Learning-based Estimation of Parameters for Spectral Windowed Regularization using Multiple Data Sets

by

Michael J. Byrne

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2023 by the
Graduate Supervisory Committee:

Rosemary Renaut, Chair
Douglas Cochran
Malena Español
Zdzislaw Jackiewicz
Rodrigo Platte

ARIZONA STATE UNIVERSITY

May 2023

ABSTRACT

During the inversion of discrete linear systems, noise in data can be amplified and result in meaningless solutions. To combat this effect, characteristics of solutions that are considered desirable are mathematically implemented during inversion. This is a process called regularization. The influence of the provided prior information is controlled by the introduction of non-negative regularization parameter(s). Many methods are available for both the selection of appropriate regularization parameters and the inversion of the discrete linear system. Generally, for a single problem there is just one regularization parameter. Here, a learning approach is considered to identify a single regularization parameter based on the use of multiple data sets described by a linear system with a common model matrix. The situation with multiple regularization parameters that weight different spectral components of the solution is considered as well. To obtain these multiple parameters, standard methods are modified for identifying the optimal regularization parameters. Modifications of the unbiased predictive risk estimation, generalized cross validation, and the discrepancy principle are derived for finding spectral windowing regularization parameters. These estimators are extended for finding the regularization parameters when multiple data sets with common system matrices are available. Statistical analysis of these estimators is conducted for real and complex transformations of data. It is demonstrated that spectral windowing regularization parameters can be learned from these new estimators applied for multiple data and with multiple windows. Numerical experiments evaluating these new methods demonstrate that these modified methods, which do not require the use of true data for learning regularization parameters, are effective and efficient, and perform comparably to a supervised learning method based on estimating the parameters using true data. The theoretical developments are validated for one and two dimensional image deblurring. It is verified that the obtained es-

timates of spectral windowing regularization parameters can be used effectively on validation data sets that are separate from the training data, and do not require known data.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

We consider solutions of linear problems described by

$$A\mathbf{x} \approx \mathbf{d}, \tag{1.0.1}$$

where $A \in \mathbb{R}^{m\times n}$ with $m \geq n$ and $\mathbf{d}$ is known, but represents $\mathbf{d} = \mathbf{b} + \boldsymbol{\eta}$, with $\mathbf{b} = A\mathbf{x}_{\text{true}}$ and $\boldsymbol{\eta}$ being a realization of a random vector representing errors/noise in $\mathbf{d}$. Consider the multivariate normal distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^m$ and positive semidefinite $\Sigma \in \mathbb{R}^{m\times m}$, which is denoted $\mathcal{N}(\boldsymbol{\mu}, \Sigma)$. If $\mathbf{X}$ is a random vector with this distribution, which is denoted $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, then the probability density function of $\mathbf{X}$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^m \det(\Sigma)}} \exp\left(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathsf{T}}\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right).$$

In this work, we primarily consider $\boldsymbol{\eta}$ being a realization of a random vector $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$, in which case $\boldsymbol{\eta}$ is considered a realization of white noise. Even for invertible square matrices A, direct matrix inversion of eq. (1.0.1) when A is ill-conditioned is not recommended due to the noise in the data. Rather, regularization is imposed in which desired characteristics of a solution are described mathematically and incorporated into the problem formulation, with the aim to produce a more well-posed problem.

The generalized Tikhonov regularized solution, $\mathbf{x}(\alpha)$, Tikhonov (1963), is

$$\mathbf{x}(\alpha) = \underset{\mathbf{x}\in\mathbb{R}^n}{\arg\min}\left\{\|A\mathbf{x}-\mathbf{d}\|_2^2 + \alpha^2\|L\mathbf{x}\|_2^2\right\}, \quad \alpha > 0,\ L \in \mathbb{R}^{q\times n}. \tag{1.0.2}$$

Here, the scalar $\alpha > 0$ is a regularization parameter and L is a $q \times n$ matrix representation of a linear operator. The term $\|L\mathbf{x}\|_2^2$ is an example of a penalty function

Vogel (2002), and L is called the penalty matrix. If $L = I_n$, the $n \times n$ identity matrix, then regularization via eq. (1.0.2) is called standard or zeroth-order Tikhonov regularization Aster *et al.* (2013). Other standard choices of L include approximations of first and second order derivative operators Ng *et al.* (1999); Strang (1999); Vogel (2002).

The quality of $\mathbf{x}(\alpha)$ depends on the choice of both α and L. There are a number of methods for selecting α when L has been fixed. The Morozov discrepancy principle Morozov (1966) assumes that the variance of the noise in the data is known and finds α as the root of a function. The unbiased predictive risk estimator, Mallows (1973), which also requires knowledge of the variance of the noise distribution of the data, yields the regularization parameter as the minimizer of a function. The method of generalized cross validation, Wahba (1977, 1990), which does not require the noise distribution be known, also yields the regularization parameter as a minimizer of a function. Some methods do not solve minimization or root-finding problems; for example, the regularization parameter found using the L-curve method occurs at the point of maximum curvature of a function Hansen (1992); Hansen and O'Leary (1993). As noted, many of these techniques, including the unbiased predictive risk, Morozov's discrepancy principle, and the generalized cross validation methods, have statistical foundations James *et al.* (2013). Throughout this work the 2-norm is utilized, as in eq. (1.0.2), though other choices of p-norms are possible but not considered here.

There has been a considerable amount of research on finding sets of regularization parameters for a pre-selected set of penalty matrices, a process called windowed regularization Brezinski *et al.* (2003); Chung and Español (2017); Gazzola and Novati (2013); Lu and Pereverzev (2011); Wood (2002). Approaches to windowed regularization using versions of the L-curve and discrepancy principle methods can be found in Belge *et al.* (2002) and Wang (2012), respectively. A windowed generalized

cross validation method was also considered in Modarresi and Golub (2007a,b). Windowing, either in the data domain or the frequency domain, can also be applied to determine multiple regularization parameters, and windowing wavelet coefficients was considered in Easley *et al.* (2014); Stephanakis and Kollias (1998). Examples of windowed regularization in other frequency domains, such as those generated by discrete trigonometric transforms or the singular value decomposition, have been presented in Chung *et al.* (2011b, 2015) and Kalke and Siltanen (2013). There is also recent work on learning seminorms as regularization operators Holler and Kunisch (2022).

Techniques for utilizing and analyzing multiple data sets permeate a multitude of scientific fields as diverse as geoscience Bergen *et al.* (2019); Zobitz *et al.* (2020), and the detection of cancers Sidey-Gibbons and Sidey-Gibbons (2019). A comprehensive overview of data-driven approaches for finding solutions of inverse problems can be found in Arridge *et al.* (2019), while specific examples of applying multiple data sets for the solution of inverse problems include Afkham *et al.* (2021); Chung *et al.* (2011a); Chung and Español (2017); Haber and Tenorio (2003); Kunisch and Pock (2013); Taroudaki and O'Leary (2015); Vito *et al.* (2005).

1.1 Overview

The functions pertaining to new windowed multidata methods are presented in Chapter 5 and constitute the majority of the main contributions of this document. The corresponding numerical results are contained in Chapter 6. However, the mathematical background must be established not only for the parameter methods but also for the underlying theory/tools that are utilized in the development of such methods. We discuss two primary approaches to regularization: a learning approach involving multiple data sets and the application of spectral windowing. Before these approaches can be discussed in detail, however, a background on regularization and matrix the-

ory is necessary and provided in Chapter 2. The specific type of problems in this work that are solved using both the learning and windowing approaches to general Tikhonov regularization involve the deblurring of images. A variety of considerations are associated with deblurring, both in the continuous and discrete settings: the process by which the blurring occurs, the amount of blur, and how blurring applied at the boundaries of the solutions is handled. In the discrete setting, consideration must also be given to noise that may be present in the available data, particularly the distribution of the noise and how the distribution changes under various transformations. While the generalized singular value decomposition is a powerful technique for analyzing the conditioning and regularization of a discrete problem (see Section 2.3), other methods must be used for efficient obtainment of meaningful solutions. All of these considerations are important for the numerical implementation of the regularization process and will be discussed in Chapter 3. Section 3.5 deviates somewhat from the rest of Chapter 3 in that the focus shifts to the statistics of noise under various transformations; the content of Section 3.5 was originally formulated in an effort to better understand the parameter methods as applied to downsampled data. The standard parameter methods, meaning the methods used for selecting a scalar regularization parameter for a single data set, are discussed in Chapter 4. A method using true solutions to determine parameters, referred to as the mean squared error method and used as a baseline method from which the other methods are compared, is discussed first. More discussion is then allocated to the unbiased predictive risk estimator method than to the generalized cross validation method and discrepancy principle. The reason for this is twofold: the generalized cross validation method for spectral windows is fundamentally distinct from the standard version and thus much of the discussion regarding the generalized cross validation method is presented in Chapter 5. The other reason is that the discrepancy principle method, when com-

pared to the unbiased predictive risk estimator method, is considerably more simple; in fact, the function associated with the unbiased predictive risk estimator method is the same as the function associated with discrepancy principle method, minus one term. Some variations of the standard methods, including summation representations involving spectral data components, are also contained in Chapter 4. As mentioned previously, the majority of the main contributions are presented in Chapter 5. The progression of the presentation proceeds by first forming the relevant terms of the function in the context of spectral windowing and their use with multiple data sets. Chapter 6 contains the 1D and 2D numerical test problems and corresponding results from which the new methods are evaluated. Lastly, Chapter 7 introduces the concept of downsampling in the context of selecting regularization parameters and some results are presented. Chapter 7 is considerably shorter than Chapter 5 and Chapter 6, serving mostly to present preliminary results that could serve as a starting point for future work. Conclusions and closing remarks are finally provided in Chapter 8.

1.2 Main Contributions

A main contribution of this work is demonstrating and validating how to use standard techniques, in particular the unbiased predictive risk estimator, Morozov's discrepancy principle, and the generalized cross validation methods, to learn regularization parameters weighting spectral components given multiple data sets, without any knowledge of the true solution. These are chosen as representative methods that either require (unbiased predictive risk estimator), or do not require (generalized cross validation), prior knowledge of the statistics of the noise in the data. Morozov's discrepancy principle falls into the former of these categories. In addition to developing these modified parameter selection methods, results are presented regarding their relationship(s) with the original method(s) (i.e. the non-modified methods).

Most significantly, it is demonstrated through this work how multiple data sets can be used in conjunction with windowed regularization using the windowing formulation introduced in Chung *et al.* (2011b). Windowed versions of the generalized cross validation method are presented in Chung *et al.* (2011b); Modarresi and Golub (2007b); here these formulations are extended for equivalent versions of the unbiased predictive risk estimator and discrepancy principle methods. Numerical results for the restoration of 2D signals demonstrate that these new windowed regularization parameter estimators can be used for multiple data sets and that their performance competes with a learning approach in which the training stage requires the knowledge of true data, which is not the case here. Furthermore, parameters that have been obtained from one set of training images can be used on a separate set of validation (testing) images distinct from the original set, provided that the signal-to-noise ratios are close and the data is collected using similar measurement procedures.

1.3 Summary of Notation

Due to the consideration of multiple data sets as well as multiple parameters for the purpose of regularizing these data sets, the notation in this document is extensive. Tables containing the notation used throughout this work are found in Appendix A; these tables also include references to pertinent equations/results where appropriate.

Chapter 2

MATRIX FACTORIZATIONS AND DECOMPOSITIONS

We begin with the motivation for regularization, namely the concepts of ill-posed and ill-conditioned problems. The singular value decomposition (SVD) is introduced in Section 2.1 to illustrate these concepts, and standard Tikhonov regularization is introduced in Section 2.2. The generalized singular value decomposition (GSVD) is discussed in Section 2.3, which is used for generalized Tikhonov regularization as the focus of this work.

The concept of a well-posed problem is due to Hadamard (1904). Given an operator $A : \mathcal{H}_1 \to \mathcal{H}_2$, where $\mathcal{H}_1$ and $\mathcal{H}_2$ are Hilbert spaces, the equation $Af = g$ is said to be well-posed if

(i) for each $g \in \mathcal{H}_2$ there exists a solution $f \in \mathcal{H}_1$ to $Af = g$,

(ii) the solution f is unique, and

(iii) if $Af_* = g_*$ and $Af = g$, then $\|f - f_*\|_{\mathcal{H}_1} \to 0$ whenever $\|g - g_*\|_{\mathcal{H}_2} \to 0$,

where $\|\cdot\|_{\mathcal{H}}$ denotes the norm induced by the inner product on $\mathcal{H}$. In order, condition (i) requires that a solution f exists, condition (ii) requires f be unique, and condition (iii) requires f to be stable under perturbations in g. If one of these conditions is not met, then $Af = g$ is said to be an ill-posed problem. The discrete problem $A\mathbf{x} = \mathbf{b}$ (where $\mathcal{H}_1 = \mathbb{R}^n$ and $\mathcal{H}_2 = \mathbb{R}^m$) can be ill-posed under a number of circumstances: singularity of A violates conditions (i) and (ii), and a poor condition number of non-singular A violates condition (iii). If $m \neq n$, certainly condition (i) is violated. As suggested by the term "condition number" of A, denoted $\kappa(A)$ and defined as

$\kappa(A) = \|A\|\|A^{-1}\|$ for any consistent matrix norm $\|\cdot\|$ on $\mathbb{R}^{n\times n}$ Leon (2010), discrete systems that are ill-posed are often considered ill-conditioned.

Even for linear systems that are well-posed, obtaining a solution via direct matrix inversion (assuming the practicality of forming the inverse A^{-1}) can be problematic if the data available is flawed in some way. In other words, if the data available is $\mathbf{d} \approx \mathbf{b}$, then forming $A^{-1}\mathbf{d}$ can produce a "solution" that is unlike $\mathbf{x}$ in any meaningful way. Specifically, if $\mathbf{d} = \mathbf{b} + \boldsymbol{\eta}$, then $A^{-1}\mathbf{d} = \mathbf{x} + A^{-1}\boldsymbol{\eta}$ and so the obtained solution differs from $\mathbf{x}$ by $A^{-1}\boldsymbol{\eta}$. The condition number of A is helpful for quantifying the ill-posedness of a discrete problem. A lower bound on $\kappa(A)$ is readily obtained by noting $\|A\|\|A^{-1}\| \geq \|AA^{-1}\| = \|I_m\| = 1$. There is no upper bound on $\kappa(A)$: for singular A it is the convention that $\kappa(A) = \infty$, in which case the associated discrete problem is ill-posed. A matrix with a condition number close to 1 is considered well-conditioned, while a matrix with a large condition number is considered ill-conditioned. The difficulty in working with ill-conditioned matrices is that errors in the data can be amplified to produce erroneous solutions, i.e. $\|A^{-1}\boldsymbol{\eta}\|$ can be excessively large. Finding a solution of the system $A\mathbf{x} \approx \mathbf{d}$ is considered an ill-conditioned problem if A is an ill-conditioned matrix.

To combat the difficulties associated with an ill-posed problem, the process of regularization essentially introduces additional information that is sufficient to make the problem more tractable. Even if a problem is well-posed yet numerically ill-conditioned, regularization processes are useful in constructing meaningful solutions. Perhaps the most famous means of regularization is attributed to Tikhonov (1963). Tikhonov regularization forms the foundation of the work presented here and will be introduced in the context of the singular value decomposition of the matrix A.

2.1 Singular Value Decomposition

To illustrate the discussion we introduce the integral equation

$$g(x) = (k * f)(x) =: \int_{-\infty}^{\infty} k(x-t)f(t)\ dt. \tag{2.1.1}$$

Equation (2.1.1) represents the one-dimensional continuous convolution of functions $k, f \in L_1(\mathbb{R})$, where k is called the convolution kernel. The process of determining f from known g and k is called deconvolution (there are problems where k is unknown as well, which are examples of blind deconvolution). The convolution integral eq. (2.1.1) is a specific example of the Fredholm equation of the first kind (Debnath and Mikusiński, 2005, p. 218):

$$g(x) = \int_a^b k(x,t)f(t)\ dt. \tag{2.1.2}$$

Equation (2.1.1) is obtained from eq. (2.1.2) if the kernel is spatially invariant, i.e. $k(x,t) = k(x-t)$. The properties of the integrands have a direct impact on the properties of g; for example, if $k(x-t)$ is integrable and $f(t)$ is bounded and locally integrable then $g(x)$ is a continuous function Debnath and Mikusiński (2005).

Deblurring is often described as a subclass of deconvolution-type problems in which the effect of convolution by specific kernels is the smoothing or blurring of the original function f. A Gaussian kernel is common in blurring/deblurring problems for this reason. The form of a Gaussian kernel comes from the probability density function of the Gaussian distribution,

$$k(t) = \frac{1}{\sqrt{2\pi\xi}} \exp\left(\frac{-(t-\mu)^2}{2\xi}\right), \tag{2.1.3}$$

where μ is the mean and ξ is the variance. The mean is the center of the Gaussian distribution, as well as the abscissa of the absolute maximum of the probability density function $k(t)$. The variance ξ is a measure of dispersion of the distribution; as

ξ increases, the width of the graph of $k(t)$ increases. The standard deviation $\sqrt{\xi}$ is also a measure of dispersion. The scale factor $1/\sqrt{2\pi\xi}$ ensures that $\int_{\mathbb{R}} k(t)\,dt = 1$, an essential property of a continuous probability distribution defined on the entire real line. For Gaussian kernels, however, this scale factor may be dropped since having a unitary integral is not required of kernels in general. The continuous deconvolution problem can be discretized using a quadrature method to obtain a linear system eq. (1.0.1). Depending upon the support of the continuous functions, boundary conditions are often necessary to effectively describe the continuous problem in a discrete setting. Figure 2.1 shows an example of a discretized deconvolution problem which leads to a problem of the form $A\mathbf{x} \approx \mathbf{d}$ as given in eq. (1.0.1) and where $\mathbf{d}$ is the true data corrupted by noise. An in-depth discussion of translating a continuous convolution problem to a linear system is given in Chapter 3.

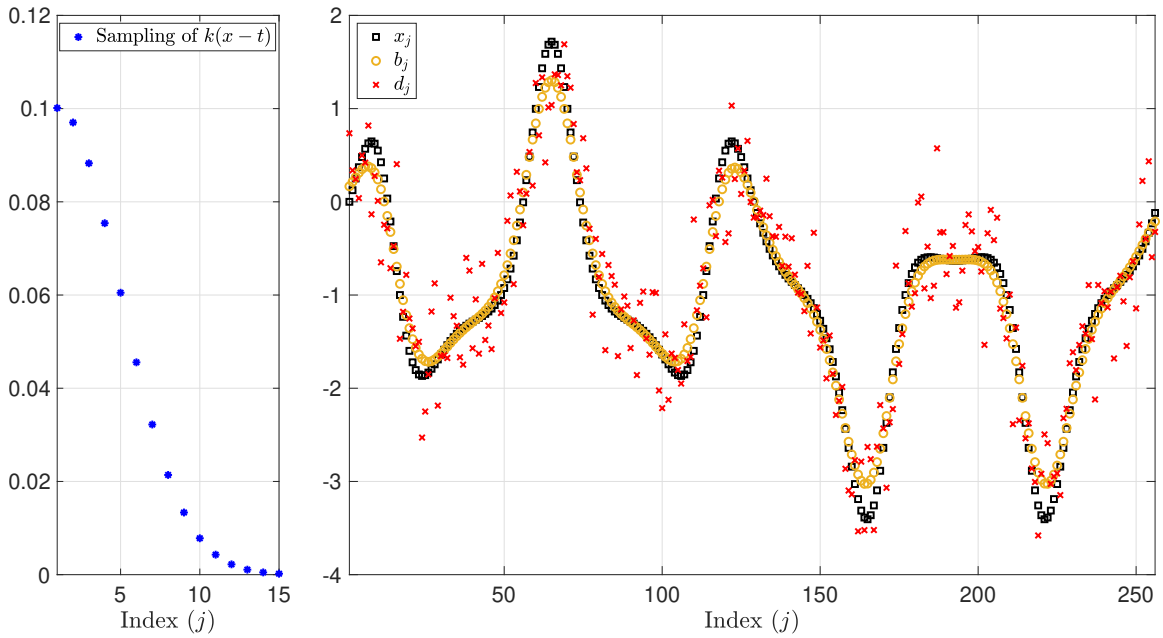


Figure 2.1: Truncated vectors as part of an example of discretizing a 1D deconvolution problem. The data vector $\mathbf{d}$ is the sum of $\mathbf{b}$ and $\boldsymbol{\eta}$, where $\boldsymbol{\eta}$ is a realization of white noise. Quadrature is used with a sampling of the Gaussian kernel eq. (2.1.3) with $\xi = 16$ to form the system matrix A. The function f is discretized to form vector $\mathbf{x}$ and $\mathbf{b}$ is formed by $\mathbf{b} = A\mathbf{x}$. The full vectors have 256 elements and are truncated for the sake of visual clarity.

The singular value decomposition (SVD) is a useful tool for analyzing the properties of the solutions of eq. (1.0.1), even if it is not used in practice for large scale problems, Golub and Van Loan (2013). Consider the full SVD of A given by $A = USV^{\mathsf{T}}$, where the $m \times m$ matrix U and the $n \times n$ matrix V are orthogonal. The columns of U and V denoted by $U_{\cdot,j}$ and $V_{\cdot,j}$, respectively, are known as the left and right singular vectors of A. S is a $m \times n$ matrix with non-zero entries s_j on the leading diagonal ordered by $s_1 \geq s_2 \geq \ldots \geq s_{r_A} > 0$. The s_j are the singular values of A, and $r_A := \text{rank}(A) \leq \min\{m, n\}$ is the number of non-zero singular values.

The condition number of a matrix A is unbounded and considered infinite for singular matrices. This is made clear through the equivalent definition of $\kappa(A)$:

$$\kappa(A) = \frac{s_{\max}}{s_{\min}} \tag{2.1.4}$$

where $s_{\max}$ and $s_{\min}$ are the largest and smallest singular values of A, respectively. The condition number of A can be arbitrarily large since $s_{\min}$ can be arbitrarily small. Since $s_{\min} = 0$ for singular A, the convention that $\kappa(A) = \infty$ in such a case is appropriate.

If A is square and invertible ($r_A = m = n$), then $A^{-1} = VS^{-1}U^{\mathsf{T}}$ and the solution $A^{-1}\mathbf{d}$ can be written as

$$A^{-1}\mathbf{d} = VS^{-1}U^{\mathsf{T}}\mathbf{d} = \sum_{j=1}^{n} \frac{(U_{\cdot,j})^{\mathsf{T}}\mathbf{d}}{s_j} V_{\cdot,j} = \sum_{j=1}^{n} \frac{\widehat{d}_j}{s_j} V_{\cdot,j}, \tag{2.1.5}$$

where $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$ is the vector of spectral coefficients of $\mathbf{d}$. When A is not square, or is square and not invertible, the pseudoinverse $A^{\dagger} = VS^{\dagger}U^{\mathsf{T}}$ (also called the Moore-Penrose inverse Penrose (1955)) can be used where $S^{\dagger} \in \mathbb{R}^{n \times m}$ is formed by reciprocating the nonzero diagonal elements of S^{T}. In such a case, forming $A^{\dagger}\mathbf{d}$ is equivalent to setting the upper limit of summation in eq. (2.1.5) to be r_A. Properties of the pseudoinverse are given in appendix B.

Generally, the summands in eq. (2.1.5) are numerically unstable for small s_j when the coefficients $\left|\widehat{d}_j\right|$ do not decay as fast as the s_j; this describes the discrete Picard condition Hansen (1998). The discrete Picard condition can be evaluated visually using a Picard plot that displays the rate of coefficient decay. Figure 2.2 is an example of a Picard plot that demonstrates the relationship between the noise, the width of a Gaussian blur, and the resulting numerical instabilities related to obtaining meaningful solutions. The discretizations used to generate Figure 2.2 are presented in Figure 2.1. Full details on Picard plots and the associated discrete Picard condition are given in Hansen (1990).

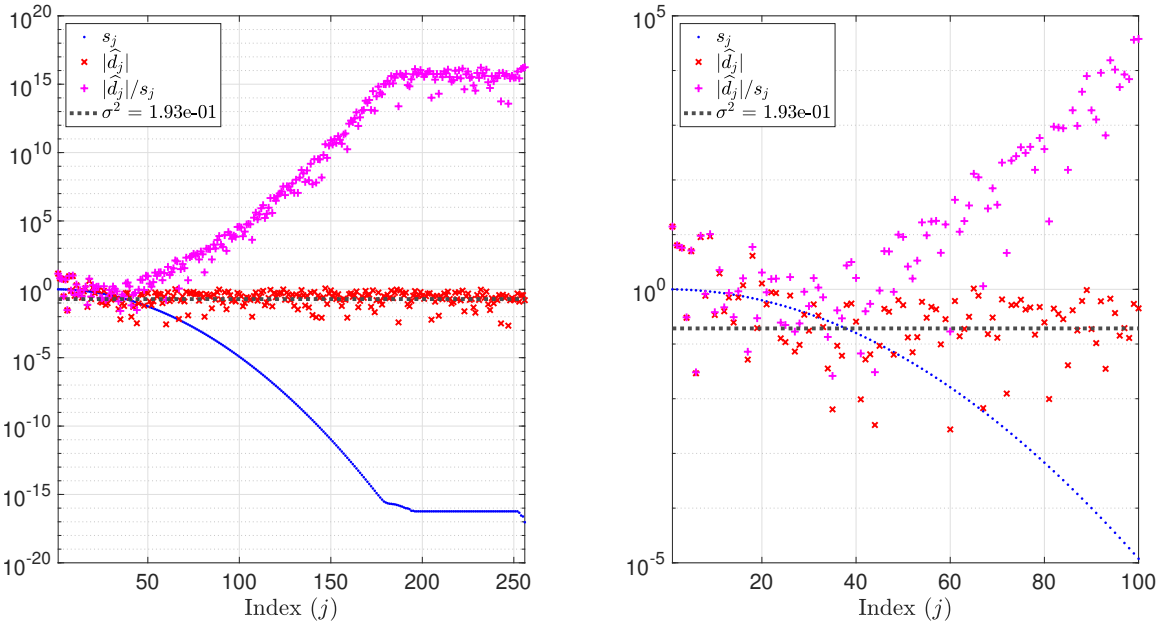


Figure 2.2: Picard plots generated from test function shown in fig. 2.1. The right plot is the same as the left, except limited to the first 100 terms. The terms $|\widehat{d}_j|/s_j|$ decrease and then increase in magnitude. This is due to the fact that the singular values s_j steadily decrease while the $|\widehat{d}_j|$ level off just below the noise variance σ^2.

For an ill-conditioned matrix A, forming eq. (2.1.5) will often result in a meaningless solution. A common approach to overcome numerical instabilities is to multiply the summands in (2.1.5) by filter functions $\phi_j(\alpha) = \phi(s_j, \alpha)$ that depend upon s_j and a non-negative regularization parameter α. By doing so, an approximate solution is

obtained:

$$\mathbf{x}(\alpha) := \sum_{j=1}^{n} \phi_j(\alpha) \frac{\widehat{d}_j}{s_j} V_{\cdot,j} = V \Phi(\alpha) S^{\dagger} \widehat{\mathbf{d}}, \tag{2.1.6}$$

where $\Phi(\alpha)$ is diagonal with entries $\Phi_{j,j}(\alpha) = \phi_j(\alpha)$. The most desired property of the filter functions is that $\phi_j(\alpha)/s_j \approx 1$ for large values of s_j and $\phi_j(\alpha)/s_j \approx 0$ for small values of s_j. Perhaps the simplest filter function is

$$\phi_j(\alpha) = \begin{cases} 1, & s_j^2 > \alpha \\ 0, & s_j^2 \leq \alpha. \end{cases}$$

Using this function in (2.1.6) gives the approximate solution

$$\mathbf{x}(\alpha) = \sum_{s_j^2 > \alpha} \frac{\widehat{d}_j}{s_j} V_{\cdot,j}$$

which corresponds to the solution obtained using a truncated SVD of the matrix A (Vogel, 2002, p. 3-5).

The use of filter functions $\phi_j(\alpha)$ is in itself a means of regularization: the desire to overcome the negative effects of small singular values in eq. (2.1.5) is mathematically accomplished by introducing $\phi_j(\alpha)$ in an effort to obtain meaningful solutions $\mathbf{x}(\alpha)$. By selecting a specific type of filter function, Tikhonov regularization Tikhonov (1963) is realized.

2.2 Tikhonov Regularization

Using the filter factors

$$\phi_j(\alpha) = \frac{s_j^2}{s_j^2 + \alpha^2}, \tag{2.2.7}$$

which are known as Tikhonov filter factors, eq. (2.1.6) is equivalent to the solution of the damped least squares problem

$$\mathbf{x}(\alpha) = \underset{\mathbf{x} \in \mathbb{R}^n}{\arg\min} \left\{ \|A\mathbf{x} - \mathbf{d}\|_2^2 + \alpha^2 \|\mathbf{x}\|_2^2 \right\}, \tag{2.2.8}$$

Aster *et al.* (2013). The solution eq. (2.2.8) is also equivalent to the ordinary least squares problem

$$\mathbf{x}(\alpha) = \min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \left\| \begin{bmatrix} A \\ \\ \alpha I_n \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{d} \\ \\ \mathbf{0} \end{bmatrix} \right\|_2^2 \right\}. \tag{2.2.9}$$

For large values of α, (2.2.7) is close to zero and for small values of α (and nonzero values of s_j), it is close to one. Another property of (2.2.7) is that for fixed nonzero s_j, it is monotone decreasing in α. Since the expression $1 - \phi_j(\alpha)$ arises a number of times in this work, we introduce

$$\psi_j(\alpha) = 1 - \phi_j(\alpha) = \frac{\alpha^2}{s_j^2 + \alpha^2}. \tag{2.2.10}$$

In contrast to eq. (2.2.7), eq. (2.2.10) is close to one for large values of α and is monotone increasing in α for fixed s_j; see fig. 2.3. Notice also that eq. (2.2.10) leads to the complementary diagonal matrix $\Psi(\alpha) = I_n - \Phi(\alpha)$.

The use of the Tikhonov filter function generates an approximate solution of eq. (2.2.9). Replacing $\|\mathbf{x}\|_2^2$ with a more general term $\|L\mathbf{x}\|_2^2$ with $L \in \mathbb{R}^{q\times n}$ produces eq. (1.0.2), which is equivalent to

$$\mathbf{x}(\alpha) = \arg\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \left\| \begin{bmatrix} A \\ \\ \alpha L \end{bmatrix} \mathbf{x} - \begin{bmatrix} \mathbf{d} \\ \\ \mathbf{0} \end{bmatrix} \right\|_2^2 \right\}. \tag{2.2.11}$$

Equation (2.2.8) follows eq. (1.0.2) by selecting L to be I_n. From the normal equations for eq. (2.2.11), we can also write

$$\mathbf{x}(\alpha) = (A^\mathsf{T}A + \alpha^2 L^\mathsf{T}L)^{-1} A^\mathsf{T}\mathbf{d} = A^\sharp(\alpha)\mathbf{d}, \tag{2.2.12}$$

where $A^\sharp(\alpha)$ is called the generalized inverse matrix Hansen (1998). The notation $A^\sharp(\alpha)$ is chosen to indicate the dependence upon α. If the matrix L in eq. (1.0.2) is

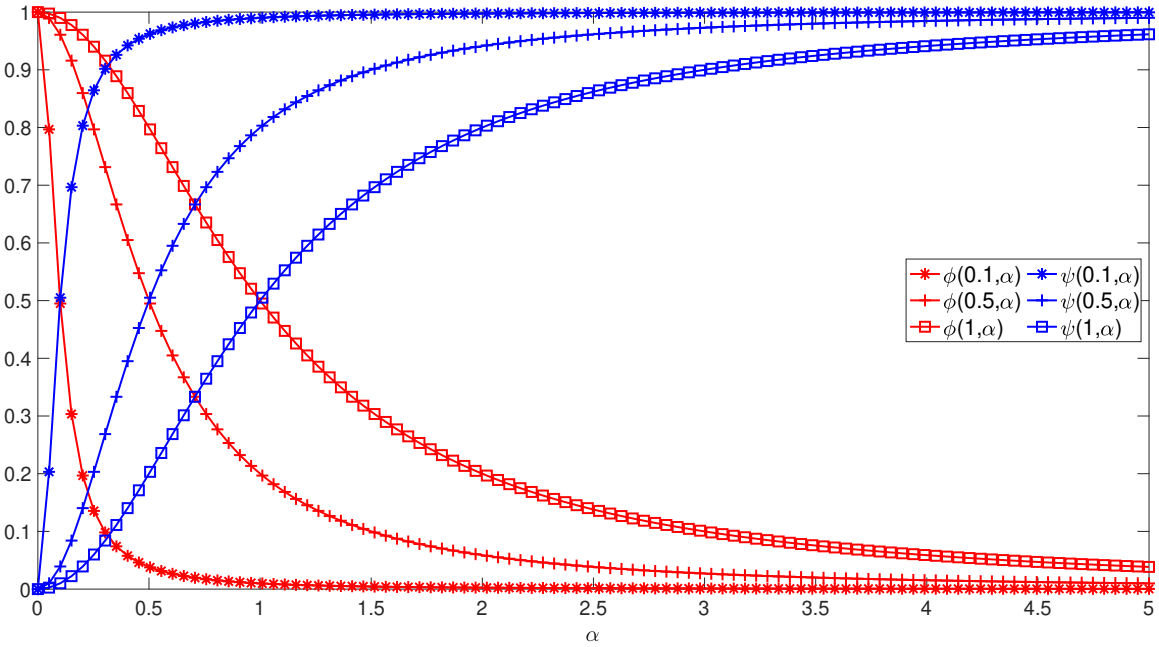


Figure 2.3: The behavior of the Tikhonov filter function $\phi_j(\alpha)$ and corresponding $\psi_j(\alpha) = 1 - \phi_j(\alpha)$ is displayed for selected singular values $s_j = 0.1, 0.5$, and 1. For fixed s_j, $\phi_j(\alpha) \to 0$ and $\psi_j(\alpha) \to 1$ as $\alpha \to \infty$. Notice also that $\phi_j(\alpha) = \psi_j(\alpha) = 1/2$ for $\alpha = s_j$.

non-singular, then the substitutions $\mathbf{y} = L\mathbf{x}$ and $B = AL^{-1}$ give

$$\mathbf{y}(\alpha) = \underset{\mathbf{y}\in\mathbb{R}^n}{\arg\min} \|B\mathbf{y} - \mathbf{d}\|_2^2 + \alpha^2\|\mathbf{y}\|_2^2. \tag{2.2.13}$$

This is known as the standard form of the regularization problem. Once $\mathbf{y}$ is obtained from (2.2.13), the final solution is recovered by $\mathbf{x} = L^{-1}\mathbf{y}$.

There are many examples of matrices L that are singular, such as finite difference matrices that approximate derivative operators. The case with singular L is addressed by eq. (2.2.15); an extension of the transformation that rendered (2.2.13) is now stated from (Hansen, 1998, p. 38) by introducing an A-weighted generalized inverse of L:

$$L_A^\dagger := \left(I - \left(A\left(I - L^\dagger L\right)\right)^\dagger A\right) L^\dagger, \quad I \in \mathbb{R}^{n\times n}. \tag{2.2.14}$$

The definition of $L_A^\dagger$ is somewhat natural, as evidenced by Prop. 2.2.1.

Proposition 2.2.1 ((Hansen, 1998, p. 38)) *Given $A \in \mathbb{R}^{m\times n}$, $L \in \mathbb{R}^{q\times n}$, and $L_A^\dagger$*

defined by (2.2.14), *if* $q \geq n$ *and* $\text{rank}(L) = n$ *then* $L_A^\dagger = L^\dagger$.

Proof. Since $\text{rank}(L) = n$, part (ii) of Prop. B.0.1 gives $L^\dagger = \left(L^\intercal L\right)^{-1} L^\intercal$. Thus $L^\dagger L = I_n$, and so

$$L_A^\dagger = \left(I_n - \left(A\left(I_n - I_n\right)\right)^\dagger A\right) L^\dagger = \left(I_n - \left(\mathbf{0}_{m\times n}\right)^\dagger A\right) L^\dagger.$$

From part (iii) of Prop. B.0.1, $\left(\mathbf{0}_{m\times n}\right)^\dagger = \mathbf{0}_{n\times m}$ and therefore

$$L_A^\dagger = \left(I_n - \mathbf{0}_{n\times m} A\right) L^\dagger = \left(I_n - \mathbf{0}_{n\times n}\right) L^\dagger = L^\dagger.$$

□

By letting $\mathbf{x}_0 = \left(A\left(I_n - L^\dagger L\right)\right)^\dagger \mathbf{d}$, we can define

$$\mathbf{y} = L\mathbf{x}, \quad B = AL_A^\dagger, \quad \mathbf{c} = \mathbf{d} - A\mathbf{x}_0. \tag{2.2.15}$$

These substitutions transform eq. (1.0.2) into the standard form eq. (2.2.13), while the solution $\mathbf{x}$ can be obtained from

$$\mathbf{x} = L_A^\dagger \mathbf{y} + \mathbf{x}_0. \tag{2.2.16}$$

The transformation proceeds by the simultaneous addition and subtraction of $A\mathbf{x}_0$ as follows

$$A\mathbf{x} - \mathbf{d} = A\mathbf{x} - A\mathbf{x}_0 - \mathbf{d} + A\mathbf{x}_0 = A\left(\mathbf{x} - \mathbf{x}_0\right) - \left(\mathbf{d} - A\mathbf{x}_0\right) = AL_A^\dagger \mathbf{y} - \mathbf{c} = B\mathbf{y} - \mathbf{c}.$$

Prop. 2.2.2 shows that $\mathbf{x}_0$ indeed lies within the nullspace of L. For example, $\text{null}(L)$ is the set of constant vectors $\mathbf{x}_0 \in \mathbb{R}^n$ for $L \in \mathbb{R}^{(n-1)\times n}$ defined by

$$L_{j,j} = 1, \quad L_{j,j+1} = -1, \qquad j = 1, \ldots, n-1.$$

Proposition 2.2.2 ((Hansen, 1998, p. 39)) *Given* $A \in \mathbb{R}^{m\times n}$, $L \in \mathbb{R}^{q\times n}$, *and* $\mathbf{d} \in \mathbb{R}^m$, *let*

$$\mathbf{x}_0 = \left(A\left(I_n - L^\dagger L\right)\right)^\dagger \mathbf{d}.$$

Then $\mathbf{x}_0 \in \text{null}(L)$.

Proof. Let $B = I_n - L^\dagger L$. Part (ii) of Prop. B.0.2 states that $L^\dagger L$ is the orthogonal projection onto range(L^T). Thus B is the orthogonal projection onto null(L). From part (iv) of Prop. B.0.2, we can write $L\mathbf{x}_0$ as $L\mathbf{x}_0 = L(AB)^\dagger \mathbf{d} = LB(AB)^\dagger \mathbf{d}$. Substituting back $B = I_n - L^\dagger L$ into the product LB, we have

$$LB = L\left(I_n - L^\dagger L\right) = L - LL^\dagger L.$$

$L^\dagger$ satisfies Moore-Penrose condition (I) in appendix B, meaning $L - LL^\dagger L = L - L = \mathbf{0}$ where $\mathbf{0} \in \mathbb{R}^{q\times n}$. Therefore

$$L\mathbf{x}_0 = LB(AB)^\dagger \mathbf{d} = \mathbf{0}(AB)^\dagger \mathbf{d} = \mathbf{0} \in \mathbb{R}^q,$$

and so $\mathbf{x}_0 \in$ null(L). □

Recasting regularization can also be accomplished in a convenient way using the generalized singular value decomposition (GSVD). In addition, it is no longer possible to use the SVD to express the solution eq. (2.2.11); rather, the use of the GSVD is invoked.

2.3 Generalized Singular Value Decomposition

Assuming null(A) $\cap$ null(L) = $\{\mathbf{0}\}$ for real matrices A and L of size $m \times n$ and $q \times n$, respectively, which will be subsequently referred to as the invertibility condition, the decompositions

$$A = U\Delta X^\mathsf{T}, \quad L = V\Lambda X^\mathsf{T} \tag{2.3.17}$$

exist where U is an $m \times m$ orthogonal matrix, V is a $q \times q$ orthogonal matrix, and X is an $n \times n$ non-singular matrix. Λ is a $q \times n$ diagonal matrix with non-negative diagonal elements in decreasing order on the principal diagonal Λ_{jj}, $1 \leq j \leq q^* = \min(q, n)$, and the only elements of the $m \times n$ matrix Δ that are possibly non-zero are

$$0 \leq \Delta_{(1,k+1)} \leq \Delta_{(2,k+2)} \leq \ldots \leq \Delta_{(\min(m,n),k+\min(m,n))} \leq 1 \tag{2.3.18}$$

with $k = \max\{n - m, 0\}$. This construction is just one version of the GSVD, Golub and Van Loan (2013). When $m \geq n$, we have $k = 0$ and eq. (2.3.18) reduces to the possibly non-zero elements of Δ being located on the main diagonal. Note that the invertibility condition is independent of the assumption that $m \geq n$. Using the GSVD with eq. (2.2.12) and the full rank condition yields the solution

$$\mathbf{x}(\alpha) = Y\Big(\Delta^{\mathsf{T}}\Delta + \alpha^2\Lambda^{\mathsf{T}}\Lambda\Big)^{-1}\Delta^{\mathsf{T}}\widehat{\mathbf{d}}, \tag{2.3.19}$$

where Y is the inverse of X^{T} and $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$.

Certainly $\Lambda^{\mathsf{T}}\Lambda$ is diagonal, but the structure of $\Delta^{\mathsf{T}}\Delta$ may not be obvious if $m < n$. To explain the structure of $\Delta^{\mathsf{T}}\Delta$, the notion of a k-diagonal matrix is introduced first. A k-diagonal matrix $\Delta \in \mathbb{R}^{m\times n}$ with $m \leq n$ and $k = n - m$ is a matrix whose only non-zero elements are $\Delta_{j,j+k}$ for $1 \leq j \leq m$; note that when $m = n$, A is a diagonal matrix. Lem. 2.3.1 presents some well-known basic properties (see Aster *et al.* (2013) for instance) of k-diagonal matrices.

Lemma 2.3.1 *For any matrices $M \in \mathbb{R}^{m\times m}$ and $N \in \mathbb{R}^{n\times n}$ and any k-diagonal matrix $\Delta \in \mathbb{R}^{m\times n}$ with $m < n$ and $k = n - m$, partition Δ as $\Delta = [\mathbf{0}_{m\times k}\ D]$, where $\mathbf{0}_{m\times k}$ is an $m \times k$ zero matrix and D is a $m \times m$ diagonal matrix. Similarly, partition N as*

$$M = \begin{bmatrix} N_{k\times k} & N_{k\times m} \\ \\ N_{m\times k} & Q \end{bmatrix},$$

where F is an $m\times m$ matrix. Then, the structure of the matrix conjugations $\Delta^{\mathsf{T}}M\Delta \in \mathbb{R}^{n\times n}$ and $\Delta N\Delta^{\mathsf{T}} \in \mathbb{R}^{m\times m}$ are described by the following.

(a) The matrix $\Delta^{\mathsf{T}} M \Delta$ is block diagonal with

$$\Delta^{\mathsf{T}} M \Delta = \begin{bmatrix} \mathbf{0}_{k\times k} & \mathbf{0}_{k\times m} \\ \\ \mathbf{0}_{m\times k} & D^{\mathsf{T}} M D \end{bmatrix}.$$

(b) $\Delta N \Delta^{\mathsf{T}} = D Q D^{\mathsf{T}}$.

Proof. For part (a), the partitioning of Δ yields

$$\Delta^{\mathsf{T}} M \Delta = \begin{bmatrix} \mathbf{0}_{k\times m} \\ \\ D^{\mathsf{T}} \end{bmatrix} M \begin{bmatrix} \mathbf{0}_{m\times k} & D \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{k\times m} \\ \\ D^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{m\times k} & MD \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{k\times k} & \mathbf{0}_{k\times m} \\ \\ \mathbf{0}_{m\times k} & D^{\mathsf{T}} M D \end{bmatrix}.$$

Similarly, the proof of part (b) uses the partitioning of both Δ and N:

$$\Delta N \Delta^{\mathsf{T}} = \begin{bmatrix} \mathbf{0}_{m\times k} & D \end{bmatrix} \begin{bmatrix} N_{k\times k} & N_{k\times m} \\ \\ N_{m\times k} & Q \end{bmatrix} \begin{bmatrix} \mathbf{0}_{k\times m} \\ \\ D^{\mathsf{T}} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{m\times k} & D \end{bmatrix} \begin{bmatrix} N_{k\times m} D^{\mathsf{T}} \\ \\ Q D^{\mathsf{T}} \end{bmatrix} = D Q D^{\mathsf{T}}.$$

□

It then follows from part (a) of Lem. 2.3.1 that $\Delta^{\mathsf{T}} \Delta$ is diagonal, which is stated as the following corollary.

Corollary 2.3.1 *The $n \times n$ matrix $\Delta^{\mathsf{T}} \Delta$ in eq. (2.3.19) is diagonal.*

Proof. The $m \times n$ matrix Δ is diagonal if $m \geq n$, in which case it is immediate that $\Delta^{\mathsf{T}} \Delta$ is diagonal. If $m < n$, then Δ is a k-diagonal matrix with $k = n - m$. Part (a) of Lem. 2.3.1 can be applied with $M = I$, which implies that the structure of $\Delta^{\mathsf{T}} \Delta$ is

$$\Delta^{\mathsf{T}} \Delta = \begin{bmatrix} \mathbf{0}_{k\times k} & \mathbf{0}_{k\times m} \\ \\ \mathbf{0}_{m\times k} & D^{\mathsf{T}} D \end{bmatrix}$$

with partition $\Delta = [\mathbf{0}_{m\times k} \; D]$ and diagonal matrix D. The matrix $D^{\mathsf{T}} D$ is thus diagonal, and therefore $\Delta^{\mathsf{T}} \Delta$ is diagonal. □

A convenient property of $\Lambda^\mathsf{T}\Lambda$ and $\Delta^\mathsf{T}\Delta$ is that $\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda = I_n$ (Aster *et al.*, 2013, p. 104). A consequence of Cor. 2.3.1 is that the matrix $\Delta^\mathsf{T}\Delta + \alpha^2\Lambda^\mathsf{T}\Lambda$ in eq. (2.3.19) is diagonal.

Defining $\boldsymbol{\delta} = \sqrt{\text{diag}(\Delta^\mathsf{T}\Delta)}$ and $\boldsymbol{\lambda} = \sqrt{\text{diag}(\Lambda^\mathsf{T}\Lambda)}$ (the square roots being applied element-wise), generalized singular values $\gamma_j = \delta_j/\lambda_j$ for $j = 1, \ldots, n$ can be formed for non-zero λ_j. In contrast to the decreasing ordering of the singular values s_j, the γ_j are ordered in increasing order, $\gamma_1 \leq \ldots \leq \gamma_n$, which is due to the order of the diagonal elements of $\text{diag}(\Delta^\mathsf{T}\Delta)$ and $\text{diag}(\Lambda^\mathsf{T}\Lambda)$. As noted in (Aster *et al.*, 2013, p. 107), there are situations in which some generalized singular values γ_j are arbitrarily small (meaning δ_j is arbitrarily small). Likewise, γ_j can be arbitrarily large (occurring when λ_j is small). These situations can either be attributed to the actual matrices or the numerical calculation of the GSVD. In the case where $\lambda_j = 0$, the corresponding γ_j is considered infinite; with $q^* = \min\{q, n\}$, there are q^* non-zero λ_j. To account for these possibilities, a preset tolerance $\tau > 0$ can be set for the magnitudes of δ_j and λ_j, and the filter factors can be defined appropriately. Specifically, for $j = 1, \ldots, n$ we redefine $\phi_j(\alpha)$ and $\psi_j(\alpha)$ by

$$\phi_j(\alpha) = \begin{cases} 0, & \delta_j < \tau \\ 1, & \lambda_j < \tau \\ \dfrac{\gamma_j^2}{\gamma_j^2 + \alpha^2}, & \text{otherwise,} \end{cases} \qquad \psi_j(\alpha) = \begin{cases} 1, & \delta_j < \tau \\ 0, & \lambda_j < \tau \\ \dfrac{\alpha^2}{\gamma_j^2 + \alpha^2}, & \text{otherwise.} \end{cases} \quad . \tag{2.3.20}$$

Note that we retain the definition of the complement $\psi_j(\alpha) = 1 - \phi_j(\alpha)$, as well as the diagonal matrices $\Phi(\alpha)$ and $\Psi(\alpha)$. Assuming that L has full rank q^*, then certainly $\phi_j(\alpha) = 1$ for $j > q^*$. Sums of $\phi_j(\alpha)$ and $\psi_j(\alpha)$ for $j = 1, \ldots, n$ are used repeatedly in Chapters 4 and 5. It is convenient to note that

$$\sum_{j=1}^{n} \phi_j(\alpha) = (n - q^*) + \sum_{j=1}^{q^*} \phi_j(\alpha), \quad \sum_{j=1}^{n} \psi_j(\alpha) = \sum_{j=1}^{q^*} \psi_j(\alpha).$$

The first derivatives of $\phi_j(\alpha)$ and $\psi_j(\alpha)$ (used in Chapter 4) are then

$$\frac{d\phi_j}{d\alpha} = \begin{cases} 0, & \delta_j < \tau \\ 0, & \lambda_j < \tau \\ \dfrac{-2\alpha\gamma_j^2}{(\gamma_j^2+\alpha^2)^2}, & \text{otherwise,} \end{cases} \qquad \frac{d\psi_j}{d\alpha} = \begin{cases} 0, & \delta_j < \tau \\ 0, & \lambda_j < \tau \\ \dfrac{2\alpha\gamma_j^2}{(\gamma_j^2+\alpha^2)^2}, & \text{otherwise.} \end{cases} \tag{2.3.21}$$

Using eq. (2.3.20), the generalized Tikhonov regularized solution can be written as

$$\mathbf{x}(\alpha) = \sum_{j=1}^{n} \frac{\delta_j^2}{\delta_j^2+\alpha^2\lambda_j^2}\frac{\widehat{d}_j}{\delta_j}Y_{\cdot,j} = \sum_{j=1}^{n} \frac{\gamma_j^2}{\gamma_j^2+\alpha^2}\frac{\widehat{d}_j}{\delta_j}Y_{\cdot,j} = \sum_{j=1}^{n} \phi_j(\alpha)\frac{\widehat{d}_j}{\delta_j}Y_{\cdot,j}, \tag{2.3.22}$$

where summands with $\delta_j < \tau$ are treated as zero.

Returning to matrix form and using $\Delta^{\mathsf{T}} = \Delta^{\mathsf{T}}\Delta\Delta^{\dagger}$ (see Prop. B.0.3), eq. (2.3.19) is replaced by

$$\mathbf{x}(\alpha) = Y\Phi(\alpha)\Delta^{\dagger}\widehat{\mathbf{d}}, \tag{2.3.23}$$

and by eq. (2.3.20) we have $\Phi(\alpha) = \left(\Delta^{\mathsf{T}}\Delta + \alpha^2\Lambda^{\mathsf{T}}\Lambda\right)^{-1}\Delta^{\mathsf{T}}\Delta$. Note that eq. (2.3.23) corresponds to eq. (2.1.6) when $L = I_n$. Equation (2.3.23) also allows use to write the generalized inverse as $A^{\sharp}(\alpha) = Y\Phi(\alpha)\Delta^{\dagger}U^{\mathsf{T}}$.

2.4 Summary

Tikhonov regularization is an effective means by which solutions can be obtained from ill-condition linear systems. Tools such as the SVD and GSVD can be used to analyze and implement the regularization process. While the GSVD is useful for analyzing problems with a matrix pair (A, L), there are efficient alternatives to the GSVD if A and L have some special structure. In the context of image deblurring, for example, the system matrix A often has structure that lends itself to decompositions related to specific discrete trigonometric transforms; the 2D problems considered in this work fall into such categories. Chapter 3 presents the numerical background for such systems.

Chapter 3

PROBLEM FORMULATIONS

We first discuss some matrix structures and how discretizations of continuous problems are connected with matrix structure. For image restoration problems, this involves relationships with the imposed boundary conditions. Some discrete trigonometric transforms and related matrix transformation are discussed, and some results regarding transformations of white noise will be presented in Section 3.5.

3.1 Matrix Structure

Matrices associated with blurring can be uniquely structured based upon which boundary conditions are imposed. These structures will be discussed in greater depth since they are used for numerical implementation of later techniques. Versions of such matrices associated with 1D problems will be presented first, followed by their 2D analogs.

Toeplitz matrices arise from blurring problems in which zero boundary conditions are imposed. A matrix T is called a Toeplitz matrix if it is constant along each diagonal, though by definition Toeplitz matrices need not be square. If a Toeplitz matrix is square with dimension $n \times n$, then it is uniquely determined by $2n - 1$ entries. If a Toeplitz matrix is symmetric, then it is determined by n entries.

Some notation for symmetric Toeplitz matrices will be introduced that will assist with the later discussion, particularly in Section 3.3. Given a column vector $\mathbf{x}$ of length n, let Toeplitz($\mathbf{x}$) denote the $n \times n$ symmetric Toeplitz matrix with $\mathbf{x}$ as its first column. The discrete convolution $\mathbf{x} * \mathbf{y}$ of $\mathbf{x}$ and $\mathbf{y}$ of equal length can be cast as the matrix-vector product $X\mathbf{y}$ using $X = \text{Toeplitz}(\mathbf{x})$; this is equivalent to `conv(x,y)`

in MATLAB®.

A type of matrix closely related to Toeplitz matrices is a Hankel matrix, which is a matrix that is constant along each anti-diagonal. Similar to square Toeplitz matrices, a square Hankel matrix with dimension $n \times n$ is uniquely defined by $2n-1$ entries and symmetric Hankel matrices are determined by n entries. Extending the notation for symmetric Toeplitz matrices, let Hankel($\mathbf{x}$) be the $n \times n$ symmetric Hankel matrix with the n-vector $\mathbf{x}$ as its first column.

A specific type of square Hankel matrix is the exchange matrix J_n, defined by the matrix which has ones along the main anti-diagonal and zero elsewhere. The exchange matrix is so-named for the property that the product of J_n with a vector $\mathbf{x} \in \mathbb{C}^n$ has the effect of "exchanging" (or "reversing") the entries of $\mathbf{x}$. For example, if $\mathbf{x} = [1, 2, 3]^\mathsf{T}$ then $J_3\mathbf{x} = [3, 2, 1]^\mathsf{T}$. As a result of this property, the exchange matrix is an involution, i.e. $J_n^2 = I_n$. Another useful property of J_n is that the product of a Toeplitz matrix with the exchange matrix is a Hankel matrix. Lem. 3.1.1 outlines a specific case where the Toeplitz and Hankel matrices are symmetric.

Lemma 3.1.1 (Ng *et al.* (1999)) *Let $\mathbf{x} \in \mathbb{R}^n$ and let J_n be the $n \times n$ exchange matrix. Then,*

$$\text{Toeplitz}(J_n\mathbf{x})J_n = \text{Hankel}(\mathbf{x}).$$

Proof. By the symmetry of Toeplitz($J_n\mathbf{x}$)J_n, it suffices to argue that the first column of Toeplitz($J_n\mathbf{x}$)J_n is $\mathbf{x}$. Since Toeplitz($J_n\mathbf{x}$) is post-multiplied by J_n, the first column of the product is equal to the last column of Toeplitz($J_n\mathbf{x}$). By definition, the last column of Toeplitz($J_n\mathbf{x}$) is $J_n(J_n\mathbf{x}) = J_n^2\mathbf{x} = \mathbf{x}$. □

Another class of matrices related to Toeplitz matrices are circulant matrices. To describe a circulant matrix, we first define the circular right shift of a row vector $[x_1, x_2, \ldots, x_n]$ as $[x_n, x_1, \ldots, x_{n-1}]$. If a matrix X has the property that each row

is the circular right shift of the components of the preceding row, then X is called a circulant matrix. From this definition, every circulant matrix is also a Toeplitz matrix. Another observation regarding circulant matrices can be made about their indexing. If $\mathbf{x} = [x_1, \cdots, x_n]$ is the first row of a circulant matrix X, then

$$X_{j,k} = x_{(k-j \bmod n)+1} \tag{3.1.1}$$

for all $1 \leq j, k \leq n$. A significant property of square circulant matrices discussed in Section 3.3 is that they are diagonalized by the discrete Fourier transform, which can be proved using eq. (3.1.1).

As a final point regarding matrix structures, the various matrices discussed in this section can be extended for use in 2D problems. The linear systems that arise for 2D problems are more complicated in their structure. Perhaps the easiest approach to conceptualize these structures is to vectorize both the data and solution and construct a matrix-vector equation describing the system. To illustrate this process, consider a matrix $X \in \mathbb{C}^{m \times n}$ (perhaps an image to be blurred), and create the vector $\mathbf{x}$ of length $N = mn$ defined through the following vectorization operator (Smith, 2007, p. 71):

$$\text{vec}(X) = [X_{1,1}, X_{2,1}, \ldots, X_{1,2}, X_{2,2}, \ldots, X_{m-1,n}, X_{m,n}]^{\mathsf{T}} \in \mathbb{C}^N, \tag{3.1.2}$$

where $X_{j,k}$ are the entries of X. After forming $\mathbf{x} = \text{vec}(X)$, multiplication with system matrix $A \in \mathbb{C}^{M \times N}$ yields $\mathbf{b} = A\mathbf{x}$. Vectorization can be undone by the array

operator (Smith, 2007, p. 71) $\text{arr} : \mathbb{C}^{mn} \to \mathbb{C}^{m\times n}$ defined by

$$\text{arr}\,(\mathbf{x}) = \begin{bmatrix} x_1 & x_{m+1} & \dots & x_{(n-1)m+1} \\ x_2 & x_{m+2} & \dots & x_{(n-1)m+2} \\ \vdots & \vdots & \ddots & \vdots \\ x_m & x_{2m} & \dots & x_{mn} \end{bmatrix} \in \mathbb{C}^{m\times n}. \tag{3.1.3}$$

Using the array operator, a blurred image $B = \text{arr}(\mathbf{b})$ can be produced from a matrix-vector product.

In many cases, the large system matrix A can be written as the Kronecker product of two smaller matrices. Given matrices $X \in \mathbb{C}^{m\times n}$ and $Y \in \mathbb{C}^{p\times q}$, the Kronecker product $X \otimes Y \in \mathbb{C}^{mp\times nq}$ is defined by

$$X \otimes Y = \begin{bmatrix} x_{1,1}Y & x_{1,2}Y & \dots & x_{1,n}Y \\ x_{2,1}Y & x_{2,2}Y & \dots & x_{2,n}Y \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1}Y & x_{m,2}Y & \dots & x_{m,n}Y \end{bmatrix}. \tag{3.1.4}$$

In addition to system matrices, the 2D versions of discrete trigonometric transforms discussed in Section 3.3 can be expressed as Kronecker products of matrices associated with the 1D transforms. Lem. 3.4.2 in Section 3.4 includes a result that connects Kronecker product decompositions with the GSVD.

The matrix structures already discussed can be extended to two dimensions for use with $\mathbf{x} = \text{vec}(X)$. A block-circulant-circulant-block matrix A, or BCCB matrix for short Vogel (2002), is a matrix that can be partitioned into blocks where A is circulant

in regard to these blocks and each block is itself a circulant matrix. Similarly there are block-Toeplitz-Toeplitz-block (BTTB) matrices; solving a system $A\mathbf{x} = \mathbf{b}$ where A is a BTTB matrix requires $O(n^4)$ operations Kalouptsidis *et al.* (1984) unless it is embedded within a BCCB matrix. There are also block structures involving Hankel matrices and ones involving a mix of structures, such as a block-Toeplitz-Hankel-block (BTHB). BTTB + BTHB + BHTB + BHHB matrices can be used in conjunction with diagonalization properties of the discrete cosine transform Hansen *et al.* (2006) for efficient numerical implementation (discussed here in Section 3.3). See both (Vogel, 2002, p. 71-75) and Hansen *et al.* (2006) for in-depth instruction on how to construct such block matrices explicitly.

Explicit construction and manipulation of such matrices associated with 2D problem is discouraged, however, due to their computational storage cost. Indeed, the system matrix to work with the vectorization of a 256×256 image would have dimension $65{,}536 \times 65{,}536$ (a matrix having potentially 2^{32} non-zero entries, requiring approximately 34 gigabytes). Fortunately, many of the matrix structures described in this section benefit from diagonalization properties of matrices associated with discrete trigonometric transforms. Section 3.3 provides background on these transforms in the 1D setting, which is then tied back to their relevant diagonalization properties in Section 3.4 and are utilized in the numerical applications of this work.

3.2 Deblurring Background

Here we consider the deconvolution problem that was introduced in Section 2.1 in more detail. We begin by carefully describing how to obtain a matrix representation of eq. (2.1.1). To describe deconvolution problem eq. (2.1.1) in a discrete setting, a starting approach is to consider bi-infinite sequences, which are obtained by sampling

the functions at a countable number of points in $\mathbb{R}$:

$$(k_n) = (\ldots, k_{-1}, k_0, k_1, \ldots), \quad (f_n) = (\ldots, f_{-1}, f_0, f_1, \ldots).$$

The convolution (2.1.1) then has an analogous definition for sequences: the jth entry g_j of (g_n) is

$$g_j = \sum_{\ell=-\infty}^{\infty} k_{j-\ell} f_\ell. \tag{3.2.5}$$

If the kernel has compact support, then (k_n) will have finitely many nonzero terms and can be expressed as

$$(k_n) = (\ldots, 0, k_{-m+1}, \ldots, k_{-1}, k_0, k_1, \ldots, k_{m-1}, 0, \ldots). \tag{3.2.6}$$

In this representation (k_n) has potentially $2m-1$ non-zero terms. Furthermore, if the data vector $\mathbf{d}$ has length m, then (3.2.5) can be expressed as a matrix-vector product, which is efficiently represented as

$$\mathbf{b} = K_{\text{left}}\mathbf{x}_{\text{left}} + K\mathbf{x} + K_{\text{right}}\mathbf{x}_{\text{right}}. \tag{3.2.7}$$

Here

$$\mathbf{x}_{\text{left}} = \begin{bmatrix} f_{-m+3} \\ f_{-m+4} \\ \vdots \\ f_0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}, \quad \mathbf{x}_{\text{right}} = \begin{bmatrix} f_{n+1} \\ f_{n+2} \\ \vdots \\ f_{n+m} \end{bmatrix}$$

and

$$K_{\text{left}} = \begin{bmatrix} k_{m-1} & \cdots & k_1 & & \\ & \ddots & \ddots & & \\ & & & k_{m-1} & \\ & & \mathbf{0} & & \end{bmatrix}, K_{\text{right}} = \begin{bmatrix} & & & \\ & & \mathbf{0} & \\ & & & \\ k_{-m+1} & & & \\ \vdots & \ddots & & \\ k_{-1} & \cdots & k_{-m+1} & \end{bmatrix},$$

$$K = \begin{bmatrix} k_0 & \cdots & k_{-m+1} & & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \\ k_{m-1} & \ddots & \ddots & \ddots & k_{-m+1} \\ & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & & k_{m-1} & \cdots & k_0 \end{bmatrix}.$$

Note that K_{left}, K, and K_{right} are Toeplitz matrices, and we use larger $\mathbf{0}$ symbol to denote combinations of zero matrices defined by their context. The goal is to obtain a solution $\mathbf{x} = [f_1, \ldots, f_n]^\mathsf{T}$. However, it is clear that $\mathbf{x}$ depends not only upon the choice of kernel k but also on the boundaries of f, i.e. the terms f_j of (f_n) where $j \notin \{1, \ldots, n\}$. There are a number of ways in which the boundaries can be handled; for example, zero, periodic, and reflective boundary conditions are considered in Ng *et al.* (1999). The most-simplifying condition in terms of the resulting system is the zero boundary condition, in which it is assumed that $\mathbf{x}_{\text{left}} = \mathbf{x}_{\text{right}} = \mathbf{0}$. The resulting system (3.2.7) is $A\mathbf{x} = \mathbf{b}$ with $A = K$ being a Toeplitz matrix (discussed in Section 3.1).

Analogous to how Toeplitz matrices are connected with the zero boundary condition, circulant matrices arise from the following periodic boundary condition on eq. (3.2.7):

$$\begin{cases} f_0 = f_n \\ f_{-1} = f_{n-1} \\ \vdots \\ f_{-m+3} = f_{n-m} \end{cases} \quad \text{and} \quad \begin{cases} f_{n+1} = f_1 \\ f_{n+2} = f_2 \\ \vdots \\ f_{n+m} = f_n \end{cases}.$$

The resulting matrix A is circulant and completely determined by the n-vector $\mathbf{f} = [f_1, \ldots, f_n]^\mathsf{T}$.

The reflective boundary condition is defined in the context of eq. (3.2.7) by setting

$$\begin{cases} f_0 = f_1 \\ f_{-1} = f_2 \\ \vdots \\ f_{-m+3} = f_n \end{cases} \quad \text{and} \quad \begin{cases} f_{n+1} = f_n \\ f_{n+2} = f_{n-1} \\ \vdots \\ f_{n+m} = f_{n-m+1} \end{cases}.$$

Using the exchange matrix J_n, the reflective boundary condition allows for eq. (3.2.7) to be written as

$$A\mathbf{f} = ([\mathbf{0} \mid K_{\text{left}}]J_n + K + [K_{\text{right}} \mid \mathbf{0}]J_n)\,\mathbf{f} = \mathbf{b}, \tag{3.2.8}$$

where $[\mathbf{0} \mid K_{\text{left}}]$ and $[K_{\text{right}} \mid \mathbf{0}]$ are the $n \times n$ Toeplitz matrices formed by augmenting K_{left} and K_{right} with $n - m + 1$ zero columns. As a result, the matrices $[\mathbf{0} \mid K_{\text{left}}]J$ and $[K_{\text{right}} \mid \mathbf{0}]J_n$ are Hankel matrices. Their sum with K is characterized as another unique matrix structure related to the discrete cosine transform (see Section 3.3).

Up to this point, we have considered problems of the form $A\mathbf{x} \approx \mathbf{d}$ where $A \in \mathbb{R}^{m\times n}$, which approximate a deconvolution problem similar to that of eq. (2.1.1). The

convolution integral in two dimensions is

$$g(x,y) = (k * f)(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} k(x-u, y-v) f(u,v)\ du\ dv. \tag{3.2.9}$$

The more general problem of having $k(x,u,y,v)$ instead of $k(x-u, y-v)$ corresponds with the 2D analog of eq. (2.1.2) and having $k(x,u,y,v) = k(x-u, y-v)$ means that the kernel is spatially invariant. The existence of the integral is again predicated upon the functional characteristics of k and f; for example, having k and f be compactly supported results in the existence of g that is continuous and compactly supported. A connection between eq. (3.2.9) and images is made by Vogel (2002), in which g represents a continuously blurred image where light (photons) can be counted on rectangular pixels $\Omega_{j,k} \subset \mathbb{R}^2$ so that

$$g_{j,k} = \int \int_{\Omega_{j,k}} g(x,y)\ dx\ dy, \quad 1 \leq k \leq n_x, \tag{3.2.10}$$

provides a representation of $g(x,y)$ as part of a discrete image for $1 \leq j \leq n_y$ and $1 \leq k \leq n_x$. We use n_y and n_x to denote the number of pixels in the x and y directions, respectively, and $N = n_y n_x$ is the total number of pixels. Letting Ω be the union of all of the pixels $\Omega_{j,k}$, eq. (3.2.9) can be truncated to an integral over Ω. As with the 1D case, numerical quadrature can be used to approximate the truncation of eq. (3.2.9) as well as eq. (3.2.10) to form a linear system approximating the continuous problem. For 2D problems, the system matrices are block matrices which possess special structure, such as being BCCB.

Even from this simple derivation of a discrete problem from a continuous convolution, it becomes clear that the system matrices for deblurring problems can have special structure. It will be seen later that some structures can be exploited to obtain regularized solutions that do not rely upon the construction of a GSVD. Now that the groundwork has been laid for describing deconvolution-type problems and ma-

trix structures, some discrete trigonometric transforms provide efficiency in numerical implementation.

3.3 Discrete Trigonometric Transforms

Having presented a number of matrix types in Section 3.1, their use in conjunction with discrete trigonometric transforms is the next step in formulating a background for the body of this work. Discrete trigonometric transforms provide efficient means of dealing with certain matrix types, especially moving from 1D to 2D problems. The discrete Fourier transform will be discussed first, followed by the discrete cosine transform and lastly the discrete sine transform. Though these transforms will be expressed in the 1D setting, their 2D counterparts are used for all 2D problems in applications of the spectral windowing and learning approaches for regularization.

The discrete Fourier transform (DFT) will be introduced from the perspective of approximating coefficients of the Fourier series of a 2π-periodic function f (Boggess and Narcowich, 2009, p. 132-134). On the interval $[0, 2\pi]$, the jth complex Fourier coefficient of $f(t)$ is given by

$$c_j = \frac{1}{2\pi} \int_0^{2\pi} f(t) \exp(-ijt)\, dt,$$

where $i = \sqrt{-1}$. Applying the trapezoidal rule for approximating this integral with n points then produces

$$c_j \approx \frac{1}{n} \sum_{\ell=1}^{n} f\left(\frac{2\pi(\ell-1)}{n}\right) \exp\left(\frac{-2\pi ij(\ell-1)}{n}\right).$$

The DFT is a mapping $\mathcal{F} : \mathbb{C}^n \to \mathbb{C}^n$ defined by

$$\mathcal{F}(\mathbf{f})_j = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} f_\ell \exp\left(\frac{-2\pi i(j-1)(\ell-1)}{n}\right), \quad \mathbf{f} \in \mathbb{C}^n, \quad 1 \le j \le n. \tag{3.3.11}$$

The inverse DFT of a vector $\widehat{\mathbf{f}}$ is given by

$$\mathcal{F}^{-1}\left(\widehat{\mathbf{f}}\right)_j = \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \widehat{f}_\ell \exp\left(\frac{2\pi i(j-1)(\ell-1)}{n}\right) = f_j. \tag{3.3.12}$$

These definitions are non-standard; typically the factors $1/\sqrt{n}$ in both the forward and inverse DFT definitions are combined as a single factor of $1/n$ in the definition of the forward DFT.

The DFT can also be stated in terms of matrix-vector multiplication. Given an $\mathbf{f} \in \mathbb{C}^n$, $\widehat{\mathbf{f}}$ can be expressed as $F\mathbf{f}$ where the matrix $F \in \mathbb{C}^{n\times n}$ has components

$$F_{j,k} = \frac{1}{\sqrt{n}} \exp\left(\frac{-2\pi i(j-1)(k-1)}{n}\right), \quad 1 \le j,k \le n. \tag{3.3.13}$$

The matrix representing the inverse DFT is then F^{H}, where H denotes conjugate transposition. A property of F is that $F^{\mathsf{H}}F = FF^{\mathsf{H}} = (1/n)\operatorname{diag}(n) = I_n$, and so splitting the factor of $1/n$ as $(1/\sqrt{n})(1/\sqrt{n})$ for the definition of the DFT provides the benefit of F being a unitary matrix. A direct consequence of the DFT being unitary is Parseval's theorem: $\|\mathbf{f}\|_2 = \|F\mathbf{f}\|_2$ for any $\mathbf{f} \in \mathbb{C}^n$.

A significant property of circulant matrices is that they are diagonalized by the DFT. Using the definition of the $n \times n$ unitary Fourier matrix F, the property can be stated as

$$C = F^{\mathsf{H}} \operatorname{diag}(\sqrt{n}\widehat{\mathbf{c}})F, \tag{3.3.14}$$

where C is any $n\times n$ circulant matrix and $\widehat{\mathbf{c}}$ is the DFT of the first row of C (recall that the rows of a circulant matrix are circulant right shifts of a single row vector of length n). This property can be shown directly by showing that $C_{j,k} = (F^{\mathsf{H}} \operatorname{diag}(\widehat{\mathbf{c}})F)_{j,k}$ for any $1 \le j,k \le n$. Using the definition of the DFT and (3.1.1), the component of $F^{\mathsf{H}} \operatorname{diag}(\sqrt{n}\widehat{\mathbf{c}})F$ in jth row and kth column is given by

$$\begin{aligned}
&\sum_{\ell=1}^{n} \left[\frac{1}{\sqrt{n}} \exp\left(\frac{2\pi i(k-1)(\ell-1)}{n}\right)\right] \left[\sqrt{n}\widehat{c}_\ell\right] \left[\frac{1}{\sqrt{n}} \exp\left(\frac{-2\pi i(j-1)(\ell-1)}{n}\right)\right] \\
&= \frac{1}{\sqrt{n}} \sum_{\ell=1}^{n} \widehat{c}_\ell \exp\left(\frac{2\pi i(k-j)(\ell-1)}{n}\right) \\
&= c_{(k-j \bmod n)+1} \\
&= C_{j,k}.
\end{aligned}$$

It should be noted that by defining the DFT with the factor of $1/n$ as part of the inverse transformation (while keeping the factor split as $(1/\sqrt{n})(1/\sqrt{n})$ in the matrix form) allows the property to be stated simply as $C = F^{\mathsf{H}} \operatorname{diag}(\widehat{\mathbf{c}})F$.

Analogous to the DFT, the discrete cosine transform (DCT) can be introduced from the perspective of approximating the coefficients of a cosine series. Given a function $f(t)$ defined on the interval $[0, 1]$, consider the even extension of $f_e(t)$ defined by

$$f_e(t) = \begin{cases} f(t), & 0 \leq t \leq 1 \\ f(-t), & -1 \leq t < 0 \end{cases}.$$

From (Boggess and Narcowich, 2009, p. 49), the Fourier series expansion of $f_e(t)$ will only contain cosine terms and the coefficients of these terms (starting the indexing at 1) are

$$\begin{aligned} a_1 &= \int_0^1 f(t)\ dt, \\ a_j &= 2\int_0^1 f(t)\cos((j-1)\pi t)\ dt, \quad j \geq 2. \end{aligned}$$

Approximating the integral for a_j with $j \geq 2$ using the midpoint rule with n subintervals of $[0, 1]$ gives

$$a_j \approx \frac{2}{n}\sum_{k=1}^{N} f\left(\frac{1}{n}\left(k - \frac{1}{2}\right)\right)\cos\left(\frac{j\pi}{n}\left(k - \frac{1}{2}\right)\right).$$

Adjusting the scale factor and rewriting the argument of cosine yields the discrete transform. Given $\mathbf{x} \in \mathbb{R}^n$, the DCT of $\mathbf{x}$, denoted $\breve{\mathbf{x}}$, is defined as

$$\breve{\mathbf{x}}_j = \sqrt{\frac{2 - \delta_{j,1}}{n}}\sum_{k=1}^{n} x_k \cos\left(\frac{(j-1)\pi(2k-1)}{2n}\right). \tag{3.3.15}$$

Here $\delta_{j,1}$ is the Kronecker delta function

$$\delta_{j,1} = \begin{cases} 1, & j = 1 \\ 0, & j \neq 1, \end{cases}$$

which in eq. (3.3.15) has the effect of introducing a factor of $1/\sqrt{n}$ in the first component of $\check{\mathbf{x}}$ instead of $\sqrt{2/n}$. The DCT matrix, denoted by C, has entries

$$C_{j,k} = \sqrt{\frac{2-\delta_{j,1}}{n}} \cos\left(\frac{\pi(j-1)(2k-1)}{2n}\right). \tag{3.3.16}$$

By switching j and k, it is clear that C is not symmetric. However, C is orthogonal which is analogous to the DFT matrix being unitary. Fortunately, many of the statistical results in Section 3.5 demonstrate that the statistics of the DCT of white noise are simpler than those of the DFT. Another benefit of the DCT is that the implied even boundary conditions ensure continuity at the endpoints in the continuous setting.

To fully describe the set of matrices diagonalized by the DCT, let the shift operator $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ be defined as

$$\sigma(\mathbf{x}) = [x_2, \ldots, x_n, 0]^\mathsf{T}, \quad \mathbf{x} = [x_1, x_2 \ldots, x_n]^\mathsf{T} \in \mathbb{R}^n.$$

The class of matrices diagonalized by the DCT can now be stated.

Theorem 3.3.1 *(Chan* et al. *(1999); Kailath and Olshevsky (1996); Martucci (1994); Sánchez* et al. *(1995))*
Let $\mathcal{C}$ be the set of matrices that can be diagonalized by the DCT matrix C. Then

$$\mathcal{C} = \{\text{Toeplitz}(\mathbf{x}) + \text{Hankel}\,(\sigma(\mathbf{x})) \mid \mathbf{x} \in \mathbb{R}^n\}.$$

The diagonalization properties of the DCT in relation to reflective boundary conditions can now be described Ng *et al.* (1999).

Theorem 3.3.2 (Ng *et al.* (1999)) *Let the kernel sequence* (3.2.6) *satisfy $k_j = k_{-j}$ for all $j \in \mathbb{Z}$. Then, the matrix A in* (3.2.8) *can be expressed as*

$$A = \text{Toeplitz}(\mathbf{k}) + \text{Hankel}\,(\sigma(\mathbf{k}))\,,$$

where $\mathbf{k} = [k_0, k_1, \ldots, k_{m-1}, 0, \ldots, 0]^\mathsf{T}$. In other words, A is diagonalized by the DCT.

Proof. Equation (3.2.8) gives $A = [(0 \mid T_{\text{left}})J + K + (T_{\text{right}} \mid 0)J_n]$. Then $K = \text{Toeplitz}(\mathbf{k})$ by the definition of K with $k_j = k_{-j}$, and so it remains to show that $[(0 \mid T_{\text{left}}) + (T_{\text{right}} \mid 0)]J_n = \text{Hankel}\,(\sigma(\mathbf{k}))$. From the definitions of T_{left} and T_{right}, the sum $(0 \mid T_{\text{left}}) + (T_{\text{right}} \mid 0)$ is equal to $\text{Toeplitz}\,(J_n\,\sigma(\mathbf{k}))$. Thus by Lem. 3.1.1,

$$[(0 \mid T_{\text{left}}) + (T_{\text{right}} \mid 0)]J_n = \text{Toeplitz}\,(J_n\,\sigma(\mathbf{k}))\,J_n = \text{Hankel}\,(\sigma(\mathbf{k}))\,.$$

Therefore $A = \text{Toeplitz}(\mathbf{k}) + \text{Hankel}\,(\sigma(\mathbf{k}))$, and Thm. 3.3.1 states that A can be diagonalized by the DCT. □

Next, the DST will be briefly discussed. Given a function $f(t)$ defined on the interval $[0, 1]$, consider the odd extension of $f_o(t)$ defined by

$$f_o(t) = \begin{cases} f(t), & 0 \le t \le 1 \\ -f(-t), & -1 \le t < 0 \end{cases}.$$

The Fourier series expansion of $f_o(t)$ will only contain sine terms and the coefficients of these terms are

$$b_j = 2\int_0^1 f(t)\sin(j\pi t)\;dt, \quad j \ge 1.$$

The factor j instead of $j-1$ is so that the series does not always begin with a zero term; in other words, b_1 is not identically zero. Using the midpoint rule with n subintervals gives

$$b_j \approx \frac{2}{n}\sum_{k=1}^{n} f\left(\frac{1}{n}\left(k-\frac{1}{2}\right)\right)\sin\left(\frac{j\pi}{n}\left(k-\frac{1}{2}\right)\right).$$

Again a modification of the scale factor and rewriting the argument of sine produces the discrete transform; the factor j in the discrete transform ensures that the first DCT component $\breve{b}_1$ is not guaranteed to be zero. Unlike the DCT, however, the DST implies odd boundary conditions that can produce discontinuities at the boundaries in the continuous setting. For this reason the DST is not widely considered for the

numerical experiments, though the DST has been considered, such as by Perrone (2006) when implemented with anti-reflective boundary conditions.

The 2D DFT or DCT of a matrix $X \in \mathbb{C}^{m \times n}$ can be written in terms of the Kronecker products of the associated 1D matrices. For example, let C_m and C_n be the 1D DCT matrices defined by eq. (3.3.16) of size $m \times m$ and $n \times n$, respectively, and let $C = C_n \otimes C_m$. Then the 2D DCT of X is

$$\breve{X} = \operatorname{arr}\left(C \operatorname{vec}(X)\right). \tag{3.3.17}$$

Computing 2D transformations using methods similar to eq. (3.3.17) is not practical, again due to the cost of constructing the matrix C. The matrices and associated diagonalization properties of 2D transformations are expressed conveniently, however, using Kronecker product notation.

3.4 Matrix Transformations

Sections 3.2 to 3.3 discuss connections between matrix structures associated with discrete convolution-type problems and discrete trigonometric transforms. The goal of Section 3.4 is to show how these concepts can be related back to the concepts presented in Chapter 2. The theoretical tools shown here, primarily Thm. 3.4.1, are the foundation of how the spectral windowing and learning results are examined numerically. Though the results are presented mostly in the context of the DCT, the DFT is also utilized in some of the numerical results.

Assuming that reflective boundary conditions are applied, then both A and L have the same block structure (BTTB + BTHB + BHTB + BHHB matrix) and the DCT can be used to simultaneously diagonalize A and L. A precursor to converting from a simultaneous diagonalization to the GSVD presented in Section 2.3 is Lem. 3.4.1, a straightforward result that shows how we can apply a joint transformation of two

vectors of non-negative numbers to obtain a specific ordering of each new vector.

Lemma 3.4.1 *Given two vectors* $\mathbf{a} = [a_1, \ldots, a_n]^\mathsf{T}$ *and* $\mathbf{b} = [b_1, \ldots, b_n]^\mathsf{T}$ *of non-negative numbers with* $a_j > 0$ *or* $b_j > 0$ *for each* $j = 1, \ldots, n$*, define two new vectors* $\mathbf{x} = [x_1, \ldots, x_n]^\mathsf{T}$ *and* $\mathbf{y} = [y_1, \ldots, y_n]^\mathsf{T}$ *by*

$$x_j = \frac{a_j}{\sqrt{a_j^2 + b_j^2}}, \quad y_j = \frac{b_j}{\sqrt{a_j^2 + b_j^2}}.$$

Then elements of $\mathbf{x}$ *and* $\mathbf{y}$ *can be ordered such that for all* $j = 1, \ldots, n$*, if* x_j *is the* k*th largest element of* $\mathbf{x}$*, then* y_j *is the* k*th smallest element of* $\mathbf{y}$*. Equivalently, there exists a permutation matrix* P *such that the elements of* $P\mathbf{x}$ *and* $P\mathbf{y}$ *are ordered from largest to smallest and smallest to largest, respectively.*

Proof. By construction, $0 < x_j, y_j \leq 1$ and $x_j^2 + y_j^2 = 1$ for all $j = 1, \ldots, n$. Thus we can write $x_j = \sin(\theta_j)$ and $y_j = \cos(\theta_j)$ with $0 < \theta_j \leq \frac{\pi}{2}$ for all $j = 1, \ldots, n$. The desired ordering of x_j and y_j then follow from the fact that the cosine and sine functions are strictly decreasing and increasing on $\left(0, \frac{\pi}{2}\right]$, respectively. □

A technique similar to that of Lem. 3.4.1 is utilized in the cosine sine (CS) decomposition of an orthogonal matrix; see (Golub and Van Loan, 2013, p. 84-86) for example. In fact, a conversion from CS decomposition to GSVD is presented in (Golub and Van Loan, 2013, p. 502-503). Using Lem. 3.4.1, a conversion from any simultaneous diagonalization of two matrices to the GSVD can be presented in the form of Thm. 3.4.1.

Theorem 3.4.1 *Given the simultaneous diagonalization of the* $n \times n$ *symmetric matrices* $A = C^\mathsf{T}\widetilde{\Delta}C$ *and* $L = C^\mathsf{T}\widetilde{\Lambda}C$ *where* C *is orthogonal and* $\text{null}(A) \cap \text{null}(L) = \{\mathbf{0}\}$*,* A *and* L *can be expressed as* $A = U\Delta X^\mathsf{T}$ *and* $L = U\Lambda X^\mathsf{T}$*, respectively, where* U *is orthogonal,* X *is invertible,* $\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda = I_n$*,* $0 \leq \Delta_{1,1} \leq \ldots \leq \Delta_{n,n} \leq 1$*, and* $1 \geq \Lambda_{1,1} \geq \ldots \geq \Lambda_{n,n} \geq 0$*.*

Proof. We begin by setting $C^{\mathsf{T}}\widetilde{\Delta}C = U\Delta Z^{\mathsf{T}}$ and rearranging terms to obtain $\Delta = U^{\mathsf{T}}C^{\mathsf{T}}\widetilde{\Delta}CZ^{-\mathsf{T}}$. Doing the same for $\widetilde{\Lambda}$, and using $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = I_n$, we have

$$Z^{-1}C^{\mathsf{T}}\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}CZ^{-\mathsf{T}} + Z^{-1}C^{\mathsf{T}}\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}CZ^{-\mathsf{T}} = I_n$$

from which we have

$$\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda} = CZZ^{\mathsf{T}}C^{\mathsf{T}}. \tag{3.4.18}$$

Since $\text{null}(A) \cap \text{null}(L) = \{\mathbf{0}\}$, the matrix on the left is diagonal with positive entries. Thus we can form the (positive) square root $S = \sqrt{\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}}$, which is also diagonal with positive entries. Using S we can write $S^2 = SS^{\mathsf{T}} = CZZ^{\mathsf{T}}C^{\mathsf{T}}$, which implies that we can set $Z = C^{\mathsf{T}}S$; S is invertible but not necessarily orthogonal, so Z is only invertible. Using the transpose $Z^{\mathsf{T}} = SC$ and inverse S^{-1}, we can then write

$$A = C^{\mathsf{T}}\widetilde{\Delta}S^{-1}SC = C^{\mathsf{T}}\widetilde{\Delta}S^{-1}Z^{\mathsf{T}},$$
$$L = C^{\mathsf{T}}\widetilde{\Lambda}S^{-1}SC = C^{\mathsf{T}}\widetilde{\Lambda}S^{-1}Z^{\mathsf{T}}.$$

The last step is to reorder the elements of $\widetilde{\Delta}S^{-1}$ and $\widetilde{\Lambda}S^{-1}$; fortunately Lem. 3.4.1 shows that they have the opposite order regardless of the order of the elements of $\widetilde{\Delta}$ and $\widetilde{\Lambda}$. Therefore we can use a permutation matrix P so that $\Delta = P^{\mathsf{T}}\widetilde{\Delta}S^{-1}P$ and $\Lambda = P^{\mathsf{T}}\widetilde{\Lambda}S^{-1}P$ have the desired ordering. Since permutation matrices are orthogonal, we finally obtain the GSVD:

$$A = C^{\mathsf{T}}\widetilde{\Delta}S^{-1}Z^{\mathsf{T}} = C^{\mathsf{T}}PP^{\mathsf{T}}\widetilde{\Delta}S^{-1}PP^{\mathsf{T}}Z^{\mathsf{T}} = U\Delta X^{\mathsf{T}},$$
$$L = C^{\mathsf{T}}\widetilde{\Lambda}S^{-1}Z^{\mathsf{T}} = C^{\mathsf{T}}PP^{\mathsf{T}}\widetilde{\Lambda}S^{-1}PP^{\mathsf{T}}Z^{\mathsf{T}} = U\Lambda X^{\mathsf{T}},$$

where $U = C^{\mathsf{T}}P$ is orthogonal and $X = ZP$ is invertible. We can confirm that $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = I_n$ using the commutability of diagonal matrices and the diagonality

of S^{-1}:

$$\begin{aligned}\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda &= P^{\mathsf{T}}S^{-\mathsf{T}}\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}S^{-1}P + P^{\mathsf{T}}S^{-\mathsf{T}}\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}S^{-1}P \\ &= P^{\mathsf{T}}S^{-\mathsf{T}}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)S^{-1}P \\ &= P^{\mathsf{T}}\left(S^2\right)^{-1}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)P \\ &= P^{\mathsf{T}}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)P = I_n.\end{aligned}$$

□

Equipped with Thm. 3.4.1, it is clear that the terms in eq. (2.3.22) can be rewritten in terms of a DCT simultaneous diagonalization of matrices A and L, which is particularly relevant for the efficient solution of two-dimensional problems.

The invertibility condition in Thm. 3.4.1 is necessary for the property $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = I_n$. If $\text{null}(A) \cap \text{null}(L) \neq \{\mathbf{0}\}$, then the matrix on the left in eq. (3.4.18) is singular and this contradicts the invertibility of C and Z on the right side. As an alternative, we could require that $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = \widetilde{I}_n$ where $\widetilde{I}_n$ is a modified identity matrix that has some zero diagonal elements. This generalization is described in Cor. 3.4.1.

Corollary 3.4.1 *Given the simultaneous diagonalization of the $n \times n$ symmetric matrices $A = C^{\mathsf{T}}\widetilde{\Delta}C$ and $L = C^{\mathsf{T}}\widetilde{\Lambda}C$ where C is orthogonal, A and L can be expressed as $A = U\Delta X^{\mathsf{T}}$ and $L = U\Lambda X^{\mathsf{T}}$, respectively, where U is orthogonal, X is invertible, $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = \widetilde{I}_n$, and*

$$0 \leq \Delta_{1,1} \leq \ldots \leq \Delta_{n,n} \leq 1,$$

$$1 \geq \Lambda_{1,1} \geq \ldots \geq \Lambda_{n,n} \geq 0.$$

Proof. The proof proceeds similarly to that of Thm. 3.4.1 until the construction of the positive square root of $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda$. If $\text{null}(A) \cap \text{null}(L) \neq \{\mathbf{0}\}$, then there

exist zero elements along the diagonal of $\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda$. We can still define the positive square root $S = \sqrt{\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda}$, except S now has zero elements on the diagonal as well and is thus a singular matrix. As such, we define a new diagonal matrix $\widetilde{S}$ from S by reassigning the zero diagonal elements to be ones. In this way, $\widetilde{S}$ is invertible but $\widetilde{S}^2 \neq \Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda$. However, $\widetilde{S}^2$ only differs from $\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda$ in the sense that the zero elements are replaced by ones. Ultimately this is not an issue, since the zero diagonal elements of the products $\widetilde{\Delta}\widetilde{S}^{-1}$ and $\widetilde{\Lambda}\widetilde{S}^{-1}$ correspond with the zero diagonal elements of $\widetilde{\Delta}$ and $\widetilde{\Delta}$ respectively. Therefore we can write

$$A = C^\mathsf{T}\widetilde{\Delta}\widetilde{S}^{-1}\widetilde{S}C = C^\mathsf{T}\widetilde{\Delta}\widetilde{S}^{-1}Z^\mathsf{T},$$
$$L = C^\mathsf{T}\widetilde{\Lambda}\widetilde{S}^{-1}\widetilde{S}C = C^\mathsf{T}\widetilde{\Lambda}\widetilde{S}^{-1}Z^\mathsf{T},$$

where $Z = C^\mathsf{T}\widetilde{S}^\mathsf{T} = C^\mathsf{T}\widetilde{S}$. The remainder of the proof then proceeds identically to the proof of Thm. 3.4.1. □

Another situation of interest is having a simultaneous diagonalization of matrices defined by Kronecker products of square matrices with related diagonalizations; the exact situation is described in Lem. 3.4.2.

Lemma 3.4.2 *Given the simultaneous diagonalizations of the $m \times m$ matrices $A_1 = C_1^\mathsf{T}\widetilde{\Delta}_1 C_1$ and $L_1 = C_1^\mathsf{T}\widetilde{\Lambda}_1 C_1$ where C_1 is orthogonal, as well as the simultaneous diagonalizations of the $n \times n$ matrices $A_2 = C_2^\mathsf{T}\widetilde{\Delta}_2 C_2$ and $L_2 = C_2^\mathsf{T}\widetilde{\Lambda}_2 C_2$ where C_2 is orthogonal, suppose that $A = A_1 \otimes A_2$ and $L = L_1 \otimes L_2$ (both A and L are $N \times N$ matrices with $N = mn$). If $\text{null}(A) \cap \text{null}(L) = \{\mathbf{0}\}$, A and L can be expressed as $A = U\Delta X^\mathsf{T}$ and $L = U\Lambda X^\mathsf{T}$, respectively, where U is orthogonal, X is invertible, $\Delta^\mathsf{T}\Delta + \Lambda^\mathsf{T}\Lambda = I_n$, and*

$$0 \leq \Delta_{1,1} \leq \ldots \leq \Delta_{N,N} \leq 1,$$
$$1 \geq \Lambda_{1,1} \geq \ldots \geq \Lambda_{N,N} \geq 0.$$

Proof. We can expand the Kronecker products defining A and L using their diagonalizations:

$$A = A_1 \otimes A_2 = \left(C_1^{\mathsf{T}} \widetilde{\Delta}_1 C_1\right) \otimes \left(C_2^{\mathsf{T}} \widetilde{\Delta}_2 C_2\right) = (C_1 \otimes C_2)^{\mathsf{T}} \left(\widetilde{\Delta}_1 \otimes \widetilde{\Delta}_2\right) (C_1 \otimes C_2),$$
$$L = L_1 \otimes L_2 = \left(C_1^{\mathsf{T}} \widetilde{\Lambda}_1 C_1\right) \otimes \left(C_2^{\mathsf{T}} \widetilde{\Lambda}_2 C_2\right) = (C_1 \otimes C_2)^{\mathsf{T}} \left(\widetilde{\Lambda}_1 \otimes \widetilde{\Lambda}_2\right) (C_1 \otimes C_2).$$

If we let $C = C_1 \otimes C_2$, $\widetilde{\Delta} = \widetilde{\Delta}_1 \otimes \widetilde{\Delta}_2$, and $\widetilde{\Lambda} = \widetilde{\Lambda}_1 \otimes \widetilde{\Lambda}_2$, then C is orthogonal and both $\widetilde{\Delta}$ and $\widetilde{\Lambda}$ are diagonal matrices. The remainder of the proof then proceeds similarly to that of Thm. 3.4.1. □

Perhaps the simplest situation is converting a single diagonalization to an SVD, which is presented as a corollary to Thm. 3.4.1.

Corollary 3.4.2 *Given the diagonalization of the $n \times n$ symmetric matrix $A = C^{\mathsf{T}} \Delta C$ where C is orthogonal, A can be expressed as the SVD $A = U \Sigma V^{\mathsf{T}}$ where U and V are orthogonal and*

$$\Sigma_{1,1} \geq \ldots \geq \Sigma_{n,n} \geq 0.$$

If A is positive semi-definite as well, then A can be written as simply $A = U \Sigma U^{\mathsf{T}}$.

Proof. Let P be the permutation matrix that orders the diagonal elements of Δ in decreasing order in regard to their magnitudes (via the product $P^{\mathsf{T}} \Delta P$), and define another matrix $S = DP$ where D is a diagonal matrix with diagonal entries being ± 1. In particular, we assign -1 as a diagonal entry of D when the corresponding diagonal entry of Δ is negative and assign 1 otherwise. As a result, D is involutary ($D^2 = I_n$). The matrix S is therefore a generalized permutation matrix, specifically a signed permutation matrix, and has the property of orthogonality:

$$S^{\mathsf{T}} S = P^{\mathsf{T}} D^{\mathsf{T}} D P = P^{\mathsf{T}} D^2 P = P^{\mathsf{T}} P = I_n,$$
$$S S^{\mathsf{T}} = D P P^{\mathsf{T}} D^{\mathsf{T}} = D D^{\mathsf{T}} = D^2 = I_n.$$

Using P and S, we can then write A as

$$A = C^{\mathsf{T}}\Delta C = C^{\mathsf{T}}PP^{\mathsf{T}}\Delta SS^{\mathsf{T}}C = U\Sigma V^{\mathsf{T}},$$

where $U = C^{\mathsf{T}}P$, $\Sigma = P^{\mathsf{T}}\Delta S$, and $V = C^{\mathsf{T}}S$. The positivity of the diagonal elements of Σ results from the product ΔD in $\Sigma = P^{\mathsf{T}}\Delta DP$, and the elements are ordered using P. Lastly, in the case where A is positive semi-definite, all diagonal elements of Δ are non-negative; thus D is reduced to the identity matrix and $S = P$, which implies $U = V$. □

As mentioned at the beginning of this section, Thm. 3.4.1 is used to the greatest extend in the numerical implementations of the spectral windowing and learning approaches. Though the property of having $\Delta^{\mathsf{T}}\Delta + \Lambda^{\mathsf{T}}\Lambda = I_n$ is part of Thm. 3.4.1, the property is not used in the primary results of this document but is rather a carry over from the version of GSVD we use in Section 2.3. The benefit of Thm. 3.4.1 is that it provides a connection between the GSVD (which is impractical to form numerically for large scale problems) to the diagonalization of matrices that have the potential for use with numerically practical algorithms. However, we recall that our overarching concern is the regularization of problems of the form $A\mathbf{x} \approx \mathbf{d}$, where the data $\mathbf{d}$ is contaminated by noise $\boldsymbol{\eta}$. With these matrix transformations in mind, consideration is then given to what occurs under these transformations when applied to such noise, white noise in particular; this is discussed next in Section 3.5.

3.5 Transformations of White Noise

We now examine how linear transformations can alter the distribution of a white noise vector $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$. While many of these examinations do not yield much value in the evaluation of the spectral windowing and learning approaches to regularization, they are useful in realizing the complexity of applying relatively simple

transformations to white noise.

It is convenient to provide some basic results that are needed in the later discussion. The first results will be in regard to the real and imaginary parts of $\widehat{\boldsymbol{\eta}} = F\boldsymbol{\eta}$ where F is the unitary $n \times n$ DFT matrix given by 3.3.13. Let F_R and F_I be the real and imaginary parts, respectively, of F (so that $F = F_R + iF_I$). Euler's identity gives

$$\begin{aligned} [F_R]_{j,k} &= \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right), \\ [F_I]_{j,k} &= \frac{1}{\sqrt{n}} \sin\left(\frac{-2\pi(j-1)(k-1)}{n}\right) \end{aligned} \tag{3.5.19}$$

for $1 \leq j, k \leq n$. It is clear from (3.5.19) that the matrices F_R and F_I are symmetric. The rows (and by symmetry, columns) of F_R and F_I have the following orthogonality properties. For compactness, let $\mathbf{e}_j$ denote the n-vector with 1 as the jth component and zeros elsewhere, namely the jth column of the identity matrix I_n.

Lemma 3.5.1 *Let F_R and F_I be the real and imaginary parts, respectively, of the $n \times n$ DFT matrix, and let $\langle \cdot, \cdot \rangle$ denote standard inner product (dot product) of $\mathbb{R}^n$. Then for $1 \leq j, k \leq n$:*

1. $\langle [F_R]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = 0$.

2. *If n is odd, then*

$$\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = \begin{cases} 1 & j = k = 1 \\ \frac{1}{2} & j = k \neq 1 \\ \frac{1}{2} & j + k - 2 = n \\ 0 & \textit{otherwise,} \end{cases} \qquad \langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = \begin{cases} 0 & j = k = 1 \\ \frac{1}{2} & j = k \neq 1 \\ -\frac{1}{2} & j + k - 2 = n \\ 0 & \textit{otherwise.} \end{cases}$$

Equivalently,

$$\begin{aligned} F_R {F_R}^{\mathsf{T}} &= \frac{1}{2} I_n + \frac{1}{2}\left(\mathbf{e}_1 \mathbf{e}_1^{\mathsf{T}} + J_n\right), \\ F_I {F_I}^{\mathsf{T}} &= \frac{1}{2} I_n - \frac{1}{2}\left(\mathbf{e}_1 \mathbf{e}_1^{\mathsf{T}} + J_n\right). \end{aligned}$$

3. If n is even, then

$$\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = \begin{cases} 1 & j = k = 1 \text{ or } j = k = (n/2)+1 \\ \frac{1}{2} & j = k \neq 1 \text{ or } j = k \neq (n/2)+1 \\ \frac{1}{2} & j \neq k \text{ and } j + k - 2 = n \\ 0 & \text{otherwise} \end{cases},$$

$$\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = \begin{cases} 0 & j = k = 1 \text{ or } j = k = (n/2)+1 \\ \frac{1}{2} & j = k \neq 1 \text{ or } j = k \neq (n/2)+1 \\ -\frac{1}{2} & j \neq k \text{ and } j + k - 2 = n \\ 0 & \text{otherwise} \end{cases}.$$

Equivalently,

$$F_R F_R{}^{\mathsf{T}} = \frac{1}{2} I_n + \frac{1}{2} \left(\mathbf{e}_1 \mathbf{e}_1^{\mathsf{T}} + \mathbf{e}_{\frac{n}{2}+1} \mathbf{e}_{\frac{n}{2}+1}^{\mathsf{T}} + J_n \right),$$
$$F_I F_I{}^{\mathsf{T}} = \frac{1}{2} I_n - \frac{1}{2} \left(\mathbf{e}_1 \mathbf{e}_1^{\mathsf{T}} + \mathbf{e}_{\frac{n}{2}+1} \mathbf{e}_{\frac{n}{2}+1}^{\mathsf{T}} + J_n \right).$$

Notice that in both the even and odd cases, $F_R F_R{}^{\mathsf{T}} + F_I F_I{}^{\mathsf{T}} = I$.

Proof. Property (i) will first be established. Letting F_R and F_I be as required, (3.5.19) gives

$$\begin{aligned} \langle [F_R]_{j,\cdot}, [F_I]_{k,\cdot} \rangle &= \frac{1}{n} \sum_{\ell=1}^{n} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n} \right) \sin\left(\frac{-2\pi(\ell-1)(k-1)}{n} \right) \\ &= \frac{1}{n} \sum_{\ell=1}^{n} \cos\left(-\theta_{j\ell}\right) \sin\left(-\theta_{k\ell}\right), \end{aligned} \tag{3.5.20}$$

where $\theta_{j\ell} = 2\pi(j-1)(\ell-1)/n$. Applying the appropriate product-to-sum trigonometric identity and using the fact that sine is an odd function allows for the sum in

(3.5.20) to be evaluated as

$$\begin{aligned}\frac{1}{n}\sum_{\ell=1}^{n}\cos\left(-\theta_{j\ell}\right)\sin\left(-\theta_{k\ell}\right) &= \frac{1}{2n}\sum_{\ell=1}^{n}\left[\sin\left(-\theta_{(j+k-1)\ell}\right)-\sin\left(-\theta_{(j-k+1)\ell}\right)\right]\\ &= \frac{1}{2n}\sum_{\ell=1}^{n}\left[\sin\left(\theta_{(j-k+1)\ell}\right)-\sin\left(\theta_{(j+k-1)\ell}\right)\right]\\ &= \frac{1}{2n}\left(s_1-s_2\right),\end{aligned}$$

where

$$s_1=\sum_{\ell=1}^{n}\sin\left(\theta_{(j-k+1)\ell}\right),\quad s_2=\sum_{\ell=1}^{n}\sin\left(\theta_{(j+k-1)\ell}\right).$$

Both s_1 and s_2 are of the form $\sum_{\ell=1}^{n}\sin(\theta_{p\ell})$, where p is an integer (either $p=j-k+1$ or $p=j+k-1$). This form can be written as

$$\sum_{\ell=1}^{n}\sin\left(\theta_{p\ell}\right)=\sum_{\ell=0}^{n-1}\sin\left(\theta_{p(\ell+1)}\right)=\operatorname{Im}\left(\sum_{\ell=0}^{n-1}\exp\left(i\theta_{p(\ell+1)}\right)\right).$$

If $p-1$ is a multiple of n, then $p=mn+1$ for some $m\in\mathbb{Z}$. In such a case,

$$\exp(i\theta_{p(\ell+1)})=\exp\left(\frac{2\pi i((mn+1)-1)((\ell+1)-1)}{n}\right)=\exp(2\pi im\ell)=1,$$

and so

$$\operatorname{Im}\left(\sum_{\ell=0}^{n-1}\exp\left(i\theta_{p(\ell+1)}\right)\right)=\operatorname{Im}\left(\sum_{\ell=0}^{n-1}1\right)=0.$$

If $p-1$ is not a multiple of n, then

$$\operatorname{Im}\left(\sum_{\ell=0}^{n-1}\exp\left(i\theta_{p(\ell+1)}\right)\right)=\operatorname{Im}\left(\frac{1-\exp(i\theta_{(n+1)p})}{1-\exp(i\theta_p)}\right).$$

However, $\theta_{(n+1)p}=2\pi n(p-1)/n=2\pi(p-1)$. Since $p-1$ is an integer, $\exp(i\theta_{(n+1)p})=\exp(2\pi i(p-1))=1$, meaning that the preceding equation is equal to zero. Therefore

$$\begin{aligned}s_1&=\sum_{\ell=1}^{n}\sin\left(\theta_{(j-k)\ell}\right)=\operatorname{Im}\left(\sum_{\ell=0}^{n-1}\exp\left(i\theta_{(j-k)(\ell+1)}\right)\right)=0,\\ s_2&=\sum_{\ell=1}^{n}\sin\left(\theta_{(j+k-1)\ell}\right)=\operatorname{Im}\left(\sum_{\ell=0}^{n-1}\exp\left(i\theta_{(j+k-1)(\ell+1)}\right)\right)=0\end{aligned}$$

for all $1\le j,k\le n$, which implies $\langle[F_R]_{j,\cdot},[F_I]_{k,\cdot}\rangle=(s_1-s_2)/2n=0$.

The results of properties (ii) and (iii) regarding F_R will now be proved; the results regarding F_I will be handled afterwards. Utilizing (3.5.19) again,

$$\begin{aligned}\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle &= \frac{1}{n} \sum_{\ell=1}^{n} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right) \cos\left(\frac{-2\pi(\ell-1)(k-1)}{n}\right) \\ &= \frac{1}{n} \sum_{\ell=1}^{n} \cos\left(-\theta_{j\ell}\right) \cos\left(-\theta_{k\ell}\right).\end{aligned}$$

By the fact that cosine is an even function and applying another product-to-sum identity gives $\langle [F_R]_{j\cdot}, [F_R]_{k\cdot} \rangle = (c_1 + c_2)/2n$, where

$$c_1 = \sum_{\ell=1}^{n} \cos\left(\theta_{(j-k+1)\ell}\right), \quad c_2 = \sum_{\ell=1}^{n} \cos\left(\theta_{(j+k-1)\ell}\right). \tag{3.5.21}$$

Both c_1 and c_2 are of the form $\sum_{\ell=1}^{n} \cos(\theta_{p\ell})$, where p is an integer (either $p = j-k+1$ or $p = j+k-1$). This form can be written as

$$\sum_{\ell=1}^{n} \cos\left(\theta_{p\ell}\right) = \sum_{\ell=0}^{n-1} \cos\left(\theta_{p(\ell+1)}\right) = \mathrm{Re}\left(\sum_{\ell=0}^{n-1} \exp\left(i\theta_{p(\ell+1)}\right)\right).$$

If $p-1$ is a multiple of n, then $p = mn+1$ for some $m \in \mathbb{Z}$. Thus,

$$\exp(i\theta_{p(\ell+1)}) = \exp\left(\frac{2\pi i((mn+1)-1)((\ell+1)-1)}{n}\right) = \exp(i2\pi m\ell) = 1,$$

and so

$$\mathrm{Re}\left(\sum_{\ell=0}^{n-1} \exp\left(i\theta_{p(\ell)}\right)\right) = \mathrm{Re}\left(\sum_{\ell=0}^{n-1} 1\right) = n.$$

If $p-1$ is not a multiple of n, then

$$\mathrm{Re}\left(\sum_{\ell=0}^{n-1} \exp\left(i\theta_{p(\ell+1)}\right)\right) = \mathrm{Re}\left(\frac{1-\exp(i\theta_{(n+1)p})}{1-\exp(i\theta_p)}\right).$$

However, $\theta_{(n+1)p} = 2\pi(p-1)$. Since p is an integer, $\exp(i\theta_{(n+1)p}) = \exp(2\pi i(p-1)) = 1$, meaning that the preceding equation is equal to zero. Thus 0 and n are the two possible values of c_1 and c_2, which depend upon whether or not $p-1 = j-k$ or $p-1 = j+k-2$ are multiples of n.

First consider property (ii), the case where n is even.

- If $j = k = 1$ or $j = k = (n/2) + 1$, then $j - k = 0$ and $j + k - 2$ is either 0 or n. In other words, both $j - k$ and $j + k$ are multiples of n, and so $c_1 = c_2 = n$. This implies that $\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = (c_1 + c_2)/2n = (n + n)/2n = 1$.

- If $j = k \neq 1$ or $j = k \neq (n/2) + 1$, then $j - k = 0$ but $j + k - 2$ is not a multiple of n. Thus $c_1 = n$ and $c_2 = 0$, and so $\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = (c_1 + c_2)/2n = (n + 0)/2n = 1/2$.

- If $j \neq k$ and $k = n-j+2$, then $j-k$ is not a multiple of n but $j+k-2 = n$. Thus $c_1 = 0$ and $c_2 = n$, and so $\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = (c_1 + c_2)/2n = (0 + n)/2n = 1/2$.

- If $j \neq k$ and $k \neq n - j + 2$, then neither $j - k$ nor $j + k - 2$ are multiples of n. Thus $c_1 = c_2 = 0$, and so $\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = (c_1 + c_2)/2n = (0 + 0)/2n = 0$.

For property (iii), the case where n is odd, $n/2$ need not be considered and the preceding argument holds.

For the results regarding F_I,

$$\begin{aligned}\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle &= \frac{1}{n} \sum_{\ell=1}^{n} \sin\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right) \sin\left(\frac{-2\pi(\ell-1)(k-1)}{n}\right) \\ &= \frac{1}{n} \sum_{\ell=1}^{n} \sin\left(-\theta_{j\ell}\right) \sin\left(-\theta_{k\ell}\right).\end{aligned}$$

By the fact that sine is an odd function and applying another product-to-sum identity gives $\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = (c_1 - c_2)/2n$, where c_1 and c_2 are defined by (3.5.21). Again 0 and n are the two possible values of c_1 and c_2, which depend upon whether or not $j - k$ and $j + k - 2$ are multiples of n.

First consider property (ii), the case where n is even.

- If $j = k = 1$ or $j = k = (n/2) + 1$, then $j - k = 0$ and $j + k - 2$ is either 0 or n. In other words, both $j - k$ and $j + k - 2$ are multiples of n, and so $c_1 = c_2 = n$. This implies that $\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = (c_1 - c_2)/2n = (n - n)/2n = 0$.

- If $j = k \neq 1$ or $j = k \neq (n/2) + 1$, then $j - k = 0$ but $j + k - 2$ is not a multiple of n. Thus $c_1 = n$ and $c_2 = 0$, and so $\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = (c_1 - c_2)/2n = (n - 0)/2n = 1/2$.

- If $j \neq k$ and $k = n-j+2$, then $j-k$ is not a multiple of n but $j+k-2 = n$. Thus $c_1 = 0$ and $c_2 = n$, and so $\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = (c_1 - c_2)/2n = (0 - n)/2n = -1/2$.

- If $j \neq k$ and $k \neq n - j + 2$, then neither $j - k$ nor $j + k - 2$ are multiples of n. Thus $c_1 = c_2 = 0$, and so $\langle [F_I]_{j,\cdot}, [F_I]_{k,\cdot} \rangle = (c_1 - c_2)/2n = (0 + 0)/2n = 0$.

For property (iii), the case where n is odd, $n/2$ need not be considered and again the preceding argument holds. □

There are some important consequences of Lem. 3.5.1. First, it provides a means to visualize the structure of the matrices $F_R^2 = F_R F_R{}^{\mathsf{T}}$ and $F_I^2 = F_I F_I{}^{\mathsf{T}}$. For example, if $n = 6$ then the matrices F_R^2 and F_I^2 are

$$F_R^2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}, \quad F_I^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Second, Lem. 3.5.1 shows that while the rows (and columns) of F form an orthonormal basis for $\mathbb{C}^n$, the orthogonality of the rows of F_R and F_I are limited by conditions on the row indices. Another consequence is that while F is invertible and therefore has

full rank, the matrices F_R and F_I are rank-deficient. Just as the components of F_R and F_I depend upon the parity of n, Lem. 3.5.2 illustrates that the ranks of F_R and F_I depend upon n as well.

Lemma 3.5.2 *Let F_R and F_I be the real and imaginary parts, respectively, of the $n \times n$ DFT matrix. If n is even, then* $\operatorname{rank}(F_R) = (n/2)+1$ *and* $\operatorname{rank}(F_I) = (n/2)-1$. *If n is odd, then* $\operatorname{rank}(F_R) = (n+1)/2$ *and* $\operatorname{rank}(F_I) = (n-1)/2$.

Proof. Let the matrices F_R, and F_I be as required. By applying various properties of the rank of a matrix, all of the effort can be dedicated to finding the rank of F_R. Since the rank of a matrix is equal to the dimension of the row space (or column space), the approach that will be taken is to determine the dimension of the row space of F_R using Lem. 3.5.2. The dimension of the row space will be determined by counting the number of linearly independent rows of F_R.

Assuming that n is even, let $1 \leq j, k \leq (n/2)+1$ with $j \neq k$. From Lem. 3.5.1,

$$\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot} \rangle = \begin{cases} 1/2 & j+k-2 = n \\ 0 & \text{otherwise} \end{cases}.$$

Now for $1 \leq j, \ell \leq (n/2)+1$, we have that $j+k-2 \neq n$ because equality is only achieved if both j and ℓ are $(n/2)+1$ and this would violate the condition that $j \neq \ell$. Thus the first $(n/2)+1$ rows of A form an orthogonal set of vectors, and so they are linearly independent and $\operatorname{rank}(F_R) \geq \frac{n}{2}+1$.

It remains to show that the bound is tight, which we do by showing that the remaining $(n/2)-1$ rows of F_R are copies of the preceding rows. Let $1 \leq j \leq (n/2)+1$, and consider row $\ell = n-j+2$ (assuming $\ell \geq 1$). From (3.5.19) and the parity of

cosine, the components of row ℓ are

$$\begin{aligned}
[F_R]_{\ell,k} &= \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(\ell-1)(k-1)}{n}\right) \\
&= \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(n-j+1)(k-1)}{n}\right) \\
&= \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n} + 2\pi(k-1)\right).
\end{aligned}$$

Since the column index k is an integer, the periodicity of cosine gives $[F_R]_{\ell,k} = [F_R]_{j,k}$. Thus rows j and $n-j+2$ of the matrix F_R are the same for all $1 \leq j \leq (n/2)+1$, and therefore $\text{rank}(F_R) = (n/2)+1$. The argument that $\text{rank}(F_R) = (n+1)/2$ for odd n is identical with the exception that j and ℓ are restricted so $j \neq \ell$ and $1 \leq j, \ell \leq (n+1)/2$.

A similar argument for the rank of F_I could be made, but for the sake of brevity a different approach can be taken which uses some matrix rank inequalities and the result from Lem. 3.5.1 that $F_R {F_I}^{\mathsf{T}}$ is a zero matrix (regardless of the parity of n). Recalling that an invertible matrix has full rank and using the subadditive property of rank,

$$n = \text{rank}(F) = \text{rank}(F_R + iF_I) \leq \text{rank}(F_R) + \text{rank}(iF_I) = \text{rank}(F_R) + \text{rank}(F_I).$$

Applying Sylvester's rank inequality to the product $F_R {F_I}^{\mathsf{T}}$ gives

$$\text{rank}(F_R) + \text{rank}(F_I) - n = \text{rank}(F_R) + \text{rank}({F_I}^{\mathsf{T}}) - n \leq \text{rank}(F_R {F_I}^{\mathsf{T}}) = 0,$$

which implies that $\text{rank}(F_R) + \text{rank}(F_I) \leq n$. Thus $\text{rank}(F_R) + \text{rank}(F_I) = n$, and since the rank of F_R has been determined for both even and odd n, the rank of F_I immediately follows. □

The rank deficiency of F_R and F_I can be used to show, however, that the nullspaces of F_R and F_I do not have non-trivial intersection; this is shown by Lem. 3.5.3.

Lemma 3.5.3 *Let F_R and F_I be the real and imaginary parts, respectively, of the $n \times n$ DFT matrix. Then* $\text{null}(F_R) \cap \text{null}(F_I) = \{\mathbf{0}\}$.

Proof. Regardless of the parity of n, Lem. 3.5.2 implies that both F_R and F_I are singular matrices. Suppose there exists some non-zero vector $\mathbf{u} \in \text{null}(F_R) \cap \text{null}(F_I)$. Then,

$$F\mathbf{u} = (F_R + iF_I)\,\mathbf{u} = F_R\mathbf{u} + iF_I\mathbf{u} = \mathbf{0},$$

which is a contradiction since the DFT matrix F is non-singular. □

The rank deficiency of F_R and F_I has statistical significance as well. Given $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$ and using the properties of the multivariate normal distribution,

$$\text{Re}(\widehat{\boldsymbol{\eta}}) = F_R\eta \sim \mathcal{N}(\mathbf{0}, \sigma^2 F_R {F_R}^{\mathsf{T}}), \quad \text{Im}(\widehat{\boldsymbol{\eta}}) = F_I\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 F_I {F_I}^{\mathsf{T}}). \tag{3.5.22}$$

$F_R {F_R}^{\mathsf{T}}$ and $F_I {F_I}^{\mathsf{T}}$ are rank-deficient and hence they do not have density functions in the traditional sense; their density functions exist in $\text{rank}(F_R {F_R}^{\mathsf{T}})$ and $\text{rank}(F_I {F_I}^{\mathsf{T}})$-dimensional subspaces of $\mathbb{R}^n$. These density functions can be expressed using the pseudoinverses of the covariance matrices or, equivalently, by defining new transformations based on the rank of $F_R {F_R}^{\mathsf{T}}$ and $F_I {F_I}^{\mathsf{T}}$ (Rao, 1973, p. 527-528).

Instead of dealing with density functions of multivariate distributions, the distribution of the components of $\text{Re}(\widehat{\boldsymbol{\eta}})$ and $\text{Im}(\widehat{\boldsymbol{\eta}})$ will be determined individually. As a step towards this goal, the independence of the components of $\text{Re}(\widehat{\boldsymbol{\eta}})$ and $\text{Im}(\widehat{\boldsymbol{\eta}})$ can be established by applying the following result regarding independence of linear combinations of random variables.

Theorem 3.5.1 (Lukacs & King, 1954) *Let $Z_1, \ldots, Z_n$ be independently distributed random variables, and assume that the nth moment of each Z_ℓ exists, i.e. $\mathbb{E}(Z_\ell^n)$ exists for each $\ell = 1, \ldots, n$. The necessary and sufficient conditions for the existence of two statistically independent linear forms $\sum_{\ell=1}^n a_\ell Z_\ell$ and $\sum_{\ell=1}^n b_\ell Z_\ell$ are:*

1. *Each random variable which has a nonzero coefficient in both forms is normally distributed.*

2. $\sum_{\ell=1}^{n} a_\ell b_\ell \sigma_\ell^2 = 0$, *where* σ_ℓ^2 *denotes the variance of* Z_ℓ $(\ell = 1, \ldots, n)$.

Thm. 3.5.1 can be directly applied to show that the components of $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$ retain independence under the DFT, which is stated as Lem. 3.5.4.

Lemma 3.5.4 *Let* $\boldsymbol{\eta}$ *be a random* n*-vector with* $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$, $\mathbf{X} = \mathrm{Re}(\widehat{\boldsymbol{\eta}})$ *and* $\mathbf{Y} = \mathrm{Im}(\widehat{\boldsymbol{\eta}})$. *Then* X_j *and* Y_k *are independent random variables for each* $1 \leq j, k \leq n$.

Proof. Let $\mathbf{X}$ and $\mathbf{Y}$ be as required. For each $1 \leq j, k \leq n$, the components X_j and Y_k are linear combinations of the components of $\boldsymbol{\eta}$, with coefficients given by (3.5.19):

$$X_j = \sum_{\ell=1}^{n} \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right) \eta_\ell, \quad Y_k = \sum_{\ell=1}^{n} \frac{1}{\sqrt{n}} \sin\left(\frac{-2\pi(k-1)(\ell-1)}{n}\right) \eta_\ell.$$

Since the covariance matrix of $\boldsymbol{\eta}$ is $\sigma^2 I_n$, the components of $\boldsymbol{\eta}$ are independent. Furthermore, $\eta_\ell \sim \mathcal{N}(0, \sigma^2)$ and so each component has an nth moment. Thus the first condition of Thm. 3.5.1 is satisfied. As for the final condition, the following sum must be shown to be equal to zero:

$$\sum_{\ell=1}^{n} \left[\frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right)\right] \left[\frac{1}{\sqrt{n}} \sin\left(\frac{-2\pi(k-1)(\ell-1)}{n}\right)\right] \sigma_\ell^2.$$

The η_ℓ are identically distributed with $\sigma_\ell^2 = \sigma^2$ for all $1 \leq \ell \leq n$. Thus

$$\begin{aligned}
&\sum_{\ell=1}^{n} \left[\frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right)\right] \left[\frac{1}{\sqrt{n}} \sin\left(\frac{-2\pi(k-1)(\ell-1)}{n}\right)\right] \sigma_\ell^2 \\
&= \sigma^2 \langle [F_R]_{j,\cdot}, [F_I]_{k,\cdot} \rangle \\
&= 0,
\end{aligned}$$

where the last equality follows from Lem. 3.5.1. Therefore X_j and Y_k are independent for each $1 \leq j, k \leq n$. □

Having established the independence of the components of $\mathrm{Re}(\widehat{\boldsymbol{\eta}})$ and $\mathrm{Im}(\widehat{\boldsymbol{\eta}})$, their individual distributions will be determined.

Lemma 3.5.5 *Let* $\mathbf{X} = \mathrm{Re}(\widehat{\boldsymbol{\eta}})$ *and* $\mathbf{Y} = \mathrm{Im}(\widehat{\boldsymbol{\eta}})$*, where* $\boldsymbol{\eta}$ *is a random* n*-vector with* $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$*. Also let* $J = \{1, \frac{n}{2} + 1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Then for* $1 \leq j \leq n$*, the distribution of* X_j *and* Y_j *is as follows.*

$$X_j \sim \begin{cases} \mathcal{N}(0, \sigma^2), & j \in J \\ \mathcal{N}(0, \sigma^2/2), & j \notin J \end{cases}, \quad Y_j \sim \begin{cases} 0, & j \in J \\ \mathcal{N}(0, \sigma^2/2), & j \notin J \end{cases}.$$

(The notation $Y_j \sim 0$ *means that* Y_j *is a constant random variable with value* 0*.)*

Proof. Let $\mathbf{X}$ and $\mathbf{Y}$ be as required. Using (3.5.19), X_j can be written as

$$X_j = \sum_{k=1}^{n} \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right) \eta_k.$$

Since the covariance matrix of $\boldsymbol{\eta}$ is $\sigma^2 I_n$, the components of $\boldsymbol{\eta}$ are independent and identically distributed $\mathcal{N}(0, \sigma^2)$. Thus by the properties of a sum of independent normal random variables (Casella and Berger, 2002, p. 184),

$$\begin{aligned} X_j &\sim \mathcal{N}\left(0, \sigma^2 \sum_{k=1}^{n} \left(\frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right)\right)^2\right) \\ &= \mathcal{N}\left(0, \sigma^2 \langle [F_R]_{j,\cdot}, [F_R]_{j,\cdot} \rangle\right). \end{aligned}$$

Lem. 3.5.1 then provides the cases for evaluation of the inner product. Determination of the distribution of Y_k is similar. □

Before discussing the distribution of the components of $|\widehat{\boldsymbol{\eta}}|^2$, distributions other than the Gaussian distribution will be introduced. First, the gamma distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$ has the probability density function

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta), \quad x > 0. \tag{3.5.23}$$

A specific case of the gamma distribution is the exponential distribution, for which $\alpha = 1$; the probability density function of the exponential distribution is

$$f(x) = \frac{1}{\beta} \exp(-x/\beta), \quad x > 0, \tag{3.5.24}$$

where again β is considered the scale parameter. Another specific case of the gamma distribution occurs when $\alpha = p/2$ and $\beta = 2$, where p is a positive integer. The resulting distribution is the chi-squared distribution with p degrees of freedom, commonly denoted $\chi^2(p)$. The probability density function of the $\chi^2(p)$ distribution is

$$f(x) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} \exp(-x/2), \quad x > 0. \tag{3.5.25}$$

A generalization of the chi-squared distribution is the noncentral chi-squared distribution. As suggested by the term "noncentral", the noncentral chi-squared distribution has a noncentrality parameter λ in addition to p degrees of freedom; the distribution is denoted by ${\chi'}^2(p, \lambda)$. One representation of the probability density function of the ${\chi'}^2(p, \lambda)$ distribution is found in (Casella and Berger, 2002, p. 166), which is

$$f(x) = \sum_{\ell=0}^{\infty} \frac{x^{(p/2)+\ell-1} \exp(-x/2)}{\Gamma((p/2)+\ell)2^{(p/2)+\ell}} \frac{\lambda^\ell \exp(-\lambda)}{\ell!}, \quad x > 0. \tag{3.5.26}$$

Here the notation for the noncentrality parameter is not to be confused with λ defined in Section 2.3. When $\lambda = 0$, the term λ^ℓ in the series (3.5.26) forces the series to simplify to the $\ell = 0$ term

$$\frac{x^{(p/2)-1} \exp(-x/2)}{\Gamma(p/2)2^{p/2}},$$

which matches (3.5.25). In other words, a zero noncentrality parameter means that the noncentral chi-squared distribution is reduced to the (central) chi-squared distribution, and we use in general ${\chi'}^2(p, \lambda)$.

With the pertinent probability distributions introduced, Thm. 3.5.2 summarizes the results regarding the distribution of the components of $|\widehat{\boldsymbol{\eta}}|^2$.

Theorem 3.5.2 *Let $\boldsymbol{\eta}$ be a random n-vector with $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$. Also let $J = \{1, \frac{n}{2}+1\}$ if n is even and $J = \{1\}$ if n is odd. Then, for $1 \leq j \leq n$, the distribution of $|\widehat{\eta}_j|^2 = \mathrm{Re}(\widehat{\eta}_j)^2 + \mathrm{Im}(\widehat{\eta}_j)^2$ is as follows.*

$$|\widehat{\eta}_j|^2 \sim \begin{cases} \text{gamma}(1/2, 2\sigma^2), & j \in J \\ \text{exponential}(\sigma^2), & j \notin J \end{cases}.$$

Proof. Let $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$, $\mathbf{X} = \mathrm{Re}(\widehat{\boldsymbol{\eta}})$, and $\mathbf{Y} = \mathrm{Im}(\widehat{\boldsymbol{\eta}})$. The proof relies on determining the distribution of the components of $\mathbf{X}^2$ and $\mathbf{Y}^2$ (here the exponent indicates that the components are individually squared) and then looking at their sum. Determination of the distribution of the components of $\mathbf{X}^2$ and $\mathbf{Y}^2$ is carried out by applying a result regarding univariate one-to-one transformations (Casella and Berger, 2002, p. 53). Without loss of generality, we assume that n is odd.

First consider the components of $\mathbf{X}^2$. Since n is odd, the two cases of X_j^2 to be considered are for $j = 1$ and $j \neq 1$, both of which are handled by Lem. 3.5.5. If $j = 1$, then $X_j \sim \mathcal{N}(0, \sigma^2)$. Letting $g(x) = x^2$, the transformation $U = g(X_j) = X_j^2$ is not one-to-one on the entire sample space $\mathbb{R}$. However, $\mathbb{R}$ can be partitioned as $A_0 \cup A_1 \cup A_2$ with $A_0 = \{0\}$, $A_1 = (-\infty, 0)$, and $A_2 = (0, \infty)$. Defining $g_1(x) = x^2$ and $g_2(x) = x^2$ with $g_1^{-1}(u) = -\sqrt{u}$ and $g_2^{-1}(u) = \sqrt{u}$,

1. $g(x) = g_1(x) = g_2(x)$ for all $x \in A_1 \cup A_2$,

2. g_1 and g_2 are monotone on A_1 and A_2, respectively,

3. $g_1(A_1) = g_2(A_2) = (0, \infty)$, and

4. the derivatives g_1^{-1} and g_2^{-1} are continuous on $(0, \infty)$.

Lastly the set A_0 is of no concern since $P(X_j \in A_0) = P(X_j = 0) = 0$. Then using the probability density function of the normal distribution, the density function of U

is

$$\begin{aligned} f_U(u) &= \frac{1}{\sqrt{2\pi\sigma^2}}\left[\exp\left(\frac{-(g_1^{-1}(u))^2}{2\sigma^2}\right)\left|\frac{d}{du}g_1^{-1}(u)\right| + \exp\left(\frac{-(g_2^{-1}(u))^2}{2\sigma^2}\right)\left|\frac{d}{du}g_2^{-1}(u)\right|\right] \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-u}{2\sigma^2}\right)\left|\frac{-1}{2\sqrt{u}}\right| + \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-u}{2\sigma^2}\right)\left|\frac{1}{2\sqrt{u}}\right| \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\frac{1}{\sqrt{u}}\exp\left(\frac{-u}{2\sigma^2}\right). \end{aligned}$$

This is the probability density function of the gamma distribution with shape parameter $1/2$ and scale parameter $2\sigma^2$. The same argument holds for $j \neq 1$, with scale parameter instead being σ^2. The argument can also be applied for Y_j^2 when $j \neq 1$ since X_j and Y_j are identically distributed in this case. When $j = 1$, Y_j is a constant random variable with $Y_j = 0$, and so $Y_j^2 = 0$ as well.

Now that the distribution of X_j^2 and Y_j^2 is known, the distribution of their sum can be established. If $j = 1$, then $X_j^2 + Y_j^2$ has the same distribution as just X_j^2. The situation is more interesting when $j \neq 1$ since Y_j^2 is no longer a constant random variable. By Lem. 3.5.4, X_j and Y_j are independent for all $1 \leq j \leq n$. As a consequence, X_j^2 and Y_j^2 are independent for all $1 \leq j \leq n$. Let $f_{X_j^2}$ and $f_{Y_j^2}$ denotes the probability density functions of X_j^2 and Y_j^2, respectively. Since the probability density function of a sum of two independent continuous random variables is equal to the convolution of their individual density functions (Casella and Berger, 2002, p. 215), the density function of $V = X_j^2 + Y_j^2$ is given by

$$f_V(v) = \int_{-\infty}^{\infty} f_{X_j^2}(w) f_{Y_j^2}(v - w)\, dw.$$

Fortunately X_j^2 and Y_j^2 are non-negative, meaning that the interval of integration of the convolution can be reduced; $f_{X_j^2}(w) = 0$ for $w < 0$ and $f_{Y_j^2}(v - w) = 0$ for $w > v$ implies an interval of integration of $[0, v]$. Using the density functions and the

substitution $t = w/v$ then gives

$$\begin{aligned} f_V(v) &= \int_0^v \left[\frac{1}{\sqrt{\pi\sigma^2}}\frac{1}{\sqrt{w}}\exp\left(\frac{-w}{\sigma^2}\right)\right]\left[\frac{1}{\sqrt{\pi\sigma^2}}\frac{1}{\sqrt{v-w}}\exp\left(\frac{-(v-w)}{\sigma^2}\right)\right]\,dw \\ &= \frac{1}{\pi\sigma^2}\exp\left(\frac{-v}{\sigma^2}\right)\int_0^v \frac{1}{\sqrt{w}}\frac{1}{\sqrt{v-w}}\,dw \\ &= \frac{1}{\pi\sigma^2}\exp\left(\frac{-v}{\sigma^2}\right)\int_0^1 \frac{1}{\sqrt{vt}}\frac{1}{\sqrt{v-vt}}v\,dt \\ &= \frac{1}{\pi\sigma^2}\exp\left(\frac{-v}{\sigma^2}\right)\int_0^1 \frac{1}{\sqrt{t}}\frac{1}{\sqrt{1-t}}\,dt. \end{aligned}$$

The last integral represents $B(1/2, 1/2)$, the beta function (Abramowitz and Stegun, 1972, p. 258) evaluated at $(1/2, 1/2)$. Since $B(1/2, 1/2) = \pi$, the probability density function of V_j is

$$f_V(v) = \frac{1}{\sigma^2}\exp\left(\frac{-v}{\sigma^2}\right).$$

This is the density function of the exponential distribution with scale parameter σ^2. □

The statistics of the DFT of white noise are now established, which allows the focus to be shifted towards analyzing the combination $\mathbf{b}+\boldsymbol{\eta} = \mathbf{d}$. By the properties of the multivariate normal distribution, if $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$ then $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$ because $\mathbf{b}$ is a constant vector. Thus, the distribution of the components of $\mathrm{Re}\left(\widehat{\mathbf{d}}\right)$ and $\mathrm{Im}\left(\widehat{\mathbf{d}}\right)$ are readily obtained by extending previous results.

Lemma 3.5.6 (Extension of Lem. 3.5.5) *Let* $\mathbf{X} = \mathrm{Re}\left(\widehat{\mathbf{d}}\right)$ *and* $\mathbf{Y} = \mathrm{Im}\left(\widehat{\mathbf{d}}\right)$, *where* $\mathbf{d}$ *is a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$. *Also let* $J = \{1, \frac{n}{2}+1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Then, for* $1 \le j \le n$, *the distribution of* X_j *and* Y_j *is as follows.*

$$X_j \sim \begin{cases} \mathcal{N}\left(\mathrm{Re}\left(\widehat{b}_j\right), \sigma^2\right), & j \in J \\ \mathcal{N}\left(\mathrm{Re}\left(\widehat{b}_j\right), \sigma^2/2\right), & j \notin J \end{cases}, \quad Y_j \sim \begin{cases} 0, & j \in J \\ \mathcal{N}\left(\mathrm{Im}\left(\widehat{b}_j\right), \sigma^2/2\right), & j \notin J \end{cases}.$$

Proof. Let $\mathbf{X}$ and $\mathbf{Y}$ be as required. Using (3.5.19), X_j can be written as

$$X_j = \sum_{k=1}^{n} \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right) d_k.$$

Since the covariance matrix of $\mathbf{d}$ is $\sigma^2 I$, the components d_k are independent and distributed $\mathcal{N}(b_k, \sigma^2)$ for all $1 \leq k \leq n$. Thus by the properties of a sum of independent normal random variables (Casella and Berger, 2002, p. 184),

$$\begin{aligned} X_j &\sim \mathcal{N}\left(\sum_{k=1}^{n} \frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right) b_k, \right. \\ &\qquad \left. \sigma^2 \sum_{k=1}^{n} \left(\frac{1}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(k-1)}{n}\right)\right)^2\right) \\ &= \mathcal{N}\left(\mathrm{Re}\left(\widehat{b}_j\right), \sigma^2 \langle [F_R]_{j\cdot}, [F_R]_{j\cdot} \rangle\right) \\ &= \mathcal{N}\left(\mathrm{Re}\left(\widehat{b}_j\right), \sigma^2 [F_R {F_R}^{\mathsf{T}}]_{jj}\right). \end{aligned}$$

Lem. 3.5.1 then provides the cases for evaluation of the inner product. Determination of the distribution of Y_k is similar. □

In contrast, however, to the results regarding the components of $|\widehat{\boldsymbol{\eta}}|^2$, the distribution of the components of $\left|\widehat{\mathbf{d}}\right|^2$ is somewhat complicated. To illustrate this, consider $U = X_1^2$, where $\mathbf{X} = \mathrm{Re}\left(\widehat{\mathbf{d}}\right)$ for odd-length $\mathbf{d}$ distributed $\mathcal{N}(\mathbf{b}, \sigma^2 I_n)$. To simplify notation, let $\boldsymbol{\mu} = \mathrm{Re}\left(\widehat{\mathbf{b}}\right)$ so that by Lem. 3.5.6, $X_1 \sim \mathcal{N}(\mu_1, \sigma^2)$. Applying the same transformation technique from Thm. 3.5.2, the probability density function of U is

$$\begin{aligned} f_U(u) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(-\sqrt{u} - \mu_1)^2}{2\sigma^2}\right) \left|\frac{-1}{2\sqrt{u}}\right| + \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\sqrt{u} - \mu_1)^2}{2\sigma^2}\right) \left|\frac{1}{2\sqrt{u}}\right| \\ &= \frac{1}{2\sqrt{2\pi\sigma^2}} \frac{1}{\sqrt{u}} \left[\exp\left(\frac{-(\sqrt{u} + \mu_1)^2}{2\sigma^2}\right) + \exp\left(\frac{-(\sqrt{u} - \mu_1)^2}{2\sigma^2}\right)\right]. \end{aligned}$$

Unfortunately, this density function is not easily identified. However, rescaling X_1 as X_1/σ before applying the square transformation results in a random variable that has a more tractable density function. The density function of $V = (X_1/\sigma)^2$ can be

equated to that of the noncentered chi-squared distribution with 1 degree of freedom and noncentrality parameter $\lambda = (\mu_1/\sigma)^2/2$.

Before the distribution of the scaled components of $\widehat{\mathbf{d}}$ is stated, Lem. 3.5.7 establishes the independence of the (squared) real and imaginary parts.

Lemma 3.5.7 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$*,* $J = \{1, \frac{n}{2} + 1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Define the diagonal matrix* M *by*

$$M_{j,j} = \begin{cases} 1/\sigma & j \in J \\ \sqrt{2}/\sigma & j \notin J \end{cases}. \tag{3.5.27}$$

Then, $\mathrm{Re}\left((M\widehat{\mathbf{d}})_j\right)^2$ *and* $\mathrm{Im}\left((M\widehat{\mathbf{d}})_k\right)^2$ *are independent for all* $1 \leq j, k \leq n$*.*

Proof. Let $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I)$, $\mathbf{X} = \mathrm{Re}\left(M\widehat{\mathbf{d}}\right)$, and $\mathbf{Y} = \mathrm{Im}\left(M\widehat{\mathbf{d}}\right)$. Since $\widehat{\mathbf{d}} = \widehat{\mathbf{b}} + \widehat{\boldsymbol{\eta}}$, X_j^2 and Y_k^2 can be expressed as

$$X_j^2 = \left[M_{j,j}\left(\mathrm{Re}(\widehat{b}_j) + \mathrm{Re}(\widehat{\eta}_j)\right)\right]^2, \quad Y_k^2 = \left[M_{k,k}\left(\mathrm{Im}(\widehat{b}_k) + \mathrm{Im}(\widehat{\eta}_k)\right)\right]^2.$$

X_j^2 and Y_k^2 are thus functions of only the random variables $\mathrm{Re}(\widehat{\eta}_j)$ and $\mathrm{Im}(\widehat{\eta}_k)$, respectively. $\mathrm{Re}(\widehat{\eta}_j)$ and $\mathrm{Im}(\widehat{\eta}_k)$ are independent by Lem. 3.5.4, and therefore $X_j^2 = \mathrm{Re}\left((M\widehat{\mathbf{d}})_j\right)^2$ and $Y_k^2 = \mathrm{Im}\left((M\widehat{\mathbf{d}})_k\right)^2$ are independent as well. □

The distribution of the scaled components of $\widehat{\mathbf{d}}$ can now be stated.

Theorem 3.5.3 (Extension of Thm. 3.5.2) *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$*,* $J = \{1, \frac{n}{2} + 1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Define the diagonal matrix* M *by*

$$M_{j,j} = \begin{cases} 1/\sigma & j \in J \\ \sqrt{2}/\sigma & j \notin J \end{cases}.$$

Then for $1 \leq j \leq n$*, the distribution of* $\left|(M\widehat{\mathbf{d}})_j\right|^2 = \mathrm{Re}\left((M\widehat{\mathbf{d}})_j\right)^2 + \mathrm{Im}\left((M\widehat{\mathbf{d}})_j\right)^2$ *is*

$$\left|(M\widehat{\mathbf{d}})_j\right|^2 \sim \begin{cases} {\chi'}^2\left(1, \frac{1}{2}\left(\frac{\mathrm{Re}(\widehat{b}_j)}{\sigma}\right)^2\right), & j \in J \\ {\chi'}^2\left(2, \frac{|\widehat{b}_j|^2}{\sigma^2}\right), & j \notin J \end{cases},$$

where ${\chi'}^2(k, \lambda)$ *denotes the noncentral chi-squared distribution with* k *degrees of freedom and noncentrality parameter* λ*.*

Proof. Let $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I)$, $\mathbf{X} = \mathrm{Re}\left(M\widehat{\mathbf{d}}\right)$, and $\mathbf{Y} = \mathrm{Im}\left(M\widehat{\mathbf{d}}\right)$. Again without loss of generality, we assume that n is odd so that the two cases to be considered are $j = 1$ and $j \neq 1$.

First, the components of $\mathbf{X}^2$ will be determined. If $j = 1$, then $\mathrm{Re}(\widehat{d}_j) \sim \mathcal{N}(\mu_j, \sigma^2)$ from Lem. 3.5.6, where $\boldsymbol{\mu} = \mathrm{Re}(\widehat{\mathbf{b}})$ for readability. Thus, by the properties of the normal distribution (Casella and Berger, 2002, p. 184), $\mathrm{Re}\left(\widehat{d}_j/\sigma\right) = \mathrm{Re}\left((M\widehat{\mathbf{d}})_j\right) = X_j \sim \mathcal{N}(\sqrt{2\lambda}, 1)$ where $\lambda = (\mu_j/\sigma)^2/2$. Applying the transformation technique used in Thm. 3.5.2, the probability density function of $V = X_j^2$ is

$$f_V(v) = \frac{1}{2\sqrt{2\pi}}\frac{1}{\sqrt{v}}\left[\exp\left(\frac{-(\sqrt{v}+\sqrt{2\lambda})^2}{2}\right) + \exp\left(\frac{-(\sqrt{v}-\sqrt{2\lambda})^2}{2}\right)\right].$$

Expanding the arguments of the exponential terms allows for the function to be rewritten as

$$\begin{aligned} f_V(v) &= \frac{1}{2\sqrt{2\pi}}\frac{1}{\sqrt{v}}\left[\exp\left(\frac{-(\sqrt{v}+\sqrt{2\lambda})^2}{2}\right) + \exp\left(\frac{-(\sqrt{v}-\sqrt{2\lambda})^2}{2}\right)\right] \\ &= \frac{1}{2\sqrt{2\pi}}\frac{1}{\sqrt{v}}\left[\exp\left(-\frac{v}{2} - \sqrt{2\lambda v} - \lambda\right) + \exp\left(-\frac{v}{2} + \sqrt{2\lambda v} - \lambda\right)\right] \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{v}}\exp\left(\frac{-v}{2} - \lambda\right)\left[\frac{\exp(-\sqrt{2\lambda v}) + \exp(\sqrt{2\lambda v})}{2}\right] \\ &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{v}}\exp\left(\frac{-v}{2} - \lambda\right)\cosh(\sqrt{2\lambda v}). \end{aligned}$$

Hyperbolic cosine is an entire function with $\cosh(z) = \sum_{\ell=0}^{\infty} z^{2\ell}/(2\ell)!$ as its Taylor

expansion. From (Abramowitz and Stegun, 1972, p. 255),

$$\Gamma\left(\ell+\frac{1}{2}\right)=\frac{1\cdot 3\cdot 5\cdot 7\cdot\cdots\cdot(2\ell-1)}{2^{\ell}}\Gamma\left(\frac{1}{2}\right)=\frac{(2\ell-1)!!}{2^{\ell}}\sqrt{\pi} \tag{3.5.28}$$

for all integers ℓ. Thus, as an intermediate step, the double factorial $(2\ell-1)!!$ must be related to $(2\ell)!$ in order to modify the density function to the desired form. This is accomplished by using the relation

$$(2\ell-1)!!2^{\ell}\ell!=(2\ell)!, \tag{3.5.29}$$

which is validated by noting that

$$\begin{aligned}
(2\ell-1)!!2^{\ell}\ell! &= \left[(2\ell-1)(2\ell-3)(2\ell-5)\cdots(1)\right]2^{\ell}\left[(\ell)(\ell-1)(\ell-2)\cdots(1)\right] \\
&= \left[(2\ell-1)(2\ell-3)(2\ell-5)\cdots(1)\right]\left[(2\ell)(2(\ell-1))(2(\ell-2))\cdots(2)\right] \\
&= \left[(2\ell-1)(2\ell-3)(2\ell-5)\cdots(1)\right]\left[(2\ell)(2\ell-2)(2\ell-4)\cdots(2)\right] \\
&= (2\ell)(2\ell-1)(2\ell-2)(2\ell-3)(2\ell-4)(2\ell-5)\cdots(2)(1) \\
&= (2\ell)!.
\end{aligned}$$

In light of (3.5.29), (3.5.28) becomes the identity $\Gamma(\ell+1/2)=(2\ell)!\sqrt{\pi}/4^{\ell}\ell!$. Therefore,

$$\begin{aligned}
f_V(v) &= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{v}}\exp\left(\frac{-v}{2}-\lambda\right)\sum_{\ell=0}^{\infty}\frac{(\sqrt{2\lambda v})^{2\ell}}{(2\ell)!} \\
&= \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{v}}\exp\left(\frac{-v}{2}-\lambda\right)\sum_{\ell=0}^{\infty}\frac{(2\lambda v)^{\ell}\sqrt{\pi}}{\Gamma(n+1/2)4^{n}n!} \\
&= \sum_{\ell=0}^{\infty}\exp\left(\frac{-v}{2}-\lambda\right)\frac{(\lambda v)^{\ell}v^{-1/2}}{\Gamma(n+1/2)2^{n}n!\sqrt{2}} \\
&= \sum_{\ell=0}^{\infty}\frac{v^{\ell-1/2}\exp(-v/2)\lambda^{\ell}\exp(-\lambda)}{\Gamma(\ell+1/2)2^{\ell+1/2}\ell!},
\end{aligned}$$

matching the probability density function (3.5.26) of the noncentral chi-squared distribution with 1 degree of freedom and noncentrality parameter $\lambda=(\mu_j/\sigma)^2/2=\left(\mathrm{Re}(\widehat{b}_j)/\sigma\right)^2/2$.

If $j \neq 1$, then Lem. 3.5.6 gives that $\mathrm{Re}\left(\widehat{d}_j\right) \sim \mathcal{N}(\mathrm{Re}(\widehat{b}_j), \sigma^2/2)$. This implies that $\mathrm{Re}\left((\sqrt{2}\widehat{d}/\sigma)_j\right) = \mathrm{Re}\left((M\widehat{\mathbf{d}})_j\right) = X_j \sim \mathcal{N}(\sqrt{2\lambda}, 1)$, now with $\lambda = (\mu_j/\sigma)^2$. Thus the same argument can be applied so that $X_j^2 \sim \chi'^{\,2}(1, (\mu_j/\sigma)^2) = \chi'^{\,2}\left(1, (\mathrm{Re}(\widehat{b}_j)/\sigma)^2\right)$.

With the components of $\mathbf{X}^2$ examined, focus can directed towards the components of $\mathbf{Y}^2$. If $j = 1$, then Lem. 3.5.6 gives that $\mathrm{Im}\left(\widehat{d}_j\right) = 0$ (a constant random variable). This implies that $\mathrm{Im}\left(\widehat{d}_j/\sigma\right) = \mathrm{Im}\left((M\widehat{\mathbf{d}})_j\right) = Y_j = 0$, meaning $Y_j^2 = 0$ as well. If $j \neq 1$, then $\mathrm{Im}\left(\widehat{d}_j\right) \sim \mathcal{N}(\mathrm{Im}(\widehat{b}_j), \sigma^2/2)$. Thus by the argument for X_j, $Y_j^2 \sim \chi'^{\,2}\left(1, (\mathrm{Im}(\widehat{b}_j)/\sigma)^2\right)$.

The final part of the proof is to establish the distribution of $X_j^2 + Y_j^2$ for the two cases of j. If $j = 1$, then $Y_j^2 = 0$ and so $X_j^2 + Y_j^2$ has the same distribution of X_j^2, which is $\chi'^{\,2}\left(1, (\mathrm{Re}(\widehat{b}_j)/\sigma)^2/2\right)$. If $j \neq 1$, then X_j^2 and Y_j^2 are distributed $\chi'^{\,2}\left(1, (\mathrm{Re}(\widehat{b}_j)/\sigma)^2\right)$ and $\chi'^{\,2}\left(1, (\mathrm{Im}(\widehat{b}_j)/\sigma)^2\right)$, respectively. Their independence is given by Lem. 3.5.7, and so the reproductive property of the noncentral chi-squared distribution (Rao, 1973, p. 182) produces

$$X_j^2 + Y_j^2 \sim \chi'^{\,2}\left(1+1, \left(\frac{\mathrm{Re}(\widehat{b}_j)}{\sigma}\right)^2 + \left(\frac{\mathrm{Im}(\widehat{b}_j)}{\sigma}\right)^2\right) = \chi'^{\,2}\left(2, \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right).$$

□

From Thm. 3.5.3, there are conditions which simplify the distribution of $\left|(M\widehat{\mathbf{d}})_j\right|^2$. If $j = 1$ (or $(n/2)+1$ when n is even), the noncentrality parameter of the noncentral chi-squared distribution is $\left(\mathrm{Re}(\widehat{b}_1)\right)^2/2\sigma^2$. By the definition of the DFT,

$$\mathrm{Re}\left(\widehat{b}_1\right) = \frac{1}{\sqrt{n}}\sum_{\ell=1}^{n}\cos\left(\frac{-2\pi(1-1)(\ell-1)}{n}\right)b_\ell = \frac{1}{\sqrt{n}}\sum_{\ell=1}^{n} b_\ell.$$

It is certainly possible that $\sum_{\ell=1}^{n} b_\ell = 0$. A situation that would lend itself to this possibility would be a function $g(t)$ whose integral over the interval being considered is zero, e.g. $\sin(2\pi t)$ defined on the interval $[0, 1]$. Even if $\mathrm{Re}\left(\widehat{b}_1\right)$ is nonzero, $\lambda = \left(\mathrm{Re}(\widehat{b}_1)\right)^2/2\sigma^2$ can be near zero when $\mathrm{Re}\left(\widehat{b}_1\right)$ is small or the variance σ^2 is large.

Extending these observations for general indices j, a zero noncentrality parameter means

$$\left|(M\widehat{\mathbf{d}})_j\right|^2 \sim \begin{cases} \chi'^{\,2}(1,0) = \chi^2(1), & j = 1 \\ \chi'^{\,2}(2,0) = \chi^2(2), & \text{otherwise} \end{cases}, \tag{3.5.30}$$

again for n being even; the same holds for odd n, with the inclusion of the $(n/2)+1$ case. Recalling that M in Thm. 3.5.3 is a diagonal matrix, (3.5.30) can be restated as

$$\left|(M\widehat{\mathbf{d}})_j\right|^2 = \begin{cases} |\widehat{d}_j|^2/\sigma^2 \sim \chi^2(1), & j = 1 \\ 2|\widehat{d}_j|^2/\sigma^2 \sim \chi^2(2), & \text{otherwise} \end{cases}.$$

Fortunately, this result agrees with Thm. 3.5.2. The connection can be stated as a lemma.

Lemma 3.5.8 *Let $\sigma^2 > 0$, $Z_1 \sim \chi^2(1)$, and $Z_2 \sim \chi^2(2)$. Then $V_1 = \sigma^2 Z_1 \sim$ gamma$(1/2, 2\sigma^2)$ and $V_2 = \sigma^2 Z_1/2 \sim$ exponential(σ^2).*

Proof. Let $\sigma^2 > 0$, $Z_1 \sim \chi^2(1)$, $Z_2 \sim \chi^2(2)$, and define $V_1 = \sigma^2 Z_1$ and $V_2 = \sigma^2 Z_1/2$. In addition, let $g_1(z) = \sigma^2 z$ and $g_2(z) = \sigma^2 z/2$. Then $V_1 = g_1(Z_1)$ and $V_2 = g_2(Z_2)$, both g_1 and g_2 are monotone on the sample space $(0, \infty)$ of chi-squared random variables, and $g_1^{-1}(v) = v/\sigma^2$ and $g_2^{-1}(v) = 2v/\sigma^2$. Using g_1^{-1} and the probability density function of the $\chi^2(1)$ distribution, the density function of V_1 is

$$f_{V_1}(v) = \frac{1}{\Gamma(1/2)\sqrt{2}} \left(\frac{v}{\sigma^2}\right)^{-1/2} \exp\left(\frac{-v}{2\sigma^2}\right) \left|\frac{1}{\sigma^2}\right| = \frac{1}{\Gamma(1/2)\sqrt{2\sigma^2}} v^{-1/2} \exp\left(\frac{-v}{2\sigma^2}\right),$$

which is the density function of the gamma$(1/2, 2\sigma^2)$ distribution. Similarly using g_2^{-1} and the probability density function of the $\chi^2(2)$ distribution, the density function of V_2 is

$$f_{V_2}(v) = \frac{1}{\Gamma(1) \cdot 2} \exp\left(\frac{-2v}{2\sigma^2}\right) \left|\frac{2}{\sigma^2}\right| = \frac{1}{\sigma^2} \exp\left(\frac{-v}{\sigma^2}\right),$$

which is the density function of the exponential(σ^2) distribution. An alternative method for showing $V_2 \sim$ exponential(σ^2) would be to note that the $\chi^2(2)$ distribution is the same as the exponential(1/2) distribution and then perform the scalar transformation. □

Ultimately there is a trade-off between the frequency content of the blurred function and the variance in the added noise. If the blurred function has a small amount of high frequency content (relative to the variance of the noise), then the statistics of corresponding terms $\left|(M\widehat{\mathbf{d}})_j\right|^2$ will resemble those of chi-squared random variables.

Since Thm. 3.5.3 is used in Section 4.1 and Section 4.2, some statistics regarding the noncentral chi-squared distribution will be stated for later convenience.

Theorem 3.5.4 *Let* $X \sim \chi'^{\,2}(p, \lambda)$. *Then* $\mathbb{E}(X) = p + \lambda$ *and* $\mathrm{Var}(X) = 2p + 4\lambda$.

Proof. Let $X \sim \chi'^{\,2}(p, \lambda)$. As noted in (Casella and Berger, 2002, p. 167), the probability density function (3.5.26) of the noncentral chi-squared distribution can be considered a mixture distribution; for the hierarchy $X|Y \sim \chi^2(p + 2Y)$ and $Y \sim$ Poisson(λ), the marginal distribution of X is (3.5.26). Then by the properties of expected value and the chi-squared and Poisson distributions,

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|Y)) = \mathbb{E}(p + 2Y) = p + 2\,\mathbb{E}(Y) = p + 2\lambda.$$

The variance of X is calculated in a similar way:

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}(\mathrm{Var}(X|Y)) + \mathrm{Var}(\mathbb{E}(X|Y)) \\
&= \mathbb{E}(2p) + \mathrm{Var}(p + 2Y) \\
&= 2p + 4\,\mathrm{Var}(Y) \\
&= 2p + 4\lambda.
\end{aligned}$$

□

Using the results from Thm. 3.5.3, Thm. 3.5.4 has the following corollary.

Corollary 3.5.1 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$*,* $J = \{1, \frac{n}{2} + 1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Define the diagonal matrix* M *by*

$$M_{j,j} = \begin{cases} 1/\sigma & j \in J \\ \sqrt{2}/\sigma & j \notin J \end{cases}.$$

Then

$$\mathbb{E}\left(\left|(M\widehat{\mathbf{d}})_j\right|^2\right) = \begin{cases} \frac{1}{2}\left(2 + \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right), & j \in J \\ 2 + \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}, & j \notin J \end{cases}$$

and

$$\mathrm{Var}\left(\left|(M\widehat{\mathbf{d}})_j\right|^2\right) = \begin{cases} 2\left(1 + \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right), & j \in J \\ 4\left(1 + \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right), & j \notin J \end{cases}.$$

The final statistical result presented here is the covariance of $\left|M_j\widehat{d}_j\right|^2$ and $\left|M_k\widehat{d}_k\right|^2$. A step in Thm. 3.5.5 is to determine covariances between components of $M\widehat{\mathbf{d}} = M\widehat{\mathbf{b}} + M\widehat{\boldsymbol{\eta}}$. Lem. 3.5.7 establishes that the real and imaginary parts of $M\widehat{\boldsymbol{\eta}}$ are independent, but Lem. 3.5.9 establishes the relationship between any two real or imaginary components of $M\widehat{\boldsymbol{\eta}}$.

Lemma 3.5.9 *Let* $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$*,* F_R *and* F_I *be the real and imaginary parts, respectively, of the* $n \times n$ *DFT matrix,* $J = \{1, \frac{n}{2} + 1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Define the diagonal matrix* M *by*

$$M_{j,j} = \begin{cases} 1/\sigma & j \in J \\ \sqrt{2}/\sigma & j \notin J \end{cases}.$$

Then $(MF_R\boldsymbol{\eta})_j$ *and* $(MF_R\boldsymbol{\eta})_k$ *are independent for all* $1 \leq j < k \leq n$ *with* $j+k-2 \neq n$. *When* $j+k-2 = n$, $(MF_R\boldsymbol{\eta})_j = (MF_R\boldsymbol{\eta})_k$. *The same result holds for the components of* $MF_I\boldsymbol{\eta}$.

Proof. Let $1 \leq j < k \leq n$ with $j + k - 2 \neq n$. The components $(MF_R\boldsymbol{\eta})_j$ and $(MF_R\boldsymbol{\eta})_k$ are

$$(MF_R\boldsymbol{\eta})_j = \sum_{\ell=1}^{n} \frac{M_{j,j}}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right)\eta_\ell,$$
$$(MF_R\boldsymbol{\eta})_k = \sum_{\ell=1}^{n} \frac{M_{k,k}}{\sqrt{n}} \cos\left(\frac{-2\pi(k-1)(\ell-1)}{n}\right)\eta_\ell.$$

As stated in Lem. 3.5.4, the η_ℓ are independent and their nth moments exist for each $1 \leq \ell \leq n$. Thus the first condition of Thm. 3.5.1 is satisfied. The second condition follows from

$$\begin{aligned}
&\sum_{\ell=1}^{n}\left[\frac{M_{j,j}}{\sqrt{n}} \cos\left(\frac{-2\pi(j-1)(\ell-1)}{n}\right)\right]\left[\frac{M_{k,k}}{\sqrt{n}} \cos\left(\frac{-2\pi(k-1)(\ell-1)}{n}\right)\right] \\
&= M_{j,j}M_{k,k}\langle [F_R]_{j,\cdot}, [F_R]_{k,\cdot}\rangle \\
&= 0,
\end{aligned}$$

where the last equality is given by Lem. 3.5.1. The proof of $(MF_R\boldsymbol{\eta})_j = (MF_R\boldsymbol{\eta})_k$ when $j + k - 2 = n$ is contained in the proof of Lem. 3.5.2. The same argument applies to $(MF_I\boldsymbol{\eta})_j$ and $(MF_I\boldsymbol{\eta})_k$. □

The covariance of $|M_j\widehat{d}_j|^2$ and $|M_k\widehat{d}_k|^2$ for $1 \leq j < k \leq n$ can now be determined.

Theorem 3.5.5 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$, $J = \{1, \frac{n}{2}+1\}$ *if* n *is even and* $J = \{1\}$ *if* n *is odd. Define the diagonal matrix* M *by*

$$M_{j,j} = \begin{cases} 1/\sigma & j \in J \\ \sqrt{2}/\sigma & j \notin J \end{cases}.$$

Then for $1 \leq j < k \leq n$,

$$\operatorname{Cov}\left(\left|M_j\widehat{d}_j\right|^2, \left|M_k\widehat{d}_k\right|^2\right) = \begin{cases} 4\left(1 + \dfrac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right), & j+k-2 = n \\ 0, & j+k-2 \neq n \end{cases}.$$

Proof. By the linearity of covariance and rewriting the magnitudes squared in terms of real and imaginary parts,

$$\begin{aligned}
&\operatorname{Cov}\left(|M_j\widehat{d}_j|^2, |M_k\widehat{d}_k|^2\right) \\
&= \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2 + \operatorname{Im}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2 + \operatorname{Im}(M_k\widehat{d}_k)^2\right) \\
&= \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) + \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Im}(M_k\widehat{d}_k)^2\right) \\
&+ \operatorname{Cov}\left(\operatorname{Im}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) + \operatorname{Cov}\left(\operatorname{Im}(M_j\widehat{d}_j)^2, \operatorname{Im}(M_k\widehat{d}_k)^2\right).
\end{aligned}$$

By Lem. 3.5.7, $\operatorname{Re}(M_j\widehat{d}_j)^2$ and $\operatorname{Im}(M_k\widehat{d}_k)^2$ are independent for all $1 \leq j,k \leq n$, meaning that their covariance is zero. Thus, the covariance simplifies to

$$\begin{aligned}
\operatorname{Cov}\left(|M_j\widehat{d}_j|^2, |M_k\widehat{d}_k|^2\right) &= \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) \\
&+ \operatorname{Cov}\left(\operatorname{Im}(M_j\widehat{d}_j)^2, \operatorname{Im}(M_k\widehat{d}_k)^2\right).
\end{aligned}$$

A further simplification can be made by using $\widehat{\mathbf{d}} = \widehat{\mathbf{b}} + \widehat{\boldsymbol{\eta}}$ and expanding the squared terms:

$$\begin{aligned}
&\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) \\
&= \operatorname{Cov}\left([\operatorname{Re}(M_j\widehat{b}_j) + \operatorname{Re}(M_j\widehat{\eta}_j)]^2, [\operatorname{Re}(M_k\widehat{b}_k) + \operatorname{Re}(M_k\widehat{\eta}_k)]^2\right) \\
&= \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{b}_j)^2, \operatorname{Re}(M_k\widehat{b}_k)^2\right) + \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{b}_j)^2, 2\operatorname{Re}(M_k\widehat{b}_k)\operatorname{Re}(M_k\widehat{\eta}_k)\right) \\
&+ \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{b}_j)^2, \operatorname{Re}(M_k\widehat{\eta}_k)^2\right) + \operatorname{Cov}\left(2\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_j\widehat{\eta}_j), \operatorname{Re}(M_k\widehat{b}_k)^2)\right) \\
&+ \operatorname{Cov}\left(2\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_j\widehat{\eta}_j), 2\operatorname{Re}(M_k\widehat{b}_k)\operatorname{Re}(M_k\widehat{\eta}_k)\right) \\
&+ \operatorname{Cov}\left(2\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_j\widehat{\eta}_j), \operatorname{Re}(M_k\widehat{\eta}_k)^2)\right) + \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, \operatorname{Re}(M_k\widehat{b}_k)^2\right) \\
&+ \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, 2\operatorname{Re}(M_k\widehat{b}_k)\operatorname{Re}(M_k\widehat{\eta}_k)\right) + \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, \operatorname{Re}(M_k\widehat{\eta}_k)^2\right).
\end{aligned}$$

While this certainly appears to be the opposite of a simplification, any covariance term with an argument not containing $\widehat{\eta}_j$ or $\widehat{\eta}_k$ is zero. Five of the nine covariance terms are thus removed, leaving

$$\begin{aligned}\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) &= \operatorname{Cov}\left(2\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_j\widehat{\eta}_j), 2\operatorname{Re}(M_k\widehat{b}_k)\operatorname{Re}(M_k\widehat{\eta}_k)\right)\\ &+ \operatorname{Cov}\left(2\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_j\widehat{\eta}_j), \operatorname{Re}(M_k\widehat{\eta}_k)^2)\right)\\ &+ \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, 2\operatorname{Re}(M_k\widehat{b}_k)\operatorname{Re}(M_k\widehat{\eta}_k)\right)\\ &+ \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, \operatorname{Re}(M_k\widehat{\eta}_k)^2\right).\end{aligned}$$

The constant coefficients can be factored and combined so that

$$\begin{aligned}\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{d}_j)^2, \operatorname{Re}(M_k\widehat{d}_k)^2\right) &= [4\operatorname{Re}(M_j\widehat{b}_j)\operatorname{Re}(M_k\widehat{b}_k)]\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j), \operatorname{Re}(M_k\widehat{\eta}_k)\right)\\ &+ [2\operatorname{Re}(M_j\widehat{b}_j)]\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j), \operatorname{Re}(M_k\widehat{\eta}_k)^2)\right)\\ &+ [2\operatorname{Re}(M_k\widehat{b}_k)]\operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, \operatorname{Re}(M_k\widehat{\eta}_k)\right)\\ &+ \operatorname{Cov}\left(\operatorname{Re}(M_j\widehat{\eta}_j)^2, \operatorname{Re}(M_k\widehat{\eta}_k)^2\right).\end{aligned}$$

Let F_R be the real part of the Fourier matrix (3.3.13). Then $\operatorname{Re}(M_j\widehat{\eta}_j) = (MF_R\boldsymbol{\eta})_j$ for all $1 \leq j \leq n$. Lem. 3.5.7 states that $(MF_R\boldsymbol{\eta})_j$ and $(MF_R\boldsymbol{\eta})_k$ are independent for all $1 \leq j < k \leq n$ with $j + k - 2 \neq n$ and $(MF_R\boldsymbol{\eta})_j = (MF_R\boldsymbol{\eta})_k$ when $j + k - 2 = n$. Thus all of the covariance terms above are zero when $j + k - 2 \neq n$. As for the $j + k - 2 = n$ case, Thm. 3.5.3 yields the final result:

$$\operatorname{Cov}\left(\left|M_j\widehat{d}_j\right|^2, \left|M_k\widehat{d}_k\right|^2\right) = 4\left(1 + \frac{\left|\widehat{b}_j\right|^2}{\sigma^2}\right), \qquad j + k - 2 = n.$$

□

The primary challenge in dealing with statistics of the DFT applied to independent and identically distributed white noise $\boldsymbol{\eta}$ is that the covariance matrices of $F_R\boldsymbol{\eta}$ and $F_I\boldsymbol{\eta}$ are singular. Fortunately, this is not the case for an orthogonal transformation such as the DCT.

Let C be an $n \times n$ orthogonal matrix, $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$, and $\breve{\boldsymbol{\eta}} = C\boldsymbol{\eta}$. By the orthogonality of C and properties of the multivariate normal distribution,

$$\breve{\boldsymbol{\eta}} \sim \mathcal{N}(C\mathbf{0}, \sigma^2 C^{\mathsf{T}} C) = \mathcal{N}\left(\mathbf{0}, \sigma^2 I_n\right). \tag{3.5.31}$$

Thus the distribution of $\boldsymbol{\eta}$ is preserved under C. More generally,

$$\breve{\mathbf{d}} \sim \mathcal{N}\left(\breve{\mathbf{d}}, \sigma^2 I_n\right)$$

where $\mathbf{d} = \mathbf{b} + \boldsymbol{\eta}$. Therefore the orthogonal version of Thm. 3.5.3 is simpler.

Theorem 3.5.6 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$ *and let* $\breve{\mathbf{d}} = C\mathbf{d}$ *where* $C \in \mathbb{R}^{n \times n}$ *is orthogonal. Then for all* $1 \leq j \leq n$,

$$\frac{1}{\sigma^2}\breve{d}_j^2 \sim \chi'\left(1, \frac{1}{2\sigma^2}\left(\breve{b}_j\right)^2\right).$$

Proof. For $1 \leq j \leq n$, the properties of the univariate normal distribution give

$$\frac{1}{\sigma}\breve{b}_j \sim \mathcal{N}\left(\sqrt{2\lambda}, 1\right), \quad \lambda = \frac{1}{2\sigma^2}\left(\breve{b}_j\right)^2.$$

The techniques utilized in the proof of Thm. 3.5.2 can now be applied to yield

$$\frac{1}{\sigma^2}\left(\breve{b}_j\right)^2 \sim \chi'\left(1, \lambda\right) = \chi'\left(1, \frac{1}{2\sigma^2}\left(\breve{b}_j\right)^2\right).$$

□

Using Thm. 3.5.6 and the fact that the sum of independent noncentral chi-squared random variables is also a noncentral chi-squared random variable, Cor. 3.5.2 describes the distribution of the 2-norm of a Gaussian random vector under an orthogonal transformation.

Corollary 3.5.2 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I)$ *and let* $\breve{\mathbf{d}} = C\mathbf{d}$ *where* $C \in \mathbb{R}^{n\times n}$ *is orthogonal. Then*

$$\left\|\breve{\mathbf{d}}\right\|_2^2 \sim \chi'\left(n, \frac{1}{2\sigma^2}\left\|\breve{\mathbf{b}}\right\|_2^2\right).$$

Lastly, the expected value and variance of the components of $\breve{\mathbf{d}}$ then follow directly from Thm. 3.5.4.

Lemma 3.5.10 *Let* $\mathbf{d}$ *be a random* n*-vector with* $\mathbf{d} \sim \mathcal{N}(\mathbf{b}, \sigma^2 I_n)$ *and let* $\breve{\mathbf{d}} = C\mathbf{d}$ *where* $C \in \mathbb{R}^{n\times n}$ *is orthogonal. Then for all* $1 \le j \le n$*,*

$$\mathbb{E}\left(\frac{1}{\sigma^2}\left(\breve{d}_j\right)^2\right) = 1 + \frac{1}{2\sigma^2}\left(\breve{b}_j\right)^2, \quad \mathrm{Var}\left(\frac{1}{\sigma^2}\left(\breve{d}_j\right)^2\right) = 2 + \frac{2}{\sigma^2}\left(\breve{b}_j\right)^2.$$

3.6 Summary

In summary, the connections between blurred data, boundary conditions, resulting matrix structure, and noise content are important for understanding the numerical implementations and consequence of regularization. These connections allow for quantitative analysis of the learning and spectral windowing approaches in later chapters. In addition, matrix decompositions associated with the DFT and DCT are utilized for efficiency in numerical implementations.

Chapter 4

PARAMETER ESTIMATION METHODS

Perhaps the simplest method of selecting a regularization parameter is to find α that minimizes

$$F_{\text{MSE}}(\alpha) = \|\mathbf{x}(\alpha) - \mathbf{x}\|_2^2, \tag{4.0.1}$$

where MSE stands for mean squared error. Though a mean is not being taken in eq. (4.0.1), the use of the broader term MSE becomes apparent once multiple data sets are considered (Chapter 6). The primary disadvantage of using this method is that the true solution $\mathbf{x}$ must be known, or in other words, this is a supervised learning method. Not only would knowing the true solution render the process of finding a regularized solution pointless, but in practice a true solution is not known. This motivates the use of other methods, which do not rely upon knowledge of a true solution. The three such methods considered are the unbiased predictive risk estimator method (Section 4.1), the discrepancy principle method (Section 4.2), and the generalized cross validation method (Section 4.3). Since true solutions are known for use in the numerical examples, the method defined using eq. (4.0.1) will be used as a benchmark for comparing the other three methods. Here we introduce these methods for the single data set, scalar parameter case; their spectral windowed multidata extensions are introduced in Chapter 5. Since Chapter 5 contains versions of the parameter functions in summation notation in addition to matrix-vector notation, summation forms of the parameter functions are also presented in this chapter. A summation version of the UPRE function is derived at the end of Section 4.1, and this derivation is used to form summation version of the MDP and GCV functions in

their corresponding sections.

4.1 Unbiased Predictive Risk Estimator

The unbiased predictive risk estimator (UPRE) method Mallows (1973) is derived by considering the quantity

$$\mathbf{p}(\alpha) = A(\mathbf{x}(\alpha) - \mathbf{x}). \tag{4.1.2}$$

The quantity $\mathbf{p}(\alpha)$ is known as the predictive error, and is an alternative to solution error defined as $\mathbf{x}(\alpha) - \mathbf{x}$. Given the above definition, the mean squared norm of the predictive error is

$$\frac{1}{m}\|\mathbf{p}(\alpha)\|_2^2 = \frac{1}{m}\|A(\mathbf{x}(\alpha) - \mathbf{x})\|_2^2$$

which is called the predictive risk. As a first step in deriving the UPRE method, the noise $\boldsymbol{\eta}$ is assumed to be a random vector, instead of a realization of a random vector. Direct consequences of this assumption are that $\mathbf{b}$ and $\mathbf{x}(\alpha)$ are random vectors and the predictive risk $(1/m)\|\mathbf{p}(\alpha)\|_2^2$ is a random variable.

Next, we assume that the regularized solution $\mathbf{x}(\alpha)$ depends linearly on the data. This means that we can write $\mathbf{x}(\alpha) = R(\alpha)\mathbf{d}$, where $R(\alpha)$ is called the regularization matrix Vogel (2002) that is dependent upon α. The $m \times m$ influence matrix $A(\alpha)$, also called the data resolution matrix Aster *et al.* (2013), is then defined as $A(\alpha) = AR(\alpha)$. Using the influence matrix with $\mathbf{x}(\alpha) = R(\alpha)\mathbf{d}$, the predictive error can be rewritten:

$$\begin{aligned}
\mathbf{p}(\alpha) &= A\mathbf{x}(\alpha) - A\mathbf{x} \\
&= A(\alpha)\mathbf{d} - A\mathbf{x} \\
&= A(\alpha)(A\mathbf{x} + \boldsymbol{\eta}) - A\mathbf{x} \\
&= (A(\alpha) - I_m)A\mathbf{x} + A(\alpha)\boldsymbol{\eta}.
\end{aligned}$$

By the assumption that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$, the Trace Lemma (Lem. 4.1.1) can be

utilized to obtain an expression for the expected value (denoted $\mathbb{E}(\cdot)$) of predictive risk.

Lemma 4.1.1 (Trace Lemma (Vogel, 2002, p. 98)) *Let $f \in \mathcal{H}$, where $\mathcal{H}$ is a deterministic real Hilbert space, let $\boldsymbol{\eta}$ be a discrete noise vector with $\boldsymbol{\eta} \sim \mathcal{N}(0, \sigma^2 I_m)$, and let $B : \mathbb{R}^m \to \mathcal{H}$ be a bounded linear operator. Then*

$$\mathbb{E}(\|f + B\boldsymbol{\eta}\|_{\mathcal{H}}^2) = \|f\|_{\mathcal{H}}^2 + \sigma^2 \operatorname{trace}(B^* B)$$

where B^* denotes the adjoint of B.

Applying the Trace Lemma to the expression for predictive risk yields

$$\begin{aligned}\mathbb{E}\left(\frac{1}{m}\|\mathbf{p}(\alpha)\|_2^2\right) &= \frac{1}{m}\,\mathbb{E}\left(\|(A(\alpha) - I)A\mathbf{x} + A(\alpha)\boldsymbol{\eta}\|_2^2\right) \\ &= \frac{1}{m}\|(A(\alpha) - I_m)A\mathbf{x}\|_2^2 + \frac{\sigma^2}{m}\operatorname{trace}(A(\alpha)^{\mathsf{T}} A(\alpha)).\end{aligned}$$

If generalized Tikhonov regularization is used, then the influence matrix is

$$A(\alpha) = A(A^{\mathsf{T}}A + \alpha^2 L^{\mathsf{T}}L)^{-1}A^{\mathsf{T}} = AA^{\sharp}(\alpha), \tag{4.1.3}$$

with $A^{\sharp}(\alpha)$ from eq. (2.2.12). In other words, we have $R(\alpha) = A^{\sharp}(\alpha)$ for generalized Tikhonov regularization. The matrix $(A^{\mathsf{T}}A + \alpha^2 L^{\mathsf{T}}L)^{-1}$ is symmetric as a result of $A^{\mathsf{T}}A$ and $\alpha^2 L^{\mathsf{T}}L$ being individually symmetric, and thus $A(\alpha)$ is symmetric. Then, the expected value of predictive risk is simplified to

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{p}(\alpha)\|_2^2\right) = \frac{1}{m}\|(A(\alpha) - I_m)A\mathbf{x}\|_2^2 + \frac{\sigma^2}{m}\operatorname{trace}(A(\alpha)^2). \tag{4.1.4}$$

The last step in the derivation of the UPRE method is to introduce the regularized residual, which is defined as $\mathbf{r}(\alpha) = A\mathbf{x}(\alpha) - \mathbf{d}$. The regularized residual is important because it is also used in the generalized cross validation and discrepancy principle methods. Using $A(\alpha)$, the expression for $\mathbf{r}(\alpha)$ can also be written as

$$\mathbf{r}(\alpha) = (A(\alpha) - I_m)\mathbf{d} = (A(\alpha) - I_m)(A\mathbf{x} + \boldsymbol{\eta}) = (A(\alpha) - I_m)A\mathbf{x} + (A(\alpha) - I_m)\boldsymbol{\eta}.$$

By the Trace Lemma and the expression for $\mathbf{r}(\alpha)$,

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2\right) = \frac{1}{n}\|(A(\alpha) - I_m)A\mathbf{x}\|_2^2 + \frac{\sigma^2}{m}\,\mathrm{trace}((A(\alpha) - I_m)^\mathsf{T}(A(\alpha) - I_m)).$$

For symmetric $A(\alpha)$, the term $(A(\alpha) - I_m)^\mathsf{T}(A(\alpha) - I_m)$ becomes

$$(A(\alpha) - I_m)^2 = A(\alpha)^2 - 2A(\alpha) + I_m,$$

and so by the linearity of the trace operator,

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2\right) = \frac{1}{m}\|(A(\alpha) - I_m)A\mathbf{x}\|_2^2 + \frac{\sigma^2}{n}\,\mathrm{trace}(A(\alpha)^2) - \frac{2\sigma^2}{m}\,\mathrm{trace}(A(\alpha)) + \sigma^2. \tag{4.1.5}$$

By comparing eq. (4.1.4) and eq. (4.1.5),

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{p}(\alpha)\|_2^2\right) = \mathbb{E}\left(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2\right) + \frac{2\sigma^2}{m}\,\mathrm{trace}(A(\alpha)) - \sigma^2,$$

yielding the UPRE function

$$F_{\mathrm{UPRE}}(\alpha) = \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 + \frac{2\sigma^2}{m}\,\mathrm{trace}(A(\alpha)) - \sigma^2. \tag{4.1.6}$$

The UPRE method finds α such that $F_{\mathrm{UPRE}}(\alpha)$ is minimized. Note that the UPRE method does rely on knowledge of σ^2.

To derive a summation form of eq. (4.1.6), we apply the GSVD of (A, L) to the influence matrix $A(\alpha) = A(A^\mathsf{T}A + \alpha^2 L^\mathsf{T}L)^{-1}A^\mathsf{T}$ yielding

$$\begin{aligned}
A(\alpha) &= A\Big(A^\mathsf{T}A + \alpha^2 L^\mathsf{T}L\Big)^{-1}A^\mathsf{T} \\
&= U\Delta X^\mathsf{T}\Big(X\Delta^\mathsf{T}U^\mathsf{T}U\Delta X^\mathsf{T} + \alpha^2 X\Lambda^\mathsf{T}V^\mathsf{T}V\Lambda X^\mathsf{T}\Big)^{-1}X\Delta^\mathsf{T}U^\mathsf{T} \\
&= U\Delta X^\mathsf{T}\Big[X\left(\Delta^\mathsf{T}\Delta + \alpha^2\Lambda^\mathsf{T}\Lambda\right)X^\mathsf{T}\Big]^{-1}X\Delta^\mathsf{T}U^\mathsf{T} \\
&= U\Delta\Big(\Delta^\mathsf{T}\Delta + \alpha^2\Lambda^\mathsf{T}\Lambda\Big)^{-1}\Delta^\mathsf{T}U^\mathsf{T}.
\end{aligned}$$

Using the similarity invariance and cyclic properties of the trace,

$$\mathrm{trace}\left(A(\alpha)\right) = \mathrm{trace}\left((\Delta^\mathsf{T}\Delta + \alpha^2\Lambda^\mathsf{T}\Lambda)^{-1}\Delta^\mathsf{T}\Delta\right) = \mathrm{trace}\left(\Phi(\alpha)\right),$$

expressed in summation notation from eq. (2.3.20) as

$$\operatorname{trace}\left(A(\alpha)\right) = \sum_{j=1}^{n} \phi_j(\alpha) = (n - q^*) + \sum_{j=1}^{q^*} \phi_j(\alpha). \tag{4.1.7}$$

Note that though $A(\alpha)$ is an $m \times m$ diagonal matrix, $[A(\alpha)]_{j,j} = 0$ for $j = n+1 : m$.

Next, the definition of $\mathbf{r}(\alpha)$ with eq. (2.3.23) and $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$ gives

$$\begin{aligned}
\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 = \frac{1}{m}\|A\mathbf{x}(\alpha) - \mathbf{d}\|_2^2 &= \frac{1}{m}\left\|U\Delta X^{\mathsf{T}} Y \Phi(\alpha)\Delta^{\dagger}\widehat{\mathbf{d}} - \mathbf{d}\right\|_2^2 \\
&= \frac{1}{m}\left\|U\Delta\Phi(\alpha)\Delta^{\dagger}U^{\mathsf{T}}\mathbf{d} - UU^{\mathsf{T}}\mathbf{d}\right\|_2^2 \\
&= \frac{1}{m}\left\|\left(\Delta\Phi(\alpha)\Delta^{\dagger} - I_m\right)\widehat{\mathbf{d}}\right\|_2^2, \\
&= \frac{1}{m}\left\|\begin{bmatrix} \Phi(\alpha) - I_n & \mathbf{0}_{n\times(m-n)} \\ \mathbf{0}_{(m-n)\times n} & -I_{(m-n)\times(m-n)} \end{bmatrix}\widehat{\mathbf{d}}\right\|_2^2,
\end{aligned}$$

which can be written using summation notation as

$$\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 = \frac{1}{m}\sum_{j=1}^{n} \psi_j^2(\alpha)\widehat{d}_j^2 + \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2. \tag{4.1.8}$$

Combining eq. (4.1.7) and eq. (4.1.8) produces a summation form of the UPRE function:

$$F_{\text{UPRE}}(\alpha) = \frac{1}{m}\sum_{j=1}^{n} \psi_j^2(\alpha)\widehat{d}_j^2 + \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2 + \frac{2\sigma^2}{m}\sum_{j=1}^{n} \phi_j(\alpha) - \sigma^2. \tag{4.1.9}$$

Since the UPRE method relies on finding a minimum of eq. (4.1.9), the constant terms $\frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2$ and $-\sigma^2$ can be ignored during implementation.

The problem of finding a minimizer of eq. (4.1.9) can be recast as a root-finding problem by using the derivative of the UPRE function, which is

$$\begin{aligned}
\frac{d}{d\alpha}F_{\text{UPRE}}(\alpha) &= \frac{2}{m}\sum_{j=1}^{n} \psi_j(\alpha)\frac{d\psi_j(\alpha)}{d\alpha}\widehat{d}_j^2 + \frac{2\sigma^2}{m}\sum_{j=1}^{n} \frac{d\phi_j(\alpha)}{d\alpha} \\
&= \frac{4\alpha}{m}\left[\sum_{j=1}^{n} \psi_j(\alpha)\frac{\gamma_j^2}{(\gamma_j^2 + \alpha^2)^2}\widehat{d}_j^2 - \sigma^2\sum_{j=1}^{n} \frac{\gamma_j^2}{(\gamma_j^2 + \alpha^2)^2}\right]
\end{aligned} \tag{4.1.10}$$

using the derivatives in eq. (2.3.21). It is clear that $\alpha = 0$ is a root of eq. (4.1.10) regardless of the data and operator spectra, so this root should be ignored (since a solution generated for $\alpha = 0$ is in fact a non-regularized solution).

Alternatively, a form of the UPRE function can also be derived in the case where a simultaneous diagonalization of A and L is available instead of the GSVD. As the illustrative example, we assume that A and L can be simultaneously diagonalized by the DFT, i.e. $A = F^{\mathsf{H}}\Delta F$ and $L = F^{\mathsf{H}}\Lambda F$ for the unitary DFT matrix F defined by eq. (3.3.13). In this new context, $\widehat{\mathbf{d}} = F\mathbf{d}$, though we retain the assumption that $\mathbf{d}$ is a vector of real numbers. We begin with the DFT of the predictive error defined by eq. (4.1.2):

$$\widehat{\mathbf{p}}(\alpha) = FA(\mathbf{x}(\alpha) - \mathbf{x}) = \Delta\left(\widehat{\mathbf{x}}(\alpha) - \widehat{\mathbf{x}}\right).$$

Since the influence matrix $A(\alpha)$ is defined as $A(\alpha) = AR(\alpha) = F^{\mathsf{H}}\Delta FR(\alpha)$ with regularization matrix $R(\alpha)$, conjugation of $A(\alpha)$ by F gives $FA(\alpha)F^{\mathsf{H}} = \Delta FR(\alpha)F^{\mathsf{H}}$. Combined with $\widehat{\mathbf{x}}(\alpha) = FR(\alpha)\mathbf{d} = FR(\alpha)F^{\mathsf{H}}\widehat{\mathbf{d}}$ and $\widehat{\mathbf{d}} = F(A\mathbf{x} + \boldsymbol{\eta}) = \Delta\widehat{\mathbf{x}} + \widehat{\boldsymbol{\eta}}$, the predictive error is rewritten as

$$\begin{aligned}
\widehat{\mathbf{p}}(\alpha) &= \Delta\widehat{\mathbf{x}}(\alpha) - \Delta\widehat{\mathbf{x}} \\
&= FA(\alpha)F^{\mathsf{H}}\widehat{\mathbf{d}} - \Delta\widehat{\mathbf{x}} \\
&= FA(\alpha)F^{\mathsf{H}}(\Delta\widehat{\mathbf{x}} + \widehat{\boldsymbol{\eta}}) - \Delta\widehat{\mathbf{x}} \\
&= \left[F(A(\alpha) - I_m)F^{\mathsf{H}}\right]\Delta\widehat{\mathbf{x}} + \left[FA(\alpha)F^{\mathsf{H}}\right]\widehat{\boldsymbol{\eta}}. \qquad (4.1.11)
\end{aligned}$$

Since the components of $\widehat{\mathbf{x}}$ and $\widehat{\boldsymbol{\eta}}$ are not guaranteed to be real, the Trace Lemma as previously stated can no longer be directly applied. Instead, Lem. 4.1.1 can be modified to accommodate the existence of complex components arising from the application of the DFT; the proof is a modification of the proof of Lem. 4.1.1 in (Vogel, 2002, p. 98).

Lemma 4.1.2 (DFT Trace Lemma) *Let $f \in \mathcal{H}$, where $\mathcal{H}$ is a deterministic complex Hilbert space, let $\boldsymbol{\eta}$ be a discrete white noise vector with $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$, and let $B : \mathbb{C}^m \to \mathcal{H}$ be a bounded linear operator. Furthermore, let $\widehat{\boldsymbol{\eta}} = F\boldsymbol{\eta}$ where F is the $n \times n$ unitary DFT matrix. Then*

$$\mathbb{E}(\|f + B\widehat{\boldsymbol{\eta}}\|_{\mathcal{H}}^2) = \|f\|_{\mathcal{H}}^2 + \sigma^2 \operatorname{trace}(B^{\mathsf{H}} B) \tag{4.1.12}$$

where B^{H} denotes the adjoint of B.

Proof. By the linearity of inner products and the expected value operator,

$$\begin{aligned}
\mathbb{E}(\|f + B\widehat{\boldsymbol{\eta}}\|_{\mathcal{H}}^2) &= \mathbb{E}(\langle f + B\widehat{\boldsymbol{\eta}}, f + B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}}) \\
&= \mathbb{E}(\|f\|_{\mathcal{H}}^2) + \mathbb{E}\left(\langle f, B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}} + \overline{\langle f, B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}}}\right) + \mathbb{E}(\langle B\widehat{\boldsymbol{\eta}}, B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}}) \\
&= \mathbb{E}(\|f\|_{\mathcal{H}}^2) + 2\,\mathbb{E}\left(\operatorname{Re}(\langle f, B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}})\right) + \mathbb{E}(\langle B\widehat{\boldsymbol{\eta}}, B\widehat{\boldsymbol{\eta}} \rangle_{\mathcal{H}}).
\end{aligned}$$

This difference between the real and DFT versions of the Trace Lemma comes from the fact that the inner product on a complex Hilbert space is a sesquilinear form instead of a bilinear form. Again the term $\mathbb{E}(\|f\|_{\mathcal{H}}^2)$ can be reduced to $\|f\|_{\mathcal{H}}^2$ because f is an element of a deterministic Hilbert space. The inner products can be rewritten using the adjoint of B:

$$\mathbb{E}(\|f + B\widehat{\boldsymbol{\eta}}\|_{\mathcal{H}}^2) = \|f\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left(\operatorname{Re}\left((\widehat{\boldsymbol{\eta}})^{\mathsf{H}} B^{\mathsf{H}} f\right)\right) + \mathbb{E}\left(\widehat{\boldsymbol{\eta}}^{\mathsf{H}} B^{\mathsf{H}} B\widehat{\boldsymbol{\eta}}\right). \tag{4.1.13}$$

Recall that Lem. 3.5.5 described the distributions of the real and imaginary components of $\widehat{\boldsymbol{\eta}}$. While statistics of the full vector $\widehat{\boldsymbol{\eta}} = \operatorname{Re}(\widehat{\boldsymbol{\eta}}) + i \operatorname{Im}(\widehat{\boldsymbol{\eta}})$ could be utilized from the distributions described by Lem. 3.5.5, the simpler approach is to write $\widehat{\boldsymbol{\eta}} = F\boldsymbol{\eta}$ and use the fact that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$ directly. As a result, eq. (4.1.13) can

be written as

$$\begin{aligned}
\mathbb{E}(\|f + B\widehat{\boldsymbol{\eta}}\|_{\mathcal{H}}^2) &= \|f\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left(\mathrm{Re}\left((F\boldsymbol{\eta})^{\mathsf{H}} B^{\mathsf{H}} f\right)\right) + \mathbb{E}\left((F\boldsymbol{\eta})^{\mathsf{H}} B^{\mathsf{H}} BF\boldsymbol{\eta}\right) \\
&= \|f\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left(\mathrm{Re}\left(\boldsymbol{\eta}^{\mathsf{H}} F^{\mathsf{H}} B^{\mathsf{H}} f\right)\right) + \mathbb{E}(\boldsymbol{\eta}^{\mathsf{H}} F^{\mathsf{H}} B^{\mathsf{H}} BF\boldsymbol{\eta}) \\
&= \|f\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left(\mathrm{Re}\left(\boldsymbol{\eta}^{\mathsf{T}} F^{\mathsf{H}} B^{\mathsf{H}} f\right)\right) + \mathbb{E}(\boldsymbol{\eta}^{\mathsf{T}} F^{\mathsf{H}} B^{\mathsf{H}} BF\boldsymbol{\eta}) \\
&= \|f\|_{\mathcal{H}}^2 + 2\sum_{j=1}^{m} \mathbb{E}(\boldsymbol{\eta}_j^{\mathsf{T}})\,\mathrm{Re}\left((F^{\mathsf{H}} B^{\mathsf{H}} f)_j\right) + \sum_{j=1}^{m}\sum_{k=1}^{m} (F^{\mathsf{H}} B^{\mathsf{H}} BF)_{j,k}\,\mathbb{E}(\eta_j \eta_k).
\end{aligned}$$

Since $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$, the expected value of η_j is zero, as is the expected value of $\eta_j \eta_k$ when $j \neq k$. Therefore the second term above is zero and the third term is a summation expression for $\sigma^2\,\mathrm{trace}(F^{\mathsf{H}} B^{\mathsf{H}} BF)$. Lastly since F is unitary and the trace operation is invariant under similarity transformations, $\sigma^2\,\mathrm{trace}(F^{\mathsf{H}} B^{\mathsf{H}} BF) = \sigma^2\,\mathrm{trace}(B^{\mathsf{H}} B)$. □

This DFT version of the Trace Lemma is suited to rewrite eq. (4.1.11) since the first term of eq. (4.1.11) is deterministic and the second term matches the form $B\widehat{\boldsymbol{\eta}}$ where $B = FA(\alpha)F^{\mathsf{H}}$ and $\widehat{\boldsymbol{\eta}} = F\boldsymbol{\eta}$ with DFT matrix F and $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$. Applying the DFT version of the Trace Lemma to eq. (4.1.11) yields

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{p}}(\alpha)\right\|_2^2\right) &= \frac{1}{m}\,\mathbb{E}\left(\left\|[F(A(\alpha) - I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}} + [FA(\alpha)F^{\mathsf{H}}]\widehat{\boldsymbol{\eta}}\right\|_2^2\right) \\
&= \mathbb{E}\left(\frac{1}{m}\left\|[F(A(\alpha) - I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}}\right\|_2^2\right) + \frac{\sigma^2}{m}\,\mathrm{trace}\left((FA(\alpha)F^{\mathsf{H}})^{\mathsf{H}} FA(\alpha)F^{\mathsf{H}}\right) \\
&= \mathbb{E}\left(\frac{1}{m}\left\|[F(A(\alpha) - I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}}\right\|_2^2\right) + \frac{\sigma^2}{m}\,\mathrm{trace}\left(A(\alpha)^{\mathsf{H}} A(\alpha)\right).
\end{aligned}$$

The DFT of the regularized residual $\mathbf{r}(\alpha)$ is

$$\widehat{\mathbf{r}}(\alpha) = F(A\mathbf{x}(\alpha) - \mathbf{d}) = FF^{\mathsf{H}}\Delta F\mathbf{x}(\alpha) - F\mathbf{d} = \Delta\widehat{\mathbf{x}}(\alpha) - \widehat{\mathbf{d}},$$

which is a complex vector. Using $FA(\alpha)F^{\mathsf{H}}$, the expression for $\widehat{\mathbf{r}}(\alpha)$ can be rewritten as

$$\widehat{\mathbf{r}}(\alpha) = F(A(\alpha) - I_m)F^{\mathsf{H}}\left(\Delta\widehat{\mathbf{x}} + \widehat{\boldsymbol{\eta}}\right) = \left[F(A(\alpha) - I_m)F^{\mathsf{H}}\right]\Delta\widehat{\mathbf{x}} + \left[F(A(\alpha) - I_m)F^{\mathsf{H}}\right]\widehat{\boldsymbol{\eta}}.$$

Applying the DFT version of the Trace Lemma again produces

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) &= \mathbb{E}\left(\frac{1}{m}\left\|[F(A(\alpha)-I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}}\right\|_2^2\right) \\
&\qquad + \frac{\sigma^2}{m}\operatorname{trace}\left(F(A(\alpha)-I_m)^{\mathsf{H}}(A(\alpha)-I)F^{\mathsf{H}}\right) \\
&= \mathbb{E}\left(\frac{1}{m}\left\|[F(A(\alpha)-I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}}\right\|_2^2\right) \\
&\qquad + \frac{\sigma^2}{m}\operatorname{trace}\left(F(A(\alpha)^{\mathsf{H}}A(\alpha)-(A(\alpha)^{\mathsf{H}}+A(\alpha))+I_m)F^{\mathsf{H}}\right) \\
&= \mathbb{E}\left(\frac{1}{m}\left\|[F(A(\alpha)-I_m)F^{\mathsf{H}}]\Delta\widehat{\mathbf{x}}\right\|_2^2\right) + \frac{\sigma^2}{m}\operatorname{trace}\left(A(\alpha)^{\mathsf{H}}A(\alpha)\right) \\
&\qquad - \frac{2\sigma^2}{m}\operatorname{trace}\left(\operatorname{Re}(A(\alpha))\right) + \sigma^2.
\end{aligned}$$

Thus $\mathbb{E}(\frac{1}{m}\|\widehat{\mathbf{p}}(\alpha)\|_2^2)$ can again be expressed as

$$\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{p}}(\alpha)\right\|_2^2\right) = \mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) + \frac{2\sigma^2}{m}\operatorname{trace}\left(\operatorname{Re}(A(\alpha))\right) - \sigma^2. \tag{4.1.14}$$

Now to obtain a greater understanding of the norm term in eq. (4.1.14) in the DFT case, we look to the statistics from Chapter 3, specifically Cor. 3.5.1. Though not immediately needed for either the UPRE or MDP methods, the variance of $\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2$ can be determined as well using results from Chapter 3. The expectation and variance of $\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2$ are provided in Lem. 4.1.3.

Lemma 4.1.3 *Let the factors M_j for $j = 1, \ldots, m$ be defined by*

$$M_j = \begin{cases} 1/\sigma & j = 1, \frac{m}{2}+1 \\ \sqrt{2}/\sigma & \textit{otherwise} \end{cases},$$

where the condition $\frac{m}{2}+1$ is ignored if m is odd. Then $\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right)$ from eq. (4.1.14) can be written as

$$\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) = \sum_{j=1}^{m}\frac{\sigma^2}{2}\left(2+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)(\psi_j(\alpha))^2, \tag{4.1.15}$$

and $\mathrm{Var}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right)$ *can be written as*

$$\mathrm{Var}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) = 2\sigma^4 \sum_{j=1}^{m}\left(1+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)\left(\psi_j(\alpha)\right)^4. \tag{4.1.16}$$

Proof. Using the factors M_j for $j = 1, \ldots, m$, $\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right)$ from eq. (4.1.14) can be written as

$$\begin{aligned}\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) &= \mathbb{E}\left(\sum_{j=1}^{m}|M_j\widehat{d}_j|^2\left(\frac{\psi_j(\alpha)}{M_j}\right)^2\right)\\ &= \sum_{j=1}^{m}\mathbb{E}\left(|M_j\widehat{d}_j|^2\right)\left(\frac{\psi_j(\alpha)}{M_j}\right)^2,\end{aligned}$$

the last equality having been obtained through the linearity of expectation Casella and Berger (2002). Cor. 3.5.1 then gives

$$\begin{aligned}\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) &= \sum_{j=1}^{m}\frac{\sigma^2}{2}\left(2+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)\left(\psi_j(\alpha)\right)^2\\ &= \sum_{j=1}^{m}\frac{\sigma^2}{2}\left(2+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)\left(\psi_j(\alpha)\right)^2.\end{aligned}$$

From the definition of variance,

$$\begin{aligned}\mathrm{Var}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) &= \mathrm{Var}\left(\sum_{j=1}^{m}|M_j\widehat{d}_j|^2\left(\frac{\psi_j(\alpha)}{M_j}\right)^2\right)\\ &= \sum_{j=1}^{m}\sum_{\ell=1}^{m}\left(\frac{\psi_j(\alpha)\psi_\ell(\alpha)}{M_jM_\ell}\right)^2\mathrm{Cov}\left(|M_j\widehat{d}_j|^2, |M_\ell\widehat{d}_\ell|^2\right).\end{aligned} \tag{4.1.17}$$

By Thm. 3.5.5, the covariance terms are only nonzero when $j = \ell$ or $j + \ell - 2 = n$. In either case,

$$\left(\frac{\psi_j(\alpha)\psi_\ell(\alpha)}{M_jM_\ell}\right)^2\mathrm{Cov}\left(|M_j\widehat{d}_j|^2, |M_\ell\widehat{d}_\ell|^2\right) = \left(\frac{\psi_j(\alpha)}{M_j}\right)^4\mathrm{Var}\left(|M_j\widehat{d}_j|^2\right),$$

Let $J = \{1, (n/2)+1\}$ if n is even and $J = \{1\}$ if n is odd. Cor. 3.5.1 states that

$$\left(\frac{\psi_j(\alpha)}{M_j}\right)^4\mathrm{Var}\left(|M_j\widehat{d}_j|^2\right) = 2\sigma^4\left(1+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)\left(\psi_j(\alpha)\right)^4, \qquad j \in J$$

and similarly

$$\left(\frac{\psi_j(\alpha)}{M_j}\right)^4 \mathrm{Var}\left(|M_j\widehat{d}_j|^2\right) = \sigma^4\left(1+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)(\psi_j(\alpha))^4, \qquad j \notin J$$

If $j \in J$, $\ell = j$ is the only $1 \le \ell \le n$ such that $j+\ell-2 = n$. On the other hand, if $j \notin J$ then there are two values of ℓ such that $j+\ell-2=n$: $\ell = j$ and $\ell = n-j+2$. Thus eq. (4.1.17) is

$$\mathrm{Var}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right) = 2\sigma^4\sum_{j=1}^{m}\left(1+\frac{|\widehat{b}_j|^2}{\sigma^2}\right)(\psi_j(\alpha))^4.$$

□

A similar derivation can be applied to a simultaneous diagonalization by a real matrix instead of a complex matrix. The illustrative example of this situation is diagonalization via the DCT, i.e. $A = C^{\mathsf{T}}\Delta C$ and $L = C^{\mathsf{T}}\Lambda C$ where C is the orthogonal DCT matrix defined by eq. (3.3.16). Recall that Thm. 3.3.1 describes the class of matrices diagonalized by the DCT. Since the DCT maps real vectors to real vectors, the standard Trace Lemma can be utilized; the DCT version of the UPRE function is

$$F_{\mathrm{UPRE}}(\alpha) = \frac{1}{m}\left\|\breve{\mathbf{r}}(\alpha)\right\|_2^2 + \frac{2\sigma^2}{m}\,\mathrm{trace}\left(A(\alpha)\right) - \sigma^2, \tag{4.1.18}$$

where $\breve{\mathbf{r}}(\alpha) = C\mathbf{r}(\alpha) = CA\mathbf{x}(\alpha) - \breve{\mathbf{d}}$ and $\breve{\mathbf{d}} = C\mathbf{d}$. The corresponding summation form is

$$F_{\mathrm{UPRE}}(\alpha) = \frac{1}{m}\sum_{j=1}^{n}\psi_j^2(\alpha)\breve{d}_j^2 + \frac{1}{m}\sum_{j=n+1}^{m}\breve{d}_j^2 + \frac{2\sigma^2}{m}\sum_{j=1}^{n}\phi_j(\alpha) - \sigma^2. \tag{4.1.19}$$

4.2 Discrepancy Principle

As a start to a stochastic derivation of the discrepancy principle method Morozov (1966) (for a deterministic derivation, see (Vogel, 2002, p. 8-9)), consider the case where $\mathbf{x}(\alpha) \approx \mathbf{x}$. Here we first assume in the discussion either the GSVD or some

real orthogonal transformation, like the DCT, is used so that the coefficients of the transformed data is real. In this case,

$$\mathbf{r}(\alpha) = A\mathbf{x}(\alpha) - \mathbf{d} \approx A\mathbf{x} - \mathbf{d} = \boldsymbol{\eta},$$

with a direct consequence being that $\mathbb{E}(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2) \approx \mathbb{E}(\frac{1}{m}\|\boldsymbol{\eta}\|_2^2) = \sigma^2$ when $\boldsymbol{\eta}$ is real. Thus using the discrepancy principle we find α such that $\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 = \sigma^2$. A similarity exists between the discrepancy principle and the UPRE method in that the variance of the noise in the data must be known for both methods.

Implementation of this method requires finding a solution of $F_{\text{MDP}}(\alpha) = 0$, where $F_{\text{MDP}}(\alpha)$ is defined to be

$$F_{\text{MDP}}(\alpha) = \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 - \sigma^2. \tag{4.2.20}$$

In other words, implementation of the discrepancy principle method is equivalent to finding a root of $F_{\text{MDP}}(\alpha)$. A summation form of the discrepancy principle function is obtained directly from eq. (4.1.8) by substituting the regularized residual term:

$$F_{\text{MDP}}(\alpha) = \frac{1}{m}\left(\sum_{j=1}^{n} \widehat{d}_j^2\psi_j^2(\alpha) + \sum_{j=n+1}^{m} \widehat{d}_j^2\right) - \sigma^2. \tag{4.2.21}$$

As with the UPRE method described in Section 4.1, summation versions of eq. (4.2.21) for the DFT or DCT instead of the GSVD are readily obtained.

Comparing the UPRE function $F_{\text{UPRE}}(\alpha)$ on eq. (4.1.9) with $F_{\text{MDP}}(\alpha)$, it can be seen that

$$F_{\text{MDP}}(\alpha) = F_{\text{UPRE}}(\alpha) - \frac{2\sigma^2}{m}\sum_{j=1}^{n}\phi_j(\alpha). \tag{4.2.22}$$

The function $F_{\text{MDP}}(\alpha)$ will be near zero when the sum in eq. (4.2.21) is close to σ^2. Furthermore, $F_{\text{MDP}}(\alpha)$ is monotone increasing on $(0, \infty)$ because for all $\alpha > 0$,

$$\frac{d}{d\alpha}F_{\text{MDP}}(\alpha) = \frac{2\alpha}{m}\sum_{j=1}^{n}\psi_j(\alpha)\frac{\gamma_j^2}{\left(\gamma_j^2 + \alpha^2\right)^2}\widehat{d}_j^2 \geq 0.$$

The monotonicity of $\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2$ does not guarantee, however, the existence of a zero of $F_{\text{MDP}}(\alpha)$. If the selected value of σ^2 is too large, then it is possible that $F_{\text{MDP}}(\alpha) < 0$ for all $\alpha > 0$ and a root will not exist. This can be attributed to the well-known limiting behavior of $\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2$, which we present as Lem. 4.2.1 to refer to it later.

Lemma 4.2.1

$$\lim_{\alpha\to\infty} \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 \leq \frac{1}{m}\|\mathbf{d}\|_2^2.$$

Proof. Writing $\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2$ in terms of filter functions eq. (2.3.20),

$$\begin{aligned}\lim_{\alpha\to\infty} \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 &= \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2 + \lim_{\alpha\to\infty}\frac{1}{m}\sum_{j=1}^{n}\psi_j^2(\alpha)\widehat{d}_j^2 \\ &= \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2 + \frac{1}{m}\sum_{j=1}^{n}\left(\lim_{\alpha\to\infty}\psi_j^2(\alpha)\right)\widehat{d}_j^2.\end{aligned}$$

Expanding the denominator of $(\psi_j(\alpha))^2$ makes the limit clear:

$$\lim_{\alpha\to\infty}\psi_j^2(\alpha) = \lim_{\alpha\to\infty}\begin{cases}1, & \delta_j < \tau \\ 0, & \lambda_j < \tau \\ \dfrac{\alpha^4}{\alpha^4 + 2\gamma_j^2\alpha^2 + \gamma_j^4}, & \text{otherwise,}\end{cases} = \begin{cases}0, & \lambda_j < \tau \\ 1, & \text{otherwise.}\end{cases}$$

Therefore

$$\begin{aligned}\lim_{\alpha\to\infty} \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 &= \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2 + \frac{1}{m}\sum_{j=1}^{n}\left(\lim_{\alpha\to\infty}\psi_j^2(\alpha)\right)\widehat{d}_j^2 \\ &\leq \frac{1}{m}\sum_{j=n+1}^{m} \widehat{d}_j^2 + \frac{1}{m}\sum_{j=1}^{n}\widehat{d}_j^2 \\ &= \frac{1}{m}\left\|\widehat{\mathbf{d}}\right\|_2^2 \\ &= \frac{1}{m}\left\|\mathbf{d}\right\|_2^2,\end{aligned}$$

with the last equality following from $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$ with orthogonal U. □

Lem. 4.2.1 implies that if the selected value of σ^2 is larger than $\frac{1}{m}\|\mathbf{d}\|_2^2$, $F_{\text{MDP}}(\alpha)$ will not have a root for $\alpha > 0$ and the MDP method fails to select a regularization

parameter. Sometimes a safety parameter $\epsilon > 0$ is introduced to modify the MDP function to

$$F_{\text{MDP}}(\alpha) = \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 - \epsilon\sigma^2. \tag{4.2.23}$$

to account for root-finding difficulties Aster *et al.* (2013); Gazzola *et al.* (2019), though selecting an appropriate value of ϵ is an ad hoc process and depends on the confidence of σ^2 as the true noise. The original MDP function is recovered from eq. (4.2.23) when $\epsilon = 1$.

4.3 Generalized Cross Validation

The UPRE and MDP methods require knowledge of the variance σ^2 of the noise vector $\boldsymbol{\eta}$. In contrast, the generalized cross validation (GCV) method Wahba (1977, 1990) does not require knowledge of σ^2. Again assuming real data coefficients, the GCV function is

$$F_{\text{GCV}}(\alpha) = \frac{\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2}{\left[\frac{1}{m}\operatorname{trace}(I_m - A(\alpha))\right]^2}, \tag{4.3.24}$$

where $\mathbf{r}(\alpha)$ is the regularized residual defined in the derivation of the UPRE method. Similarities between the GCV and UPRE methods are that both functions are estimators of the predictive risk, and the regularization parameter α is chosen as the minimizers of these functions.

By the linearity of the trace operator,

$$\operatorname{trace}(I_m - A(\alpha)) = \operatorname{trace}(I_m) - \operatorname{trace}(A(\alpha)) = m - \operatorname{trace}(A(\alpha)).$$

Then by eq. (4.1.7),

$$\operatorname{trace}(I_m - A(\alpha)) = (m-n) + \sum_{j=1}^{n} \psi_j(\alpha) = (m-n) + \sum_{j=1}^{q^*} \psi_j(\alpha). \tag{4.3.25}$$

Substituting eq. (4.1.8) and eq. (4.3.25) into eq. (4.3.24) produces a summation form

of the GCV function:

$$F_{\mathrm{GCV}}(\alpha) = \frac{\frac{1}{m}\left(\sum_{j=1}^{n} \widehat{d}_j^2 \psi_j^2(\alpha) + \sum_{j=n+1}^{m} \widehat{d}_j^2\right)}{\left[\frac{1}{m}((m-n) + \sum_{j=1}^{q^*} \psi_j(\alpha))\right]^2} = \frac{m\left(\sum_{j=1}^{n} \widehat{d}_j^2 \psi_j^2(\alpha) + \sum_{j=n+1}^{m} \widehat{d}_j^2\right)}{\left[(m-n) + \sum_{j=1}^{q^*} \psi_j(\alpha)\right]^2}. \tag{4.3.26}$$

As with the UPRE and MDP methods, the summation version of eq. (4.3.26) for the DFT and DCT are of the same form, with the primary differences being whether the spectral coefficients of the data are real or complex. However, unlike the UPRE and MDP methods, a rescaling of the data through a diagonal scaling matrix M with $M_{j,j} = 1/\sigma$ as in Lem. 4.1.3 cannot be accomplished for the GCV since that method assumes that variance of the noise is unknown.

4.4 Summary

There are a variety of methods used to select parameters controlling the amount of regularization. In general, methods can be assigned to one of two classes: methods that rely on knowledge of the variance of the noise present in the data, and methods that do not rely on the variance. Common methods falling into the first class are the unbiased predictive risk estimator and discrepancy principle, while a method which falls into the latter class is the generalized cross validation method. Though these methods are distinct, all three share a common term which is a norm of the regularized residual.

Chapter 5

WINDOWED MULTI-DATA FUNCTIONS

The single data set, scalar parameter regularization parameter functions in introduced in Chapter 4 will now be extended to use with multiple data sets and multiple regularization parameters. This chapter presents the background material for the development of these methods, starting with an introduction to windowed regularization in Section 5.1 in which we use an approach in which the parameters are determined by windows on the combined spectrum of the operators A and L. The application of scalar parameter methods to multiple data sets is discussed in Section 5.2, and the background of the full windowed multidata approach is given in Section 5.3. Though complex versions of transformed data were considered in Chapter 4 (specifically $\widehat{\mathbf{d}} = F\mathbf{d}$ for the DFT), here and in the remainder of this document we assume that the components of $\widehat{\mathbf{d}}$ are real. For instance, $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$ for orthogonal U as in the GSVD.

5.1 Windowed Tikhonov Regularization

A more general approach to regularization replaces the single regularization parameter by a vector $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_p]^{\mathsf{T}}$. Here, we follow the approach in Chung *et al.* (2011b) by defining P vectors $\mathbf{w}^{(p)} \in \mathbb{R}^n$ that contain non-negative weights which satisfy

$$\sum_{p=1}^{P} w_j^{(p)} = 1, \quad j = 1, \ldots, n. \tag{5.1.1}$$

Defining windows $W^{(p)} = \text{diag}\left(\mathbf{w}^{(p)}\right)$ for $p = 1, \ldots, P$, we have $\sum_{p=1}^{P} W^{(p)} = I_n$. Considering first the Tikhonov regularization with $L = I_n$ (where we use the SVD of A which is $A = USV^{\mathsf{T}}$), a regularization parameter α_p can be selected for each

window $W^{(p)}$ so that a windowed regularized solution can be constructed as

$$\mathbf{x}_{\text{win}}(\boldsymbol{\alpha}) = \sum_{p=1}^{P} V\left[S^{\mathsf{T}}S + \alpha_p^2\right]^{-1} W^{(p)} S^{\mathsf{T}}\widehat{\mathbf{d}}, \tag{5.1.2}$$

which corresponds to eq. (2.1.6) when $P = 1$. In this framework, the windows must be selected before choosing corresponding regularization parameters.

First we consider non-overlapping windows, $W^{(p)}$, for which the components of their corresponding weight vectors $\mathbf{w}^{(p)}$ satisfy

$$w_j^{(p)} \in \{0, 1\}, \qquad j = 1, \ldots, n, \quad p = 1, \ldots, P. \tag{5.1.3}$$

The condition given by eq. (5.1.3) means that for each $j = 1, \ldots, n$, there is exactly one $p \in \{1, \ldots, P\}$ such that $w_j^{(p)} = 1$. When working with non-overlapping windows, the pigeonhole principle Dummit and Foote (2004) can be used to show that there will exist indices p such that $\mathbf{w}^{(p)} = \mathbf{0}$ if $P > n$. Perhaps the simplest way of choosing the components of $\mathbf{w}^{(p)}$ is to first choose $P+1$ partition values $\omega^{(0)} \geq \ldots \geq \omega^{(P)}$ such that $\omega^{(0)} \geq s_1$ and $s_n > \omega^{(P)}$, then for $p = 1, \ldots, P$, we set

$$w_j^{(p)} = \begin{cases} 1, & \omega^{(p-1)} \geq s_j > \omega^{(p)} \\ 0, & \text{otherwise.} \end{cases} \tag{5.1.4}$$

There are some advantages of using eq. (5.1.4). One advantage is that singular values of similar magnitude are grouped together. Another advantage is that the windowed estimation functions to be discussed in Chapter 6 decouple into linear combinations of functions of single parameters. Choosing $\omega^{(1)}, \ldots, \omega^{(P-1)}$ to be the $P - 1$ linearly spaced or logarithmically spaced points between s_1 and s_n and then setting $\omega^{(0)} = s_1$ and $\omega^{(P)} < s_n$ is an example of how to use eq. (5.1.4).

Partition values $\omega^{(0)} \geq \ldots \geq \omega^{(P)}$ can also be used to generate overlapping windows. For example, cosine windows are defined in Chung *et al.* (2011b) by using

midpoints, $\omega_{\text{mid}}^{(p)}$, so that

$$w_j^{(p)} = \begin{cases} \cos^2\left(\dfrac{\frac{\pi}{2}\left(s_j - \omega_{\text{mid}}^{(p)}\right)}{\omega_{\text{mid}}^{(p-1)} - \omega_{\text{mid}}^{(p)}}\right) & \omega_{\text{mid}}^{(p-1)} \geq s_j > \omega_{\text{mid}}^{(p)}, \\ \cos^2\left(\dfrac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - s_j\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right) & \omega_{\text{mid}}^{(p)} \geq s_j > \omega_{\text{mid}}^{(p+1)}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.1.5}$$

for $p = 2, \ldots, P-1$. The first and Pth weight vectors can be defined by

$$w_j^{(1)} = \begin{cases} 1 & \omega^{(0)} \geq s_j > \omega_{\text{mid}}^{(1)}, \\ \cos^2\left(\dfrac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(1)} - s_j\right)}{\omega_{\text{mid}}^{(1)} - \omega_{\text{mid}}^{(2)}}\right) & \omega_{\text{mid}}^{(1)} \geq s_j > \omega_{\text{mid}}^{(2)}, \\ 0 & \text{otherwise} \end{cases} \tag{5.1.6}$$

and

$$w_j^{(P)} = \begin{cases} \cos^2\left(\dfrac{\frac{\pi}{2}\left(s_j - \omega_{\text{mid}}^{(P)}\right)}{\omega_{\text{mid}}^{(P-1)} - \omega_{\text{mid}}^{(P)}}\right) & \omega_{\text{mid}}^{(P-1)} \geq s_j > \omega_{\text{mid}}^{(P)}, \\ 1 & \omega_{\text{mid}}^{(P)} \geq s_j > \omega^{(P)}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.1.7}$$

respectively. Prop. 5.1.1 shows that such cosine windows satisfy eq. (5.1.1).

Proposition 5.1.1 *The cosine windows given by Chung* et al. *(2011b), presented in eqs.* (5.1.5) *to* (5.1.7)*, satisfy*

$$\sum_{p=1}^{P} w_j^{(p)} = 1, \quad j = 1, \ldots, n.$$

Proof. Let $j \in \{1, \ldots, n\}$ be fixed. If either $\omega^{(0)} \geq s_j > \omega_{\text{mid}}^{(1)}$ or $\omega_{\text{mid}}^{(P)} \geq s_j > \omega^{(P)}$, the result follows immediately from eq. (5.1.6) or eq. (5.1.7), respectively. Now suppose that $\omega_{\text{mid}}^{(1)} \geq s_j > \omega_{\text{mid}}^{(P)}$. Then there exists exactly one value $p \in \{1, \ldots, P-1\}$ such that $\omega_{\text{mid}}^{(p)} \geq s_j > \omega_{\text{mid}}^{(p+1)}$. By definition of the cosine windows, $\mathbf{w}^{(p)}$ and $\mathbf{w}^{(p+1)}$

are the only weight vectors such that their jth components are nonzero. It must be shown that $w_j^{(p)} + w_j^{(p+1)} = 1$. From eq. (5.1.5), the jth component of $\mathbf{w}^{(p)}$ is

$$w_j^{(p)} = \cos^2\left(\frac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - s_j\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right)$$

and the jth component of $\mathbf{w}^{(p+1)}$ is

$$w_j^{(p+1)} = \cos^2\left(\frac{\frac{\pi}{2}\left(s_j - \omega_{\text{mid}}^{(p+1)}\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right).$$

Using the identity $\cos(\theta) = \sin\left(\frac{\pi}{2} - \theta\right)$, we have that

$$\begin{aligned} w_j^{(p+1)} &= \sin^2\left(\frac{\pi}{2} - \frac{\frac{\pi}{2}\left(s_j - \omega_{\text{mid}}^{(p+1)}\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right) \\ &= \sin^2\left(\frac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}\right) - \frac{\pi}{2}\left(s_j - \omega_{\text{mid}}^{(p+1)}\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right) \\ &= \sin^2\left(\frac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - s_j\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right). \end{aligned}$$

Therefore,

$$w_j^{(p)} + w_j^{(p+1)} = \cos^2\left(\frac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - s_j\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right) + \sin^2\left(\frac{\frac{\pi}{2}\left(\omega_{\text{mid}}^{(p)} - s_j\right)}{\omega_{\text{mid}}^{(p)} - \omega_{\text{mid}}^{(p+1)}}\right) = 1.$$

□

For generalized Tikhonov regularization ($L \neq I_n$), a windowed regularized solution similar to eq. (5.1.2) can be obtained using the GSVD,

$$\mathbf{x}_{\text{win}}(\boldsymbol{\alpha}) = \sum_{p=1}^{P} Y\left[\Delta^{\mathsf{T}}\Delta + \alpha_p^2 \Lambda^{\mathsf{T}}\Lambda\right]^{-1} W^{(p)} \Delta^{\mathsf{T}} \widehat{\mathbf{d}}. \tag{5.1.8}$$

In terms of the filter functions eq. (2.3.20), the windowed solution eq. (5.1.8) can be written as

$$\mathbf{x}_{\text{win}}(\boldsymbol{\alpha}) = \sum_{j=1}^{n} \left(\sum_{p=1}^{P} \phi_j\left(\alpha_p\right) w_j^{(p)}\right) \frac{\widehat{d}_{j+k}}{\delta_j} Y_{\cdot,j}.$$

Care must be taken when selecting the weight vectors $\mathbf{w}^{(p)}$ in light of the fact that the generalized singular values are arranged in ascending order. We extend the notation $\Phi(\alpha)$ and $\Psi(\alpha)$ introduced in Section 2.3 to define the $n \times n$ diagonal matrices $\Phi(\alpha_p)$ and $\Psi(\alpha_p)$ by

$$\Phi_{j,j}(\alpha_p) = \phi_j(\alpha_p), \quad \Psi_{j,j}(\alpha_p) = \psi_j(\alpha_p), \quad j = 1:n.$$

Then by again applying the identity $\Delta^{\mathsf{T}} = \Delta^{\mathsf{T}}\Delta\Delta^{\dagger}$ with the fact that $W^{(p)}$ is a diagonal matrix, eq. (5.1.8) is replaced by the general formulation

$$\begin{aligned}
\mathbf{x}_{\text{win}}(\boldsymbol{\alpha}) &= Y \sum_{p=1}^{P} \left[\Delta^{\mathsf{T}}\Delta + \alpha_p^2 \Lambda^{\mathsf{T}}\Lambda\right]^{-1} W^{(p)} \Delta^{\mathsf{T}}\Delta\Delta^{\dagger}\widehat{\mathbf{d}} \\
&= Y \sum_{p=1}^{P} W^{(p)} \left[\Delta^{\mathsf{T}}\Delta + \alpha_p^2 \Lambda^{\mathsf{T}}\Lambda\right]^{-1} \Delta^{\mathsf{T}}\Delta\Delta^{\dagger}\widehat{\mathbf{d}} \\
&= Y \sum_{p=1}^{P} W^{(p)} \Phi(\alpha_p) \Delta^{\dagger}\widehat{\mathbf{d}}.
\end{aligned}$$

5.2 Multiple Data Sets

We now consider the situation where we have a collection of data sets $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ where

$$\mathbf{d}^{(r)} = \mathbf{b}^{(r)} + \boldsymbol{\eta}^{(r)} = A^{(r)}\mathbf{x}^{(r)} + \boldsymbol{\eta}^{(r)}, \quad \boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}^{(r)}, \Sigma^{(r)}). \tag{5.2.9}$$

The vectors $\mathbf{d}^{(r)}$, $\mathbf{b}^{(r)}$, $\boldsymbol{\eta}^{(r)}$, and $\mathbf{0}^{(r)}$ have length m_r, while the vector $\mathbf{x}^{(r)}$ has length n_r. The system matrices $A^{(r)}$ and covariance matrices $\Sigma^{(r)}$ are thus $m_r \times n_r$ and $m_r \times m_r$, respectively. We also assume that the random vectors $\{\boldsymbol{\eta}^{(r)}\}_{r=1}^{R}$ are mutually independent. For given regularization parameters $\alpha^{(r)}$ and penalty matrices $L^{(r)}$ of dimension $q_r \times n_r$, Tikhonov regularization can be performed to produce regularized solutions $\mathbf{x}(\alpha^{(r)})$ that minimize the functionals $T^{(r)}(\mathbf{x}; \alpha^{(r)}) := \|A^{(r)}\mathbf{x} - \mathbf{d}^{(r)}\|_2^2 + \left(\alpha^{(r)}\right)^2 \|L^{(r)}\mathbf{x}\|_2^2$. Typically, a parameter selection method is utilized to select the regularization parameter for each system. Instead, applying the notation introduced

in Section 1.3, let $\widetilde{\mathbf{d}}$ be the vector formed by vertically concatenating the data sets $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ and define the functional

$$\widetilde{T}\left(\widetilde{\mathbf{x}};\widetilde{\alpha}\right) = \|\widetilde{A}\widetilde{\mathbf{x}} - \widetilde{\mathbf{d}}\|_2^2 + \widetilde{\alpha}^2\|\widetilde{L}\widetilde{\mathbf{x}}\|_2^2, \tag{5.2.10}$$

where $\widetilde{\alpha}$ is a single regularization parameter. Notice, by the definition of the 2-norm and the construction of eq. (5.2.10), we could also write

$$\widetilde{T}\left(\widetilde{\mathbf{x}};\widetilde{\alpha}\right) = \sum_{r=1}^{R}\left(\|A^{(r)}\mathbf{x}^{(r)} - \mathbf{d}^{(r)}\|_2^2 + \widetilde{\alpha}^2\|L^{(r)}\mathbf{x}^{(r)}\|_2^2\right) = \sum_{r=1}^{R} T^{(r)}\left(\mathbf{x}^{(r)};\widetilde{\alpha}\right).$$

The advantage of regularizing via eq. (5.2.10) is that we only have to select one parameter instead of R parameters (one for each data set). Assumption 1 summarizes the set-up established in eq. (5.2.9).

Assumption 1 *For $r = 1, \ldots, R$, assume that $\mathbf{b}^{(r)} = A^{(r)}\mathbf{x}^{(r)}$, $\mathbf{d}^{(r)} = \mathbf{b}^{(r)} + \boldsymbol{\eta}^{(r)}$, and $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}^{(r)}, \Sigma^{(r)})$ with the $\boldsymbol{\eta}^{(r)}$ being mutually independent. The vectors $\mathbf{b}^{(r)}$, $\mathbf{d}^{(r)}$, and $\boldsymbol{\eta}^{(r)}$ are of length m_r and $\mathbf{x}^{(r)}$ is of length n_r.*

Now using the GSVD of each pair $(A^{(r)}, L^{(r)}) = \left(U^{(r)}\Delta^{(r)}\left(X^{(r)}\right)^{\mathsf{T}}, V^{(r)}\Lambda^{(r)}\left(X^{(r)}\right)^{\mathsf{T}}\right)$ we obtain $\left(\widetilde{A}, \widetilde{L}\right) = \left(\widetilde{U}\widetilde{\Delta}\widetilde{X}^{\mathsf{T}}, \widetilde{V}\widetilde{\Lambda}\widetilde{X}^{\mathsf{T}}\right)$ yielding the solution

$$\widetilde{\mathbf{x}}(\widetilde{\alpha}) = \widetilde{Y}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}}\widehat{\widetilde{\mathbf{d}}}, \tag{5.2.11}$$

where $\widetilde{Y}$ is the inverse of $\widetilde{X}$ and $\widehat{\widetilde{\mathbf{d}}} = \widetilde{U}^{\mathsf{T}}\widetilde{\mathbf{d}}$, which is analogous to eq. (2.3.19) for a single system. The filter function notation eq. (2.3.20) can be used to define

$$\phi_j^{(r)}(\alpha) = \begin{cases} 0, & \delta_j^{(r)} < \tau \\ 1, & \lambda_j^{(r)} < \tau \\ \dfrac{\left(\gamma_j^{(r)}\right)^2}{\left(\gamma_j^{(r)}\right)^2 + \alpha^2}, & \text{otherwise,} \end{cases} \qquad \psi_j^{(r)}(\alpha) = \begin{cases} 1, & \delta_j^{(r)} < \tau \\ 0, & \lambda_j^{(r)} < \tau \\ \dfrac{\alpha^2}{\left(\gamma_j^{(r)}\right)^2 + \alpha^2}, & \text{otherwise} \end{cases}$$

for $r = 1, \ldots, R$, as well as the $n_r \times n_r$ diagonal matrices

$$\Phi^{(r)}(\alpha) = \operatorname{diag}\left(\phi_1^{(r)}(\alpha), \ldots, \phi_n^{(r)}(\alpha)\right), \quad \Psi^{(r)}(\alpha) = \operatorname{diag}\left(\psi_1^{(r)}(\alpha), \ldots, \psi_n^{(r)}(\alpha)\right).$$

Here we are using the same tolerance τ for each data set. Concatenating $\Phi^{(r)}(\alpha)$ and $\Psi^{(r)}(\alpha)$ via the block diagonal representation, eq. (5.2.11) becomes $\widetilde{\mathbf{x}}(\widetilde{\alpha}) = \widetilde{Y}\widetilde{\Phi}(\alpha)\widetilde{\Delta}^{\dagger}\widehat{\widetilde{\mathbf{d}}}$.

Some of the parameter estimation methods considered in this document rely upon the statistical properties of the noise in the data, so the statistics of $\widetilde{\boldsymbol{\eta}}$ will be addressed. Though eq. (5.2.9) indicates that the random vectors $\{\boldsymbol{\eta}^{(r)}\}_{r=1}^{R}$ are assumed to have zero mean, we can relax this assumption so that $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\boldsymbol{\mu}^{(r)}, \Sigma^{(r)})$ for all $r = 1, \ldots, R$. The distribution of $\widetilde{\boldsymbol{\eta}}$ is then given by Lem. 5.2.1, which follows from the properties of the multivariate normal distribution.

Lemma 5.2.1 *Let $\{\boldsymbol{\eta}^{(r)}\}_{r=1}^{R}$ be a collection of mutually independent random vectors with $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\boldsymbol{\mu}^{(r)}, \Sigma^{(r)})$ for each $r = 1, \ldots, R$. Then $\widetilde{\boldsymbol{\eta}} \sim \mathcal{N}(\widetilde{\boldsymbol{\mu}}, \widetilde{\Sigma})$.*

We conclude Section 5.2 by enumerating and discussing the additional underlying assumptions that will be utilized in Chapter 6.

Assumption 2 *Given $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}^{(r)}, \Sigma^{(r)})$ for $r = 1, \ldots, R$, we assume $\Sigma^{(r)} = \sigma_r^2 I_{m_r}$ (a constant diagonal matrix).*

Assumption 3 *For all $r = 1, \ldots, R$, we assume that $m_r = m$. In other words, we assume that the size of each data vector $\mathbf{d}^{(r)}$ is the same.*

Assumption 4 *We assume that there exist matrices $\Delta \in \mathbb{R}^{m \times n}$ and $\Lambda \in \mathbb{R}^{q \times n}$ such that $A^{(r)} = U^{(r)}\Delta\left(X^{(r)}\right)^{\mathsf{T}}$ and $L^{(r)} = V^{(r)}\Lambda\left(X^{(r)}\right)^{\mathsf{T}}$ for $r = 1, \ldots, R$, where $U^{(r)}$ and $V^{(r)}$ are orthogonal and $X^{(r)}$ is invertible.*

Assumption 5 *For $r = 1, \ldots, R$, assume that $A^{(r)} = A \in \mathbb{R}^{m \times n}$ and $L^{(r)} = L \in \mathbb{R}^{q \times n}$.*

It should be noted that Assumption 2 could be relaxed so that $\Sigma^{(r)}$ is any diagonal matrix $D^{(r)}$, in which case a whitening transformation $C^{(r)}$ could applied so that $\xi^{(r)} = C^{(r)}\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}_{m\times 1}, I_m)$. For example, one could use the zero-phase component analysis (ZCA) transformation $C^{(r)} = \left(\Sigma^{(r)}\right)^{-1/2}$ Bell and Sejnowski (1997).

The strength of the assumptions presented increase in accordance with their numbering. Assumption 1 describes the most general situation involving multiple data sets considered in this document, where the size of each problem and the covariance of each zero-centered multivariate Gaussian noise vector are potentially distinct. Assumption 2 then requires that the covariance matrices be constant diagonal matrices, though the sizes of the problems remain unrestricted. In contrast, Assumption 3 specifies that each data set much be the same size m; however, the size of the solutions may still be district (n_r can differ for different values of r). A consequence of Assumption 4 is that the pairs $(A^{(r)}, L^{(r)})$ all have the same generalized singular values. Though Assumption 4 is strong, it is implied by Assumption 5. Assumption 5 is the strongest, in that it not only implies Assumptions 3 and 4 but demands the solution size be the same and the system and penalty matrices be the same for all problems. A useful consequence of Assumption 5 is that $\Phi^{(r)}(\alpha) = \Phi(\alpha)$ and $\Psi^{(r)}(\alpha) = \Psi(\alpha)$ for all $r = 1, \ldots, R$ and $\alpha > 0$. As a final note, under Assumption 5 we can form $X = [\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(R)}]$ and $D = [\mathbf{d}^{(1)}, \ldots, \mathbf{d}^{(R)}]$ (horizontal concatenations) and write

$$\begin{aligned}\left\|\widetilde{A}\widetilde{\mathbf{x}} - \widetilde{\mathbf{d}}\right\|_2^2 + \|L\widetilde{\mathbf{x}}\|_2^2 &= \sum_{r=1}^{R} \left\|A^{(r)}\mathbf{x}^{(r)} - \mathbf{d}^{(r)}\right\|_2^2 + \left\|L^{(r)}\mathbf{x}^{(r)}\right\|_2^2 \\ &= \|AX - D\|_F^2 + \|LX\|_F^2\end{aligned}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Moreover, if we consider the averages $\overline{\mathbf{x}} = \frac{1}{R}\sum_{r=1}^{R}\mathbf{x}^{(r)}$ and $\overline{\mathbf{d}} = \frac{1}{R}\sum_{r=1}^{R}\mathbf{d}^{(r)}$ then

$$\left\|A\overline{\mathbf{x}} - \overline{\mathbf{d}}\right\|_2^2 + \|L\overline{\mathbf{x}}\|_2^2 \leq \frac{1}{R^2}\left(\|AX - D\|_F^2 + \|LX\|_F^2\right).$$

5.3 Windowed Regularization for Multiple Data Sets

We now consider the case where windowed regularization is applied to multiple data sets. Letting $\boldsymbol{\alpha}^{(r)} = [\alpha_1^{(r)}, \alpha_2^{(r)}, \ldots, \alpha_{P_r}^{(r)}]$ be the P_r regularization parameters used for windowed regularization applied to the rth system described by eq. (5.2.9), we can independently construct regularized solutions

$$\mathbf{x}_{\text{win}}^{(r)}\left(\boldsymbol{\alpha}^{(r)}\right) = \sum_{p=1}^{P_r} Y^{(r)}\left[\left(\Delta^{(r)}\right)^{\mathsf{T}}\Delta^{(r)} + \left(\alpha_p^{(r)}\right)^2\left(\Lambda^{(r)}\right)^{\mathsf{T}}\Lambda^{(r)}\right]^{-1} W^{(r,p)}\left(\Delta^{(r)}\right)^{\mathsf{T}}\widehat{\mathbf{d}}^{(r)}, \tag{5.3.12}$$

where $W^{(r,p)}$ is the pth window for the rth system. Each system can have its own set of windows, meaning that there are a total of $\sum_{r=1}^{R} P_r$ regularization parameters. The primary assumption we make moving forward, however, is that the number of windows $W^{(r,p)}$ for each system is the same, i.e $P_r = P$ for all $r = 1, \ldots, R$. This implies that $\boldsymbol{\alpha}^{(r)}$ are all vectors of length P and there are a total of RP parameters. A stronger assumption would require that the windows are the same across all data sets, which is described by Assumption 6 and used in Chapter 6.

Assumption 6 *For all $r = 1, \ldots, R$ and $p = 1, \ldots, P$, assume that $W^{(r,p)} = W^{(p)}$.*

Analogous to the single parameter case for MD, we define $\widetilde{\mathbf{x}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ as the vertical concatenation of the $\mathbf{x}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})$ for $r = 1, \ldots, R$, where $\widetilde{\boldsymbol{\alpha}} \in \mathbb{R}^P$ is a single vector of parameters. We then have

$$\widetilde{\mathbf{x}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \sum_{p=1}^{P} \widetilde{Y}\left[\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + (\widetilde{\alpha}_p)^2\,\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right]^{-1}\widetilde{W}^{(p)}\widetilde{\Delta}^{\mathsf{T}}\widehat{\widetilde{\mathbf{d}}}, \tag{5.3.13}$$

where $\widetilde{W}^{(p)} = \text{diag}(W^{(1,p)}, \ldots, W^{(R,p)})$ has R diagonal blocks. As the final extension

of notation from Section 5.2, we define

$$\phi_j^{(r)}(\alpha_p) = \begin{cases} 0, & \delta_j^{(r)} < \tau \\ 1, & \lambda_j^{(r)} < \tau \\ \dfrac{\left(\gamma_j^{(r)}\right)^2}{\left(\gamma_j^{(r)}\right)^2 + \alpha_p^2}, & \text{otherwise,} \end{cases} \quad \psi_j^{(r)}(\alpha_p) = \begin{cases} 1, & \delta_j^{(r)} < \tau \\ 0, & \lambda_j^{(r)} < \tau \\ \dfrac{\alpha_p^2}{\left(\gamma_j^{(r)}\right)^2 + \alpha_p^2}, & \text{otherwise} \end{cases} \tag{5.3.14}$$

for $r = 1, \ldots, R$ and $p = 1, \ldots, P$. With the diagonal matrices $\Phi^{(r)}(\widetilde{\alpha}_p)$ and $\Psi^{(r)}(\widetilde{\alpha}_p)$ whose diagonal elements are

$$\Phi_{j,j}^{(r)}(\widetilde{\alpha}_p) = \phi_j^{(r)}(\alpha_p), \quad \Psi_{j,j}^{(r)}(\widetilde{\alpha}_p) = \psi_j^{(r)}(\alpha_p), \quad j = 1:n$$

the appropriate block matrices can be formed:

$$\widetilde{\Phi}(\widetilde{\alpha}_p) = \operatorname{diag}\left(\Phi^{(1)}(\widetilde{\alpha}_p), \ldots, \Phi^{(R)}(\widetilde{\alpha}_p)\right), \quad \widetilde{\Psi}(\widetilde{\alpha}_p) = \operatorname{diag}\left(\Psi^{(1)}(\widetilde{\alpha}_p), \ldots, \Psi^{(R)}(\widetilde{\alpha}_p)\right).$$

Equation (5.3.13) can then be written as

$$\widetilde{\mathbf{x}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \widetilde{Y} \sum_{p=1}^{P} \widetilde{W}^{(p)} \widetilde{\Phi}(\widetilde{\alpha}_p) \widetilde{\Delta}^{\dagger} \widehat{\widetilde{\mathbf{d}}}. \tag{5.3.15}$$

In an effort to streamline notation, we use $\Phi_j^{(r)}(\widetilde{\alpha}_p)$ and $\Psi_j^{(r)}(\widetilde{\alpha}_p)$ to denote the jth element along the main diagonal of the matrices $\Phi^{(r)}(\widetilde{\alpha}_p)$ and $\Psi^{(r)}(\widetilde{\alpha}_p)$, respectively, for $j = 1:n$.

5.4 Windowed Multi-data Residual and Traces

For the windowed multidata case, we extend the notation introduced in Section 5.1 for windowed regularization involving a single data point to define a windowed regularized residual $\mathbf{r}_{\text{win}}(\boldsymbol{\alpha}) = A\mathbf{x}_{\text{win}}(\boldsymbol{\alpha}) - \mathbf{d}$ where the windowed regularized $\mathbf{x}_{\text{win}}(\boldsymbol{\alpha})$ is given by eq. (5.1.8) assuming the use of the GSVD presented in Section 2.3. Similarly,

we generalize the windowed regularization matrix defined in Chung *et al.* (2011b) so that for windowed spectral regularization, the influence matrix $A_{\text{win}}(\boldsymbol{\alpha})$ is

$$A_{\text{win}}(\boldsymbol{\alpha}) = U\Delta \sum_{p=1}^{P} \left(\Delta^{\mathsf{T}}\Delta + \alpha_p^2 \Lambda^{\mathsf{T}}\Lambda\right)^{-1} W^{(p)} \Delta^{\mathsf{T}} U^{\mathsf{T}} = U\Delta \sum_{p=1}^{P} W^{(p)} \Phi\left(\alpha_p\right) \Delta^{\dagger} U^{\mathsf{T}}.$$

If instead we have multiple data sets but each is used with a scalar α, we apply the set-up from Section 5.2 to produce expressions extended from the single data case in Chapter 4 to obtain

$$\widetilde{\mathbf{r}}(\widetilde{\alpha}) = \widetilde{A}\widetilde{\mathbf{x}}(\widetilde{\alpha}) - \widetilde{\mathbf{d}}, \quad \widetilde{A}(\widetilde{\alpha}) = \widetilde{U}\widetilde{\Delta}\Big(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + (\widetilde{\alpha})^2\, \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1} \widetilde{\Delta}^{\mathsf{T}}\widetilde{U}^{\mathsf{T}} = \widetilde{U}\widetilde{\Delta}\widetilde{\Phi}\left(\widetilde{\alpha}\right) \widetilde{\Delta}^{\dagger}\widetilde{U}^{\mathsf{T}},$$

using the GSVD representation of $\widetilde{\mathbf{x}}(\widetilde{\alpha})$ in eq. (5.2.11) and where the diagonal blocks of $\widetilde{A}(\widetilde{\alpha})$ are

$$\begin{aligned} A^{(r)}(\widetilde{\alpha}) &= U^{(r)}\Delta^{(r)}\Big(\Big(\Delta^{(r)}\Big)^{\mathsf{T}}\Delta^{(r)} + \widetilde{\alpha}^2\Big(\Lambda^{(r)}\Big)^{\mathsf{T}}\Lambda^{(r)}\Big)^{-1}\Big(\Delta^{(r)}\Big)^{\mathsf{T}}\Big(U^{(r)}\Big)^{\mathsf{T}} \\ &= U^{(r)}\Delta^{(r)}\Phi^{(r)}\left(\widetilde{\alpha}\right)\Big(\Delta^{(r)}\Big)^{\dagger}\Big(U^{(r)}\Big)^{\mathsf{T}}, \quad r = 1, \ldots, R. \end{aligned}$$

For the most general windowed multidata case, we combine the notations of Sections 5.1 to 5.3 to construct a windowed regularized residual $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \widetilde{A}\widetilde{\mathbf{x}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - \widetilde{\mathbf{d}}$, and windowed influence matrix defined by

$$\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \widetilde{U}\widetilde{\Delta} \sum_{p=1}^{P} \left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + (\widetilde{\alpha}_p)^2\, \widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1} \widetilde{W}^{(p)} \widetilde{\Delta}^{\mathsf{T}}\widetilde{U}^{\mathsf{T}} = \widetilde{U}\widetilde{\Delta} \sum_{p=1}^{P} \widetilde{W}^{(p)} \widetilde{\Phi}\left(\widetilde{\alpha}_p\right) \widetilde{\Delta}^{\dagger}\widetilde{U}^{\mathsf{T}}.$$

$\widetilde{\mathbf{x}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ can be represented by eq. (5.3.13), while the diagonal blocks of $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ are

$$A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}}) = U^{(r)}\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)}\Phi^{(r)}\left(\widetilde{\alpha}_p\right)\Big(\Delta^{(r)}\Big)^{\dagger}\Big(U^{(r)}\Big)^{\mathsf{T}}. \tag{5.4.16}$$

Both $W^{(r,p)}$ and $\Phi^{(r)}(\widetilde{\alpha}_p)$ are $n_r \times n_r$ matrices.

We will now present representations of the necessary norm and trace terms that are used in discussions of the UPRE, MDP, and GCV methods modified for the windowed multidata case, which is the most general situation. The first representation is given by Thm. 5.4.1, which concerns the norm of the regularized residual.

Theorem 5.4.1 (Norm of the windowed residual) *Under Assumption 1, the norm of* $\widetilde{\mathbf{r}}_{win}(\widetilde{\boldsymbol{\alpha}})$ *is*

$$\|\widetilde{\mathbf{r}}_{win}(\widetilde{\boldsymbol{\alpha}})\|_2^2 = \sum_{r=1}^{R} \left\|\mathbf{r}_{win}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2$$

where for each $r = 1, \ldots, R$ *we have*

$$\left\|\mathbf{r}_{win}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 = \begin{cases} \sum_{j=1}^{m_r} \left(\left[\sum_{p=1}^{P} w_{j+k_r}^{(r,p)} \Psi_{j+k_r}^{(r)}(\widetilde{\alpha}_p) \right] \hat{d}_j^{(r)} \right)^2, & m_r \leq n_r, \\ \sum_{j=1}^{n_r} \left(\left[\sum_{p=1}^{P} w_j^{(r,p)} \Psi_j^{(r)}(\widetilde{\alpha}_p) \right] \hat{d}_j^{(r)} \right)^2 + \sum_{j=n_r+1}^{m_r} \left(\hat{d}_j^{(r)} \right)^2, & m_r > n_r. \end{cases}$$

with $k_r = n_r - m_r$.

Proof. We first let $M = \sum_{r=1}^{R} m_r$, which is the length of $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$. Substituting eq. (5.3.15) into the definition of $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$, we have

$$\begin{aligned} \widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) &= \widetilde{U}\widetilde{\Delta} \sum_{p=1}^{P} \widetilde{W}^{(p)} \widetilde{\Phi}(\widetilde{\alpha}_p) \widetilde{\Delta}^{\dagger} \widehat{\widetilde{\mathbf{d}}} - \widetilde{\mathbf{d}} \\ &= \left[\widetilde{U}\widetilde{\Delta} \sum_{p=1}^{P} \widetilde{W}^{(p)} \widetilde{\Phi}(\widetilde{\alpha}_p) \widetilde{\Delta}^{\dagger} \widetilde{U}^{\mathsf{T}} - I_M \right] \widetilde{\mathbf{d}} \\ &= \widetilde{U} \left[\widetilde{\Delta} \sum_{p=1}^{P} \widetilde{W}^{(p)} \widetilde{\Phi}(\widetilde{\alpha}_p) \widetilde{\Delta}^{\dagger} - I_M \right] \widehat{\widetilde{\mathbf{d}}}. \end{aligned}$$

Using the 2-norm and the block structure of the matrices, we can then write

$$\begin{aligned} \|\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2 &= \sum_{r=1}^{R} \left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 \\ &= \sum_{r=1}^{R} \left\| U^{(r)} \left[\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}(\widetilde{\alpha}_p) \left(\Delta^{(r)}\right)^{\dagger} - I_{m_r} \right] \widehat{\mathbf{d}}^{(r)} \right\|_2^2 \\ &= \sum_{r=1}^{R} \left\| \left[\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}(\widetilde{\alpha}_p) \left(\Delta^{(r)}\right)^{\dagger} - I_{m_r} \right] \widehat{\mathbf{d}}^{(r)} \right\|_2^2. \end{aligned}$$

We must now consider the two cases for each $r = 1, \ldots, R$. If $m_r \leq n_r$, then let $k_r = n_r - m_r$ and so that

$$\begin{aligned} &\left\| \left[\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}(\widetilde{\alpha}_p) \left(\Delta^{(r)}\right)^{\dagger} - I_{m_r} \right] \widehat{\mathbf{d}}^{(r)} \right\|_2^2 \\ &= \sum_{j=1}^{m_r} \left(\left[\sum_{p=1}^{P} w_{j+k_r}^{(r,p)} \Psi_{j+k_r, j+k_r}^{(r)}(\widetilde{\alpha}_p) \right] \hat{d}_j^{(r)} \right)^2. \end{aligned}$$

If $m_r > n_r$ instead, then let $k_r = m_r - n_r$. The matrix within the norm has the block form

$$\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}\left(\widetilde{\alpha}_p\right) \left(\Delta^{(r)}\right)^{\dagger} - I_{m_r} = \begin{bmatrix} -\sum_{p=1}^{P} W^{(r,p)} \Psi^{(r)}\left(\widetilde{\alpha}_p\right) & \mathbf{0}_{n_r \times k_r} \\ \mathbf{0}_{k_r \times n_r} & -I_{k_r} \end{bmatrix}.$$

Thus, the norm becomes

$$\begin{aligned} &\left\| \begin{bmatrix} -\sum_{p=1}^{P} W^{(r,p)} \Psi^{(r)}\left(\widetilde{\alpha}_p\right) & \mathbf{0}_{n_r \times k_r} \\ \mathbf{0}_{k_r \times n_r} & -I_{k_r} \end{bmatrix} \widehat{\mathbf{d}}^{(r)} \right\|_2^2 \\ &= \sum_{j=1}^{n_r} \left(\left[\sum_{p=1}^{P} w_j^{(r,p)} \Psi_j^{(r)}\left(\widetilde{\alpha}_p\right) \right] \hat{d}_j^{(r)} \right)^2 + \sum_{j=n_r+1}^{m_r} \left(\hat{d}_j^{(r)}\right)^2. \end{aligned}$$

□

We can also develop an analog of the windowed regularized residual as applied to the average $\overline{\mathbf{d}} = \frac{1}{R}\sum_{r=1}^{R} \mathbf{d}^{(r)}$ under Assumptions 5 to 6. Defining the windowed regularized residual applied to $\overline{\mathbf{d}}$ as $\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha}) = A\overline{\mathbf{x}}_{\text{win}}(\widetilde{\alpha}) - \overline{\mathbf{d}}$ for the windowed regularized solution $\overline{\mathbf{x}}_{\text{win}}(\widetilde{\alpha}) = Y \sum_{p=1}^{P} W^{(p)} \Phi\left(\widetilde{\alpha}_p\right) \Delta^{\dagger} \widehat{\overline{\mathbf{d}}}$ with $\widehat{\overline{\mathbf{d}}} = U^{\mathsf{T}} \overline{\mathbf{d}}$, the following corollary then applies.

Corollary 5.4.1 *Under Assumptions 1 to 6, for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$ we have that*

$$\|\overline{\mathbf{r}}_{win}(\boldsymbol{\alpha})\|_2^2 \leq \frac{1}{R} \|\widetilde{\mathbf{r}}_{win}(\boldsymbol{\alpha})\|_2^2.$$

Proof. Under Assumption 5, $m_r = m$, $n_r = n$, and $(A^{(r)}, L^{(r)}) = (A, L)$ for all $r = 1, \ldots, R$. Without loss of generality, suppose that $m \leq n$ and let $k = n - m$. Then we immediately have, by their definitions,

$$\begin{aligned} \|\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\|_2^2 &= \left\| U \left[\Delta \sum_{p=1}^{P} W^{(p)} \Phi\left(\alpha_p\right) \Delta^{\dagger} - I_m \right] \widehat{\overline{\mathbf{d}}} \right\|_2^2 \\ &= \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k}\left(\alpha_p\right) \right]^2 \left(\hat{\overline{d}}_j\right)^2. \end{aligned}$$

We also have

$$\left(\hat{\bar{d}}_j\right)^2 = \left(\frac{1}{R}\sum_{r=1}^{R}\hat{d}_j^{(r)}\right)^2 \leq \frac{1}{R}\sum_{r=1}^{R}\left(\hat{d}_j^{(r)}\right)^2, \quad j = 1, \ldots, m.$$

Thus, through a change of summation

$$\|\bar{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\|_2^2 \leq \frac{1}{R}\sum_{r=1}^{R}\left(\sum_{j=1}^{m}\left(\left[\sum_{p=1}^{P} w_{j+k}^{(p)}\Psi_{j+k}\left(\alpha_p\right)\right]\hat{d}_j^{(r)}\right)^2\right) = \frac{1}{R}\left\|\tilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2$$

where the last equality follows from Thm. 5.4.1 in conjunction with Assumption 6. □

Prop. 5.4.1 describes how $\|\tilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2$ can be decomposed when working with non-overlapping windows.

Proposition 5.4.1 *For a given $r \in \{1, \ldots, R\}$, if the weight vectors $\{\mathbf{w}^{(r,p)}\}_{p=1}^{P}$ satisfy eq. (5.1.3), then $\|\mathbf{r}_{win}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2$ can be written as*

$$\left\|\mathbf{r}_{win}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 = \sum_{p=1}^{P}\left\|\mathbf{r}_{win}^{(r,p)}\left(\widetilde{\alpha}_p\right)\right\|_2^2$$

where

$$\left\|\mathbf{r}_{win}^{(r,p)}\left(\widetilde{\alpha}_p\right)\right\|_2^2 = \begin{cases} \sum_{j=1}^{m_r}\left(w_{j+k_r}^{(r,p)}\Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right)\hat{d}_j^{(r)}\right)^2, & m_r \leq n_r, \\ \sum_{j=1}^{n_r}\left(w_j^{(r,p)}\Psi_j^{(r)}\left(\widetilde{\alpha}_p\right)\hat{d}_j^{(r)}\right)^2 + \sum_{j=n_r+1}^{m_r}\left(\hat{d}_j^{(r)}\right)^2, & m_r > n_r. \end{cases}$$

with $k_r = n_r - m_r$.

Proof. Given $r \in \{1, \ldots, R\}$, we assume without loss of generality that $m_r \leq n_r$. Using Thm. 5.4.1, we can write $\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2$ as

$$\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 = \sum_{j=1}^{m_r}\left[\sum_{p=1}^{P} w_{j+k_r}^{(r,p)}\Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right)\right]^2\left(\hat{d}_j^{(r)}\right)^2$$

with $k_r = n_r - m_r$. Since the weight vectors $\{\mathbf{w}^{(r,p)}\}_{p=1}^{P}$ satisfy (5.1.3), for each index j there exists exactly one index p such that $w_j^{(r,p)} \neq 0$. Thus, the sum over p has only

one nonzero summand for each index j, meaning we can write

$$\left[\sum_{p=1}^{P} w_{j+k_r}^{(r,p)} \Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right)\right]^2 = \sum_{p=1}^{P} \left(w_{j+k_r}^{(r,p)} \Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right)\right)^2.$$

Therefore $\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2$ can be rewritten through a change of summation so that

$$\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 = \sum_{p=1}^{P} \left[\sum_{j=1}^{m_r} \left(w_{j+k_r}^{(r,p)} \Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right)\right)^2 \left(\hat{d}_j^{(r)}\right)^2\right] = \sum_{p=1}^{P} \left\|\mathbf{r}_{\text{win}}^{(r,p)}(\widetilde{\alpha}_p)\right\|_2^2$$

with

$$\left\|\mathbf{r}_{\text{win}}^{(r,p)}(\widetilde{\alpha}_p)\right\|_2^2 = \sum_{j=1}^{m_r} \left(w_{j+k_r}^{(r,p)} \Psi_{j+k_r}^{(r)}\left(\widetilde{\alpha}_p\right) \hat{d}_j^{(r)}\right)^2.$$

□

Prop. 5.4.1 means that if regularization is being performed with non-overlapping windows, then the norm of the regularized residual can be written as a sum of norms of residuals specific to each window.

As a function of $\boldsymbol{\alpha} \in \mathbb{R}_+^P$, the limiting behavior of $\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\|_2^2$ provides insight into how best to deal with minimizing functions that involve $\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\|_2^2$. Lem. 5.4.1 shows that $\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\|_2^2$ can be bounded above by the scaled norm of the data.

Lemma 5.4.1 *Given $P \in \{1, \ldots, N\}$ windows,*

$$\lim_{\|\boldsymbol{\alpha}\|_2 \to \infty} \|\mathbf{r}_{win}(\boldsymbol{\alpha})\|_2^2 \leq P^2 \|\mathbf{d}\|_2^2, \quad \boldsymbol{\alpha} \in \mathbb{R}_+^P.$$

Proof. From Thm. 5.4.1 with $R = 1$ and $A \in \mathbb{R}^{m \times n}$, we have that

$$\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\|_2^2 = \begin{cases} \sum_{j=1}^{m} \left(\left[\sum_{p=1}^{P} w_{j+k}^{(p)} \psi_{j+k}\left(\alpha_p\right)\right] \hat{d}_j\right)^2, & m \leq n, \\ \sum_{j=1}^{n} \left(\left[\sum_{p=1}^{P} w_j^{(p)} \psi_j\left(\alpha_p\right)\right] \hat{d}_j\right)^2 + \sum_{j=n+1}^{m} \left(\hat{d}_j\right)^2, & m > n. \end{cases}$$

with $k = n - m$. Without loss of generality we assume that $m \leq n$, so

$$\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\|_2^2 = \sum_{j=1}^{m} \left(\left[\sum_{p=1}^{P} w_{j+k}^{(p)} \psi_{j+k}\left(\alpha_p\right)\right] \hat{d}_j\right)^2 = \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \psi_{j+k}\left(\alpha_p\right)\right]^2 \hat{d}_j^2.$$

Applying the Cauchy-Schwarz inequality immediately produces

$$\sum_{j=1}^{m}\left[\sum_{p=1}^{P} w_{j+k}^{(p)}\psi_{j+k}\left(\alpha_p\right)\right]^2 \hat{d}_j^2 \leq \sum_{j=1}^{m}\left[\sum_{p=1}^{P}\left(w_{j+k}^{(p)}\right)^2\right]\left[\sum_{p=1}^{P}\psi_{j+k}^2\left(\alpha_p\right)\right]\hat{d}_j^2.$$

Since $0 \leq w_{j+k}^{(p)} \leq 1$ for all $p = 1, \ldots, P$ and $\sum_{p=1}^{P} w_{j+k}^{(p)} = 1$ by eq. (5.1.1), we can form the inequality

$$\frac{1}{P} = \frac{1}{P}\left(\sum_{p=1}^{P} w_{j+k}^{(p)}\right)^2 \leq \sum_{p=1}^{P}\left(w_{j+k}^{(p)}\right)^2 \leq 1,$$

where the upper bound is attained for non-overlapping windows. Thus we have

$$\left\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 \leq \sum_{j=1}^{m}\left[\sum_{p=1}^{P}\psi_{j+k}^2\left(\alpha_p\right)\right]\hat{d}_j^2.$$

From eq. (5.3.14), we know that $\lim_{\alpha_p\to\infty}\psi_j^2(\alpha_p) \leq 1$ for all $j = 1:n$. Therefore we have

$$\lim_{\|\boldsymbol{\alpha}\|_2\to\infty}\left\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 \leq \sum_{j=1}^{m}\left[\sum_{p=1}^{P}\lim_{\alpha_p\to\infty}\psi_{j+k}^2\left(\alpha_p\right)\right]\hat{d}_j^2 \leq P\sum_{j=1}^{m}\hat{d}_j^2 = P\left\|\mathbf{d}\right\|_2^2,$$

with the last equality following from $\widehat{\mathbf{d}} = U^{\mathsf{T}}\mathbf{d}$ with orthogonal U. □

In practice, Lem. 5.4.1 does not provide a tight bound in the case of overlapping windows since there are weight terms in the sum over p that are strictly less than one. However, Lem. 5.4.1 does provide some insight into the numerical behavior of the functions using $\left\|\mathbf{r}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2$.

Next, Thm. 5.4.2 provides a representation of a general trace term that is needed in Sections 5.5 to 5.7.

Theorem 5.4.2 *Under Assumptions 1 to 3, the trace of* $\widetilde{\Sigma}\widetilde{A}_{win}(\widetilde{\boldsymbol{\alpha}})$ *is*

$$\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{win}(\widetilde{\boldsymbol{\alpha}})\right) = \sum_{r=1}^{R}\left(\sigma_r^2\sum_{j=k_r+1}^{n_r}\left[\sum_{p=1}^{P} w_j^{(r,p)}\Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right]\right)$$

where $k_r = \min\{0, n_r - m_r\}$.

Proof. The diagonal block structures of both $\widetilde{\Sigma} = \text{diag}\left(\Sigma^{(1)}, \ldots, \Sigma^{(R)}\right)$ and $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ allow the trace to be written as a sum of traces:

$$\text{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right) = \sum_{r=1}^{R} \text{trace}\left(\Sigma^{(r)} A_{\text{win}}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right).$$

From Assumption 2, we have that $\Sigma^{(r)} = \sigma_r^2 I_{m_r}$ for all $r = 1, \ldots, R$. Furthermore, representation eq. (5.4.16) of the diagonal blocks of $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ and the similarity invariance of the trace operation allow us to write

$$\begin{aligned}\text{trace}\left(\Sigma^{(r)} A_{\text{win}}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right) &= \sigma_r^2 \,\text{trace}\left(A_{\text{win}}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right) \\ &= \sigma_r^2 \,\text{trace}\left(\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}\left(\widetilde{\alpha}_p\right)\left(\Delta^{(r)}\right)^{\dagger}\right).\end{aligned}$$

Using $k_r = \min\{0, n_r - m_r\}$, we have

$$\sigma_r^2 \,\text{trace}\left(\Delta^{(r)} \sum_{p=1}^{P} W^{(r,p)} \Phi^{(r)}\left(\widetilde{\alpha}_p\right)\left(\Delta^{(r)}\right)^{\dagger}\right) = \sigma_r^2 \sum_{j=k_r+1}^{n_r}\left[\sum_{p=1}^{P} w_j^{(r,p)} \Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right].$$

Therefore,

$$\text{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right) = \sum_{r=1}^{R}\left(\sigma_r^2 \sum_{j=k_r+1}^{n_r}\left[\sum_{p=1}^{P} w_j^{(r,p)} \Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right]\right).$$

□

With the inclusion of Assumption 5 and Assumption 6, we can make a statement regarding the traces of the covariance and influence matrices similar to that of Cor. 5.4.1. We first define $\overline{\Sigma}$ as the covariance matrix of the averaged noise $\overline{\boldsymbol{\eta}} = \frac{1}{R}\sum_{r=1}^{R} \boldsymbol{\eta}^{(r)}$ for the averaged data $\overline{\mathbf{d}}$. Since the random vectors $\{\boldsymbol{\eta}^{(r)}\}_{r=1}^{R}$ are mutually independent and have mean zero, $\mathbb{E}(\overline{\boldsymbol{\eta}}) = \mathbf{0}_{m\times 1}$ and the evaluation of $\overline{\Sigma}$ is reduced to

$$\begin{aligned}\overline{\Sigma} = \text{Cov}\left(\overline{\boldsymbol{\eta}}\,\overline{\boldsymbol{\eta}}^{\mathsf{T}}\right) &= \frac{1}{R^2}\sum_{r,\ell}\mathbb{E}\left(\boldsymbol{\eta}^{(r)}\left[\boldsymbol{\eta}^{(\ell)}\right]^{\mathsf{T}}\right) \\ &= \frac{1}{R^2}\sum_{r=1}^{R}\mathbb{E}\left(\boldsymbol{\eta}^{(r)}\left[\boldsymbol{\eta}^{(r)}\right]^{\mathsf{T}}\right) \\ &= \frac{1}{R^2}\sum_{r=1}^{R}\Sigma^{(r)}. \qquad (5.4.17)\end{aligned}$$

Cor. 5.4.2 describes the analog of Cor. 5.4.1 for matrix traces with averaged noise.

Corollary 5.4.2 *Under Assumptions 1 to 6, we have that*

(i) $\operatorname{trace}\left(\overline{\Sigma}A_{win}\left(\boldsymbol{\alpha}\right)\right) = \frac{1}{R^2}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{win}\left(\boldsymbol{\alpha}\right)\right)$ *for all* $\boldsymbol{\alpha} \in \mathbb{R}_+^P$,

(ii) $\operatorname{trace}\left(\overline{\Sigma}\right) = \frac{1}{R^2}\operatorname{trace}\left(\widetilde{\Sigma}\right)$.

Proof. We prove (item i) only, since (item ii) follows directly from eq. (5.4.17) and the block structure of $\widetilde{\Sigma}$. Using eq. (5.4.17), for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$ we have

$$\operatorname{trace}\left(\overline{\Sigma}A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right) = \operatorname{trace}\left(\left(\frac{1}{R^2}\sum_{r=1}^{R}\Sigma^{(r)}\right)A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right) = \frac{1}{R^2}\sum_{r=1}^{R}\operatorname{trace}\left(\Sigma^{(r)}A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right).$$

Assumption 2 then gives

$$\frac{1}{R^2}\sum_{r=1}^{R}\operatorname{trace}\left(\Sigma^{(r)}A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right) = \frac{1}{R^2}\sum_{r=1}^{R}\operatorname{trace}\left(\sigma_r^2 A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right) = \frac{1}{R^2}\operatorname{trace}\left(A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right)\sum_{r=1}^{R}\sigma_r^2.$$

Shifting to $\frac{1}{R^2}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}\left(\boldsymbol{\alpha}\right)\right)$, Assumptions 5 and 6 imply $A_{\text{win}}^{(r)}(\boldsymbol{\alpha}) = A_{\text{win}}(\boldsymbol{\alpha})$ for all $r = 1, \ldots, R$. This fact combined with the proof of Thm. 5.4.2 yields

$$\begin{aligned}\frac{1}{R^2}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}\left(\boldsymbol{\alpha}\right)\right) &= \frac{1}{R^2}\sum_{r=1}^{R}\operatorname{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}\left(\boldsymbol{\alpha}\right)\right)\\ &= \frac{1}{R^2}\operatorname{trace}\left(A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right)\sum_{r=1}^{R}\sigma_r^2\\ &= \operatorname{trace}\left(\overline{\Sigma}A_{\text{win}}\left(\boldsymbol{\alpha}\right)\right).\end{aligned}$$

□

In contrast to Prop. 5.4.1, it is not necessary to have non-overlapping windows in order to decompose trace terms into traces for each window. The windowed decomposition of the trace is described in Prop. 5.4.2.

Proposition 5.4.2 *Under Assumptions 1 to 3, for all* $r = 1, \ldots, R$ *the trace of* $\Sigma^{(r)}A_{win}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)$ *can be written as*

$$\operatorname{trace}\left(\Sigma^{(r)}A_{win}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right) = \sigma_r^2\sum_{p=1}^{P}\operatorname{trace}\left(A_{win}^{(r,p)}\left(\widetilde{\alpha}^{(p)}\right)\right)$$

where

$$A_{win}^{(r,p)}\left(\widetilde{\alpha}^{(p)}\right) = U^{(r)}\Delta^{(r)}W^{(r,p)}\Phi^{(r)}\left(\widetilde{\alpha}_p\right)\left(\Delta^{(r)}\right)^{\dagger}\left(U^{(r)}\right)^{\mathsf{T}}.$$

Proof. From Thm. 5.4.2, we can write trace $\left(\Sigma^{(r)}A_{\text{win}}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right)$ as

$$\text{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right) = \sigma_r^2\sum_{j=k_r+1}^{n_r}\left[\sum_{p=1}^{P}w_j^{(r,p)}\Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right]$$

with $k_r = \min\{0, n_r - m_r\}$ for each $r = 1, \ldots, R$. Changing the order of summation gives

$$\begin{aligned}\sigma_r^2\sum_{j=k_r+1}^{n_r}\left[\sum_{p=1}^{P}w_j^{(r,p)}\Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right] &= \sigma_r^2\sum_{p=1}^{P}\left[\sum_{j=k_r+1}^{n_r}w_j^{(r,p)}\Phi_j^{(r)}\left(\widetilde{\alpha}_p\right)\right]\\ &= \sigma_r^2\sum_{p=1}^{P}\text{trace}\left(A_{\text{win}}^{(r,p)}\left(\widetilde{\alpha}_p\right)\right).\end{aligned}$$

where $A_{\text{win}}^{(r,p)}\left(\widetilde{\alpha}_p\right) = U^{(r)}\Delta^{(r)}W^{(r,p)}\Phi^{(r)}\left(\widetilde{\alpha}_p\right)\left(\Delta^{(r)}\right)^{\dagger}\left(U^{(r)}\right)^{\mathsf{T}}$. □

5.5 Windowed Multi-data UPRE Functions

We first derive the UPRE function for Tikhonov regularization under the more general condition that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \Sigma)$. To this end, we use the following lemma, which is a generalization of the Trace Lemma stated in (Vogel, 2002, p. 98).

Lemma 5.5.1 *Let $\mathbf{x} \in \mathbb{R}^m$ be a constant vector, $\boldsymbol{\eta}$ be a real random n-vector with $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, $B \in \mathbb{R}^{m\times n}$, and let $\langle\cdot,\cdot\rangle$ be the standard Euclidean inner product. Then*

$$\mathbb{E}\left(\|\mathbf{x} + B\boldsymbol{\eta}\|_2^2\right) = \|\mathbf{x}\|_2^2 + 2\sum_{j=1}^{n}\left(\mathbf{x}^{\mathsf{T}}B\right)_j\mu_j + \text{trace}\left(B\Sigma B^{\mathsf{T}}\right).$$

Proof. By the linearity of the expected value operator and inner product,

$$\mathbb{E}\left(\|\mathbf{x} + B\boldsymbol{\eta}\|_2^2\right) = \mathbb{E}\left(\langle\mathbf{x} + B\boldsymbol{\eta}, \mathbf{x} + B\boldsymbol{\eta}\rangle\right) = \mathbb{E}\left(\|\mathbf{x}\|_2^2\right) + 2\,\mathbb{E}\left(\langle\mathbf{x}, B\boldsymbol{\eta}\rangle\right) + \mathbb{E}\left(\|B\boldsymbol{\eta}\|_2^2\right).$$

$\mathbb{E}(\|\mathbf{x}\|_2^2) = \|\mathbf{x}\|_2^2$ because $\mathbf{x}$ is a constant vector. Moreover, the definition of the Euclidean inner product can be used to write $\mathbb{E}(\langle \mathbf{x}, B\boldsymbol{\eta}\rangle)$ as $\mathbb{E}(\sum_{j=1}^n (\mathbf{x}^\mathsf{T}B)_j \eta_j) = \sum_{j=1}^n (\mathbf{x}^\mathsf{T}B)_j \, \mathbb{E}(\eta_j)$. Thus,

$$\begin{aligned}\mathbb{E}(\|\mathbf{x} + B\boldsymbol{\eta}\|_2^2) &= \|\mathbf{x}\|_2^2 + 2\sum_{j=1}^n (\mathbf{x}^\mathsf{T}B)_j \, \mathbb{E}(\eta_j) + \mathbb{E}(\|B\boldsymbol{\eta}\|_2^2) \\ &= \|\mathbf{x}\|_2^2 + 2\sum_{j=1}^n (\mathbf{x}^\mathsf{T}B)_j \mu_j + \mathbb{E}(\|B\boldsymbol{\eta}\|_2^2).\end{aligned}$$

Focusing on $\mathbb{E}(\|B\boldsymbol{\eta}\|_2^2)$, we can write

$$\mathbb{E}(\|B\boldsymbol{\eta}\|_2^2) = \mathbb{E}\left(\sum_{j=1}^n (B\boldsymbol{\eta})_j^2\right) = \sum_{j=1}^n \mathbb{E}((B\boldsymbol{\eta})_j^2) = \sum_{j=1}^n \mathbb{E}(y_j^2)$$

where $\mathbf{y} = B\boldsymbol{\eta}$. Since $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, we have $\mathbf{y} \sim \mathcal{N}(B\boldsymbol{\mu}, B\Sigma B^\mathsf{T})$ Rao (1973). Lastly, $\mathbb{E}(y_j^2) = \mathrm{Var}(y_j) = (B\Sigma B^\mathsf{T})_{j,j}$ for each $j = 1, \ldots, n$. Therefore $\sum_{j=1}^n \mathbb{E}(y_j^2) = \sum_{j=1}^n \mathrm{Var}(y_j) = \mathrm{trace}(B\Sigma B^\mathsf{T})$ and

$$\mathbb{E}(\|\mathbf{x} + B\boldsymbol{\eta}\|_2^2) = \|\mathbf{x}\|_2^2 + 2\sum_{j=1}^n (\mathbf{x}^\mathsf{T}B)_j \mu_j + \mathrm{trace}(B\Sigma B^\mathsf{T}).$$

□

Returning to the assumption that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, application of Lem. 5.5.1 to the norm of $\mathbf{p}(\alpha)$ and noting that $\mathbb{E}(\boldsymbol{\eta}) = \boldsymbol{\mu} = \mathbf{0}$ yields

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{p}(\alpha)\|_2^2\right) = \frac{1}{m}\|(A(\alpha) - I_m)\, A\mathbf{x}\|_2^2 + \frac{1}{m}\,\mathrm{trace}\left(A(\alpha)\Sigma A^\mathsf{T}(\alpha)\right). \tag{5.5.18}$$

The regularized residual $\mathbf{r}(\alpha)$ can be rewritten as

$$\mathbf{r}(\alpha) = (A(\alpha) - I_m)\, A\mathbf{x} + (A(\alpha) - I_m)\, \boldsymbol{\eta}, \tag{5.5.19}$$

and so applying Lem. 5.5.1 to the norm of $\mathbf{r}(\alpha)$ yields

$$\mathbb{E}\left(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2\right) = \frac{1}{m}\|(A(\alpha) - I_m)\, A\mathbf{x}\|_2^2 + \frac{1}{m}\,\mathrm{trace}\left((A(\alpha) - I_m)^\mathsf{T}\Sigma\,(A(\alpha) - I_m)\right).$$

The trace term can be expanded as

$$\begin{aligned}&\operatorname{trace}\left((A(\alpha)-I_m)^{\mathsf{T}}\Sigma\,(A(\alpha)-I_m)\right)\\&=\operatorname{trace}\left(A^{\mathsf{T}}(\alpha)\Sigma A(\alpha)\right)-\operatorname{trace}\left(A^{\mathsf{T}}(\alpha)\Sigma\right)-\operatorname{trace}\left(\Sigma A(\alpha)\right)+\operatorname{trace}\left(\Sigma\right).\end{aligned}$$

The cyclic property of the trace operator and the fact that Σ and $A(\alpha) = A(A^{\mathsf{T}}A + \alpha^2 L^{\mathsf{T}}L)^{-1}A^{\mathsf{T}}$ are symmetric matrices give

$$\begin{aligned}\operatorname{trace}\left((A(\alpha)-I_m)^{\mathsf{T}}\Sigma\,(A(\alpha)-I_m)\right)&=\operatorname{trace}\left(A(\alpha)\Sigma A^{\mathsf{T}}(\alpha)\right)\\&\quad-2\operatorname{trace}\left(\Sigma A(\alpha)\right)+\operatorname{trace}\left(\Sigma\right),\end{aligned}$$

and so eq. (5.5.18) can be expressed as

$$\mathbb{E}\left(\frac{1}{m}\left\|\mathbf{p}(\alpha)\right\|_2^2\right)=\mathbb{E}\left(\frac{1}{m}\left\|\mathbf{r}(\alpha)\right\|_2^2\right)+\frac{2}{m}\operatorname{trace}\left(\Sigma A(\alpha)\right)-\frac{1}{m}\operatorname{trace}\left(\Sigma\right). \tag{5.5.20}$$

Analogous to the standard UPRE function, we can then redefine $F_{\mathrm{UPRE}}(\alpha)$ as

$$F_{\mathrm{UPRE}}(\alpha)=\frac{1}{m}\left\|\mathbf{r}(\alpha)\right\|_2^2+\frac{2}{m}\operatorname{trace}\left(\Sigma A(\alpha)\right)-\frac{1}{m}\operatorname{trace}\left(\Sigma\right). \tag{5.5.21}$$

The standard UPRE function eq. (4.1.6) is recovered from eq. (5.5.21) if $\Sigma = \sigma^2 I_m$ (which is Assumption 2 with $R = 1$). Lem. 5.5.1 is used in the derivation the main result as well, described by Prop. 5.5.1.

Proposition 5.5.1 *Under Assumption 1, the UPRE function* $\widetilde{F}_{\underset{win}{UPRE}}(\alpha)$ *for the data sets* $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ *and windows* $\{\{W^{(r,p)}\}_{p=1}^{P}\}_{r=1}^{R}$ *is*

$$\widetilde{F}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}})=\frac{1}{M}\sum_{r=1}^{R}m_r F^{(r)}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}}), \tag{5.5.22}$$

where $M = \sum_{r=1}^{R} m_r$ *and*

$$F^{(r)}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}})=\frac{1}{m_r}\left\|\mathbf{r}^{(r)}_{win}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2+\frac{2}{m_r}\operatorname{trace}\left(\Sigma^{(r)}A^{(r)}_{win}(\widetilde{\boldsymbol{\alpha}})\right)-\frac{1}{m_r}\operatorname{trace}\left(\Sigma^{(r)}\right). \tag{5.5.23}$$

Proof. Defining the regularization matrix $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \widetilde{A}\widetilde{Y}\widetilde{\Phi}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\widetilde{\Delta}^{\dagger}\widetilde{U}^{\mathsf{T}}$ and $M = \sum_{r=1}^{R} m_r$, we have

$$\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{d}} = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{b}} + \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\boldsymbol{\eta}},$$
$$\widetilde{\mathbf{p}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{b}} + \widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\widetilde{\boldsymbol{\eta}}.$$

Both the regularized residual $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ and predictive error $\widetilde{\mathbf{p}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ are linear combinations of a deterministic term $\mathbf{f}$ ($\mathbf{f} = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{b}}$ in both cases) and a noise term $B\widetilde{\boldsymbol{\eta}}$ (where $B = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)$ in the case of the regularized residual and $B = \widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ for predictive error). Since Lem. 5.2.1 implies $\widetilde{\boldsymbol{\eta}} \sim \mathcal{N}(\widetilde{\mathbf{0}}, \widetilde{\Sigma})$ with $\widetilde{\mathbf{0}} \in \mathbb{R}^M$, Lem. 5.5.1 can be applied to write the expectations of the squared norms of $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ and $\widetilde{\mathbf{p}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$. Starting with the predictive error,

$$\mathbb{E}\left(\|\widetilde{\mathbf{p}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2\right) = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{b}} + \operatorname{trace}\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\widetilde{\Sigma}\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right)^{\mathsf{T}}\right). \tag{5.5.24}$$

Similarly, the regularized residual is expressed as

$$\mathbb{E}\left(\|\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2\right) = \left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\mathbf{b}} + \operatorname{trace}\left(\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)\widetilde{\Sigma}\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) - I_M\right)^{\mathsf{T}}\right).$$

The trace term that makes up part of the norm of the regularized residual can be expanded as

$$\operatorname{trace}\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\widetilde{\Sigma}\left(\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right)^{\mathsf{T}}\right) - 2\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right) + \operatorname{trace}\left(\widetilde{\Sigma}\right),$$

where the cyclic property of the trace and the symmetry of $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ was utilized. Combining the expanded trace term with eq. (5.5.24), the predictive risk can be written as

$$\mathbb{E}\left(\frac{1}{M}\|\widetilde{\mathbf{p}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2\right) = \frac{1}{M}\left(\|\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2 + 2\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right) - \operatorname{trace}\left(\widetilde{\Sigma}\right)\right).$$

Distributing the scale factor, the windowed multi-data UPRE function is

$$\widetilde{F}_{\substack{\text{UPRE}\\\text{win}}}(\alpha) = \frac{1}{M}\|\widetilde{\mathbf{r}}_{\text{win}}(\alpha)\|_2^2 + \frac{2}{M}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\right) - \frac{1}{M}\operatorname{trace}\left(\widetilde{\Sigma}\right). \tag{5.5.25}$$

As used in the proofs of Thm. 5.4.1 to 5.4.2, the structure of $\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$, $\widetilde{\Sigma}$, and $\widetilde{A}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})$ allow for eq. (5.5.25) to be written in terms of sums over r:

$$\widetilde{F}_{\substack{\text{UPRE}\\\text{win}}}(\alpha) = \frac{1}{M}\sum_{r=1}^{R}\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 + \frac{2}{M}\sum_{r=1}^{R}\text{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right) - \frac{1}{M}\sum_{r=1}^{R}\text{trace}\left(\Sigma^{(r)}\right).$$

The scale factors can be rewritten so that

$$\begin{aligned}\widetilde{F}_{\substack{\text{UPRE}\\\text{win}}}(\alpha) &= \frac{1}{M}\sum_{r=1}^{R}\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 + \frac{2}{M}\sum_{r=1}^{R}\text{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right) - \frac{1}{M}\sum_{r=1}^{R}\text{trace}\left(\Sigma^{(r)}\right)\\ &= \frac{1}{M}\sum_{r=1}^{R}\left(\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 + 2\,\text{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right) - \text{trace}\left(\Sigma^{(r)}\right)\right)\\ &= \frac{1}{M}\sum_{r=1}^{R} m_r F_{\substack{\text{UPRE}\\\text{win}}}^{(r)}(\widetilde{\boldsymbol{\alpha}}),\end{aligned}$$

where for $r = 1, \ldots, R$,

$$F_{\text{UPRE}}^{(r)}(\alpha) = \frac{1}{m_r}\left\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right\|_2^2 + \frac{2}{m_r}\text{trace}\left(\Sigma^{(r)}A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\right) - \frac{1}{m_r}\text{trace}\left(\Sigma^{(r)}\right).$$

□

The windowed multidata UPRE method for real $\widehat{\mathbf{d}} = U^{\mathsf{H}}\mathbf{d}$ defines

$$\widetilde{\boldsymbol{\alpha}}_{\text{UPRE}} = \underset{\widetilde{\boldsymbol{\alpha}}\in\mathbb{R}_+^P}{\arg\min}\,\widetilde{F}_{\substack{\text{UPRE}\\\text{win}}}(\widetilde{\boldsymbol{\alpha}}).$$

Thm. 5.4.1 provides a filter function representation for $\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2$, while Thm. 5.4.2 provides a representation for $\text{trace}(\Sigma^{(r)}A_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}}))$ under Assumption 2. Additionally, eq. (5.5.23) is equivalent to eq. (4.1.6) under Assumption 2. $M = mR$ under Assumption 3, in which case $\widetilde{F}_{\substack{\text{UPRE}\\\text{win}}}(\widetilde{\boldsymbol{\alpha}})$ is the average of the functions $F_{\substack{\text{UPRE}\\\text{win}}}^{(r)}(\widetilde{\boldsymbol{\alpha}})$.

If the windows being considered for a certain data set are non-overlapping (meaning eq. (5.1.3) is satisfied), then we can write eq. (5.5.23) as a sum of functions of a scalar argument; this is accomplished using Prop. 5.4.1 and presented as Cor. 5.5.1.

Corollary 5.5.1 *Under Assumption 1 with windows* $\{\{W^{(r,p)}\}_{p=1}^{P}\}_{r=1}^{R}$ *which satisfy Prop. 5.4.1 for some* $r \in \{1, \ldots, R\}$*, the UPRE function* $F_{\substack{UPRE\\win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})$ *applied to data*

set $\mathbf{d}^{(r)}$ *can be written as*

$$F^{(r)}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}}) = \sum_{p=1}^{P}\left(\frac{1}{m_r}\left\|\mathbf{r}^{(r,p)}_{win}\left(\widetilde{\alpha}_p\right)\right\|_2^2 + \frac{2\sigma_r^2}{m_r}\operatorname{trace}\left(A^{(r,p)}_{win}\left(\widetilde{\alpha}^{(p)}\right)\right)\right), \tag{5.5.26}$$

where the trailing trace term in eq. (5.5.23) *is ignored.*

The proof of Cor. 5.5.1 follows via direct application of Prop. 5.4.1 and Prop. 5.4.2 (recall that Prop. 5.4.2 does not rely on having non-overlapping windows). The fact that the trailing trace term $\frac{1}{m_r}\operatorname{trace}\left(\Sigma^{(r)}\right)$ from eq. (5.5.23) is ignored does not affect the estimation of parameters $\widetilde{\alpha}_p$ because such terms do not affect minimization. Attempting to decompose $\frac{1}{m_r}\operatorname{trace}\left(\Sigma^{(r)}\right)$ as a sum over p raises the question of how much variance to attribute to the data in the spectral domain for each spectral window, which is unclear. If the windows $\{\{W^{(r,p)}\}_{p=1}^{P}\}_{r=1}^{R}$ satisfy Prop. 5.4.1 for all $r = 1, \ldots, R$, then we obtain

Corollary 5.5.2 *Under Assumption 1 with windows* $\{\{W^{(r,p)}\}_{p=1}^{P}\}_{r=1}^{R}$ *which satisfy Prop. 5.4.1 for all* $r = 1, \ldots, R$*, the UPRE function* $\widetilde{F}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}})$ *applied to data sets* $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ *can be written as*

$$\widetilde{F}_{\underset{win}{UPRE}}(\widetilde{\boldsymbol{\alpha}}) = \sum_{p=1}^{P}\left[\sum_{r=1}^{R}\left(\frac{1}{m_r}\left\|\mathbf{r}^{(r,p)}_{win}\left(\widetilde{\alpha}_p\right)\right\|_2^2 + \frac{2\sigma_r^2}{m_r}\operatorname{trace}\left(A^{(r,p)}_{win}\left(\widetilde{\alpha}^{(p)}\right)\right)\right)\right], \tag{5.5.27}$$

where the trailing trace terms from eq. (5.5.23) *is ignored.*

The proof of Cor. 5.5.2 follows from using Cor. 5.5.1 for each $r = 1, \ldots, R$ and changing the order of summation. The advantage of having eq. (5.5.27) is that even in the multidata setting, the components of a parameter vector which minimizes $\widetilde{F}_{\underset{win}{\mathrm{UPRE}}}(\widetilde{\boldsymbol{\alpha}})$ can be found individually by minimizing P functions of a scalar argument.

We conclude Section 5.5 by showing a relationship between the windowed multidata UPRE method that uses eq. (5.5.22) and the UPRE method as applied to averaged data $\overline{\mathbf{d}}$, denoted $\overline{F}_{\underset{win}{\mathrm{UPRE}}}(\boldsymbol{\alpha})$.

Proposition 5.5.2 *Under Assumptions 1 to 6, let*

$$\overline{F}_{\underset{win}{UPRE}}(\boldsymbol{\alpha}) = \frac{1}{m}\left\|\overline{\mathbf{r}}_{win}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{m}\operatorname{trace}\left(\overline{\Sigma}A_{win}(\boldsymbol{\alpha})\right) - \frac{1}{m}\operatorname{trace}\left(\overline{\Sigma}\right).$$

Then for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$,

$$\overline{F}_{\underset{win}{UPRE}}(\boldsymbol{\alpha}) \leq \widetilde{F}_{\underset{win}{UPRE}}(\boldsymbol{\alpha}) + \left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right). \tag{5.5.28}$$

Proof. Using Cor. 5.4.2 and $M = mR$, $\overline{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha})$ can be written as

$$\begin{aligned}
\overline{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha}) &= \frac{1}{m}\left\|\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{m}\operatorname{trace}\left(\overline{\Sigma}A_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{m}\operatorname{trace}\left(\overline{\Sigma}\right) \\
&= \frac{1}{m}\left\|\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{R^2m}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{R^2m}\operatorname{trace}\left(\widetilde{\Sigma}\right) \\
&= \frac{1}{m}\left\|\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right).
\end{aligned}$$

Cor. 5.4.1 is then used to bound the norm of the regularized residual:

$$\begin{aligned}
\overline{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha}) &= \frac{1}{m}\left\|\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right) \\
&\leq \frac{1}{Rm}\left\|\widetilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right) \\
&= \frac{1}{M}\left\|\widetilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\right\|_2^2 + \frac{2}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right).
\end{aligned}$$

$\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) \geq 0$ for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$ from the representation given by Thm. 5.4.2, and so we write

$$\frac{2}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) \leq \frac{2}{M}\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right).$$

As for the term containing $\operatorname{trace}\left(\widetilde{\Sigma}\right)$, we subtract and add $\left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right)$ so that

$$\begin{aligned}
-\frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right) &= -\frac{1}{RM}\operatorname{trace}\left(\widetilde{\Sigma}\right) - \left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right) + \left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right) \\
&= -\operatorname{trace}\left(\widetilde{\Sigma}\right)\left[\frac{1}{RM} + \frac{R-1}{RM}\right] + \left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right) \\
&= -\frac{1}{M}\operatorname{trace}\left(\widetilde{\Sigma}\right) + \left(\frac{R-1}{RM}\right)\operatorname{trace}\left(\widetilde{\Sigma}\right).
\end{aligned}$$

Finally we can write

$$\begin{aligned}\overline{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha}) &\leq \frac{1}{M}\left[\|\widetilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})\|_2^2 + 2\,\text{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right) - \text{trace}\left(\widetilde{\Sigma}\right)\right] + \left(\frac{R-1}{RM}\right)\text{trace}\left(\widetilde{\Sigma}\right) \\ &= \widetilde{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha}) + \left(\frac{R-1}{RM}\right)\text{trace}\left(\widetilde{\Sigma}\right).\end{aligned}$$

□

Prop. 5.5.2 shows that the windowed multidata UPRE function $\widetilde{F}_{\underset{\text{win}}{\text{UPRE}}}$ is truly distinct from simply applying the UPRE method to the average of the data, and the bound eq. (5.5.28) provides a description of the relationship between the two modalities. However, Prop. 5.5.2 does not provide information regarding the parameters found using $\overline{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha})$ compared with those found using $\widetilde{F}_{\underset{\text{win}}{\text{UPRE}}}(\boldsymbol{\alpha})$.

5.6 Windowed Multi-data MDP Functions

The MDP function for the more general assumption that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ will first be presented. Since $\mathbb{E}(\frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2) = E(\frac{1}{m}\|\boldsymbol{\eta}\|_2^2)$ and $E(\|\boldsymbol{\eta}\|_2^2) = \text{trace}(\Sigma)$, the MDP function can be redefined to be

$$F_{\text{MDP}}(\alpha) = \frac{1}{m}\|\mathbf{r}(\alpha)\|_2^2 - \frac{1}{m}\,\text{trace}\,(\Sigma)\,. \tag{5.6.29}$$

Equation (5.6.29) can then be applied to data $\{\mathbf{d}^{(r)}\}_{r=1}^R$ and windows $\{\{W^{(r,p)}\}_{p=1}^P\}_{r=1}^R$ to obtain the main MDP result, which is described by Prop. 5.6.1.

Proposition 5.6.1 *Under Assumption 1, the MDP function $\widetilde{F}_{\underset{win}{MDP}}(\widetilde{\boldsymbol{\alpha}})$ for the data sets $\{\mathbf{d}^{(r)}\}_{r=1}^R$ and windows $\{\{W^{(r,p)}\}_{p=1}^P\}_{r=1}^R$ is*

$$\widetilde{F}_{\underset{win}{MDP}}(\widetilde{\boldsymbol{\alpha}}) = \frac{1}{M}\sum_{r=1}^R m_r F^{(r)}_{\underset{win}{MDP}}(\widetilde{\boldsymbol{\alpha}}), \tag{5.6.30}$$

where $M = \sum_{r=1}^R m_r$ and

$$F^{(r)}_{\underset{win}{MDP}}(\widetilde{\boldsymbol{\alpha}}) = \frac{1}{m_r}\|\mathbf{r}^{(r)}_{win}(\widetilde{\boldsymbol{\alpha}})\|_2^2 - \frac{1}{m_r}\,\text{trace}\left(\Sigma^{(r)}\right), \quad r = 1, \ldots, R. \tag{5.6.31}$$

Proof. Using Lem. 5.2.1 for $\widetilde{\boldsymbol{\eta}}$ with $M = \sum_{r=1}^{R} m_r$, the MDP function eq. (5.6.29) can be applied to define

$$\widetilde{F}_{\substack{\text{MDP}\\\text{win}}}(\widetilde{\boldsymbol{\alpha}}) = \frac{1}{M}\|\widetilde{\mathbf{r}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}})\|_2^2 - \frac{1}{M}\,\text{trace}\left(\widetilde{\Sigma}\right). \tag{5.6.32}$$

Defining the individual MDP functions by eq. (5.6.31), the MDP function eq. (5.6.32) for the large system can be rewritten by exploiting the block structure of $\widetilde{A}(\alpha)$ and $\widetilde{\Sigma}$ as follows:

$$\begin{aligned}
\widetilde{F}_{\substack{\text{MDP}\\\text{win}}}(\widetilde{\boldsymbol{\alpha}}) &= \frac{1}{M}\sum_{r=1}^{R}\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2 - \frac{1}{M}\sum_{r=1}^{R}\text{trace}\left(\Sigma^{(r)}\right)\\
&= \frac{1}{M}\sum_{r=1}^{R}\left(\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2 - \text{trace}\left(\Sigma^{(r)}\right)\right)\\
&= \frac{1}{M}\sum_{r=1}^{R} m_r\left(\frac{1}{m_r}\|\mathbf{r}_{\text{win}}^{(r)}(\widetilde{\boldsymbol{\alpha}})\|_2^2 - \frac{1}{m_r}\,\text{trace}\left(\Sigma^{(r)}\right)\right)\\
&= \frac{1}{R}\sum_{r=1}^{R} F_{\substack{\text{MDP}\\\text{win}}}^{(r)}(\widetilde{\boldsymbol{\alpha}}).
\end{aligned}$$

□

The windowed multidata MDP method then defines $\widetilde{\boldsymbol{\alpha}}_{\text{MDP}}$ as the zero of $\widetilde{F}_{\substack{\text{MDP}\\\text{win}}}(\widetilde{\boldsymbol{\alpha}})$. Analogous to the UPRE method, eq. (5.6.31) is equivalent to eq. (4.2.20) under Assumption 2. A safety parameter ϵ can also be included in the trace terms of eq. (5.6.29) and eq. (5.6.32) for more control over selected parameters as described in Section 4.2. If ϵ is included in the trace term of eq. (5.6.32), Prop. 5.6.1 suggests a partition $\epsilon = \sum_{r=1}^{R}\epsilon_r$ among the trace terms of the individual functions in eq. (5.6.31); however, this situation is not considered here.

As with the windowed multidata UPRE method, the windowed multidata MDP method is distinct from the MDP method as applied to the averaged data $\overline{\mathbf{d}}$.

Proposition 5.6.2 *Under Assumptions 1 to 6, let*

$$\overline{F}_{\substack{MDP\\win}}(\boldsymbol{\alpha}) = \frac{1}{m}\,\|\overline{\mathbf{r}}_{win}(\boldsymbol{\alpha})\|_2^2 - \frac{1}{m}\,\text{trace}\left(\overline{\Sigma}\right).$$

Then for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$,

$$\overline{F}_{\substack{MDP \\ win}}(\boldsymbol{\alpha}) \leq \widetilde{F}_{\substack{MDP \\ win}}(\boldsymbol{\alpha}) + \left(\frac{R-1}{RM}\right) \operatorname{trace}\left(\widetilde{\Sigma}\right).$$

The proof of Prop. 5.6.2 is nearly identical to that of Prop. 5.5.2, with the exclusion of a term involving $\operatorname{trace}\left(\widetilde{\Sigma}\widetilde{A}_{\text{win}}(\boldsymbol{\alpha})\right)$.

To conclude the theoretical discussion of the MDP method, we note that it would be desirable to have results for the MDP method analogous to Cor. 5.5.1 and Cor. 5.5.2 when non-overlapping windows are utilized. However, the difference in how these functions are implemented to estimate regularization parameters (namely, the UPRE method is a minimization problem and the MDP method is a root-finding problem) prevents such results from being generated. Indeed, the constant terms that were ignored in the formulation of Cor. 5.5.1 and Cor. 5.5.2 are the same such terms used in the windowed multidata MDP function; ignoring these terms would affect the root-finding process on which the MDP method relies.

5.7 Windowed Multi-data GCV Functions

We first highlight that, in contrast to the UPRE and MDP methods, the windowed multidata GCV function is distinct in its form from that of the single parameter version applied to a single data set. The first step in the derivation of a windowed multidata GCV function is to define $\widetilde{\mathbf{y}}(\widetilde{\alpha}) = \widetilde{X}^{\mathsf{T}}\widetilde{\mathbf{x}}(\widetilde{\alpha})$, where 5.2.11 gives

$$\widetilde{\mathbf{y}}_{\text{win}}(\widetilde{\alpha}) = \left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}}\widehat{\widetilde{\mathbf{d}}}. \tag{5.7.33}$$

Note that eq. (5.7.33) corresponds to the solution of the normal equations

$$\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)\widetilde{\mathbf{y}} = \widetilde{\Delta}^{\mathsf{T}}\widehat{\widetilde{\mathbf{d}}}$$

for the problem

$$\min_{\widetilde{\mathbf{y}} \in \mathbb{R}^M} \left\| \widetilde{\Delta}\widetilde{\mathbf{y}} - \widehat{\widetilde{\mathbf{d}}} \right\|_2^2 + \widetilde{\alpha}^2 \left\| \widetilde{\Lambda}\widetilde{\mathbf{y}} \right\|_2^2.$$

Following the approaches in Golub *et al.* (1979) and Chung *et al.* (2011b), we introduce the block diagonal matrix $\widetilde{C} = \text{diag}(C^{(1)}, \ldots, C^{(R)})$ where $C^{(r)}$ is the $m_r \times m_r$ unitary matrix that diagonalizes the circulants for $r = 1, \ldots, R$. Now considering the system $G\tilde{\mathbf{y}} \approx \widetilde{C}\widehat{\widetilde{\mathbf{d}}}$ with $\widetilde{G} = \widetilde{C}\widetilde{\Delta}$, the normal equations are

$$\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)\tilde{\mathbf{y}} = \widetilde{G}^{\mathsf{H}}\widetilde{C}\widehat{\widetilde{\mathbf{d}}}$$

Note that $\widetilde{G}^{\mathsf{H}}\widetilde{G}$ could be written as simply $\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}$ since $\widetilde{C}$ is itself unitary. Under Assumption 6, the windowed solution of the new system is

$$\tilde{\mathbf{y}}_{\text{win}}(\widetilde{\boldsymbol{\alpha}}) = \sum_{p=1}^{P}\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{W}^{(p)}\widetilde{G}^{\mathsf{H}}\widetilde{C}\widehat{\widetilde{\mathbf{d}}} \tag{5.7.34}$$

where $\widetilde{W}^{(p)} = \text{diag}(W^{(p)}, \ldots, W^{(p)})$ has R blocks along the diagonal. As in Chung *et al.* (2011b), we introduce the following resolution matrices for $p = 1, \ldots, P$:

$$\widetilde{G}(\widetilde{\alpha}_p) = \widetilde{G}\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{G}^{\mathsf{H}} = \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p),$$
$$\widetilde{G}_{\text{win}}(\widetilde{\alpha}_p) = \widetilde{G}\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{W}^{(p)}\widetilde{G}^{\mathsf{H}} = \widetilde{G}\widetilde{G}^{\sharp}_{\text{win}}(\widetilde{\alpha}_p).$$

The difference between $\widetilde{G}(\widetilde{\alpha}_p)$ and $\widetilde{G}_{\text{win}}(\widetilde{\alpha}_p)$ is the presence of the window matrix $\widetilde{W}^{(p)}$. The diagonal entries of these resolution matrices will be used later, motivating Lem. 5.7.1.

Lemma 5.7.1 *Under Assumption 1 and Assumption 6, the diagonal entries of the matrices* $I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$ *and* $I_M - \widetilde{G}\widetilde{G}^{\sharp}_{win}(\widetilde{\alpha}_p)$*, denoted by* μ_p *and* ν_p*, respectively, are given by*

$$\mu_p = \frac{1}{M}\left(M - N + \sum_{j=1}^{N}\left(1 - \widetilde{\Phi}_{jj}(\widetilde{\alpha}_p)\right)\right),$$
$$\nu_p = \frac{1}{M}\left(M - N + \sum_{j=1}^{N}\left(1 - \left(\widetilde{\Phi}_{win}\right)_{jj}(\widetilde{\alpha}_p)\right)\right),$$

where $M = \sum_{r=1}^{R} m_r$ *and* $N = \sum_{r=1}^{R} n_r$.

Proof. Let $M = \sum_{r=1}^{R} m_r$ and $N = \sum_{r=1}^{R} n_r$. Expanding $\widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$ with the definition for $\widetilde{G}$, we have

$$\begin{aligned}
\widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p) &= \widetilde{G}\Big(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1}\widetilde{G}^{\mathsf{H}} \\
&= \widetilde{C}\widetilde{\Delta}\Big(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1}\widetilde{\Delta}^{\mathsf{T}}\widetilde{C}^{\mathsf{H}} \\
&= \widetilde{C}\begin{bmatrix} \widetilde{\Phi}(\widetilde{\alpha}_p) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\widetilde{C}^{\mathsf{H}}
\end{aligned}$$

where $\widetilde{\Phi}(\widetilde{\alpha}_p)$ is an $N \times N$ matrix. Performing the subtraction $I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$ then yields

$$I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p) = \widetilde{C}\widetilde{C}^{\mathsf{H}} - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p) = \widetilde{C}\begin{bmatrix} I_N - \widetilde{\Phi}(\widetilde{\alpha}_p) & \mathbf{0} \\ \mathbf{0} & I_{M-N} \end{bmatrix}\widetilde{C}^{\mathsf{H}}.$$

The middle term in this product is diagonal, and thus $I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$ is circulant. Therefore, the diagonal entries of $I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$ are constant and we can write

$$\operatorname{diag}\Big(I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)\Big) = \mu_p I_M.$$

Since $\widetilde{C}$ is unitary,

$$\begin{aligned}
\operatorname{trace}\Big(I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)\Big) &= \operatorname{trace}\left(\begin{bmatrix} I_N - \widetilde{\Phi}(\widetilde{\alpha}_p) & \mathbf{0} \\ \mathbf{0} & I_{M-N} \end{bmatrix}\right) \\
&= \sum_{j=1}^{N}\Big(1 - \widetilde{\Phi}_{jj}(\widetilde{\alpha}_p)\Big) + (M-N),
\end{aligned}$$

and so

$$\mu_p = \frac{1}{M}\left(M - N + \sum_{j=1}^{N}\Big(1 - \widetilde{\Phi}_{jj}(\widetilde{\alpha}_p)\Big)\right), \quad p = 1, \ldots, P.$$

The process for obtaining the diagonal entries of $I_M - \widetilde{G}\widetilde{G}^{\sharp}_{\text{win}}(\widetilde{\alpha}_p)$ is similar. $\square$

The spectral windowed multidata GCV function is at last described by Prop. 5.7.1, the proof of which utilizes Lem. 5.7.1.

Proposition 5.7.1 *Under Assumptions 1 to 5 where $m \geq n$ and L has full rank q^*, $\widetilde{\alpha}_p > 0$, $j = 1 : q^*$, and $0 = \delta_1 = \cdots = \delta_\ell < \delta_{\ell+1} \cdots \leq \delta_n$. The windowed $\widetilde{F}_{\substack{GCV \\ win}}$ function for the generalized Tikhonov regularization is given by*

$$\widetilde{F}_{\substack{GCV \\ win}}(\widetilde{\boldsymbol{\alpha}}) = \frac{1}{M}\left(\sum_{j=q^*+1}^{M}\left(1+\left(\sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}\right)\right)^2\left(\widehat{\widetilde{\mathbf{d}}}\right)_j^2 + \sum_{j=1}^{q^*}\left(1+\left(\sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}\right)-\left(\sum_{p=1}^{P}\frac{1}{\mu_p}\frac{\gamma_j^2 w_j^{(p)}}{\gamma_j^2+\widetilde{\alpha}_p^2}\right)\right)^2\left(\widehat{\widetilde{\mathbf{d}}}\right)_j^2\right)$$

where

$$\mu_p = \frac{1}{M}\left(M - N + \sum_{j=1}^{N}\left(1-\widetilde{\Phi}_{jj}\left(\widetilde{\alpha}_p\right)\right)\right),$$
$$\nu_p = \frac{1}{M}\left(M - N + \sum_{j=1}^{N}\left(1-\left(\widetilde{\Phi}_{win}\right)_{jj}\left(\widetilde{\alpha}_p\right)\right)\right).$$

Equivalently, $\widetilde{F}_{\substack{GCV \\ win}}$ can be written in matrix form as

$$\frac{1}{M}\left\|\left(I_M - \sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}+\widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{W}^{(p)}\widetilde{\Delta}^{\mathsf{T}} + \sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}I_M\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2. \tag{5.7.35}$$

Proof. The Allen PRESS estimates Golub *et al.* (1979) of $\widetilde{\alpha}_p$ minimize

$$F\left(\widetilde{\boldsymbol{\alpha}}\right) = \frac{1}{M}\sum_{k=1}^{M}\left(\left(\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\right)_k - \left(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}\left(\widetilde{\boldsymbol{\alpha}}\right)\right)_k\right)^2, \tag{5.7.36}$$

where $\widetilde{\mathbf{y}}_{\text{win}}^{(k)}\left(\widetilde{\boldsymbol{\alpha}}\right)$ is $\widetilde{\mathbf{y}}_{\text{win}}\left(\widetilde{\boldsymbol{\alpha}}\right)$ without the kth entry. To obtain $\widetilde{\mathbf{y}}_{\text{win}}^{(k)}\left(\widetilde{\boldsymbol{\alpha}}\right)$, the kth row of $\widetilde{G}$ is removed in eq. (5.7.34); in Chung *et al.* (2011a), this is accomplished through multiplication by the symmetric matrix $E_k = I_M - \mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}$ as $E_k\widetilde{G}$. Before expanding terms of eq. (5.7.34) using the product $E_k\widetilde{G}$, it will be convenient to define $\mathbf{c}_k = \widetilde{C}^{\mathsf{H}}\mathbf{e}_k$

which is the kth column of $\widetilde{C}^{\mathsf{H}}$. Expanding $\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{G}$ gives

$$\begin{aligned}\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{G} &= \widetilde{\Delta}^{\mathsf{T}}\widetilde{C}^{\mathsf{H}}\left(I_M - \mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}\right)\left(I_M - \mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}\right)\widetilde{C}\widetilde{\Delta} \\ &= \left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{C}^{\mathsf{H}} - \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{e}_k^{\mathsf{T}}\right)\left(\widetilde{C}\widetilde{\Delta} - \mathbf{e}_k\mathbf{c}_k^{\mathsf{T}}\widetilde{\Delta}\right) \\ &= \widetilde{\Delta}^{\mathsf{T}}\widetilde{C}^{\mathsf{H}}\widetilde{C}\widetilde{\Delta} - \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{c}_k^{\mathsf{H}}\widetilde{\Delta} - \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{c}_k^{\mathsf{H}}\widetilde{\Delta} + \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{e}_k^{\mathsf{T}}\mathbf{e}_k\mathbf{c}_k^{\mathsf{T}}\widetilde{\Delta} \\ &= \widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} - \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{c}_k^{\mathsf{H}}\widetilde{\Delta},\end{aligned}$$

where the final equality is obtained from $\widetilde{C}$ being unitary and $\mathbf{e}_k^{\mathsf{T}}\mathbf{e}_k = 1$. Thus the inverse term in eq. (5.7.34) using $E_k\widetilde{G}$ can be written as

$$\left(\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1} = \left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda} - \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k\mathbf{c}_k^{\mathsf{H}}\widetilde{\Delta}\right)^{-1} = \left(H(\widetilde{\alpha}_p) - \mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}\right)^{-1}$$

with $\mathbf{h}_k = \widetilde{\Delta}^{\mathsf{T}}\mathbf{c}_k$ and $H(\widetilde{\alpha}_p) = \widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}$. The Sherman-Morrison formula Golub and Van Loan (2013) then gives

$$\left(H(\widetilde{\alpha}_p) - \mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}\right)^{-1} = (H(\widetilde{\alpha}_p))^{-1} + \frac{(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}}{1 - \mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k}. \tag{5.7.37}$$

Expanding the denominator in eq. (5.7.37) using the definitions of $H(\widetilde{\alpha}_p)$, $\mathbf{h}_k$, $\mathbf{c}_k$, and subsequently $\widetilde{G}$ and $\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)$, yields

$$\begin{aligned}1 - \mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k &= 1 - \mathbf{e}_k^{\mathsf{T}}\widetilde{C}\widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}}\widetilde{C}^{\mathsf{H}}\mathbf{e}_k \\ &= 1 - \mathbf{e}_k^{\mathsf{T}}\widetilde{G}\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{G}^{\mathsf{H}}\mathbf{e}_k \\ &= \mathbf{e}_k^{\mathsf{T}}\left(I_M - \widetilde{G}\left(\widetilde{G}^{\mathsf{H}}\widetilde{G} + \widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{G}^{\mathsf{H}}\right)\mathbf{e}_k \\ &= \mathbf{e}_k^{\mathsf{T}}\left(I_M - \widetilde{G}\widetilde{G}^{\sharp}(\widetilde{\alpha}_p)\right)\mathbf{e}_k.\end{aligned}$$

Thus $1 - \mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k = \mu_p$ (independent of k) by Lem. 5.7.1, and so

$$\left(H(\widetilde{\alpha}_p) - \mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}\right)^{-1} = (H(\widetilde{\alpha}_p))^{-1} + \frac{1}{\mu_p}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}.$$

As a final step before fully expanding the expression for $(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}}))_k$ from eq. (5.7.34),

note that

$$\begin{aligned}
\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\Big(\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{G}+\widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1} &= \mathbf{h}_k^{\mathsf{H}}\Big(H(\widetilde{\alpha}_p)-\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}\Big)^{-1}\\
&= \mathbf{h}_k^{\mathsf{H}}\left((H(\widetilde{\alpha}_p))^{-1}+\frac{1}{\mu_p}(H(\widetilde{\boldsymbol{\alpha}}_p))^{-1}\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\right)\\
&= \mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}+\frac{1}{\mu_p}\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\\
&= \left(1+\frac{1}{\mu_p}\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\mathbf{h}_k\right)\mathbf{h}_k^{\mathsf{H}}(H(\widetilde{\alpha}_p))^{-1}\\
&= \frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\Big(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}+\widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1}.
\end{aligned}$$

Equation (5.7.34) can now be used to express $(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}}))_k$ as follows:

$$\begin{aligned}
\Big(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}})\Big)_k &= \mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}})\\
&= \sum_{p=1}^{P}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\Big(\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{G}+\widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1}\widetilde{W}^{(p)}\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\\
&= \sum_{p=1}^{P}\frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\Big(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta}+\widetilde{\alpha}_p^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\Big)^{-1}\widetilde{W}^{(p)}\widetilde{G}^{\mathsf{H}}E_k^{\mathsf{T}}E_k\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\\
&= \sum_{p=1}^{P}\frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)\left(I_M-\mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\\
&= \sum_{p=1}^{P}\frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)\,\widetilde{C}\widehat{\widetilde{\mathbf{d}}}-\frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)\,\mathbf{e}_k\mathbf{e}_k^{\mathsf{T}}\widetilde{C}\widehat{\widetilde{\mathbf{d}}}.
\end{aligned}$$

The term $\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)\,\mathbf{e}_k$ is the kth diagonal element of $\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)$, which is constant (independent of k) by Lem. 5.7.1 and equal to $1-\nu_p$. Thus

$$\begin{aligned}
\Big(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}})\Big)_k &= \sum_{p=1}^{P}\frac{1}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)\,\widetilde{C}\widehat{\widetilde{\mathbf{d}}}-\frac{1-\nu_p}{\mu_p}\mathbf{e}_k^{\mathsf{T}}\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\\
&= \mathbf{e}_k^{\mathsf{T}}\left(\sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)-\frac{1-\nu_p}{\mu_p}I_M\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}}.
\end{aligned}$$

Subtracting $(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}}))_k$ from $(\widetilde{C}\widehat{\widetilde{\mathbf{d}}})_k$ can now be written as

$$\begin{aligned}
(\widetilde{C}\widehat{\widetilde{\mathbf{d}}})_k-\Big(\widetilde{G}\widetilde{\mathbf{y}}_{\text{win}}^{(k)}(\widetilde{\boldsymbol{\alpha}})\Big)_k &= \mathbf{e}_k^{\mathsf{T}}\widetilde{C}\widehat{\widetilde{\mathbf{d}}}-\mathbf{e}_k^{\mathsf{T}}\left(\sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)-\frac{1-\nu_p}{\mu_p}I_M\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\\
&= \mathbf{e}_k^{\mathsf{T}}\left(I_M-\sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{G}\widetilde{G}_{\text{win}}^{\sharp}(\widetilde{\alpha}_p)+\sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}I_M\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}},
\end{aligned}$$

and therefore 5.7.36 is equivalent to

$$F\left(\widetilde{\boldsymbol{\alpha}}\right) = \frac{1}{M}\left\|\left(I_M - \sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{G}\widetilde{G}^{\sharp}_{\text{win}}\left(\widetilde{\alpha}_p\right) + \sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}I_M\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\right\|_2^2.$$

The windowed multidata GCV function is finally obtained by rewriting the norm as follows:

$$\begin{aligned}
F\left(\widetilde{\boldsymbol{\alpha}}\right) &= \frac{1}{M}\left\|\left(\widetilde{C}\widetilde{C}^{\mathsf{H}} - \sum_{p=1}^{P}\frac{1}{\mu_p}\widetilde{C}\begin{bmatrix}\widetilde{\Phi}_{\text{win}}\left(\widetilde{\alpha}_p\right) & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\widetilde{C}^{\mathsf{H}} + \sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}\widetilde{C}\widetilde{C}^{\mathsf{H}}\right)\widetilde{C}\widehat{\widetilde{\mathbf{d}}}\right\|_2^2\\
&= \frac{1}{M}\left\|\left(I_M - \sum_{p=1}^{P}\frac{1}{\mu_p}\begin{bmatrix}\widetilde{\Phi}_{\text{win}}\left(\widetilde{\alpha}_p\right) & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix} + \sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}I_M\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2\\
&= \frac{1}{M}\left(\sum_{j=q^*+1}^{M}\left(1+\left(\sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}\right)\right)^2\hat{d}_j^2\right.\\
&\left.+\sum_{j=1}^{q^*}\left(1+\left(\sum_{p=1}^{P}\frac{1-\nu_p}{\mu_p}\right)-\left(\sum_{p=1}^{P}\frac{1}{\mu_p}\frac{\gamma_j^2 w_j^{(p)}}{\gamma_j^2+\widetilde{\alpha}_p^2}\right)\right)^2\hat{d}_j^2\right).
\end{aligned}$$

□

The windowed multidata GCV method defines $\widetilde{\boldsymbol{\alpha}}_{\text{GCV}} = \arg\min_{\widetilde{\boldsymbol{\alpha}}>0}\widetilde{F}_{\underset{\text{win}}{\text{GCV}}}(\widetilde{\boldsymbol{\alpha}})$. However, the single parameter multidata GCV function is analogous to the standard GCV function eq. (4.3.24) in its form and can be obtained from Prop. 5.7.1 with $P = 1$; this is shown by 5.7.2.

Proposition 5.7.2 *Under Assumptions 1 to 4, the GCV function $\widetilde{F}_{GCV}(\widetilde{\alpha})$ for the data sets $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ is*

$$\widetilde{F}_{GCV}(\widetilde{\alpha}) = \frac{\frac{1}{M}\|\widetilde{\mathbf{r}}(\widetilde{\alpha})\|_2^2}{\left[1-\frac{1}{M}\operatorname{trace}\left(\widetilde{A}(\widetilde{\alpha})\right)\right]^2}, \tag{5.7.38}$$

where $M = \sum_{r=1}^{R} m_r$.

Proof. Using eq. (5.7.35) from Prop. 5.7.1 with $P = 1$, we have

$$\widetilde{F}_{\mathrm{GCV}}(\widetilde{\alpha}) = \frac{1}{M}\left\|\left(I_M - \frac{1}{\mu}\widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}} + \frac{1-\nu}{\mu}I_M\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2.$$

Here $\widetilde{W}^{(p)}$ is replaced with I_M, as is μ_p and ν_p with μ and ν, respectively. Next, $\widetilde{\Phi}_{\mathrm{win}} = \widetilde{\Phi}$ and so $\nu = \mu$; replacing ν with μ allows the function to be written as

$$\begin{aligned}
\widetilde{F}_{\mathrm{GCV}}(\widetilde{\alpha}) &= \frac{1}{M}\left\|\left(I_M - \frac{1}{\mu}\widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}} + \frac{1-\mu}{\mu}I_M\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2 \\
&= \frac{1}{M}\left\|\left(\frac{1}{\mu}I_M - \frac{1}{\mu}\widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}}\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2 \\
&= \frac{1}{M\mu^2}\left\|\left(I_M - \widetilde{\Delta}\left(\widetilde{\Delta}^{\mathsf{T}}\widetilde{\Delta} + \widetilde{\alpha}^2\widetilde{\Lambda}^{\mathsf{T}}\widetilde{\Lambda}\right)^{-1}\widetilde{\Delta}^{\mathsf{T}}\right)\widehat{\widetilde{\mathbf{d}}}\right\|_2^2 \\
&= \frac{1}{M\mu^2}\left\|\widetilde{\mathbf{r}}(\widetilde{\alpha})\right\|_2^2.
\end{aligned}$$

The last step is to rewrite μ^2 with regard to matrix traces:

$$\begin{aligned}
\mu^2 &= \left[\frac{1}{M}\left(M - N + \sum_{j=1}^{N}\left(1 - \widetilde{\Phi}_{jj}\left(\widetilde{\alpha}\right)\right)\right)\right]^2 \\
&= \left[1 + \frac{1}{M}\sum_{j=1}^{N}\widetilde{\Phi}_{jj}\left(\widetilde{\alpha}\right)\right]^2 \\
&= \left[1 - \frac{1}{M}\operatorname{trace}\left(\widetilde{A}(\widetilde{\alpha})\right)\right]^2.
\end{aligned}$$

□

Another distinction between the GCV method and the UPRE and MDP methods is that the single parameter multidata GCV function cannot be readily expresses as a linear combination of individual single parameter GCV functions

$$F^{(r)}_{\mathrm{GCV}}(\alpha) = \frac{\frac{1}{m_r}\|\mathbf{r}^{(r)}(\alpha)\|_2^2}{\left[1 - \frac{1}{m_r}\operatorname{trace}\left(A^{(r)}(\alpha)\right)\right]^2}, \quad r = 1, \ldots, R \tag{5.7.39}$$

when only Assumption 1 and Assumption 3 are considered. Assumption 5 is required to obtain a result similar to Prop. 5.5.1.

Proposition 5.7.3 *Under Assumptions 1 to 5, the GCV function $\widetilde{F}_{GCV}(\widetilde{\alpha})$ for the data sets $\{\mathbf{d}^{(r)}\}_{r=1}^{R}$ is*

$$\widetilde{F}_{GCV}(\widetilde{\boldsymbol{\alpha}}) = \frac{1}{R}\sum_{r=1}^{R} F_{GCV}^{(r)}(\widetilde{\alpha}),$$

where $F_{GCV}^{(r)}(\widetilde{\alpha})$ is defined by eq. (5.7.39).

Proof. Assumption 5 implies that $A^{(r)}(\widetilde{\alpha}) = A(\widetilde{\alpha})$ for each $r = 1, \ldots, R$, as well as $M = \sum_{r=1}^{R} m_r = mR$. Having the same influence matrix for each r is crucial because of the trace term contained in the denominator of the multidata GCV function eq. (5.7.38). From Thm. 5.4.1 to 5.4.2, we can then write

$$\begin{aligned}\widetilde{F}_{\mathrm{GCV}}(\widetilde{\alpha}) = \frac{\frac{1}{M}\sum_{r=1}^{R} \|\mathbf{r}^{(r)}(\widetilde{\alpha})\|_2^2}{\left[1 - \frac{1}{M}\sum_{r=1}^{R} \operatorname{trace}\left(A(\widetilde{\alpha})\right)\right]^2} &= \frac{\frac{1}{R}\sum_{r=1}^{R} \frac{1}{m}\|\mathbf{r}^{(r)}(\widetilde{\alpha})\|_2^2}{\left[1 - \frac{1}{m}\operatorname{trace}\left(A(\widetilde{\alpha})\right)\right]^2} \\ &= \frac{1}{R}\sum_{r=1}^{R} F_{\mathrm{GCV}}^{(r)}(\widetilde{\alpha}).\end{aligned}$$

□

As with the windowed multidata situations, the single parameter multidata GCV method defines $\widetilde{\alpha}_{\mathrm{GCV}} = \arg\min_{\widetilde{\alpha}>0} \widetilde{F}_{\mathrm{GCV}}(\widetilde{\alpha})$. Note that eq. (5.7.39) is equivalent to eq. (4.3.24) without Assumption 2, which is to be expected since the GCV method does not rely on knowledge of the noise variance.

In conclusion, a result analogous to Prop. 5.5.2 and Prop. 5.6.2 can be obtained for the windowed multidata GCV function using Cor. 5.4.1 and Cor. 5.4.2.

Proposition 5.7.4 *Under Assumptions 1 to 6, let*

$$\overline{F}_{GCV}(\alpha) = \frac{\frac{1}{m}\|\overline{\mathbf{r}}(\alpha)\|_2^2}{\left[1 - \frac{1}{m}\operatorname{trace}\left(A(\alpha)\right)\right]^2}.$$

Then for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$,

$$\overline{F}_{GCV}(\alpha) \leq \widetilde{F}_{GCV}(\alpha).$$

Proof. Cor. 5.4.1 provides the bound on the numerator of $\overline{F}_{\mathrm{GCV}}(\alpha)$. Assumption 5 implies $M = mR$ and $A^{(r)}(\alpha) = A(\alpha)$ for all $r = 1, \ldots, R$, and so the denominator of eq. (5.7.38) becomes

$$\begin{aligned}\left[1 - \frac{1}{M}\operatorname{trace}\left(\widetilde{A}(\alpha)\right)\right]^2 &= \left[1 - \frac{1}{M}\sum_{r=1}^{R}\operatorname{trace}\left(A^{(r)}(\alpha)\right)\right]^2 \\ &= \left[1 - \frac{1}{M}\sum_{r=1}^{R}\operatorname{trace}\left(A(\alpha)\right)\right]^2 \\ &= \left[1 - \frac{R}{M}\operatorname{trace}\left(A(\alpha)\right)\right]^2 \\ &= \left[1 - \frac{1}{m}\operatorname{trace}\left(A(\alpha)\right)\right]^2.\end{aligned}$$

Combining numerator and denominator yields

$$\overline{F}_{\mathrm{GCV}}(\alpha) = \frac{\frac{1}{m}\|\overline{\mathbf{r}}(\alpha)\|_2^2}{\left[1 - \frac{1}{m}\operatorname{trace}\left(A(\alpha)\right)\right]^2} \leq \frac{\frac{1}{mR}\|\widetilde{\mathbf{r}}_{\mathrm{win}}(\alpha)\|_2^2}{\left[1 - \frac{1}{m}\operatorname{trace}\left(A(\alpha)\right)\right]^2} = \widetilde{F}_{\mathrm{GCV}}(\alpha).$$

□

As a closing remark regarding the single parameter multidata GCV method, it is important to emphasize that Assumption 5 was needed for result Prop. 5.7.3. In contrast, no such assumption is necessary for the corresponding multidata UPRE and MDP results, which are Prop. 5.5.1 and 5.6.1 respectively. For the same reason, we are unable to separate the windowed multidata GCV function given in Prop. 5.7.3 even for non-overlapping windows. Such a barrier prevents the development of GCV results analogous to Cor. 5.5.1 and Cor. 5.5.2.

5.8 Summary

In summary, Chapter 5 introduced the backgrounds of both spectral windowing and multidata approaches independently in Section 5.1 and Section 5.2 respectively. The two approaches were combined in Section 5.3, though this section is brief. The

majority of the necessary tools for the development of the windowed multidata methods are presented in Section 5.4. The new methods were individually presented, noting various assumptions that are required for each corresponding function.

Chapter 6

RESULTS FOR WINDOWED MULTI-DATA SETS

To evaluate the effectiveness of the windowed multidata parameter selection methods, 1D and 2D test problems were considered. Selected 2D results were presented in Byrne and Renaut (2023). For both problems, data vectors/images were considered as a set of training data and resulting parameters were then applied to a validation data set. All numerical results use data where a true solution has been blurred and white noise has been added to the blurred image. In both the 1D and 2D cases, we assume that the data is real. In particular, the GSVD is used for the 1D problem. Reflective boundary conditions are assumed in the case of the 2D problem, and therefore the DCT is the primary tool for the 2D results. Details of how the blurring is accomplished are discussed at the beginnings of Sections 6.1 to 6.2. The construction of the noise and how the parameters are evaluated is discussed first.

The noise added to all blurred images are realizations of white noise, i.e., the assumption is that $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_m)$. The signal-to-noise ratio (SNR) is used as a measurement for noise content and is given by

$$\text{SNR} = 10 \log_{10} \left(\frac{\mathcal{P}_{\text{signal}}}{\mathcal{P}_{\text{noise}}} \right). \tag{6.0.1}$$

In the discrete setting, the average power $\mathcal{P}$ of a vector $\mathbf{x}$ of length n is defined as $\frac{1}{m}\|\mathbf{x}\|_2^2$. Using this definition for vectors $\mathbf{b}$ and $\boldsymbol{\eta}$, $\mathcal{P}_{\text{signal}} = \frac{1}{m}\|\mathbf{b}\|_2^2$ and $\mathcal{P}_{\text{noise}} = \frac{1}{m}\|\boldsymbol{\eta}\|_2^2$ and so the quotient in the logarithm is $\|\mathbf{b}\|_2^2/\|\boldsymbol{\eta}\|_2^2$. The quotient can also be expressed as $\|\mathbf{b}\|_2^2/\|\mathbf{d}-\mathbf{b}\|_2^2$, which is the multiplicative inverse of the relative error squared of $\mathbf{d}$. In the 2D setting, the noise can be considered realizations of a random matrix instead of a vector; in this case, eq. (6.0.1) can be used where the average power terms

are written using the Frobenius norm.

In the 1D setting, the noise vector $\boldsymbol{\eta}$ can be constructed by first taking an n-vector $\mathbf{e}$ drawn from the multivariate standard normal distribution and multiplying the vector by a constant σ. Doing so ensures that $\boldsymbol{\eta}$ has variance σ^2 since $\mathrm{Var}(\eta) = \mathrm{Var}(\sigma\,\mathbf{e}) = \sigma^2\ \mathrm{Var}(\mathbf{e})$, and $\mathbf{e}$ has unit variance. Thus it is useful to rearrange the equation defining SNR into an equation that provides a way of finding the necessary variance for a given SNR value. The rearrangement is shown below, with $\|\boldsymbol{\eta}\|_2^2$ replaced by $\mathbb{E}(\|\boldsymbol{\eta}\|_2^2)$:

$$\mathbb{E}(\|\boldsymbol{\eta}\|_2^2) = \frac{\|\mathbf{b}\|_2^2}{10^{(\mathrm{SNR}/10)}}.$$

Using the properties of expected value and the fact that $\mathbb{E}(\|\boldsymbol{\eta}\|_2^2) = \mathbb{E}(\|\sigma\mathbf{e}\|_2^2)$, the term on the left hand side of the equation can be changed as

$$\mathbb{E}(\|\boldsymbol{\eta}\|_2^2) = \sigma^2 \sum_{j=1}^{n} \mathbb{E}(\mathbf{e}_i^2) = \sigma^2 \sum_{j=1}^{n} \Big(\mathbb{E}(\mathbf{e}_i)^2 + \mathrm{Var}(\mathbf{e}_i)\Big) = \sigma^2 \sum_{j=1}^{n} \Big(0^2 + 1\Big) = \sigma^2\, n.$$

Utilizing this change, the following equation for variance is obtained:

$$\sigma^2 = \frac{\|\mathbf{b}\|^2}{n \cdot 10^{(\mathrm{SNR}/10)}}. \tag{6.0.2}$$

This equation is used for the numerical construction of the noise vectors/matrices.

For a basis of comparison for the parameter estimation methods, parameters were also selected as minimizers of the supervised learning function

$$\underset{\text{win}}{\widetilde{F}_{\mathrm{MSE}}}(\boldsymbol{\alpha}) = \frac{1}{R}\,\|\widetilde{\mathbf{x}}_{\mathrm{win}}(\boldsymbol{\alpha}) - \widetilde{\mathbf{x}}\|_2^2 = \frac{1}{R}\sum_{r=1}^{R} \underset{\text{win}}{F_{\mathrm{MSE}}^{(r)}}(\boldsymbol{\alpha}), \tag{6.0.3}$$

where

$$\underset{\text{win}}{F_{\mathrm{MSE}}^{(r)}}(\boldsymbol{\alpha}) = \left\|\mathbf{x}_{\mathrm{win}}^{(r)}(\boldsymbol{\alpha}) - \mathbf{x}^{(r)}\right\|_2^2. \tag{6.0.4}$$

Equation (6.0.3) is the windowed multidata version of the MSE function given in eq. (4.0.1), where a mean of the individual squared error terms, given by eq. (6.0.4), is

being computed. As with eq. (4.0.1), this definition requires that the true solutions, $\{\mathbf{x}^{(r)}\}_{r=1}^{R}$, be known for generating the training data. Regularization parameters chosen as minimizers of eq. (6.0.3) are optimal in the sense of minimizing the mean squared error of the regularized solutions $\mathbf{x}_{\text{win}}^{(r)}(\boldsymbol{\alpha})$; the use of eq. (6.0.3) was considered in Chung and Español (2017). One could also find minimizers $\boldsymbol{\alpha}^{(r)}$ of eq. (6.0.4) for each $r = 1, \ldots, R$, which would produce parameters that are optimal for their own data set; this corresponds to using eq. (4.0.1) in the scalar parameter case. In the results we use MSE to indicate results that are found using the learning function given by eq. (6.0.4).

6.1 One-dimensional Test Problem

Results for the 1D problem serve as a proof of concept for the multidata windowed methods and use the MRI data built into MATLAB®, which can be accessed by the command `load mri.mat`. Five horizontal slices were selected and reformatted as a single image (fig. 6.1) of dimensions 256×536. The default dimension of each horizontal MRI slice accessible using `load mri.mat` is 128×128. Linear interpolation was used to double the number of rows; the number of columns of the concatenated MRI slices was trimmed to to eliminate leading and trailing zero columns. The columns of the image were then multiplied by a symmetric Toeplitz matrix, approximating a Fredholm integral equation of the first kind with the Gaussian kernel eq. (2.1.3). The resulting MRI image is vertically blurred and realizations of normal noise with the corresponding variances were added to blurred columns; see fig. 6.1. For the consideration of multiple data sets, the first $R = 40$ columns of fig. 6.1 served as the training set, while another 40 columns serve as a validation set. Another validation was generated from built-in MATLAB® images, shown in fig. 6.2. The images in fig. 6.2 were horizontally concatenated and the columns blurred in the same way as the training

set; fig. 6.3 shows an example. The first 40 columns of this second validation set were used to construct regularized solutions from the parameters obtained via the training set, and the relative errors were computed against the true solutions.

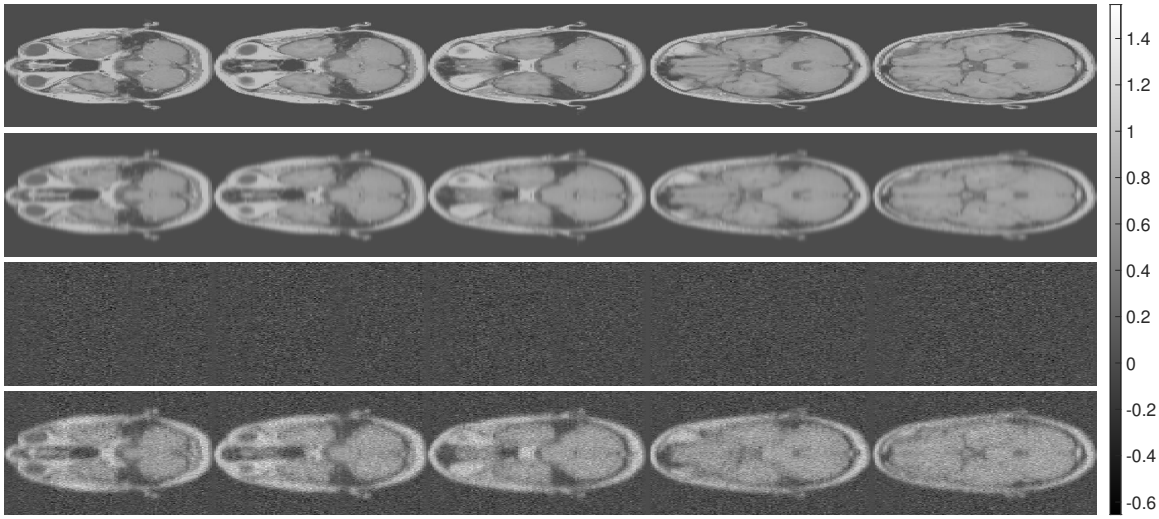


Figure 6.1: MRI data formed by reformatting the MATLAB® built-in MRI data. Each column was blurred using a Gaussian kernel with $\xi = 16$. White noise was added to produce a specific SNR (in this case, an SNR of 10; see eq. (6.0.1)). The resulting vectors are shown at the bottom. The dimension of all four images is 256×536.

For the penalty matrix L, the standard one-dimensional first order finite difference matrix L_1 of size 255×256 was used:

$$L_1 = \begin{bmatrix} -1 & 1 & & 0 \\ & \ddots & \ddots & \\ 0 & & -1 & 1 \end{bmatrix},$$

where the negative ones are positioned along the main diagonal. Left multiplication of a vector by L_1 approximates the computation of a first derivative. The resulting system matrix in eq. (2.2.11) has full column rank, and so applying the normal equations in terms of the GSVD results in unique solutions. The identity matrix I_m was also considered as a penalty matrix, though this corresponds to standard Tikhonov

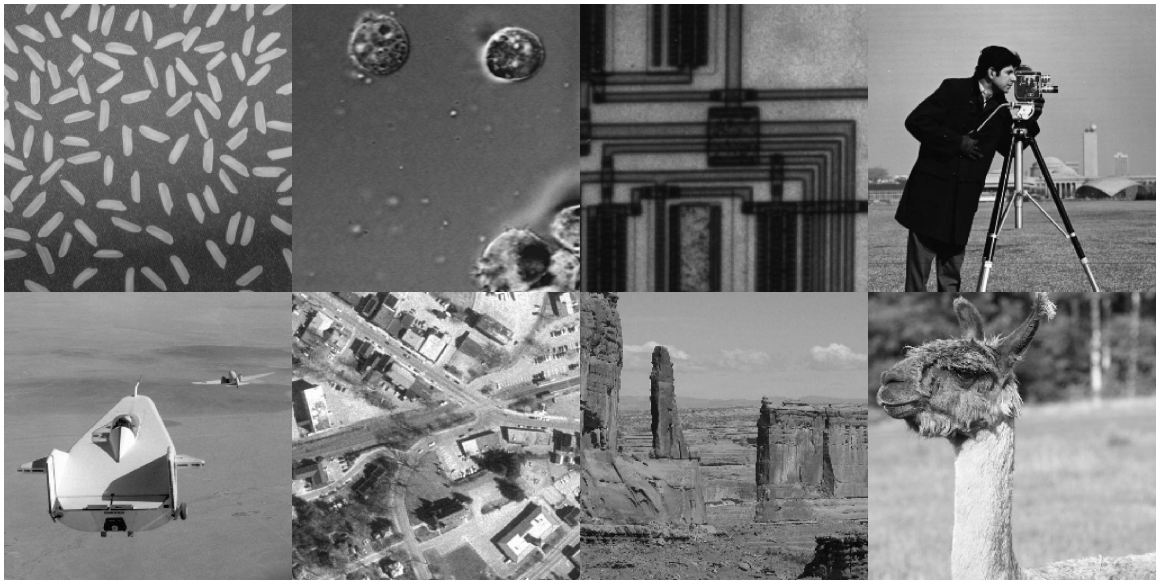


Figure 6.2: The second validation set, consisting of built-in MATLAB® images. From left to right starting in the top row, the images are: `rice.png`, `AT3_1m4_01.tif`, `circuit.tif`, `cameraman.tif`, `liftingbody.png`, `westconcordorthophoto.png`, `parkavenue.jpg`, and `llama.jpg`.

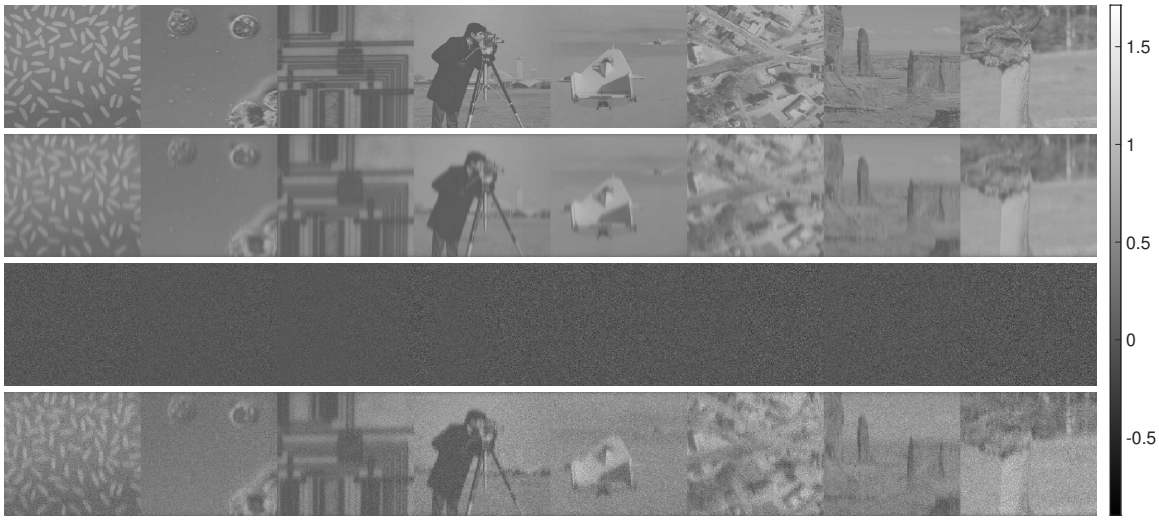


Figure 6.3: 1D validation data formed by reformatting the MATLAB® built-in images from fig. 6.2. Each column was blurred using a Gaussian kernel with $\xi = 16$. White noise was added to produce a specific SNR (in this case, an SNR of 10). The resulting vectors are shown at the bottom. The dimension of all four images is 256 $\times$ 2,048.

regularization. One and two windows were used and the window types considered in the two-window case were logarithmic and logarithmic cosine (non-overlapping and overlapping, respectively). An SNR of 10 and 25 was used in combination with $\xi = 4, 16$. An additional blue value of 36 is considered in the 2D problem discussed in Section 6.2. In regards to the value of ξ, the values $\xi = 4$, 16, and 36 correspond to blurring that is referred to as "mild," "medium," and "severe" in Gazzola *et al.* (2019), respectively. The GSVD was used for all 1D configurations, while the DCT was used for the 2D results in Section 6.2.

To begin, fig. 6.4 shows the spectral weights for $P = 1, 2, 3$ with logarithmic cosine windows (see Section 5.1, specifically eqs. (5.1.5) to (5.1.7)). Regularization parameters were found using eq. (4.0.1) on data with an SNR of 25 and the identity as the penalty matrix for only 10 data sets (in this example, parameters are found for each data set; this is windowed regularization for a single data set instead of full windowed multidata regularization). Figure 6.5 shows how the parameters are distributed as the window number changes. The relative errors of the corresponding regularized solutions are shown in fig. 6.6. The choice to focus on using eq. (4.0.1) with only 10 data sets initially was motivated to first determine whether windowed regularization can produce solutions that outperform single parameter regularization.

Figure 6.6 demonstrates that for this example, there is benefit to using two and three spectral windows. While there is some benefit in using three windows as compared to one window, $P = 2$ provides the best relative errors; using three windows does not improve upon the solutions generated from two windows. Furthermore, for every increase in window size there is additional cost of finding the regularization parameters because the number of variables contributing to the minimization problem increases. For these reasons, $P = 2$ is chosen as the maximum number of spectral windows in this document, in both the 1D and 2D problems. A plot similar to that

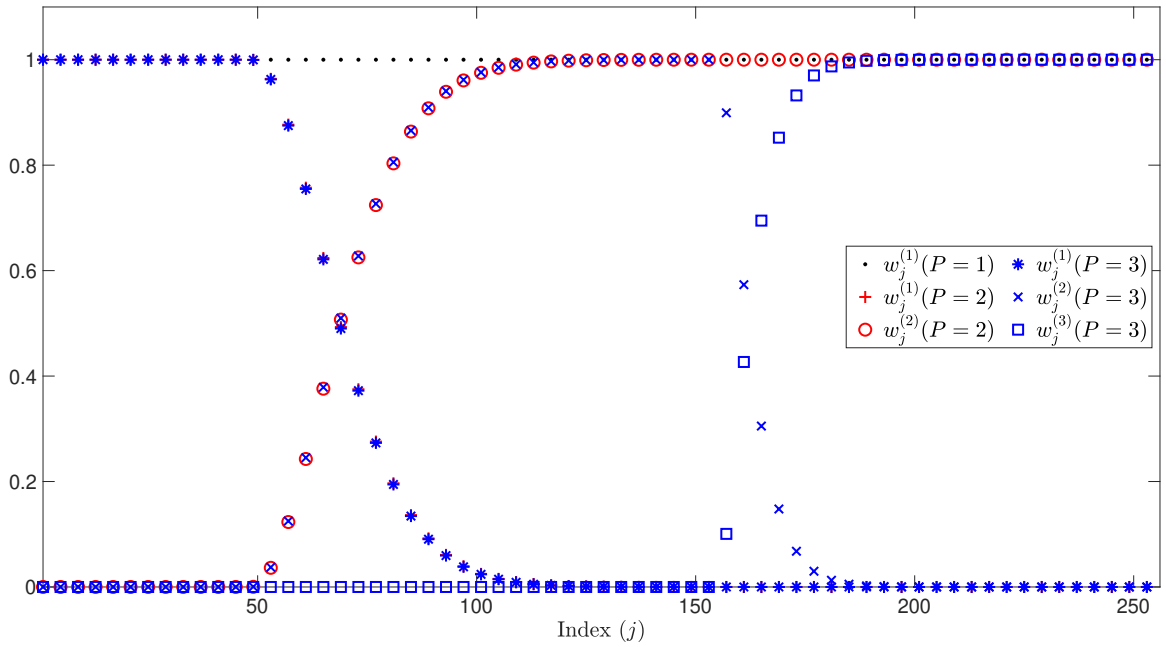


Figure 6.4: Spectral weights determined using logarithmic cosine windows a symmetric Toeplitz system matrix with $\xi = 16$. The overlap of the windows for $P > 1$ is clear.

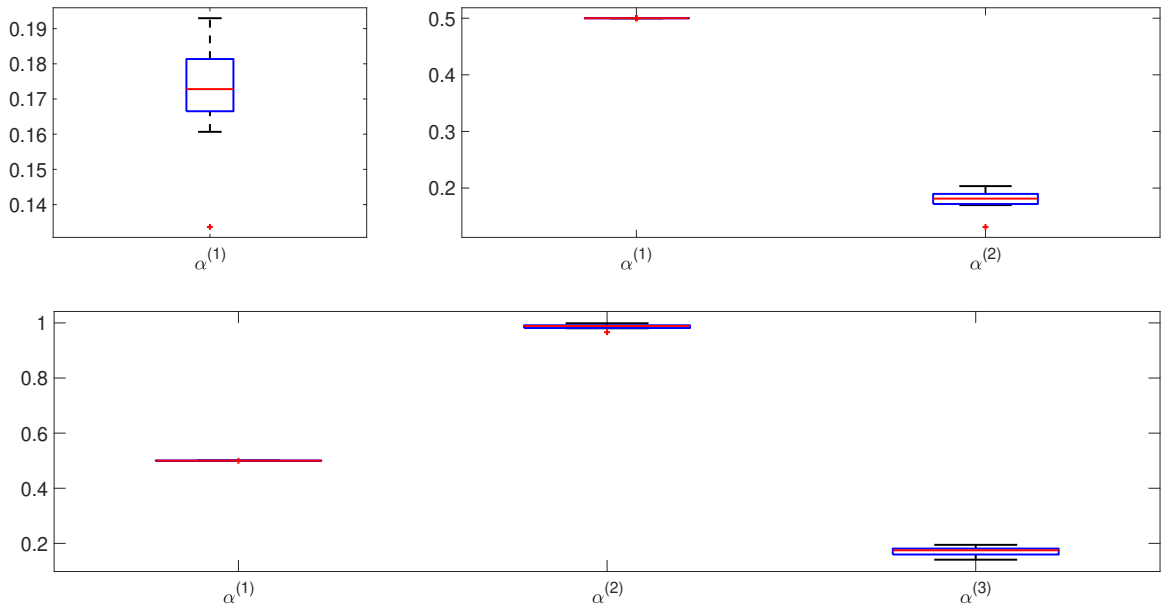


Figure 6.5: MSE parameters obtained using a symmetric Toeplitz system matrix with $\xi = 4$, data having an SNR of 25, and the identity penalty matrix. See fig. 6.6 for corresponding relative errors.

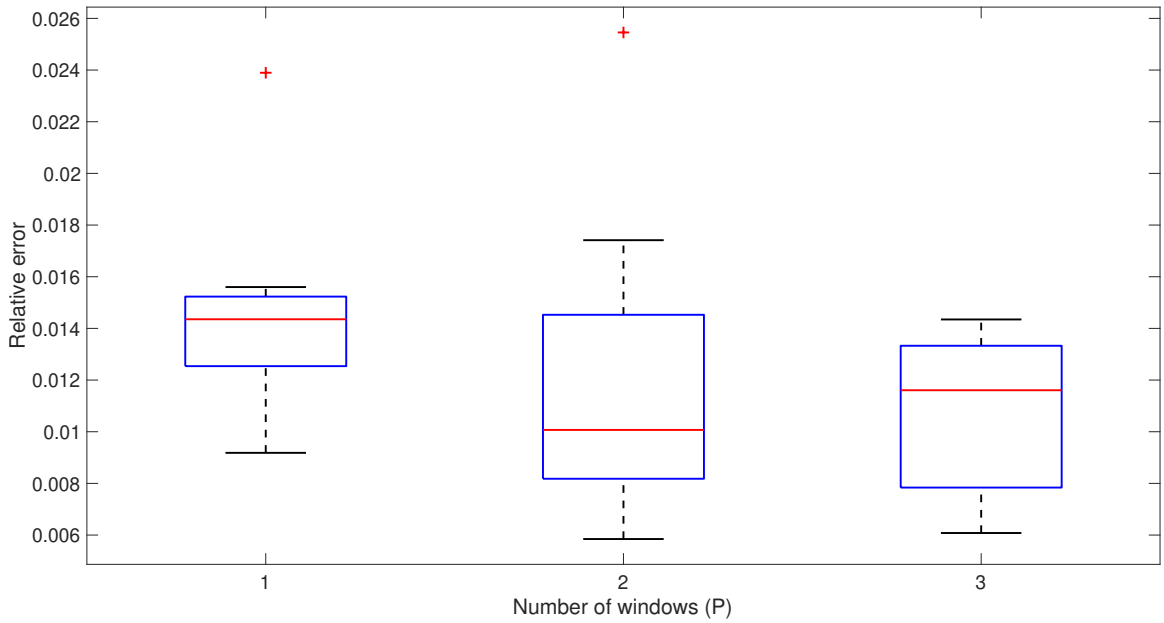


Figure 6.6: Relative errors obtained using MSE parameters for a symmetric Toeplitz system matrix with $\xi = 4$, data having an SNR of 25, and the identity penalty matrix.

of fig. 6.5 is shown for the 2D problem in Section 6.2.

Moving to all of the parameter selection methods considered in this work, fig. 6.7 illustrates a comparison of the methods in terms of the regularization parameters α selected as the number of training vectors increases. When the number of training vectors is small (e.g. between 1 and 10 from fig. 6.7), the parameters determined by all four methods change significantly. This can be attributed to the fact that all of the methods find a parameter that is either a root or a minimum of an average of functions. For a small number of training vectors, each additional vector has more influence over the behavior of the multidata function. The parameters stabilize as the number of training vectors reaches a certain point; in the problem being considered here, the parameters stabilize by about 10 training vectors.

Before looking at the relative errors of the regularized solutions, another observation regarding fig. 6.7 can be made. While the parameters determined using the learning function eq. (6.0.3) and multidata MDP method are close (even for safety

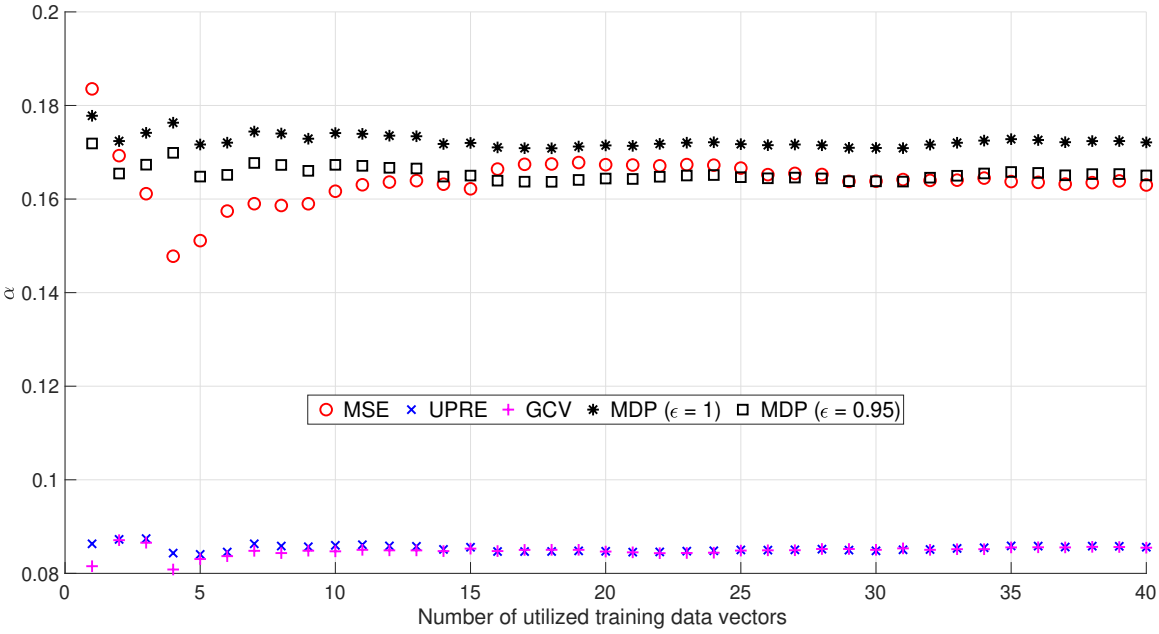


Figure 6.7: Trend of parameters selected by each multidata method as the number of training vectors (R) increases. The parameters stabilize as the number of training vectors increases. A safety parameter of $\epsilon = 0.95$ for the MDP method was chosen manually to produce results similar to the other methods.

parameter $\epsilon = 1$), the multidata UPRE and GCV methods produce parameters that are similar as well. Modification of ϵ can produce even better parameters, though this only accomplished for fig. 6.7. For the remainder of the document, the MDP methods are assumed to use $\epsilon = 1$.

Figure 6.8 shows the relative errors of the regularized solutions corresponding to the parameters from each method. The relative errors obtained by the learning and multidata UPRE and GCV methods are quite similar; the means of the relative errors are close, and there is a collection of upper outliers. In contrast, the multidata MDP method better matches the errors produced by the learning method in this case. To summarize the 1D results, table 6.1 shows a collection of relative errors for multiple values R, window types/numbers, and each windowed multidata method.

One observation regarding table 6.1 is that the MSE method consistently outperforms the other methods, but for this configuration of SNR and blur amount, the

Table 6.1: Averaged percent relative errors of the multidata windowed regularized solutions for $\xi = 36$ and an SNR of 10 with one window, two logarithmically spaced windows and two logarithmically spaced cosine windows with the L_1 penalty matrix. The result with least error for given R, method, and validation set is highlighted in bold face.

R	Win	Training				Validation 1				Validation 2			
		MSE	UPRE	GCV	MDP	MSE	UPRE	GCV	MDP	MSE	UPRE	GCV	MDP
1	None	**4.41**	4.44	**4.41**	4.62	**4.18**	4.23	4.19	4.27	3.11	3.24	3.14	**2.65**
	Log	**4.41**	4.44	4.45	4.62	**4.18**	4.23	4.24	4.27	3.11	3.24	3.28	**2.65**
	LogCos	**4.41**	4.44	4.46	4.62	**4.18**	4.23	4.24	4.27	3.11	3.23	3.28	**2.65**
10	None	4.62	**4.45**	**4.45**	6.32	4.27	4.17	**4.16**	5.47	**2.65**	2.77	2.78	2.75
	Log	**4.62**	6.99	8.48	6.32	**4.27**	7.03	8.71	5.47	**2.65**	6.04	7.93	2.75
	LogCos	4.62	**4.45**	7.24	6.32	4.27	**4.17**	7.03	5.47	**2.65**	2.77	7.07	2.75
20	None	4.62	**4.45**	**4.45**	6.32	4.27	4.17	**4.16**	5.47	**2.65**	2.77	2.78	2.75
	Log	**4.41**	4.48	4.45	5.37	**4.19**	4.27	4.24	4.8	3.13	3.33	3.27	**2.60**
	LogCos	**4.41**	4.48	10.03	5.37	**4.19**	4.27	9.76	4.8	3.13	3.33	10.26	**2.6**
30	None	4.39	**4.41**	**4.41**	5.47	**4.16**	4.19	4.19	4.87	3.03	3.13	3.14	**2.61**
	Log	**4.41**	**4.41**	**4.41**	5.37	**4.19**	**4.19**	**4.19**	4.8	3.13	3.13	3.12	**2.6**
	LogCos	**4.39**	4.41	8.4	5.47	**4.16**	4.19	8.17	4.87	3.03	3.13	8.41	**2.61**
40	None	**4.39**	4.4	4.4	5.81	**4.15**	4.16	4.17	5.11	2.97	3.04	3.08	**2.65**
	Log	**4.41**	4.4	4.4	5.37	4.19	**4.16**	**4.16**	4.8	3.13	3.04	3.03	**2.60**
	LogCos	**4.39**	4.4	8.16	5.81	**4.15**	4.16	7.94	5.11	2.97	3.04	8.15	**2.65**

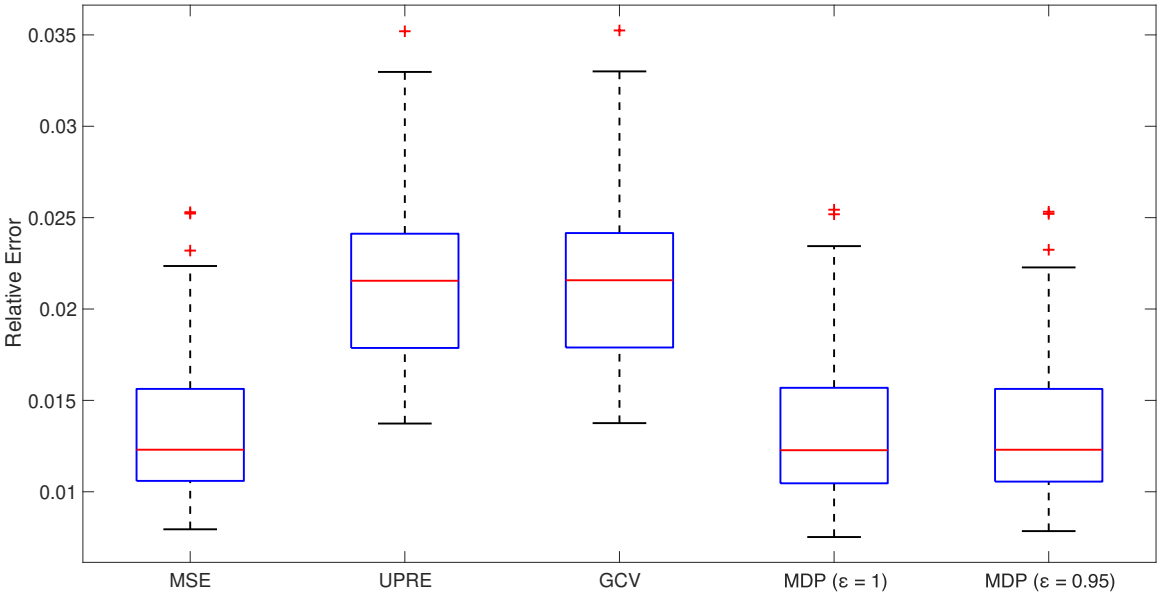


Figure 6.8: Boxplots of the relative errors of the regularized solutions constructed from parameters selected by each method. The effectiveness of the multidata MDP method is affected by the choice of safety parameter ϵ; here $\epsilon = 0.95$ yields comparable results to the MSE method.

greater performance is marginal. Another observation is that for the second validation set, the windowed multidata MDP method produced parameters that resulted in the lowest relative errors. In general, the relative errors associated with the second validation set differ from those of the training set and first validation set. This phenomenon is investigated further in Section 6.2.

The 1D test problem demonstrates that the multidata methods have potential for selecting viable regularization parameters that can be applied for multiple sets of data. In the 1D experiments, the windowed multidata UPRE and GCV methods performed similarly. The windowed multidata MSE method performs best on average, which is to be expected since this method relies on knowledge of the true solutions. The windowed multidata MDP method can also perform competitively, though this method often requires some fine tuning of the safety parameter ϵ.

The data sets of the 2D test problem consist of images of size 256×256. This problem utilizes the images in fig. 6.9 of the planet Mercury obtained by the MESSENGER space probe [1] . To obtain the 16 images from the 512×512 Mercury images in fig. 6.9, 8 images were randomly chosen and two 256×256 subimages of each image were selected as the northwest and southeast corners. A total of 16 images were used and split into training and validation sets containing 8 images each. The second validation set is shown in fig. 6.2.

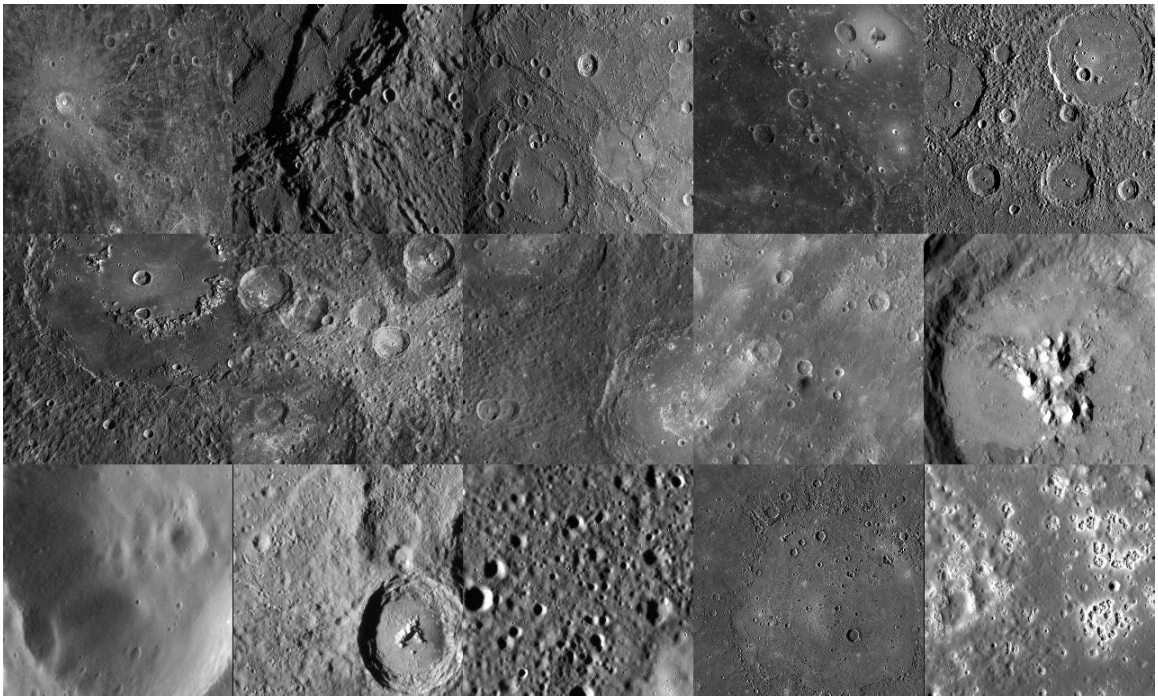


Figure 6.9: Selected Images for the MESSENGER 2D Test Problem. Available courtesy of NASA/JPL-Caltech NASA and JPL-Caltech (2016).

A 256×256 point spread function was formed using a discretization of the zero centered, circularly symmetric 2D Gaussian kernel,

$$k(x, y) = \exp\left(-\frac{x^2 + y^2}{2\xi}\right). \tag{6.2.5}$$

[1]The MESSENGER images are available to the public courtesy of NASA/JPL-Caltech NASA and JPL-Caltech (2016)

The scale factor which makes $\int_{\mathbb{R}^2} k(x,y)\,dx\,dy = 1$ has been omitted since the appropriate scaling is accomplished numerically. As with the 1D Gaussian kernel eq. (2.1.3), the parameter ξ controls the width of the Gaussian kernel. Choosing $k(x,y)$ to be circularly symmetric is for convenience; a Gaussian kernel with different width parameters for the x and y directions can still be used to construct a PSF that is doubly symmetric for diagonalization via the DCT Hansen *et al.* (2006). The corresponding PSFs were discretely convolved with each image as a means of blurring. SNR values of 10 and 25 were used to construct realizations of normal noise that were added to the blurred images to create the data. For one choice of the penalty matrix L, we used the appropriately structured version of the discrete negative Laplacian operator, which is an approximation of the continuous Laplacian operator $-\nabla^2$ Debnath and Mikusiński (2005); LeVeque (2007). Given an open subset $\Omega \subseteq \mathbb{R}^2$ and an integer $k \geq 2$, we can denote $\mathcal{C}^k$ as the space of all real-valued functions on Ω that have continuous partial derivatives of order k Debnath and Mikusiński (2005). The negative Laplacian operator $-\nabla^2 : \mathcal{C}^k(\Omega) \to \mathcal{C}^{k-2}(\Omega)$ is then defined by

$$-\nabla^2 f = -\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right), \quad f \in \mathcal{C}^k(\Omega). \tag{6.2.6}$$

However, the derivatives can be approximated by finite differences so that the following stencil can be used to represent the effect of $-\nabla^2$ in a discrete convolution:

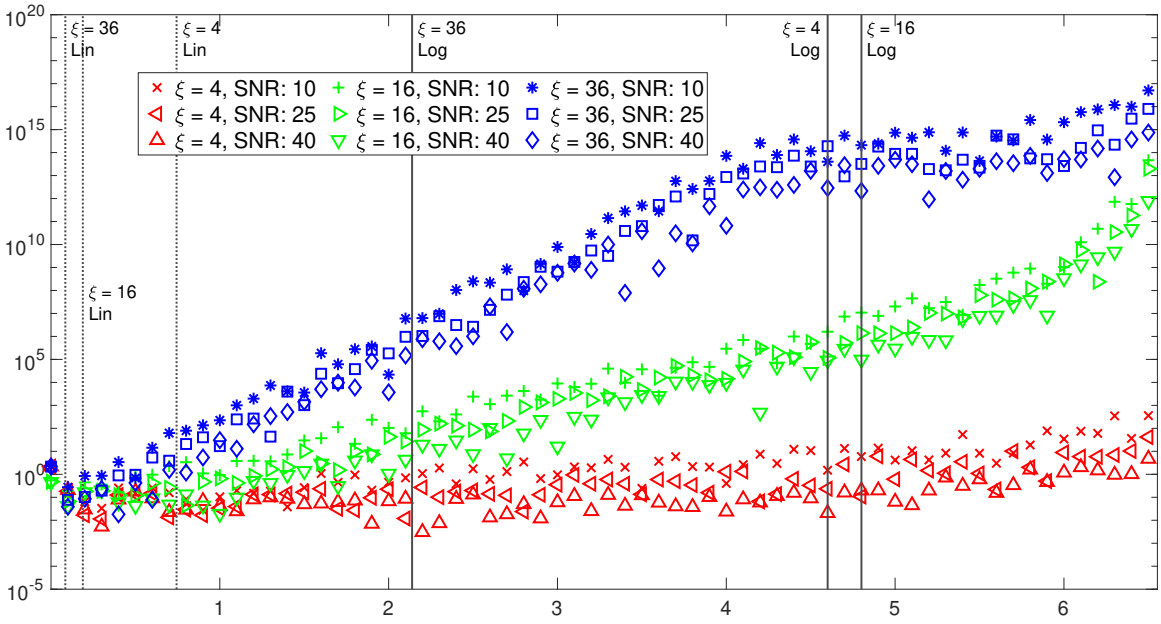
$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{6.2.7}$$

Using eq. (6.2.7), a discrete approximation L_2 of eq. (6.2.6) can be constructed using the process outlined in Vogel (2002). For the second penalty matrix we used $L = I_m$.

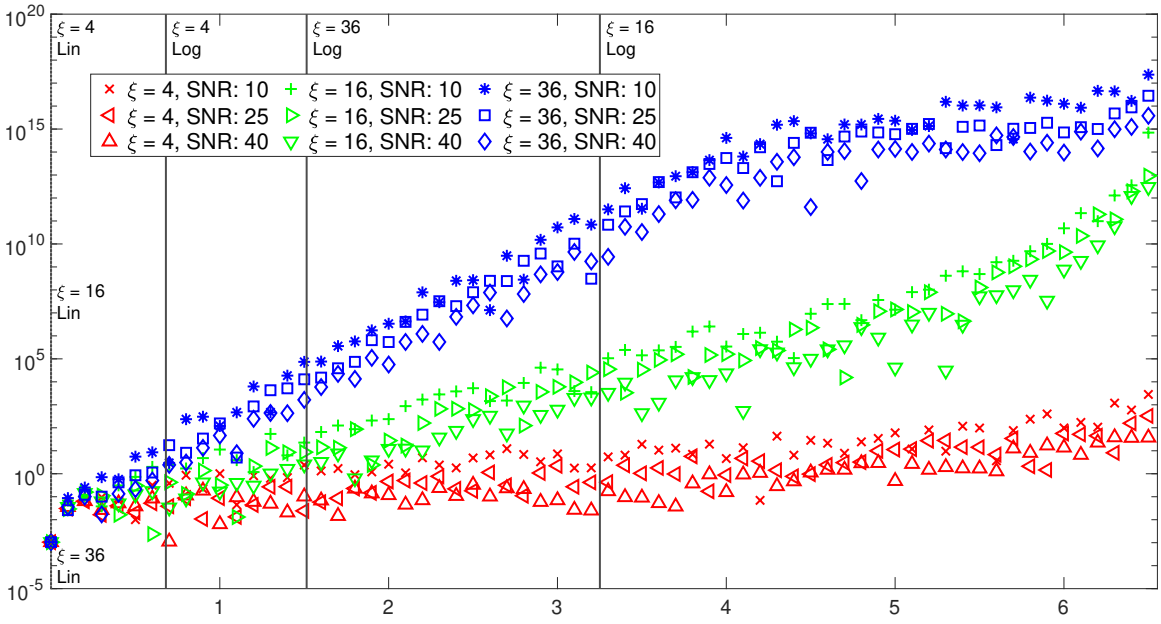
The structure of resulting matrices A and L allows for simultaneous diagonalization using the DFT/DCT for numerical efficiency (see Section 3.4).

The learning methods were evaluated for both the scalar and spectral windowing cases using training data sets of sizes $R = 1$ to 8. The learned parameters in each case were then used to construct regularized solutions for data from two independent validation sets. For the windowed regularization we considered both non-overlapping linear/logarithmic windows and overlapping linear/logarithmic cosine windows. The decision to use linear spacing for $L = I_m$ and logarithmic spacing for $L = L_2$ is supported by how the ordered spectral components decay; see fig. 6.10. As in Chung *et al.* (2011b), the windowed GCV function in Prop. 5.7.1 was replaced by the P independent GCV functions eq. (5.7.39) for simplicity when considering non-overlapping windows. Parameters were also obtained for the separable UPRE method given by eq. (5.5.27). For the spectral windowing with overlapping windows, the minimizations were initialized using the parameters obtained by the non-overlapping methods. Overall, in terms of the choice to initialize the parameters for the overlapping windows with parameters obtained from the separable case, we note that the windowed UPRE and GCV methods corresponding to overlapping windows performed better when the minimizations were initialized using the parameter obtained by the non-overlapping methods. Results without this initialization are not considered.

Considering first the scalar parameter multidata case, the resulting parameters appear to stabilize as the number of data sets is increased. Figure 6.11 demonstrates this effect and shows that the amount of stabilization appears to be connected to the homogeneity of the training set. Sets constructed from fig. 6.9 are homogeneous in the sense that they all contain images of the surface of Mercury. In contrast, fig. 6.2 consisted of entirely distinct images. By changing which sets are used for training or validation influences the resulting parameters as R increases. The corresponding



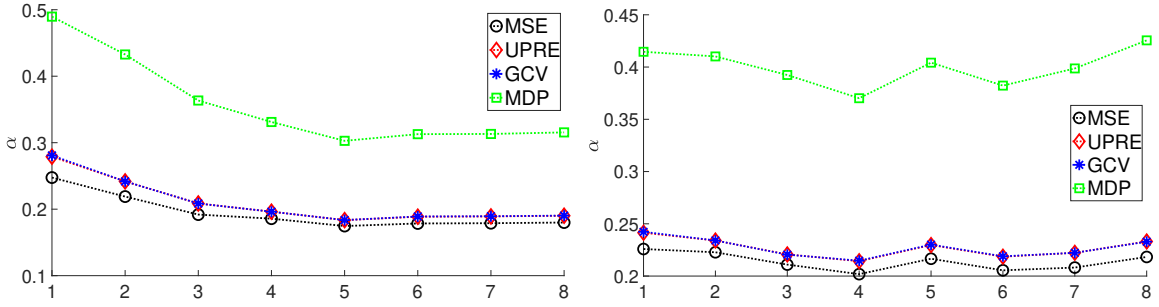
(a) $L = I$



(b) $L = L_2$

Figure 6.10: Plot of the ratios $\left|\widehat{d}_j\right| / \gamma_j$, $j = 1 : 200 : n$, for a fixed data set selected from the first image in fig. 6.9 using $L = I$ and $L = L_2$, in figs. 6.10a and 6.10b, respectively. Here we assume j increasing corresponds to γ_j in descending order. The vertical lines represent the partition points of two (linear or logarithmic) windows, which depend only on the blur amount ξ.

relative errors of the regularized solutions are shown in fig. 6.12. While the box plots in fig. 6.12a and fig. 6.12b simply look as if they are the same but reordered, the box plots appear similar because the resulting parameters are approximately the same ($\alpha \approx 0.2$). These experiments suggest that it is sufficient to use only a small number of images, relative to the total available, to obtain meaningful results. The use of fig. 6.2 as a training set was only considered to produce fig. 6.11 and fig. 6.12; its use as a validation set is retained through the remainder of the results.

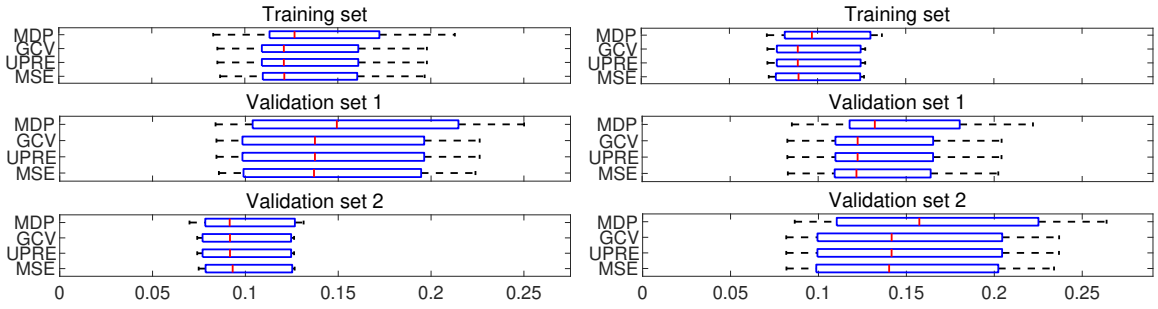


(a) Parameters α against R (Set 1)

(b) Parameters α against R (Set 3)

Figure 6.11: Figure 6.11a illustrates the change in scalar α as the number of data sets increases, here with Set 1 as the training set and in fig. 6.11b with Set 3 (see fig. 6.12). Figure 6.11a is an example of how scalar regularization parameters can stabilize as the number of data sets in the multidata methods increases. In contrast, fig. 6.11b shows less stabilization with increasing R when the training set is changed. For both plots, $\xi = 16$, $L = L_2$, and an SNR of 25 was used.

In regards to the spectral windowing, typically two windows were sufficient (corresponding to the use of just two parameters in the windowed estimators) to obtain meaningful solutions. The observed benefit of using greater than two windows was minor, an example of which is shown for the windowed UPRE method in fig. 6.13a. Another advantage of using two windows is that there is a greater computational cost of finding more parameters than is necessary for meaningful regularized solutions; this is especially true for overlapping windows where decoupling is not an option. Extending the number of windows also has the effect of reducing the influence of one or more parameters. For example, fig. 6.13b shows that one of the three parameters

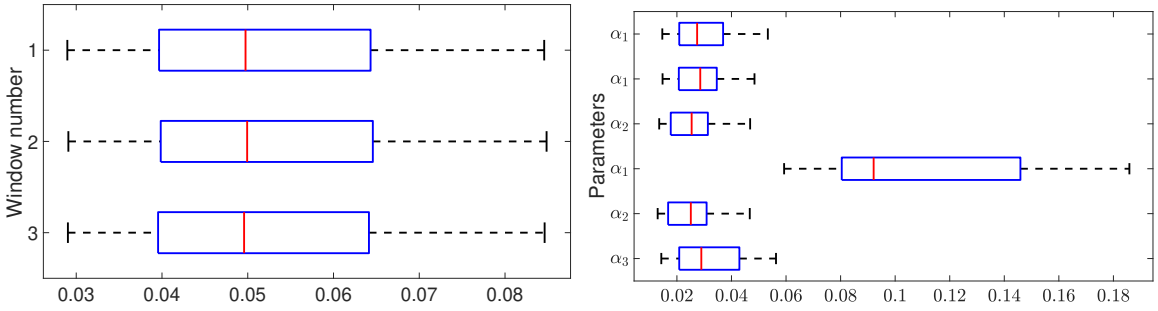


(a) Relative errors (Set 1)

(b) Relative errors (Set 3)

Figure 6.12: Relative errors of regularized solutions obtained for scalar α from each multidata method with $R = 5$ data sets. Here Set 1 and Set 2 were constructed from fig. 6.9, while Set 3 was constructed from fig. 6.2. In fig. 6.12a, Set 1 served as the training set and the resulting parameters were used to construct solutions for the data from Sets 2 and 3. Figure 6.12b shows results where training was done using Set 3 instead and Sets 1 and 2 served as validation sets. For both plots, $\xi = 16$, $L = L_2$, and an SNR of 25 was used.

obtained from the windowed `UPRE` method with three windows is more variable and larger in magnitude than that other two parameters.



(a) Relative errors

(b) Parameters for $P = 1:3$

Figure 6.13: Parameters and corresponding relative errors from the UPRE method as the number of windows is increased from one to three. Figure 6.13a shows that there is little benefit in using an increasing number of windows. Figure 6.13b shows that past two windows, the new regularization parameters are more variable. For both plots, logarithmically spaced windows were used with $\xi = 4$, $L = L_2$, and an SNR of 40.

The results presented in Chung *et al.* (2011b) also suggested that there is little to be gained when using more than two windows, even when using the learning approach, method MSE, to find the parameters. On the other hand, the presented framework is

valid for more windows, should there be situations in which the use of two windows seems insufficient based on numerical experiments.

It should be noted also, that when using the windowed multidata MDP method, there is an additional tuning safety parameter, which is required and makes the presentation of results for the MDP method much less interesting. In figs. 6.11 to 6.12 the safety parameter is 1, though the MDP results can be altered to better match those of the MSE error method by adjusting the safety parameter as needed.

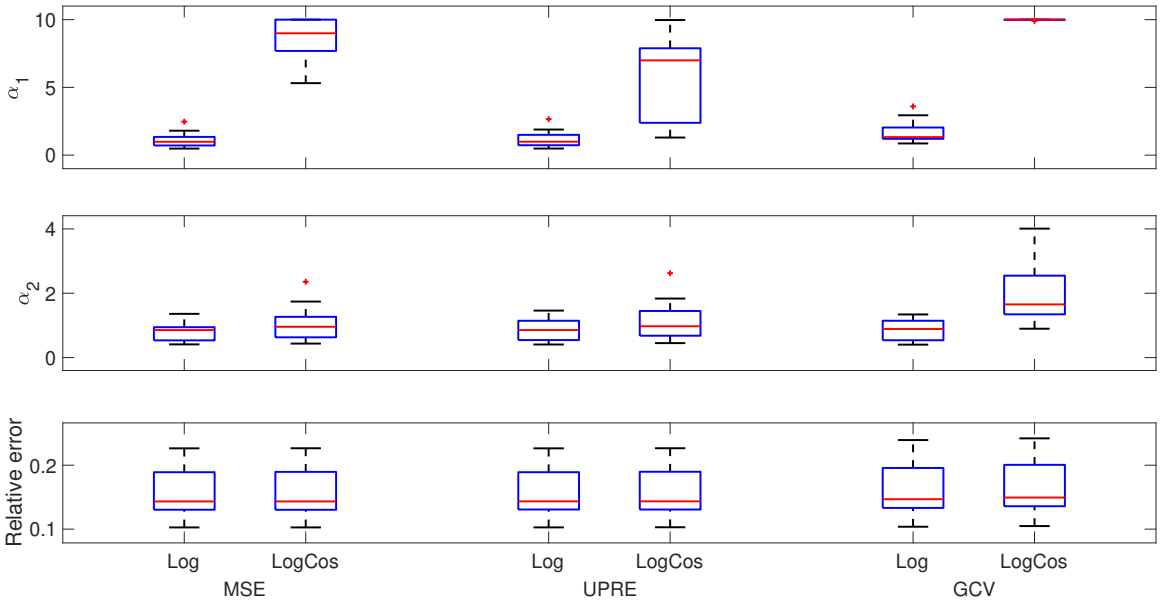


Figure 6.14: Parameters and corresponding relative errors obtained from using logarithmic vs logarithmic-cosine windows with the MSE, UPRE, and GCV methods. In the case of the logarithmic (non-overlapping) windows, the independent versions of the UPRE and GCV functions were used, eqs. (5.5.27) and (5.7.39), respectively. For both window versions, $\xi = 4$, $L = L_2$, and an SNR of 10 was used.

The use of overlapping or non-overlapping windows influences the degree of interdependence between the two parameters. Figure 6.14 presents the results of using overlapping and non-overlapping logarithmic windows with $L = L_2$. When using non-overlapping windows, the ranges of both parameters are smaller than those for non-overlapping windows. For overlapping windows, the behavior of α_1 exhibited in fig. 6.14 shows the parameters grouping near 10. The grouping behavior can be

explained by the choice of an upper bound during the minimization process; in the case of fig. 6.14, the upper bound was chosen near 10. The calculated gradients of the $F_{\substack{\texttt{MSE}\\\texttt{win}}}(\boldsymbol{\alpha})$, $F_{\substack{\texttt{UPRE}\\\texttt{win}}}(\boldsymbol{\alpha})$, and $F_{\substack{\texttt{GCV}\\\texttt{win}}}(\boldsymbol{\alpha})$ are too small to resolve a minimum in the direction of α_2 and thus the minimization process determines the minimizers near the specified boundary. However, using overlapping windows also increased the magnitude of α_2, most significantly in the case of the GCV method.

In regards to the multidata windowed methods, which select P parameters using R data sets, the parameters converge as R increases. Table 6.2 details the mean percent relative errors of solutions obtain using parameters from each multidata windowed method, where one and two (both overlapping and non-overlapping) windows were used. Even for the limited number of training sets (2 through 8), the errors decrease as R increases. For most numerical configurations tested, the use of overlapping vs non-overlapping windows provides minor benefit with regards to the relative errors of the regularized solutions.

It is interesting to note that the average relative errors of solutions obtained for parameters applied to the second validation set (fig. 6.2) were less than those of either the training or first validation set. The superior (reduced) errors calculated for the second validation set are consistent throughout most numerical configurations. Additionally, the relative errors for the second validation set show greater variability than those for either the training, or first validation, set. Furthermore, the relative errors are indeed least in each case when training is performed using known data, namely with the MSE, but the results with both UPRE and GCV learning methods are not significantly larger when using windowed regularization. This demonstrates that windowed regularization parameters can be learned from training data without knowledge of the true solutions. The results obtained using UPRE are in most cases slightly improved as compared to those using GCV, and hence UPRE would be pre-

Table 6.2: Averaged percent relative errors of the multidata windowed regularized solutions for $\xi = 36$ and an SNR of 10 with one window, two linearly spaced windows and two linearly spaced cosine windows with the identity penalty matrix. The result with least error for given R, method, and validation set is highlighted in bold face.

R	Win	Training			Validation 1			Validation 2		
		MSE	UPRE	GCV	MSE	UPRE	GCV	MSE	UPRE	GCV
2	None	**21.32**	25.83	25.87	**23.76**	27.58	27.62	**17.57**	23.27	23.32
	Lin	19.64	**19.54**	**19.54**	22.85	**22.67**	**22.67**	14.94	**14.88**	14.89
	LinCos	**19.54**	19.93	19.94	**22.83**	23.38	23.39	**14.78**	15.10	15.11
4	None	**21.29**	26.09	26.12	**23.71**	27.82	27.86	**17.58**	23.57	23.62
	Lin	**19.44**	19.53	19.55	22.39	**22.36**	22.37	**14.93**	15.16	15.17
	LinCos	**19.32**	19.33	19.94	22.32	**22.29**	23.39	**14.75**	14.79	15.11
6	None	**21.29**	26.03	26.07	**23.71**	27.77	27.81	**17.58**	23.51	23.56
	Lin	**19.44**	19.49	19.50	22.41	**22.36**	22.37	**14.91**	15.05	15.07
	LinCos	**19.32**	**19.32**	19.94	**22.34**	**22.34**	23.39	14.73	**14.72**	15.11
8	None	**21.29**	26.02	26.06	**23.71**	27.76	27.80	**17.57**	23.50	23.55
	Lin	**19.44**	19.48	19.49	22.41	**22.36**	22.37	**14.90**	15.04	15.06
	LinCos	**19.32**	**19.32**	19.94	**22.34**	**22.34**	23.39	**14.72**	**14.72**	15.11

ferred if information about the noise in the data is available. Finally, to illustrate the performance of the approach, fig. 6.15 presents two examples of images from the second validation set that have differing relative errors. Examples of results of the GCV and MDP methods are shown in fig. 6.16 and fig. 6.17, respectively.

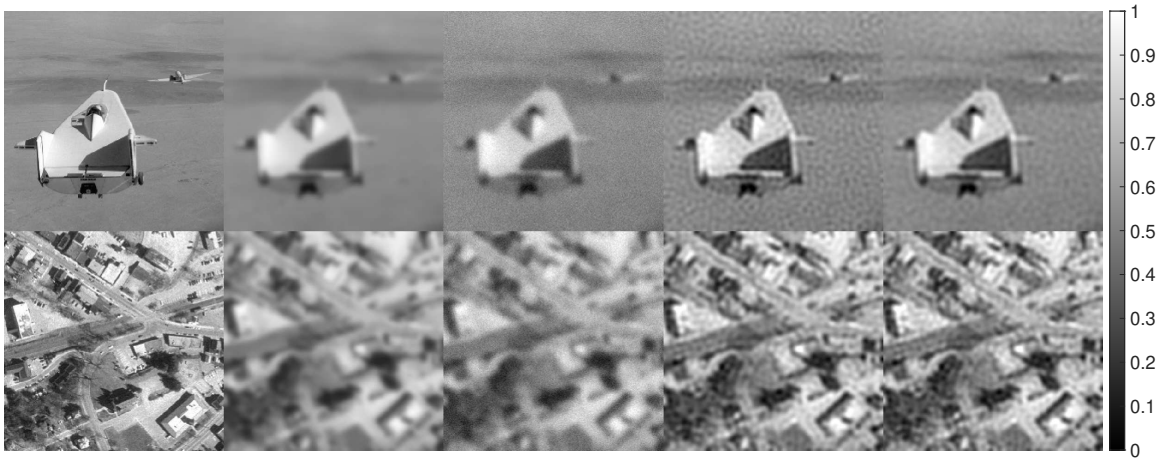


Figure 6.15: Two samples from the second validation set, with $\xi = 36$, an SNR of 25, two log cosine windows and the Laplacian penalty matrix. From left to right for each sample are the true solution, the blurred image, the blurred image after noise was added, the regularized solution obtained using the multidata windowed UPRE method with $R = 8$ (the entire training set), and the regularized solutions using parameters that are optimal for the individual image. The multidata windowed UPRE solutions have relative errors of 8.55% and 14.81% for the top and bottom samples, respectively, while the optimal solutions have relative errors of 8.04% and 14.80%.

6.3 Summary

In both the 1D and 2D cases, the windowed multidata methods can be competitive with the method of finding parameters using the true solutions. Typically, the benefit of using multiple windows is maximized after about two or three windows; beyond that, the benefit is outweighed by the computational cost of minimization/root-finding problems. The use of overlapping or non-overlapping windows influences the degree of interdependence between the two parameters. In addition, parameters selected from different data sets often produce differing amounts of error in their solutions;

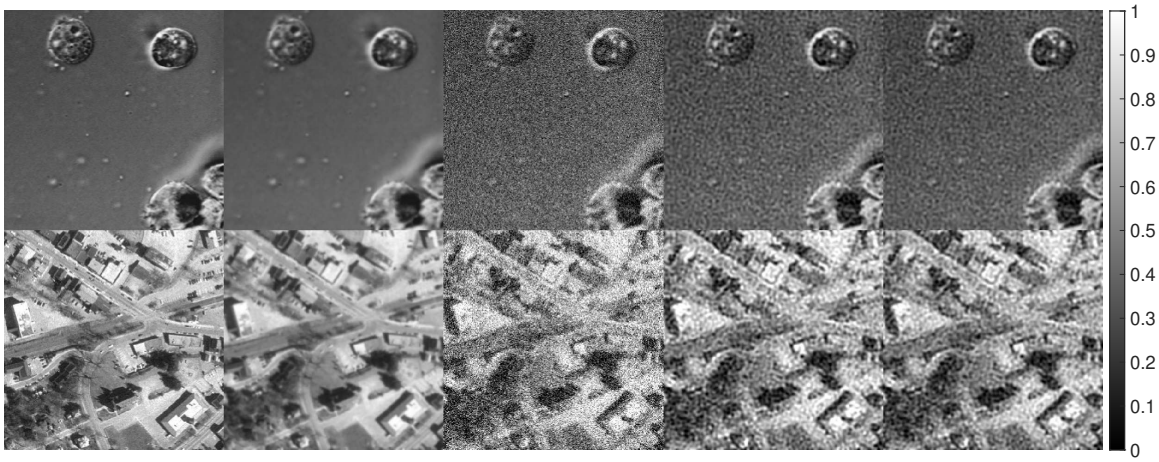


Figure 6.16: Two samples from the second validation set, with $\xi = 4$, an SNR of 10, two linear cosine windows and the identity penalty matrix. From left to right for each sample are the true solution, the blurred image, the blurred image after noise was added, the regularized solution obtained using the multidata windowed GCV method with $R = 8$ (the entire training set), and the regularized solutions using parameters that are optimal for the individual image. The multidata windowed GCV solutions have relative errors of 14.47% and 16.87% for the top and bottom samples, respectively.

Figure 6.17: Two samples from the second validation set, with $\xi = 16$, an SNR of 40, two log windows and the Laplacian penalty matrix. From left to right for each sample are the true solution, the blurred image, the blurred image after noise was added, the regularized solution obtained using the multidata windowed MDP method ($\epsilon = 1$) with $R = 8$ (the entire training set), and the regularized solutions using parameters that are optimal for the individual image. The multidata windowed MDP solutions have relative errors of 10.43% and 7.09% for the top and bottom samples, respectively.

this is seen by comparing the results of using the second validation set fig. 6.2 versus the results of using the Mercury images fig. 6.9.

A limitation of the numerical results is that there are better options than the 2-norm regularization term for the process of image deblurring. For example, total variation regularization is often used to produce deblurred images, though the implementation of this method is iterative in nature. One direction of future work is then to adapt the windowed multidata methods presented here for use with iterative methods. Another direction of future work could be to fix a parameter in the windowed framework and determine the other parameters afterwards. Though this methodology would require a useful estimate of the fixed parameter, it could allow for fine tuning of remaining parameters to obtain more accurate solutions.

Chapter 7

DOWNSAMPLING FUNCTIONS

While analysis has been conducted regarding the convergence of predictive and estimation error for Tikhonov regularization as the number of sample points becomes large (Vogel, 2002, p. 109-126), efforts to analyze the the effects of reducing the number of sample points can be expanded. A basic methodology of downsampling for the selection of regularization parameters has been presented in Hansen (2015); Renaut *et al.* (2017). This chapter serves as a bridge between methods discussed in this work and directions for future investigation. The effectiveness of regularization parameter functions discussed in Chapter 4 and methods introduced in Chapter 5 applied to downsampled data will be discussed briefly. An introduction to the concept of downsampling, as well as its effects on noise, are given in Section 7.1. Results from Chapter 3 are used to suggest specific approaches to expanded work using downsampling. Preliminary numerical results are given in Section 7.2.

7.1 Downsampling Background

To formalize the concept of downsampling, consider $\mathbf{z} = [z_1, z_2, \ldots, z_n]$. A vector $\mathbf{y}$ is called a downsampling of $\mathbf{z}$ if $\mathbf{y} = [z_{\mathcal{I}_1}, z_{\mathcal{I}_2}, \ldots, z_{\mathcal{I}_m}]$, where $m \leq n$ and $\mathcal{I}_j : \{1, \ldots, m\} \to \{1, \ldots, n\}$ is a strictly increasing function. This definition is analogous to the definition of a subsequence except with a finite number of terms.

Given an n-vector $\mathbf{x}$, an $m \times n$ matrix E can be constructed such that $E\mathbf{x}$ is a downsampling of $\mathbf{x}$. The structure of E is such that

$$E = [\mathbf{e}_{\mathcal{I}_1}, \mathbf{e}_{\mathcal{I}_2}, \ldots, \mathbf{e}_{\mathcal{I}_m}]^\mathsf{T} \tag{7.1.1}$$

with columns $\mathbf{e}_{\mathcal{I}_j} \in \mathbb{R}^n$ and $\mathcal{I}_j : \{1, \ldots, m\} \to \{1, \ldots, n\}$ being strictly increasing. In

other words, E is an identity matrix with rows removed. An immediate consequence of eq. (7.1.1) is that E is a semi-orthogonal matrix, which is described by Lem. 7.1.1.

Lemma 7.1.1 *Let E be an $m \times n$ matrix defined by eq. (7.1.1). Then $EE^\mathsf{T} = I_m$ but $E^\mathsf{T}E \neq I_n$.*

Proof. Using eq. (7.1.1), for $1 \leq j, k \leq m$ we have

$$\left[EE^\mathsf{T}\right]_{j,k} = \mathbf{e}_{\mathcal{I}_j}^\mathsf{T}\mathbf{e}_{\mathcal{I}_k} = \begin{cases} 1, & \mathcal{I}_j = \mathcal{I}_k, \\ 0, & \mathcal{I}_j \neq \mathcal{I}_k. \end{cases}$$

Since $\mathcal{I}_j$ is strictly increasing, $\mathcal{I}_j = \mathcal{I}_k$ only when $j = k$. Thus,

$$\left[EE^\mathsf{T}\right]_{j,k} = \begin{cases} 1, & j = k, \\ 0, & j \neq k, \end{cases}$$

and so $EE^\mathsf{T} = I_m$. In contrast,

$$E^\mathsf{T}E = \sum_{\ell=1}^{m} \mathbf{e}_{\mathcal{I}_\ell}\mathbf{e}_{\mathcal{I}_\ell}^\mathsf{T}.$$

For each $\ell = 1, \ldots, m$, $\mathbf{e}_{\mathcal{I}_\ell}\mathbf{e}_{\mathcal{I}_\ell}^\mathsf{T}$ is an $n \times n$ diagonal matrix of zeros except for entry $(\mathcal{I}_\ell, \mathcal{I}_\ell)$, which is a 1. Since $m < n$, there exists some $j \in \{1, \ldots, n\}$ such that $\left[E^\mathsf{T}E\right]_{j,j} = 0$. Thus, $E^\mathsf{T}E \neq I_n$. □

If a vector $\mathbf{x}$ has $n = 2^\rho$ points, a natural downsampling is to select every other component of $\mathbf{x}$. The resulting downsampled vector would then have length $n/2 = 2^{\rho-1}$ and the downsampling matrix E would have dimension $n/2 \times n$.

For any real $m \times n$ matrix M and n-vector $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I_n)$

$$\operatorname{Var}(M\boldsymbol{\eta}) = M \operatorname{Var}(\boldsymbol{\eta}) M^\mathsf{T} = \sigma^2 MIM^\mathsf{T} = \sigma^2 MM^\mathsf{T} \tag{7.1.2}$$

where MM^T is an $m \times m$ matrix. Certainly for arbitrary M, MM^T can differ from an $m \times m$ identity matrix, which would mean that the new noise vector $M\boldsymbol{\eta}$ no longer

represents white noise. However, the variance of the noise vector does not change when the vector is downsampled using an $m \times n$ matrix E defined by eq. (7.1.1). From (7.1.2), the variance of a noise vector $E\boldsymbol{\eta}$ downsampled from $\boldsymbol{\eta}$ is

$$\mathrm{Var}(E\boldsymbol{\eta}) = E\,\mathrm{Var}(\boldsymbol{\eta})E^{\mathsf{T}} = \sigma^2 EE^{\mathsf{T}} = \sigma^2 I_m.$$

Therefore theoretical variance is unchanged across downsampling resolutions.

To accomplish a downsampling of an image $X \in \mathbb{R}^{n\times n}$ instead of a vector, we can pre- and post-multiply X by E and E^{T}, respectively. The downsampled image EXE^{T} is then of size $m \times m$. Figure 7.1 demonstrates how an image can be downsampled to reduced image size at the cost of image clarity. The consideration of square images is simply a convenience; the same downsampling matrix can be applied to downsample each individual dimension, though distinct matrices could be used to downsample non-square images.

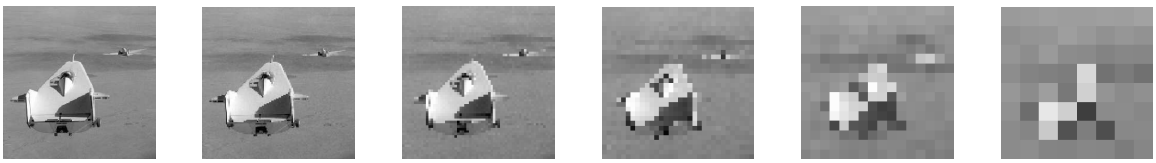


Figure 7.1: Downsamplings of the lifting body image from MATLAB®. The benefit of having starting images of size 256×256 is the reductions in problem size by powers of 2 can be accomplishes all the way down to the extreme of having a single pixel. Reducing each dimension by a factor of 2 reduces the overall problem size by 4.

While the variance of the noise is preserved across downsampling resolution in theory, numerically there is some fluctuation. As the downsampling resolutions decrease, i.e. the length of the downsampled vectors decreases, the sample variances differ. Figure 7.2 demonstrates this phenomenon by showing boxplots of sample variance versus downsampling resolutions.

Though downsampling of data can be easily explained through the use of a semi-orthogonal matrix such as E, the same approach cannot be applied to an original

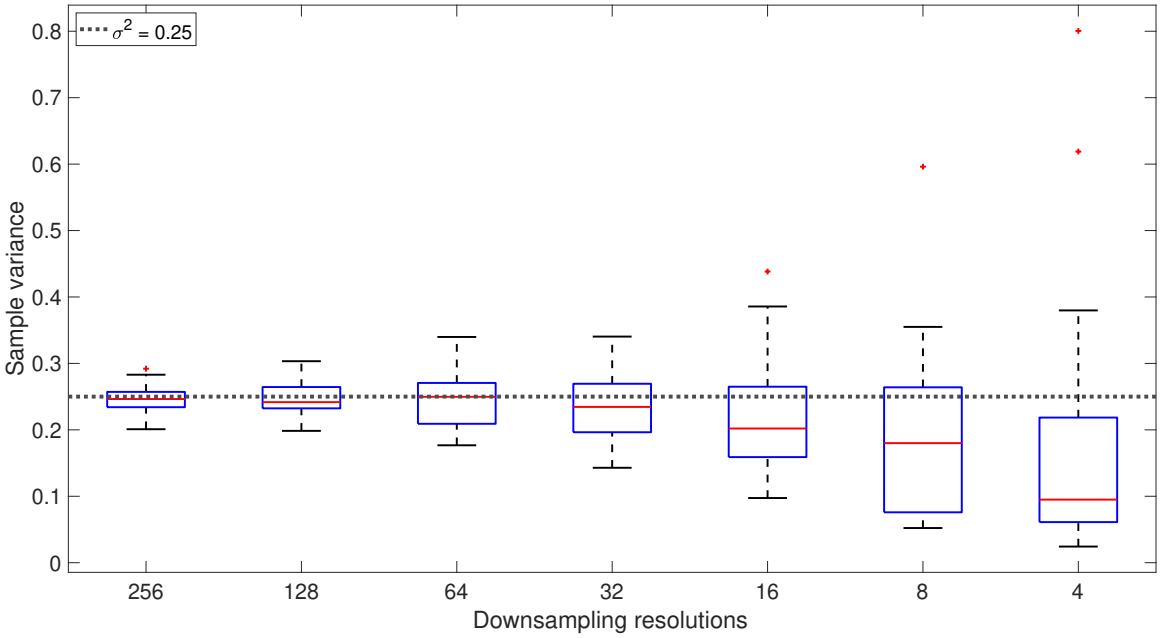


Figure 7.2: Boxplots generated from the sample variances of downsampled noise vectors. This figure illustrates that as the lengths of the downsampled vectors decrease, the variance in the computed sample variances increases.

system matrix A. Instead of downsampling A directly, downsampling perhaps should be applied to the kernel that was used to generate A in order to obtain a system matrix for the downsampled problem.

One approach to producing useful regularization parameters from downsampled data could be to analyze statistical term involved in the parameter methods. For example, both eq. (4.1.15) and (4.1.16) involve $\widehat{b}_j$, the DFT components of the blurred vector/image. Since the UPRE and MDP rely on $\mathbb{E}\left(\frac{1}{m}\left\|\widehat{\mathbf{r}}(\alpha)\right\|_2^2\right)$, perhaps analysis of eq. (4.1.15) or even (4.1.16) could yield improved methods. Perhaps one description of the effects of downsampling to be applied is the Downsampling Theorem (Smith, 2007, Ch. 7), which provides a relation between the DFT of a vector with the DFT of a downsampling. A direction of future research is to determine whether a result such as the Downsampling Theorem can be used to downsample data to effectively select regularization parameters.

We use the same MRI data from Section 6.1 to use as a 1D downsampling test problem. The 256×1 columns of the full MRI image were downsampled by powers of 2 down to columns of size 8×1. Figure 7.3 shows the visual result of this downsampling process. As with the 1D problem in section 6.1, the GSVD is utilized in the regularization process.

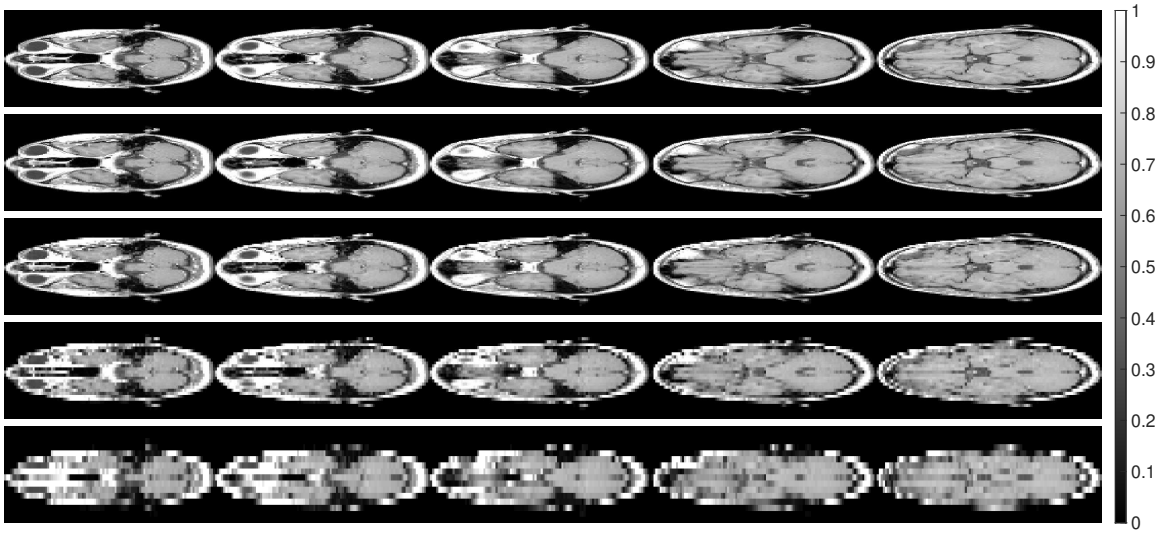


Figure 7.3: Downsampled MRI data formed by reformatting the MATLAB® built-in MRI data. The columns of each image was downsampled by a factor of 2, starting with the full columns of length 256. The number of rows in the last image is 8; the number of columns of all five images is 536.

The kernel associated with the symmetric Toeplitz matrix used in Section 6.1 was downsampled and system and penalty matrices were constructed for each problem size to be used with the GSVD. As with the previous 1D problem, a blur amount of $\xi = 4$ and an SNR of 25 was considered, as well as the finite-difference matrix L_1. As a selected example, the standard UPRE method eq. (4.1.6) was used to determine regularization parameters. The primary concern was the stability of the parameters across downsampling resolutions. The reason for this was that instability in obtained parameters across downsamplings would almost certainly result in regularized solu-

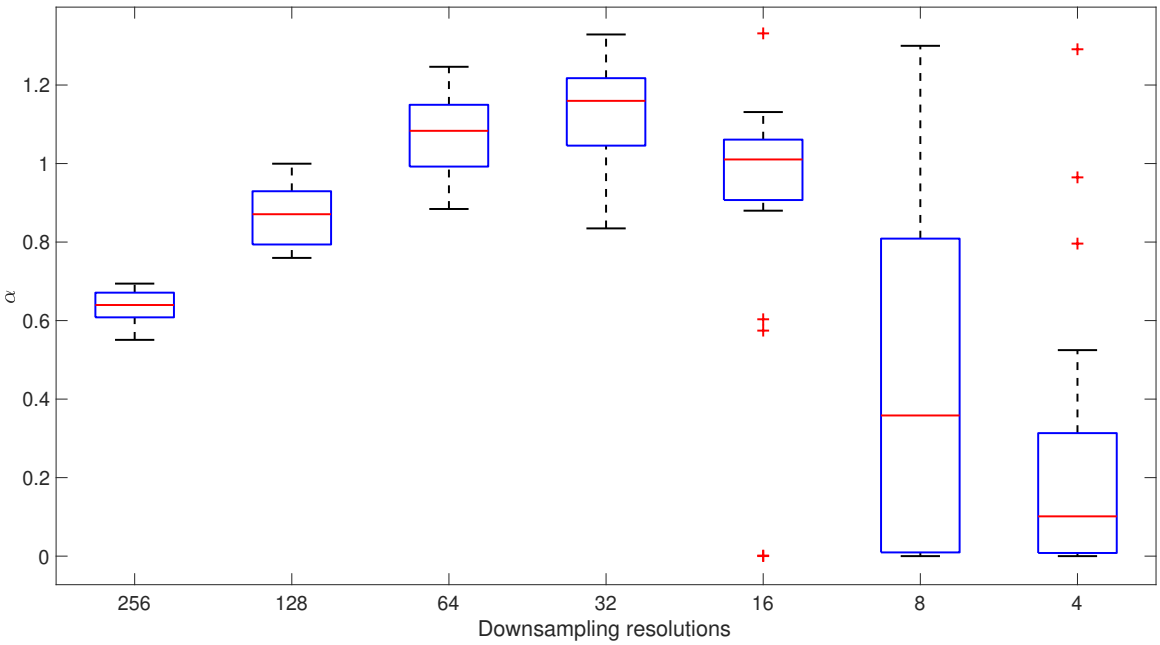


Figure 7.4: Parameters obtained using the UPRE method across downsamples. The increase in parameters stabilizes and reverses around a resolution of 32. A blur amount of $\xi = 4$, an SNR of 25, and $L = L_1$ were used to generate the results.

tions having high relative errors. Figure 7.4 shows some of the stabilizing behavior that is possible in a downsampled setting.

Results such as fig. 7.4 demonstrate the possibility of leveraging downsampling techniques to reduce problem size (and subsequently computational cost) and still obtain regularization parameters that produce meaningful solutions. In addition to applying methods developed in this work to iterative methods, the investigation of downsampling could prove to be a fruitful direction of future work.

7.3 Summary

Regularization parameters from the methods discussed in this work can be obtained from downsampled data. These parameters are then used to generate regularized solutions of the full (non-downsampled) problems. In certain cases, the parameters found from downsampled data are competitive with parameters found using

full data. The competitiveness in parameters found from downsampling motivates future application of downsampling to more complex problems, such as iterative-based methods.

Chapter 8

CONCLUSIONS AND FUTURE WORK

We have shown that the UPRE, MDP, and GCV methods can be extended to accommodate regularization parameter estimation using multiple data sets and for both single and multiple parameters, for generalized Tikhonov regularization. The UPRE and MDP are representative methods that assumes the knowledge of the variance of mean zero Gaussian noise in the data, while no additional assumptions are required for the GCV estimator. The most general forms of functions associated with these methods are eqs. (5.5.25), (5.6.30) and (5.7.38), respectively. While the corresponding functions for the spectrally windowed multidata UPRE and MDP can be written as an average of the individual functions associate with each data set, this is not possible for the spectrally windowed multidata GCV method without stronger assumptions (see Section 5.7). None of these methods require knowledge of true solutions unlike the learning approach defined by eq. (6.0.3). The presented numerical experiments for 2D signal restoration demonstrate that the windowed multidata methods can perform competitively with the learning approach that requires knowledge of true signals for training the data. Further, it is also demonstrated that the parameters obtained from a specific training set of validation images can also be used for a set of different testing images, provided that the general noise characteristics are the same.

The spectrally windowed multidata approach extends immediately for any estimator which relies only on an approximation for the regularized residual and the trace of the influence matrix, and the general idea can be modified to address estimators requiring other terms, such as an augmented regularized residual use for the χ^2 estimator described in Mead (2008); Mead and Renaut (2009). Although the derivations

are presented for the case in which there is a known mutual decomposition of the model and penalty matrices (A and L), because it is only needed to obtain estimates of the required terms, the approach can be extended for any iterative method which yields suitable estimates, e.g. Chung *et al.* (2008); Renaut *et al.* (2018).

Moving beyond the spectrally windowed multidata approach, the method of downsampling represents an effort to reduce problem size. The preliminary results in Chapter 7 show that the statistics of the parameter function terms rely on the structure of the underlying blurred vector/matrix. The DFT and DCT versions of the UPRE, GCV, and MDP functions in Chapter 4 are useful when the two transforms are to be used in the inversion process. The corresponding statistical results of the functions in Chapter 7, such as eq. (4.1.16), could be useful from an analytic standpoint.

There are many possible directions of future work regarding downsampling. One direction is to further quantify the statistics of the regularization parameters when downsampling is applied. Fortunately the DFT and DCT are topics which have been thoroughly investigated in a number of settings. Another means of quantification could be to utilize the statistics of the parameter estimation functions and try to estimate intervals in which α will be found for each downsampling level. If some information about the frequency content (in regard to the GSVD, DFT, or DCT) of the underlying solution is available, then the information can be used to select an appropriate downsampling level to capture the overall behavior of the solution; the ideal result would be to explicitly show how this process effects the resulting value of α.

As a final and overarching direction for future work, three-dimensional problems should be considered. Examples of settings that produce three-dimensional inverse problems include computed tomography and subsurface imaging by gravitational measurements Aster *et al.* (2013). The primary challenges of working with three-

dimensional problems are the numerical structure of the problems themselves, such as how to correctly express the problems as matrix-vector product, and the computational costs of solving the problems. However, three-dimensional problems could be the types of problems where downsampling is most useful. To illustrate this intuition, suppose a unit cube in $\mathbb{R}^3$ is discretized by using 4 sample points in each dimension, then downsampling to 2-point discetizations reduces the number of total data points from 64 to 8. While the quantitative effects of downsampling should be determined before moving to higher-dimensional problems, three-dimensional problems possess the most potential for demonstrating these effects.

Abramowitz, M. and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (United States Department of Commerce, National Bureau of Standards, 1972).

Afkham, B. M., J. Chung and M. Chung, "Learning regularization parameters of inverse problems via deep neural networks", Inverse Problems **37** (2021).

Arridge, S., P. Maass, O. Öktem and C.-B. Schönlieb, "Solving inverse problems using data-driven models", Acta Numerica **28**, 1–174 (2019).

Aster, R. C., B. Borchers and C. H. Thurber, *Parameter Estimation and Inverse Problems* (Elsevier, Amsterdam, 2013), 2nd edn.

Belge, M., M. E. Kilmer and E. L. Miller, "Efficient determination of multiple regularization parameters in a generalized L-curve framework", Inverse Problems **18**, 1161–1183 (2002).

Bell, A. J. and T. J. Sejnowski, "The "independent components" of natural scenes are edge filters", Vision Research **37**, 3327–3338 (1997).

Bergen, K. J., P. A. Johnson, M. V. de Hoop and G. C. Beroza, "Machine learning for data-driven discovery in solid earth geoscience", Science **363** (2019).

Boggess, A. and F. J. Narcowich, *A First Course in Wavelets with Fourier Analysis* (John Wiley & Sons, Inc., 2009), 2nd edn.

Brezinski, C., M. Redivo-Zaglia, G. Rodriguez and S. Seatzu, "Multi-parameter regularization techniques for ill-conditioned linear systesm", Numerische Mathematik **94**, 203–228 (2003).

Byrne, M. J. and R. A. Renaut, "Learning spectral windowing parameters for regularization using unbiased predictive risk and generalized cross validation techniques for multiple data sets", Inverse Problems and Imaging **17**, 4, 841–869 (2023).

Casella, G. and R. L. Berger, *Statistical inference* (Wadsworth, 2002), 2nd edn.

Chan, R. H., T. F. Chan and C.-K. Wong, "Cosine transform based preconditioners for total variation deblurring", IEEE Transactions on Image Processing **8**, 10, 1472–1478 (1999).

Chung, J., M. Chung and D. P. O'Leary, "Designing optimal spectral filters for inverse problems", SIAM Journal on Scientific Computing **33**, 3131–3135 (2011a).

Chung, J., G. Easley and D. O'leary, "Windowed spectral regularization of inverse problems", SIAM J. Scientific Computing **33**, 3175–3200 (2011b).

Chung, J. and M. I. Español, "Learning regularization parameters for general-form Tikhonov", Inverse Problems **33**, 7, 074004, URL `http://stacks.iop.org/0266-5611/33/i=7/a=074004` (2017).

Chung, J., J. G. Nagy and D. P. O'Leary, "A weighted GCV method for Lanczos hybrid regularization", Electronic Transactions on Numerical Analysis **28**, 149–167 (2008).

Chung, J. M., M. E. Kilmer and D. P. O'Leary, "A framework for regularization via operator approximation", SIAM Journal on Scientific Computing **37**, 2, B332–B359 (2015).

Debnath, L. and P. Mikusiński, *Introduction to Hilbert Spaces with Applications* (Elsevier, 2005), 3rd edn.

Dummit, D. S. and R. M. Foote, *Abstract algebra* (John Wiley & Sons, Inc., 2004), 3 edn.

Easley, G. R., D. Labate and V. M. Patel, "Directional multiscale processing of images using wavelets with composite dilations", Journal of Mathematical Imaging and Vision **48**, 13–34 (2014).

Gazzola, S., P. C. Hansen and J. G. Nagy, "IR Tools: A MATLAB package for iterative regularization methods and large-scale problems", Numerical Algorithms **81**, 773–811 (2019).

Gazzola, S. and P. Novati, "Multi-parameter Arnoldi-Tikhonov methods", Electronic Transactions on Numerical Analysis **40**, 452–475 (2013).

Golub, G. and C. Van Loan, *Matrix Computations*, Johns Hopkins Studies in the Mathematical Sciences (Johns Hopkins University Press, 2013).

Golub, G. H., M. Heath and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter", Technometrics **21**, 2, 215–223 (1979).

Haber, E. and L. Tenorio, "Learning regularization functionals—a supervised training approach", Inverse Problems **19**, 611–626 (2003).

Hadamard, J., "Sur les problèmes aux dérivées partielles et leur signification physique", Princeton University Bulletin **13**, 4, 49–52 (1904).

Hansen, J. K., "Downsampling for efficient parameter choice in ill-posed deconvolution problems", Arizona State University (2015).

Hansen, P. C., "The discrete Picard condition for discrete ill-posed problems", BIT Numerical Mathematics **30**, 4, 658–672, URL `https://doi.org/10.1007/BF01933214` (1990).

Hansen, P. C., "Analysis of discrete ill-posed problems by means of the L-curve", SIAM Review **34**, 561–580 (1992).

Hansen, P. C., *Rank-Deficient and Discrete Ill-Posed Problems* (Society for Industrial and Applied Mathematics, Philadelphia, 1998), URL `http://epubs.siam.org/doi/abs/10.1137/1.9780898719697`.

Hansen, P. C., J. G. Nagy and D. P. O'Leary, *Deblurring Images: Matrices, Spectra, and Filtering*, Fundamentals of Algorithms (Society for Industrial and Applied Mathematics, Philadelphia, 2006), URL http://epubs.siam.org/doi/abs/10.1137/1.9780898718874.

Hansen, P. C. and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems", SIAM J. Sci. Comput. **14**, 1487–1503 (1993).

Holler, G. and K. Kunisch, "Learning nonlocal regularization operators", Mathematical Control and Related Fields **12**, 1, 81–114, URL /article/id/d9c6a881-159a-4623-a2ed-e7661aac13d2 (2022).

James, G., D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning with Applications in R* (Springer, 2013).

Kailath, T. and V. Olshevsky, "Displacement structure approach to discrete-trigonometric-transform based preconditioners of G.Strang type and of T.Chan type", Calcolo **33**, 191–208 (1996).

Kalke, M. and S. Siltanen, "Adaptive frequency-domain regularization for sparse-data tomography", Inverse Problems in Science and Engineering **21**, 1099–1124 (2013).

Kalouptsidis, N., G. Carayannis and D. Manolakis, "Algorithms for block Toeplitz matrices with Toeplitz entries", Signal Processing **6**, 1, 77–81 (1984).

Kunisch, K. and T. Pock, "A bilevel optimization approach for parameter learning in variational models", SIAM Journal on Imaging Sciences **6**, 2, 938–983, URL https://doi.org/10.1137/120882706 (2013).

Leon, S. J., *Linear Algebra with Applications* (Prentice Hall, 2010), 8th edn.

LeVeque, R. J., *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems* (SIAM, 2007).

Lu, S. and S. V. Pereverzev, "Multi-parameter regularization and its numerical realization", Numerische Mathematik **118**, 1–31 (2011).

Mallows, C. L., "Some comments on c_p", Technometrics **15**, 4, 661–675 (1973).

Martucci, S. A., "Symmetric convolution and the discrete sine and cosine transforms", IEEE Transactions on Signal Processing **42**, 5, 1038–1051 (1994).

Mead, J. L., "Parameter estimation: A new approach to weighting a priori information", Journal of Inverse and Ill-posed Problems **16**, 175–193 (2008).

Mead, J. L. and R. A. Renaut, "A Newton root-finding algorithm for estimating the regularization parameter for solving ill-conditioned least squares problems", Inverse Problems **25**, 2, 025002, URL http://stacks.iop.org/0266-5611/25/i=2/a=025002 (2009).

Modarresi, K. and G. Golub, "Multi-level approach to numerical solution of inverse problems", in "CSC 2007: SIAM Workshop on Combinatorial Scientific Computing", (IEEE Computer Society, 2007a).

Modarresi, K. and G. Golub, "Using multiple generalized cross-validation as a method for varying smoothing effects", in "CSC 2007: SIAM Workshop on Combinatorial Scientific Computing", SIAM (IEEE Computer Society, 2007b).

Morozov, V. A., "On the solution of functional equations by the method of regularization", Soviet Mathematics Doklady **7**, 414–417 (1966).

NASA and JPL-Caltech, "Photojournal: Mercury", URL `https://photojournal.jpl.nasa.gov/targetFamily/Mercury` (2016).

Ng, M. K., R. H. Chan and W.-C. Tang, "A fast algorithm for deblurring models with Neumann boundary conditions", SIAM Journal on Scientific Computing **21**, 3, 851–866 (1999).

Penrose, R., "A generalized inverse for matrices", Mathematical Proceedings of the Cambridge Philosophical Society **51**, 3, 406–413 (1955).

Perrone, L., "Kronecker product approximations for image restoration with anti-reflective boundary conditions", Numerical Linear Algebra with Applications **13** (2006).

Rao, C. R., *Linear Statistical Inference and its Applications* (John Wiley & Sons, Inc., 1973), 2nd edn.

Renaut, R. A., M. Horst, Y. Wang, D. Cochran and J. Hansen, "Efficient estimation of regularization parameters via downsampling and the singular value expansion", BIT Numerical Mathematics **57**, 2, 499–529 (2017).

Renaut, R. A., S. Vatankhah and A. Helmstetter, "Regularization parameter estimation for hybrid RSVD solvers of large scale inverse problems", In Preparation (2018).

Sánchez, V., P. García, A. M. Peinado, J. C. Segura and A. Rubio, "Diagonalizing properties of the discrete cosine transforms", IEEE Transactions on Signal Processing **43**, 11, 2631–2641 (1995).

Sidey-Gibbons, J. A. M. and C. J. Sidey-Gibbons, "Machine learning in medicine: a practical introduction", BMC Medical Research Methodology **19** (2019).

Smith, J. O., III, *Mathematics of the Discrete Fourier Transform (DFT), with Audio Applications* (W3K Publishing, 2007), 2nd edn.

Stephanakis, I. M. and S. Kollias, "Generalized-cross-validation estimation of the regularization parameters of the subbands in wavelet domain regularized image restoration", in "Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284)", vol. 2, pp. 938–940 vol.2 (1998).

Strang, G., "The discrete cosine transform", SIAM Review **41**, 1, 135–147 (1999).

Taroudaki, V. and D. P. O'Leary, "Near-optimal spectral filtering and error estimation for solving ill-posed problems", SIAM Journal on Scientific Computing **37**, 6, A2947–A2968, URL `https://doi.org/10.1137/15M1019581` (2015).

Tikhonov, A. N., "Regularization of incorrectly posed problems", Soviet Mathematics Doklady **4**, 1624–1627 (1963).

Vito, E. D., L. Rosasco, A. Caponnetto, U. Giovannini and F. Odone, "Learning from examples as an inverse problem", Journal of Machine Learning Research **6**, 883–904 (2005).

Vogel, C., *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, Philadelphia, 2002), URL `http://epubs.siam.org/doi/abs/10.1137/1.9780898717570`.

Wahba, G., "Practical approximate solutions to linear operator equations when the data are noisy", SIAM Journal on Numerical Analysis **14**, 4, 651–667 (1977).

Wahba, G., *Spline Models for Observational Data*, chap. Estimating the Smoothing Parameter, pp. 52–62, CBMS-NSF Regional Conference Series in Applied Mathematics (SIAM, 1990).

Wang, Z., "Multi-parameter Tikhonov regularization and model function approach to the damped Morozov principle for choosing regularization parameters", Journal of Computational and Applied Mathematics **236**, 1815–1832 (2012).

Wood, S. N., "Modelling and smoothing parameter estimation with multiple quadratic penalties", Journal of the Royal Statistical Society: Series B **62**, 413–428 (2002).

Zobitz, J. M., T. Quaife and N. K. Nichols, "Efficient hyper-parameter determination for regularised linear BRDF parameter retrieval", International Journal of Remote Sensing **41**, 1437 – 1457 (2020).

APPENDIX A

NOTATION

Table A.1: Abbreviations of methods, decompositions, and transforms used throughout this work.

Notation	Description
UPRE	Unbiased predictive risk estimator; see Section 4.1 for standard UPRE method and Section 5.5 for windowed multidata UPRE method
MDP	Morozov discrepancy principle; see Section 4.2 for standard MDP method and Section 5.6 for windowed multidata MDP method
GCV	Generalized cross validation; see Section 4.3 for standard GCV method and Section 5.7 for windowed multidata GCV method
MSE	Mean squared error; see eq. (6.0.3) and eq. (6.0.4)
SVD	Singular value decomposition; see Section 2.1
GSVD	Singular value decomposition; see Section 2.3
BCCB	Block-circulant matrix with circulant blocks; either "C" can be replaced with "H" or "T", meaning Hankel and Toeplitz, respectively, all of which are used in Vogel (2002)
DFT	Discrete Fourier transform; see Section 3.3, specifically eq. (3.3.11) and eq. (3.3.13)
DCT	Discrete cosine transform; see Section 3.3, specifically eq. (3.3.15) and eq. (3.3.16)
DST	Discrete sine transform; see Section 3.3
$\text{vec}(\cdot)$	Vectorization function defined by eq. (3.1.2)
$\text{arr}(\cdot)$	Array operation defined by eq. (3.1.3)
$A \otimes B$	Kronecker product of arbitrary matrices A and B defined by eq. (3.1.4)

Table A.2: General notation used throughout the work. The superscript (r) can be dropped when considering a single data set.

Notation	Description
$A^{(r)}$	System matrix of size $m_r \times n_r$ of the rth data set
$\mathbf{x}^{(r)}$	True solution vector of size $n_r \times 1$ of the rth data set
$\mathbf{b}^{(r)}$	Product $A^{(r)}\mathbf{x}^{(r)}$ of size $m_r \times 1$
$\boldsymbol{\eta}^{(r)}$	Additive noise vector of size $m_r \times 1$ associated with the rth data set
$\mathbf{d}^{(r)}$	Data vector of size $m_r \times 1$ of the rth data set; result of the sum $\mathbf{b}^{(r)} + \boldsymbol{\eta}^{(r)}$
R	Maximum number of data sets being considered at one time, i.e. $r = 1, \ldots, R$

Table A.3: Notation used for general Tikhonov regularization. The superscript (r) and subscript r can be dropped when considering a single data set.

Notation	Description
α_r	Regularization parameter for the rth data set
$\mathbf{r}^{(r)}(\alpha_r)$	Regularized residual of size $m_r \times 1$ for the rth data set
$A^{(r)}(\alpha_r)$	Regularization matrix of size $m_r \times m_r$ for the rth data set; see eq. (4.1.3)
$\mathbf{x}(\alpha_r)$	Regularized solution of size $n_r \times 1$ of the rth data set; see eq. (2.3.19)
$L^{(r)}$	Penalty matrix of size $q_r \times n_r$ of the rth data set
$\phi_j^{(r)}(\alpha_r)$, $\psi_j^{(r)}(\alpha_r)$	Filter function and complement, respectively, used with the rth data set ($j = 1 : n_r$); see eq. (2.3.20)
$\Phi(\alpha_r)$, $\Psi(\alpha_r)$	Filter matrix and complement, respectively, of size $n_r \times n_r$

Table A.4: Notation used for spectral windowing applied to general Tikhonov regularization. The superscript (r) and subscript r can be dropped when considering a single data set.

Notation	Description
$\mathbf{w}^{(r,p)}$	Weight vector associated with data set r and window p; see Section 5.1, eqs. (5.1.1) and (5.1.3)
$W^{(r,p)}$	Window matrix for the rth data set and window p; $W^{(r,p)} = \operatorname{diag}\left(\mathbf{w}^{(r,p)}\right)$; see Section 5.1
$\mathbf{x}_{\text{win}}^{(r)}\left(\boldsymbol{\alpha}^{(r)}\right)$	Windowed regularized solution for the rth data set; see eq. (5.3.12)
P_r	Number of windows to be used with the rth data set; see Section 5.1

Table A.5: Notation for multidata constructions

Notation	Description
$\widetilde{A}$	Block diagonal system matrix of size $M \times N$, where $M = \sum_{r=1}^{R} m_r$ and $N = \sum_{r=1}^{R} n_r$; the rth diagonal block is $A^{(r)}$ for $r = 1, \ldots, R$.
$\widetilde{\mathbf{x}}$	Vertical concatenation of the true solutions $\{\mathbf{x}^{(r)}\}_{r=1}^{R}$; the length of $\widetilde{\mathbf{x}}$ is N
$\widetilde{\mathbf{b}}$	Vertical concatenation of the vectors $\{\mathbf{b}^{(r)}\}_{r=1}^{R}$; the length of $\widetilde{\mathbf{b}}$ is M
$\widetilde{\boldsymbol{\eta}}$	Vertical concatenation of the noise vectors $\{\boldsymbol{\eta}^{(r)}\}_{r=1}^{R}$; the length of $\widetilde{\mathbf{b}}$ is M
$\widetilde{\mathbf{d}}$	Result of the sum $\widetilde{\mathbf{b}} + \widetilde{\boldsymbol{\eta}}$
$\widetilde{\mathbf{x}}(\widetilde{\alpha})$	Regularized solution generated from $\widetilde{\mathbf{d}}$; see eq. (5.2.11)

Table A.6: Notation for parameter selection functions

Notation	Description
F_{UPRE}, F_{MDP}, F_{GCV}, F_{MSE}	Single parameter UPRE, MDP, GCV, and MSE functions applied to one data set, all of which are functions of α; see eq. (4.1.6) for UPRE, eq. (4.2.20) for MDP, eq. (4.3.24) for GCV, eq. (4.0.1) for MSE
$\overline{F}_{\text{UPRE}}$, $\overline{F}_{\text{MDP}}$, $\overline{F}_{\text{GCV}}$	Single parameter UPRE, MDP, and GCV functions applied to one data set, all of which are functions of α; see Prop. 5.5.2 for UPRE, Prop. 5.6.2 for MDP, and Prop. 5.7.4 for GCV
$\underset{\text{win}}{\widetilde{F}_{\text{UPRE}}}$, $\underset{\text{win}}{\widetilde{F}_{\text{MDP}}}$, $\underset{\text{win}}{\widetilde{F}_{\text{GCV}}}$, $\underset{\text{win}}{\widetilde{F}_{\text{MSE}}}$	Windowed UPRE, MDP, GCV, and MSE functions applied to R data sets, all of which are functions of $\widetilde{\boldsymbol{\alpha}}$; see Prop. 5.5.1 for UPRE, Prop. 5.6.1 for MDP, and Prop. 5.7.3 for GCV

Table A.7: List of assumptions used in Chapter 5. See the discussion at the end of Section 5.2 for how the assumptions relate with one another.

Number	Statement
Assumption 1	For $r = 1, \ldots, R$, assume that $\mathbf{b}^{(r)} = A^{(r)}\mathbf{x}^{(r)}$, $\mathbf{d}^{(r)} = \mathbf{b}^{(r)} + \boldsymbol{\eta}^{(r)}$, and $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}^{(r)}, \Sigma^{(r)})$ with the $\boldsymbol{\eta}^{(r)}$ being mutually independent. The vectors $\mathbf{b}^{(r)}$, $\mathbf{d}^{(r)}$, and $\boldsymbol{\eta}^{(r)}$ are of length m_r and $\mathbf{x}^{(r)}$ is of length n_r.
Assumption 2	Given $\boldsymbol{\eta}^{(r)} \sim \mathcal{N}(\mathbf{0}^{(r)}, \Sigma^{(r)})$ for $r = 1, \ldots, R$, we assume $\Sigma^{(r)} = \sigma_r^2 I_{m_r}$ (a constant diagonal matrix).
Assumption 3	For all $r = 1, \ldots, R$, we assume that $m_r = m$. In other words, we assume that the size of each data vector $\mathbf{d}^{(r)}$ is the same.
Assumption 4	We assume that there exist matrices $\Delta \in \mathbb{R}^{m \times n}$ and $\Lambda \in \mathbb{R}^{q \times n}$ such that $A^{(r)} = U^{(r)}\Delta\left(X^{(r)}\right)^{\mathsf{T}}$ and $L^{(r)} = V^{(r)}\Lambda\left(X^{(r)}\right)^{\mathsf{T}}$ for $r = 1, \ldots, R$, where $U^{(r)}$ and $V^{(r)}$ are orthogonal and $X^{(r)}$ is invertible.
Assumption 5	For $r = 1, \ldots, R$, assume that $A^{(r)} = A \in \mathbb{R}^{m \times n}$ and $L^{(r)} = L \in \mathbb{R}^{q \times n}$.
Assumption 6	For all $r = 1, \ldots, R$ and $p = 1, \ldots, P$, assume that $W^{(r,p)} = W^{(p)}$.

APPENDIX B

PROPERTIES OF THE PSEUDOINVERSE

The pseudoinverse is the unique matrix $A^\dagger \in \mathbb{R}^{n\times m}$ that satisfies the following Moore-Penrose conditions (Golub and Van Loan, 2013, p. 290):

$$\begin{aligned} &\text{(I)}\ \ AA^\dagger A = A, && \text{(III)}\ \ \left(AA^\dagger\right)^\mathsf{T} = AA^\dagger, \\ &\text{(II)}\ \ A^\dagger AA^\dagger = A^\dagger, && \text{(IV)}\ \ \left(A^\dagger A\right)^\mathsf{T} = A^\dagger A. \end{aligned}$$

Prop. B.0.1 describes some special cases of the pseudoinverse.

Proposition B.0.1 *Let $A \in \mathbb{R}^{m\times n}$ and let $A^\dagger \in \mathbb{R}^{n\times m}$ be the pseudoinverse of A.*

(i) If $\operatorname{rank}(A) = m$*, then* $A^\dagger = A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}$ *and so $A^\dagger$ is the right inverse of A.*

(ii) If $\operatorname{rank}(A) = n$*, then* $A^\dagger = \left(A^\mathsf{T}A\right)^{-1}A^\mathsf{T}$ *and so $A^\dagger$ is the left inverse of A.*

(iii) If $A = \mathbf{0}_{m\times n}$, then $A^\dagger = \mathbf{0}_{n\times m}$.

(iv) If $\operatorname{rank}(A) = m = n$*, then $A^\dagger = A^{-1}$.*

Proof. We prove part (i). The proofs of parts (ii) and (iii) are analogous to that of part (i), and part (iv) follows from the definition of $S^\dagger$ using $A = USV^\mathsf{T}$.

Proof of part (i): Since A has full row rank, AA^T is invertible. Let $B = A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}$.

The Moore-Penrose conditions are satisfied as follows:

(I): $ABA = \left(AA^\mathsf{T}\right)\left(AA^\mathsf{T}\right)^{-1}A = A$,

(II): $BAB = A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}\left(AA^\mathsf{T}\right)\left(AA^\mathsf{T}\right)^{-1} = A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1} = B$,

(III): $(AB)^\mathsf{T} = \left(\left(AA^\mathsf{T}\right)\left(AA^\mathsf{T}\right)^{-1}\right)^\mathsf{T} = I_{m\times m} = \left(AA^\mathsf{T}\right)\left(AA^\mathsf{T}\right)^{-1} = AB$,

(IV): $(BA)^\mathsf{T} = \left(A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}A\right)^\mathsf{T} = A^\mathsf{T}\left(AA^\mathsf{T}\right)^{-1}A = BA$.

As a final remark regarding part (iii) of Prop. B.0.1, any matrix other than $\mathbf{0}_{n\times m}$ would fail to satisfy (II) of the Moore-Penrose conditions. □

The pseudoinverse can also be used with orthogonal projections onto fundamental subspaces; some applications are described in Prop. B.0.2.

Proposition B.0.2 *Let $A \in \mathbb{R}^{m\times n}$, $P \in \mathbb{R}^{n\times n}$ be an orthogonal projection matrix, $A^\dagger \in \mathbb{R}^{n\times m}$ be the pseudoinverse of A, and let $r_A = \operatorname{rank}(A)$.*

(i) $AA^\dagger$ is the orthogonal projection onto $\operatorname{range}(A)$.

(ii) $A^\dagger A$ is the orthogonal projection onto $\operatorname{range}\left(A^\mathsf{T}\right)$.

(iii) $P^\dagger = P$.

(iv) $P(AP)^\dagger = (AP)^\dagger$.

(v) $(PA)^\dagger P = (PA)^\dagger$.

Proof. We prove parts (i) and (iv). The proof of part (iii) follows immediately from the Moore-Penrose conditions and the properties of orthogonal projections. The proof of parts (ii) and (v) are similar to that of parts (i) and (iv), respectively.

Proof of part (i): Let $B = AA^\dagger$. From (Golub and Van Loan, 2013, p. 82), we must show that B is idempotent ($B^2 = B$), symmetric, and that range(B) = range(A). We can use the SVD $A = USV^\mathsf{T}$ to demonstrate all three properties. Condition (II) of the Moore-Penrose conditions gives the idempotence of B:

$$B^2 = \left(AA^\dagger\right)\left(AA^\dagger\right) = A\left(A^\dagger AA^\dagger\right) = AA^\dagger = B.$$

For symmetry, we have that

$$B^\mathsf{T} = \left(AA^\dagger\right)^\mathsf{T} = \left(USS^\dagger U^\mathsf{T}\right)^\mathsf{T} = U\left(SS^\dagger\right)^\mathsf{T} U^\mathsf{T}.$$

By the definition of $S^\dagger$, the matrix $SS^\dagger \in \mathbb{R}^{m\times m}$ is

$$SS^\dagger = \begin{bmatrix} I_{r_A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Thus $SS^\dagger$ is symmetric and

$$B^\mathsf{T} = U\left(SS^\dagger\right)^\mathsf{T} U^\mathsf{T} = USS^\dagger U^\mathsf{T} = B,$$

showing that B is symmetric. If we carry out the multiplication $B = USS^\dagger U^\mathsf{T}$, then

$$B = U_{r_A} U_{r_A}^\mathsf{T}$$

where U_{r_A} consists of the first r_A columns of U. Since the first r_A columns of U form an orthonormal basis for range(A) (Leon, 2010, p. 340), B being equal to $U_{r_A} U_{r_A}^\mathsf{T}$ shows that range(B) = range(A). Therefore, B is the orthogonal projection onto range(A).

Proof of part (iv): Let $B = AP$ and $C = P(AP)^\dagger$. It will be shown that C satisfies the Moore-Penrose conditions.

(I): Substituting B and C into the product BCB, we have

$$BCB = APP(AP)^\dagger AP = AP(AP)^\dagger AP.$$

By definition, $(AP)^\dagger$ satisfies Moore-Penrose condition (I); thus $AP(AP)^\dagger AP = AP = B$.

(II): Checking the product CBC, we have

$$CBC = P(AP)^\dagger APP(AP)^\dagger = P(AP)^\dagger AP(AP)^\dagger.$$

$(AP)^\dagger$ satisfies Moore-Penrose condition (II), and so

$$P\left((AP)^\dagger AP(AP)^\dagger\right) = P(AP)^\dagger = C.$$

(III): First, we have that

$$(BC)^\mathsf{T} = \left(APP(AP)^\dagger\right)^\mathsf{T} = \left(AP(AP)^\dagger\right)^\mathsf{T} = AP(AP)^\dagger,$$

where the last equality results from the symmetry of $AP(AP)^\dagger$. Writing $P = P^2$, we have

$$(BC)^\mathsf{T} = AP^2(AP)^\dagger = (AP)\,P(AP)^\dagger = BC.$$

(IV): We begin with

$$(CB)^\mathsf{T} = \left(P(AP)^\dagger AP\right)^\mathsf{T} = PA^\mathsf{T}\left(PA^\mathsf{T}\right)^\dagger P.$$

Writing the leading P as $P = P^2$ and noting that $PA^\mathsf{T}\left(PA^\mathsf{T}\right)^\dagger$ is symmetric, we have

$$\begin{aligned} PA^\mathsf{T}\left(PA^\mathsf{T}\right)^\dagger P = P\left(PA^\mathsf{T}\left(PA^\mathsf{T}\right)^\dagger\right)^\mathsf{T} P = P(AP)^\dagger AP^2 = P(AP)^\dagger AP \\ = CB. \end{aligned}$$

Therefore $C = P(AP)^\dagger$ satisfies the Moore-Penrose conditions, meaning that $P(AP)^\dagger$ is the pseudoinverse of $B = AP$.

□

Proposition B.0.3 *Let $A \in \mathbb{R}^{m\times n}$ and $A^\dagger \in \mathbb{R}^{n\times m}$ be the pseudoinverse of A. Then*

$$A^\mathsf{T} = A^\mathsf{T} AA^\dagger.$$

Proof. Beginning with condition (III) and using the properties of matrix transposition,

$$\begin{aligned} \left(AA^\dagger\right)^\mathsf{T} &= AA^{-1} \\ \left(A^\dagger\right)^\mathsf{T} A^\mathsf{T} &= AA^{-1} \\ A^\mathsf{T}\left(A^\dagger\right)^\mathsf{T} A^\mathsf{T} &= A^\mathsf{T} AA^{-1} \\ \left(AA^\dagger A\right)^\mathsf{T} &= A^\mathsf{T} AA^{-1}. \end{aligned}$$

Condition (I) yields the final result:

$$\begin{aligned} \left(AA^\dagger A\right)^\mathsf{T} &= A^\mathsf{T} AA^{-1} \\ A^\mathsf{T} &= A^\mathsf{T} AA^{-1}. \end{aligned}$$

□

APPENDIX C

COMPARISON OF DERIVATIVES OF RESIDUALS

As with Cor. 5.4.1, a comparison of the norms of $\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})$ and $\widetilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})$ can be made with respect to their partial derivatives; this comparison is described in Cor. C.0.1.

Corollary C.0.1 *Under Assumptions 1 to 6, for all $\boldsymbol{\alpha} \in \mathbb{R}_+^P$ we have that*

$$\frac{\partial}{\partial \alpha^{(p)}} \left\| \overline{\mathbf{r}}_{win}(\boldsymbol{\alpha}) \right\|_2^2 \leq \frac{1}{R^2} \frac{\partial}{\partial \alpha^{(p)}} \left\| \widetilde{\mathbf{r}}_{win}(\boldsymbol{\alpha}) \right\|_2^2, \quad p = 1, \ldots, P.$$

Proof. Without loss of generality, suppose that $m \leq n$ and let $k = n - m$. By the definition of $\overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha})$, for fixed $p = 1, \ldots, P$ we have

$$\begin{aligned}
\frac{\partial}{\partial \alpha^{(p)}} \left\| \overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha}) \right\|_2^2 &= \frac{\partial}{\partial \alpha^{(p)}} \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k} \left(\alpha^{(p)} \right) \right]^2 \left(\hat{\bar{d}}_j \right)^2 \\
&= 2 \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k} \left(\alpha^{(p)} \right) \right] \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \frac{\partial \Psi_{j+k} \left(\alpha^{(p)} \right)}{\partial \alpha^{(p)}} \right] \left(\hat{\bar{d}}_j \right)^2.
\end{aligned}$$

Similarly, for $\mathbf{r}_{\text{win}}^{(r)}(\boldsymbol{\alpha})$ with $r = 1, \ldots, R$ we have

$$\begin{aligned}
\frac{\partial}{\partial \alpha^{(p)}} \left\| \mathbf{r}_{\text{win}}^{(r)}(\boldsymbol{\alpha}) \right\|_2^2 &= \frac{\partial}{\partial \alpha^{(p)}} \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k} \left(\alpha^{(p)} \right) \right]^2 \left(\hat{d}_j^{(r)} \right)^2 \\
&= 2 \sum_{j=1}^{m} \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k} \left(\alpha^{(p)} \right) \right] \left[\sum_{p=1}^{P} w_{j+k}^{(p)} \frac{\partial \Psi_{j+k} \left(\alpha^{(p)} \right)}{\partial \alpha^{(p)}} \right] \left(\hat{d}_j^{(r)} \right)^2.
\end{aligned}$$

Thus, through a change of summation

$$\left\| \overline{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha}) \right\|_2^2 \leq \frac{1}{R^2} \sum_{r=1}^{R} \left(\sum_{j=1}^{m} \left(\left[\sum_{p=1}^{P} w_{j+k}^{(p)} \Psi_{j+k} \left(\alpha^{(p)} \right) \right] \hat{d}_j^{(r)} \right)^2 \right) = \frac{1}{R^2} \left\| \widetilde{\mathbf{r}}_{\text{win}}(\boldsymbol{\alpha}) \right\|_2^2$$

where the last equality follows from Thm. 5.4.1 in conjunction with Assumption 6. □