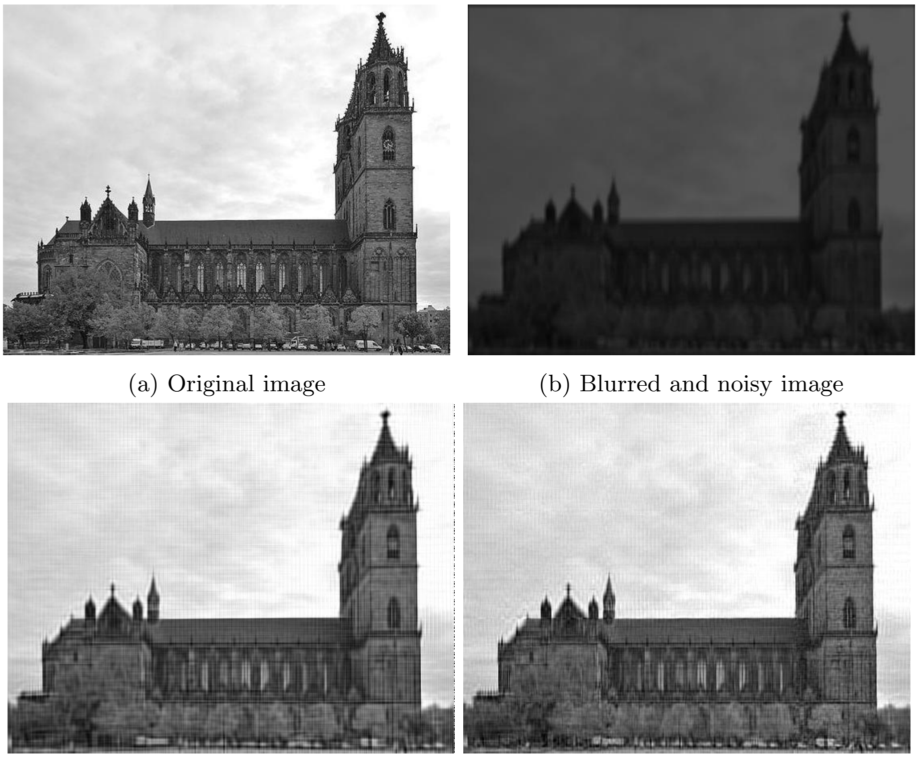
(a) Original image (b) Blurred and noisy image





(c) Restored image (exact), $\lambda_{\text{opt}} = 0.070$ (d) Restored image (appr.), $\lambda_{\text{opt}} = 0.298$

Figure 3: Uniform blur ($r_1 = 4$) and atmospheric blur ($\sigma = 7, r_2 = 5$) for $n_r = 50$.



(a) Original image

(b) Blurred and noisy image

(c) Restored image (exact), $\lambda_{\text{opt}} = 0.336$

(d) Restored image (appr.), $\lambda_{\text{opt}} = 0.055$

Figure 4: Uniform blur ($r_1 = 5$) and atmospheric blur ($\sigma = 7, r_2 = 2$) for $n_r = 50$.