Superoptimization - Part I : An explicit state-space approach to the one-block super-optimal distance problem

J. Kiskiras\textsuperscript{1}, I.M. Jaimoukha\textsuperscript{2} and G.D. Halikias\textsuperscript{1}

1 Abstract

An explicit state-space approach is presented for solving the super-optimal Nehari-extension problem. The approach is based on the all-pass dilation technique developed in [JL93] which offers considerable advantages compared to traditional methods relying on a diagonalisation procedure via a Schmidt pair of the Hankel operator associated with the problem. As a result, all derivations presented in this work rely only on simple linear-algebraic arguments. Further, when the simple structure of the one-block problem is taken into account, this approach leads to a detailed and complete state-space analysis which clearly illustrates the structure of the optimal solution and allows for the removal of all technical assumptions (minimality, multiplicity of largest Hankel singular value, positive-definiteness of the solutions of certain Riccati equations) made in previous work [LHG89],[HLG93]. The advantages of the approach are illustrated with a numerical example. Finally, the paper presents a short survey of super-optimization, the various techniques developed for its solution and some of its applications in the area of modern robust control.

Keywords: super-optimal Nehari-extension problems, Hankel operator, all-pass dilations, $\mathcal{H}_\infty$ - optimal control, maximally robust stabilization.

\textsuperscript{1}Control Engineering Research Centre, School of Engineering and Mathematical Sciences, City University, Northampton Square, London EC1V 0HB, U.K.
\textsuperscript{2}Control and Power Group, Department of Electrical and Electronic Engineering, Imperial College, London SW8 2BT, U.K.
2 Notation

Here we define the main notation used in the paper. Additional notation is introduced in subsequent sections as needed. All systems considered in this paper are assumed linear, time invariant and finite dimensional. Let $\mathcal{R}^{p \times m}(s)$ denote the space of proper $p \times m$ rational matrix functions in $s$ with real coefficients. Associated with $P(s) \in \mathcal{R}^{p \times m}(s)$ of McMillan degree $n$ is a state-space realization:

$$P(s) = C(sI - A)^{-1}B + D$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ and $D \in \mathbb{R}^{p \times m}$. For $P(s) \in \mathcal{R}(s)^{p \times m}$ let $P(s)^\sim := P'(-\bar{s})$ denotes the para-hermitian conjugate of $P(s)$. Let $P(s)$ be partitioned in $2 \times 2$ sub-blocks $P_{ij}(s)$, $i = 1, 2$, $j = 1, 2$. Then a state space realization of $P(s)$ can be written as:

$$P(s) := \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix} = \begin{bmatrix} A & B_1 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix}$$

and

$$P_{ij} = C_i(sI - A)^{-1}B_j + D_{ij}$$

is a state-space realization of $P_{ij}(s)$. A lower linear fractional transformation of $P(s)$ and $K(s)$ is defined as

$$\mathcal{F}_l(P, K) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21}$$

where $K(s)$ is of dimension $m \times p$ if $P_{22}(s)$ has dimension $p \times m$ and the indicated inverse exists. Similarly we define the upper linear fractional transformation of $P(s)$ and $K(s)$ as:

$$\mathcal{F}_u(P, K) = P_{22} + P_{21}K(I - P_{11}K)^{-1}P_{12}$$

for a compatible partitioning of $P(s)$ with $K(s)$ and provided that the indicated inverse exists. The space $\mathcal{RL}_\infty$ consists of all proper real-rational transfer matrix functions which are analytic on the imaginary axis. $\mathcal{RH}^+_{\infty}$ and $\mathcal{RH}^-_{\infty}$ are the subspaces of $\mathcal{RL}_\infty$ consisting of all real-rational proper matrix functions which are analytic in the closed right-half plane and closed left-half plane, respectively. Thus $\mathcal{RL}_\infty = \mathcal{RH}^+_{\infty} \oplus \mathcal{RH}^-_{\infty}$ where $\oplus$ denotes direct sum of subspaces. The norm $\| \cdot \|_\infty$ denotes either the $L_\infty$-norm of a function in $L_\infty$ or the $H_\infty$-norm of a function in $H_\infty$, depending on context. $\mathcal{RH}_\infty(k)$ is the subset of $\mathcal{RL}_\infty$ consisting of all functions with no more
than \( k \) poles in the right-half plane. If \( \Gamma \) is an operator, then \( \|\Gamma\| \) denotes its induced norm. Here we make use of the induced norm of the Hankel operator with symbol \( G \) defined in section 4, which will also be denoted as \( \sigma_1(\Gamma G) \) or as \( \|G\|_H \), where \( \sigma_1 \) denotes the largest singular value of \( \Gamma G \). A square matrix function \( G \in \mathcal{RL}_\infty^m \) is called \( \gamma \)-allpass if \( GG^\sim = G^\sim G = \gamma^2 I \). A square all-pass function with \( \gamma = 1 \) is called inner if it lies in \( \mathcal{RH}_\infty^+ \) and anti-inner if it lies in \( \mathcal{RH}_\infty^- \).

For a matrix \( A \in \mathbb{R}^{m \times n} \) we define \( \mathcal{R}(A) \) to be the range of \( A \) and \( \mathcal{N}(A) \) the null-space (kernel) of \( A \), respectively. \( \mathcal{R}(A) \) and \( \mathcal{N}(A) \) are subspaces of \( \mathcal{R}^m \) and \( \mathcal{R}^n \), respectively, whose corresponding dimensions are denoted as \( \text{rank}(A) \) and \( \text{null}(A) \). For a square matrix \( A \), \( \lambda(A) \) denotes the spectrum of \( A \), i.e. the set of its eigenvalues, and \( \rho(A) \) is the spectral radius of \( A \).

The acronyms ARE, CIF, LFT, LTI and SODP stand for \textit{algebraic Riccati equation}, \textit{complementary inner factorization}, \textit{linear fractional transformation}, \textit{linear time-invariant} and \textit{super-optimal distance problem}, respectively.

## 3 Introduction

In Nehari approximation problems we seek to minimize

\[
\inf_{Q \in \mathcal{H}_\infty^+ \times \mathbb{R}^{m \times m}} \|R + Q\|_\infty
\]

where \( R \in \mathcal{RL}_\infty^p \times m \) (or \( R \in \mathcal{RH}_\infty^p \times m \) without loss of generality). Throughout this paper we study the matrix case \( \min(p, m) > 1 \). Further, depending on the kind of application \( Q \) may be further constrained to have a zero block row and/or column. Then the problem is said to be a \textit{two-block} or a \textit{four-block} distance problem. In this work only \textit{one-block} problems are considered.

By introducing the new notation \( s_1^\infty(R) = \|R\|_\infty \) the approximation problem posed in (1) above can be rewritten as:

\[
s_1(R) := \inf_{Q \in \mathcal{H}_\infty^+ \times \mathbb{R}^{m \times m}} s_1^\infty(R + Q)
\]

where \( s_1(R) \) will be referred to as the optimal level of \( R \). The set of all optimal approximations of \( R \) is defined by

\[
S_1(R) := \{Q \in \mathcal{H}_\infty^+ \times \mathbb{R}^{m \times m} : s_1^\infty(R + Q) = s_1 \}
\]
Note that $s_1(R) := \sigma_1(R^\sim)$, the Hankel norm of $R^\sim$. Since, in general, the solution of this problem is not unique, we can define a stronger version of optimality, by requiring that the sequence of the suprema (taken over $\omega \in \mathbb{R} \cup \{\infty\}$) of all singular values of the “error” system $(R+Q)(j\omega)$ is minimized lexicographically. This stronger version of the problem was first proposed by Young and was defined as super-optimization. The main motivation, arising from esthetic considerations, was to restore uniqueness to the solution of the matrix Nehari problem, by showing in [You86] the existence of a unique super-optimal approximation $Q_{\text{sup}}$. Nevertheless, in the present work and also others (e.g. [PF85]) it is argued that super-optimization fits naturally within the modern robust control-theoretic framework, and can be used to define hierarchical optimization problems in which additional performance and stability objectives can be addressed [PF85], [GHJ00].

3.1 Background theory: Hankel operators and Schmidt vectors

In this section we first give some preliminary definitions related to Hankel operators and their Schmidt vectors and outline some of their properties related to super-optimal approximations. This preliminary material is based on [ZDG96], [GL95] and [Fra87]. Given $G \in \mathcal{RL}_\infty^{p \times m}$, the Hankel operator with symbol $G$ is defined as:

$$
\Gamma_G = P_+ M_G|_{\mathcal{H}_2^\perp} : \mathcal{H}_2^\perp \to \mathcal{H}_2,
\quad \Gamma_G f := (P_+ M_G)f = P_+ (Gf) \quad \text{for} \quad f \in \mathcal{H}_2^\perp
$$

Here $M_G$ denotes the multiplication operator and $P_+ $, $P_-$ denote the orthogonal projections from $\mathcal{L}_2$ to $\mathcal{H}_2$ and $\mathcal{H}_2^\perp$, respectively. Since $G \in \mathcal{RL}_\infty$ is analytic on a vertical strip containing the imaginary axis, we can define its two-sided Laplace transform, $g(t) \in \mathcal{L}_2(-\infty, \infty)$, containing both causal and anti-causal parts. Here $\mathcal{L}_2(-\infty, \infty)$ denotes the space of all square-integrable functions with support $(\infty, \infty)$. The equivalent definition of the Hankel operator in the time-domain is:

$$
\Gamma_g : \mathcal{L}_2(-\infty, 0] \to \mathcal{L}_2[0, \infty), \quad \Gamma_g f = P_+(g*f), \quad \text{for} \quad f \in \mathcal{L}_2(-\infty, 0]
$$

where $*$ denotes convolution. Thus

$$(\Gamma_g f)(t) = \left\{ \begin{array}{ll} 
\int_{-\infty}^0 g(t-\tau)f(\tau)d\tau & t \geq 0 \\
0 & t < 0
\end{array} \right.$$ 

Clearly, the anti-causal part of the “impulse response” of $G(s)$ does not affect $(\Gamma_g f)(t)$, and hence it can be assumed without loss of generality that
$G(s) \in \mathcal{RH}_\infty$ with $G(\infty) = 0$. Further, due to the isometric isomorphism property between the $L_2$ spaces in the time and frequency domains [ZDG96], $\|G\| = \|\Gamma_g\|$ and we can use the definitions of the Hankel operator in the two domains interchangeably. It further follows from the definition that the Hankel operator may be written as the composition of the controllability and observability operators, defined via a state-space realization of $G = (A,B,C)$ (assumed minimal without loss of generality) as

$$\Psi_c : \mathcal{L}_2[-\infty,0] \to \mathcal{R}^n, \quad \Psi_c u := \int_{-\infty}^0 e^{-At} Bu(t) \, dt$$

and

$$\Psi_o : \mathcal{R}^n \to \mathcal{L}_2[0,\infty), \quad \Psi_o x_0 := Ce^{-At}x_0, \quad t \geq 0$$

where $n$ denotes the dimension of $A$, i.e. $\Gamma_G = \Psi_o \Psi_c$. Thus $\Gamma_g$ may be thought as the operator mapping “past” inputs $u(t)$ to “future” outputs $y(t)$ via the initial state $x_0$ [Glo84]. The adjoint operator of $\Gamma_G$ can be shown to be [ZDG96]:

$$\Gamma_g^* : \mathcal{H}_2 \to \mathcal{H}_2^\perp, \quad \Gamma_g^* = P_j \bar{M}_G$$

and hence [ZDG96],

$$\Gamma_g^* = (\Psi_o \Psi_c)^* = \Psi_c^* \Psi_o^* : \mathcal{L}_2[0,\infty) \to \mathcal{L}_2(-\infty,0]$$

where $\Psi_c^*$ and $\Psi_o^*$ denote the adjoint operators of $\Psi_c$ and $\Psi_o$, respectively:

$$\Psi_c^* : \mathcal{R} \to \mathcal{L}_2(-\infty,0], \quad \Psi_c^* x_0 = B'e^{-A't}x_0, \quad \tau \leq 0$$

and

$$\Psi_o^* : \mathcal{L}_2[0,\infty) \to \mathcal{R}^n, \quad \Psi_o^* y(t) = \int_{0}^{\infty} e^{A't} C'y(t) \, dt, \quad t \geq 0$$

Now

$$\Psi_c \Psi_c^* x_0 = \left( \int_{-\infty}^{0} e^{-At} BB'e^{-A't} \, d\tau \right) x_0 = \left( \int_{0}^{\infty} e^{A't} BB'e^{A't} \, d\tau \right) x_0 := P x_0$$

where $P$ is the controllability gramian of the pair $(A,B)$, which satisfies the Lyapunov equation

$$AP + PA' + BB' = 0$$

Thus $P$ is the matrix representation of $\Psi_c \Psi_c^*$. Similarly,

$$\Psi_o \Psi_o^* x_0 = \left( e^{A't} C'C e^{-At} \, dt \right) x_0 := Q x_0$$
where \( Q \) is the observability gramian of the pair \((A, C)\), which satisfies the Lyapunov equation
\[
A'Q + QA + C'C = 0
\]
Now the operators \( \Gamma^*_g \Gamma_g \) and \( \Gamma_g \Gamma^*_g \) have matrix representations \( \Gamma^*_g \Gamma_g = \Psi^*_c \Psi^*_o \Psi_o \Psi_c \) and \( \Gamma_g \Gamma^*_g = \Psi_o \Psi_c \Psi^*_c \Psi^*_o \), respectively. Thus their non-zero eigenvalues satisfy:
\[
\lambda_i(\Gamma^*_g \Gamma_g) = \lambda_i(\Gamma_g \Gamma^*_g) = \lambda_i(\Psi^*_c \Psi^*_o \Psi_o \Psi_c) = \lambda_i(\Psi_o \Psi_c \Psi^*_c \Psi^*_o) = \lambda_i(PQ) =: \sigma^2_i(\Gamma_G)
\]
The \( \sigma_i(\Gamma_G) \)'s are the singular values of \( \Gamma_G \) (Hankel singular values of \( G \)). Let these be ordered as \( \sigma_1 = \ldots = \sigma_r > \sigma_{r+1} \geq \ldots \geq \sigma_n > 0 \) where \( n \) is the McMillan degree of \( G \). Then, \( \sigma_1 = \|\Gamma_G\| \) is the Hankel norm of \( G \). Next, let \( u_i(t) \in \mathcal{L}_2(-\infty,0], u_i(t) \neq 0 \), be an eigenvector of \( \Gamma^*_g \Gamma_g \) corresponding to the eigenvalue \( \sigma^2_i \). Then
\[
\Gamma^*_g \Gamma_g u_i = \Psi^*_c \Psi^*_o \Psi_o \Psi_c u_i = \sigma^2_i u_i
\]
Pre-multiplying by \( \Psi_c \) and defining \( x_i = \Psi_c u_i \in \mathbb{R}^n \) gives \( PQ x_i = \sigma^2_i x_i \). Define \( v_i = (1/\sigma_i)\Gamma_g u_i \in \mathcal{L}_2[0,\infty) \). Then the pair \((u_i, v_i)\) satisfies
\[
\Gamma_g u_i = \sigma_i v_i \text{ and } \Gamma^*_g v_i = \sigma_i u_i
\]
and is called a Schmidt pair of \( \Gamma_G \). Thus
\[
u_i(t) = \Psi^*_c \left( \frac{1}{\sigma_i} Q x_i \right) = \sigma^{-1}_i B' e^{-A' t} Q x_i \in \mathcal{L}_2(-\infty,0]
\]
and
\[
v_i(t) = \Psi_o x_i = C e^{A t} x_i \in \mathcal{L}_2[0,\infty)
\]
Let \( \{u_1, u_2, \ldots, u_r\} \) and \( \{v_1, v_2, \ldots, v_r\} \) be a collection of \( r \) linearly independent eigenvectors of \( \Gamma^*_g \Gamma_g \) and \( \Gamma_g \Gamma^*_g \), respectively, corresponding to the eigenvalue \( \sigma^2_1 \). Then
\[
U(t) = \begin{bmatrix} u_1 & \ldots & u_r \end{bmatrix} (t) = \sigma^{-1}_1 B' e^{-A' t} Q \begin{bmatrix} x_1 & \ldots & x_r \end{bmatrix} \in \mathcal{L}_2^{m \times r}(-\infty,0]
\]
and
\[
V(t) = \begin{bmatrix} v_1 & \ldots & v_r \end{bmatrix} (t) = C e^{A t} \begin{bmatrix} x_1 & \ldots & x_r \end{bmatrix} \in \mathcal{L}_2^{p \times r}[0,\infty)
\]
Taking the (bilateral) Laplace transform shows that
\[
U(s) = -B'(sI + A)^{-1} \Xi \in \mathcal{RH}_2^{1+m \times r}, \quad \Xi = \sigma^{-1}_1 Q \begin{bmatrix} x_1 & x_2 & \ldots & x_r \end{bmatrix}
\]
and
\[ V(s) = C(sI - A)^{-1}\Theta \in \mathcal{H}_2^{p \times r}, \quad \Theta = \begin{bmatrix} x_1 & x_2 & \ldots & x_r \end{bmatrix} \]

We can now invoke Nehari’s theorem:

**Theorem 3.1.**
\[
\inf_{Q \in \mathcal{H}_\infty} \|G - Q\| = \|\Gamma_G\| = \sigma_1
\]

**Proof.** See [Fra87], [Glo84], [Pel03].

It can be shown that the infimum in (4) is attained; further [Glo89]:
\[
\text{rank}_{\mathbb{R}(s)}[U(s)] = \text{rank}_{\mathbb{R}(s)}[V(s)] := l \leq \min(p, q, r) \quad (5)
\]

and
\[
(G - Q)U(s) = \sigma_1 V(s) \quad (6)
\]

for every (optimal) \(Q\) which achieves the infimum in (4). Equation (6) may be used to show that in the scalar case the optimal Nehari extension is unique and is given by \(Q = G - \sigma_1 V(s)/U(s)\). In the matrix case the equation has been used to derive the parametrization of all optimal solutions of the Nehari extension problem [Glo89], and has also inspired most methods used to solve the super-optimal distance problem, typically based on the construction of all-pass diagonalising transformations of \(G - Q\) using \(U(s)\) and \(V(s)\).

### 3.2 Statement of the problem

A formal definition of the problem follows. Firstly, define
\[
s_i^{\infty}(R) := \sup_{\omega \in \mathbb{R}} \sigma_i[R(j\omega)], \quad i = 1, 2, \ldots, \min(p, m).
\]

If \(p\) and \(m\) are both greater than 1, then we define recursively the first and subsequent super-optimal levels of \(R\) as
\[
s_i(R) := \inf_{Q \in S_{i-1}(R)} s_i^{\infty}(R + Q) \quad i = 1, 2, \ldots, \min(p, m) \quad (7)
\]

and the set of all \(i\)-th level super-optimal approximations of \(R\) as
\[
S_i(R) := \{Q \in S_{i-1}(R) : s_i^{\infty}(R + Q) = s_i(R)\} \quad i = 1, 2, \ldots, \min(p, m).
\]

In other words, we seek among all super-optimal approximations at the \((i - 1)\)-th level \(S_{i-1}(R)\) a set for which \(s_i(R)\) is minimized (it turns out that
the infimum in (7) is always attained). This set is not a singleton in general (apart from the case of \( i = \min(p, m) \)), but forms a subset of all \((i - 1)\)-th level super-optimal approximations of \( R, S_{i-1}(R) \). Due to the lexicographic nature of the problem, it is clear that every element of \( S_i(R) \) is also an element of \( S_{i-1}(R) \), i.e. that the super-optimal approximation sets nest as:

\[
S_0(R) \supseteq S_1(R) \supseteq \ldots \supseteq S_i(R) \supseteq \ldots \supseteq S_{\min(p,m)}(R)
\]

Note that for \( i = 1 \), (7) is taken to be a Nehari extension problem and hence we define \( S_0(R) := \mathcal{H}_\infty^{+p \times m} \). The super-optimal approximation problem ([SODP]) considered in this paper can be formally defined as follows:

**Problem 3.1.** [SODP]. Given a \( G \in \mathcal{RH}_\infty^{-p \times m} \), find the (unique) matrix-function \( Q_{\sup} \in \mathcal{H}_\infty^{+p \times m} \) which minimizes the sequence

\[
s_\infty(G + Q) = (s_1^\infty(G + Q), s_2^\infty(G + Q), \ldots, s_k^\infty(G + Q))
\]

with respect to the lexicographic ordering, where \( k = \min(p, m) \).

The approach followed here involves the reduction of the lexicographic minimization into a hierarchy of ordinary \( \mathcal{H}_\infty \)-optimization (Nehari-extension) problems of progressively reduced input-output dimensions, whose solution is well known in the literature [Glo84], [Glo89], [ZDG96], [GL95]. In particular, for the case of \( i = 2 \) in (7), two all-pass system matrices \( V^\sim \) and \( W \) are constructed (depending on \( R \)) which diagonalise every optimal “error system” \( R + Q, Q \in S_1(R) \), i.e.

\[
V^\sim(R + Q)W = \left( \begin{array}{cc} s_1(R)\alpha(s) & 0 \\ 0 & \hat{R} + \overline{Q} \end{array} \right)
\]

in which \( \hat{R} \in \mathcal{RH}_\infty^{-l \times (p-l)} \), \( \overline{Q} \in \mathcal{RH}_\infty^{+l \times (m-l)} \), \( l \geq 1 \) (generically \( l = 1 \)). Note that \( \alpha(s) \) is anti-inner of dimension \( l \times l \); also \( \alpha(s) \) and \( \hat{R}(s) \) are fixed (i.e. they do not depend on \( Q \in S_1(R) \)). It is further shown that \( \|\hat{R}^\sim\|_H < s_1(R) \) and that as \( Q \) varies over \( S_1(R) \), \( \overline{Q} \) varies over the set of all \( s_1(R) \) sub-optimal Nehari approximations of \( \hat{R} \), i.e. over the set

\[
S(\hat{R}, s_1(R)) := \{ \Psi \in \mathcal{H}_\infty^{+(p-l) \times (m-l)} : \|\hat{R} + \Psi\|_\infty \leq s_1(R) \}
\]

Thus (in the generic case \( l = 1 \)),

\[
s_2(R) = \inf_{Q \in S_1(R)} s_2^\infty(V^\sim(R + Q)W) = \inf_{\overline{Q} \in S(\hat{R}, s_1(R))} s_1^\infty(\hat{R} + \overline{Q})
\]
and so in this case (as all optimal Nehari approximations of $\hat{R}$ are also $s_1(R)$-suboptimal)

$$s_2(R) = s_1(\hat{R})$$

A recursive application of this procedure generates all super-optimal levels.

### 3.3 Overview

The paper considers the super-optimal Nehari-extension problem for real-rational continuous-time systems. All results are established via simple linear algebraic methods. The main steps of the algorithm are first developed purely at a transfer-function level, although this construction is subsequently supported via a detailed state-space analysis. The main features of our approach and the contribution of the work are briefly described below:

- We remove all main assumptions made in previous state-space based solutions to the problem. Specifically: (i) The state-space realization of the system which is approximated ($R(s)$) is not assumed to be minimal or balanced (although no unstable modes are allowed in its realization); (ii) The largest Hankel singular value of $R(s)$ is here assumed to have arbitrary multiplicity; and (iii) no assumption is made about the invertibility of the controllability and observability gramians of certain systems’ realizations arising at an intermediate step of the algorithm; these invertibility conditions were previously assumed to facilitate the state-space analysis of the algorithm and (unnecessarily) qualified the derived degree bound of the super-optimal approximation [LHG89].

- We adopt the approach in [JL93] which develops the solution of the two-block super-optimal distance problem via matrix dilations/all-pass embedding techniques. This approach provides conceptual and computational advantages over existing methods, e.g. [TGP88], [Kwa86], [LHG89], whose starting point is invariably the diagonalisation of the Nehari optimal solution set with the help of the Schmidt-pair of the Hankel operator $\Gamma$ associated with the problem. Although the diagonalising property of the Schmidt pair provides clear intuition for defining and solving the super-optimal distance problem, it is in fact computationally redundant. This is because all information about the Schmidt vectors and their properties is already implicitly contained in the structure of the “generator” of all optimal solutions of the Nehari-extension problem (contrast, for example, the state-space construc-
tion and the relevant proofs of [Glo89] for the suboptimal and optimal cases).

The approach based on Schmidt vectors requires a number of frequency-dependent scalings for constructing the diagonalising transformations (including a non-proper preliminary scaling), while the all-pass embedding approach followed here constructs the diagonalisation transformations directly from the off-diagonal blocks of an optimal and a sub-optimal Nehari generator via the solution of two Riccati equations (which are “dual” to each other). In fact, it turns out that for the purpose of describing the algorithm at a transfer-function level, it is not the scaled Schmidt vectors that are relevant, but their inner complements, and these are constructed directly using the all-pass embedding approach. An additional advantage of this approach is its straightforward treatment of repeated Hankel-singular values; in this case intricate questions about the dimension of spectral eigenspaces of \( \Gamma^{\sim} \) do not arise (in a sense, they have already been addressed in the construction of the optimal Nehari generator [Glo89], [GCR88] which is used by this method). Further, it is worth noting that working with state-space descriptions rather than polynomials enhances the computational robustness of the algorithm (the construction involves numerous pole-zero cancellations which are identified in our state-space analysis and, potentially, numerical ill-conditioned Riccati equations). Although our work concentrates on the solution of the two-level problem, the proposed algorithm can be easily extended to solve the full problem (lexicographic minimization of all super-optimal levels) using a standard recursive argument [LHG89], [HLG93], [JL93], [HJ98a].

- In this paper the method of [JL93] is specialized to the 1-block SODP. The special structure of this problem allows for a detailed state-space pole-zero cancellation analysis, arising in part from the duality between two spectral factorization-type Riccati equations, a relationship that fails to hold in the two-block case (where one additional constant term appears in one Riccati equation). The duality is formally established by linking the two Hamiltonians corresponding to the two Riccati equations, which shows that the two diagonalising transformations used in the construction share the same set of modes. The connection is even more transparent when we analyze symmetric problems, where the two Riccati equation solutions are identical, a fact that is used to establish the symmetry of the super-optimal approximation.
for this class of problems. The state-space analysis for the general (non-symmetric) problem is complete, in the sense that concrete state-space realizations are derived for all systems defined in the algorithm; it is further shown by numerical examples that these realizations are minimal for problems without a special structure. The state-space analysis can be used to derive degree bounds of the super-optimal approximation and establish certain interlacing inequalities between super-optimal levels and Hankel singular values [LHG88], [LHG89] without imposing additional unnecessary assumptions and using only standard algebraic-type arguments.

• The paper briefly discusses applications of super-optimization in control theory. Early references report applications in the areas of disturbance rejection [Kwa86], robust stabilization [KN89], [Nym95] and hierarchical $\mathcal{H}_\infty$ design [HJ98a], [HJW97]. Here we concentrate on some more recent applications of super-optimization in the areas of robust stabilization and structured-singular value approximations [GHJ00], [JHMG]. A detailed overview of these results (including new results on co-prime factor perturbation models) will be presented in a future publication forming part II of this work.

3.4 Brief survey of literature

The first published results in super-optimization can be found in [You86] and are based on operator theoretic methods. In subsequent years, linear-algebraic algorithms for the real-rational problem appeared in a series of papers [PF85], [PTG89], [TGP88], [LHG88], [LHG89], [GTP90]. These all relied on state-space methods and addressed the problem both in continuous and discrete-time settings. A parallel approach using a polynomial framework was developed in references [Kwa86], [KN89]. Investigations on cancellation analysis, degree-bounds and “interlacing inequalities” between Hankel singular values and super-optimal levels can be found in [LHG88], [LHG89] and [Pel03]. Generalizations of super-optimization to the two-block and four-block problems were first reported in [GTP89], [Nym94] and [JL93]. Reference [GTP89] follows the early state-space approach for solving the two-block $\mathcal{H}_\infty$ problem, by reducing it to an equivalent one-block problem via a spectral and an inner-outer factorization. In contrast, the approach of [Nym94] is based on the “equalization-principle”, widely used in early $\mathcal{H}_\infty$ polynomial methods [Kwa86], while [JL93] relies on a state-space all-pass dilation technique, proposed in [GLD+91] for solving the general-distance
$\mathcal{H}_\infty$ problem. An interesting state-feedback approach based on Riccati inequalities, in the spirit of recent LMI developments, can be found in [Foo04]. Extensions of super-optimization to the Hankel-norm approximation (AAK) problem, originating with the work of [PY96], [Tre95] were further developed in an algorithmic state-space setting in [HLG93] and [HJ98b]. Despite its similarity to its Nehari counterpart, the super-optimal Hankel-norm problem is considerably more intricate; it is known that in pathological cases, even uniqueness of the super-optimal approximation can be lost [Tre95],[HJ98b], which was the original motivating factor for introducing super-optimization.

Applications of super-optimization in control theory were first reported in the areas of disturbance rejection [Kwa86] and robust stabilization [Nym95]. The stronger version of optimality resulting from super-optimal approximations has been used in [Hal93], [HJW97], [HJ98a], [DH98] to address hierarchical optimization problems in an $\mathcal{H}_\infty$ or mixed-norm setting. In [Nym99] a multidirectional gap-metric is defined for multivariable systems under gap and coprime-factor perturbations using super-optimization ideas. In [Gom95] an inverse-robust stabilization problem is addressed: Given a super-optimal controller, determine the set of plants which it stabilizes. Reference [GHJ00] applies super-optimization techniques in the area of maximal robust-stabilization of LTI systems under additive perturbations: Explicit expressions for the improved robust stability radius are derived by imposing structure on the perturbation set via a uniform frequency constraint in the most-critical direction which is identified. The method is also used in [GHJ00], [JHMG] to derive an upper bound on the structured singular value for multivariable systems in the case of complex structured block-diagonal perturbations, which is tighter than the convex upper bound provided by the “D-iteration”. In this context, the multiplicity of the largest Hankel singular value becomes a crucial consideration, which motivates the detailed analysis of the general problem presented in this paper. An overview of these results and extensions to the case of normalized coprime-factor uncertainty models will be reported in Part II of this work.

4 The 1-block Super-Optimal Distance Problem

The approach for solving the SODP adopted in this paper is based on all-pass dilation techniques. First the system to be approximated, $R(s)$, is embedded in an all-pass system $H(s)$ of higher dimensions (note that $R(s)$ is taken to lie in $\mathcal{H}_\infty$ for compatibility with the existing $\mathcal{H}_\infty$ optimal-control
literature). This acts as a “generator” of the optimal solution set of the Nehari extension problem, as all solutions can be obtained via a LFT of $H(s)$ with the ball of $H_\infty$ of radius $s_1^2$ (i.e. the set of all stable $s_1^2$-contractions) [Glo89]. Next, a sub-block of the optimal generator $H(s)$ is dilated to define a new square all-pass system $\overline{H}(s)$, of lower dimensions compared to those of $H(s)$. Exploiting the all-pass nature of $H(s)$ and $\overline{H}(s)$ and the fact that they share a common block, two diagonalising transformations of $H(s)$ can be defined from certain sub-blocks of $H(s)$ and $\overline{H}(s)$. The diagonalisation is analogous to the partial singular-value decomposition of constant matrices and makes the minimization of the second super-optimal level transparent.

First, the general solution of the optimal Nehari-extension problem is given under minimal assumptions:

**Theorem 4.1 (Optimal Nehari approximation).** Consider $R \in \mathcal{RH}_\infty^{-,p \times m}$ with realization $R = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}$ where $\lambda(A) \subseteq \mathbb{C}_+$. Then there exists $Q_a \in \mathcal{RH}_\infty^{+(p+m-r) \times (p+m-r)}$ such that all $Q \in \mathcal{H}_\infty^{+,p \times m}$ such that $\|R + Q\|_\infty = \|R^+\|_H = s_1$ (Nehari optimal approximations of $R$) are given by

$$Q = \mathcal{F}_l(Q_a, s_1^2 BH(p-l) \times (m-l))$$

where in which $r$ denotes the multiplicity of the largest Hankel singular value of $R^+$, $l$ is defined in (5), and

$$Q_a := \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{bmatrix} A_q & B_{q1} & B_{q2} \\ C_{q1} & D_{11} & D_{12} \\ C_{q2} & D_{21} & 0 \end{bmatrix}$$

(8)

The corresponding “error” system is given by

$$H := \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} R + Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{bmatrix} A_H & B_H \\ C_H & D_H \end{bmatrix}$$

(9)

where $\|H_{22}\|_\infty < s_1$ and $Q_{ij} \in \mathcal{H}_\infty^+$, for $i, j \in \{1, 2\}$. Further, $HH^\sim = H^\sim H = s_1^2 I$ and the following set of equations is satisfied

$$P_HQ_H = Q_HP_H = s_1^2 I$$
$$D_HD_H' = D_H'D_H = s_1^2 I$$
$$A_H'Q_H + Q_HA_H + C'_H C_H = 0$$
$$A_HP_H + P_HA_H' + B_HB_H' = 0$$
$$D'_HC_H + B'_HQ_H = 0$$
$$D_HB_H' + C_HP_H = 0$$

(10)
Here $P_H$ and $Q_H$ are the gramians of the realization of $H$ given in (10).

Proof. See [Glo84]; see also [JL93] and [GLD+91] for a more general setting. 

**Remark 4.1.** The realization of $R$ need not be assumed minimal. However, we require that $\lambda(A) \subseteq C^+$. If $R$ has McMillan degree $n$, it can be shown [Glo86] that $Q_a$ given in (8) has degree $n-r$; in addition, $\sigma_i(Q_a) = \sigma_{i+r}(R^\sim), i = 1, 2, \ldots, n-r$ [Glo86], [GL95].

**Remark 4.2.** Integer parameter $l$ which is used to define the input and output dimension of $Q_{22}$ is the normal rank of the Laplace transform of the matrix formed by the $r$ Schmidt vectors of $\Sigma_R^\sim$ corresponding to $\sigma_1$, defined in equation (5). In the notation of Theorem 4.1 $R^\sim = (-A', C', -B')$ and hence $U(s)$ and $V(s)$ are given as

$$U(s) = -C(sI - A)^{-1}E \in \mathcal{R}(s)_{2m \times r}, \ E = \sigma_1^{-1}P \begin{bmatrix} x_1 & x_2 & \ldots & x_r \end{bmatrix}$$

and

$$V(s) = -B'(sI + A')^{-1}\Theta \in \mathcal{H}_{2n \times r}, \ \Theta = \begin{bmatrix} x_1 & x_2 & \ldots & x_r \end{bmatrix}$$

where $P$ and $Q$ are the controllability and observability matrices of $R = (A, B, C)$ and the $x_i$'s are $r$ linearly independent eigenvectors of $QP$ corresponding to the eigenvalue $\sigma_1^2$. In particular, if $(A, B, C)$ is balanced, $P = Q = -\text{diag}(\sigma_1 I, \Sigma_2)$, and thus $E = E_r$ and $\Theta = \sigma_1^2 E_r$ (where $E_r$ denotes the first $r$-columns of the $n \times n$ unit matrix), so that $U(s) = C(sI - A)^{-1}E_r \in \mathcal{H}_{2}^{1,m \times r}$ and $V(s) = -s^2 B'(sI + A')^{-1}E_r \in \mathcal{H}_{2}$. Thus,

$$\text{rank}_{R(s)}U^\sim(s) \geq \lim_{s \to \infty} [sU^\sim(s)] = \text{rank} \ (CE_r)$$

and

$$\text{rank}_{R(s)}V(s) \geq \lim_{s \to \infty} [sV(s)] = \text{rank} \ (E_r' B)$$

It is shown in [Glo86] that these two inequalities are actually equalities; further, the normal rank of $U(s)$ and $V(s)$ is equal, since $\text{Rank} \ (CE_r) = \text{Rank} \ (E_r' B)$, as can be verified by the equality $E_r'C CE_r = E_r'BB'E_r$, which follows easily from the all-pass equations (10). Thus $l \leq \min(p, m, r)$ and $l$ can be easily determined from the balanced realization of $R$.

**Remark 4.3.** In the present work, the gramians of $H$ are not considered to be balanced. The above set of equations is known as the set of "all-pass"
Proposition 4.1. Let $H_{22}$ be defined in 4.1 with $\|H_{22}\|_\infty < s_1$. Then,
1. There exists a square transfer matrix $\mathcal{P}_{21} \in \mathcal{RH}_\infty$ such that $\mathcal{P}_{21}^{-1} \mathcal{P}_{21} = s_1^2 I - H_{22} \tilde{H}_{22}$ and $\mathcal{P}_{21}^{-1} \in \mathcal{RH}_\infty$.

2. There exists a square transfer matrix $\mathcal{P}_{12} \in \mathcal{RH}_\infty$ such that $\mathcal{P}_{12}^{-1} \mathcal{P}_{12} = s_1^2 I - H_{22} \tilde{H}_{22}$ and $\mathcal{P}_{12}^{-1} \in \mathcal{RH}_\infty$.

3. The system
$$\mathcal{H} = \begin{pmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix} := \begin{pmatrix} -\mathcal{P}_{12} H_{22} \tilde{H}_{21} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{pmatrix}$$

is in $\mathcal{RL}_\infty$ and is $s_1$-allpass. Further, let $-\mathcal{P}_{12} H_{22} \tilde{H}_{21} = \hat{R} + \mathcal{Q}_{11}$ where $\hat{R} \in \mathcal{RH}_\infty^ -$ and $\mathcal{Q}_{11} \in \mathcal{RH}_\infty^ +$. Then $\|\hat{R}\|_H < s_1$.

**Proof.** For parts (1) and (2) see [ZDG96], Corollary 13.22. The proof follows from a detailed construction involving elements from the theory of algebraic Riccati equations and spectral factorization, which is briefly discussed in the following section. The proof that $\mathcal{H}$ is in $\mathcal{L}_\infty$ and is $s_1$-allpass follows from [Glo86] and can be verified directly by showing that $\tilde{H} \tilde{H}^{-} = s_1^2 I$. Finally, to show that $\|\hat{R}\|_H < s_1$, note that since $\mathcal{P}_{12}$ (or $\mathcal{P}_{21}$) is a unit of $\mathcal{H}_\infty$, and $\mathcal{H}$ is $s_1$-allpass, then $\|\mathcal{P}_{11}\|_\infty < s_1$. Write $\mathcal{P}_{11} = \hat{R} + \mathcal{Q}_{11}$ where $\hat{R} \in \mathcal{H}_\infty^ -$ and $\mathcal{Q}_{11} \in \mathcal{H}_\infty^ +$. Then, using Nehari’s theorem
$$\|\hat{R}\|_H = \inf_{X \in \mathcal{H}_\infty} \|\hat{R} + X\|_\infty \leq \|\hat{R} + \mathcal{Q}_{11}\|_\infty = \|\mathcal{P}_{11}\|_\infty < s_1$$
which completes the proof. □

**Remark 4.4.** Since $s_1 = \sigma_1(\hat{R})$ the inequality of part (3) says that $\sigma_1(\hat{R}) < \sigma_1(\mathcal{R}^-)$. As shown later in this section this can be strengthened to $\sigma_1(\hat{R}) < \sigma_{r+1}(\mathcal{R}^-)$, where $r$ is the multiplicity of the largest Hankel singular value of $\mathcal{R}^-$. A detailed state-space construction of $\mathcal{H}$ and its properties are given in Theorem 4.2 below.

**Theorem 4.2.** Consider
$$H_{22} = Q_{22} \equiv \begin{bmatrix} A_q & B_{q2} \\ C_{q2} & 0 \end{bmatrix} \in \mathcal{H}_\infty^{+(m-l)\times(p-l)}, \quad \|Q_{22}\|_\infty < s_1$$
defined in Theorem 4.1. Then there exist unique stabilizing solutions \( P_2 \) and \( Q_2 \) to the following algebraic Riccati equations:

\[
\begin{align*}
A_q P_2 + P_2 A_q' + B_q B_q' + s_1^{-2} P_2 C_q' C_q P_2 &= 0 \\
A_q' Q_2 + Q_2 A_q + C_q' C_q + s_1^{-2} Q_2 B_q B_q' Q_2 &= 0
\end{align*}
\] (12)

respectively. Define:

\[
R := Q_2 P_2 - s_1^2 I
\] (13)

Then \( R \) is non-singular. Further, there exists a \( Q_a \in H_{+}^{(p+m-2)\times(p+m-2)} \) with realization

\[
Q_a := \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} = \begin{bmatrix} A_q & B_q \\ C_{q_1} & 0 \end{bmatrix} \begin{bmatrix} B_{q_1} & B_{q_2} \\ 0 & s_1 I \end{bmatrix}
\] (14)

where

\[
\begin{align*}
\mathcal{C}_{q_1} &= -s_1^{-1} B_{q_2} Q_2 \\
\mathcal{B}_{q_1} &= -s_1^{-1} P_2 C_q'
\end{align*}
\] (15)

so that \( Q = F_l(Q_a, s^2) B(H_{+}^{(p-l)\times(m-l)}) \) is the set of all \( s_1 \)-suboptimal Nehari extensions of a system \( \hat{R} \in H_{-}^{(p-l)\times(m-l)} \) defined as:

\[
\hat{R} = \begin{bmatrix} \hat{A} & \hat{B} \\ C & 0 \end{bmatrix}
\] (16)

in which

\[
\begin{align*}
\hat{A} &= -(A_q + s_1^{-2} P_2 C_q' C_q) = -A_q' - s_1^{-2} C_q' C_q P_2 \\
\hat{B} &= -s_1^{-1} C_q' \\
\hat{C} &= s_1^{-1} B_{q_2} R
\end{align*}
\] (17)

The corresponding “error system”

\[
\mathcal{H} = \hat{R}_a + Q_a = \begin{pmatrix} \hat{R} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}
\] (18)

is \( s_1 \)-allpass and has a realization

\[
\mathcal{H} := \begin{pmatrix} \mathcal{H}_{11} & \mathcal{H}_{12} \\ \mathcal{H}_{21} & \mathcal{H}_{22} \end{pmatrix} = \begin{pmatrix} \hat{A} & 0 \\ 0 & A_q \end{pmatrix} \begin{pmatrix} \hat{B} & 0 \\ 0 & B_{q_1} \end{pmatrix} \begin{pmatrix} \mathcal{C} & \mathcal{C}_{q_1} \\ C_{q_2} & 0 \end{pmatrix} \begin{pmatrix} 0 & s_1 I \\ s_1 I & 0 \end{pmatrix}
\] (19)

17
which satisfies the following set of all-pass equations:

\[
\begin{align*}
A_\Pi'Q_\Pi + Q_\Pi A_\Pi + C_\Pi'C_\Pi &= 0 \\
A_\Pi'P_\Pi + P_\Pi A_\Pi' + B_\Pi'B_\Pi &= 0 \\
D_\Pi'C_\Pi + B_\Pi'Q_\Pi &= 0 \\
D_\Pi'B_\Pi' + C_\Pi P_\Pi' &= 0 \\
D_\Pi'D_\Pi = D_\Pi'\Pi D_\Pi &= s_1^2 I \\
P_\Pi Q_\Pi = Q_\Pi P_\Pi &= s_1^2 I
\end{align*}
\]

(20)

in which \(Q_\Pi\) and \(P_\Pi\) are the gramians of the realization of \(\Pi\) given in (19).

**Proof.** The proof is based on [Glo84]; see also [JL93] and [GLD+91] for a more general setting. Here we outline the sequence of logical arguments.

The existence of solutions of the two Riccati equations (12) follows from standard theory of spectral factorization and the bounded real-lemma (see Lemma 4.1 in the next section) and relies on the fact that \(\|Q_{22}\|_\infty < s_1\). Details and additional properties of the two solution are included in the following section. Since the two stabilising solutions are chosen, \(\hat{A}\) defined in equation (17) is anti-stable and thus \(\hat{R} \in \mathcal{H}_\infty\). Systems \(Q_a\) and \(\hat{R}\) correspond to the stable and anti-stable projections of \(\Pi\) given in Proposition 4.1 which also shows that \(\Pi\) is \(s_1\)-all pass. For a state-space based proof one needs to verify the all-pass equations given in (20) and expanded in (21) below; this is straightforward using the realizations given in Theorem 4.1 and the two Riccati equations (12). To show that \(\Pi\) is non-singular, first note that \(P_2\) and \(Q_2\) are the controllability and observability gramians, respectively, of the realization of \(Q_a\) given in equation (14), so that \(\sigma_1^2(Q_{a}) = \lambda_{\max}(P_2Q_2)\). A standard argument (e.g. see the early part of the proof of Theorem 4.4 which does not rely on any state-space arguments) shows that \(\sigma_1(Q_a) \leq \sigma_{r+1}(\hat{R}) < \sigma_1(R^-) = s_1\). Thus \(\rho(P_2Q_2) < s_1^2\) and thus \(\Pi\) is nonsingular. Finally, the fact that \(Q_a\) generates all \(s_1\)-suboptimal Nehari extensions of \(\hat{R}\) follows from the inertia properties of \(A\) and \(\hat{A}\) and the all pass-nature of \(\Pi\) [Glo86]; the proof reduces to showing that the invariant zeros of the realizations of \(Q_{12}\) (or \(Q_{21}\)) given in (19) lie in the open right-half plane, which follows readily by a simple calculation using the fact that \(\lambda(\hat{A}) \subseteq C_+\). \(\square\)

**Remark 4.5.** Expanding the compact form of the all-pass equations given
in Theorem 4.2 we get

\[
\begin{align*}
(i) & \quad \begin{bmatrix} \hat{A}' & 0 \\ 0 & A_q' \end{bmatrix} \begin{bmatrix} \hat{Q}_1 & -\hat{R} \\ -\hat{R} & \hat{Q}_2 \end{bmatrix} + \begin{bmatrix} Q_1 & -\hat{R}' \\ -\hat{R}' & Q_2 \end{bmatrix} \begin{bmatrix} \hat{A} & 0 \\ 0 & A_q \end{bmatrix} + \begin{bmatrix} \hat{C}' & 0 \\ 0 & C_{q_1} \end{bmatrix} = 0 \\
(ii) & \quad \begin{bmatrix} \hat{A} & 0 \\ 0 & A_q \end{bmatrix} \begin{bmatrix} \hat{P}_1 & I \\ I & \hat{P}_2 \end{bmatrix} + \begin{bmatrix} \hat{P}_1 & I \\ I & \hat{P}_2 \end{bmatrix} \begin{bmatrix} \hat{A}' & 0 \\ 0 & A_q' \end{bmatrix} + \begin{bmatrix} \hat{B} & 0 \\ 0 & B_{q_1} \end{bmatrix} = 0 \\
(iii) & \quad \begin{bmatrix} 0 & s_1 I \\ s_1 I & 0 \end{bmatrix} \begin{bmatrix} \hat{C} & C_{q_1} \\ 0 & C_{q_2} \end{bmatrix} + \begin{bmatrix} \hat{B}' & B_{q_1} \\ B_{q_1} & \hat{B}' \end{bmatrix} \begin{bmatrix} Q_1 & -\hat{R}' \\ -\hat{R}' & Q_2 \end{bmatrix} = 0 \\
(iv) & \quad \begin{bmatrix} 0 & s_1 I \\ s_1 I & 0 \end{bmatrix} \begin{bmatrix} \hat{B}' & B_{q_1} \\ B_{q_1} & \hat{B}' \end{bmatrix} + \begin{bmatrix} \hat{C} & C_{q_1} \\ 0 & C_{q_2} \end{bmatrix} \begin{bmatrix} \hat{P}_1 & I \\ I & \hat{P}_2 \end{bmatrix} = 0 \\
(v) & \quad \begin{bmatrix} Q_2 \hat{R}' & 0 \\ 0 & \hat{P}_2 \hat{R}' \end{bmatrix} \begin{bmatrix} \hat{P}_2 \hat{R}' & -\hat{R}' \\ -\hat{R}' & \hat{P}_2 \hat{R}' \end{bmatrix} = \begin{bmatrix} s_1^2 I & 0 \\ 0 & s_1^2 I \end{bmatrix}
\end{align*}
\]

(21)

where \( \hat{P}_1 = Q_2 \hat{R}' \) and \( \hat{Q}_1 = \hat{P}_2 \hat{R}' \).

The following theorem constructs a diagonalising transformation of \( H \) and solves the level-two SODP.

**Theorem 4.3.** Let \( H \) and \( \overline{H} \) be as defined in Theorems 4.1 and 4.2, respectively. Then

\[ \| R^\sim \|_H = s_1(R) = s_2(R) = \ldots = s_l(R) > s_{l+1}(R) = \| \hat{R}^\sim \|_H \]

Further,

\[ S_1(R) = S_2(R) = \ldots = S_l(R) = F_l(Q_a, s_1^2 B H_{\infty}^{(p-l) \times (m-l)}) \]

and

\[ S_{l+1}(R) = F_l(Q_a, F_a(\hat{Q}_a, S_1(\hat{R}))) \subseteq S_l(R) \]

where \( Q_a \) and \( \hat{Q}_a \) are defined in Theorems 4.1 and 4.2.

**Proof.** We adapt the proof of [JL93] Theorem 3 to our setting. First note that since \( HH^\sim = H^\sim H = s_1^2 I \) and \( \overline{H} \overline{H}^\sim = \overline{H}^\sim \overline{H} = s_1^2 I \), it follows that

\[ H_{11} H_{21}^\sim = -H_{21} H_{22}^\sim, \quad \overline{\Pi}_{11} = -\overline{\Pi}_{12} \overline{\Pi}_{22} \overline{\Pi}_{21}^\sim, \]

\[ \overline{\Pi}_{21} \overline{\Pi}_{21}^\sim = s_1^2 I - H_{22} H_{22}^\sim = H_{21} H_{21}^\sim \]

(22)

(23)
and
\[ \mathcal{H}_{12} \sim \mathcal{H}_{12} = s_1^2 I - H_{22}^2 H_{22} = H_{12}^2 H_{12} \] 
(24)

Define
\[ V_\perp := H_{12} \mathcal{H}_{12}^{-1} \quad \text{and} \quad W_\perp := H_{21} \mathcal{H}_{21}^{-1} \] 
(25)

Then (23) implies that
\[ V_\perp V_\perp = I_{p-l} \quad \text{and} \quad W_\perp W_\perp = I_{m-l} \] 
(26)

It can be readily verified from a state-space calculation (see next section) that \( V_\perp \in \mathcal{H}_\infty^{+, (p-l) \times p} \) and \( W_\perp \in \mathcal{H}_\infty^{-(m-l) \times m} \). Thus there exist complementary inner and co-inner factors, respectively, such that
\[ V := (v \ V_\perp) \in \mathcal{H}_\infty^{+, p \times p} \quad \text{and} \quad W := (w \ W_\perp) \in \mathcal{H}_\infty^{-, m \times m} \]
are square-inner and square anti-inner, respectively [ZDG96], [GL95]. Thus, using (22) and the definitions (25), we obtain
\[ H_{21} W_\perp = H_{21} H_{12} H_{21} \mathcal{H}_{21}^{-1} = \mathcal{H}_{21} \mathcal{H}_{21}^{-1} = H_{21} \] 
(27)

and
\[ V_\perp H_{11} W_\perp = V_\perp H_{11} H_{12} \mathcal{H}_{12}^{-1} = -V_\perp H_{12} H_{22} \mathcal{H}_{21}^{-1} \]
\[ = -H_{12} H_{22} \mathcal{H}_{21}^{-1} = \mathcal{H}_{11} \] 
(28)

It follows that
\[
\begin{pmatrix}
V^\sim & 0 \\
0 & I
\end{pmatrix}
\begin{pmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{pmatrix}
\begin{pmatrix}
W & 0 \\
0 & I
\end{pmatrix} =
\begin{pmatrix}
v^\sim H_{11} w & v^\sim H_{11} W_\perp & v^\sim H_{12} \\
V^\sim H_{11} w & \mathcal{H}_{11} & \mathcal{H}_{12} \\
H_{21} w & \mathcal{H}_{21} & \mathcal{H}_{22}
\end{pmatrix}
\] 
(29)

Now, since \( V \) and \( W \) are all-pass and \( H \) is \( s_1 \)-allpass, the system on the RHS of equation (29) is \( s_1 \)-allpass. But since \( \mathcal{H} \) is also \( s_1 \)-allpass (Theorem 4.2), we have that \( v^\sim H_{11} W_\perp = 0, v^\sim H_{12} = 0, V^\sim H_{11} w = 0, H_{21} w = 0, \) and \( v^\sim H_{11} w \) is \( s_1 \)-allpass and can be written as \( v^\sim H_{11} w = s_1 \alpha(s) \), for some \( l \times l \) all-pass matrix-function \( \alpha(s) \) (generically \( l = 1 \) and hence \( \alpha(s) \) is scalar).

Taking linear fractional transformations with the set \( s_1^2 \mathcal{BH}_\infty^{(p-l) \times (m-l)} \) and using the results of Theorem 4.2 and Theorem 4.1 shows that:
\[ V^\sim [\mathcal{F}_1(H, s_1^2 \mathcal{BH}_\infty^{(p-l) \times (m-l)})] W = \begin{pmatrix}
s_1 \alpha & 0 \\
0 & \mathcal{F}_1(\mathcal{H}, s_1^2 \mathcal{BH}_\infty^{(p-l) \times (m-l)})
\end{pmatrix} \] 
(30)
or equivalently,
\[
V^{{\sim}}[R + S_1(R)]W = \begin{pmatrix}
  s_1\alpha & 0 \\
  0 & \hat{R} + S(\hat{R}, s_1)
\end{pmatrix}
\]  
(31)

Since \(\alpha(s) \in R^{l \times l}(s)\) and is all-pass (in fact anti-inner as shown in the next section), it follows that:
\[
\|R^{{\sim}}\|_H = s_1(R) = s_2(R) = \ldots = s_l(R) > s_{l+1}(R) = \|\hat{R}^{{\sim}}\|_H
\]
and
\[
S_1(R) = S_2(R) = \ldots = S_l(R) = \mathcal{F}_l(Q, s_1^2 \mathcal{B} \mathcal{H}_\infty^{(p-l) \times (m-l)})
\]
which is the set of all optimal Nehari extensions of \(R\). Further, since all optimal Nehari extensions of \(\hat{R}\) are also \(s_1\)-suboptimal extensions of \(\hat{R}\), i.e. \(S_1(\hat{R}) \subseteq S(\hat{R}, s_1)\), it follows that
\[
s_{l+1}(R) = s_1(\hat{R}) = \|R^{{\sim}}\|_H
\]
and
\[
R + S_2(R) = (v \ V_\perp) \begin{pmatrix}
  s_1\alpha & 0 \\
  0 & \hat{R} + S_1(\hat{R})
\end{pmatrix} \begin{pmatrix}
  w^{{\sim}}_\perp \\
  W^{{\sim}}_\perp
\end{pmatrix}
= (v \ V_\perp) \begin{pmatrix}
  s_1\alpha & 0 \\
  0 & \hat{R} + Q
\end{pmatrix} \begin{pmatrix}
  w^{{\sim}}_\perp \\
  W^{{\sim}}_\perp
\end{pmatrix}
= R + Q_{11} + V_\perp (S_1(\hat{R}) - Q)W^{{\sim}}_\perp
\]  
(32)
by observing that
\[
V^{{\sim}}H_{11}W = \begin{pmatrix}
  s_1\alpha & 0 \\
  0 & P_{11}
\end{pmatrix} \Rightarrow R + Q_{11} = V \begin{pmatrix}
  s_1\alpha & 0 \\
  0 & \hat{R} + Q_{11}
\end{pmatrix} W^{{\sim}}
\]
Using the definitions of of \(V_\perp\) and \(W^{{\sim}}_\perp\) in (25) and cancelling \(R\) from both sides of equation (32), we can write:
\[
S_2(R) = Q_{11} + Q_{12}Q_{12}^{-1}(S_1(\hat{R}) - Q)Q_{21}^{-1}Q_{21}
= \mathcal{F}_l \begin{pmatrix}
  Q_{11} - Q_{12}Q_{12}^{-1}Q_{11}Q_{21}^{-1}Q_{21} & Q_{12}Q_{12}^{-1}Q_{21}^{-1}Q_{21}
\end{pmatrix}, S_1(\hat{R})
=: \mathcal{F}_l(K, S_1(\hat{R}))
\]
where
\[ K := \begin{pmatrix} Q_{11} - Q_{12} Q_{12}^{-1} Q_{11}^{-1} Q_{21} & Q_{12} Q_{12}^{-1} Q_{21} \\ Q_{21}^{-1} Q_{21} & 0 \end{pmatrix} = \mathcal{F}_I(Q_a, \overline{Q}_a^{-1}) \]
using a simple calculation. This completes the proof.

The following Theorem establishes bounds on the super-optimal levels. The proof is similar to a parallel result in [LHG89], but the assumption involving the multiplicity of the largest Hankel singular value of \( R^\sim \) is removed.

**Theorem 4.4 (Super-optimal level bounds).** The \((l + 1)\)-th super-optimal level is bounded above by the \((r + 1)\)-th Hankel singular value of \( R^\sim \), i.e.

\[ \sigma_1(\hat{R}^\sim) = s_{l+1}(R) \leq \sigma_{r+1}(R^\sim) < s_1(R) = s_2(R) = \ldots = s_l(R) = \sigma_1(R^\sim) \]

**Proof.** The proof follows from the following sequence of inequalities:

\[ \sigma_i(R^\sim) = \sigma_i(Q_a) \quad i = 1, 2, \ldots, n - r \]

\[ = \inf_{\Psi \in \mathcal{H}_\infty^{i-1}} \|Q_a + \Psi\|_\infty \]

\[ = \inf_{\Psi \in \mathcal{H}_\infty^{i-1}} \|R + Q_a + \Psi\|_\infty \]

\[ \geq \inf_{\Psi \in \mathcal{H}_\infty^{i-1}} \left\| \begin{pmatrix} V_{\perp}^\sim & 0 \\ 0 & I \end{pmatrix} (R + Q_a + \Psi) \begin{pmatrix} W_{\perp} & 0 \\ 0 & I \end{pmatrix} \right\|_\infty \]

\[ \geq \inf_{\Psi \in \mathcal{H}_\infty^{i-1}} \|\hat{R}_a + \overline{Q}_a + \hat{\Psi}\|_\infty \]

\[ \geq \inf_{\Psi \in \mathcal{H}_\infty^{i-1}} \|\overline{Q}_a + \hat{\Psi}\|_\infty \]

\[ = \sigma_i(\overline{Q}_a) \]

The first equality follows from Theorem 4.1. The second equality is a statement of the AAK Theorem [Glo86], while the third equality holds since \( R \in \mathcal{H}_\infty \) and can be absorbed in \( \Psi \). The first inequality follows from the fact that \( V_{\perp} \) and \( W_{\perp} \) are contractive, while the second inequality follows from Theorem 4.3 and the fact that \( V_{\perp}^\sim \) and \( W_{\perp} \) are both in \( \mathcal{R}\mathcal{H}_\infty \). Finally, the third inequality follows from the fact that \( \hat{R} \in \mathcal{R}\mathcal{H}_\infty \), while the last equality is a restatement of the AAK Theorem.

Setting \( i = 1 \) in the above inequality shows that \( \sigma_{r+1}(R^\sim) \geq \sigma_1(\overline{Q}_a) \). Now, using (21), it follows that

\[ \sigma_i^2(\hat{R}^\sim) = \lambda_i(\hat{P}_1 \overline{Q}_1) = \lambda_i(\overline{Q}_2 \hat{R}^\sim \overline{P}_2 R) = \lambda_i(\overline{Q}_2 P_2) = \sigma_i^2(\overline{Q}_a) \]

22
and so $\hat{R}^-$ and $\overline{Q_a}$ have identical Hankel singular values. In particular, $s_{l+1}(R) = \sigma_1(\hat{R}^-) \leq \sigma_{r+1}(R^-)$ using the result of Theorem 4.3.

**Remark 4.6.** The result of Theorem 4.4 may be propagated to establish upper bounds for the subsequent super-optimal levels $s_i(R)$, $i > l + 1$.

**Remark 4.7.** The early part of the proof (which does not rely on any state-space based arguments) may be used to show that $\sigma_1(\overline{Q_a}) \leq \sigma_{r+1}(R^-) < \sigma_1(R^-) = s_1$, from which it follows immediately that $R$ defined in Theorem 4.2 is non-singular.

### 4.1 State-space analysis

In this section we develop a state-space analysis of the solution to the super-optimal distance problem. First, some background material is presented related to algebraic Riccati equations and spectral factorizations.

Let $A$, $Q$ and $R$ be real $n$-by-$n$ matrices with $Q$ and $R$ symmetric. The Algebraic Riccati equation (ARE) is the matrix equation:

$$A'X + XA + RX + Q = 0$$

Associated with this equation, the Hamiltonian matrix is defined as:

$$H := \begin{bmatrix} A & R \\ -Q & -A' \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$$

Introduce the matrix:

$$J := \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

Then $J' = J^{-1}$ or $J^2 = -I_n$. It follows easily that $J^{-1}HJ = -JHJ = -H'$ and hence the spectrum of $H$ is symmetric with respect to the imaginary axis. A solution of the ARE is called stabilizing if the matrix $A + RX$ is stable (i.e. $\lambda(A + RX) \subseteq \mathbb{C}_-$) and in this case we write $H \in \text{dom}(\text{Ric})$. Note that if a stabilizing solution exists then it is unique and in this case $H$ has no eigenvalues on the imaginary axis. For necessary and sufficient conditions for the existence of a stabilizing solution see [ZDG96], [Kim97] and [Fra87].

We start our state-space analysis by quoting the following well-known result ("Bounded-real lemma"): 

...
Lemma 4.1. Let $G \in \mathcal{RH}_\infty$ with $G(s) = C(sI - A)^{-1}B$ and assume that $(A, B)$ and $(C, A)$ are stabilisable and detectable, respectively. Then, the following conditions are equivalent:

1. $\|G\|_\infty < \gamma$

2. The Hamiltonian $H = \begin{bmatrix} A & \gamma^{-2}BB' \\ -C'C & A' \end{bmatrix}$ has no pure imaginary eigenvalues

3. $H \in \text{dom}(\text{Ric})$

Proof. 1 $\iff$ 2. See [ZDG96], lemma 4.7.

2 $\iff$ 3. See [ZDG96], Theorem 13.6. \hfill $\square$

As an immediate consequence of the above Lemma we get the following result:

Proposition 4.2. The algebraic Riccati equations (12) (Theorem 4.2) have (unique) positive-semidefinite stabilising solutions $P_2$ and $Q_2$ respectively.

Proof. Since $A_q$ is asymptotically stable, the conditions of stabilizability and detectability of Lemma 4.1 are trivially satisfied. Further, the fact that $\|Q_{22}\|_\infty < s_1$ (see Theorem 4.1) shows that the two Hamiltonians associated with equations (12) are free of imaginary axis eigenvalues and that (unique) stabilizing solutions $P_2$ and $Q_2$ to these two equations exist. The fact that $P_2 \geq 0$ and $Q_2 \geq 0$ follows from [ZDG96]. \hfill $\square$

Our next result shows that the two Riccati equations (12) are intimately related.

Proposition 4.3. Let $P_2$ be the stabilizing solution of $\text{Ric1}$,

$$A_qP_2 + P_2A'_q + s_1^{-2}P_2C'_qC_qP_2 + B_qB'_q = 0$$

so that $\lambda(A_q + s_1^{-2}P_2C'_qC_q) \subseteq \mathcal{C}_-$ and its associated Hamiltonian

$$H_1 = \begin{bmatrix} A'_q & s_1^{-2}C'_qC_q \\ -B_qB'_q & -A_q \end{bmatrix}$$

Let also $Q_2$ be the stabilizing solution of $\text{Ric2}$:

$$A'_qQ_2 + Q_2A_q + s_1^{-2}Q_2B_qB'_qQ_2 + C'_qC_q = 0$$
so that $\lambda(A_q + s_1^{-2}B_{q2}B'_q\bar{Q}_2) \subseteq C_-$ and its associated Hamiltonian

$$H_2 = \begin{bmatrix} A_q & s_1^{-2}B_{q2}B'_q \\ -C_{q2}'C_{q2} & -A'_q \end{bmatrix}$$  \(34\)

Then $H_1$ and $H_2$ have identical spectra. In particular there exist a similarity transformation $\mathbf{R}'$ so that

$$(A_q + s_1^{-2}P_2C'_qC_{q2}) = \mathbf{R}'(A_q + s_1^{-2}B_{q2}B'_q\bar{Q}_2)(\mathbf{R}')^{-1}$$  \(35\)

where $\mathbf{R}$ is defined (13).

Proof. Take

$$T = \begin{bmatrix} 0 & s_1^{-1}I \\ s_1I & 0 \end{bmatrix} \Rightarrow T^{-1} = \begin{bmatrix} I & 0 \\ 0 & s_1^{-1}I \end{bmatrix}$$

Note that $T = T^{-1}$. Then by inspection the first claim is true. Define

$$T_P := \begin{bmatrix} I & 0 \\ -P_2 & I \end{bmatrix} \Rightarrow T_P^{-1} = \begin{bmatrix} I & 0 \\ P_2 & I \end{bmatrix}$$

and observe that

$$\begin{bmatrix} I & 0 \\ -P_2 & I \end{bmatrix} \begin{bmatrix} A'_q & s_1^{-2}C_{q2}'C_{q2} \\ -B_{q2}B'_q & -A_q \end{bmatrix} \begin{bmatrix} I & 0 \\ P_2 & I \end{bmatrix} = \begin{bmatrix} A'_q + s_1^{-2}C_{q2}'C_{q2}P_2 & s_1^{-2}C_{q2}'C_{q2} \\ 0 & -(A_q + s_1^{-2}P_2C_{q2}'C_{q2}) \end{bmatrix} =: \tilde{H}_1$$

Similarly, define

$$T_Q := \begin{bmatrix} I & 0 \\ -Q_2 & I \end{bmatrix} \Rightarrow T_Q^{-1} = \begin{bmatrix} I & 0 \\ Q_2 & I \end{bmatrix}$$

so that

$$\begin{bmatrix} I & 0 \\ -Q_2 & I \end{bmatrix} \begin{bmatrix} A_q & s_1^{-2}B_{q2}B'_q \\ -C_{q2}'C_{q2} & -A'_q \end{bmatrix} \begin{bmatrix} I & 0 \\ Q_2 & I \end{bmatrix} = \begin{bmatrix} A_q + s_1^{-2}B_{q2}B'_q\bar{Q}_2 & s_1^{-2}B_{q2}B'_q \\ 0 & -(A'_q + s_1^{-2}Q_2B_{q2}B'_q) \end{bmatrix} =: \tilde{H}_2$$

$$\begin{cases} H_1 = -TH_2T^{-1} \\ \tilde{H}_1 = T_PH_1T_P^{-1} \\ \tilde{H}_2 = T_QH_2T_Q^{-1} \Rightarrow H_2 = T_Q^{-1}\tilde{H}_2T_Q \end{cases}$$


Using the three above equations

\[ H_1 = -TT_Q^{-1} H_T T_Q T^{-1} \Rightarrow \tilde{H}_1 = -T_T T_Q^{-1} \tilde{H}_2 T_Q T^{-1} T_P \]

Further,

\[ \tilde{H}_1(T QT^{-1} T_P^{-1})^{-1} = -T_T T_Q^{-1} \tilde{H}_2 \Rightarrow \tilde{H}_1 T_T T_Q^{-1} = -T_T T_Q^{-1} \tilde{H}_2 \]

and

\[ (T_T T_Q^{-1})^{-1} \tilde{H}_1 = -\tilde{H}_2 T_T T_Q^{-1} T_P^{-1} \Rightarrow \tilde{H}_1(T_T T_Q^{-1}) = -(T_T T_Q^{-1}) \tilde{H}_2 \]

with

\[ T_T T_Q^{-1} = \begin{bmatrix} I & 0 \\ -P_2 & I \end{bmatrix} \begin{bmatrix} 0 & s_1^{-1}I \\ s_1I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ Q_2 & I \end{bmatrix} = s_1^{-1} \begin{bmatrix} \overline{Q}_2 & I \\ -\overline{R} & -P_2 \end{bmatrix} \]

and

\[ T_Q T^{-1} T_P^{-1} = \begin{bmatrix} I & 0 \\ -Q_2 & I \end{bmatrix} \begin{bmatrix} 0 & s_1^{-1}I \\ s_1I & 0 \end{bmatrix} \begin{bmatrix} I & 0 \\ P_2 & I \end{bmatrix} = s_1^{-1} \begin{bmatrix} \overline{P}_2 & I \\ -\overline{R} & -Q_2 \end{bmatrix} \]

Writing equation (36) in full:

\[
\begin{bmatrix}
A'_q + s_1^{-2} C'_q C_q \overline{P}_2 & s_1^{-2} C'_q C_q \\
0 & -\left(A_q + s_1^{-2} \overline{P}_2 C'_q C_q\right)
\end{bmatrix}
\begin{bmatrix}
\overline{Q}_2 & I \\
-\overline{R} & -\overline{P}_2
\end{bmatrix}
= \begin{bmatrix}
\overline{Q}_2 & I \\
-\overline{R} & -\overline{P}_2
\end{bmatrix}
\]

From the (2, 1) partition of the above equation, we have

\[
(A_q + s_1^{-2} \overline{P}_2 C'_q C_q) \overline{R}' = \overline{R}' (A_q + s_1^{-2} B_q Q_2 B_q' \overline{Q}_2). \]

So,

\[
(A_q + s_1^{-2} \overline{P}_2 C'_q C_q) = \overline{R}' (A_q + s_1^{-2} B_q Q_2 B_q' \overline{Q}_2)(\overline{R}')^{-1}
\]

which proves the second claim.

\[ \square \]

**Remark 4.8.** Note that this proposition implies that the “A” matrices of the state space realizations of \(V_\perp\) and \(W_\perp\) have the same spectrum.

**Proposition 4.4.** Define

\[ V_\perp := H_{12} H_{12}^{-1} \quad \text{and} \quad W_\perp := H_{21} H_{21}^{-1} \]
Then, $V_{\perp}$ and $W_{\sim}$ have the following realizations:

\[
V_{\perp} = s \begin{bmatrix}
A_q & s_1^{-1}B_{q2}C_{q1} & s_1^{-1}B_{q2}C_{q1}
\end{bmatrix}
\]

\[
W_{\sim} = s \begin{bmatrix}
A_q & s_1^{-1}B_{q1}C_{q2} & B_{q1} & s_1^{-1}B_{q1}D_{21}
\end{bmatrix}
\]

with corresponding controllability and observability gramians:

\[
Y_v = -(\overline{R}^t)^{-1}\overline{P}_2, \quad X_v = Q_2 - \overline{Q}_2
\]

\[
Y_w = P_2 - \overline{P}_2, \quad X_w = -\overline{P}_1.
\]

In particular, the following matrix inequalities hold: $P_2 \geq P_2$ and $Q_2 \geq \overline{Q}_2$.

**Proof.** The proof is broken into separate sections. The first part is verified by simple state-space calculations. This is followed by deriving the gramians for the derived realizations of $V_{\perp}$ and $W_{\sim}$.

1. **State-space realizations:**

\[
V_{\perp} = Q_{12}\overline{Q}_{12}^{-1} = \begin{bmatrix}
A_q & s_1^{-1}B_{q2}C_{q1} & s_1^{-1}B_{q2}C_{q1}
\end{bmatrix}
\]

\[
W_{\sim} = \begin{bmatrix}
A_q & s_1^{-1}B_{q1}C_{q2} & B_{q1} & s_1^{-1}B_{q1}D_{21}
\end{bmatrix}
\]

where $T := \begin{bmatrix} I & -I \\ 0 & I \end{bmatrix}$. By removing the uncontrollable terms,

\[
V_{\perp} = s \begin{bmatrix}
A_q & s_1^{-1}B_{q2}C_{q1} & s_1^{-1}B_{q2}
\end{bmatrix}
\]

\[
W_{\sim} = s \begin{bmatrix}
A_q & s_1^{-1}B_{q1}C_{q2} & B_{q1} & s_1^{-1}B_{q1}D_{21}
\end{bmatrix}
\]
Similarly,

\[
W_\perp = Q_2' \overline{Q}_{21}^{-1} = \begin{bmatrix}
-A_q' & C_{q2} \\ -B_{q1}' & D_{21}'
\end{bmatrix} \begin{bmatrix}
-A_q + s_1^{-1}C_{q2}'B_{q1}' & s_1^{-1}C_{q2}' \\
-s_1^{-1}B_{q1}' & s_1^{-1}D_{21}'
\end{bmatrix} = \begin{bmatrix}
-A_q' & s_1^{-1}C_{q2}'B_{q1}' & s_1^{-1}C_{q2}' \\
0 -A_q + s_1^{-1}C_{q2}'B_{q1}' & s_1^{-1}C_{q2}' \\
-B_{q1}' & s_1^{-1}D_{21}' & s_1^{-1}D_{21}'
\end{bmatrix}
\]

2a. **Controllability and Observability gramian of** \(V_\perp\): Take the controllability gramian of \(V_\perp\). That is

\[
(A_q - s_1^{-1}B_{q2}\overline{C}q_1)Y_v + Y_v(A_q' - s_1^{-1}\overline{C}q_1 B_{q2}') + s_1^{-2}B_{q2}B_{q2}' = 0
\]

or

\[
(A_q + s_1^{-1}B_{q2}B_{q2}' \overline{Q}_{2}q_2)Y_v + Y_v(A_q' + s_1^{-1}\overline{Q}B_{q2}B_{q2}') + s_1^{-2}B_{q2}B_{q2}' = 0
\]

Multiply from the left by \(\overline{R}'\) and multiply from the right by \(\overline{R}\).

\[
\overline{R}'(A_q + s_1^{-2}B_{q2}B_{q2}' \overline{Q}_{2}q_2)Y_v + \overline{R}'Y_v(A_q' + s_1^{-2}\overline{Q}B_{q2}B_{q2}') + s_1^{-2}\overline{R}' B_{q2}B_{q2}' \overline{R} = 0
\]

From proposition 6.1.3, we have

\[
(A_q + s_1^{-2}B_{q2}B_{q2}' \overline{Q}_{2}q_2) = (\overline{R}')^{-1}(A_q + s_1^{-2}\overline{P}_2 C_{q2}' \overline{C}_{q2})\overline{R}'
\]

Thus,

\[
(A_q + s_1^{-2}\overline{P}_2 C_{q2}' \overline{C}_{q2})\overline{Y}_v \overline{R} + \overline{R}'Y_v(\overline{R}'(A_q' + s_1^{-2}C_{q2}' \overline{C}_{q2}\overline{P}_2) + s_1^{-2}\overline{R}' B_{q2}B_{q2}' \overline{R} = 0
\]

(37)

Compare this with the all-pass equation:

\[
\hat{A}'_q \overline{Q}_{1} + \overline{Q}_{1} \hat{A} + \overline{C}'\hat{C} = 0
\]

written out in full as:

\[
(-A_q - s_1^{-2}\overline{P}_2 C_{q2}' \overline{C}_{q2})\overline{Q}_{1} + \overline{Q}_{1}(-A_q' - s_1^{-2}C_{q2}' \overline{C}_{q2}\overline{P}_2) + s_1^{-2}\overline{R}' B_{q2}B_{q2}' \overline{R} = 0
\]

(38)
In the sequel we show that
\[ Y_v = -(\mathbf{T'})^{-1}\mathbf{T}_2 \]
since the matrix \( A_q + s_1^{-2}\mathbf{P}_2 C'_q C_{q2} \) is asymptotically stable (i.e. its spectrum lies in the open left half plane).

Recall first the following Lyapunov equation, derived from all-pass equations (11(iii))2,2,
\[ A'_q Q_2 + Q_2 A_q + C'_q C_{q1} + C'_q C_{q2} = 0 \]
Further, from all-pass equation (21(i))2,2 we get the Riccati equation
\[ A'_q \mathbf{Q}_2 + \mathbf{Q}_2 A_q + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} \mathbf{Q}_2 + C'_q C_{q2} = 0 \]
Now, let the observability gramian of the derived realization of \( V_\perp \) be \( X_v \). Then:
\[
X_v(A_q - s_1^{-1} B_{q2} C_{q1}) + (A'_q - s_1^{-1} C'_{q1} B'_{q2}) X_v \\
+ (C'_q - s_1^{-1} C_{q1} D'_1)(C_{q1} - s_1^{-1} D_{12} C_{q1}) = 0
\]
Substituting \( C_{q1} = -s_1^{-1} B'_{q2} \mathbf{Q}_2 \) gives:
\[
X_v(A_q + s_1^{-2} B_{q2} B'_{q2} \mathbf{Q}_2) + (A'_q + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2}) X_v \\
+ (C'_q + s_1^{-2} \mathbf{Q}_2 B_{q2} D'_1)(C_{q1} + s_1^{-2} D_{12} B'_{q2} \mathbf{Q}_2) = 0
\]
which can be expanded as:
\[
X_v A_q + s_1^{-2} X_v B_{q2} B'_{q2} \mathbf{Q}_2 + A'_q X_v + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} X_v + C'_q C_{q1} \\
+ s_1^{-2} \mathbf{Q}_2 B_{q2} D'_1 C_{q1} + s_1^{-2} C'_{q1} D_{12} B'_{q2} \mathbf{Q}_2 + s_1^{-4} \mathbf{Q}_2 B_{q2} D'_{12} D_{12} B'_{q2} \mathbf{Q}_2 = 0
\]
But from (11(ii))22 we get that \( D'_{12} D_{12} = s_1^2 I \). Hence,
\[
X_v A_q + s_1^{-2} X_v B_{q2} B'_{q2} \mathbf{Q}_2 + A'_q X_v + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} X_v + C'_q C_{q1} \\
+ s_1^{-2} \mathbf{Q}_2 B_{q2} D'_1 C_{q1} + s_1^{-2} C'_{q1} D_{12} B'_{q2} \mathbf{Q}_2 + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} \mathbf{Q}_2 = 0
\]
In the sequel we show that \( X_v = Q_2 - \mathbf{Q}_2 \) is the unique solution of the above equation. The term on the left hand side of the equation can be written as:
\[
(Q_2 - \mathbf{Q}_2) A_q + s_1^{-2}(Q_2 - \mathbf{Q}_2) B_{q2} B'_{q2} \mathbf{Q}_2 + A'_q (Q_2 - \mathbf{Q}_2) \\
+ s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} (Q_2 - \mathbf{Q}_2) + C'_{q1} C_{q1} + s_1^{-2} \mathbf{Q}_2 B_{q2} D'_{12} C_{q1} \\
+ s_1^{-2} C'_{q1} D_{12} B'_{q2} \mathbf{Q}_2 + s_1^{-2} \mathbf{Q}_2 B_{q2} B'_{q2} \mathbf{Q}_2
\]
29
or, equivalently as:

\[ Q_2 A_q - \overline{q}_2 A_q + s_1^{-2} Q_2 B_q B'_q \overline{q}_2 - s_1^{-1} \overline{q}_2 B_q B'_q \overline{q}_2 + A'_q Q_2 - A'_q \overline{q}_2 + s_1^{-2} \overline{q}_2 B_q B'_q Q_2 - s_1^{-2} \overline{q}_2 B_q B'_q \overline{q}_2 + C'_q C_q + s_1^{-2} \overline{q}_2 B_q D'_q C_q + s_1^{-2} C'_q D_1 B_q \overline{q}_2 + s_1^{-2} C'_q D_1 B_q \overline{q}_2 \]

By subtracting the Riccati from the Lyapunov equation we get

\[ s_1^{-2} Q_2 B_q B'_q \overline{q}_2 + s_1^{-2} \overline{q}_2 B_q B'_q Q_2 + s_1^{-2} \overline{q}_2 B_q D'_q C_q + s_1^{-2} C'_q D_1 B_q \overline{q}_2 = 0 =: \text{RHS} \]

using all-pass equation (11(\(v\)))22.

2b. Controllability and Observability gramians of \(W_\perp\): First note that

\[ W_\perp \triangleq \begin{bmatrix} -A'_q + s_1^{-1} C'_q \overline{q}_2 & s_1^{-1} C'_q \\ s_1^{-1} D_2 \overline{q}_1 - B'_q & s_1^{-1} D'_q_2 \end{bmatrix} \]

implies that

\[ W_\perp \triangleq \begin{bmatrix} A_q - s_1^{-2} \overline{q}_2 C_q \overline{q}_2 & B_q - s_1^{-1} \overline{q}_1 D_2 \overline{q}_2 \\ s_1^{-1} C_q & s_1^{-1} D_2 \overline{q}_2 \end{bmatrix} \]

Hence, the controllability gramian \(Y_w\) of this realization satisfies:

\[(A_q - s_1^{-1} \overline{q}_1 C_q \overline{q}_2) Y_w + Y_w (A'_q - s_1^{-1} C'_q \overline{q}_2) + (B_q - s_1^{-1} \overline{q}_1 D_2) (B'_q - s_1^{-1} D'_q \overline{q}_1) = 0 \]

Substituting

\[ \overline{q}_1 := -s_1^{-1} \overline{p}_2 C'_q \]

this can be written as:

\[(A_q + s_1^{-2} \overline{p}_2 C'_q C_q \overline{q}_2) Y_w + Y_w (A'_q + s_1^{-2} C'_q C_q \overline{p}_2) + B_q B'_q + s_1^{-2} B_q D_2 C_q \overline{q}_2 + s_1^{-2} \overline{q}_2 C'_q D_2 B'_q + s_1^{-4} \overline{p}_2 C'_q D_2 C_q \overline{p}_2 = 0 \]

From all-pass equation (11(\(ii\)))22 we get that \(D'_2 D_2 = D_2 D'_2 = s_1^2 I\).

When this is substituted in the above equation we get:

\[ A_q Y_w + s_1^{-2} \overline{p}_2 C'_q C_q Y_w + Y_w A'_q + s_1^{-2} Y_w C'_q C_q \overline{p}_2 + B_q B'_q + s_1^{-2} B_q D_2 C_q \overline{p}_2 + s_1^{-2} \overline{p}_2 C'_q D_2 B'_q + s_1^{-2} \overline{p}_2 C'_q C_q \overline{p}_2 = 0 \]
Assume $Y_w = P_2 - \overline{P}_2$. Next we show that with this assumption $Y_w$ satisfies the above equation.

\[
\text{LHS : } = A_qP_2 - A_q\overline{P}_2 + s_1^{-2}\overline{P}_2 C_q'q_2C_qP_2 - s_1^{-2}\overline{P}_2 C_q'q_2\overline{P}_2 + P_2 A_q' - \overline{P}_2 A_q' + s_1^{-2}P_2 C_q'q_2C_qP_2 + B_q1B_q' + s_1^{-2}B_q1D_21C_q2\overline{P}_2 + s_1^{-2}\overline{P}_2 C_q'q_2D_21B_q' + s_1^{-2}\overline{P}_2 C_q'q_2\overline{P}_2 = A_qP_2 + P_2 A_q' - A_q\overline{P}_2 - \overline{P}_2 A_q' - s_1^{-2}\overline{P}_2 C_q'q_2\overline{P}_2 + B_q1B_q' = 0 =: \text{RHS}
\]

Now, (11(vi))$_{2,2}$ shows that $B_q1D_21 + P_2 C_q'2 = 0$. Thus,

\[
A_qP_2 + P_2 A_q' - A_q\overline{P}_2 - \overline{P}_2 A_q' - s_1^{-2}\overline{P}_2 C_q'q_2\overline{P}_2 + B_q1B_q' = 0 =: \text{RHS}
\]

The last equation is derived by subtracting the all-pass equation (21(ii))$_{2,2}$ from (11(iv))$_{2,2}$. This gives:

\[
A_qP_2 + P_2 A_q' + B_q1B_q' - A_q\overline{P}_2 - \overline{P}_2 A_q' - \overline{B}_q1\overline{B}_q' = 0
\]

Using $\overline{B}_q1 := -s_1^{-1}\overline{P}_2 C_q'$ this can be written as:

\[
A_qP_2 + P_2 A_q' - A_q\overline{P}_2 - \overline{P}_2 A_q' - s_1^{-1}\overline{P}_2 C_q'q_2\overline{P}_2 + B_q1B_q' = 0
\]

Next we find the observability gramian of the realization of $W_{\perp}$. This is the unique solution of the Lyapunov equation:

\[
(A_q' - s_1^{-1}C_q'q_2\overline{B}_q')X_w + X_w(A_q - s_1^{-1}\overline{B}_q1C_q) + s_1^{-1}C_q'q_2C_q = 0
\]

or, equivalently

\[
(A_q' + s_1^{-2}C_q'q_2\overline{P}_2)X_w + X_w(A_q + s_1^{-2}\overline{P}_2 C_q'q_2) + s_1^{-1}C_q'q_2C_q = 0
\]

Now by definition,

\[
\overline{Q}_1 := \overline{P}_2\overline{R} \quad \text{and} \quad \hat{P}_1 := \overline{Q}_2(\overline{R}')^{-1}
\]

Further, the all-pass equation (21(ii))$_{11}$,

\[
\hat{A}\hat{P}_1 + \hat{P}_1\hat{A}' + \hat{B}\hat{B}' = 0
\]

implies that

\[
(-A_q' - s_1^{-2}C_q'q_2\overline{P}_2)\hat{P}_1 + \hat{P}_1(-A_q - s_1^{-2}\overline{P}_2 C_q'q_2) + s_1^{-2}C_q'q_2C_q = 0
\]

Therefore, $X_w = -\hat{P}_1$. \qed
$V_\perp$ and $W_\perp^\sim$ constructed in proposition 4.4 are parts of inner matrix functions. Theorem 4.3 relies on the construction of two inner complements $v$ and $w^\sim$ so that $(v V_\perp)$ and $(w^\sim W_\perp^\sim)$ are square inner. To find realizations for $v$ and $w$, we can apply Lemma 13.31 from [ZDG96] which uses the gramians of the realizations of $V_\perp$ and $W_\perp^\sim$. This is outlined next, together with concrete realizations of $v$ and $w^\sim$.

**Corollary 4.1.** Let $V_\perp, W_\perp^\sim$ be as defined in proposition 4.4. Then there exists a complementary inner factor of $v$ and a complementary co-inner factor of $w$, respectively, such that

$$V(s) := (v V_\perp)(s), \quad W(s) := (w^\sim W_\perp^\sim)(s)$$

are square inner. Further, $V \in \mathcal{RH}_{\infty}^{p \times p}$ and $W \in \mathcal{RH}_{\infty}^{m \times m}$. Concrete realizations of $v^\sim$ and $w$ are given as:

$$v^\sim = s \begin{bmatrix} -A'_q - s_1^{-2}Q_2 B q_2 B'_q \frac{C'_{q_1} + s_1^{-2}Q_2 B q_2 D'_1}{(D'_1)} \frac{1}{(D'_2)} \end{bmatrix}$$

and

$$w = s \begin{bmatrix} -A'_q - s_1^{-2}C'_{q_2} C q_2 P_2 \frac{(P_2 - P_2)^\dagger B q_1 D'_{12}}{D'_{12}} \\ -B q_1 - s_1^{-2}D'_{21} C q_2 P_2 \frac{1}{D'_{21}} \end{bmatrix}$$

respectively.

**Proof.** This follows immediately from Lemma 13.31 in [ZDG96].

**Observation 4.1.** The pair $(v, w)$ as constructed in corollary 4.1 forms a scaled Schmidt pair corresponding to the largest Hankel singular value of $R^\sim$.

In the final part of this section we develop a state space realisation of the allpass system $\alpha(s)$ defined in the proof of Theorem 4.3 and show that it is anti-inner. The proof is based on a lengthy state space calculation and numerous pole-zero cancellations. We first need the following two results.

**Proposition 4.5.** Let $Q$, $P$ be the observability and the controllability gramians, respectively, of a system having state space realization $G s = (A, B, C)$. Then, (i) $\mathcal{N}(Q) \subseteq \mathcal{N}(C)$ and (ii) $\mathcal{N}(P) \subseteq \mathcal{N}(B')$.  

32
Proof. (i) Let \( \xi_o \in \text{Ker}(Q), \xi_o \neq 0 \). Then, \( Q\xi_o = 0 \). Consider the Lyapunov equation:

\[
A'Q + QA + C'C = 0 \Rightarrow \xi_o'(A'Q + QA + C'C)\xi_o = 0 \Rightarrow C\xi_o = 0
\]

and hence \( \mathcal{N}(Q) \subseteq \mathcal{N}(C) \). A similar argument proves part (ii). \( \square \)

**Proposition 4.6.** In previously defined notation:

(i) \( [I - (Q_2 - \overline{Q}_2)]^\dagger (Q_2 - \overline{Q}_2) C_{q1}' D_{12}^\dagger = 0 \), and 

(ii) \( [I - (P_2 - P_2)]^\dagger (P_2 - P_2) B_{q1} D_{21}^\dagger = 0 \).

Proof. (i) First note that from Proposition 4.4 \( (Q_2 - \overline{Q}_2) \) is the observability gramian of \( (A_q + s_1^{-2}B_{q2}B_{q2}\overline{Q}_2, C_{q1} + s_1^{-2}D_{12}B_{q2}'\overline{Q}_2) \). It follows, using Proposition 4.5 that \( \mathcal{N}[Q_2 - \overline{Q}_2] \subseteq \mathcal{N}[C_{q1} + s_1^{-2}D_{12}B_{q2}'\overline{Q}_2] \), or equivalently, \( \mathcal{R}[C_{q1} + s_1^{-2}\overline{Q}_2 B_{q2} D_{12}'] \subseteq \mathcal{R} [Q_2 - \overline{Q}_2] \). Thus,

\[
\mathcal{R}[C_{q1} + s_1^{-2}\overline{Q}_2 B_{q2} D_{12}'] D_{12}^\dagger = \mathcal{R}[C_{q1}' D_{12}^\dagger] \subseteq \mathcal{R}[C_{q1} + s_1^{-2}\overline{Q}_2 B_{q2} D_{12}']
\]

and hence \( \mathcal{R}[C_{q1}' D_{12}^\dagger] \subseteq \mathcal{R} [Q_2 - \overline{Q}_2] \). The result now follows on noting that \( [I - (Q_2 - \overline{Q}_2)]^\dagger (Q_2 - \overline{Q}_2) \) projects orthogonally onto \( \mathcal{N}[Q_2 - \overline{Q}_2] \). Part (ii) follows dually on noting that \( P_2 - \overline{P}_2 \) is the controllability gramian of the realization of \( W_{\perp}^\perp \) given in Proposition 4.4. \( \square \)

**Proposition 4.7.** The \( s_1 \)-allpass system \( s_1\alpha(s) \in \mathcal{R} C_{l\times l}^{\infty} \) defined in the proof of Theorem 4.3 can be written as a parallel system interconnection \( s_1\alpha(s) = \alpha_1(s) + \alpha_2(s) \),

\[
s_1\alpha(s) = \begin{bmatrix}
A & 0 \\
0 & -A_q' - s_1^{-2}C_{q2}'C_{q2}\overline{P}_2 \\
C_{\alpha_1} & C_{\alpha_2} \\
\end{bmatrix}
\begin{bmatrix}
B_{\alpha_1} \\
B_{\alpha_2} \\
\end{bmatrix}
\]

in which

\[
B_{\alpha_1} := BD_{21}^\dagger + P_3(P_2 - \overline{P}_2)^\dagger B_{q1} D_{21}^\dagger \\
B_{\alpha_2} := (P_2 - \overline{P}_2)^\dagger B_{q1} D_{21}^\dagger \\
C_{\alpha_1} := -(D_{12}^\dagger)'C_{q1} (Q_2 - \overline{Q}_2)^\dagger Q_3' + (D_{12}^\dagger)'C \\
C_{\alpha_2} := -(D_{12}^\dagger)'C_{q1} (Q_2 - \overline{Q}_2)^\dagger R
\]

In particular, \( \alpha \in \mathcal{H}_{\infty}^{l\times l} \) and \( \deg(\alpha) \leq 2n - r \).
Proof. The proof follows a sequence of detailed state-space calculations. First,

\[ v^* H_{11} = \begin{bmatrix} -A'_{1} - s^{-2}Q_{2}B_{q2}B'_{q2} & C'_{q1} + s^{-2}Q_{2}B_{q2}D'_{12} \\ (D'_{12})^{T}C_{q1}(Q_{2} - \mathcal{Q}_{2})^{T} & (D'_{12})^{T} \end{bmatrix} \begin{bmatrix} A & 0 & B \\ 0 & A_{q} & B_{q1} \\ C & C_{q1} & D_{11} \end{bmatrix} \]

\[ = \begin{bmatrix} -A'_{1} - s^{-2}Q_{2}B_{q2}B'_{q2} & C'_{q1} + Z_{1}C & C'_{q1} + Z_{1}C_{q1} & C'_{q1}D_{11} \\ 0 & A & 0 & A_{q} \\ 0 & 0 & A_{q} & B_{q1} \\ (D'_{12})^{T}C_{q1}(Q_{2} - \mathcal{Q}_{2})^{T} & (D'_{12})^{T}C & (D'_{12})^{T}C_{q1} & (D'_{12})^{T}D_{11} \end{bmatrix} \]

in which \( Z_{1} = s^{-2}Q_{2}B_{q2}D'_{12} \), using \( D'_{12}D_{11} = 0 \). Further, using a similarity transformation \( T \):

\[ T := \begin{bmatrix} I & 0 & Q_{2} - \mathcal{Q}_{2} \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \]

we can write:

\[ v^* H_{11} = \begin{bmatrix} -A'_{1} - s^{-2}Q_{2}B_{q2}B'_{q2} & C'_{q1}C + Z_{1}C & C'_{q1} + Z_{1}C_{q1} & (Q_{2} - \mathcal{Q}_{2})B_{q1} + C'_{q1}D_{11} \\ 0 & A & 0 & B \\ (D'_{12})^{T}C_{q1}(Q_{2} - \mathcal{Q}_{2})^{T} & (D'_{12})^{T}C & (D'_{12})^{T}C_{q1} & (D'_{12})^{T}D_{11} \end{bmatrix} \]

\[ = \begin{bmatrix} \Phi_{1} & \Phi_{2} & \Phi_{5} \\ 0 & A & B \\ \Phi_{3} & \Phi_{4} & \Phi_{6} \end{bmatrix} \]

We next form:

\[ v^* H_{11}w = \begin{bmatrix} \Phi_{1} & \Phi_{2} & \Phi_{5} \\ 0 & A & B \\ \Phi_{3} & \Phi_{4} & \Phi_{6} \end{bmatrix} \begin{bmatrix} -A'_{1} - s^{-2}C'_{q2}C_{q2}P_{2} \\ -B'_{q1} - s^{-2}D'_{21}C_{q2}P_{2} \\ (P_{2} - P_{2})^{T}B_{q1}D_{21} \end{bmatrix} \]

\[ = \begin{bmatrix} \Phi_{1} & \Phi_{2} & \Phi_{5}(-B'_{q1} - s^{-2}D'_{21}C_{q2}P_{2}) \\ 0 & A & B(-B'_{q1} - s^{-2}D'_{21}C_{q2}P_{2}) \\ \Phi_{3} & \Phi_{4} & \Phi_{6}(-B'_{q1} - s^{-2}D'_{21}C_{q2}P_{2}) \end{bmatrix} \begin{bmatrix} \Phi_{5}D_{21} \\ \Phi_{5}D_{21} \\ \Phi_{6}D_{21} \end{bmatrix} \]
Now,

\[
A(1, 1) = -A_q' - s_1^2\overline{Q}_2 B_{q2} B_{q2}'
\]

\[
A(1, 2) = C_{q1}' C + s_1^2\overline{Q}_2 B_{q2} D_{12}' C
\]

\[
= -A_q' Q_3' - Q_3 A - s_1^2\overline{Q}_2 B_{q2} B_{q2}' Q_3
\]

by using the all-pass equations (11(v))_{2,1} and (11(iii))_{2,1}. In addition,

\[
A(1, 3) = -(Q_2 - \overline{Q}_2) B_{q1} B_{q1}' - s_1^2(Q_2 - \overline{Q}_2) B_{q1} D_{21}' C_{q2} \overline{P}_2 - C_{q1}' D_{11} B_{q1}'
\]

\[
- s_1^2 C_{q1}' D_{11} D_{21}' C_{q2} \overline{P}_2
\]

\[
= -(Q_2 - \overline{Q}_2) B_{q1} B_{q1}' - s_1^2(Q_2 - \overline{Q}_2) B_{q1} D_{21}' C_{q2} \overline{P}_2 - C_{q1}' D_{11} B_{q1}'
\]

on noticing that \(D_{11} D_{21}' = 0\) (from all-pass (11(ii))_{1,2}). Moreover,

\[
A(2, 3) = -B B_{q1}' - s_1^2 B D_{21}' C_{q2} \overline{P}_2
\]

\[
A(3, 3) = -A_q' - s_1^2 C_{q1}' C_{q2} \overline{P}_2
\]

\[
B(1) = (Q_2 - \overline{Q}_2) B_{q1} D_{21}^\dagger + C_{q1}' D_{11} D_{21}^\dagger
\]

\[
B(2) = B D_{21}^\dagger
\]

\[
B(3) = (\overline{P}_2 - P_2)^\dagger B_{q1} D_{21}^\dagger
\]

\[
C(1) = (D_{12}^\dagger)' C_{q1}(Q_2 - \overline{Q}_2)^\dagger
\]

\[
C(2) = (D_{12}^\dagger)' C
\]

\[
C(3) = -(D_{12}^\dagger)' D_{11} B_{q1}' - s_1^2(D_{12}^\dagger)' D_{11} D_{21}' C_{q2} \overline{P}_2
\]

\[
= -(D_{12}^\dagger)' D_{11} B_{q1}'
\]

The expression for \(C(3)\) is due to the fact that \(D_{11} D_{21}' = 0\) (from all-pass (11(ii))_{1,2}). Apply now, the similarity transformation

\[
T = \begin{bmatrix}
I & Q_3' & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix}
\Rightarrow T^{-1} = \begin{bmatrix}
I & -Q_3 & 0 \\
0 & I & 0 \\
0 & 0 & I
\end{bmatrix}
\]

Then, we have that \(A(1, 2) = -A(1, 1) Q_3' + A(1, 2) + Q_3' A(2, 2) = 0\). Further,

\[
A(1, 3) = A(1, 3) + Q_3' A(2, 3)
\]

\[
= -(Q_2 - \overline{Q}_2) B_{q1} B_{q1}' - s_1^2(Q_2 - \overline{Q}_2) B_{q1} D_{21}' C_{q2} \overline{P}_2
\]

\[
- C_{q1}' D_{11} B_{q1}' + Q_3'(-B B_{q1}' - s_1^2 B D_{21}' C_{q2} \overline{P}_2)
\]

\[
= -(Q_2 - \overline{Q}_2) B_{q1} B_{q1}' - s_1^2(Q_2 - \overline{Q}_2) B_{q1} D_{21}' C_{q2} \overline{P}_2
\]

\[
- C_{q1}' D_{11} B_{q1}' - Q_3' B B_{q1}' - s_1^2 Q_3' B D_{21}' C_{q2} \overline{P}_2
\]

\[
= 0
\]
Next note that the all-pass equation \((D'_{H}C_{H} + B'_{H}Q_{H} = 0)_{(1,2)}\) implies that:

\[
B'Q_3 + Q_1Q_2 = -D'_{11}C_{q1} - D'_{21}C_{q2}
\]
\[
\Rightarrow Q'_3B = -Q'_2B_{q1} - C'_{q1}D_{11} - C'_{q2}D_{21}
\]
\[
\Rightarrow Q'_3BB'_q = -Q'_2B_{q1}B'_q - C'_{q1}D_{11}B'_q - C'_{q2}D_{21}B'_q
\]
\[
\Rightarrow Q'_3BD'_{21} = -Q'_2B_{q1}D'_{21} - C'_{q1}D_{11}D'_{21} - C'_{q2}D_{21}D'_{21} = Q'_2P'_2C'_{q2} - s^2Q_{1}C_{q2}
\]

Note from all-pass \((11(ii))\) that \(D_{11}D'_{21} = 0\) and that \(D_{21}D'_{21} = s^2I\). Further, taking \((D_{H}B'_{H} + C_{H}P'_{H} = 0)_{(2,2)}\) we deduce that \(BD'_{21} = P'_2C'_{q2}\). Thus,

\[
A(1,3) = -(Q_2 - Q'_2)B_{q1}B'_q - s_1^{-2}(Q_2 - Q'_2)B_{q1}D'_{21}C_{q2}P_2 - C'_{q1}D_{11}B'_q
\]
\[
+ Q'_2B_{q1}B'_q + C'_{q1}D_{11}B'_q + C'_{q2}D_{21}B'_q - s_1^{-2}Q'_2P'_2C'_{q2}C_{q2}P_2
\]
\[
+ C'_{q2}C_{q2}P_2
\]
\[
= Q_2B_{q1}B'_q - s_1^{-2}(Q_2 - Q'_2)B_{q1}D'_{21}C_{q2}P_2 + C'_{q2}D_{21}B'_q
\]
\[
- s_1^{-2}Q'_2P'_2C'_{q2}C_{q2}P_2 + C'_{q2}C_{q2}P_2
\]
\[
= Q_2B_{q1}B'_q - s_1^{-2}Q_2B_{q1}D'_{21}C_{q2}P_2 + s_1^{-2}Q'_2B_{q1}D'_{21}C_{q2}P_2
\]
\[
+ C'_{q2}D_{21}B'_q - s_1^{-2}Q'_2P'_2C'_{q2}C_{q2}P_2 + C'_{q2}C_{q2}P_2
\]
\[
= Q_2B_{q1}B'_q - s_1^{-2}Q_2B_{q1}D'_{21}C_{q2}P_2 + s_1^{-2}Q'_2B_{q1}D'_{21}C_{q2}P_2
\]
\[
- C'_{q2}C_{q2}P_2 - s_1^{-2}Q'_2P'_2C'_{q2}C_{q2}P_2 + C'_{q2}C_{q2}P_2
\]

by observing that \(D_{21}B'_q = -C_{q2}P_2\) (from all-pass \((11(vi))_{(22)}\)). Now, take the last two terms of the above equation:

\[
-s_1^{-2}Q'_2P'_2C'_{q2}C_{q2}P_2 + C'_{q2}C_{q2}P_2 = s_1^{-2}(s^2I - Q'_2P'_2)C'_{q2}C_{q2}P_2
\]
\[
= s_1^{-2}(Q'_3P_3)C'_{q2}C_{q2}P_2
\]
\[
= -s_1^{-2}Q'_2BD'_{21}C_{q2}P_2
\]

where all-pass \((11(vi))_{(21)}\) gives \(D_{21}B' = -C_{q2}P'_3 \Rightarrow P'_3C'_{q2} = -BD'_{21}\). Thus, we have

\[
A(1,3) = \overline{Q}_2B_{q1}B'_q - s_1^{-2}Q_2B_{q1}D'_{21}C_{q2}P_2 + s_1^{-2}\overline{Q}_2B_{q1}D'_{21}C_{q2}P_2
\]
\[
- C'_{q2}C_{q2}P_2 - s_1^{-2}Q'_2BD'_{21}C_{q2}P_2
\]
\[
= \overline{Q}_2B_{q1}B'_q + (-s_1^{-2}Q_2B_{q1} + s_1^{-2}\overline{Q}_2B_{q1} - s_1^{-2}Q'_3B)D'_{21}C_{q2}P_2
\]
\[
- C'_{q2}C_{q2}P_2
\]

by taking the term \(D'_{21}C_{q2}P_2\) as common factor. Now, recall from all-pass \((11(v))_{(12)}\) \(B'Q_3 + B'_qQ_2 + D_{11}C_{q1} + C'_{q2}D_{21} = 0\) or equivalently \(Q'_3B +
\[ Q_2 B_q = -C'_q D_{1q} - C'_q D_{21}. \] In addition, from all-pass equations we get that \( D_{11} D'_{21} = 0 \) and that \( D_{21} D'_{21} = s^2 I \). Hence, we have

\[
A(1, 3) = \overline{Q}_2 B_q B'_q - C'_q C_{q2} P_2 + \left(-Q_2 B_q + \overline{Q}_2 B_q - Q_3 B\right) s^{-2} D'_{21} C_{q2} P_2
\]

\[
= \overline{Q}_2 B_q B'_q - C'_q C_{q2} P_2 + s^{-2} \overline{Q}_2 B_q D'_{21} C_{q2} P_2 + C'_q C_{q2} P_2
\]

but \( B_q D'_{21} = -P_2 C'_{q2} \) and \( B'_q B_q = -A_q P_2 - P_2 A'_q - B'_q B_q \). Substituting,

\[
A(1, 3) = \overline{Q}_2 B_q B'_q - C'_q C_{q2} P_2 + s^{-2} \overline{Q}_2 B_q D'_{21} C_{q2} P_2 + C'_q C_{q2} P_2
\]

\[
= \overline{Q}_2(-A_q P_2 - P_2 A'_q - B'_q B_q) - C'_q C_{q2} P_2 - s^{-2} \overline{Q}_2 P_2 C'_q C_{q2} P_2
\]

\[
+ C'_q C_{q2} P_2
\]

\[
= \overline{Q}_2(-P_2 - P_2 A'_q - B'_q B_q - s^{-2} P_2 C'_q C_{q2} P_2) - C'_q C_{q2} P_2
\]

\[
+ C'_q C_{q2} P_2
\]

\[
= \overline{Q}_2(\overline{P}_2 - P_2) A'_q + A_q (\overline{P}_2 - P_2) + s^{-2} (\overline{P}_2 - P_2) C'_q C_{q2} P_2
\]

\[
- C'_q C_{q2} P_2 + C'_q C_{q2} P_2 =: \Phi
\]

since \( -B_q B'_q = A_q \overline{P}_2 + \overline{P}_2 A'_q + s^{-2} \overline{P}_2 C'_q C_{q2} P_2 \). Further,

\[
C T^{-1} = \begin{bmatrix} I & -Q_3^1 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}
\]

\[
= \begin{bmatrix} I & -Z_2 Q_3^1 & (D_{12}^1)' D_{11} B_q' \\ 0 & I & 0 \end{bmatrix}
\]

where \( Z_2 := (D_{12}^1)' C_{q1} (Q_2 - \overline{Q}_2) \) and

\[
TB = \begin{bmatrix} I & Q_3^1 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}
\]

\[
= \begin{bmatrix} (Q_2 - \overline{Q}_2) B_q D_{21}^1 + C'_q D_{11} D_{21}^1 \\ (\overline{P}_2 - P_2)^\dagger B_q D_{21}^1 \end{bmatrix}
\]

\[
= \begin{bmatrix} (Q_2 B_q - \overline{Q}_2 B_q + C'_q D_{11} + Q_3 B) D_{21}^1 \\ (\overline{P}_2 - P_2)^\dagger B_q D_{21}^1 \end{bmatrix}
\]

\[
= \begin{bmatrix} -C'_q D_{21} - \overline{Q}_2 B_q D_{21}^1 \\ (\overline{P}_2 - P_2)^\dagger B_q D_{21}^1 \end{bmatrix}
\]

\[
= \begin{bmatrix} -\overline{Q}_2 B_q D_{21}^1 \\ (\overline{P}_2 - P_2)^\dagger B_q D_{21}^1 \end{bmatrix}
\]

37
So, \( v^\sim H_{11} w \) has a realization:

\[
\begin{bmatrix}
- A_q' - s_1^{-2} \overline{Q}_2 B_q B_{q2}' & 0 & \Phi \\
0 & A & - B B_{q1}' - s_1^{-2} B D_{21}' C_{q2} \overline{P}_2 \\
0 & 0 & - A_q' - s_1^{-2} C_{q2}' C_q \overline{P}_2 \\
\end{bmatrix}
= \begin{bmatrix}
- \overline{Q}_2 B_{q1} D_{21}^\perp \\
- Z_2 Q_3' + (D_{12}^\perp)' C \\
0 \\
- Z_2 Q_3' + (D_{12}^\perp)' C \\
\end{bmatrix}
= \begin{bmatrix}
- (P_2 - P_2)' B_q D_{21}^\perp \\
0 \\
0 \\
0 \\
\end{bmatrix}
\]

where we have used the transformation

\[
T = \begin{bmatrix}
I & 0 & \overline{Q}_2 (P_2 - P_2) \\
0 & I & 0 \\
0 & 0 & I \\
\end{bmatrix} \Rightarrow T^{-1} = \begin{bmatrix}
I & 0 & - \overline{Q}_2 (P_2 - P_2) \\
0 & I & 0 \\
0 & 0 & I \\
\end{bmatrix}
\]

Also,

\[
C(3) = - (D_{12}^\perp)' C_{q1} (Q_2 - \overline{Q}_2)' \overline{Q}_2 (P_2 - P_2) - (D_{12}^\perp)' D_{11} B_{q1}' \\
B(1) = - \overline{Q}_2 B_{q1} D_{21}^\perp + \overline{Q}_2 (P_2 - P_2)' (P_2 - P_2) B_q D_{21}^\perp \\
= - \overline{Q}_2 \left\{ I - (P_2 - P_2)' (P_2 - P_2) \right\} B_q D_{21}^\perp = 0
\]
due to corollary 4.6. Further,

\[
A(1, 3) = \overline{Q}_2 (P_2 - P_2) \left[ - A_q' - s_1^{-2} C_{q2}' C_q \overline{P}_2 \right] - [ - A_q' - s_1^{-2} \overline{Q}_2 B_q B_{q2}' ] \overline{Q}_2 (P_2 - P_2) \\
+ \overline{Q}_2 (P_2 - P_2) A_q' + \overline{Q}_2 A_q (P_2 - P_2) + s_1^{-2} \overline{Q}_2 (P_2 - P_2) C_{q2}' C_q \overline{P}_2 \\
+ C_{q2}' C_q \overline{P}_2 (P_2 - P_2) \\
= 0
\]

Hence, \( v^\sim H_{11} w \) has a realization:

\[
\begin{bmatrix}
A & - B B_{q1}' - s_1^{-2} B D_{21}' C_{q2} \overline{P}_2 \\
0 & - A_q' - s_1^{-2} C_{q2}' C_q \overline{P}_2 \\
- Z_2 Q_3' + (D_{12}^\perp)' C & - Z_2 Q_2 (P_2 - P_2) - (D_{12}^\perp)' D_{11} B_{q1}' \\
\end{bmatrix}
= \begin{bmatrix}
B_{D_{21}}^\perp \\
(P_2 - P_2)' B_q D_{21}^\perp \\
- (P_2 - P_2)' B_q D_{21}^\perp \\
\end{bmatrix}
\]

Finally, applying the similarity transformation \( T := \begin{bmatrix} I & P_3 \\ 0 & I \end{bmatrix} \) shows that:

\[
s_1 \alpha(s) = \begin{bmatrix}
A & 0 & B_{\alpha_1} \\
0 & - A_q' - s_1^{-2} C_{q2}' C_q \overline{P}_2 & B_{\alpha_2} \\
C_{\alpha_1} & C_{\alpha_2} & (D_{12}^\perp)' D_{11} D_{21}^\perp \\
\end{bmatrix}
\]

38
where

\[ B_{\alpha 1} := BD_{\perp 21}^1 + P_3(\overline{P}_2 - P_2)^\dagger B_{q1}D_{\perp 21}^1 \]
\[ B_{\alpha 2} := (\overline{P}_2 - P_2)^\dagger B_{q1}D_{\perp 21}^1 \]
\[ C_{\alpha 1} := -(D_{12}^1)^\dagger C_{q1}(Q_2 - \overline{Q}_2)^\dagger Q_3^\prime + (D_{12}^1)^\dagger C \]
\[ C_{\alpha 2} := -(D_{12}^1)^\dagger C_{q1}(Q_2 - \overline{Q}_2)^\dagger R \]

This completes the proof. \(\square\)

5 Application: Super-optimally robust controllers

The problem of robust stabilization of feedback systems subject to unstructured (or structured) uncertainties has been studied extensively by the control community for more than two decades. In general, for the multivariable case, the maximum robust stability radius for unstructured perturbations reduces to a Nehari extension problem, or equivalently, to the minimization of the \(\mathcal{H}_\infty\) norm of an appropriate closed-loop transfer-matrix. In the multivariable case the solution is not unique, and in general there is a continuum of controllers which achieves the largest achievable robust stability radius, say \(\epsilon^{\star}\). Thus a natural question arising is how to choose between them. In a certain sense, all optimal controllers are equivalent, as it is well-known that destabilizing perturbations of norm \(\epsilon^{\star}\) can be constructed on the boundary of the uncertainty set for each individual optimal controller (or, indeed, for each suboptimal controller achieving a robust stabilization radius \(\epsilon < \epsilon^{\star}\)). Thus, the only way of distinguishing between the multitude of optimal controllers is to impose additional structure on the uncertainty set. Note that in practice, known uncertainty structure is often ignored (or subsumed under a more conservative model uncertainty description) to simplify the design problem which otherwise would be intractable.

A natural way of imposing uncertainty structure for the multivariable maximum robust stabilisation problem is to use directionality information. It can be shown [GHJ00] that for the set of all optimal controllers a “weakest perturbation direction” exists, along which uniformly destabilizing perturbations of norm \(\epsilon^{\star}\) can be constructed, i.e. boundary perturbations which destabilizes the feedback loop for every optimal (“maximally robust”) controller. In the additive perturbation case, this direction is described by the Schmidt pair of a Hankel operator with symbol the nominal plant, corresponding to its smallest singular value. Thus, it is futile to attempt to extend the robust stabilization radius in this direction, using an optimal controller.
By restricting the projection of the uncertainty set in this “weakest perturbation direction”, it may be shown [GHJ00] that the robust stability radius increases maximally from $\epsilon^\star$ in every other direction, provided the super-optimal controller is used. The increase depends on the gap between the first two super-optimal levels and the severity of the imposed parametric constraint. Thus, the super-optimal controller guarantees the stabilization not only of all perturbations within the ball of the largest possible radius $\epsilon^\star$, but potentially also a much wider class outside this ball.

A second alternative is to consider block-diagonal structures of the type arising in $\mu$-analysis, a problem that is computationally intractable. In this case, our method can be used to provide easily-computable upper bounds on the structured singular value, by optimally embedding the block-structured perturbation set within the structured set described in the previous paragraph [GHJ00]. Extensions of this method may be employed to obtain tight upper bounds of $\mu$ which outperform the standard convex upper bounds provided by “D-iteration” and LMI techniques [JHMG]. In these type of problems it is vital to assume that the multiplicity of the largest Hankel singular value is larger than one, which is part motivates the detailed analysis of the general SODP problem presented in this work.

In [GHJ00] the problem was considered in the framework of the first two super-optimal levels. The main objective of future work is to extend this analysis to multiple super-optimal levels and to additional unstructured uncertainty models (co-prime, multiplicative, etc), or even models with a general LFT description. Further, research will also be carried out in the area of the graph-metric and its connection to super-optimization for which a few preliminary results have been published [Nym95].

6 Conclusions

By means of conclusions we summarize the main contributions of this work:

- We have presented an explicit solution to the 1-block SODP which is easily implementable using state-space techniques. All assumptions made in previous work (minimal realization of the system which is approximated, non-repeated largest singular value of the associated Hankel operator, invertibility of certain gramians arising at intermediate steps of the algorithm) have been removed.

- The solution methodology is based on all-pass dilation techniques [JL93] and provides considerable conceptual and numerical simplifica-
tions compared to existing methods. In particular, the diagonalisation of the optimal solution set, normally carried out via the Schmidt pair of the Hankel operator associated with the problem, is now performed by exploiting the properties of an optimal and a suboptimal Nehari generator, which are constructed directly from the data of the problem. As a result, all preliminary steps requiring a sequence of Schmidt vector scalings are completely avoided and related technical issues do not arise.

- By exhibiting the intimate relation between the stabilising solutions of two algebraic Riccati equations, a detailed state-space analysis of the algorithm is developed. Realizations of all systems defined in the construction of the superoptimal solution are obtained; these are shown via computational examples to be generically minimal, i.e. for problems without a special structure.

- We have briefly discussed applications of super-optimization in the areas of robust stability and design. These will be developed further in the second part of this work.

- The solution to the SODP has been presented in a tutorial form which is largely self-contained.

In future research we will attempt to establish further links between super-optimization and robust stabilisation. A number of preliminary results in this direction have already been obtained for unstructured co-prime factorization models as discussed in section 5. These and other related issues will be reported in a future publication.

References


