A Fast Method for Finding the Global Solution of the Regularized Structured Total Least Squares Problem for Image Deblurring

Amir Beck[*] Aharon Ben-Tal[†] and Christian Kanzow[‡]

November 15, 2007

Abstract

Given a linear system $\mathbf{A}\mathbf{x} \approx \mathbf{b}$ over the real or complex field where both $\mathbf{A}$ and $\mathbf{b}$ are subject to noise, the total least squares (TLS) problem seeks to find a correction matrix and a correction righthand side vector of minimal norm which makes the linear system feasible. To avoid ill-posedness, a regularization term is added to the objective function; this leads to the so-called regularized TLS (RTLS) problem. A further complication arises when the matrix $\mathbf{A}$ and correspondingly the correction matrix must have a specific structure. This is modelled by the regularized structured TLS (RSTLS) problem. In general this problem is nonconvex and hence difficult to solve. However, the RSTLS problem arising from image deblurring applications under reflexive or periodic boundary conditions possess a special structure where all relevant matrices are simultaneously diagonalizable (SD). In this paper we introduce an algorithm for finding the *global* optimum of the RSTLS problem with this SD structure. The devised method is based on decomposing the problem into single variable problems and then transforming them into one-dimensional unimodal real-valued minimization problems which can be solved globally. Based on uniqueness and attainment properties of the RSTLS solution we show that a constrained version of the problem possess a strong duality result and can thus be solved via a sequence of RSTLS problems.

1 Introduction

Given a linear system $\mathbf{A}\mathbf{x} \approx \mathbf{b}$ over the real or complex field where both the matrix $\mathbf{A}$ and the righthand side vector $\mathbf{b}$ are subjected to noise, the total least squares (TLS) problem seeks to

[*]Department of Industrial Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel. E-mail: becka@ie.technion.ac.il.

[†]MINERVA Optimization Center, Department of Industrial Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel. E-mail: abental@ie.technion.ac.il. Partly supported by the Technion VPR fund for promotion of Research, grant #2005519

[‡]University of Würzburg, Institute of Mathematics, Am Hubland, 97074 Würzburg, Germany. E-mail: kanzow@mathematik.uni-wuerzburg.de

minimize the sum of square norms of the perturbations to both the model matrix and vector $\|\mathbf{E}\|^2 + \|\mathbf{w}\|^2$ subject to the condition that the perturbed system holds: $(\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w}$. Although this problem is nonconvex it can be solved efficiently and globally using a spectral decomposition of the augmented matrix $(\mathbf{A}, \mathbf{b})$, see [12, 18].

In many applications, the matrix $\mathbf{A}$ has a specific linear structure, *e.g.,* Toeplitz or Hankel, which imposes a requirement on the perturbation matrix $\mathbf{E}$ to possess a corresponding special structure. The TLS solution does not take into account this requirement, and consequently the structured TLS (STLS)[1] became of intensive research, see e.g., [1, 25, 30, 24, 21, 20]. The general formulation of the STLS problem is

$$\text{(STLS):} \quad \begin{array}{ll} \min_{\mathbf{E},\mathbf{x},\mathbf{w}} & \|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 \\ \text{s.t.} & (\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w}, \\ & \mathbf{E} \in \mathcal{L}, \end{array}$$

where $\mathcal{L}$ is a linear subspace.

The STLS problem is a nonconvex problem and thus finding its global solution is in general a difficult task. There are only a few exceptions to this state of affairs. For block circulant structures with unstructured blocks the corresponding STLS problem can be solved by decomposing the problem into several smaller TLS problems using the discrete Fourier transform [3]. Another tractable case arises when some of the columns of $\mathbf{A}$ are error free while the other are subjected to noise. This problem is called the *generalized TLS* (GTLS) problem or *mixed LS-TLS* problem and its solution can be obtained by computing a QR factorization of $\mathbf{A}$ and then solving a TLS problem of reduced dimension [17]. A more general problem is the *restricted TLS* problem introduced in [19]. Here it is assumed that $(\mathbf{E}, \mathbf{w}) = \mathbf{D}_1 \tilde{\mathbf{E}} \mathbf{C}_1$, where $\mathbf{D}_1$ and $\mathbf{C}_1$ are known matrices and $\tilde{\mathbf{E}}$ is unknown. As was shown in [19], by choosing the matrices $\mathbf{D}_1$ and $\mathbf{C}_1$ appropriately, the restricted TLS problem contains as special cases any weighted least squares (LS), generalized LS, TLS, and generalized TLS problems. The restricted TLS problem can be solved using the restricted singular value decomposition [32].

In this paper we consider yet another tractable class of STLS problems in which the global solution can efficiently be found. We deal with structures in which all the matrices in $\mathcal{L}$ can be diagonalized by a certain fixed orthogonal (or unitary in the complex case) matrix. These structures are called *simultaneously diagonalizable (SD)* structures. The motivation for considering such structures stems from image deblurring problems with spatially invariant point spread functions (PSF). For two-dimensional image deblurring problems it is well known that the model matrix can be diagonalized by a two-dimensional discrete Fourier transform matrix when *periodic boundary conditions* are assumed. For *reflexive boundary conditions* with symmetric PSF the model matrix can be diagonalized by a two-dimensional discrete cosine transform matrix. Similar structures can be found in one-dimensional deconvolution problems. Section 2 contains a brief review of these structures.

A characteristic feature of image deblurring problems is that the matrix $\mathbf{A}$ is ill-conditioned and as a result the STLS solution has usually a huge norm and as such is meaningless. Regularization is required in order to stabilize the solution. For the unstructured TLS problem

---

[1]In some papers the STLS problem is also called *constrained total least squares* (CTLS).

several regularization methods are well known. Among them are truncation methods [9, 14], Tikhonov regularization [11, 4] in which a quadratic penalty is added to the objective function, or a quadratic constraint bounding the size of the solution norm is added to the problem [31, 29, 11, 6, 5].

For the STLS problem, Tikhonov regularization, seems to be the most popular method. The resulting problem is called the regularized STLS problem (RSTLS) and is given by

$$\text{(RSTLS):} \quad \begin{array}{ll} \min_{\mathbf{E},\mathbf{x},\mathbf{w}} & \|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 + \rho\|\mathbf{Lx}\|^2 \\ \text{s.t.} & (\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w}, \\ & \mathbf{E} \in \mathcal{L}. \end{array}$$

Common choices for $\mathbf{L}$ are the identity or a matrix approximating the first or second order derivative operator [13, 11, 15].

The RSTLS problem for structures arising in image deblurring was studied in several works. In [23] periodic boundary conditions are considered. Using the discrete Fourier transform the problem is decomposed into many complex-valued single-variable problems. The complex univariate problems are solved as two-variable nonconvex problems over the real domain using Davidon-Fletcher-Powell optimization algorithm.

In [26] an iterative algorithm of quasi-Newton form is applied for the RSTLS problem for reflexive boundary conditions that exploits the diagonalization properties of the associated matrices. The work [28] extends the structured total least norm (STLN) algorithm [30] to include regularization and image deblurring examples are discussed. This approach was also advocated in [10] for image deblurring problems with separable PSFs and in [22] for problems with zero boundary conditions.

In all the above mentioned works the optimization problems needed to be solved are nonconvex and consequently the devised algorithms are not guaranteed to converge to a global optimum but rather to a stationary point. The main contribution of the present paper is the introduction of a method capable of obtaining the *global minimum* of the RSTLS problem for SD structures.

The paper is organized as follows. In Section 2 we present a precise problem formulation followed by a brief review of the essential ingredients from image deblurring. The decomposition of the RSTLS problem into single-variable real or complex-valued problems is discussed in Section 3. These univariate problems are not necessarily unimodal but we show in Section 4 that they can be transformed into single-variable real-valued unimodal problems. Attainment and uniqueness conditions are also obtained. In Section 5 we concentrate on circulant structures and show that when the data is real-valued then there exists at least one real-valued optimal solution (although the corresponding single variable problems are complex-valued). In Section 6 we tackle the constrained version of the RSTLS problem, called CSTLS, and show that based on the derived uniqueness properties and on a strong duality result, the constrained problem can be solved by a sequence of RSTLS problems. The paper ends in Section 7 with detailed descriptions of the numerical algorithms and a demonstration of our method as applied to an image deblurring problem. A MATLAB implementation and documentation of the RSTLS and CSTLS methods for image deblurring

problems with either periodic or reflexive boundary conditions can be found in

1.1   Notation

A vector or matrix is called real-valued (complex-valued) if all its entries are real (complex). For a complex scalar a, the complex conjugate is denoted by $\bar{a}$. Given a matrix $\mathbf{A}$ (a vector $\mathbf{v}$), the complex conjugate is denoted by $\mathbf{A}^*$ ($\mathbf{v}^*$). For a real-valued matrix $\mathbf{Q}$, the complex conjugate $\mathbf{Q}^*$ translates to the usual transpose $\mathbf{Q}^T$, and unitarity translates to orthogonality: $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$. The root of -1 is denoted by $\mathbf{i} = \sqrt{-1}$. For a given vector $\mathbf{v}$, $\|\mathbf{v}\|$ denotes the Euclidean norm of $\mathbf{v}$ and for a matrix $\mathbf{A}$, $\|\mathbf{A}\|$ denotes the Frobenius norm of the matrix. The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted by $\mathbf{A} \otimes \mathbf{B}$.

2   RSTLS for Simultaneously Diagonalizable Structures

2.1   Problems Formulation

The RSTLS problem can be written as follows:

$$
\text{(RSTLS):} \quad
\begin{aligned}
\min \quad & \|\mathbf{E}\|^2 + \|\mathbf{w}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 \\
\text{s.t.} \quad & (\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{w}, \\
& \mathbf{E} \in \mathcal{L}, \\
& \mathbf{x} \in \mathbb{F}^n, \mathbf{w} \in \mathbb{F}^m,
\end{aligned}
\tag{2.1}
$$

where $\mathbf{A} \in \mathbb{F}^{m \times n}, \mathbf{b} \in \mathbb{F}^m$ with $\mathbb{F}$ being either the real or the complex number field ($\mathbb{R}$ or $\mathbb{C}$ respectively). The parameter ρ is a positive real number and the set $\mathcal{L}$ is a linear subspace of the set of all $m \times n$ matrices $\mathbb{F}^{m \times n}$. As was discussed in the introduction, this formulation was considered in several papers, see e.g., [23, 26, 28, 10, 22].

In this paper we consider the case in which $m = n$ and $\mathcal{L}$ is a linear subspace of the set of all $n \times n$ matrices diagonalizable by a given unitary matrix. That is, $\mathcal{L} = \mathcal{L}_{\mathbf{Q}}$ where

$$
\mathcal{L}_{\mathbf{Q}} = \{\mathbf{Q}^*\text{diag}(\boldsymbol{\lambda})\mathbf{Q} : \boldsymbol{\lambda} \in \mathbb{F}^n\},
\tag{2.2}
$$

with $\mathbf{Q}$ being a given unitary matrix (i.e., $\mathbf{Q}^*\mathbf{Q} = \mathbf{I}$). Such a structure is called a *simultaneously diagonalizable* (SD) structure. In our derivations we also assume that $\mathbf{A}, \mathbf{L} \in \mathcal{L}_{\mathbf{Q}}$. This particular structure is also discussed in e.g. [23, 26].

In Section 3 we will show that as opposed to most structures, the RSTLS problem with an SD structure can be solved globally and efficiently. Before doing so we will describe some image deblurring examples in which SD structures appear.

2.2   SD Structures Associated with Image Deblurring

We will now present four classes of SD structures that arise naturally in image deblurring problems. In addition to two-dimensional images, we will also consider one-dimensional

signals and refer to them as "one-dimensional images". Before examining the four classes, we briefly review some essential facts and notation from image processing.

Many image deblurring problems can be modelled as $\mathbf{g} = \mathbf{Sf}$, where $\mathbf{g} \in \mathbb{R}^n$ is the blurred image and $\mathbf{f} \in \mathbb{R}^n$ is the unknown true image, whose size is assumed to be the same as the one of $\mathbf{g}$. The matrix $\mathbf{S}$ describes the blur operator. In the case of spatially invariant blurs, $\mathbf{Sf}$ is usually a convolution of a corresponding point spread function (PSF) and the true image $\mathbf{f}$.

The structure of the matrix $\mathbf{S}$ depends on the choice of boundary conditions, that is, the underlying assumptions on the image outside the field of view. Three very popular boundary conditions are: (i) *zero boundary conditions* in which all pixels outside the borders are assumed to be zero (ii) *periodic boundary conditions* in which it is assumed that the image repeats itself in all directions (iii) *reflexive (Neumann) boundary conditions* in which it is assumed that the scene outside of the boundaries is an image mirror of the image boundaries.

Let us illustrate the three types of boundary conditions. First, in the one-dimensional case. Consider the image

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

then for zero, periodic and reflexive boundary conditions the larger image looks like

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ \hline 1 \\ 2 \\ 3 \\ \hline 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 2 \\ 3 \\ \hline 1 \\ 2 \\ 3 \\ \hline 1 \\ 2 \\ 3 \end{pmatrix} \quad \begin{pmatrix} 3 \\ 2 \\ 1 \\ \hline 1 \\ 2 \\ 3 \\ \hline 3 \\ 2 \\ 1 \end{pmatrix}$$

respectively. In the two-dimensional case if we consider the image

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

then for zero, periodic and reflexive boundary conditions the larger image looks like

$$\left(\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 5 & 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7 & 8 & 9 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right), \left(\begin{array}{ccc|ccc|ccc} 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \\ 4 & 5 & 6 & 4 & 5 & 6 & 4 & 5 & 6 \\ 7 & 8 & 9 & 7 & 8 & 9 & 7 & 8 & 9 \\ \hline 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \\ 4 & 5 & 6 & 4 & 5 & 6 & 4 & 5 & 6 \\ 7 & 8 & 9 & 7 & 8 & 9 & 7 & 8 & 9 \\ \hline 1 & 2 & 3 & 1 & 2 & 3 & 1 & 2 & 3 \\ 4 & 5 & 6 & 4 & 5 & 6 & 4 & 5 & 6 \\ 7 & 8 & 9 & 7 & 8 & 9 & 7 & 8 & 9 \end{array}\right), \left(\begin{array}{ccc|ccc|ccc} 9 & 8 & 7 & 7 & 8 & 9 & 9 & 8 & 7 \\ 6 & 5 & 4 & 4 & 5 & 6 & 6 & 5 & 4 \\ 3 & 2 & 1 & 1 & 2 & 3 & 3 & 2 & 1 \\ \hline 3 & 2 & 1 & 1 & 2 & 3 & 3 & 2 & 1 \\ 6 & 5 & 4 & 4 & 5 & 6 & 6 & 5 & 4 \\ 9 & 8 & 7 & 7 & 8 & 9 & 9 & 8 & 7 \\ \hline 9 & 8 & 7 & 7 & 8 & 9 & 9 & 8 & 7 \\ 6 & 5 & 4 & 4 & 5 & 6 & 6 & 5 & 4 \\ 3 & 2 & 1 & 1 & 2 & 3 & 3 & 2 & 1 \end{array}\right)$$

respectively. The structure of the matrix $\mathbf{S}$ depends on the underlying boundary conditions. Here we consider spatially invariant blurs which, as was already mentioned, imply that the

blur is a convolution of given PSF with the true (larger) image. For one-dimensional problems the PSF is just a vector $\mathbf{p} \in \mathbb{R}^d$ with an associated center $c \in \{1, 2, \ldots, d\}$. The convolution operation is then:

$$g_i = \sum_{j=1}^{d} p_j f_{i+c-j}, \quad i = 1, \ldots, n,$$

where $\mathbf{f} \in \mathbb{R}^d$ is the true image. Notice that the above formula uses values of $\mathbf{f}$ beyond the boundaries (indices smaller than 1 and larger than n), but these values are determined by the boundary conditions. For example, consider a one-dimensional image of length three: $\mathbf{f} = (f_1, f_2, f_3)^T$, and let the PSF array be $\mathbf{p} = (p_1, p_2, p_3)^T$ with $c = 2$. Then the blurred image $\mathbf{g}$ depends on the true image $\mathbf{f}$ via the relation $\mathbf{g} = \mathbf{S}\mathbf{f}$ where

$$\mathbf{S} = \begin{pmatrix} p_2 & p_1 & 0 \\ p_3 & p_2 & p_1 \\ 0 & p_3 & p_2 \end{pmatrix}, \begin{pmatrix} p_2 & p_1 & p_3 \\ p_3 & p_2 & p_1 \\ p_1 & p_3 & p_2 \end{pmatrix}, \begin{pmatrix} p_2 + p_3 & p_1 & 0 \\ p_3 & p_2 & p_1 \\ 0 & p_3 & p_2 + p_1 \end{pmatrix}$$

for zero, periodic and reflexive boundary conditions. Note that the above three matrices have different structures (Toeplitz, circulant and Toeplitz-plus-Hankel). We now discuss four SD structures arising from one and two-dimensional problems with either periodic or reflexive boundary conditions[2]

1. **Circulant[8]**. For one-dimensional images with periodic boundary conditions the structure of the model matrix is circulant, i.e., has the form

$$\mathbf{S} = \begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ s_n & s_1 & \cdots & s_{n-1} \\ \vdots & \vdots & & \vdots \\ s_2 & s_3 & \cdots & s_1 \end{pmatrix}.$$

All $n \times n$ circulant matrices are diagonalizable by the unitary discrete Fourier transform (DFT) matrix $\mathbf{F}_n$ given by

$$\mathbf{F}_n = \left(\frac{1}{\sqrt{n}} \omega^{(j-1)(k-1)} \right)_{j,k=1}^{n},$$

where $\omega = e^{\frac{2\pi \mathbf{i}}{n}}$. Multiplications of the DFT matrix $\mathbf{F}_n$ with vectors, as well as eigenvalues computation of circulant matrices can be done very efficiently using the fast fourier transform (FFT) with complexity of $O(n \log n)$.

2. **Block Circulant with Circulant Blocks [2]** For two-dimensional images of size $m \times n$ with periodic boundary conditions the model matrix has a block circulant matrix with circulant blocks (BCCB) structure:

$$\mathbf{S} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 & \cdots & \mathbf{C}_n \\ \mathbf{C}_n & \mathbf{C}_1 & \cdots & \mathbf{C}_{n-1} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}_2 & \mathbf{C}_3 & \cdots & \mathbf{C}_1 \end{pmatrix},$$

---

[2]We do not consider in this paper the zero boundary conditions as it does not lead to an SD structure.

where $\mathbf{C}_1, \ldots, \mathbf{C}_n$ are $m \times m$ circulant matrices. All BCCB matrices of the above size are diagonalizable by the unitary two-dimensional DFT matrix $\mathbf{F}_n \otimes \mathbf{F}_m$. As in the circulant case, computations with BCCB matrices can be performed by using the FFT.

3. **Toeplitz-plus-Hankel [27]** For one-dimensional images with reflexive boundary conditions and symmetric PSF, the matrix $\mathbf{S}$ has a Toeplitz-plus-Hankel structure of the form [27]

$$T(\mathbf{s}) + H(\mathbf{s}),$$

where for a given vector $\mathbf{s} = (s_1, \ldots, s_n)^T \in \mathbb{R}^n$, $T(\mathbf{s})$ is the symmetric Toeplitz matrix whose first column is $\mathbf{s}$ and $H(\mathbf{s})$ is the Hankel matrix whose first and last columns are $(s_1, s_2, \ldots, s_n, 0)^T$ and $(0, s_n, \ldots, s_2, s_1)^T$ respectively. All Toeplitz-plus-Hankel matrices of the above form are diagonalizable by the orthogonal discrete cosine transform (DCT) matrix $\mathbf{C}_n$ given by

$$\mathbf{C}_n = \left(\sqrt{(2 - \delta_{k1})/n} \cos \frac{\pi(2j-1)(k-1)}{2n} \right)_{j,k=1}^n,$$

where for two indices i, j, δ_{ij} denotes the Kronecker sign. Multiplications of the DCT matrix $\mathbf{C}_n$ with vectors, as well as eigenvalues computation of circulant matrices can be done very efficiently using the fast cosine transform (FCT) with complexity of $O(n \log n)$.

4. **BTTB+BTHB+BHTB+BHHB structure [16, 27].** For two-dimensional images of size $m \times n$ with reflexive boundary conditions and symmetric PSF the matrix $\mathbf{S}$ is a sum of a BTTB (block Toeplitz with Toeplitz blocks), BTHB (block Toeplitz with Hankel blocks), BHTB (block Hankel with Toeplitz blocks) and BHHB (block Hankel with Hankel blocks) matrices. All matrices of this form are diagonalizable by the orthogonal two-dimensional DCT matrix $\mathbf{C}_n \otimes \mathbf{C}_m$.

We have thus described four SD structures arising from one and two-dimensional deblurring problems. The first two classes correspond to $\mathbb{F} = \mathbb{C}$ (since the DFT matrix is complex valued) and the last two classes correspond to $\mathbb{F} = \mathbb{R}$. Coming back to the RSTLS problem, we note that it is very natural to assume that the boundary conditions also apply to the regularization operator and we can thus assume that $\mathbf{L} \in \mathcal{L}_{\mathbf{Q}}$.

3   Decomposition of the RSTLS Problem for SD Structures

We begin by showing that the RSTLS problem (2.1) with an SD structure can be decomposed into n one-dimensional minimization problems.

Theorem 3.1 *Consider the RSTLS problem (2.1) with $m = n$ and $\mathcal{L} = \mathcal{L}_{\mathbf{Q}}$ (see (2.2)) where $\mathbf{Q} \in \mathbb{F}^{n \times n}$ is a given unitary matrix. Suppose that $\mathbf{A}, \mathbf{L} \in \mathcal{L}_{\mathbf{Q}}$ and let $\boldsymbol{\alpha}, \mathbf{l}$ be the eigenvalues of $\mathbf{A}$ and $\mathbf{L}$ defined by the relations*

$$\mathbf{Q}\mathbf{A}\mathbf{Q}^* = diag(\boldsymbol{\alpha}), \quad \mathbf{Q}\mathbf{L}\mathbf{Q}^* = diag(\mathbf{l}). \tag{3.1}$$

Then any solution to the RSTLS problem is given by $\mathbf{x} = \mathbf{Q}^*\hat{\mathbf{x}}$ *where for every* $i = 1, \ldots, n$, *the i-th component of* $\hat{\mathbf{x}}$, $\hat{x}_i$, *is an optimal solution to the one dimensional problem*

$$\min_{\hat{x}_i} \left\{ \frac{|\alpha_i \hat{x}_i - \hat{b}_i|^2}{1 + |\hat{x}_i|^2} + \rho |l_i|^2 |\hat{x}_i|^2 \right\}, \tag{3.2}$$

where $\hat{\mathbf{b}} = \mathbf{Q}\mathbf{b}$. *The optimal matrix* $\mathbf{E}$ *is given by*

$$\mathbf{E} = \mathbf{Q}^* diag(\mathbf{r})\mathbf{Q}, \tag{3.3}$$

where

$$r_i = -\frac{\overline{\hat{x}_i}(\alpha_i \hat{x}_i - \hat{b}_i)}{1 + |\hat{x}_i|^2}. \tag{3.4}$$

Proof: Using the relation $\mathbf{w} = (\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b}$, we can rewrite (2.1) as the following problem in the variables $\mathbf{E}$ and $\mathbf{x}$:

$$\min_{\mathbf{E},\mathbf{x}} \{ \|\mathbf{E}\|^2 + \|(\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b}\|^2 + \rho\|\mathbf{L}\mathbf{x}\|^2 : \mathbf{E} \in \mathcal{L}_{\mathbf{Q}}, \mathbf{x} \in \mathbb{F}^n \},$$

which, by the unitarity property of $\mathbf{Q}$, is the same as

$$\min_{\mathbf{E},\mathbf{x}} \{ \|\mathbf{Q}\mathbf{E}\mathbf{Q}^*\|^2 + \|\mathbf{Q}(\mathbf{A} + \mathbf{E})\mathbf{Q}^*\mathbf{Q}\mathbf{x} - \mathbf{Q}\mathbf{b}\|^2 + \rho\|\mathbf{Q}\mathbf{L}\mathbf{Q}^*\mathbf{Q}\mathbf{x}\|^2 : \mathbf{E} \in \mathcal{L}_{\mathbf{Q}}, \mathbf{x} \in \mathbb{F}^n \}. \tag{3.5}$$

Since $\mathbf{E} \in \mathcal{L}_{\mathbf{Q}}$, we can make the change of variables $\mathbf{Q}\mathbf{E}\mathbf{Q}^* = \mathrm{diag}(\mathbf{r})$ where $\mathbf{r} \in \mathbb{F}^n$ is an unknown variables vector. Combing this along with (3.1) we conclude that (3.5) can be reformulated as

$$\min_{\mathbf{r},\hat{\mathbf{x}}} \{ \|\mathrm{diag}(\mathbf{r})\|^2 + \|\mathrm{diag}(\boldsymbol{\alpha} + \mathbf{r})\hat{\mathbf{x}} - \hat{\mathbf{b}}\|^2 + \rho\|\mathrm{diag}(\mathbf{l})\hat{\mathbf{x}}\|^2 : \mathbf{r}, \hat{\mathbf{x}} \in \mathbb{F}^n \},$$

where $\hat{\mathbf{x}} = \mathbf{Q}\mathbf{x}$, and more explicitly as

$$\min_{\mathbf{r},\hat{\mathbf{x}}} \left\{ \sum_{i=1}^{n} (|r_i|^2 + |(\alpha_i + r_i)\hat{x}_i - \hat{b}_i|^2 + \rho|l_i|^2|\hat{x}_i|^2) : \mathbf{r}, \hat{\mathbf{x}} \in \mathbb{F}^n \right\}.$$

The above optimization problem is separable with respect to the pairs of variables

$$(r_1, \hat{x}_1), (r_2, \hat{x}_2), ..., (r_n, \hat{x}_n),$$

implying that for every i, the optimal $(r_i, \hat{x}_i)$ is the solution to the two dimensional problem

$$\min_{r_i, \hat{x}_i} \left\{ |r_i|^2 + |(\alpha_i + r_i)\hat{x}_i - \hat{b}_i|^2 + \rho|l_i|^2|\hat{x}_i|^2 : r_i, \hat{x}_i \in \mathbb{F} \right\}. \tag{3.6}$$

Next, we fix $\hat{x}_i$ and minimize with respect to r_i. The result is:

$$r_i = -\frac{\overline{\hat{x}_i}(\alpha_i \hat{x}_i - \hat{b}_i)}{1 + |\hat{x}_i|^2}.$$

Substituting the above expression back into the objective function of (3.6) with some simple algebraic manipulations, we arrive at the following equivalent problem in the single variable $\hat{x}_i$:

$$\min_{\hat{x}_i} \left\{ \frac{|\alpha_i \hat{x}_i - \hat{b}_i|^2}{1 + |\hat{x}_i|^2} + \rho|l_i|^2|\hat{x}_i|^2 \right\},$$

establishing the result. $\square$

4 Solution and Analysis of the RSTLS Problem for SD Structures

In this section we study the 1D problems (3.2) arising in the decomposition of the RSTLS problem. We show in Section 4.1 that although these problems are not unimodal[3], they can be transformed into (strictly) unimodal problems and consequently solved efficiently and globally. This is especially crucial in image deblurring applications in which there are hundreds of thousands or even millions of 1D problems to be solved. Based on uniqueness and attainment properties of the 1D problems, corresponding conditions for the RSTLS problem are established in Section 4.2.

4.1 Solution of the Single Variable Problem

Our goal in this section is to analyze the one-dimensional problem (3.2) and to devise an efficient solution method for solving it. Consider the problem

$$\min_{x \in \mathbb{F}} \left\{ f(x) = \frac{|ax - b|^2}{1 + |x|^2} + |c|^2 |x|^2 \right\}, \tag{4.1}$$

where $a, b, c \in \mathbb{F}$. If $c \neq 0$ then the objective function is coercive and consequently its minimum is attained. The objective function of (4.1) is not unimodal (c.f. Figure 1) and thus finding its global minimum efficiently is in principle a hard task. We will show in the next result that it can be solved via the minimization problem

$$\min_{y \geq 0} \left\{ g(y) \equiv \frac{|a|^2 y - 2|ab|\sqrt{y} + |b|^2}{1 + y} + |c|^2 y \right\} \tag{4.2}$$

in the real variable y. Before stating the result we briefly recall that for a real number $x \in \mathbb{R}$ the sign function is defined by

$$\mathrm{sgn}\,(x) \equiv \begin{cases} 1 & x > 0, \\ 0 & x = 0, \\ -1 & x < 0, \end{cases}$$

and for a complex number $z \in \mathbb{C}$ the sign function is given by

$$\mathrm{sgn}\,(z) \equiv \begin{cases} \frac{z}{|z|} & z \neq 0, \\ 0 & z = 0. \end{cases}$$

Lemma 4.1 (Equivalence of Problems (4.1) and (4.2)) *Consider problem (4.1) with $a, b, c \in \mathbb{F}$. Then*
(i). If $ab \neq 0$ then $\tilde{y}$ is an optimal solution of (4.2) if and only if $\tilde{x} = \mathrm{sgn}\,(\bar{a}b)\sqrt{\tilde{y}}$ is an optimal solution of (4.1).
(ii). If $ab = 0$ then $\tilde{y}$ is an optimal solution of (4.2) if and only if $\tilde{x} = z\sqrt{\tilde{y}}$ is an optimal solution of (4.1) for every $z \in \mathbb{F}$ satisfying $|z| = 1$.

---

[3]A function $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ being a closed interval, is (strictly) unimodal if it has a unique local minimizer on I and is (strictly) decreasing from the left boundary of the interval to this unique minimum and (strictly) increasing from the minimum to the right boundary of the interval.

Proof: Let $\tilde{x}$ be an optimal solution of (4.1). Then by the optimality of $\tilde{x}$ we have

$$f(\tilde{x}) \leq f(z\tilde{x}) \text{ for every } z \in \mathbb{F} \text{ satisfying } |z| = 1,$$

which is the same as

$$\frac{|a\tilde{x} - b|^2}{1 + |\tilde{x}|^2} + |c|^2|\tilde{x}|^2 \leq \frac{|a(z\tilde{x}) - b|^2}{1 + |z\tilde{x}|^2} + |c|^2|z\tilde{x}|^2.$$

The latter inequality reduces to

$$\Re((1 - z)a\bar{b}\tilde{x}) \geq 0. \tag{4.3}$$

We will now show that $a\bar{b}\tilde{x}$ is a nonnegative real number. This is obviously true if $\tilde{x} = 0$. Otherwise, we split the analysis into two cases:

Case I: If $ab \neq 0$, then substituting

$$z = \frac{\overline{a\bar{b}\tilde{x}}}{|a\bar{b}\tilde{x}|}$$

into (4.3) yields

$$\Re(a\bar{b}\tilde{x}) \geq |a\bar{b}\tilde{x}|,$$

implying that $a\bar{b}\tilde{x}$ is a nonnegative real number and, in particular, that $\text{sgn}(\tilde{x}) = \text{sgn}(\bar{a}b)$.

Case II: If $ab = 0$, the function f satisfies $f(zx) = f(x)$ for every $x, z \in \mathbb{F}$ such that $|z| = 1$ and thus $z\tilde{x}$ is also an optimal solution for every z satisfying $|z| = 1$.

A conclusion from the above two cases is that if the minimum of (4.1) is attained at a nonzero solution then there must be at least one optimal solution $\tilde{x}$ for which $\text{sgn}(\tilde{x}) = \text{sgn}(\bar{a}b)$; consequently, we can make the change of variables $x = \text{sgn}(\bar{a}b)\sqrt{y}$ which transforms problem (4.1) into (4.2). $\square$

Remark 4.1 *Consider problem (4.1) with $\mathbb{F} = \mathbb{C}$ but with real data, i.e., $a, b, c \in \mathbb{R}$. Then a direct consequence of Lemma 4.1 is that if the optimal set of (4.1) is nonempty, then there must exist at least one real-valued optimal solution.*

The following simple lemma establishes some key properties of problem (4.2). In particular, it is shown that problem (4.2) is strictly unimodal (in all interesting cases) and thus can be solved efficiently. This is in fact the main motivation for transforming problem (4.1) into (4.2).

Lemma 4.2 (Properties of Problem (4.2)) *Consider problem (4.2) with $a, b, c \in \mathbb{F}$. Then*
(i). The objective function $g(y)$ of (4.2) is quasiconvex[4] over $[0, \infty)$.
(ii). If $c \neq 0$ and $\tilde{y}$ is an optimal solution of (4.2) then $\tilde{y} \leq \frac{|b|^2}{|c|^2}$.
(iii). The solution of (4.2) is attained and unique if and only if $(a, c) \neq (0, 0)$.
(iv). If $(a, c) \neq (0, 0)$, then the objective function $g(y)$ of (4.2) is strictly unimodal over $[0, \infty)$.

---

[4]A function $f : I \to \mathbb{R}$ ($I \subseteq \mathbb{R}$ being an interval) is quasiconvex if all its level sets $\{x \in I : f(x) \leq \alpha\}$ are convex.

Proof: (i). We need to show that the level set $\{y : g(y) \le \alpha\}$ is convex. Indeed,

$$\{y \ge 0 : g(y) \le \alpha\} = \{y \ge 0 : (|a|^2 + |c|^2 - \alpha)y - 2|ab|\sqrt{y} + |c|^2 y^2 + |b|^2 - \alpha \le 0\}.$$

The latter is the zero level set of a convex function and hence convex.

(ii). Note that for $y \ge 0$

$$g(y) = \frac{(|a|\sqrt{y} - |b|)^2}{1 + y} + |c|^2 y \ge |c|^2 y.$$

Therefore, for $y > \frac{|b|^2}{|c|^2}$ we have

$$g(y) \ge |c|^2 y > |b|^2 = g(0),$$

showing that there are no optimal solutions for (4.2) larger than $\frac{|b|^2}{|c|^2}$.

(iii). First consider the case $(a, c) = (0, 0)$. Then $g(y) = |b|^2/(1 + y)$. Hence it follows that g either does not attain a minimum (if $b \ne 0$) or that the minimum (namely all $y \ge 0$) is nonunique (if $b = 0$). Now consider the case $(a, c) \ne (0, 0)$. We split the analysis into two subcases.

Case I: If $c \ne 0$ then $\lim_{y \to \infty} g(y) = \infty$, implying the attainment of the minimum. To show the uniqueness of the minimum in this case, assume in contradiction that the optimal solution of (4.2) is not unique. Then since the optimal set is convex (by quasiconvexity) we conclude that the optimal set is an interval $I \subseteq [0, \infty)$ with a nonempty interior. Denote the optimal value by f^*. Then

$$g(y) = f^* \text{ for every } y \in I,$$

which can be explicitly written as

$$(|a|^2 + |c|^2 - f^*)y - 2|ab|\sqrt{y} + |c|^2 y^2 + |b|^2 - f^* = 0 \text{ for every } y \in I.$$

Making the change of variables $z = \sqrt{y}$, we obtain

$$(|a|^2 + |c|^2 - f^*)z^2 - 2|ab|z + |c|^2 z^4 + |b|^2 - f^* = 0 \text{ for every } z \in J, \quad (4.4)$$

where $J = \{z : z^2 \in I\}$ is an interval with a nonempty interior. However, (4.4) is impossible since an univariate quartic equation has at most four roots and thus can not have an infinite number of roots.

Case II: Suppose $c = 0$. Then $a \ne 0$, and it is easy to see that g attains a unique minimum at $|b|^2/|a|^2$.

(iv). Since $(a, c) \ne (0, 0)$, we know from part (iii) that g attains a unique global minimum on the interval $[0, \infty)$. Hence it remains to show that it is strictly decreasing from the origin to this minimum, and strictly increasing when we go from this minimum to plus infinity. Suppose this is not true. Then the function g must have a stationary point in $(0, \infty)$ which is different from the unique minimum. However, we will show that for any $y > 0$ such that $g'(y) = 0$, we automatically have $g''(y) > 0$, hence this stationary point y is at least a local
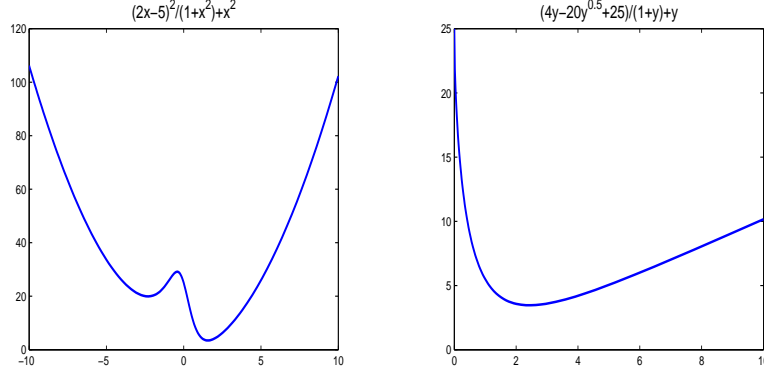
11

Figure 1: The objective function of problem (4.5) (left) and (4.6) (right)

minimum and, therefore, must be equal to the unique minimum of g on the interval $[0, \infty)$. Elementary differentiation gives

$$g'(y) = \frac{(|a|\sqrt{y} - |b|)(|a|\frac{1}{\sqrt{y}} + |b|)}{(1+y)^2} + |c|^2.$$

Now let $\tilde{y} > 0$ be such that $g'(\tilde{y}) = 0$. Then

$$
\begin{aligned}
g''(\tilde{y}) \quad &= \quad -\frac{2}{(1+\tilde{y})^3}(|a|\sqrt{\tilde{y}} - |b|)\left(|a|\frac{1}{\sqrt{\tilde{y}}} + |b|\right) + \frac{|a||b|}{2(1+\tilde{y})^2\sqrt{\tilde{y}}}\left(1 + \frac{1}{\tilde{y}}\right) \\
&\stackrel{g'(\tilde{y})=0}{=} \quad \frac{2|c|^2}{1+\tilde{y}} + \frac{|a||b|}{2(1+\tilde{y})^2\sqrt{\tilde{y}}}\left(1 + \frac{1}{\tilde{y}}\right) > 0,
\end{aligned}
$$

where the positivity of the last expression comes from the fact that $(a, c) \neq (0, 0)$, hence this last expression can be equal to zero only if both $c = 0$ and $b = 0$, but then, taking into account that $\tilde{y}$ is also a stationary point, we would obtain $a = 0$ as well, in contrast to $(a, c) \neq (0, 0)$. $\square$

The most important property of the function g is its strict unimodality (as stated in Lemma 4.2 (iv)). The strict unimodality property implies that there are no non-global local minima and thus enables us to invoke efficient one-dimensional solvers for (strictly) unimodal functions that are guaranteed to converge to the global minimum. The following example illustrates this property.

Example 1: Consider problem (4.1) with $\mathbb{F} = \mathbb{R}$, $a = 2, b = 5, c = 1$. In this case, problems (4.1) and (4.2) are given by

$$\min_x \left\{ \frac{(2x-5)^2}{1+x^2} + x^2 \right\}, \tag{4.5}$$

and

$$\min_{y \geq 0} \left\{ \frac{4y - 20\sqrt{y} + 25}{1+y} + y \right\} \tag{4.6}$$

respectively. The plots of the two functions are given in Figure 1.

Clearly, the objective function in (4.5) is not unimodal and indeed possesses a non-global local minimizer. The global solution of (4.5) is $\tilde{x} = 1.5606$ (given in four digits accuracy).
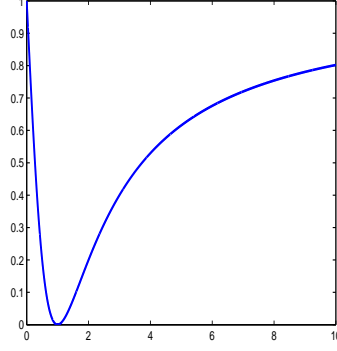
12

Figure 2: The function $\frac{(\sqrt{y}-1)^2}{1+y}$ from Remark 4.2

The objective function in (4.6) is, as guaranteed by Lemma 4.1, an unimodal function. The global minimum is $\tilde{y} = 2.4354$ and the relation $\tilde{x} = \sqrt{\tilde{y}}$ holds. $\square$

Remark 4.2 *A natural question here is whether g is even more than quasiconvex, namely convex. The answer to this question is negative. For example, for $a = b = 1, c = 0$, the function g is clearly nonconvex as can be seen from Figure 2. Note, however, that this figure also illustrates the quasiconvexity of g.*

Combining Lemmata 4.1 and 4.2, we are now able to state the basic properties of problem (4.1).

Lemma 4.3 (Uniqueness for Problem (4.1)) *The optimal solution of problem (4.1) is uniquely attained if and only if one of the following two conditions holds:*
(i). $a \neq 0$.
(ii). $a = 0, c \neq 0$ and $|b| \leq |c|$.

Proof: We will split the analysis into four cases:
Case I: $a \neq 0, b \neq 0$.
By Lemma 4.2 (iii), since $a \neq 0$, the optimal solution of (4.2) is uniquely attained. Moreover, since $ab \neq 0$, then by Lemma 4.1, there is a one to one correspondence between optimal solutions of (4.1) and (4.2) (via the relation $\tilde{x} = \text{sgn}\,(\bar{a}b)\sqrt{\tilde{y}}$), implying the uniqueness and attainment of the optimal solution of (4.1).
Case II: $a \neq 0, b = 0$.
The objective function of (4.2) in this case is strictly increasing implying that the unique optimal solution of (4.2) is $\tilde{y} = 0$ and hence that the unique optimal solution of (4.1) is $\tilde{x} = 0$.
Case III: $a = 0, b \neq 0$.
By Lemma 4.2 (iii), to guarantee the uniqueness and attainment of the optimal solution of (4.2) we must further assume that $c \neq 0$. The solution of (4.1) is unique if and only if the optimal solution $\tilde{y}$ of (4.2) is zero (otherwise, $z\sqrt{\tilde{y}}$ will be an optimal solution of (4.1) for every z satisfying $|z| = 1$). By the unimodality of g, the optimal solution is 0 if and only if

13

$g'(0) \geq 0$ which is equivalent to $|b| \leq |c|$.

Case IV: $a = 0, b = 0$.

Again, as in the previous case, we further assume that $c \neq 0$. Here it is evident that the unique optimal solution is $\tilde{x} = 0$.

Combining the four cases we obtain the result. $\square$

4.2 Uniqueness and Attainment of the RSTLS Solution

The result in Section 4.1 collectively can be summed up in the following result.

Theorem 4.1 *Consider the RSTLS problem (2.1) with $m = n$ and $\mathcal{L} = \mathcal{L}_{\mathbf{Q}}$ (see (2.2)) where $\mathbf{Q} \in \mathbb{F}^{n \times n}$ is a given unitary matrix. Let $\hat{\mathbf{b}} = \mathbf{Q}^*\mathbf{b}$. Suppose further that $\mathbf{A}, \mathbf{L} \in \mathcal{L}_{\mathbf{Q}}$ and let $\boldsymbol{\alpha}, \mathbf{l}$ be the eigenvalues of $\mathbf{A}$ and $\mathbf{L}$ given by the relations:*

$$\mathbf{Q}^*\mathbf{A}\mathbf{Q} = diag(\boldsymbol{\alpha}), \quad \mathbf{Q}^*\mathbf{L}\mathbf{Q} = diag(\mathbf{l}). \tag{4.7}$$

Then the solution to the RSTLS problem is uniquely attained if and only if for each $i = 1, \ldots, n$ one of the following two conditions is satisfied:

(i). $\alpha_i \neq 0$.

(ii). $\alpha_i = 0, l_i \neq 0$ and $|\hat{b}_i| \leq \sqrt{\rho}|l_i|$.

Proof: Note that the optimal $\mathbf{E}$ is uniquely defined via the optimal $\mathbf{x}$ by (3.3) and (3.4). Therefore, uniqueness and/or attainment properties of the optimal solution of (2.1) amounts to uniqueness and/or attainment of the single variable problems (3.2), which combined with Lemma 4.3 establishes the result. $\square$

Theorem 4.1 provides conditions for the optimal solution of the RSTLS problem to be uniquely attained. Based on this, we can derive a simpler condition:

$$\text{Null}(\mathbf{A}) \cap \text{Null}(\mathbf{L}) = \{\mathbf{0}\}, \tag{4.8}$$

which is sufficient for attainment of the optimal solution and necessary for the unique attainment of the optimal solution, as shown in the following theorem.

Theorem 4.2 *Consider the setting of Theorem 4.1. Then*

(i). If the optimal solution of (2.1) is uniquely attained then condition (4.8) is satisfied.

(ii). If condition (4.8) is satisfied then the optimal solution set of (2.1) is nonempty.

(iii). If $\mathbf{A}$ is nonsingular then the solution of (2.1) is uniquely attained.

Proof: (i). Note that by Theorem 4.1 a necessary condition for the optimal solution of (2.1) to be uniquely attainted is that $|\alpha_i|^2 + |l_i|^2 \neq 0$ for every i, that is, α_i and l_i are not both zero for any given i. The eigenvalues of the matrix $\mathbf{A}^*\mathbf{A} + \mathbf{L}^*\mathbf{L}$ are exactly $|\alpha_i|^2 + |l_i|^2$ implying that $\mathbf{A}^*\mathbf{A} + \mathbf{L}^*\mathbf{L}$ is nonsingular; therefore,

$$\text{Null}(\mathbf{A}) \cap \text{Null}(\mathbf{L}) = \text{Null}(\mathbf{A}^*\mathbf{A} + \mathbf{L}^*\mathbf{L}) = \{\mathbf{0}\}.$$

(ii). Assume that condition (4.8) holds. By Theorem 3.1, it is enough to show that for every $i = 1, \ldots, n$ the one-dimensional problem (3.2) has at least one optimal solution. Now, by Lemma 4.1 it is sufficient to establish the attainment of the solution of

$$\min_{y \geq 0} \left\{ \frac{|\alpha_i|^2 y - 2|\alpha_i \hat{b}_i| \sqrt{y} + |\hat{b}_i|^2}{1 + y} + \rho |l_i|^2 y \right\} \qquad (4.9)$$

for every $i = 1, \ldots, n$ where $\alpha_i, \hat{b}_i$ and l_i are defined in the premise of Theorem 3.1. By Lemma 4.2 (iii), this is guaranteed if $(\alpha_i, l_i) \neq (0, 0)$ for every i, which, as shown in the proof of (i), is equivalent to condition (4.8).

(iii). It follows from the nonsingularity of $\mathbf{A}$ that all its eigenvalues are nonzero, which, by Theorem 4.1, implies that the solution of (2.1) is uniquely attained. $\square$

The following example shows by suitable counterexamples that the assumptions used in Theorem 4.2 are sufficient, but not necessary for the corresponding statements to be true.

Example 2 (i) Consider problem (2.1) with $n = m = 1, A = (0), L = (1), b = (2), \rho = 1$, and $\mathbb{F} = \mathbb{R}$. Then condition (4.8) holds, but problem (2.1) has the two solutions $(E, x) = (1, 1)$ and $(E, x) = (-1, -1)$. This shows that the unique attainment of a solution of problem (2.1) is sufficient for condition (4.8) to hold, but not necessary.

(ii) Consider problem (2.1) with $n = m = 1, A = (0), L = (0), b = (0), \rho = 1$, and $\mathbb{F} = \mathbb{R}$. Then every vector (E, x) with $E = 0$ and $x \in \mathbb{R}$ arbitrary is a solution of problem (2.1), although condition (4.8) does not hold. Hence this condition is sufficient for problem (2.1) to have a nonempty solution set, but not necessary.

It is interesting to compare the above conditions to the corresponding attainment/uniqueness conditions for the regularized least squares problem:

$$\text{(RLS):} \quad \min \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \rho \|\mathbf{L}\mathbf{x}\|^2.$$

The optimal solution of (RLS), as opposed to the solution of the RSTLS problem, is always attained; it is unique if and only if condition (4.8) holds. This is in contrast to the RSTLS problem where condition (4.8) is only a necessary condition for unique attainment of the solution.

5 The RSTLS Problem with Circulant Structure

The RSTLS problem (2.1) with $\mathcal{L} = \mathcal{L}_{\mathbf{F}_n}$ ($\mathbf{F}_n$ being the $n \times n$ DFT matrix) corresponds to problems with circulant structured matrices. Here the underlying number field is $\mathbb{F} = \mathbb{C}$ since the matrix $\mathbf{F}_n$ is complex-valued. However, in many applications the data $\mathbf{A}, \mathbf{b}$ and $\mathbf{L}$ is real-valued. The main result in this section is that if the optimal set of the RSTLS problem is nonempty, then there exists at least one real-valued optimal solution. Therefore, there is no drawback in analyzing the RSTLS problem over the complex field even when the data is real-valued.

Theorem 5.1 *Consider the RSTLS problem with $\mathbb{F} = \mathbb{C}, \mathcal{L} = \mathcal{L}_{\mathbf{F}_n}$ with $\mathbf{F}_n$ being the $n \times n$ DFT matrix. Assume that $\mathbf{A}$, $\mathbf{b}$ and $\mathbf{L}$ are real-valued, that is $\mathbf{A} \in \mathbb{R}^{n \times n}, \mathbf{b} \in \mathbb{R}^n, \mathbf{L} \in \mathbb{R}^{n \times n}$. If the optimal set of (RSTLS) is nonempty, then there exists at least one optimal real-valued solution.*

Proof: We will require the following notation:

$$\mathcal{A} = \{\mathbf{z} \in \mathbb{C}^n : z_1 \in \mathbb{R}, z_{j+1} = \overline{z_{n+1-j}} \text{ for every } j = 1, \ldots, n-1\}.$$

To simplify the notation we omit the subscript in the $n \times n$ DFT matrix and denote it by $\mathbf{F}$ rather than by $\mathbf{F}_n$. The proof is based on the following three claims:
(i) Let $\mathbf{w} = \mathbf{F}\mathbf{v}$ for some $\mathbf{v} \in \mathbb{R}^n$. Then $\mathbf{w} \in \mathcal{A}$.
(ii) Let $\boldsymbol{\alpha}$ be the vector of eigenvalues of a real-valued circulant matrix $\mathbf{A}$. Then $\boldsymbol{\alpha} \in \mathcal{A}$.
(iii) Let $\mathbf{z} \in \mathcal{A}$. Then $\mathbf{F}^*\mathbf{z} \in \mathbb{R}^n$.
Proof of (i). First,

$$w_1 = (\mathbf{F}\mathbf{v})_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_i,$$

proving that $w_1 \in \mathbb{R}$. Next, for every $j = 1, \ldots, n-1$ we have

$$w_{j+1} = (\mathbf{F}\mathbf{v})_{j+1} = \sum_{i=1}^{n} F_{j+1,i} v_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \omega^{j(i-1)} v_i. \tag{5.1}$$

On the other hand,

$$\overline{w_{n+1-j}} = \overline{(\mathbf{F}\mathbf{v})_{n+1-j}} = \sum_{i=1}^{n} \overline{F_{n+1-j,i}} v_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \overline{\omega^{(n-j)(i-1)}} v_i = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \omega^{j(i-1)} v_i \stackrel{(5.1)}{=} w_{j+1}.$$

Proof of (ii). Let $(s_1, s_2, \ldots, s_n)$ be the first row of $\mathbf{A}$. The j-th eigenvalue of the circulant matrix $\mathbf{A}$ is given by $\alpha_j = \sum_{i=1}^{n} \omega^{(i-1)(j-1)} s_i$. Then

$$\alpha_1 = \sum_{i=1}^{n} s_i \in \mathbb{R}$$

and

$$\overline{\alpha_{n+1-j}} = \overline{\sum_{i=1}^{n} \omega^{(i-1)(n-j)} s_i} = \sum_{i=1}^{n} \omega^{(i-1)j} s_i = \alpha_{j+1}$$

for every $j = 1, \ldots, n-1$. Thus, $\boldsymbol{\alpha} \in \mathcal{A}$.

Proof of (iii). For every $i = 1, 2, \ldots, n$:

$$
\begin{aligned}
\sqrt{n}(\mathbf{F}^*\mathbf{w})_i \quad &= \quad \sqrt{n}\sum_{j=1}^{n}\overline{F_{j,i}}w_j = \sum_{j=1}^{n}\omega^{-(i-1)(j-1)}w_j \\
&\stackrel{\mathbf{w}\in\mathcal{A}}{=} \quad w_1 + \sum_{j=2}^{n}\omega^{-(i-1)(j-1)}\overline{w_{n+2-j}} = w_1 + \sum_{j=2}^{n}\omega^{(i-1)(n+1-j)}\overline{w_{n+2-j}} \\
&\stackrel{k\leftarrow n+2-j}{=} \quad w_1 + \sum_{k=2}^{n}\omega^{(i-1)(k-1)}\overline{w_k} = \sqrt{n}\sum_{k=1}^{n}F_{k,i}\overline{w_k} = \\
&= \quad \overline{\sqrt{n}\sum_{k=1}^{n}\overline{F_{k,i}}w_k} = \sqrt{n}\,\overline{(\mathbf{F}^*\mathbf{w})_i}.
\end{aligned}
$$

By Theorem 3.1, an optimal solution of the RSTLS problem is given by $\mathbf{x} = \mathbf{F}^*\hat{\mathbf{x}}$ where $\hat{x}_i$, the i-th component of $\hat{\mathbf{x}}$, is an optimal solution of (3.2). Recall that $\hat{\mathbf{b}} = \mathbf{F}\mathbf{b}$ for real-valued $\mathbf{b}$ and that $\boldsymbol{\alpha}, \mathbf{l}$ are the eigenvalues vectors of the real-valued circulant matrices $\mathbf{A}$ and $\mathbf{L}$ respectively. Therefore, by properties (i) and (ii), $\hat{\mathbf{b}}, \boldsymbol{\alpha}, \mathbf{l} \in \mathcal{A}$. Hence, $\hat{x}_1$ is the solution of (3.2) with $i = 1$ and with real data, which by Remark 4.1 implies that $\hat{x}_1$ is real. Moreover, for every $j = 1, \ldots, n-1$, $\hat{x}_j$ and $\hat{x}_{n+1-j}$ are the optimal solutions of

$$
\min_{\hat{x}_{j+1}}\left\{\frac{|\alpha_{j+1}\hat{x}_{j+1} - \hat{b}_{j+1}|^2}{1 + |\hat{x}_{j+1}|^2} + \rho|l_j|^2|\hat{x}_j|^2\right\} , \min_{\hat{x}_{n+1-j}}\left\{\frac{|\alpha_{j+1}\overline{\hat{x}_{n+1-j}} - \hat{b}_{j+1}|^2}{1 + |\overline{\hat{x}_{n+1-j}}|^2} + \rho|l_{j+1}|^2|\overline{\hat{x}_{n+1-j}}|^2\right\}
$$

respectively. Therefore, we can always choose the optimal solutions of these problems to satisfy $\hat{x}_{n+1-j} = \overline{\hat{x}_{j+1}}$. Thus, for the mentioned choice $\hat{\mathbf{x}} \in \mathcal{A}$ and by property (iii) this proves that $\mathbf{x} = \mathbf{F}^*\hat{\mathbf{x}}$ is real-valued. $\square$

Remark 5.1 *It can be shown using the same methodology employed in the proof of Theorem 5.1 that there always exists a real-valued solution for the RSTLS problem with $\mathbf{Q} = \mathbf{F}_n \otimes \mathbf{F}_m$ (BCCB structure) whenever $\mathbf{A}, \mathbf{L}$ and $\mathbf{b}$ are real-valued.*

The following two examples demonstrate the validity of Theorem 5.1.
Example 3: Let $\mathbf{Q} = \mathbf{F}_3$ (3×3 circulant matrices) and

$$
\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \mathbf{L} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}, \rho = 1.
$$

Then

$$
\boldsymbol{\alpha} = \mathrm{diag}(\mathbf{F}_3\mathbf{A}\mathbf{F}_3^*) = \begin{pmatrix} 6 \\ -1.5 - 0.866025\mathbf{i} \\ -1.5 + 0.866025\mathbf{i} \end{pmatrix},
$$

$$
\hat{\mathbf{b}} = \mathbf{F}_3\mathbf{b} = \begin{pmatrix} 8.660254 \\ -0.866025 + 0.5\mathbf{i} \\ -0.866025 - 0.5\mathbf{i} \end{pmatrix}, \mathbf{l} = \begin{pmatrix} 0 \\ 1.5 - 0.866025\mathbf{i} \\ 1.5 + 0.866025\mathbf{i} \end{pmatrix}.
$$

17

The vector, $\hat{\mathbf{x}}$ consisting of the optimal solutions the three arising optimization problems is:

$$\hat{\mathbf{x}} = \begin{pmatrix} 1.443375 \\ 0.143941 - 0.249314\mathbf{i} \\ 0.143941 + 0.249314\mathbf{i} \end{pmatrix}$$

and the optimal solution

$$\mathbf{x} = \mathbf{F}_3^* \hat{\mathbf{x}} = \begin{pmatrix} 0.999543 \\ 0.999543 \\ 0.500913 \end{pmatrix}$$

is indeed real.

Example 4: Consider the RSTLS problem with $\mathbf{Q} = \mathbf{F}_3$ (3×3 circulant matrices) and

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \mathbf{L} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}, \rho = 1.$$

Then

$$\boldsymbol{\alpha} = \mathrm{diag}(\mathbf{F}_3 \mathbf{A} \mathbf{F}_3^*) = \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}, \hat{\mathbf{b}} = \mathbf{F}_3 \mathbf{b} = \begin{pmatrix} 6.928203 \\ -1.732050 + \mathbf{i} \\ -1.732050 - \mathbf{i} \end{pmatrix}.$$

In this example the optimal solutions of the arising one-dimensional problems are not unique and they consist of the collection of vectors $\hat{\mathbf{x}}$ of the form:

$$\hat{\mathbf{x}} = \begin{pmatrix} 2.309401 \\ 0.393319 z_1 \\ 0.393319 z_2 \end{pmatrix},$$

where z_1 and z_2 are complex numbers satisfying $|z_1| = |z_2| = 1$. Correspondingly, the set of optimal solutions of (RSTLS) consists of all vectors $\mathbf{F}_3^* \hat{\mathbf{x}}$ where $\hat{\mathbf{x}}$ is of the above form and is thus equal to

$$\{(a + z_1 b + z_2 c, a + b z_1 \bar{\omega} + c z_2 \omega, a + b z_1 \omega + c z_2 \bar{\omega})^T : |z_1| = |z_2| = 1\},$$

where $a = 2.309401, b = 0.393319$ and $\omega = e^{\frac{2\pi \mathbf{i}}{3}}$. The above set certainly contains complex-valued optimal solutions, but if we choose $z_1 = \bar{z}_2$ we obtain a subset of real-valued optimal solutions.

$$\left\{ \frac{1}{\sqrt{3}} (a + 2\cos(\theta)c, a + 2\cos(\theta + 2\pi/3)c, a + 2\cos(\theta - 2\pi/3)b)^T : 0 \le \theta \le 2\pi \right\}.$$

6 Solution of the CSTLS Problem with SD structure

When the regularization is made by adding a constraint rather than by penalization, the problem becomes

$$\text{(CSTLS):} \quad \begin{array}{ll} \min & \|\mathbf{E}\|^2 + \|(\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b}\|^2 \\ \text{s.t.} & \|\mathbf{L}\mathbf{x}\|^2 \le \alpha, \\ & \mathbf{E} \in \mathcal{L}_{\mathbf{Q}}, \\ & \mathbf{x} \in \mathbb{F}^n, \end{array} \qquad (6.1)$$

where $\alpha > 0$. We will show that the CSTLS problem can be solved by a sequence of RSTLS problems using a dual approach. We assume throughout this section that $\mathbf{A}$ is nonsingular. This assumption prevails in many image deblurring problems, although the matrix is often extremely ill-conditioned.

The Lagrangian dual problem of (6.1) is given by

$$\max_{\lambda \geq 0} q(\lambda), \tag{6.2}$$

where

$$q(\lambda) = \begin{array}{ll} \min & \|\mathbf{E}\|^2 + \|(\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b}\|^2 + \lambda(\|\mathbf{L}\mathbf{x}\|^2 - \alpha) \\ \text{s.t.} & \mathbf{E} \in \mathcal{L}_{\mathbf{Q}}, \mathbf{x} \in \mathbb{F}^n. \end{array} \tag{6.3}$$

Therefore, evaluating a value of the dual objective function amounts to solving a single RSTLS problem which can be solved efficiently as shown the in previous sections. Since $\mathbf{A}$ is nonsingular then by Theorem 4.2 (iii), the optimal solution of (6.3) is uniquely attained for all $\lambda \geq 0$ and we denote it by $(\mathbf{x}_\lambda, \mathbf{E}_\lambda)$. The function q has several important properties which are summarized in Lemma 6.1 below. The differentiability property of q (part (ii) of Lemma 6.1), relies on the uniqueness property and on the following well known result [7, Proposition 6.1.1].

Theorem 6.1 *Let f, g be continuous functions defined on a compact set X. Let*

$$h(\lambda) \equiv \min_{\mathbf{x} \in X} \{f(\mathbf{x}) + \lambda g(\mathbf{x})\}, \quad \lambda \in [\lambda_1, \lambda_2]$$

and assume that there exists a unique minimizer $\mathbf{x}_\lambda$ to the above optimization problem for every $\lambda \in [\lambda_1, \lambda_2]$ denoted by $\mathbf{x}_\lambda$. Then h is differentiable for every $\lambda \in (\lambda_1, \lambda_2)$ and $h'(\lambda) = g(\mathbf{x}_\lambda)$.

In our case the compactness assumption is not satisfied; however, this difficulty can be avoided. We will use the following notation

$$\begin{array}{rcl} s(\mathbf{x}, \mathbf{E}) & = & \|\mathbf{E}\|^2 + \|(\mathbf{A} + \mathbf{E})\mathbf{x} - \mathbf{b}\|^2, \\ t(\mathbf{x}, \mathbf{E}) & = & \|\mathbf{L}\mathbf{x}\|^2 - \alpha, \\ Y & = & \{(\mathbf{x}, \mathbf{E}) : \mathbf{x} \in \mathbb{F}^n, \mathbf{E} \in \mathcal{L}_{\mathbf{Q}}\}. \end{array}$$

Then, in this notation, the CSTLS problem can be written as

$$\min_{\mathbf{x}, \mathbf{E}} \{s(\mathbf{x}, \mathbf{E}) : t(\mathbf{x}, \mathbf{E}) \leq 0, (\mathbf{x}, \mathbf{E}) \in Y\}. \tag{6.4}$$

Lemma 6.1 *Consider the function q given by (6.3). Then*
(i). q is concave over $[0, \infty)$.
(ii). $q(\lambda)$ is differentiable for every $\lambda > 0$ and $q'(\lambda) = \|\mathbf{L}\mathbf{x}_\lambda\|^2 - \alpha$.
(iii). $\lim_{\lambda \to \infty} q(\lambda) = -\infty$.

Proof: (i). $q(\lambda)$ is the pointwise minimum of functions which are linear in λ and hence concave.

(ii). Let $\tilde{\lambda} > 0$ and let $\lambda_2 > \lambda_1 > 0$ be two positive numbers for which $\tilde{\lambda} \in (\lambda_1, \lambda_2)$. The dual objective can be written as

$$q(\lambda) = \min\{s(\mathbf{x}, \mathbf{E}) + \lambda t(\mathbf{x}, \mathbf{E}) : (\mathbf{x}, \mathbf{E}) \in Y\}. \tag{6.5}$$

From the nonsingularity of $\mathbf{A}$ and Theorem 4.2 (iii) it follows that there exists a unique minimizer to the above problem which we denote by $(\mathbf{x}_\lambda, \mathbf{E}_\lambda)$. By Theorem 3.1 it follows that $\mathbf{x}_\lambda = \mathbf{Q}^* \mathbf{y}^\lambda$ where the i-th component of $\mathbf{y}^\lambda$, y_i^λ, is the solution to

$$\min_{y_i} \left\{ \frac{|\alpha_i y_i - \hat{b}_i|^2}{1 + |y_i|^2} + \rho\lambda |l_i|^2 |y_i|^2 \right\}.$$

If $l_i = 0$ then $y_i^\lambda = \frac{\hat{b}_i}{\alpha_i}$ ($\alpha_i \neq 0$ for every i as an eigenvalue of a nonsingular matrix). Otherwise,

$$\rho\lambda |l_i|^2 |y_i^\lambda|^2 \leq \frac{|\alpha_i y_i^\lambda - \hat{b}_i|^2}{1 + |y_i^\lambda|^2} + \rho\lambda |l_i|^2 |y_i^\lambda|^2 \leq \frac{|\alpha_i 0 - \hat{b}_i|^2}{1 + 0^2} + \rho\lambda |l_i|^2 0^2 = |\hat{b}_i|^2,$$

so that $|y_i^\lambda|^2 \leq \frac{|\hat{b}_i|^2}{\rho\lambda |l_i|^2}$. Consequently, for every $\lambda \in [\lambda_1, \lambda_2]$

$$|y_i^\lambda|^2 \leq \begin{cases} \left|\frac{\hat{b}_i}{\alpha_i}\right|^2 & l_i = 0, \\ \frac{|\hat{b}_i|^2}{\rho\lambda_1 |l_i|^2} & l_i \neq 0. \end{cases}$$

Hence, $\mathbf{y}^\lambda$ is bounded for every $\lambda \in [\lambda_1, \lambda_2]$ showing that $\mathbf{x}_\lambda = \mathbf{Q}^* \mathbf{y}^\lambda$ is also bounded over $[\lambda_1, \lambda_2]$, that is, there exists $\beta > 0$ for which $\|\mathbf{x}_\lambda\| \leq \beta, \lambda \in [\lambda_1, \lambda_2]$. Moreover, by the relation between the optimal $\mathbf{E}$ and the optimal $\mathbf{x}$ given by (3.3) and (3.4), it follows that $\mathbf{E}_\lambda$ is also bounded over $[\lambda_1, \lambda_2]$, namely, there exists $\gamma > 0$ for which $\|\mathbf{E}_\lambda\| \leq \gamma$. The dual objective function can thus be written as

$$q(\lambda) = \min\{s(\mathbf{x}, \mathbf{E}) + \lambda t(\mathbf{x}, \mathbf{E}) : (\mathbf{x}, \mathbf{E}) \in \tilde{Y}\},$$

where
$$\tilde{Y} = \{(\mathbf{x}, \mathbf{E}) : \mathbf{x} \in \mathbb{F}^n, \mathbf{E} \in \mathcal{L}_\mathbf{Q}, \|\mathbf{x}\| \leq \beta, \|\mathbf{E}\| \leq \gamma\}$$

is a compact set. Therefore, by Theorem 6.1, q is differentiable over (λ_1, λ_2) and in particular at $\tilde{\lambda}$ and $q'(\tilde{\lambda}) = t(\mathbf{x}_{\tilde{\lambda}}, \mathbf{E}_{\tilde{\lambda}}) = \|\mathbf{L}\mathbf{x}_{\tilde{\lambda}}\|^2 - \alpha$.

(iii). Since $\mathbf{E} = \mathbf{0}, \mathbf{x} = \mathbf{0}$ is feasible for (6.3) we obtain that

$$q(\lambda) \leq \|\mathbf{b}\|^2 - \lambda\alpha,$$

establishing that $q(\lambda) \to -\infty$ as $\lambda \to \infty$. $\square$

We will now show that despite the nonconvexity of the CSTLS problem, strong duality holds.

Theorem 6.2 (Strong duality for CSTLS) *Let $\lambda^* > 0$ be a maximizer of (6.2). Then $q(\lambda^*)$ is equal to the optimal value of the primal problem (6.1) and $(\mathbf{x}_{\lambda^*}, \mathbf{E}_{\lambda^*})$ is the optimal solution of (6.1).*

Proof: Since $\lambda^* > 0$ is the optimal solution of (6.2) and q is differentiable by Lemma 4.2 (ii), we have $\|\mathbf{L}\mathbf{x}_{\lambda^*}\|^2 - \alpha = q'(\lambda^*) = 0$. Therefore, $\mathbf{x}_{\lambda^*}$ is a feasible solution of the primal problem (6.1) and

$$q(\lambda^*) = s(\mathbf{x}_{\lambda^*}, \mathbf{E}_{\lambda^*}) + \lambda^*(\|\mathbf{L}\mathbf{x}_{\lambda^*}\|^2 - \alpha) = s(\mathbf{x}_{\lambda^*}, \mathbf{E}_{\lambda^*}),$$

which, from basic duality theory, implies that $\lambda^*, (\mathbf{x}_{\lambda^*}, \mathbf{E}_{\lambda^*})$ are the dual and primal optimal solutions. $\square$

The optimal λ^* is a root of the nondecreasing function $q'(\lambda)$ and can thus be found via a simple bisection procedure.

7 Implementation and Numerical Results

7.1 Implementation

The core of the numerical method for solving the RSTLS problem is the solution of n single-variable problems of the form (4.1). Since the number of these 1D problems might be huge (for example, for a two-dimensional 1024×1024 image, there are more than one million problems), it is imperative to find the *global* solution of each of them. The method will produce an erroneous solution even if one of the 1D problems is not solved correctly.

From numerical considerations the algorithm is split into two phases. In the first phase, we find the optimal solution of (4.2) up to a moderate tolerance ε (in our experiments $\varepsilon = 10^{-4}$). That is, the output of the first phase is an interval $[\ell, u]$ with $u - \ell < \varepsilon$ in which the optimal solution of (4.2) is guaranteed to reside. The goal of the first phase is to find a "small enough" interval in which the global solution is guaranteed to reside. Since in the course of the change of variables $x = \operatorname{sgn}(\bar{a}b)\sqrt{y}$ the accuracy of the solution might be reduced from ε to $\sqrt{\varepsilon}$, a second phase is invoked in which we seek the global minimizer x^* of the problem

$$\min_x \left\{ \frac{|a|^2 x^2 - 2|ab|x + |b|^2}{1 + x^2} + |c|^2 x^2 \right\} \tag{7.1}$$

in the interval $[\sqrt{\ell}, \sqrt{u}]$ up to a tolerance ε^2. The interval $[\sqrt{\ell}, \sqrt{u}]$ is small enough so that for all practical purposes the function in (7.1) is unimodal over $[\sqrt{\ell}, \sqrt{u}]$ and the global optimal solution given by $\operatorname{sign}(\bar{a}b)x^*$ is obtained. A detailed description of the algorithm follows.

Algorithm SOLVE1D(a, b, c)

---

input	$a, b, c \in \mathbb{C}$
output	x - an optimal solution of (4.1).
comments	1. it is assumed that a and c are not both zero
	2. The functions f_1 and f_2 called in the solver are given by

$$f_1(x; a, b, c) = \frac{(|a|\sqrt{x} - |b|)^2}{1+x} + |c|^2 x.$$
$$f_2(x; a, b, c) = \frac{(|a|x - |b|)^2}{1+x^2} + |c|^2 x^2.$$

If c is equal to zero up to some tolerance than the output of the algorithm is b/a, otherwise the upper bound is chosen.

if $c < 10^{-8}$

  $x = \frac{b}{a}$

  stop

else

  $u = \left|\frac{b}{c}\right|^2$

end if

$\ell = 0$

$s = \operatorname{sgn}(\bar{a}b)$

Phase I. Activating an unimodal solver on the function f_1

while $(u - \ell) > \varepsilon$

    $x^- = \frac{2}{3}\ell + \frac{1}{3}u$

    $x^+ = \frac{1}{3}\ell + \frac{2}{3}u$

    $f^+ = f_1(x^+; a, b, c)$

    $f^- = f_1(x^-; a, b, c)$

    if $f^- \leq f^+$

      $u = x^+$

    else

      $\ell = x^-$

    end if

end while

Updating the lower and upper bounds.

$\ell = \sqrt{\ell}$

$u = \sqrt{u}$

Phase II. Activating an unimodal solver on the function f_2.

while $(u - \ell) > \varepsilon^2$

    $x^- = \frac{2}{3}\ell + \frac{1}{3}u$

    $x^+ = \frac{1}{3}\ell + \frac{2}{3}u$

    $f^+ = f_2(x^+; a, b, c)$

    $f^- = f_2(x^-; a, b, c)$

    if $f^- \leq f^+$

      $u = x^+$

    else

      $\ell = x^-$

    end if

end while
$x = s\frac{x^+ + x^-}{2}$
stop

We note that in the MATLAB implementation the minimization of the n 1D problems is done simultaneously using MATLAB's vector operations. For the exact implementation please see the (small) RSTLS MATLAB package available at

> `http://iew3.technion.ac.il/~becka/papers/rstls_package.zip`

Given the 1D solver, the solution of the RSTLS problem (2.1) is obtained via the following procedure.

Algorithm RSTLS $(\mathbf{Q}, \mathbf{A}, \mathbf{L}, \rho)$

---

input	$\mathbf{Q} \in \mathbb{F}^{n \times n}$ - a unitary matrix.
	$\mathbf{A}, \mathbf{L} \in \mathcal{L}_{\mathbf{Q}}, \mathbf{b} \in \mathbb{F}^n$.
	$\rho \in \mathbb{R}_{++}$.
output	The $\mathbf{x}$-part of the optimal solution of (2.1).

Step 1. $\hat{\mathbf{b}} = \mathbf{Q}\mathbf{b}$
Step 2. Compute the eigenvalues vectors $\boldsymbol{\alpha}, \mathbf{l}$ of $\mathbf{A}$ and $\mathbf{L}$ defined
　　　　　by the relations (3.1).
Step 3. For each $i = 1, \ldots, n$ call algorithm SOLVE1D with input $\alpha_i, \hat{b}_i, c_i$
　　　　　and obtain an output $\hat{x}_i$.
Step 4. $\mathbf{x} = \mathbf{Q}^* \hat{\mathbf{x}}$ where $\hat{\mathbf{x}} = (\hat{x}_i)_{i=1}^n$.

Based on the RSTLS algorithm, the constrained version, problem (CSTLS) is solved via a simple bisection algorithm applied to $q'(\lambda)$ where q is the dual function defined by (6.3). The bisection is over the logarithm of base 10 of the dual variable λ.

Algorithm CSTLS$(\mathbf{Q}, \mathbf{A}, \mathbf{L}, \alpha)$

---

input	$\mathbf{Q} \in \mathbb{F}^{n \times n}$ - a unitary matrix.
	$\mathbf{A}, \mathbf{L} \in \mathcal{L}_{\mathbf{Q}}, \mathbf{b} \in \mathbb{F}^n$.
	$\alpha \in \mathbb{R}_{++}$.
output	The $\mathbf{x}$-part of the optimal solution of (6.1).

Step 1. $u = 2, \ell = -4$.
Step 2. while $(u - \ell) > 0.1$
　　　　　$h = \frac{u + \ell}{2}$

call Algorithm RSTLS with input $\mathbf{Q}, \mathbf{A}, \mathbf{L}, 10^h$ and obtain an output $\tilde{\mathbf{x}}$
    if $\|\mathbf{L}\tilde{\mathbf{x}}\|^2 < \alpha$
      $u = h$
    else
      $\ell = h$
    end if
  end while
Step 3. $\mathbf{x} = \tilde{\mathbf{x}}$.

7.2 A Numerical Example

To demonstrate our approach we consider an image deblurring example. We start with the 512×512 Lena gray image (top left image of Figure 3), blur it with a Gaussian PSF of dimension 9×9 with standard deviation 6 implemented in the command psfGauss([9,9],6) from [16]. We assume that the blurring is not exactly known and that the observed PSF is a Gaussian PSF of dimension 9×9 with standard deviation 8. We then cut the margins by 20 rows and columns resulting with a 492×492 and add a Gaussian white noise with standard deviation 10^{-3} (top right image of Figure 3). Assuming reflexive boundary conditions, the poor naive solution construction (i.e., $\mathbf{A}^{-1}\mathbf{b}$) is given in the left middle image of Figure 3. This poor quality of the naive solution is not surprising since the problem is extremely ill-conditioned. In our experiments, the regularization matrix $\mathbf{L}$ represents a differential operator corresponding to the PSF

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}.$$

The constrained least squares solution, that is the solution of the problem

$$\min\{\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 : \|\mathbf{L}\mathbf{x}\|^2 \leq \alpha\},$$

is presented in the right middle image. The CSTLS constructions under periodic and reflexive boundary conditions are the left and right bottom images respectively. The parameter α is chosen as $1.2\|\mathbf{L}\mathbf{x}_{\text{true}}\|^2$. Clearly, the best reconstruction is provided by the CSTLS algorithm with reflexive boundary conditions. The artifacts in the CSTLS reconstruction with periodic boundary conditions are much more prominent. The relative error of the CSTLS reconstruction with reflexive boundary conditions, $\frac{\|\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{CSTLS-R}}\|}{\|\mathbf{x}_{\text{true}}\|}$, is 0.0961, while for periodic boundary conditions, the relative error is 0.1393. The constrained least squares solution gave the worst relative error: 0.15.
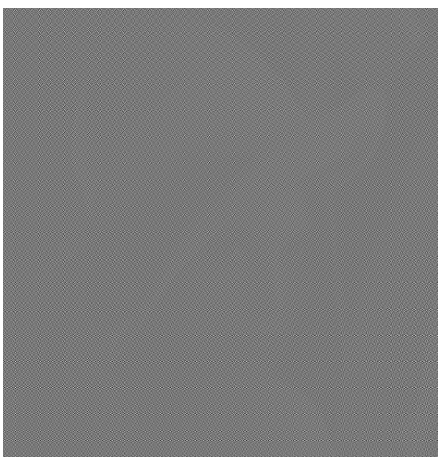
original

blurred and noisy

naive solution

constrained LS

CSTLS (periodic)

CSTLS (reflexive)

Figure 3: Deblurring of Lena

References

[1] T. J. Abatzoglou, J. M. Mendel, and G. A. Harada. The constrained total least squares technique and its applications to harmonic supperresolution. *IEEE Trans. Signal Processing*, 39(5):1070–1087, May 1991.

[2] H. Andrews and B. Hunt. *Digital Image Restoration.* Englewood Cliffs, NJ: Prentice Hall, Inc., 1977.

[3] A. Beck and A. Ben-Tal. A global solution for the structured total least squares problem with block circulant matrices. *SIAM J. Matrix Anal. Appl.*, 27(1):238–255 (electronic), 2005.

[4] A. Beck and A. Ben-Tal. On the solution of the Tikhonov regularization of the total least squares problem. *SIAM J. Optim.*, 17(1):98–118 (electronic), 2006.

[5] A. Beck, A. Ben-Tal, and M. Teboulle. Finding a global optimal solution for a quadratically constrained fractional quadratic problem with applications to the regularized total least squares. *SIAM J. Matrix Anal. Appl.*, 28(2):425–445 (electronic), 2006.

[6] A. Beck and M. Teboulle. A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid. To appear.

[7] D. P. Bertsekas. *Nonlinear Programming.* Belmont MA: Athena Scientific, second edition, 1999.

[8] D. F. Elliott and K. R. Rao. *Fast transforms.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982. Algorithms, analyses, applications.

[9] R. D. Fierro, G. H. Golub, P. C. Hansen, and D. P. O'Leary. Regularization by truncated total least squares. *SIAM J. Sci. Comp.*, 18(1):1223–1241, 1997.

[10] H. Fu and J. Barlow. A regularized structured total least squares algorithm for high-resolution image reconstruction. *Linear Algebra Appl.*, 391:75–98, 2004.

[11] G. H. Golub, P. C. Hansen, and D. P. O'Leary. Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl.*, 21(2):185–194, 1999.

[12] G. H. Golub and C. F. Van Loan. An analysis of the total least-squares problem. *SIAM J. Numer. Anal.*, 17(6):883–893, Dec. 1980.

[13] P. C. Hansen. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Stat. Comput.*, 14:1487–1503, 1993.

[14] P. C. Hansen. Regularization tools, a matlab package for analysis of discrete regularization problems. *Numerical Algorithms*, 6:1–35, 1994.

[15] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed problems: Numerical Aspects of Linear Inversion.* Philadelphia, PA: SIAM, 1998.

[16] P. C. Hansen, J. G. Nagy, and D. P. O'Leary. *Deblurring images*, volume 3 of *Fundamentals of Algorithms*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Matrices, spectra, and filtering.

[17] S. Van Huffel and J. Vandewalle. Analysis and properties of the generalized total least squares problem $AX \approx B$ when some or all columns in A are subject to error. *SIAM J. Matrix Anal. Appl.*, 10(3):294–315, 1989.

[18] S. Van Huffel and J. Vandewalle. *The Total Least-Squares Problem: Computational Aspects and Analysis*, volume 9 of *Frontier in Applied Mathematics*. Philadelphia, PA: SIAM, 1991.

[19] S. Van Huffel and H. Y. Zha. The restricted total least squares problem: formulation, algorithm, and properties. *SIAM J. Matrix Anal. Appl.*, 12(2):292–309, 1991.

[20] P. Lemmerling and S. Van Huffel. Analysis of the structured total least squares problem for Hankel/Toeplitz matrices. *Numerical Algorithms*, 27:89–114, 2001.

[21] N. Mastronardi, P. Lemmerling, and S. Van Huffel. Fast structured total least squares algorithm for solving the basic deconvolution problem. *SIAM J. Matrix Anal. Appl.*, 22(2):533–553, 2000.

[22] N. Mastronardi, P. Lemmerling, A. Kalsi, D. P. O'Leary, and S. Van Huffel. Implementation of the regularized structured total least squares algorithms for blind image deblurring. *Linear Algebra Appl.*, 391:203–221, 2004.

[23] V. Z. Mesarovic, N. P. Galatsanos, and A. K. Katsaggelos. Regularized constrained total least squares image restoration. *IEEE Trans. Image Processing*, 4(8):1096–1108, 1995.

[24] B. De Moor. Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Trans. Signal Processing*, 42(11):3104–3113, Nov. 1994.

[25] J. J. Moré. Generalizations of the trust region subproblem. *Optim. Methods Software*, 2:189–209, 1993.

[26] M. K. Ng, , R. J. Plemmons, and F. Pimentel. A new approach to constrained total least squares image restoration. *Linear Algebra Appl.*, 316(1-3):237–258, 2000. Conference Celebrating the 60th Birthday of Robert J. Plemmons (Winston-Salem, NC, 1999).

[27] M. K. Ng, H. C. Raymond H., and W. C. Tang. A fast algorithm for deblurring models with Neumann boundary conditions. *SIAM J. Sci. Comput.*, 21(3):851–866 (electronic), 1999.

[28] A. Pruessner and D. P. O'Leary. Blind deconvolution using a regularized structured total least norm algorithm. *SIAM J. Matrix Anal. Appl.*, 24(4):1018–1037 (electronic), 2003.

[29] R. A. Renaut and H. Guo. Efficient algorithms for solution of regularized total least squares. *SIAM J. Matrix Anal. Appl.*, 26(2):457–476, 2005.

[30] J. B. Rosen, H. Park, and J. Glick. Total least norm formulation and solution for structured problems. *Siam J. Matrix Anal. Appl.*, 17(1):110–126, 1996.

[31] D. Sima, S. Van Huffel, and G. H. Golub. Regularized total least squares based on quadratic eigenvalue prblem solvers. *BIT Numerical Mathematics*, 44(4):793–812, 2004.

[32] H. Y. Zha. The restricted singular value decomposition of matrix triplets. *SIAM J. Matrix Anal. Appl.*, 12(1):172–194, 1991.