Dual periodicity in $l_1$-norm minimisation problems

Robin D. Hill*

Sir Lawrence Wackett Aerospace Centre, School of Mathematical and Geospatial Science, RMIT University, Latrobe St., Melbourne, 3001, Australia

Received 25 October 2005; received in revised form 3 September 2007; accepted 12 November 2007
Available online 9 January 2008

Abstract

The topic of this paper is the discrete-time $l_1$-norm minimisation problem with convolution constraints. We find primal initial conditions for which the dual optimal solution is periodic. Periodicity of the dual optimal solution implies satisfaction of a simple linear recurrence relation by the primal optimal solution.© 2007 Elsevier B.V. All rights reserved.

Keywords: $l_1$-norm minimisation; Linear programming; Duality; Convolution constraints

1. Introduction

Suppose we are given a linear, causal, time-invariant system with rational $Z$-transform. A non-zero initial condition is also specified. The objective is to minimise the sum of the one norms of the scalar input and output signals $x = (x_k)_{k=1}^\infty$ and $y = (y_k)_{k=1}^\infty$. This problem is difficult because in general the mapping from the initial condition to the optimal $x$ and $y$ signals is nonlinear. This contrasts with the two-norm case. If the sum of the two norms is minimised, instead of the sum of the one norms, we have a linear quadratic regulator problem, where there is linear dependence on initial conditions.

The long-term behaviour of sequences $x$ and $y$ that are optimal for the one-norm problem has been until now known only for a few special cases. We would like to find explicit maps which when iterated produce the optimising sequences. As these maps are nonlinear in general, a natural first step is to find initial conditions for which the long-term behaviour of optimal solutions can be understood. We determine special primal initial conditions, which we term dual-periodic initial states, for which the dual optimising sequences are periodic, in which case the primal optimising sequences are generated by a linear recurrence relation. Examples show many dual-periodic initial states, most of them repelling; for the problems investigated thus far there appears to be at most one dual-periodic initial state that is attracting.

* Tel.: +61 3 9925 2283; fax: +61 3 9925 1784.
E-mail address: r.hill@rmit.edu.au.

A preliminary version of some of the results presented here is in [7].

For the $l_1$-norm minimisation methodology and applications see [10,1,3,9] and their cited references. Most research to date has been on finding approximate solutions. An approach that seeks exact solutions is the delay augmentation method, discussed in [3]. Convex optimisation and dynamic programming are used in [2].

2. Formulation

2.1. Terminology

Denote by $\mathbb{R}^n$ the $n$-dimensional real space. The symbol $0_{p \times k}$ denotes a $p \times k$ matrix all of whose elements are zero. A $p \times k$ matrix $M$ will sometimes have its dimension made explicit by the notation $M_{p \times k}$. The $p \times p$ identity matrix is denoted $I_p$. The set of positive integers is denoted $\mathbb{N}^+$ and $\mathbb{N} = \mathbb{N}^+ \cup \{0\}$. The $l_1$-norm of a vector sequence $y = (y_k)_{k=1}^\infty$ is defined as $\|y\|_1 = \sum_{k=1}^\infty |y_k|$ whenever the series exists. The Banach space of absolutely-summable sequences, equipped with the $l_1$-norm, is denoted $l_1$. The space of continuous linear functionals on $l_1$, that is the dual of $l_1$, is denoted $l_\infty$; it is the space of bounded sequences with the norm $\|y\|_\infty := \sup_{k} |y_k|$. For sequences $y$ and $y^*$ we write $(y, y^*) := \sum_{k=1}^\infty y_k y_k^*$ whenever the series converges. For $y, y^* \in \mathbb{R}^n$, $(y, y^*) := \sum_{k=1}^n y_k y_k^*$. The $Z$-transform of an arbitrary sequence $y = (y_k)_{k=1}^\infty$ is defined to be $\hat{y}(z) = \sum_{k=1}^\infty y_k z^{k-1}$, where $z$ lies within the radius of convergence of the series. For any vector $y$, $y = 0$
means that all components of \( y \) are equal to zero, and \(|y| \leq 1\) means \(|y_i| \leq 1\) for all components \( y_i \). The function \text{sgn} of a real number \( x \) is defined to be 1 if \( x > 0 \), \(-1\) if \( x < 0 \), and 0 if \( x = 0 \). The \text{sgn} of a vector with components \((y_i)_{i=1}^n\) is the vector with components \((\text{sgn}(y_i))_{i=1}^n\). If, for \( y = (y_k)_{k=1}^\infty \in l_\infty \), there is an integer \( n \) such that \( y_n \neq 0 \) and \( y_k = 0 \) for \( k > n \), then \( y \) has length \( n \). Given a vector \( y \) and any \( s \in \mathbb{N}^+ \), \( t \in \mathbb{N}^+ \) satisfying \( s < t \), we denote \((y_s, y_{s+1}, \ldots, y_t)\) by \( y_{(s:t)} \). If \( s, t, q, r \in \mathbb{N}^+ \), \( 1 < s < t \) and \( 1 < q < r \) then \( M_{(s,t);(q,r)} \) is a matrix composed of row \( s \) to row \( t \), and of columns \( q \) to \( r \), of the matrix \( M \) having at least \( t \) rows and at least \( r \) columns. Square brackets are used when it is important to distinguish between row and column vectors. For example, \( y = (y_1, y_2) \in \mathbb{R}^n \times \mathbb{R}^n \) can be written in matrix equations as either \( \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{2 \times 1} \) or \( \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{1 \times 2n} \). The concatenation of \((y_1, x_1) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_1}\) and \((y_2, x_2) \in \mathbb{R}^{n_2} \times \mathbb{R}^{n_2}\) is defined to be the vector \((y_3, x_3) \in \mathbb{R}^{n_1+n_2} \times \mathbb{R}^{n_1+n_2}\) for which \( y_3 = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \) and \( x_3 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \), with \( y_1, y_2, x_1, x_2 \) and \( x_3 \) all column vectors. In a similar fashion more than two ordered vector pairs can also be concatenated into a single ordered vector pair. The \textit{infinite periodic block} \((y_{(0:k)}; x_{(0:k)}) \in \mathbb{R}^{\infty \times \mathbb{R}^{\infty}} \) of \((y, x) \in \mathbb{R}^n \times \mathbb{R}^n \) is the concatenation of \((y, x) \in \mathbb{R}^n \times \mathbb{R}^n \) an infinite number of times with itself. The superscript T denotes transpose. Notation for a basis is in Section 4.1.

### 2.2. Problem description

The decision vectors for our restricted version of the general \( l_1 \) problem are two vectors in \( l_1 \), denoted \( y \) and \( x \). The convolution constraints are

\[
\hat{d} \hat{y} + \hat{n} \hat{x} = \hat{b},
\]

where \( \hat{d}, \hat{y}, \hat{x}, \) and \( \hat{b} \) are polynomials with real coefficients,

\[
\hat{n}(z) = n_1 + n_2 z + n_3 z^2 + \cdots + n_{l+1} z^l
\]

\[
\hat{d}(z) = d_1 + d_2 z + d_3 z^2 + \cdots + d_{l+1} z^l
\]

\[
\hat{b}(z) = b_1 + b_2 z + \cdots + b_{l+1} z^{l+1},
\]

\(n_{l+1}\) and \(d_{l+1}\) are not both zero, and \( l \geq 1 \) is a positive integer. We assume that neither \( \hat{n}(z) \) nor \( \hat{d}(z) \) have zeros lying on the unit circle in the complex plane, and that \( \hat{n}(z) \) and \( \hat{d}(z) \) have no zeros in common.

The \( l_1 \)-norm minimisation problem we investigate is

\[
\mathcal{P}(b) : \begin{cases} 
\min_{y \in \ell_1^1, x \in \ell_1^1} & \|y\|_1 + \|x\|_1 \\
\text{s.t.} & d \hat{y} + \hat{n} \hat{x} = b.
\end{cases}
\]

The vector \( b \) specifies an initial state for the discrete-time dynamic system represented by (1). The notation \( \mathcal{P}(\cdot) \) will be used to denote a problem for which the initial state is arbitrary. A pictorial representation of the system being optimised is given in Fig. 1.

It can be shown that a solution to \( \mathcal{P}(\cdot) \) with finite cost is guaranteed to exist. There is a stronger conjecture, namely that all optimising vectors \((y, x)\) for \( \mathcal{P}(\cdot) \) have rational \( Z \)-transforms. This conjecture remains open.

### 2.3. Linear programming formulation of problem \( \mathcal{P}(b) \)

Using block matrix notation, the problem \( \mathcal{P}(b) \) can be written as

\[
\mathcal{P}(b) : \begin{cases} 
\min_{y \in \ell_1^1} & \|y\|_1 + \|x\|_1 \\
\text{s.t.} & [D \quad N] [y \quad x] = [b_1, \ldots, b_T]^T,
\end{cases}
\]

where \( D \) is the infinite-dimensional lower-triangular toeplitz matrix having \((d_1, \ldots, d_{l+1}, 0, \ldots)\) as its first column, and \( N \), defined similarly, has first column \((n_1, \ldots, n_{l+1}, 0, 0, \ldots)\). Also \( b := [b_1, \ldots, b_T]^T \). For all of the convex optimisation problems considered in this paper the term \textit{feasible solution} means a solution which satisfies all of the constraints, both equality and inequality (if any).

### 2.4. Toeplitz and circulant matrix notation

Define

\[
D_{UT} := \begin{bmatrix} d_{l+1} & d_1 & \cdots & d_2 \\ 0 & d_{l+1} & \vdots & \vdots \\ 0 & 0 & d_{l+1} & \vdots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}_{l \times l}
\]

\[
D_{LT} := \begin{bmatrix} d_1 & 0 & 0 & 0 \\ \vdots & d_1 & 0 & \vdots \\ d_{l-1} & \vdots & d_1 & 0 \\ d_1 & \cdots & \cdots & \cdots \end{bmatrix}_{l \times l}
\]

Fix an integer \( p \geq 2l \). The North-West corner submatrix of \( D \) with dimension \( p \times p \) is \( D_{(1:p,1:p)} \). Schematically

\[
D_{(1:p,1:p)} = \begin{bmatrix} D_{LT} & D_{UT} & D_{LT} \\ D_{UT} & D_{LT} & D_{UT} \\ \vdots & \vdots & \ddots \end{bmatrix}_{p \times p}
\]

where here and later it is \textit{not} necessarily the case that \( p \) is an integer multiple of \( l \). Denote by \( D_C(p) \) the circulant matrix of dimension \( p \times p \) whose first column is \((d_1, d_2, \ldots, d_{l+1}, 0, \ldots, 0)\). That is

\[
D_C(p) := \begin{bmatrix} 1 & & & \\ D_{LT} & D_{UT} & D_{LT} & \vdots \\ \vdots & D_{UT} & D_{LT} & \vdots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}_{p \times p}
\]

The matrices \( N_{UT}, N_{LT} \) and \( N_C(p) \) are defined similarly. Non-singularity of the circulant matrices \( D_C(p) \) and \( N_C(p) \) is a
consequence of the absence of zeros of \( \dot{z}(t) \) and \( \dot{h}(t) \) on
the unit circle in the complex plane (see [5], p. 89).

For later use define
\[
S := \begin{bmatrix} D_{LT} & N_{LT} \\ D_{UT} & N_{UT} \end{bmatrix},
B := \{D_{LT} N_{UT} - N_{LT} D_{UT}\}^{-1}.
\]
The 2\( l \times 2 \) matrix \( S \) is the Sylvester matrix for the polynomials
\( d(z) \) and \( h(z) \), and the \( l \times l \) matrix \( B \) is the inverse of the
Bezoutian matrix of \( d(z) \) and \( h(z) \), with the columns reversed.
It is well known that \( S \) is non-singular, and \( B \) exists, if and only if
\( d(z) \) and \( h(z) \) are coprime.

The identity
\[
\begin{bmatrix} D_{LT} & N_{LT} \\ D_{UT} & N_{UT} \end{bmatrix} [D_{LT} N_{UT} - N_{LT} D_{UT}] =
\begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}
\]
follows from the identity \( D_{LT} N_{UT} - N_{LT} D_{UT} = N_{UT} D_{LT} - D_{UT} N_{LT} \),
which is one form of the Gohberg–Semencul formulas; see ([6], p. 206) for a proof. Then (4) implies
\[
S^{-1} = \begin{bmatrix} N_{UT} & -N_{LT} \\ -D_{UT} & D_{LT} \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}.
\]

3. Duality

We denote the dual of \( P(b) \) by \( D(b) \). The maximization problem \( D(b) \) was first described in [4] and in the next Section
we describe \( D(b) \) in the notation of this paper. The dual pairing between \( P(b) \) and \( D(b) \) is a consequence of the minimum-norm
theorem.

3.1. \( D(b) \) — the infinite dual of \( P(b) \)

**Proposition 1.** Select any \((\tilde{y}, \tilde{x})\) feasible for \( P(b) \). A dual to
\( P(b) \), denoted \( D(b) \), is
\[
D(b) := \max \left\{ \begin{array}{c} y^* \in \mathbb{R}^p \times \mathbb{R}^p \\ s.t. \ N^T y^* - D^T x^* = 0, \|y^*\|_{\infty} \leq 1, \|x^*\|_{\infty} \leq 1 \end{array} \right\}
\]
If \((y, x)\) and \((y^*, x^*)\) are feasible for \( P(b) \) and \( D(b) \), respectively, then a necessary and sufficient condition that
both be optimal is that they be aligned; that is, for all \( k \in \mathbb{N}^+ \),
\[
y_k > 0 \implies y_k^* = 1, x_k > 0 \implies x_k^* = 1
\]
\[
y_k < 0 \implies y_k^* = -1, x_k < 0 \implies x_k^* = -1
\]
\[
|y_k^*| < 1 \implies y_k = 0, |x_k^*| < 1 \implies x_k = 0.
\]
The dual objective function will now be expressed in terms of primal
and dual initial states. First some notation. For \((y, x) \in l \times l\) satisfying
\( Dy + N x = 0 \) define the primal state at time \( k \) to be
\( b^{(k)}(y, x) := D_{LT} y_{(k+1,l+1;k)} + N_{LT} x_{(k+1,l+1;k)} \).
Note that this notation is consistent in that, if \((y, x)\) is feasible
for \( P(b) \), then \( b^{(0)}(y, x) = b^{(0)} \). For \((y, x) \in \mathbb{R}^p \times \mathbb{R}^p , p \geq 2l \),
satisfying \( D_{l+1,p+1,p+1} y + N_{l+1,p+1,p+1} x = 0 \), the primal state at
time \( k \) is defined by
\[
b^{(k)}(y, x) := \begin{cases} D_{LT} y_{(k+1,l+1;k)} + N_{LT} x_{(k+1,l+1;k)} & \text{for } 0 \leq k \leq p-l \\ -D_{LT} y_{(k-l+1;k)} - N_{LT} x_{(k-l+1;k)} & \text{for } p-l + 1 \leq k \leq p \end{cases}
\]
For \((y^*, x^*) \in l \times l \) satisfying \( N^T y^* - D^T x^* = 0 \) define
the dual state at time \( k \) to be \( b^{*(k)}(y^*, x^*) := N^T y_{(k+1,k+1)} -
D^T x_{(k+1,k+1)} \). For \((y^*, x^*) \in \mathbb{R}^p \times \mathbb{R}^p , p \geq 2l \), satisfying
\( N^T (1,p-l,1,p) y^* - D^T (1,p-l,1,p) x^* = 0 \), define the dual state at
time \( k \) to be
\[
b^{*(k)}(y^*, x^*) := \begin{cases} N^T y_{(k+1,k+1)} - D^T x_{(k+1,k+1)} & \text{for } 0 \leq k \leq p-l \\ -N^T y_{(k-l+1,k+1)} + D^T x_{(k-l+1,k+1)} & \text{for } p-l + 1 \leq k \leq p \end{cases}
\]

**Proposition 2.** For any \((\tilde{y}, \tilde{x})\) feasible for \( P(b) \), and any
\((y^*, x^*) \in l \times l \) satisfying \( N^T y^* - D^T x^* = 0 \), \( \langle \tilde{y}, y^* \rangle +
\langle \tilde{x}, x^* \rangle = \langle \{B b, b^{*(0)}(y^*, x^*)\} \rangle \).

**Proof.** Select \((\tilde{y}, \tilde{x})\) feasible for \( P(b) \) with the lengths of \( \tilde{y} \)
and \( \tilde{x} \) at most \( l \). The first \( 2l \) rows of the constraints for \( P(b) \) will then be
\[
\begin{bmatrix} D_{LT} & N_{LT} \\ D_{UT} & N_{UT} \end{bmatrix} [\tilde{y}_{(1:l)}(1:l)] = [b^T, 0_{1 \times l}]^T,
\]
all later constraints being satisfied as identities. By (8) and (5), for any \((y^*, x^*) \in l \times l \),
\( \langle \tilde{y}, y^* \rangle + \langle \tilde{x}, x^* \rangle = \langle y^*_{(1:l)}, x^*_{(1:l)} \rangle S^{-1} [b^T, 0_{1 \times l}]^T =
\langle y^*_{(1:l)}, x^*_{(1:l)} \rangle [N_{UT} - N_{LT} D_{UL} - N_{LT} D_{LT}] [B \ 0] [b^T, 0_{1 \times l}]^T = b^{*(0)}(y^*, x^*) \). \( \Box \)

In view of Proposition 2, \( D(b) \) can be equivalently stated as
\[
D(b) := \max \left\{ \begin{array}{c} y^* \in \mathbb{R}^p \times \mathbb{R}^p \end{array} \right\} \left\{ \begin{array}{c} B b, b^{*(0)}(y^*, x^*) \\ \text{s.t. } N^T y^* - D^T x^* = 0, \|y^*\|_{\infty} \leq 1, \|x^*\|_{\infty} \leq 1 \end{array} \right\}
\]

3.2. \( D_{\perp p}(b; p) \) — a finite-dimensional modification of \( D(b) \)

Let \( p \geq 2l \) be an integer. We construct a finite-dimensional convex programming problem related to \( D(b) \). The decision
variables, \( y^* \) and \( x^* \), for \( D_{\perp p}(b; p) \) are \( p \)-dimensional, and
periodic boundary conditions are imposed; the final state is put equal to
the initial state, \( b^{*(p)}(y^*, x^*) = b^{*(0)}(y^*, x^*) \).
The bottom \( l \) equality constraints for \( D_{\perp p}(b; p) \), that is the last \( l \) rows of \( N^T y^* - D^T x^* = 0 \), are \( b^{*(p)}(y^*, x^*) = b^{*(0)}(y^*, x^*) \). The objective functions, as well as the inequality
constraints, for \( D_{\perp p}(b; p) \) and \( D(b) \) are kept the same.

\[
D_{\perp p}(b; p) := \max \left\{ \begin{array}{c} y^* \in \mathbb{R}^p \times \mathbb{R}^p \end{array} \right\} \left\{ \begin{array}{c} B b, b^{*(0)}(y^*, x^*) \\ \text{s.t. } N^T y^* - D^T x^* = 0, \|y^*\|_{\infty} \leq 1, \|x^*\|_{\infty} \leq 1 \end{array} \right\}
\]
The infinite periodic extension of any \((y^*, x^*)\) feasible for \(\mathcal{D}_{\text{per}}(b; p)\) is feasible for \(\mathcal{D}(b)\). Hence the optimal value of the program \(\mathcal{D}_{\text{per}}(b; p)\) must be less than or equal to the optimal value for the program \(\mathcal{D}(b)\). Our goal is to find those initial conditions \(b\) and integers \(p\) for which these optimal values are the same, for then the infinite periodic extension of any \((y^*, x^*)\) is arg max \(\mathcal{D}_{\text{per}}(b; p)\) is optimal for \(\mathcal{D}(b)\).

3.3. The (pre-)dual of \(\mathcal{D}_{\text{per}}(b; p)\), denoted \(\mathcal{P}_{\text{per}}(b; p)\)

**Proposition 3.** The program \(\mathcal{P}_{\text{per}}(b; p)\) having the property that \(\mathcal{P}_{\text{per}}(b; p)\) and \(\mathcal{D}_{\text{per}}(b; p)\) are a dual pairing is

\[
\mathcal{P}_{\text{per}}(b; p): \left\{ \min_{y \in \mathbb{R}^p, x \in \mathbb{R}^p} \sum_{k=1}^{p} |y_k| + |x_k| \right. \\
\left. \text{s.t. } D_c(p)y + N_C(p)x = [b^T 0_{1 \times (p-1)}]^T \right\}
\]

The optimal values of \(\mathcal{P}_{\text{per}}(b; p)\) and \(\mathcal{D}_{\text{per}}(b; p)\) are equal. Furthermore, if \((y, x)\) and \((y^*, x^*)\) are feasible for \(\mathcal{P}_{\text{per}}(b; p)\) and \(\mathcal{P}_{\text{per}}(b; p)\), respectively, then a necessary and sufficient condition that they both be optimal solutions is that \((y, x)\) be aligned with \((y^*, x^*)\), that is Eqs. (6) hold for \(k = 1, \ldots, p\).

A proof is in the Appendix.

4. Basis vectors and induced dynamics

An important feature of sequences \((y, x) \in \arg \min \{\mathcal{P}(b)\}\) is the pattern of the location of those values of \(k\) for which \(y_k = 0\) and \(x_k = 0\). Related also is the fact that, although \(\mathcal{P}_{\text{per}}(b; p)\) is not a linear program in standard form, it can still be shown that for any optimal solution there is an optimal basic feasible solution. Investigation of these issues requires consideration of basis vectors.

4.1. Notation for a basis

Consider the set of equations \(Ax + By = b\), or in block matrix notation

\[
\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix} = b,
\]

where \(A\) and \(B\) are any real \(p \times p\) matrices, and \(y\) and \(x\) are real \(p\)-dimensional column vectors. Let \([A \ B]_B\) be any \(p \times p\) submatrix made up of the columns of the \(p \times 2p\) matrix \([A \ B]\). The \(p\) integers \(i_1, i_2, \ldots, i_p\) from 1, 2, \ldots, \(2p\) identify the columns of \([A \ B]\) that have been retained in \([A \ B]_B\). If \([A \ B]_B\) is non-singular, the set \(\{i_1, i_2, \ldots, i_p\}\) determines the basis \([A \ B]_B\) and the vector \(B = [i_1, i_2, \ldots, i_p]\) will be referred to as a \(p\)-dimensional basis vector. The basis vector complementary to \(B\), denoted \(B^*\), is the vector whose components are the \(p\) integers from 1, 2, \ldots, \(2p\) not in \(B\). If the \(p\) components of \((y, x)\) not associated with columns of \([A \ B]_B\) are set equal to zero, the solution to the resulting set of equations is said to be a basic solution to (9) with respect to the basis vector \(B\), denoted \((y^{\text{bol}}, x^{\text{bol}})_B\), a \(2p\)-dimensional vector. The components of \((y^{\text{bol}}, x^{\text{bol}})_B\) associated with the basis vector \(B\) are called basic variables and will be denoted by \((y^{\text{bvar}}, x^{\text{bvar}})_B\), a \(p\)-dimensional vector. Thus \((y^{\text{bvar}}, x^{\text{bvar}})_B = [0_{p \times (p-1)}]^T\), where \([A \ B]_B^{-1}\) denotes the inverse of the matrix \([A \ B]_B\). The basic solution is non-redundant if none of the components of \((y^{\text{bvar}}, x^{\text{bvar}})_B\) are zero. If a convex optimisation problem with constraints (9), denoted \(\mathcal{P}(b; p)\), has an optimal solution \((y^{\text{bol}}, x^{\text{bol}})_B\), then \(B = B_{\text{opt}}(\mathcal{P}(b; p))\) is termed a \(p\)-dimensional optimal basis vector for \(\mathcal{P}(b; p)\).

We are now ready to investigate how basis vector selection influences the dynamic evolution of the primal state \(b(k)y(x)\) for \((y, x)\) satisfying the constraints for \(\mathcal{P}(b)\). Put \(A = D_{(1:p,1:p)}, B = N_{(1:p,1:p)}\), fix \(b(0) \in \mathbb{R}^p\), and denote \((y^{\text{bol}}, x^{\text{bol}})_B\) as the basic solution to \(D_{(1:p,1:p)}y + N_{(1:p,1:p)}x = [b(0), 0]^T\) with respect to the \(p\)-dimensional basis vector \(B\). Consistent with (7), the initial state \(b(0)(y^{\text{bol}}, x^{\text{bol}})_B = b(0)\) for any \(B\). The terminal state \(b(p)(y^{\text{bol}}, x^{\text{bol}})_B\), defined in the second row of (7), will be denoted \(b(p)(b(0), B)\), as it is determined by the initial state, \(b(0)\), and the basis vector \(B\).

Define the map \(G(B): \mathbb{R}^l \rightarrow \mathbb{R}^l\) by \(G(B): b(0) \mapsto b(p)(b(0), B) = G(B)b(0)\). Thus \(G(B)\) is the map from the initial to the terminal state induced by \(B\). Proposition 4 below shows that \(G(B)\) is the North-West corner of \(I_p - YZ^{-1}\), where \(Z(B) := \left[D_{(1:p,1:p)} - N_{(1:p,1:p)}\right]_{B}\) and \(Y(B) := \left[D_{C}(p) - N_{C}(p)\right]_{B}\). From now on, if the basis vector is clear from the context, \(Y\) may be written in place of \(Y(B)\), and similarly for other matrices.

**Proposition 4.** Let \(H(B) := I_p - YZ^{-1}\). Then, for any \(b(0) \in \mathbb{R}^l\) and any \(p\)-dimensional basis vector \(B\), \(G(B) := H(1_{l,l})\) satisfies \(b(p)(b(0), B) = G(B)b(0)\).

**Proof.** Put \(F(p) := \left[D_{(1:p,1:p)} - N_{(1:p,1:p)}\right]_{B}\). Then from (7)

\[
b(p)(b(0), B) = F(p) \begin{bmatrix} y^{\text{bol}} \\ x^{\text{bol}} \end{bmatrix}_B = \left\{ [F(p)]_{B1_{l,1:2p}} \begin{bmatrix} y^{\text{bvar}} \\ x^{\text{bvar}} \end{bmatrix}_B \right. \\
\left. \text{s.t. } \left[I_p - D_{C}(p) - N_{C}(p)\right]_{B} \left[D_{(1:p,1:p)} - N_{(1:p,1:p)}\right]_{B}\right)^{-1} [b(0)]_{0_{(p-1) \times 1}} \times 0_{(p-1) \times 1} \right\}_{(l,l)}
\]

\[
= \left[I_p - D_{C}(p) - N_{C}(p)\right]_{B} \left[D_{(1:p,1:p)} - N_{(1:p,1:p)}\right]_{B}\right)^{-1} [b(0)]_{0_{(p-1) \times 1}} \times 0_{(p-1) \times 1} \right\}_{(l,l)}
\]

\[
= [H(B)]_{B1_{l,1:2p}} b(0) = G(B)b(0).
\]

\(\square\)
The $l \times l$ matrix $G(B)$ is therefore the state transition matrix induced by the $p$-dimensional basis vector $B$ which takes the state at time $k = 0$ to the state at time $k = p$.

4.2. Invariant bases

Fix $p \geq 2l$, and suppose that $B$ is a $p$-dimensional basis vector for which $G(B)$ has an eigenvector, $v$, with non-negative corresponding eigenvalue $\lambda$. Recall that $[y_{bsol}^{(b)}]_{B} = (y_{bsol}^{(b0)}; y_{bsol}^{(b0)})$ is the basic solution to $D_{1(p,1)} y_{B} + N_{1(p,1)} x_{B} = \begin{bmatrix} 0 \end{bmatrix}$ (sometimes denoted $(b^{(0)}, 0)$) with respect to $B$. Put $b^{(0)} = v$, so $b^{(p)}(v, B) = \lambda v$ and $(y_{bsol}, x_{bsol})_{B} = (\lambda y_{bsol}, x_{bsol})$. Continuing in this manner, for all $n \in \mathbb{N}$, $\lambda^{n} (y_{bsol}, x_{bsol})_{B}^{(n)}$ is the basic solution to $D_{1(p,1)} y_{B} + N_{1(p,1)} x_{B} = (\lambda^{n} v, 0)$. The terminal condition for the constraints $D_{1(p,1)} y_{B} + N_{1(p,1)} x_{B} = (\lambda^{\infty} v, 0)$ matches the initial condition for the constraints $D_{1(p,1)} y_{B} + N_{1(p,1)} x_{B} = (\lambda^{0} + 1 v, 0)$, so the concatenation of all $(y_{bsol}, x_{bsol})_{B}^{(n)}$ for $n \in \mathbb{N}$ is feasible for $\mathcal{P}(v)$. In order that this concatenation be optimal for $\mathcal{P}(v)$ it is clearly necessary that $\lambda$ be further restricted to the interval $[0, 1)$, but this is not enough to ensure optimality. Theorem 6 below says that, for bases satisfying the following definition, this concatenated sequence of basic solutions is not only feasible, but also optimal for $\mathcal{P}(v)$.

Definition 1. A $p$-dimensional basis vector $B$ for $[D_{1(p,1)} N_{1(p,1)}]_{B}$ is said to be invariant for $\mathcal{P}(\cdot)$ if there is an eigenvector of $G(B)$, denoted $v^{(G(B))}$, having corresponding eigenvalue $\lambda \in [0, 1)$, with the property that $B$ is an optimal basis vector for $P_{p} v[G(B); P]$. The eigenvector $v[G(B)]$ is termed a dual-periodic initial state. A dual-periodic initial state $v[G(B)]$ is termed attracting if $|\lambda| > |\lambda_{i}|$ for all eigenvalues $\lambda_{i}$ of $G(B)$ other than $\lambda$.

Thus an invariant basis satisfies $B = B_{opt}(P_{p} v[G(B)]; P)$. Also, for $v[G(B)]$, the corresponding eigenvalue satisfies $\lambda = \rho(G(B))$, the spectral radius of $G(B)$.

To apply Definition 1 it is useful to know that optimal bases for $P_{p} v[G(B); P]$ can be characterised using the alignment condition relating optimal solutions for $P_{p} v[G(B); P]$ and $P_{p} v[G(B); P]$. For an arbitrary basis vector $B$ put $\bar{s} := \text{sgn} \left( \begin{bmatrix} D_{c} & N_{c}(p) \end{bmatrix}_{B}^{1 - 1} \right)_{s}$. Then $\bar{s}$ is a $p$-dimensional vector whose components are $-1$, $0$, or $1$. Define

$$S_{set} := \{ s = (s_{1}, \ldots, s_{p}) \in \mathbb{R}^{p} : s_{k} = \bar{s}_{k} \text{ if } \bar{s}_{k} = 1 \text{ or } -1 \text{ if } \bar{s}_{k} = 0 \}.$$ 

Clearly $\bar{s} \in S_{set} \subseteq \mathbb{R}^{p}$; also $S_{set} = [\bar{s}]$ if and only if the basic solution with respect to $B$ is non-redundant.

Proposition 5. A basis vector $B$ for $P_{p} v[G(B); P]$ is an optimal basis vector if and only if $\{|N_{c}^{x}(p)| - D_{c}^{x}(p)B_{c}|_{s} \leq 1 \text{ for some } s \in S_{set} \}$.

Proof. See Appendix. □

A candidate basis for $[D_{1(p,1)} N_{1(p,1)}]_{B}$ is that is a selection of $p$ integers from $1, 2, \ldots, 2p$ for which $[D_{1(p,1)} N_{1(p,1)}]_{B}$ is non-singular, can then be easily checked for the property of being invariant with the aid of Proposition 5.

5. Primal optimal solutions and dual-periodic initial states

Given an invariant basis vector, optimising sequences for $\mathcal{P}(\cdot)$ can be constructed if the initial condition is a dual-periodic initial state associated with the invariant basis.

Theorem 6. Suppose the $p$-dimensional basis vector $B$ is invariant for $\mathcal{P}(\cdot)$, so the dual-periodic initial state $v[G(B)]$ has corresponding eigenvalue $\lambda \in [0, 1)$. Then the concatenation of basic solutions $\left((y_{bsol}, x_{bsol})_{B}^{(n)} v[G(B)])\right)_{n=0}^{\infty}$ is optimal for $\mathcal{P}(v[G(B)])$. Furthermore, for all $(y^{*}, x^{*}) \in \text{arg max } D_{p} v[G(B)]$; $\rho$, $(x^{*}, y^{*}) \in \text{arg max } D_{p} v[G(B)]$, where $(x^{*}, y^{*})$ is the infinite periodic extension of $(y^{*}, x^{*})$.

Proof. It was shown in Section 4.2 that if $B$ is invariant then the concatenation of the sequence $\left((y_{bsol}, x_{bsol})_{B}^{(n)} v[G(B)]\right)_{n=0}^{\infty} = (y, x)$ is feasible for $\mathcal{P}(v[G(B)])$. Also $(y, x) \in l_{1} \times l_{1}$ since $|\lambda| < 1$. We claim that $(y, x)$ is aligned with $(y^{*}, x^{*})$. To see this, denote by $b_{bsol}^{(b)}$, respectively $b_{bsol}^{(p)}$, the basic solution, respectively basic variables, for $D_{c}(p) y_{B} + N_{c}(p) x_{B} = (v[G(B)], 0)$ with respect to $B$. Recalling the definitions of $Y$ and $Z$ in Section 4.1, we have, for any $b \in \mathbb{R}^{l}$,

$$\begin{bmatrix} y^{(b)}_{bsol} \end{bmatrix}_{B} = Y^{-1}(B) \begin{bmatrix} b \end{bmatrix}_{0},$$

$$\begin{bmatrix} y^{(b)}_{bsol} \end{bmatrix}_{B} = Z^{-1}(B) \begin{bmatrix} b \end{bmatrix}_{0}.$$ 

Then $Z^{-1} Y : \begin{bmatrix} y^{(b)}_{bsol} \end{bmatrix}_{B} \mapsto \begin{bmatrix} y^{(p)}_{bsol} \end{bmatrix}_{B}$. It is straightforward to show that $\begin{bmatrix} y^{(p)}_{bsol} \end{bmatrix}_{B}^{(s)}$ is an eigenvector of $Z^{-1} Y$ with corresponding eigenvalue $1 - \lambda$, and $\begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B} = (1 - \lambda) \begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B}^{(l)}$. Since $\lambda \in [0, 1)$ it follows that

$$\text{sgn} \begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B}^{(l)} = \text{sgn} \begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B}^{(l)}.$$ 

By Proposition 3 applied to $P_{p} v[G(B)]$ and $P_{p} v[G(B); P]$, $\begin{bmatrix} y^{(b)}_{bsol} \end{bmatrix}_{B}^{(n)}$ is aligned with arg max $D_{p} v[G(B); P]$. By (10) $\begin{bmatrix} y^{(b)}_{bsol} \end{bmatrix}_{B}^{(s)}$ is aligned with arg max $D_{p} v[G(B)]$. Since, for $n \in \mathbb{N}$ and $\lambda \in (0, 1)$, $\text{sgn} \begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B} = \text{sgn} \begin{bmatrix} y^{(s)}_{bsol} \end{bmatrix}_{B}^{(l)}$, we conclude that $(y, x)$ is aligned with $(y^{*}, x^{*})$. Clearly also if $\lambda = 0$.
then $(y, x)$ is aligned with $(y^{\text{ext}}_i, x^{\text{ext}}_i)$. Since $(y, x)$ is feasible for $\mathcal{P}(v(G(B)))$ and $(y^{\text{ext}}_i, x^{\text{ext}}_i)$ is feasible for $\mathcal{D}(v(G(B)))$, by Proposition 1, $(y, x) \in \arg \max \mathcal{P}(v(G(B)))$ and $(y^{\text{ext}}_i, x^{\text{ext}}_i) \in \arg \max \mathcal{D}(v(G(B)))$. \hfill $\Box$

Theorem 6 solves all problems $\mathcal{P}(v)$ whenever $v$ is a dual-periodic initial state. As an immediate corollary, dual-periodic initial states for which optimal signals are of finite length are easily identified.

**Corollary 7.** Suppose $B$ is a $p$-dimensional invariant basis vector for $\mathcal{P}(\cdot)$. If the dual-periodic initial state $v(G(B))$ has corresponding eigenvalue $\lambda = 0$, then $\mathcal{P}(v(G(B)))$ has an optimal solution of finite length given by $y = (y^{\text{sol}}, 0, \ldots)$ and $x = (x^{\text{sol}}, 0, \ldots)$, where $[D(1:p, 1:p) N(1:p, 1:p)]G^{\text{var}}_{b^{\text{var}}} = (v(G(B)), 0)$.

In the non-redundant case the non-zero components of $(x^{\text{sol}}, y^{\text{sol}})$, a vector with $2p$ components, are precisely the $p$ components of $(x^{\text{var}}, y^{\text{var}})$.

Although there are as yet no results of a general nature on the number of dual-periodic initial states for a given problem $\mathcal{P}(\cdot)$, an algorithm based on Definition 1 can in principle check all finite-dimensional candidate bases. Our computer code performs an exhaustive search in reasonable time for periods up to about $p = 16$. For the Example presented later, with $\hat{h}(z)$ and $d(z)$ cubic polynomials, $\mathcal{P}(\cdot)$ has about 100 dual-periodic initial states associated with invariant bases of period 16 or less. For the problems tested so far, the number of attracting dual-periodic initial states found is always either one or zero. Despite their relatively infrequent occurrence, the presence of an attracting dual-periodic initial state, $b_{\text{att}}$, is often apparent even in floating-point simulations. This is because there is a neighbourhood of $b_{\text{att}}$ for which the optimal solution to $\mathcal{D}(b)$, for any $b$ in the neighbourhood, is either periodic or ultimately periodic, with period the same as the optimal solution to $\mathcal{D}(b_{\text{att}})$.

### 5.1. Attracting dual-periodic initial states

Theorem 6 gives the optimal solution for a countable number of initial states, and is applicable whether or not initial states are attracting. For the case of an attracting dual-periodic initial state, $b_{\text{att}}$, there are many more initial states for which the solution to $\mathcal{P}(\cdot)$ can be constructed explicitly. In fact there is a cone, $C_{b_{\text{att}}}$, defined below, containing $b_{\text{att}}$ with the property that, for all initial states $b \in C_{b_{\text{att}}}$, $\arg \max \mathcal{D}(b) = \arg \max \mathcal{D}(b_{\text{att}})$. For such initial $b$, the time evolution of the optimal states for $\mathcal{P}(\cdot)$ will be forward asymptotic to the direction of $b_{\text{att}}$.

Suppose $b_{\text{att}}[G(B)]$ is an attracting dual-periodic initial state. Then, for $b \in \mathbb{R}^l$, $\lim_{n \to \infty} G^n b = b_{\text{att}}[G(B)]$ because $b_{\text{att}}[G(B)]$ is the dominant eigenvector of $G$. Assume that $\lambda \in (0, 1)$. We define a cone containing $b_{\text{att}}[G(B)]$, denoted $C_{b_{\text{att}}}$, having the property that, for all $b \in C_{b_{\text{att}}}$ and for all $n \in \mathbb{N}$, the sign patterns of $(y^{\text{sol}}, x^{\text{sol}}(G^n b))$ are the same. For $Z := [D(1:p, 1:p) N(1:p, 1:p)]$, put $\tilde{Z} := Z(1:p, 1:l)$. Define $s := \text{sgn}(\tilde{Z} b_{\text{att}}[G(B)])$.

**Definition 2.**

\[ C_{b_{\text{att}}} := \left\{ b \in \mathbb{R}^l : \begin{array}{l} \forall n \in \mathbb{N}, \text{sgn}(\tilde{Z} G^n b) = s, \text{ or } \forall n \in \mathbb{N}, \text{sgn}(\tilde{Z} G^n b) = -s \end{array} \right\}. \tag{11} \]

Clearly $b_{\text{att}}[G(B)] \in C_{b_{\text{att}}}$, since $b_{\text{att}}[G(B)]$ is an eigenvector of $G$, with corresponding eigenvalue $\lambda \in (0, 1)$. Then using the Cayley–Hamilton theorem it can be shown that, for all $b \in C_{b_{\text{att}}}$, there is a solution $(y, x) \in \arg \min \mathcal{P}(b)$ satisfying a linear recurrence relation of order at most $lp$. If $G$ has a complete set of linearly independent real eigenvectors the optimal solution can be derived in a simple explicit form, so we present this special case as our final result.

Suppose $b_{\text{att}}[G(B)]$ is an attracting, dual-periodic, initial state associated with the invariant basis $B$. If $G$ has a complete set of linearly independent real eigenvectors then $b \in C_{b_{\text{att}}}$ can be represented as $b = \sum_{i=1}^l V_i$, where the $V_i$ are suitably normalised eigenvectors of $G$, each with associated real eigenvalue $\lambda_i$. The basic variables $y^{\text{var}}_{b^{\text{var}}} (V_i) B = [D(1:p, 1:p) N(1:p, 1:p)]^{-1} y^{\text{var}}_{b^{\text{var}}} (V_i)$ can be decomposed into $\sum_i y^{\text{var}}_{b^{\text{var}}} (V_i) B = \sum_i y^{\text{var}}_{b^{\text{var}}} (V_i) B$. Filling in the zero-valued non-basic variables, we have $y^{\text{sol}} = \sum_i y^{\text{sol}} (V_i) B$. Put $y^{(V_i)} := y^{\text{var}}_{b^{\text{var}}} (V_i) / y^{\text{sol}}$. Then $y^{(V_i)}$ and $y^{(V_i)}$ are both $p$-dimensional vectors.

**Proposition 8.** Suppose $B$ is invariant for $\mathcal{P}(\cdot)$ and $b_{\text{att}}[G(B)]$ is an attracting dual-periodic initial state. If $V_i$ are linearly independent real eigenvectors of $G(B)$ with $\lambda_i$ the associated eigenvalues, then for $(y, x) \in I_1 \times l_1$ satisfying

\[ \hat{y}(z) = \sum_{i=1}^l \sum_{k=1}^p y_i^{(V_i)} z^{k-1} \quad \hat{x}(z) = \sum_{i=1}^l \sum_{k=1}^p y_i^{(V_i)} z^{k-1} (1 - \lambda_i z)^{-p}, \tag{12} \]

we have $(y, x) \in \arg \min \mathcal{P}(\sum_{i=1}^l V_i)$ whenever $\sum_{i=1}^l V_i \in C_{b_{\text{att}}}$. 

**Proof.** If $b \in C_{b_{\text{att}}}$ then the concatenation of the sequence $(y^{\text{sol}}, x^{\text{sol}}(G^n b))_{n=0}^{\infty}$ has the same sign pattern as the concatenation of the sequence $(y^{\text{sol}}, x^{\text{sol}}(G^n b))_{n=0}^{\infty}$. It then follows using the same argument as in the proof of Theorem 6 that the concatenation of the sequence $(y^{\text{sol}}, x^{\text{sol}}(G^n b))_{n=0}^{\infty}$ is feasible and optimal for $\mathcal{P}(b)$. Since $G^n b = \sum_{i=1}^l \lambda_i^n v_i$, we have

\[ y^{\text{sol}}(G^n b) = \sum_i \lambda_i^n y^{\text{sol}} (V_i) B = \sum_i \lambda_i^n y^{(V_i)} B \]

The $Z$-transform of the $p$-dimensional column vector $y^{(V_i)}$ is $\sum_{k=1}^p y_i^{(V_i)} z^{k-1}$. The $Z$-transform of the concatenation of the sequence $(\lambda_i^n y^{(V_i)})_{n=0}^{\infty}$, that is of the sequence
Proposition 3 gives an exact solution to the problem for the initial condition $G = \left(\cdots\right)$, where $\lambda_i, \lambda_j, \lambda_k, \ldots$ are the eigenvalues of $G$. Hence the result follows.

5.2. Example

We illustrate the results in this paper for the problem $P(\cdot)$ having the following given data:

\[
\begin{align*}
\hat{\lambda} &= (3/2)(1 + z/2)(1 + 2z/9)(1 - z/5) \\
\hat{n} &= (1 - z/3)(1 - 2z/7)(1 - 2z/5) \\
&= 1 - 107/105z + 12/35z^2 - 4/105z^3.
\end{align*}
\]

For values of $p$ up to about 15 or so it is possible to determine all of the invariant bases by an exhaustive search. There are, for example, 12 non-redundant invariant bases with fundamental period either 6 or a divisor of 6. Three of these, along with the associated eigenvectors which are dual-periodic initial states, are:

<table>
<thead>
<tr>
<th>$B$</th>
<th>$v(G(B))$</th>
<th>Fundamental period</th>
</tr>
</thead>
<tbody>
<tr>
<td>${1, 2, 3, 4, 5, 6}$</td>
<td>$(9, 13/2, 1)$</td>
<td>1</td>
</tr>
<tr>
<td>${1, 2, 4, 5, 8, 11}$</td>
<td>$(1.32, 1, 0)$</td>
<td>3</td>
</tr>
<tr>
<td>${1, 2, 4, 9, 11, 12}$</td>
<td>$(1.92, 1, -0.09)$</td>
<td>6</td>
</tr>
</tbody>
</table>

Numbers expressed as a fraction are exact. Numbers with a decimal point are approximations to irrationals. There are also 2 bases of fundamental period 5, 11 of period 7, 11 of period 8, 13 of period 9, 20 of period 10, 15 of period 11, etc. It is not known how many invariant bases there are in total, or even if their number is finite. For all of these bases, Theorem 6 gives an exact solution to the problem $P(v(G(B)))$. All except one of these dual-periodic initial states associated with invariant bases are repelling — in only one case is $v(G(B))$ a dominant eigenvector of $G(B)$. In fact, an extensive computer assisted search has uncovered only one attracting dual-periodic initial state, $v_{att} = \left[\frac{28655197}{4606468}, \frac{203277932802894}{4606468}, -1, 0\right]^T$, the dominant eigenvector of $G(B_{att})$ associated with the invariant basis vector $B_{att} = [1, 2, 4, 6, 8, 10, 12]$. There is a second dual-periodic initial state associated with $B_{att}$, namely $[\frac{28655197}{4606468}, -\frac{203277932802894}{4606468}, -1, 0]^T$, which is not attracting.

The characteristic polynomial for $G(B_{att})$ is $P(z) = z^3 - (7497163/294989611985)z^2 + (896/560480267715)z$. We illustrate (12) for the initial condition $b = [1, -1/10, 1, 1/10]^T$, which satisfies $b \in C_{v_{att}}$. By (12) the optimal solution to $P(b)$ is $\hat{\lambda}(z) = \sum_{i=1}^{3} \lambda_i \sum_{k=1}^{i} \frac{\frac{y(v_i)}{\lambda_i} \lambda_i^{-\frac{k-1}{3}}}{\lambda_i^{-\frac{k-1}{3}}}$, where $\lambda_i$ are the zeros of $P(z)$, the components of $[D(1, 1, 7); N(1, 1, 7)]_{B_{att}}^{-1} [V_{T}, 0]^T$ are the non-zero (basic) components of $y(v_i)$ and $x(v_i)$, so $[D(1, 1, 7); N(1, 1, 7)]_{B_{att}}^{-1} [V_{T}, 0]^T = [V_{T}, 0]^T$, and $V_{T}$ are the right eigenvectors of $G(B_{att})$, normalised so that $b = \sum_{i=1}^{3} V_i$. The optimal value for $P(b)$ is 49122365681221/56046601825140 $\approx$ 0.87646. This number is obtained either by summing the one norms of the inverse $Z$-transforms of $\hat{\lambda}(z)$ and $\hat{\lambda}(z)$, or by evaluating the dual cost $b(\cdot) = \hat{\lambda}(T, 7) = \hat{\lambda}(T, 7)$.

6. Conclusions

We have investigated the time evolution of optimal solutions to an $l_1$-minimisation problem with convolution constraints. A systematic procedure finds invariant bases $B$, and associated primal initial states $v$, for which the optimal dual solution is periodic. If $v$ is a dominant eigenvector of the mapping describing the time evolution of the primal state, then there is a cone containing $v$ having the property that initiating the system at any point in the cone yields an orbit that is forward asymptotic to the direction of $v$.

Acknowledgement

I thank Rodney Topor for help with the algorithm for finding invariant bases.

Appendix A. Proof of Proposition 3

Put $W := \begin{bmatrix} D_C(p) & -D_C(p) & -N_C(p) & -N_C(p) \\ -D_C(p) & D_C(p) & -N_C(p) & -N_C(p) \end{bmatrix}$. The program $P_{\per}(b; p)$ is equivalent to the following program

\[
\begin{align*}
\min & \sum_{k=1}^{p} (x_k^+ + x_k^- + x_k^+ + y_k^-) \\
\text{s. t.} & W \left[\begin{array}{c} y_k^- \\ y_k^- \\ x_k^+ \\ x_k^+ \end{array}\right]^T \\
& \geq \begin{bmatrix} b^T & 0_{1 \times (p-l)} & -b^T & 0_{1 \times (p-l)} \end{bmatrix}^T y_k^+ + y_k^- + x_k^+ + x_k^+ \geq 0.
\end{align*}
\]

The standard linear programming dual in symmetric form ([8], p. 85) is

\[
\begin{align*}
\max_{\lambda} & \lambda^T W \left[\begin{array}{c} b^T & 0_{1 \times (p-l)} & -b^T & 0_{1 \times (p-l)} \end{array}\right]^T \\
\text{s. t.} & \lambda^T W \leq 1 \times 4p \\
& \lambda \geq 0.
\end{align*}
\]

Partitioning $\lambda$ as $(\mu, -\nu)$, this becomes, successively,

\[
\begin{align*}
\max_{\mu \in \mathbb{R}_p, \nu \in \mathbb{R}_p} & \left[\begin{array}{c} \mu^T \\ -\nu^T \end{array}\right]^T \\
\text{s. t.} & \lambda^T W \leq 1 \times 4p \\
& \lambda \geq 0.
\end{align*}
\]
\[
\begin{align*}
\lambda\left[b^T : 0_{1\times(p-l)} : -b^T : 0_{1\times(p-l)}\right]^T \\
\text{s. t. } [u^T : -v^T]W \leq 1_{4\times p}, u \geq 0, v \leq 0; \\
\max_{a \in \mathbb{R}^p, s.t. v \in \mathbb{R}^p} \lambda\left[u^T : -v^T\right]^T \\
\times \left[b^T : 0_{1\times(p-l)} : -b^T : 0_{1\times(p-l)}\right]^T \\
\text{s. t. } \left\{u^T + v^T\right\}D_{C}(p) \leq 1_{1 \times p}, \\
\left\{u^T + v^T\right\}N_{C}(p) \leq 1_{1 \times p}, u \geq 0, v \leq 0; \\
\max_{a \in \mathbb{R}^p} \lambda\left[a^Tb^T : 0_{1\times(p-l)}\right]^T \\
\text{s. t. } N^T_{C}(p)y - D^T_{C}(p)x = 0, \|a^*\|_\infty \leq 1, \|u^*\|_\infty \leq 1.
\end{align*}
\]

The objective function \(\alpha^T\left[b^T : 0_{1\times(p-l)}\right]^T\) will now be expressed in terms of the primal and dual initial states. For any \((\bar{y}, \bar{x})\) satisfying \(D_{C}(p)\bar{y} + N_{C}(p)\bar{x} = \left[b^T, 0\right]^T, \alpha^T\left[b^T, 0\right]^T = \alpha^TD_{C}(p)\bar{y} + \alpha^TN_{C}(p)\bar{x} = (\bar{y}, D^T_{C}(p)\alpha) + (\bar{x}, N^T_{C}(p)\alpha) = (\bar{y}, y^* + \bar{x}, x^*)\). Then minor modifications to the argument used in the proof of Proposition 2 show that \((\bar{y}, y^*) + (\bar{x}, x^*) = b^*(0)T(y^*, x^*)Bb\), as required. It is routine to verify that the alignment condition of the Proposition is a consequence of the complementary slackness conditions ([8], p. 96) applied to (A.1) and (A.2).

**Appendix B. Proof of Proposition 5**

Given a basis vector \(B\) and its complement \(B^*\), the constraints for \(P_{per}(b; p), \left[N^T_{C}(p) : -D^T_{C}(p)\right]^{(i)} = 0\), can be partitioned as

\[
\begin{align*}
\left[N^T_{C}(p) : -D^T_{C}(p)\right]_B^{(i)}y^*_B & \quad + \left[N^T_{C}(p) : -D^T_{C}(p)\right]_{B^*}^{(i)}x^*_B^{(i)} = 0. \\
\text{Then} \\
\left[y^*_B, x^*_B\right]^{(i)} = \left[N^T_{C}(p) : -D^T_{C}(p)\right]^{-1}B^{(i)} \\
\times \left[N^T_{C}(p) : -D^T_{C}(p)\right]_B y^*_B. \\
\end{align*}
\]

Put \(\left[y^*_B, x^*_B\right] = \text{sgn}\left([D_{C}(p) \quad N_{C}(p)]^{-1}_{B} \left[s^*_B\right]\right) = s\), in which case

\[
\left[y^*_B, x^*_B\right] = \left[N^T_{C}(p) : -D^T_{C}(p)\right]^{-1}_{B^*} \\
\times \left[N^T_{C}(p) : -D^T_{C}(p)\right]_B s = \left[y^*_B, x^*_B\right]. \\
\]

Now combining \(s\) and \(\left[y^*_B, x^*_B\right]\), consistent with the partitioning of \(\left[y^*_B, x^*_B\right]\) into \(\left[y^*_B, x^*_B\right]\) and \(\left[y^*_B, x^*_B\right]\), gives a \(2p \times 1\) vector, denoted \(\left[y^*_B, x^*_B\right]\). By construction \(\left[y^*_B, x^*_B\right]\) satisfies \(N^T_{C}(p)y^* = D^T_{C}(p)x^*\); also \(\left[y^*_B, x^*_B\right]\) and \(\left[y^*_B, x^*_B\right]\) are aligned. If in addition the inequality in the Proposition statement holds then \(\left[y^*_B, x^*_B\right]\) is feasible for \(P_{per}(b; p)\). Then by Proposition 3 \(\left[y^*_B, x^*_B\right] \in \arg \min_{y^*, x^*} P_{per}(b; p)\), that is \(B\) is an optimal basis vector for \(P_{per}(b; p)\). This proves the if part. Also from (B.3) it follows that if \(B\) is optimal for \(P_{per}(b; p)\), that is \(\left[y^*_B, x^*_B\right] \in \arg \min_{y^*, x^*} P_{per}(b; p)\), then

\[
\left[N^T_{C}(p) : -D^T_{C}(p)\right]^{-1}_{B^*} \left[N^T_{C}(p) : -D^T_{C}(p)\right]_B s = 1 \\
\text{for some } s \in S_{set}. \\
\]

This concludes the proof.

**References**


