# Computational Advancements for Solving Large-scale Inverse Problems

Taewon Cho

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mathematics

Julianne Chung, Chair
Jeff Borggaard
Matthias Chung
Eileen Martin

May 11, 2021
Blacksburg, Virginia

Keywords: inverse problems, uncertainty quantification, generalized Golub–Kahan, hybrid projection methods, Tikhonov regularization, Bayesian inverse problems, sample covariance matrix, blind deconvolution, alternating optimization, tomography, respirometry

# Computational Advancements for Solving Large-scale Inverse Problems

Taewon Cho

(ABSTRACT)

For many scientific applications, inverse problems have played a key role in solving important problems by enabling researchers to estimate desired parameters of a system from observed measurements. For example, large-scale inverse problems arise in many global problems and medical imaging problems such as greenhouse gas tracking and computational tomography reconstruction. This dissertation describes advancements in computational tools for solving large-scale inverse problems and for uncertainty quantification.

Oftentimes, inverse problems are ill-posed and large-scale. Iterative projection methods have dramatically reduced the computational costs of solving large-scale inverse problems, and regularization methods have been critical in obtaining stable estimations by applying prior information of unknowns via Bayesian inference. However, by combining iterative projection methods and variational regularization methods, hybrid projection approaches, in particular generalized hybrid methods, create a powerful framework that can maximize the benefits of each method. In this dissertation, we describe various advancements and extensions of hybrid projection methods that we developed to address three recent open problems. First, we develop hybrid projection methods that incorporate mixed Gaussian priors, where we seek more sophisticated estimations where the unknowns can be treated as random variables from a mixture of distributions. Second, we describe hybrid projection methods for mean estimation in a hierarchical Bayesian approach. By including more than one prior covariance matrix (e.g., mixed Gaussian priors) or estimating unknowns and hyper-parameters

simultaneously (e.g., hierarchical Gaussian priors), we show that better estimations can be obtained. Third, we develop computational tools for a respirometry system that incorporate various regularization methods for both linear and nonlinear respirometry inversions. For the nonlinear systems, blind deconvolution methods are developed and prior knowledge of nonlinear parameters are used to reduce the dimension of the nonlinear systems. Simulated and real-data experiments of the respirometry problems are provided. This dissertation provides advanced tools for computational inversion and uncertainty quantification.

# Computational Advancements for Solving Large-scale Inverse Problems

Taewon Cho

(GENERAL AUDIENCE ABSTRACT)

For many scientific applications, inverse problems have played a key role in solving important problems by enabling researchers to estimate desired parameters of a system from observed measurements. For example, large-scale inverse problems arise in many global problems such as greenhouse gas tracking where the problem of estimating the amount of added or removed greenhouse gas at the atmosphere gets more difficult. The number of observations has been increased with improvements in measurement technologies (e.g., satellite). Therefore, the inverse problems become large-scale and they are computationally hard to solve. Another example of an inverse problem arises in tomography, where the goal is to examine materials deep underground (e.g., to look for gas or oil) or reconstruct an image of the interior of the human body from exterior measurements (e.g., to look for tumors). For tomography applications, there are typically fewer measurements than unknowns, which results in non-unique solutions. In this dissertation, we treat unknowns as random variables with prior probability distributions in order to compensate for a deficiency in measurements. We consider various additional assumptions on the prior distribution and develop efficient and robust numerical methods for solving inverse problems and for performing uncertainty quantification. We apply our developed methods to many numerical applications such as greenhouse gas tracking, seismic tomography, spherical tomography problems, and the estimation of $CO_2$ of living organisms.

# Dedication

*To Youjin Lee*

# Acknowledgments

I am deeply grateful for the various support and assistance I received from faculty, fellow graduate students, and family during my six years of graduate research and teaching at Virginia Tech.

First, I would like to thank my advisor, Prof. Julianne Chung, who has patiently guided and supported me throughout my graduate study. There are no words to describe the full extent of my appreciation for your precious advice and ongoing support.

I would also like to acknowledge the other members of my committee. To Prof. Tia Chung, you helped me achieve confidence and significant self-development in the Chung Lab. To Prof. Jeff Borggaard and Prof. Eileen Martin, I thank you both for serving on my committee and providing valuable feedback regarding my work.

I would also like to expression my appreciation for my collaborators. Dr. Jiahua Jiang organized a great collaboration for my first journal publication. Dr. Hodjat Pendar presented me with extremely motivating problems. Dr. Scot Miller and Dr. Arvind Saibaba gave me concrete guidance to choose the right direction and thus successfully complete my research.

I would like to thank the National Science Foundation for supporting my graduate research under grants DMS-1638521 to the Statistical and Applied Mathematical Science Institute and DMS-2026841, 2026830, and 2026835 to the Algorithms for Threat Detection program.

Finally, I could not have completed this work without the support of my amazing wife, Youjin Lee, my parents, and parents-in-law, all of whom provided a comfortable environment as well as happy distractions to rest outside of my research work.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Inverse problems arise in many scientific applications such as medical imaging, geophysics, biology, and atmospheric science. This dissertation develops computational tools for solving large-scale inverse problems and for uncertainty quantification. In many applications, solving linear inverse problems is an important and challenging part of inverse problems. In general, a linear inverse problem can be written by

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \boldsymbol{\epsilon} \tag{1.1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{s} \in \mathbb{R}^n$, and $\mathbf{d}, \boldsymbol{\epsilon} \in \mathbb{R}^m$. Usually, $\mathbf{A}$ represents a forward modeling, $\mathbf{d}$ contains observations or measurements, $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ represents Gaussian random noise where the covariance matrix $\mathbf{R}$ is assumed to be easy to invert and factorize (e.g., diagonal matrix), and $\mathbf{s}$ contains the desired parameters.

## 1.1 Motivating Examples of Inverse Problems

We begin this dissertation by introducing some examples of inverse problems. First, in respirometry systems [89, 90], the blurring process can be represented by a mathematical model of convolution. The goal of this process is to estimate emissions of Carbon Dioxide ($CO_2$) of living organisms using a flow-through respirometry chamber. Getting a reconstruc-

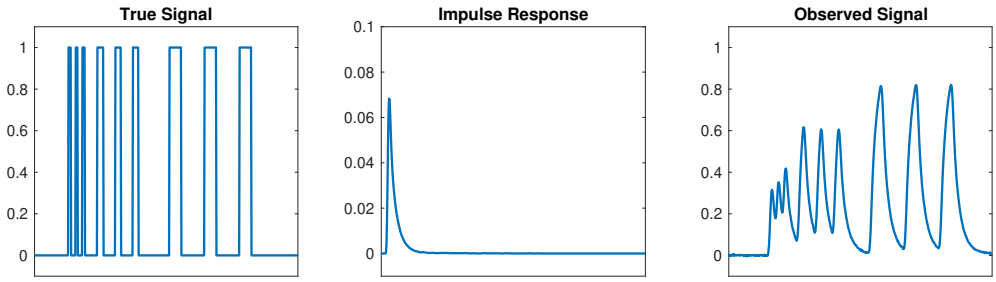


Figure 1.1: The blurred signal (right) is obtained from convolution between the true signal (left) and the impulse response function (middle).

tion which is a sharper and sparse signal from the blurred signal as described in Figure 1.1 is a typical inverse problem. For example, in image deblurring, rather than the impulse response function, there are various types of blurs (e.g., represented using the point spread function [61]). Since we only have the blurred signal (or image) as an observation, we have to consider methods to reconstruct or estimate the true signal (or image) with a known or even an unknown impulse response function.

The second example that we consider is tomography reconstruction, see Figure 1.2. In this example, we wish to obtain information from inside the human body or deep underground, but cannot access this information directly due to physical and technical limitations. Instead, we collect measurements from the surface or from the outside of the object and solve an inverse problem to reconstruct the interior. For example, in crosswell seismic tomography [2] the goal is to determine the slowness in a medium where the slowness is the reciprocal of the seismic velocity. By shooting rays from some sources to receivers, we can observe the total elapsed time of travel for each ray. From these observations, the goal of the inverse problem is to estimate the slowness properties of the the mass.

Another example of a tomographic reconstruction problem is spherical tomography [60], which is described in Figure 1.3. Such models are often used in imaging problems from photoacoustic or optoacoustic imaging applications [70], and the forward model is based on

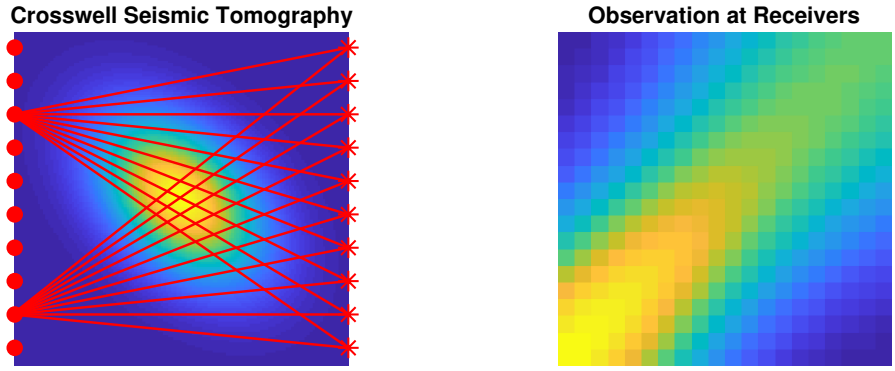


Figure 1.2: The images show for seismic tomography the true slowness of the mass (left) and the collected data at each receiver from each source (right). The red filled-in circles denote the sources and the red stars denote the receivers.

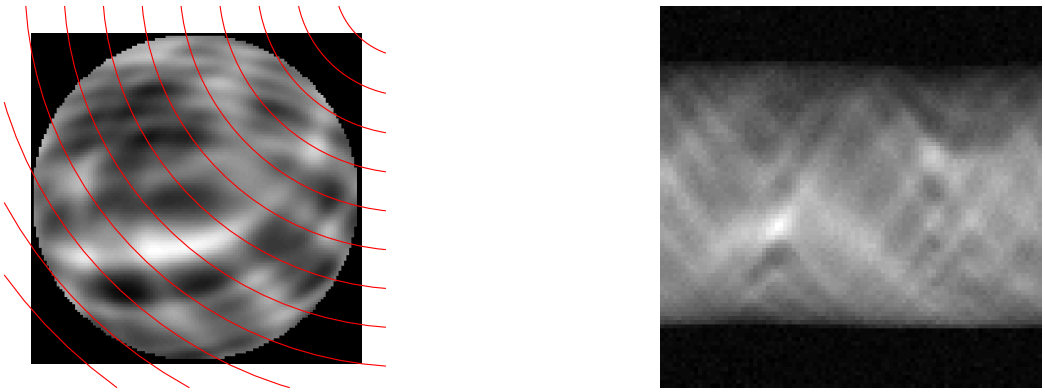

Figure 1.3: The forward process of spherical tomography (left) and the measurement sinogram (right).

the spherical Radon transform. This transform integrates the data along circles starting outside the object. The size of the observation dataset depends on the number of circles and the range of circle center's angles.

The last example that we consider is an atmospheric imaging problem where the goal is to track greenhouse gases using satellites. Estimating $CO_2$ fluxes at the Earth's surface is a prime example. As described in Figure 1.4, the true simulated $CO_2$ fluxes consist of

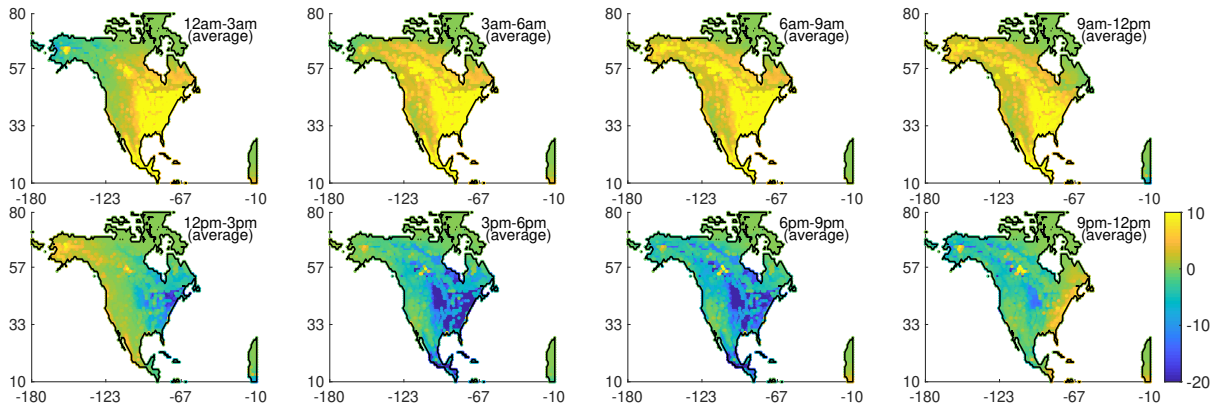


Figure 1.4: Average $CO_2$ fluxes across North America from late June through July 2015 [1].

temporal and spatial resolutions. With an atmospheric transport model used to represent the forward modeling process [73, 83], we can access limited observations using satellites [20, 46, 47, 82, 113]. Then, given statistical models and the observations, we can solve an underdetermined system to estimate changes in emissions over time.

| Problems | Input ($\mathbf{s}$) | Forward Model ($\mathbf{A}$) | Output ($\mathbf{d}$) |
|---|---|---|---|
| Respirometry | $CO_2$ emissions | Blurring Process | $CO_2$ observation |
| Seismic Tomography | Slowness of Mass | Ray Tracing | Travel Time |
| Spherical Tomography | Pressure Distribution | Radon Transform | Sinogram |
| Greenhouse Gas Tracking | $CO_2$ Fluxes | Atmospheric Transport | Satellite Observations |

Table 1.1: Linear models from motivations.

In Table 1.1, we provide a summary of these four motivating examples of inverse problems. Each of these applications can be represented using the linear model (1.1) and require efficient computational techniques for computing solutions and performing uncertainty quantification.

In general inverse problems have four main components: (i) There are inputs (e.g., the original signal, the true image, or the slowness of an object) and outputs (e.g., observed data or measurements). (ii) There is a forward model describing a system that maps the input

to the output. (iii) The forward problem is to compute the output given knowledge of the input and the forward system. (iv) The inverse problem is to compute either the input given the system and the output, or to compute both the input and the system from the given output.

## 1.2 Challenges of solving inverse problems

According to Hadamard [57], there are three conditions for a problem to be well-posed: (i) a solution exists, (ii) the solution is unique, and (iii) the solution must depend continuously on the data.

If the problem does not satisfy one or more of these conditions, it is called ill-posed. Unfortunately, most inverse problems are ill-posed problems. For some applications, that means that a solution does not exist. For others that means that there are infinitely many solutions (i.e., non-uniqueness of a solution). Furthermore, since data are not perfect, which means that errors are introduced during the measurement process, and since numerical errors can be introduced during solution computation, the potential propagation of errors during the solution process must be considered. Note that for a linear problem, a tiny perturbation in the output data can cause significantly large errors in the reconstruction of the input if the linear problem is ill-posed. To get a stable estimate of the true or exact solution for a given problem, we need to impose more conditions or constraints on the solutions. This is often referred to as regularization.

The general idea of regularization is to enforce some regularity or constraints on the solution (e.g., a smoothness or sparsity constraint). By including a penalty term with an objective function (typically the data-fit term) and using a numerical optimization process, we can suppress the unwanted parts of the computed solution. In addition, some problems may

need further constraints on the solution quantities (e.g., nonnegativity).

There are various regularization approaches. Common examples are the $\ell_2$ norm, $\ell_1$ norm, total variation (TV), constrained optimization, and iterative regularization. Choosing the type of regularization method depends on the character of the problem and the desired parameters. If the desired parameters are sparse (i.e., many parameters should be zero), then $\ell_1$-regularization will perform better than others. If the desired parameters should remove noise and reconstruct edge information, then TV is preferable. If the parameters should be bounded within a specified range, then constrained optimization techniques should be considered. For very large-scale problems where the number of observations and unknown parameters are in the millions or hundreds of millions, iterative methods may be the only computationally feasible option. For such problems, iterative regularization can be achieved by early termination of the iterative process, or a hybrid approach can be used.

One of challenging issues for any regularization approach is determining how to choose the regularization parameters in different scenarios. For example, the $\ell_2$ norm (or Tikhonov) regularization parameter can be chosen by filtering methods which is introduced later sections, and flexible methods can be used to estimate the $p$-norm regularization ($1 < p \leq 2$) parameter [34]. Many of the other types of regularization methods require the user to choose the regularization parameter in advance.

To get a good estimation, we use modern mathematical and computational tools to advance reconstruction methodologies for the inverse problem. For many applications, the ability to obtain good image or signal reconstructions from observation data requires the inclusion of a suitable prior. Priors provide a systematic and different means to describe in probabilistic terms any prior knowledge about the unknowns [7, 22] via a Bayesian formulation. Oftentimes prior knowledge will come from a characteristic property or a combination of sources, and striking a good balance of information is critical. For example, priors may be learned

from available training data, but bias in the reconstructions can have a significant impact (e.g., when the number of samples in the training set is not large enough or when the desired image is very different from the training set) [4, 33, 56]. Thus, a safer approach is to include a prior that combines learned information with conventional smoothness properties. In other scenarios (e.g., in seismic tomography [2]), the desired solution may consist of components with different smoothness properties, and the correct mixture of smoothness priors can be difficult to know *a priori*. Using mixed Gaussian priors, where the prior covariance matrix can be represented as a convex combination of matrices, is a common approach to incorporate different prior covariance matrices [5]. Another type of prior knowledge may be that the desired input signal is sparse. For example, in inverse respirometry, reconstructions that correspond to smooth priors tend to be over-smoothed or contain many unnecessary artifacts. By using a sparsity-enforcing prior, the input signal can be reconstructed accurately. Incorporating a good prior is key to solving ill-posed inverse problems. However, various computational challenges arise, especially when the number of unknowns is very large or when the regularization and mixing parameter are not known in advance.

Another challenge of solving inverse problems is that the forward model may not be known exactly. In this case, unknowns are also present in the impulse response function which models the forward operation. This is a significantly more challenging problem. One approach is to incorporate separable priors. Another approach is to treat these sets of parameters separately.

## 1.3 Overview of Contributions

The goal of this dissertation is to develop numerical methods to efficiently solve large-scale inverse problems and to perform uncertainty quantification. In Chapter 2, we provide some

background approaches, where our goal is not to provide a comprehensive overview, but rather to cover relevant material needed to motivate the new computational approaches and further developments that are described in later sections. The dissertation is divided into three main parts.

- In Chapter 3, we describe hybrid projection methods for large-scale inverse problems with mixed Gaussian priors. The main contribution here is developing a mixed Golub-Kahan process which is an extension of the generalized Golub-Kahan bidiagonalization. A distinctive feature of the proposed approach is that both the regularization parameter and the weighting parameter for the covariance matrix can be estimated automatically during the iterative process. Furthermore, sample covariance matrix and statistically learned covariance kernels can be easily incorporated. This work is published in [29].

- In Chapter 4, we extend generalized hybrid projection methods to the problem of mean estimation using a hierarchical Gaussian prior. The main contributions here are estimating unknowns and a prior mean simultaneously and giving more flexibility in the definition of the prior mean. Additionally, we describe computational tools to obtain approximations of the posterior distribution using matrices generated from generalized hybrid projection methods. This work appears in [30].

- In Chapter 5, we describe computational tools for inversion and uncertainty quantification in respirometry applications. The main contributions here are: First, providing a robust set of computational tools for the linear respirometry problem and the described methods can either include different regularizers, accelerate iterate methods for large-scale problems, automatically select regularization parameters, or provide quantification of solution uncertainties. Second, using alternating optimization for the nonlinear respirometry problem with reduced system of separable nonlinear equations

given prior knowledge. This work appears in [31].

Chapter 6 provides a summary of the dissertation and some ideas for future work.

# Chapter 2

# Background

In this chapter, we describe some background work on the methods and approaches that will be used throughout the dissertation. First, Bayesian approaches are explained to show how to incorporate prior knowledge about the unknowns when those unknowns are treated as random variables [22, 93]. Second, we describe hybrid iterative projection methods for large-scale inverse problems. We focus on generalized iterative projection methods since they can be used to include Gaussian priors where the prior covariance matrix is not explicitly constructed by access via matrix-vector multiplications [38]. Third, we concisely explain some common regularization parameter selection methods for Tikhonov regularization [7, 59]. Fourth, we describe some recent work on efficient methods for uncertainty quantification for cases where the posterior distribution can be described by a Gaussian distribution (e.g., the likelihood and the prior are Gaussian). We provide a brief summary of an efficient approximation of the posterior distribution described in [97].

## 2.1 Bayesian Inversion and Uncertainty Quantification

Assume that $\mathbf{s}$ is a Gaussian random variable with mean $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$. That means,

$$\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}, \lambda^{-2}\mathbf{Q}) \tag{2.1}$$

where $\lambda$ is a scaling parameter to be determined later. By Bayes' Theorem, the posterior probability density function is proportional to multiplication of the likelihood and the prior functions as

$$\pi_{\text{post}}(\mathbf{s}|\mathbf{d}) \propto \pi_{\text{like}}(\mathbf{d}|\mathbf{s})\pi_{\text{prior}}(\mathbf{s}) \tag{2.2}$$

where given the above assumptions, the likelihood and the prior can be written as respectively

$$\pi_{\text{like}}(\mathbf{d}|\mathbf{s}) \propto \exp\left(-\frac{1}{2}(\mathbf{A}\mathbf{s}-\mathbf{d})^\top \mathbf{R}^{-1}(\mathbf{A}\mathbf{s}-\mathbf{d})\right) \tag{2.3}$$

$$\pi_{\text{prior}}(\mathbf{s}) \propto \exp\left(-\frac{\lambda^2}{2}(\mathbf{s}-\boldsymbol{\mu})^\top \mathbf{Q}^{-1}(\mathbf{s}-\boldsymbol{\mu})\right) \tag{2.4}$$

To maximize a posterior (MAP), we minimize $-\ln \pi_{\text{post}}(\mathbf{s}|\mathbf{d})$ and thus

$$\mathbf{s}_{\text{MAP}} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\|\mathbf{s}-\boldsymbol{\mu}\|^2_{\mathbf{Q}^{-1}} \tag{2.5}$$

where $\|\mathbf{x}\|^2_{\mathbf{M}} = \mathbf{x}^\top\mathbf{M}\mathbf{x}$ for symmetric positive definite matrix $\mathbf{M}$. Then, $\mathbf{x}_{\text{MAP}}$ is equivalent to the Tikhonov solution where $\lambda$ has the roles of being a regularization parameter as well as a solution to normal equation such that

$$(\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A} + \lambda^2\mathbf{Q}^{-1})\mathbf{s} = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{d} + \lambda^2\mathbf{Q}^{-1}\boldsymbol{\mu} \tag{2.6}$$

The MAP estimate, $\mathbf{s}_\lambda$, can be written as a solution of the Tikhonov problem,

$$\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{L}_{\mathbf{R}}(\mathbf{A}\mathbf{s}-\mathbf{d})\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{L}_{\mathbf{Q}}(\mathbf{s}-\boldsymbol{\mu})\|_2^2 \tag{2.7}$$

where $\mathbf{Q}^{-1} = \mathbf{L}_{\mathbf{Q}}^{\top}\mathbf{L}_{\mathbf{Q}}$ and $\mathbf{R}^{-1} = \mathbf{L}_{\mathbf{R}}^{\top}\mathbf{L}_{\mathbf{R}}$. And the posterior distribution of $\mathbf{s}$ given $\mathbf{d}$ is the Gaussian distribution,

$$\mathbf{s} \sim \mathcal{N}(\mathbf{\Gamma}_{\text{post}}(\mathbf{A}^{\top}\mathbf{R}^{-1}\mathbf{d} + \lambda^2\mathbf{Q}^{-1}\boldsymbol{\mu}), \mathbf{\Gamma}_{\text{post}}) \tag{2.8}$$

where $\mathbf{\Gamma}_{\text{post}} = (\mathbf{A}^{\top}\mathbf{R}^{-1}\mathbf{A} + \lambda^2\mathbf{Q}^{-1})^{-1}$.

However, it is not computationally feasible to get a correct covariance matrix for a large-scale Gaussian random variable. We focus on using Gaussian random fields to represent prior information and summarize some common choices for the (unscaled) prior covariance matrix $\mathbf{Q}$. Oftentimes, the covariance matrix is generated using a covariance function (also called a kernel function). Covariance functions are crucial in many fields and encode assumptions about the form of the function that we are modeling. In most cases, the prior covariance matrix $\mathbf{Q}$ is large and dense with entries directly computed as $\mathbf{Q}_{ij} = \kappa(\mathbf{z}_i, \mathbf{z}_j)$, where $\{\mathbf{z}_i\}_{i=1}^n$ are the spatial points in the domain and $\kappa(\cdot,\cdot)$ is a covariance kernel function. Some commonly used parametric covariance functions [92] are provided in Table 2.1.

| | covariance kernel function |
|---|---|
| squared exponential | $\exp\left(-\frac{r^2}{2\ell^2}\right)$ |
| Matérn | $\frac{1}{2^{\nu-1}\Gamma(\nu)}\left(\frac{\sqrt{2\nu}r}{\ell}\right)^{\nu} K_\nu\left(\frac{\sqrt{2\nu}r}{\ell}\right)$ |
| $\gamma-$exponential | $\exp\left(-\left(\frac{r}{\ell}\right)^{\gamma}\right)$ |
| rational quadratic | $\left(1+\frac{r^2}{2\nu\ell^2}\right)^{-\nu}$ |
| sinc | $\frac{\sin(\nu r)}{\nu r}$ |

Table 2.1: Summary of commonly-used covariance functions. The covariance functions are written either as functions of $\mathbf{z}_i$ and $\mathbf{z}_j$, or as a function of $r = |\mathbf{z}_i - \mathbf{z}_j|$ and depend on $\ell$ or $\ell$ and $\nu$. $\Gamma$ is the Gamma function and $K_\nu(\cdot)$ is the modified Bessel function of the second kind of order $\nu$.

For some kernel choices, the precision matrix (i.e., the inverse of the covariance matrix)

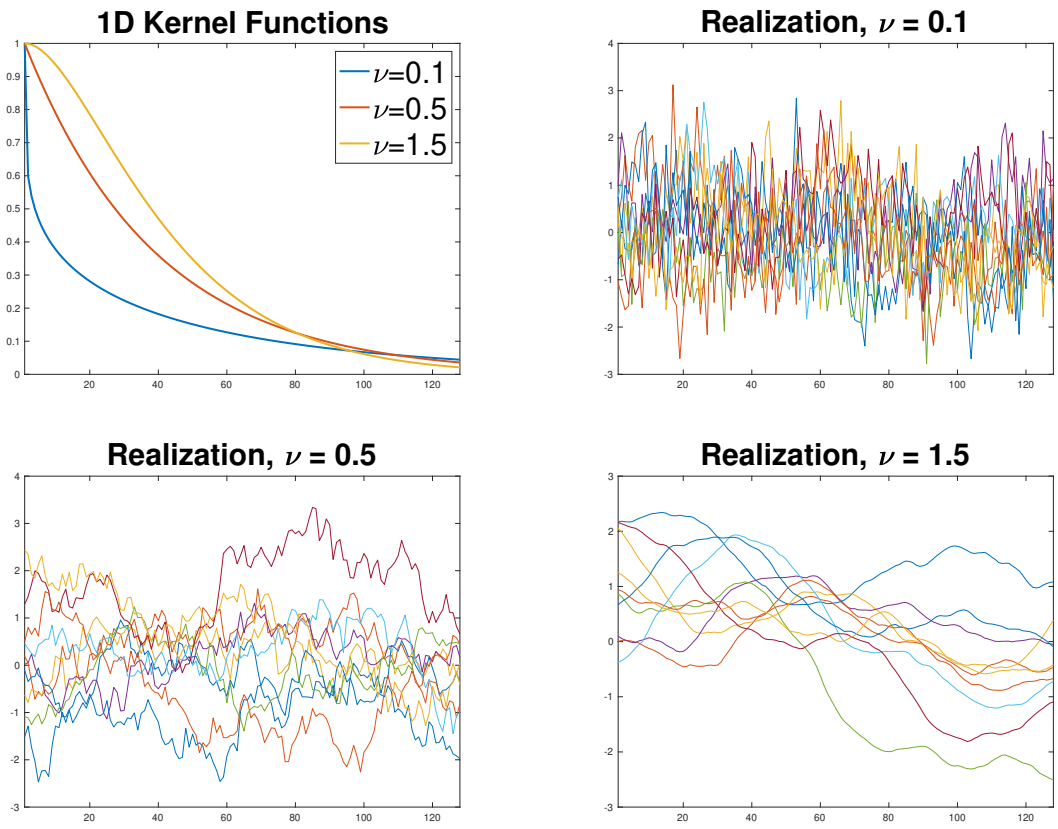


Figure 2.1: Top left shows that 1D Matérn kernel with $\nu = 0.1, 0.5$, and 1.5. The others describe the realizations of each Matérn kernel functions with 10 samples.

is sparse or structured, so working with $\mathbf{Q}^{-1}$ or its symmetric factorization has obvious computational advantages.

For example, Matérn kernels and their realization samples can be visualized in 1D (Figure 2.1) and 2D (Figure 2.2) respectively. Smaller values of $\nu$ and $\ell$ give less relationship (more independence) between neighbors in realization samples. Larger values of $\nu$ and $\ell$ increase the smoothness (strengthen relationship).

In a simple case, the covariance matrix $\mathbf{Q}$ is usually an identity matrix $\mathbf{I}$ or a symmetric positive definite matrix from which it is easy to compute $\mathbf{L}_\mathbf{Q}$ or $\mathbf{L}_\mathbf{Q}^{-1}$. In general, computing $\mathbf{Q}^{-1}$ or $\mathbf{L}_\mathbf{Q}$, is not feasible in many cases. Also, there are iterative methods that require

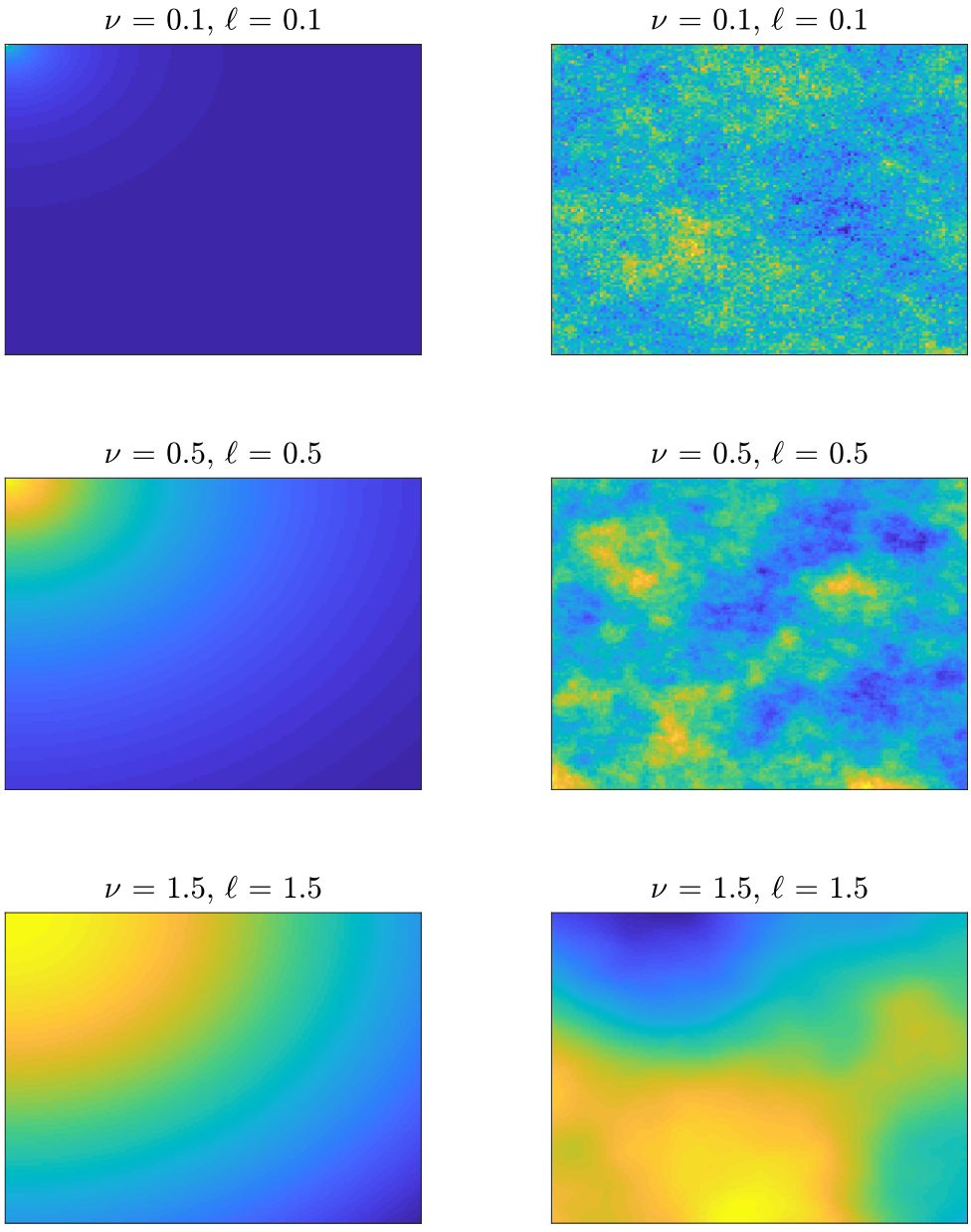


Figure 2.2: Left column shows that 2D Matérn kernel functions with $\nu, \ell = 0.1, 0.5$, and 1.5. Right column shows the realization sample of each Matérn kernel function.

matrix-vector multiplications with $\mathbf{Q}^{-1}$: a linear solve with $\mathbf{Q}$ which is expensive for large-scale problems. Using a change of variables

$$\mathbf{x} = \mathbf{Q}^{-1}(\mathbf{s} - \boldsymbol{\mu}), \quad \mathbf{b} = \mathbf{d} - \mathbf{A}\boldsymbol{\mu}, \tag{2.9}$$

(2.6) can be re-written as

$$(\mathbf{A}^\top \mathbf{R}^{-1}\mathbf{A}\mathbf{Q} + \lambda^2 \mathbf{I})\mathbf{x} = \mathbf{A}^\top \mathbf{R}^{-1}\mathbf{b}. \tag{2.10}$$

Then, the MAP estimate becomes $\mathbf{s}_\lambda = \boldsymbol{\mu} + \mathbf{Q}\mathbf{x}_\lambda$ where $\mathbf{x}_\lambda$ is solution to

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{Q}\mathbf{x} - \mathbf{b}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\|\mathbf{x}\|^2_{\mathbf{Q}}. \tag{2.11}$$

## 2.2 Generalized Hybrid Methods

We describe generalized hybrid methods, which are hybrid projection methods that build on an iterative projection method described in [3]. More details about generalized hybrid methods in the context of inverse problems can be found in [38]. For (2.11), given matrices $\mathbf{A}$, $\mathbf{Q}$, $\mathbf{R}$, and vector $\mathbf{b}$, with initial setting $\beta_1 = \|\mathbf{b}\|_{\mathbf{R}^{-1}}$, $\mathbf{u}_1 = \mathbf{b}/\beta_1$, and $\alpha_1\mathbf{v}_1 = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_1$, the $k$th iteration of the generalized Golub-Kahan (gen-GK) bidiagonalization procedure constructs vectors $\mathbf{u}_{k+1}$ and $\mathbf{v}_{k+1}$ such that

$$\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{A}\mathbf{Q}\mathbf{v}_k - \alpha_k\mathbf{u}_k,$$
$$\alpha_{k+1}\mathbf{v}_{k+1} = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k,$$

where scalars $\alpha_i, \beta_i \geq 0$ are chosen such that $\|\mathbf{u}_i\|_{\mathbf{R}^{-1}} = \|\mathbf{v}_i\|_{\mathbf{Q}} = 1$. At the $k$th iteration, we get

$$\mathbf{B}_k \equiv \begin{bmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_k \\ & & & \beta_{k+1} \end{bmatrix}, \qquad \mathbf{U}_{k+1} \equiv [\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}], \quad \text{and} \quad \mathbf{V}_k \equiv [\mathbf{v}_1, \ldots, \mathbf{v}_k],$$

where the following relations hold up to machine precision,

$$\mathbf{U}_{k+1}\beta_1\mathbf{e}_1 = \mathbf{b} \tag{2.12}$$

$$\mathbf{A}\mathbf{Q}\mathbf{V}_k = \mathbf{U}_{k+1}\mathbf{B}_k \tag{2.13}$$

$$\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{U}_{k+1} = \mathbf{V}_k\mathbf{B}_k^\top + \alpha_{k+1}\mathbf{v}_{k+1}\mathbf{e}_{k+1}^\top . \tag{2.14}$$

Furthermore, in exact arithmetic, matrices $\mathbf{U}_{k+1}$ and $\mathbf{V}_k$ satisfy the following orthogonality conditions

$$\mathbf{U}_{k+1}^\top\mathbf{R}^{-1}\mathbf{U}_{k+1} = \mathbf{I}_{k+1} \qquad \text{and} \qquad \mathbf{V}_k^\top\mathbf{Q}\mathbf{V}_k = \mathbf{I}_k. \tag{2.15}$$

Algorithm 1 describes the gen-GK bidiagonalization process. $\mathbf{A}$, $\mathbf{A}^\top$, and $\mathbf{Q}$ are not required to be constructed explicitly at each iteration since they are only involved with matrix-vector multiplications.

The classic methods to choose a regularization parameters are not practical for large-scale linear inverse problems. For example, generalized cross validation (GCV) is based on the SVD of matrix $\mathbf{A}$. To reduce the cost of choosing the regularization parameter, the gen-GK process is used. This process projects the problem onto smaller subspaces (i.e., a low-

**Algorithm 1** gen-GK process

**Require:** Matrices $\mathbf{A}$, $\mathbf{Q}$, and $\mathbf{R}$, and vector $\mathbf{b}$.

1: $\beta_1\mathbf{u}_1 = \mathbf{b}$, where $\beta_1 = \|\mathbf{b}\|_{\mathbf{R}^{-1}}$
2: $\alpha_1\mathbf{v}_1 = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_1$
3: **for** $k = 1, 2, \ldots$ **do**
4: $\quad\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{A}\mathbf{Q}_1\mathbf{v}_k - \alpha_k\mathbf{u}_k$, where $\beta_{k+1} = \|\mathbf{A}\mathbf{Q}\mathbf{v}_k - \alpha_k\mathbf{u}_k\|_{\mathbf{R}^{-1}}$
5: $\quad\alpha_{k+1}\mathbf{v}_{k+1} = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k$, where $\alpha_{k+1} = \left\|\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k\right\|_{\mathbf{Q}}$
6: **end for**

dimensional space) such as

$$\min_{\mathbf{x}_k\in\mathcal{R}(\mathbf{V}_k)} \frac{1}{2}\|\mathbf{A}\mathbf{Q}\mathbf{x}_k - \mathbf{b}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\|\mathbf{x}_k\|^2_{\mathbf{Q}} \iff \min_{\mathbf{y}_k\in\mathbb{R}^k} \frac{1}{2}\|\mathbf{B}_k\mathbf{y}_k - \beta_1\mathbf{e}_1\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{y}_k\|_2^2 \tag{2.16}$$

where $\mathbf{x}_k = \mathbf{V}_k\mathbf{y}_k$. Since $\mathbf{B}_k$ is $(k+1)\times k$, it is very easy to apply standard regularization parameter selection methods to the projected system (2.16).

From these iterative projection methods, we can reduce the problem to a smaller system where classical methods can be exploited for regularization. We will briefly review how to choose regularization parameters for a simple linear problem in the next section.

## 2.3 Choice of Regularization Parameters, $\lambda$

For simplicity, we use a simple linear model (1.1) where $\mathbf{Q} = \mathbf{I}_n$, $\mathbf{R} = \sigma^2\mathbf{I}_m$, and $\boldsymbol{\mu} = \mathbf{0}$ so that (2.16) holds as well. Consider a Tikhonov regularization problem,

$$\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s} - \mathbf{d}\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{s}\|_2^2. \tag{2.17}$$

There are various techniques to determine $\lambda$ such as the discrepancy principle (DP), the L-curve, the generalized cross validation (GCV), and the unbiased predictive risk estimation

(UPRE) [7, 59]. The closed form of solution to (2.17) is

$$\mathbf{s}_\lambda = \mathbf{A}_\lambda^\dagger \mathbf{d}$$

where $\mathbf{A}_\lambda^\dagger = (\mathbf{A}^\top \mathbf{A} + \lambda^2 \mathbf{I}_n)^{-1} \mathbf{A}^\top$ for given $\lambda > 0$. Then,

- the DP selects $\lambda$ to make the squared residual approximating $m\sigma^2$ (i.e., expectation of $\|\boldsymbol{\varepsilon}\|_2^2$).

  $$\|\mathbf{A}\mathbf{s}_\lambda - \mathbf{d}\|_2^2 \approx \tau m \sigma^2$$

  where $\tau$ is safety factor and $\tau$ is usually 1 or slightly larger than 1 [50, 59].

- the UPRE is derived from the means square error (MSE) and $\lambda$ is selected by minimizing

  $$U(\lambda) = \|\mathbf{A}\mathbf{s}_\lambda - \mathbf{b}\|_2^2 + 2\sigma^2 \mathrm{tr}(\mathbf{A}\mathbf{A}_\lambda^\dagger) - m\sigma^2 \tag{2.18}$$

  where $\mathrm{tr}(\cdot)$ is used for the trace.

- the GCV is based on a leave-one-out cross validation and GCV is highly commonly used in inverse problems. $\lambda$ is selected by minimizing the function,

  $$G(\lambda) = \frac{n\|\mathbf{A}\mathbf{s}_\lambda - \mathbf{d}\|_2^2}{[\mathrm{tr}(\mathbf{I}_m - \mathbf{A}\mathbf{A}_\lambda^\dagger)]^2} \tag{2.19}$$

Contrary to the DP and the UPRE, the GCV method does not require *a prior* estimates of the noise variance $\sigma^2$ to determine the regularization parameter $\lambda$. This is an advantage of the GCV method; however, computing the GCV regularization parameter can get costly especially for large-scale problems [69]. An alternating regularization technique is to use an iterative method (e.g., Section 2.2) to project a large-scale linear problem onto a small but growing subspace and to solve the projected problem using standard regularization

techiniques [85]. These are so-called *hybrid methods*. In [36], an implementation called HyBR combines the Golub-Kahan bidiagonalization with a weighted GCV method to solve problem (2.17). This approach is efficient and can select the regularization parameter $\lambda$ automatically. There have been many other investigations of hybrid iterative methods for $\mathbf{Q} = \mathbf{I}$ [9, 36, 37, 69, 85]. Various generalized Krylov techniques have been developed to handle the general form of the Tikhonov problem where $\mathbf{Q} \neq \mathbf{I}_n$ [65, 94], and various works have explored hybrid methods that can efficiently handle Gaussian priors by working directly with $\mathbf{Q}$ [38] or by working with mixed Gaussian priors [29].

**Theorem 2.1.** *Fix $\lambda \geq 0$. Let $\mathbf{y}_k$ be the exact solution to gen-LSQR to (2.16). Then the kth iterate, $\mathbf{s}_k = \boldsymbol{\mu} + \mathbf{Q}\mathbf{V}_k\mathbf{y}_k$, is equivalent to $\boldsymbol{\mu} + \mathbf{L}_\mathbf{Q}^{-1}\mathbf{w}_k$ where $\mathbf{w}_k$ is the kth iterate of LSQR on the following Tikhonov problems,*

$$\min_{\mathbf{w}} \left\| \begin{bmatrix} \mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{L}_\mathbf{Q}^{-1} \\ \lambda\mathbf{I} \end{bmatrix} \mathbf{w} - \begin{bmatrix} \mathbf{L}_\mathbf{R}\mathbf{b} \\ \mathbf{0} \end{bmatrix} \right\|_2^2. \tag{2.20}$$

*Proof.* See the proof of Theorem 4.2 in [38]. □

The convergence of iterative solution with a fixed parameter is described in Chapter 3 for the mixed Golub-Kahan process.

## 2.4 Efficient Methods for Uncertainty Quantification

In Bayesian inference (2.2), assuming that likelihood and prior distributions are Gaussian produce a posterior distribution that is also Gaussian. Thus, these assumptions do not require us to sample from the posterior density function. In (2.8), the posterior mean can be computed or approximated by solving MAP. However, calculating exact the posterior

covariance matrix is challenging problem. To obtain the posterior variance component-wise, the diagonal entries of the posterior covariance matrix are required. However, since the size of $\mathbf{Q}$ follows the size of the unknowns, computing explicit $\mathbf{Q}^{-1}$ and $\mathbf{\Gamma}_{\text{post}}$ in large-scale inverse problem is not possible. Instead, many approaches use low-rank approximations to estimate the diagonal entries of $\mathbf{\Gamma}_{\text{post}}$ [18, 19, 48, 97, 98, 103]. The references combines low-rank approximation of $\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A}$ with the Woodbury matrix identity. In particular, [97, 98] does not require a decomposition of $\mathbf{Q}$.

In the recent work [97], we can estimate $\mathbf{\Gamma}_{\text{post}}$ of (2.8) very efficiently by exploiting matrices generated from gen-GK bidiagonalization and relationships (2.13), (2.14), and (2.15).

At the $k$th iteration, let $\mathbf{B}_k^\top\mathbf{B}_k = \mathbf{W}_k\mathbf{\Theta}_k\mathbf{W}_k^\top$ be the eigenvalue decomposition with eigenvalues $\theta_1, \ldots, \theta_k$ and let $\mathbf{Z}_k = \mathbf{Q}\mathbf{V}_k\mathbf{W}_k$, then we can get the following low-rank approximation

$$\mathbf{Q}(\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A})\mathbf{Q} \approx \mathbf{Q}(\mathbf{V}_k\mathbf{B}_k^\top\mathbf{B}_k\mathbf{V}_k^\top)\mathbf{Q} = \mathbf{Z}_k\mathbf{\Theta}_k\mathbf{Z}_k^\top. \tag{2.21}$$

Thus, we obtain the following approximation of $\mathbf{\Gamma}_{\text{post}}$ with (2.21) and the Woodbury formula,

$$\begin{aligned}
\mathbf{Q}_{\text{post}} &= \left(\lambda^2\mathbf{Q}^{-1} + \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A}\right)^{-1} \\
&= \mathbf{Q}(\lambda^2\mathbf{Q} + \mathbf{Q}\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A}\mathbf{Q})^{-1}\mathbf{Q} \\
&\approx \mathbf{Q}(\lambda^2\mathbf{Q} + \mathbf{Z}_k\mathbf{\Theta}_k\mathbf{Z}_k^\top)^{-1}\mathbf{Q} \\
&= \mathbf{Q}(\lambda^{-2}\mathbf{Q}^{-1} - \lambda^{-2}\mathbf{Q}^{-1}\mathbf{Z}_k(\lambda^2\mathbf{\Theta}_k^{-1} + \mathbf{Z}_k^\top\mathbf{Q}^{-1}\mathbf{Z}_k)^{-1}\mathbf{Z}_k^\top\mathbf{Q}^{-1})\mathbf{Q} \\
&= \lambda^{-2}\mathbf{Q} - \lambda^{-2}\mathbf{Z}_k(\lambda^2\mathbf{\Theta}_k^{-1} + \mathbf{Z}_k^\top\mathbf{Q}^{-1}\mathbf{Z}_k)^{-1}\mathbf{Z}_k^\top \\
&= \lambda^{-2}\mathbf{Q} - \mathbf{Z}_k\mathbf{\Delta}_k\mathbf{Z}_k^\top
\end{aligned}$$

where

$$\boldsymbol{\Delta}_k \equiv \lambda^{-2} \begin{bmatrix} \frac{\theta_1}{\theta_1+\lambda^2} & & \\ & \ddots & \\ & & \frac{\theta_k}{\theta_k+\lambda^2} \end{bmatrix} \in \mathbb{R}^{k\times k}$$

If $\mathbf{Q}$ is a structured matrix (e.g., Toeplitz), we can easily obtain the diagonal entries of $\mathbf{Q}$ and thus $\mathbf{Q}_{\text{post}}$. More theoretical results about the accuracy of the approximate posterior covariance matrix are described in [97].

# Chapter 3

# Hybrid projection Methods for Large-scale Inverse Problems with Mixed Gaussian Priors

As described in Section 1.2, using mixed Gaussian priors, where the prior covariance matrix can be represented as a convex combination of matrices, is a common approach to incorporate different prior covariance matrices [5]. However, various computational challenges arise for problems where the number of unknowns is very large and the regularization and mixing parameter are not known in advance. We address these challenges by developing hybrid iterative projection methods for the efficient computation of solutions to inverse problems with mixed Gaussian priors. By exploiting a project-then-regularize framework, we enable statistical optimization tools for selecting the regularization parameter and the mixing parameter automatically, which would be very costly for the original problem.

We are interested in linear inverse problems of the form (1.1). Regularization is required to stabilize the inversion process due to ill-posedness. We follow a Bayesian framework, where we assume a prior for $\mathbf{s}$. That is, we treat $\mathbf{s}$ as a Gaussian random variable as (2.1).

In many applications, the choice of $\mathbf{Q}$ is pre-determined (e.g., using expert knowledge) and is chosen to enforce smoothness or regularity conditions on the solution [17, 38, 65].

However, in some cases, there is not enough information to determine $\mathbf{Q}$ completely or expensive procedures are needed to determine an informative subset of covariates from a set of candidates (e.g., in geophysical imaging [112, 115, 116]). These scenarios motivate us to consider mixed Gaussian priors, where the covariance matrix can be represented as a convex combination of matrices. Without loss of generality we consider prior covariance matrices of the form,

$$\mathbf{Q} = \gamma \mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2 \tag{3.1}$$

where $\mathbf{Q}_1$ is a symmetric positive definite matrix, $\mathbf{Q}_2$ is a symmetric positive semi-definite matrix, and mixing parameter $0 < \gamma \leq 1$. We consider the case where computing matrix-vector products with $\mathbf{Q}_1$ is easy, but accessing $\mathbf{Q}_1^{-1}$ or its symmetric factorization (e.g., Cholesky or eigenvalue factorization) is not feasible. Such scenarios arise, for example, when the prior covariance matrix is modeled entry-wise using covariance kernels. In such cases, the main challenge is that the resulting covariance matrices are large and dense, and factorizing or inverting them can be computationally prohibitive. However, matrix-vector multiplications can be done efficiently (e.g., via FFT embedding). A wide range of kernels, including nonseparable spatio-temporal kernels [38], can be included. We assume that matrix-vector products with $\mathbf{Q}_2$ can be done efficiently.

Covariance matrices of the form (3.1) are becoming more common, especially in modern imaging applications where data (e.g., in the form of training images) are playing a larger role in the development of reconstruction algorithms [4]. Suppose we are given a dataset consisting of $N$ samples, $\mathbf{s}^{(i)} \in \mathbb{R}^n, i = 1, 2, \ldots, N$. Then the training data can be used to obtain an unbiased estimator of an $n \times n$ sample covariance matrix,

$$\widehat{\mathbf{Q}} = \frac{1}{N}\sum_{i=1}^{N}(\mathbf{s}^{(i)} - \bar{\mathbf{s}})(\mathbf{s}^{(i)} - \bar{\mathbf{s}})^\top, \tag{3.2}$$

where $\bar{\mathbf{s}} = \frac{1}{N}\sum_{i=1}^{N} \mathbf{s}^{(i)}$ is the sample mean. Notice that $\widehat{\mathbf{Q}} = \mathbf{S}\mathbf{S}^\top$, where the symmetric factor is defined as $\mathbf{S} = \frac{1}{\sqrt{N}}\left(\begin{bmatrix}\mathbf{s}^{(1)} & \dots & \mathbf{s}^{(N)}\end{bmatrix} - \bar{\mathbf{s}} \otimes \mathbf{1}^\top\right)$ with $\mathbf{1} \in \mathbb{R}^N$ denoting the vector whose elements are all 1. For any vector $\mathbf{x} \in \mathbb{R}^n$, multiplication with $\widehat{\mathbf{Q}}$ can be done efficiently if $N << n$, e.g., using the following order of operations $\mathbf{S}(\mathbf{S}^\top\mathbf{x})$. However, notice that $\widehat{\mathbf{Q}}$ is likely positive semi-definite rather than positive definite, so it is common to use $\widehat{\mathbf{Q}} + \gamma\mathbf{I}$ where $\gamma$ is a nudging term. Such approaches are known as sample based priors [21]. Another common approach is to use a convex combination, i.e., the prior covariance matrix is given as

$$\mathbf{Q} = \gamma\mathbf{D} + (1-\gamma)\widehat{\mathbf{Q}} \tag{3.3}$$

where $\mathbf{D}$ is chosen to be the identity matrix or a suitably chosen diagonal or correlation matrix, which ensures that $\mathbf{Q}$ is positive definite, and $\gamma \in \mathbb{R}$ is called the mixing parameter. The matrix in (3.3) is called a shrinkage estimator of the covariance matrix [99]. It is worth noting that covariance matrices of the form (3.3) are also used in hybrid methods for data assimilation that combine an ensemble Kalman filter system with a variational (e.g., 3D-Var) system [5]. These methods require careful tuning of the so-called blending parameter $\gamma$, and many of the existing approaches require $\gamma$ to be fixed in advance. We do not assume this.

Previous works on combining training data with regularization techniques typically follow an optimal experimental design or empirical Bayes risk minimization framework [33, 56]. More recently, there has been significant work on using training data in the context of machine learning to learn regularization functionals (e.g., [72, 100]) or to learn the "invisible" regions (e.g., [16]). The area of data-driven machine learning is currently a hot topic [4, 76], where the main goal is to determine new ways to combine physical models with deep learning techniques. In this work, we incorporate training data in a Bayesian framework and exploit tools from numerical linear algebra not only to compute solutions efficiently but also to determine the appropriate weighting of the training data.

In this chapter we develop a hybrid iterative projection method that is based on a mixed, generalized Golub-Kahan process to approximate the MAP estimate (2.5), where $\mathbf{Q}$ is of the form (3.1). Our approach can handle a wide range of scenarios, including data-informed regularization terms that use training or test images to define the prior. We assume that $\gamma$ is not known in advance and neither the inverse nor the factorization of $\mathbf{Q}$ is available. The proposed method has two distinctive features. First, we assume that *both* $\gamma$ and $\lambda$ are unknown a priori and we estimate them during the solution process. For large-scale problems where $\gamma$ is fixed and hence $\mathbf{Q}$ is fixed in advance, generalized hybrid methods [38] have been developed where $\lambda$ can be estimated during the reconstruction process. On the other hand, previous works have used linear combinations of "simple" covriance matrices and statistical methods to estimate mixing parameter $\gamma$, but the choice of regularization parameter $\lambda$ is fixed. To the best of our knowledge, both problems have not been address simultaenously, and developing a hybrid method where $\lambda$ and $\gamma$ can be selected adaptively is not an obvious extension of existing methods. We develop an iterative hybrid approach where the problem is projected onto generalized Krylov subspaces of small but increasing dimension and the regularization parameter and mixing parameter can be simultaneously and automatically selected. Second, we describe and investigate various scenarios where training data can be used to define $\mathbf{Q}_1$ and $\mathbf{Q}_2$, so our approach can be considered a learning approach for the regularization term. There has been a large emphasis on using training data in the development of reconstruction techniques for modern imaging problems, but the ability to balance training data with physical priors remains an open problem. A key feature of the proposed hybrid methods is the ability to incorporate data-driven covariance matrices while simultaneously balancing existing priors. In our numerical experiments, we have verified that such methods can perform better than classical shrinkage algorithms, especially when the data contain freckles.

An outline for the section is as follows. In Section 3.1 we provide some background on Gaussian priors and focus on various data-driven prior covariance matrices. Then in Section 3.2, we describe mixed, generalized hybrid projection methods for approximating the MAP estimate (2.5), where $\mathbf{Q}$ is of the form (3.1). The approach consists of two-steps: (1) Project the problem onto a subspace of small but increasing dimension using an extension of the generalized Golub-Kahan bidiagonalization approach. (2) Solve the projected problem where the regularization parameter $\lambda$ and mixing parameter $\gamma$ can be selected automatically. Various regularization paremeter selection techniques will be investigated, and some theoretical results will be provided. In Section 3.3 numerical results on various image processing applications show the potential benefits and flexibility of these methods.

## 3.1 Mixed Gaussian priors

In this section, we motivate the need for mixed Gaussian priors and draw some connections to existing works on multi-parameter Tikhonov regularization and shrinkage estimation.

In many applications, the precision matrix is not readily available, and the aim is to develop computational methods that can work with $\mathbf{Q}$ directly and avoid the need for the inverse or symmetric factorization. Such covariance kernels may arise in dynamic scenarios with nonseparable, spatio-temporal priors [39, 49, 74] or from spatially-variant priors [44, 114]. It is worth mentioning that in a truly Bayesian framework, the regularization parameter and the covariance kernel parameters could be included as hyperparameters and explored using MCMC methods [7], but the computational costs of this approach would be very high.

One reason to use Gaussian mixtures as prior distributions is that it allows greater flexibility in the definition of the prior. In this chapter, we consider a mixture of two Gaussians, but one could consider more general mixtures. From a statistical viewpoint, a general formulation

with $N$ Gaussian random vectors would correspond to a sum of covariance matrices. That is, let $\mathbf{x}_1, ..., \mathbf{x}_N$ be $N$ mutually independent $n \times 1$ normal random vectors having means $\boldsymbol{\mu}_1, ...\boldsymbol{\mu}_N$ and covariance matrices $\mathbf{V}_1, ...\mathbf{V}_N$. Let $\mathbf{B}_1, ...\mathbf{B}_N$ be real $L \times n$ full rank matrices. Then the $L \times 1$ random vector

$$\mathbf{y} = \sum_{i=1}^{N} \mathbf{B}_i \mathbf{x}_i \tag{3.4}$$

has a normal distribution with mean $\mathbb{E}\mathbf{y} = \sum_{i=1}^{N} \mathbf{B}_i \boldsymbol{\mu}_i$ and covariance matrix of the form $Cov(\mathbf{y}) = \sum_{i=1}^{N} \mathbf{B}_i \mathbf{V}_i \mathbf{B}_i^\top$. Thus, a Gaussian mixture prior corresponds to an assumption that the desired solution can be represented as a linear combination of Gaussian realizations (e.g., with different smoothness properties).

In the context of inverse problems, we point out a connection between mixed Gaussian priors and multi-parameter Tikhonov regularization. The basic idea of multi-parameter Tikhonov regularization, see e.g. [10, 51, 75, 107], is to solve a problem of the form,

$$\min_{\mathbf{s}} \|\mathbf{A}\mathbf{s} - \mathbf{d}\|^2_{\mathbf{R}^{-1}} + \sum_{i=1}^{N} \lambda_i^2 \|\mathbf{L}_i \mathbf{s}\|_2^2 , \tag{3.5}$$

where $\lambda_i \in \mathbb{R}$ is the regularization parameter corresponding to regularization matrix $\mathbf{L}_i$ for $i = 1, \ldots, N$. By including multiple penalty terms, this approach can enforce different smoothness properties (e.g, at different frequency bands) and avoid difficulties in having to select just one regularization matrix. In a Bayesian framework, the multi-parameter Tikhonov solution can be interpreted as a MAP estimate, under the assumption of a Gaussian prior with mean $\mathbf{0}$ and covariance matrix $\left(\sum_{i=1}^{N} \lambda_i^2 \mathbf{L}_i^\top \mathbf{L}_i\right)^{-1}$. Notice that except for in very limited scenarios, this is not the same as using mixed Gaussian priors, since here the precision matrix (not the covariance matrix) is represented as a sum of matrices.

### 3.1.1 Data-driven prior covariance matrices

With the increasing data resolution and data dimension in many applications, an important and challenging task is to determine how to efficiently and effectively incorporate prior knowledge in the form of training data both in the solution computation process and the subsequent data analyses. In this section, we describe various examples where training data can be used to define the prior covariance matrix. For all cases, we assume that training data is provided and the sample covariance matrix (3.2) has the form $\widehat{\mathbf{Q}} = \mathbf{S}\mathbf{S}^\top$.

As described in the introduction, the most common approach is to take $\mathbf{Q}_2 = \widehat{\mathbf{Q}}$ and $\mathbf{Q}_1 = \mathbf{D}$ where $\mathbf{D}$ is easy to invert (e.g., diagonal or identity matrix). In this case, a very popular approach called shrinkage estimation of covariance matrices, or more general biased estimation, can be used to reduce the variance of the estimator. Typical shrinkage targets are diagonal matrices (e.g., including the identity matrix), and approaches to estimate the optimal shrinkage intensity $\gamma$ have been proposed by Ledoit and Wolf, Rao and Blackwell, and others [5, 28, 71, 99].

Another approach to incorporate training data is to force some structure or functional form on the prior covariance kernel function as described in Table 2.1. For kernel functions that depend on a few parameters, the training data can be used to estimate these parameters. A similar idea was considered in [56] where training data was used to learn parameters defining the regularization functional. However, that approach requires solving an expensive constrained optimization problem, and the learned regularization functional is tailored to the forward operator and the noise level. We consider the case where the training data come from a prior defined by a covariance kernel function (e.g., for simplicity, we consider Matérn kernels). We use the training data to learn the parameters defining the prior. This reduces to an optimization problem where the goal is to learn two parameters $\nu$ and $\ell$ from the

training data. Consider the optimization problem,

$$\min_{\nu>0,\ell>0} \left\| \mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}} \right\|_F^2. \tag{3.6}$$

Let $\hat{\nu}, \hat{\ell}$ denote the parameters, which can be used to define $\mathbf{Q}_1 = \mathbf{Q}(\hat{\nu}, \hat{\ell})$. This matrix can be used directly in generalized hybrid methods, or can be combined with the sample covariance matrix, i.e., $\mathbf{Q}$ as in (3.1) with $\mathbf{Q}_1 = \mathbf{Q}(\hat{\nu}, \hat{\ell})$ and $\mathbf{Q}_2 = \widehat{\mathbf{Q}}$, and solvers described in Section 3.2 can be used.

Next, we describe some computationally efficient methods to estimate $\hat{\nu}$ and $\hat{\ell}$. Notice that

$$\|\mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}}\|_F^2 = \operatorname{tr}(\mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}})^\top (\mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}}) \tag{3.7}$$

$$= \mathbb{E}(\|(\mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}})\boldsymbol{\xi}\|_2^2) \tag{3.8}$$

where $\boldsymbol{\xi}$ is a random variable such that $\mathbb{E}\boldsymbol{\xi} = \mathbf{0}$ and $\mathbb{E}(\boldsymbol{\xi}\boldsymbol{\xi}^\top) = \mathbf{I}$. Although stochastic optimization methods [101] could be use here, we follow an approximation approach where we use a Hutchinson trace estimator. That is, we let $\boldsymbol{\xi}^{(i)} \in \mathbb{R}^n$ for $i = 1, 2, \ldots, M$ be realizations of a Rademacher distribution (i.e., $\boldsymbol{\xi}$ consists of $\pm 1$ with equal probability), and we consider the approximate optimization problem,

$$\min_{\nu>0,\ell>0} \frac{1}{M} \sum_{i=1}^{M} \|(\mathbf{Q}(\nu,\ell) - \widehat{\mathbf{Q}})\boldsymbol{\xi}^{(i)}\|_2^2. \tag{3.9}$$

We used an inteiror-point method (`fmincon.m` in MATLAB) to minimize (3.9). We mention that for problems without training data, semivariogram hyperparameters were investigated in [15] to estimate Matérn parameters from the data.

## 3.2 Hybrid projection methods for mixed Gaussian priors

In this section, we describe a hybrid projection method to approximate the MAP estimate (2.5). The distinguishing factor of this approach compared to generalized Golub-Kahan (genGK) hybrid methods [38] is that we address problems where the prior covariance matrix is of the form (3.1). That is, we consider priors of the form $\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}, \lambda^{-2}(\gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2))$, and exploit a hybrid projection framework to enable tools for selecting both the regularization parameter $\lambda$ and the mixing parameter $\gamma$ simultaneously.

Using the following change of variables (2.9), we see that solving (2.5) is equivalent to solving (2.11). If $\gamma$ is known in advance, we can directly apply the genGK hybrid method and estimate $\lambda$ automatically [38]. However, if $\gamma$ is not known in advance, significant computations would be required for a different chocie of $\gamma$. For this, we develop a variant of the genGK bidiagonlization which we call a *mixed* Golub-Kahan (mixGK) process, where both $\gamma$ and $\lambda$ can be estimated during the iterative process. Each iteration of the mixGK process requires two steps. The first step is to run one iteration of the genGK bidiagonalization process with $\mathbf{Q}_1$. The second step incorporates $\mathbf{Q}_2$ so that the regularized problem can be iteratively projected onto a smaller subspace, and $\gamma$ and $\lambda$ can *both* be selected automatically. Next we describe the mixGK process in detail.

Given matrices $\mathbf{A}$, $\mathbf{R}$, $\mathbf{Q}_1$, and vector $\mathbf{b}$, the genGK bidiagonalization process generates matrices $\mathbf{B}_k$, $\mathbf{U}_{k+1}$, and $\mathbf{V}_k$ satisfying (2.12), (2.13), (2.14), and (2.15) for $\mathbf{Q}_1$. If we let $\widetilde{\mathbf{U}}_{k+1} = \mathbf{L}_\mathbf{R}\mathbf{U}_{k+1}$ where $\mathbf{R}^{-1} = \mathbf{L}_\mathbf{R}^\top\mathbf{L}_\mathbf{R}$, then $\widetilde{\mathbf{U}}_{k+1}^\top\widetilde{\mathbf{U}}_{k+1} = \mathbf{I}_{k+1}$.

Next, in order to incorporate $\mathbf{Q}_2$, we additionally compute $m \times k$ matrix $\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k$. Assuming that the columns of $\widetilde{\mathbf{U}}_{k+1}$ and $\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k$ are linearly independent, we can compute

the skinny QR factorization, $(\mathbf{I}-\widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k = \mathbf{Y}_k\mathbf{R}_k$ where $\mathbf{Y}_k \in \mathbb{R}^{m\times k}$ contains orthonormal columns and $\mathbf{R}_k \in \mathbb{R}^{k\times k}$ is upper triangular. Notice that since column vectors in $\mathbf{Y}_k$ and $\widetilde{\mathbf{U}}_{k+1}$ are orthogonal, we get the skinny QR factorization,

$$\begin{bmatrix}\widetilde{\mathbf{U}}_{k+1} & \mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\end{bmatrix} = \begin{bmatrix}\widetilde{\mathbf{U}}_{k+1} & \mathbf{Y}_k\end{bmatrix}\begin{bmatrix}\mathbf{I}_{k+1} & \widetilde{\mathbf{U}}_{k+1}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k \\ \mathbf{0} & \mathbf{R}_k\end{bmatrix}. \tag{3.10}$$

The mixGK process is summarized in Algorithm 2.

**Algorithm 2** mixed Golub-Kahan (mixGK) process

**Require:** Matrices $\mathbf{A}$, $\mathbf{R}$, $\mathbf{Q}_1$ and $\mathbf{Q}_2$, and vector $\mathbf{b}$.

1: $\beta_1\mathbf{u}_1 = \mathbf{b}$, where $\beta_1 = \|\mathbf{b}\|_{\mathbf{R}^{-1}}$
2: $\alpha_1\mathbf{v}_1 = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_1$
3: **for** $k = 1, 2, \ldots$ **do**
4: $\quad\beta_{k+1}\mathbf{u}_{k+1} = \mathbf{A}\mathbf{Q}_1\mathbf{v}_k - \alpha_k\mathbf{u}_k$, where $\beta_{k+1} = \|\mathbf{A}\mathbf{Q}_1\mathbf{v}_k - \alpha_k\mathbf{u}_k\|_{\mathbf{R}^{-1}}$
5: $\quad\alpha_{k+1}\mathbf{v}_{k+1} = \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k$, where $\alpha_{k+1} = \left\|\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{u}_{k+1} - \beta_{k+1}\mathbf{v}_k\right\|_{\mathbf{Q}_1}$
6: $\quad[\mathbf{Y}_k, \mathbf{R}_k] = qr((\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{V}_k, 0)$;
7: **end for**

Notice that in addition to the computational cost of the genGK bidiagonalization, which includes one matrix-vector product with $\mathbf{A}$, one with $\mathbf{A}^\top$, two with $\mathbf{Q}_1$, and two solves with $\mathbf{R}$, each iteration of the mixGK process requires one matrix-vector product with $\mathbf{Q}_2$ and a QR factorization in step 6. Instead of performing a standard QR factorization on an $m$-by-$k$ matrix, an efficient rank-one update strategy can be used to alleviate the computational cost. More specifically, we will describe it using mathematical induction. Let

$$(\mathbf{I} - \widetilde{\mathbf{U}}_k\widetilde{\mathbf{U}}_k^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{V}_{k-1} = \mathbf{Y}_{k-1}\mathbf{R}_{k-1} \tag{3.11}$$

be the skinny QR factorization, where $\mathbf{Y}_{k-1}^\top\mathbf{Y}_{k-1} = \mathbf{I}_{k-1}$ and $\mathbf{R}_{k-1}$ is an upper triangular

matrix. Define $\widetilde{\mathbf{U}}_{k+1} = \begin{bmatrix} \widetilde{\mathbf{U}}_k & \widetilde{\mathbf{u}}_{k+1} \end{bmatrix}$ and $\mathbf{V}_k = \begin{bmatrix} \mathbf{V}_{k-1} & \mathbf{v}_k \end{bmatrix}$. Then by (3.11), we have

$$\begin{aligned}
(\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{V}_k &= \begin{bmatrix} (\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{V}_{k-1} & (\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{v}_k \end{bmatrix} \\
&= \begin{bmatrix} (\mathbf{I} - \widetilde{\mathbf{U}}_k\widetilde{\mathbf{U}}_k^\top - \widetilde{\mathbf{u}}_{k+1}\widetilde{\mathbf{u}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{V}_{k-1} & (\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{v}_k \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{Y}_{k-1}\mathbf{R}_{k-1} - \widetilde{\mathbf{u}}_{k+1}\widetilde{\mathbf{u}}_{k+1}^\top\mathbf{Y}_{k-1}\mathbf{R}_{k-1} & (\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{v}_k \end{bmatrix}.
\end{aligned}$$

Since the first matrix is a rank-one update of a QR factorization, its QR factorization can be obtained in $\mathcal{O}(mk)$ operations [42]. That is, we have

$$\mathbf{Y}_{k-1}\mathbf{R}_{k-1} - \widetilde{\mathbf{u}}_{k+1}(\mathbf{R}_{k-1}^\top\mathbf{Y}_{k-1}^\top\widetilde{\mathbf{u}}_{k+1})^\top = \widehat{\mathbf{Y}}_{k-1}\widehat{\mathbf{R}}_{k-1}$$

where $\widehat{\mathbf{Y}}_{k-1}^\top\widehat{\mathbf{Y}}_{k-1} = \mathbf{I}_{k-1}$ and $\widehat{\mathbf{R}}_{k-1}$ is an upper triangular matrix. Finally, let $\widehat{\mathbf{v}}_k = (\mathbf{I} - \widetilde{\mathbf{U}}_{k+1}\widetilde{\mathbf{U}}_{k+1}^\top)\mathbf{L}_R\mathbf{A}\mathbf{Q}_2\mathbf{v}_k$, then one step of the Gram-Schmidt process gives the desired QR factorization,

$$\begin{bmatrix} \widehat{\mathbf{Y}}_{k-1}\widehat{\mathbf{R}}_{k-1} & \widehat{\mathbf{v}}_k \end{bmatrix} = \mathbf{Y}_k\mathbf{R}_k.$$

### 3.2.1 Solving the projected problem

Using the mixGK process described above, we now describe a hybrid iterative projection method to solve (2.11). In particular, we consider the projected problem,

$$\min_{\mathbf{x}\in\mathcal{R}(\mathbf{V}_k)} \frac{1}{2}\|\mathbf{A}\mathbf{Q}\mathbf{x} - \mathbf{b}\|_{\mathbf{R}^{-1}}^2 + \frac{\lambda^2}{2}\|\mathbf{x}\|_{\mathbf{Q}}^2 \tag{3.12}$$

where $\mathcal{R}(\cdot)$ denotes the column space. Let $\mathbf{x} = \mathbf{V}_k\mathbf{y}$ where $\mathbf{y} \in \mathbb{R}^k$. Then using the relationships from the mixGK process, we obtain the equivalent problems,

$$\min_{\mathbf{y}} \frac{1}{2} \|\gamma\mathbf{A}\mathbf{Q}_1\mathbf{V}_k\mathbf{y} + (1-\gamma)\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\mathbf{y} - \mathbf{b}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\mathbf{y}^\top\mathbf{V}_k^\top(\gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2)\mathbf{V}_k\mathbf{y} \tag{3.13}$$

$$\min_{\mathbf{y}} \frac{1}{2} \left\|\gamma\widetilde{\mathbf{U}}_{k+1}\mathbf{B}_k\mathbf{y} + (1-\gamma)\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\mathbf{y} - \mathbf{L}_\mathbf{R}\mathbf{b}\right\|_2^2 + \frac{\lambda^2\gamma}{2}\mathbf{y}^\top\mathbf{y} + \frac{\lambda^2(1-\gamma)}{2}\mathbf{y}^\top\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\mathbf{y} \tag{3.14}$$

$$\min_{\mathbf{y}} \frac{1}{2} \left\| \begin{bmatrix} \widetilde{\mathbf{U}}_{k+1} & \mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k \end{bmatrix} \begin{bmatrix} \gamma\mathbf{B}_k \\ (1-\gamma)\mathbf{I}_k \end{bmatrix} \mathbf{y} - \mathbf{L}_\mathbf{R}\mathbf{b} \right\|_2^2 + \frac{\lambda^2\gamma}{2}\|\mathbf{y}\|_2^2 + \frac{\lambda^2(1-\gamma)}{2}\mathbf{y}^\top\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\mathbf{y}. \tag{3.15}$$

Using equation (3.10) and the fact that

$$\begin{bmatrix} \widetilde{\mathbf{U}}_{k+1} & \mathbf{Y}_k \end{bmatrix} \begin{bmatrix} \beta_1\mathbf{e}_1 \\ \mathbf{0} \end{bmatrix} = \widetilde{\mathbf{U}}_{k+1}(\beta_1\mathbf{e}_1) = \mathbf{L}_\mathbf{R}\mathbf{b} \tag{3.16}$$

where $\begin{bmatrix} \widetilde{\mathbf{U}}_{k+1} & \mathbf{Y}_k \end{bmatrix}$ contains orthonormal columns (so it can be taken out of the norm), the projected, regularized problem becomes

$$\min_{\mathbf{y}} \frac{1}{2} \left\| \begin{bmatrix} \mathbf{I}_{k+1} & \widetilde{\mathbf{U}}_{k+1}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k \\ \mathbf{0} & \mathbf{R}_k \end{bmatrix} \begin{bmatrix} \gamma\mathbf{B}_k \\ (1-\gamma)\mathbf{I}_k \end{bmatrix} \mathbf{y} - \begin{bmatrix} \beta_1\mathbf{e}_1 \\ \mathbf{0} \end{bmatrix} \right\|_2^2 + \frac{\lambda^2\gamma}{2}\|\mathbf{y}\|_2^2 + \frac{\lambda^2(1-\gamma)}{2}\mathbf{y}^\top\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\mathbf{y}. \tag{3.17}$$

Note that the solution subspace for $\mathbf{x}$ does not depend on $\gamma$ and $\lambda$, but the solution of the projection problem depends on both $\gamma$ and $\lambda$. Let $\mathbf{y}_k(\lambda, \gamma)$ denote the solution to (3.17), then the $k$ iterate of the mixGK method is given as

$$\mathbf{s}_k(\lambda, \gamma) = \boldsymbol{\mu} + (\gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2)\mathbf{V}_k\mathbf{y}_k(\lambda, \gamma). \tag{3.18}$$

In Section 3.2.2 we describe some techniques for selecting $\lambda$ and $\gamma$ at each iteration, but first we provide a theoretical result. We show that for fixed regularization parameter $\lambda$ and fixed mixing parameter $\gamma$, the proposed mixGK method converges in exact arithmetic to the desired regularized solution.

**Theorem 3.1.** *Assume $\lambda > 0$ and $0 < \gamma \leq 1$. Let $\mathbf{y}_k(\lambda, \gamma)$ be the exact solution to projected problem (3.17). Then the kth iterate of the mixGK approach, written as*

$$\mathbf{s}_k = \boldsymbol{\mu} + \mathbf{Q}\mathbf{V}_k\mathbf{y}_k(\lambda, \gamma) \tag{3.19}$$

*converges to the MAP estimate given by*

$$\mathbf{s}_{\text{MAP}} = \boldsymbol{\mu} + \mathbf{Q}(\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A}\mathbf{Q} + \lambda^2\mathbf{I}_n)^{-1}\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{b}. \tag{3.20}$$

*Proof.* The proof is provided in Appendix A.1. □

### 3.2.2 Regularization parameter selection methods

In this section, we describe two extensions of existing regularization parameter selection methods that can be used for selecting $\gamma$ and $\lambda$ at each iteration of the mixGK hybrid method. For all theoretical results, we assume no breakdown of the algorithms. Notice that the solution at the $k$-th iteration can be written as

$$\mathbf{s}_k(\lambda, \gamma) = \boldsymbol{\mu} + (\gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2)\mathbf{V}_k\mathbf{y}_k(\lambda, \gamma), \tag{3.21}$$

where

$$\begin{aligned}\mathbf{y}_k(\lambda,\gamma) &= \left(\mathbf{D}_k(\gamma)^\top\mathbf{D}_k(\gamma)+\lambda^2\gamma\mathbf{I}_k+\lambda^2(1-\gamma)\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\right)^{-1}\mathbf{D}_k(\gamma)^\top\begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix}\\ &= \mathbf{C}_k(\gamma,\lambda)\begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix}\end{aligned} \tag{3.22}$$

with

$$\mathbf{D}_k(\gamma)=\begin{bmatrix}\mathbf{I}_{k+1} & \widetilde{\mathbf{U}}_{k+1}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\\ \mathbf{0} & \mathbf{R}_k\end{bmatrix}\begin{bmatrix}\gamma\mathbf{B}_k\\ (1-\gamma)\mathbf{I}_k\end{bmatrix}=\begin{bmatrix}\gamma\mathbf{B}_k+(1-\gamma)\widetilde{\mathbf{U}}_{k+1}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\\ (1-\gamma)\mathbf{R}_k\end{bmatrix} \tag{3.23}$$

$$\mathbf{C}_k(\gamma,\lambda)=\left(\mathbf{D}_k(\gamma)^\top\mathbf{D}_k(\gamma)+\lambda^2\gamma\mathbf{I}_k+\lambda^2(1-\gamma)\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\right)^{-1}\mathbf{D}_k(\gamma)^\top. \tag{3.24}$$

As with regularization parameter selection methods for standard hybrid methods, there is not one method that will work for all problems, so it is advised to try various approaches in practice.

In order to provide a comparison, we provide "optimal" parameters which are computed as

$$(\gamma_{\text{opt}},\lambda_{\text{opt}})=\underset{0<\gamma\le 1,\,\lambda}{\arg\min}\,\|\mathbf{s}_k(\gamma,\lambda)-\mathbf{s}_{\text{true}}\|_2^2\,, \tag{3.25}$$

where $\mathbf{s}_{\text{true}}$ is the true solution (that is not available in practice).

**Unbiased predictive risk estimation (UPRE).** We can select parameters $\gamma, \lambda$ such that

$$(\gamma_{\mathrm{u}}^{\mathrm{proj}}, \lambda_{\mathrm{u}}^{\mathrm{proj}}) = \underset{0<\gamma\leq 1,\,\lambda}{\arg\min}\, \mathcal{U}_{\mathrm{proj}}(\gamma,\lambda) = \frac{1}{2k+1}\|\mathbf{r}_k^{\mathrm{proj}}(\gamma,\lambda)\|_2^2 + \frac{2\sigma^2}{2k+1}\mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) - \sigma^2 \tag{3.26}$$

where $\sigma^2$ is noise level, and

$$\mathbf{r}_k^{\mathrm{proj}}(\gamma,\lambda) = \mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda) - \begin{bmatrix} \beta_1\mathbf{e}_1 \\ \mathbf{0} \end{bmatrix} \tag{3.27}$$

and

$$\begin{aligned} \mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) &= \mathrm{tr}(\mathbf{C}_k(\gamma,\lambda)\mathbf{D}(\gamma)) \\ &= \mathrm{tr}\big(\big((\mathbf{D}_k(\gamma))^\top\mathbf{D}_k(\gamma) + \lambda^2\gamma\mathbf{I}_k + \lambda^2(1-\gamma)\mathbf{V}_k^\top\mathbf{Q}_2\mathbf{V}_k\big)^{-1}(\mathbf{D}_k(\gamma))^\top\mathbf{D}_k(\gamma)\big). \end{aligned} \tag{3.28}$$

When the noise level $\sigma^2$ is not provided, a noise level estimation algorithm (e.g., based on a wavelet decomposition of the observation) can be utilized [45].

**Generalized cross validation (GCV).** Without a priori knowledge of the noise level, another option is to use an extension of the GCV method [52, 59]. The basic idea is to select parameters,

$$(\gamma_{\mathrm{g}}^{\mathrm{proj}}, \lambda_{\mathrm{g}}^{\mathrm{proj}}) = \underset{0<\gamma\leq 1,\,\lambda}{\arg\min}\, \mathcal{G}_{\mathrm{proj}}(\gamma,\lambda) = \frac{\|\mathbf{r}_k^{\mathrm{proj}}(\gamma,\lambda)\|_2^2}{(\mathrm{tr}(\mathbf{I}_{2k+1} - \mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)))^2} \tag{3.29}$$

where $\mathbf{r}_k^{\mathrm{proj}}(\gamma,\lambda)$, $\mathbf{D}_k(\gamma)$, and $\mathbf{C}_k(\gamma,\lambda)$ are same as (3.26).

Notice that $\mathbf{r}_k^{\mathrm{proj}}$ and $\mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda))$ are functions of $k$ in both the GCV and UPRE

functions. In order to prove convergence of the parameters chosen by UPRE and GCV, we begin with a lemma that shows convergence of the projected residual $\mathbf{r}_k^{\text{proj}}$ and trace term $\text{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda))$ to their full counterparts. The derivations of (3.26) and (3.29) are provided in Appendix A.2.

**Lemma 3.2.** *With (3.27), (3.28), if $k = n$, then*

$$\begin{aligned} \mathbf{r}_k^{\text{proj}} &= \mathbf{r}^{\text{full}}(\gamma,\lambda) \\ \text{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) &= \text{tr}(A(\gamma,\lambda)) \end{aligned} \tag{3.30}$$

*where*

$$\begin{aligned} \mathbf{r}^{\text{full}}(\gamma,\lambda) &= \mathbf{L_R A Q x}(\gamma,\lambda) - \mathbf{L_R b} \\ A(\gamma,\lambda) &= \mathbf{L_R A Q}(\mathbf{Q}^\top\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A Q} + \lambda^2\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{A}^\top\mathbf{L}_\mathbf{R}^\top. \end{aligned} \tag{3.31}$$

*and $\mathbf{r}^{\text{full}}(\gamma,\lambda) = \mathbf{r}_n^{\text{proj}}(\gamma,\lambda)$.*

*Proof.* The proof is provided in Appendix A.3. □

Next we provide convergence results for the UPRE and GCV selected parameters that are similar to results provided in [95] but are extended to the mixed hybrid methods. In particular, we show in Theorem 3.3 that the UPRE parameters for the projected problem converge to the UPRE parameters for the full problem. Then, we show that with an additional weighting parameter, the same result holds for GCV parameters.

**Theorem 3.3.** *From (2.11), the UPRE parameters for the full problem are given by*

$$(\gamma_{\text{u}}^{\text{full}},\lambda_{\text{u}}^{\text{full}}) = \underset{0<\gamma\le 1,\,\lambda}{\arg\min}\,\mathcal{U}_{\text{full}}(\gamma,\lambda) = \frac{1}{m}\|\mathbf{r}^{\text{full}}(\gamma,\lambda)\|_2^2 + \frac{2\sigma^2}{m}\text{tr}(A(\gamma,\lambda)) - \sigma^2. \tag{3.32}$$

*For $k = n$,*

$$(\gamma_{\text{u}}^{\text{proj}},\lambda_{\text{u}}^{\text{proj}}) = (\gamma_{\text{u}}^{\text{full}},\lambda_{\text{u}}^{\text{full}}). \tag{3.33}$$

*Proof.* Since when $k = n$, $\|\mathbf{r}_k^{\text{proj}}\|_2^2 = \|\mathbf{r}^{\text{full}}\|_2^2$ and $\text{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) = \text{tr}(A(\gamma,\lambda))$ as shown in Lemma 3.2,

$$\arg\min_{0<\gamma\leq 1,\,\lambda} \mathcal{U}_{\text{proj}}(\gamma,\lambda) = \arg\min_{0<\gamma\leq 1,\,\lambda} \mathcal{U}_{\text{full}}(\gamma,\lambda)$$

when $k = n$ for the same noise level $\sigma^2$. □

For the full problem, the GCV parameters are given by

$$(\gamma_{\text{g}}^{\text{full}}, \lambda_{\text{g}}^{\text{full}}) = \arg\min_{0<\gamma\leq 1,\,\lambda} \mathcal{G}_{\text{full}}(\gamma,\lambda) = \frac{\|\mathbf{r}^{\text{full}}(\gamma,\lambda)\|_2^2}{(\text{tr}(\mathbf{I}_m - A(\gamma,\lambda)))^2}. \tag{3.34}$$

In contrast with UPRE, $(\gamma_{\text{g}}^{\text{proj}}, \lambda_{\text{g}}^{\text{proj}})$ does not minimize (3.34) when $k = n$ because the trace of $\mathbf{I}_{2k+1} - \mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)$ is not equal to the trace of $\mathbf{I}_m - A(\gamma,\lambda)$. To compensate for this, we define a weighted GCV (WGCV) method by including an additional parameter $\omega$ in (3.29), and computing the WGCV parameters for the projected problem as,

$$(\lambda_{\text{w}}^{\text{proj}}, \gamma_{\text{w}}^{\text{proj}}) = \arg\min_{0<\gamma\leq 1,\,\lambda} \mathcal{W}(\gamma,\lambda)_{\text{proj}} = \frac{\|\mathbf{r}_k^{\text{proj}}(\gamma,\lambda)\|_2^2}{(\text{tr}(\mathbf{I}_{2k+1} - \omega\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)))^2} \tag{3.35}$$

where $\omega = \frac{2k+1}{m}$. Since

$$(\text{tr}(\mathbf{I}_{2k+1} - \omega\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)))^2 = \frac{2k+1}{m}(\text{tr}(\mathbf{I}_m - \mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)))^2, \tag{3.36}$$

$\mathcal{G}_{\text{full}}(\gamma,\lambda)$ is minimized by $(\lambda_{\text{w}}^{\text{proj}}, \gamma_{\text{w}}^{\text{proj}})$ when $k = n$. Similar modified GCV functions were considered in [36, 95].

**Remark 3.4.** It is worth mentioning that although optimization problem (2.11) resembles the widely-studied multi-parameter Tikhonov problem, there are some important distinctions.

For our proposed mixHyBR method, we have $\mathbf{Q} = \gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2$. Under the additional

assumption that $\mathbf{Q}_2$ is positive definite, (2.11) can be written as

$$\min_{\mathbf{x}} \frac{1}{2}||\mathbf{AQx} - \mathbf{b}||^2_{\mathbf{R}^{-1}} + \frac{\lambda^2\gamma}{2}||\mathbf{x}||^2_{\mathbf{Q}_1} + \frac{\lambda^2(1-\gamma)}{2}||\mathbf{x}||^2_{\mathbf{Q}_2} \tag{3.37}$$

$$= \min_{\mathbf{x}} \frac{1}{2}||\mathbf{AQx} - \mathbf{b}||^2_{\mathbf{R}^{-1}} + \frac{\widetilde{\lambda}_1^2}{2}||\mathbf{L}_1\mathbf{x}||^2 + \frac{\widetilde{\lambda}_2^2}{2}||\mathbf{L}_2\mathbf{x}||^2, \tag{3.38}$$

where $\mathbf{Q}_1 = \mathbf{L}_1^\top\mathbf{L}_1$, $\mathbf{Q}_2 = \mathbf{L}_2^\top\mathbf{L}_2$, $\widetilde{\lambda}_1 = \lambda\sqrt{\gamma}$ and $\widetilde{\lambda}_2 = \lambda\sqrt{(1-\gamma)}$. We can see that although the regularization terms are similar in the multi-parameter Tikhonov problem (3.5) and in problem (3.38), the problems are very different. In multi-parameter Tikhonov, regularizaton parameters $\lambda_1$ and $\lambda_2$ only control the corresponding regularization terms. For mixHyBR, from (3.37), we can see that the weighting parameter $\gamma$ affects both the regularization term and the data fit term, which is also evident in the projected problem (3.17). Furthermore, since mixHyBR does not require $\mathbf{Q}_2$ to be positive definite, more general covariance matrices such as data-driven matrices can be included.

## 3.3 Numerical results

In this section, we provide various numerical results from tomography to investigate our proposed hybrid method based on the mixGK process, which we denote as 'mixHyBR'. First, in Section 3.3.1 we investigate data-driven mixed Gaussian priors where we assume that training data are available, and we compare various hybrid methods to existing shrinkage algorithms. Then, we consider a seismic crosswell tomography reconstruction problem in Section 3.3.2, where we show that using a combination of covariance kernels can result in improved reconstructions. For selecting the UPRE, GCV, and WGCV regularization parameters for the mixGK approach, we solve nonlinear constrained optimization problems (3.26), (3.29), and (3.35) respectively. We use an interior-point method as implemented in

MATLAB's `fmincon.m` function with an initial guess of $\gamma_0 = 0.5$ and $\lambda_0 = 0$ at the first mixGK iteration. Then, for subsequent iterations, we used the computed values of $\gamma$ and $\lambda$ from the previous iteration as an initial guess. In general, the regularization parameter parameter selection functions may have many minimizers, but we observed that by using an adaptive strategy, we are able to guide the optimization methods toward appropriate parameters. For the stopping criteria for mixHyBR, we use a combination of approaches described in [36, 37, 38], where the iterative process is terminated if either of the following three criteria is satisfied: (i) a maximum number of iterations is reached, (ii) depending on the chosen regularization parameter selection method, the function (3.26) for UPRE, (3.29) for GCV, or (3.35) for WGCV attains a minimum or flattens out, and (iii) tolerances on residuals are achieved.

### 3.3.1 Spherical tomography example

For our first example, we use a spherical means tomography reconstruction problem from the IRTools toolbox [50, 60]. Such models are often used in imaging problems from photoacoustic or optoacoustic imaging, which is a non-ionizing biomedical imaging modality. The true image $\mathbf{s}_{\text{true}}$ consists of $128 \times 128$ pixels, and the forward model matrix $\mathbf{A}$ represents a ray-tracing operation along semi-circle curves where the angle of centers range from $0^\circ$ to $90^\circ$ at steps of $(90/64)^\circ$. The number of circles at each angle is 90. Thus the dimension of $\mathbf{A}$ is $5,760 \times 16,384$ and the sinogram is $90 \times 64$. The simulated observed sinogram was obtained as in (1.1), where we have included 3% additive Gaussian white noise, i.e., $\frac{\|\boldsymbol{\epsilon}\|}{\|\mathbf{A}\mathbf{s}_{\text{true}}\|} = 0.03$. Other conditions are chosen as the default settings provided by the toolbox; see [50] for details. In the left panel of Figure 3.1, we provide the true image along with some of the integration curves.

Next, we assume that we have a dataset of training images for this problem consisting of 49 images; four of the training images are provided in the right panel of Figure 3.1. All of the images contain a circular mask to denote the region of interest or region of visibility. The inner regions of the images are generated using a linear combination of sine-squared functions, where the coefficients are random numbers uniformly distributed between 0.5 and 1, and the random numbers in sine-squared functions are uniformly distributed between 0 and 128. Furthermore, each image is contaminated by at most 8 "freckles" generated as white disks, where 5 of them have radius 3 and the rest have radius 4. The freckles are randomly placed, where the origins of the freckles are uniformly distributed. Notice that the freckles do not appear in the true image.

Given the training dataset $\{\mathbf{s}^{(1)}, \ldots, \mathbf{s}^{(49)}\}$, we first compute the (vectorized) mean image $\bar{\mathbf{s}}$ and the sample covariance matrix $\widehat{\mathbf{Q}}$ is defined as in (3.2). Next, assuming that the prior covariance matrix represents a Matérn kernel, we solve optimization problem (3.9) to obtain "learned" Matérn parameters $\check{\nu}$ and $\check{\ell}$ and consider the covariance matrix $\mathbf{Q}_{\text{learn}} = \mathbf{Q}(\check{\nu}, \check{\ell})$.

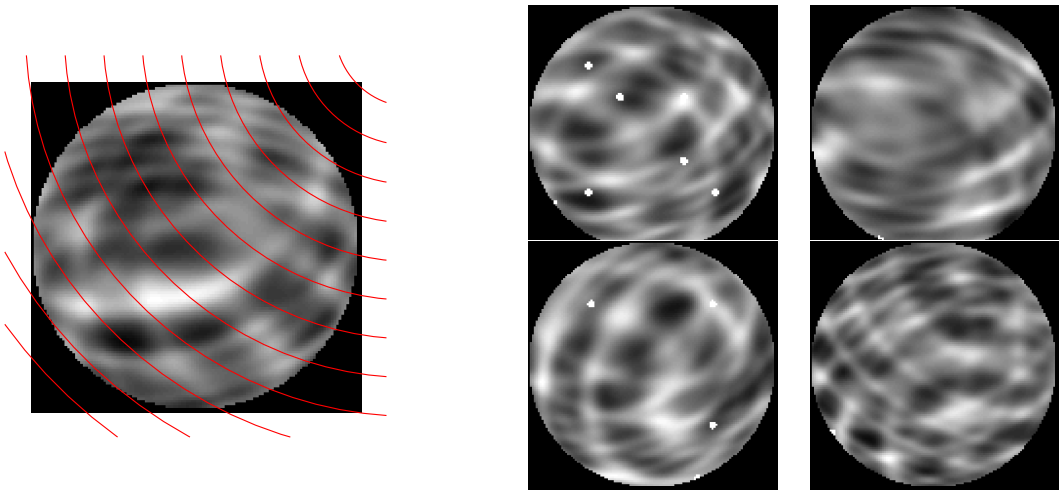

Figure 3.1: Spherical tomography example. On the left, the true image is provided, along with a few of the integration curves whose centers are located at 45°. Four sample images from the training dataset are provided on the right.

To obtain $(\check{\nu}, \check{\ell})$, we used $\nu_0 = 1$ and $\ell_0 = 1$ as initial guess.

We consider four hybrid iterative reconstruction methods, all with initial vector $\bar{\mathbf{s}}$. Given the training data, we run the genHyBR algorithm with $\mathbf{Q} = \mathbf{Q}_{\text{learn}}$ which we denote as 'genHyBR-data-driven'. We also provide results for 'mixHyBR' where $\mathbf{Q} = \gamma\mathbf{Q}_{\text{learn}} + (1-\gamma)\widehat{\mathbf{Q}}$ where $\gamma$ and $\lambda$ are selected during the iterative process. For comparison, we provide results for genHyBR with $\mathbf{Q} = \gamma\mathbf{I} + (1-\gamma)\widehat{\mathbf{Q}}$ where $\gamma$ was pre-selected using the Rao-Blackwell Ledoit and Wolf estimator (rblw) [27, 28, 71]. We also provide results for HyBR where $\mathbf{Q} = \mathbf{I}$, but remark that this approach only uses the training data for the initial (sample mean) vector.

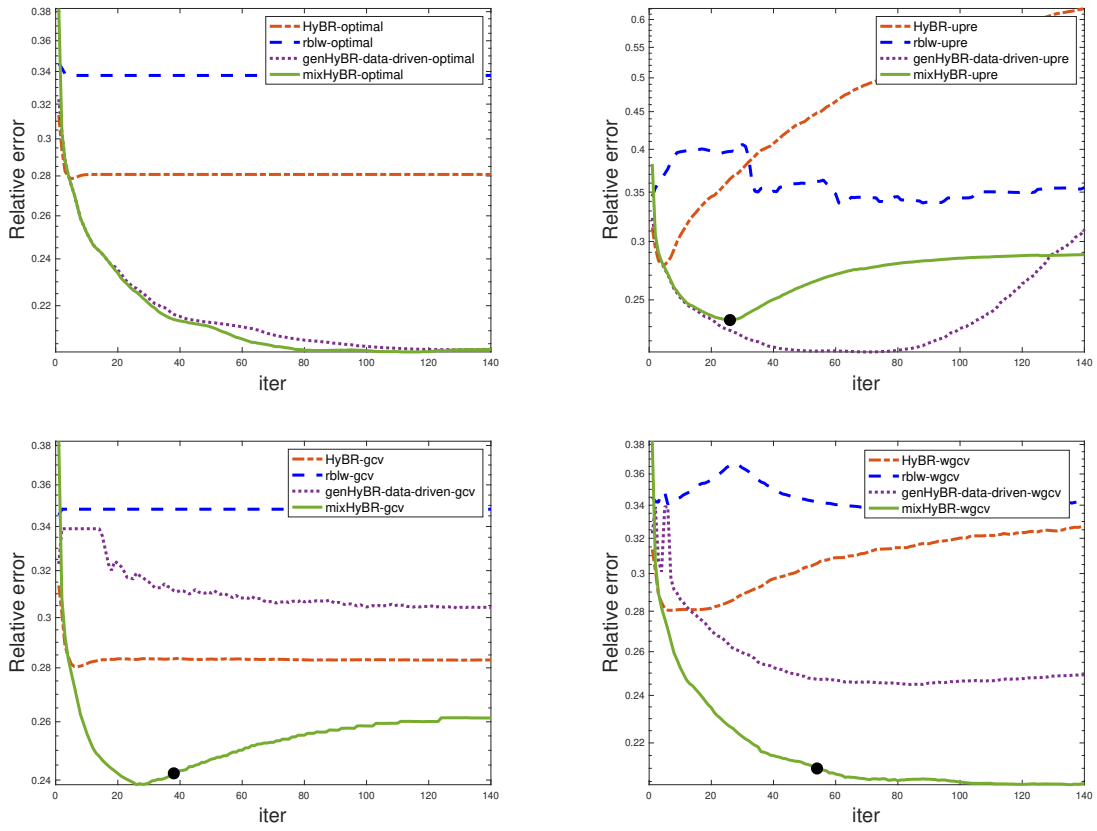


Figure 3.2: Comparison of relative reconstruction error norms for various iterative hybrid approaches for spherical tomography reconstruction. The top left plot corresponds to using the optimal regularization parameters. Other plots correspond to different methods to choose the regularization parameters, including UPRE, GCV, and WGCV.

Note that for all considered methods, the regularization parameter $\lambda$ must be selected, and we investigate various approaches to do this.

In Figure 3.2, we provide relative reconstruction error norms computed as $\|\mathbf{s}_k - \mathbf{s}_{\text{true}}\|_2 / \|\mathbf{s}_{\text{true}}\|_2$, where $\mathbf{s}_k$ is the reconstruction at the $k$th iteration. Each plot corresponds to a different method for selecting the regularization parameters. For comparison, we provide in the top left plot results corresponding to the optimal regularization parameter, although these parameters cannot be computed in practice. We observe that both genHyBR-data-driven and mixHyBR result in small error norms and that even with the optimal regularization parameter $\lambda$, the rblw approach performs poorly because of the poorly-estimated mixing parameter $\gamma$. We remark that we also compared these results to a shrinkage algorithm based on the oracle approximating shrinkage (OAS) estimator [27, 28] for obtaining $\gamma$. However, we observed very similar results as rblw, so we do not include them here.

For the automatic parameter selection methods, we observe that mixHyBR reconstructions with GCV and WGCV and genHyBR-data-driven reconstructions with UPRE have the smallest relative reconstruction error norms per iteration, compared to the other methods. Thus, we observe that including a data-driven covariance matrix, if done properly, can be beneficial. The black dots denote the (automatically-selected) stopping iteration for mixHyBR. Although one may wish to tweak the stopping criteria, all of the examples with mixHyBR resulted in a good reconstruction with the described stopping criteria. For a better comparison of the different parameter selection methods, we provide all relative reconstruction errors for mixHyBR in Figure 3.3, where it is evident that relative errors for WGCV are very close to those for the optimal regularization parameter for this example.

Absolute error images, computed as $|\mathbf{s}_k - \mathbf{s}_{\text{true}}|$, reshaped as an image, and displayed in inverted colormap, are provided in Figure 3.4. For better comparison, all error images have been put on the same scale, and dark regions corresponds to larger absolute errors. Rel-

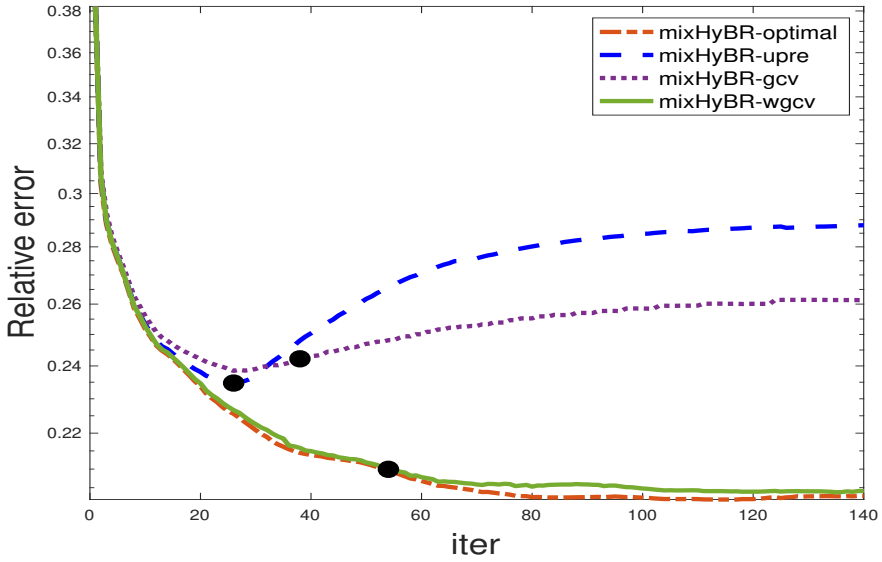


Figure 3.3: Relative reconstruction error norms per iteration of mixHyBR, for various regularization parameter choice methods. Black dots denote the automatically computed stopping iteration.

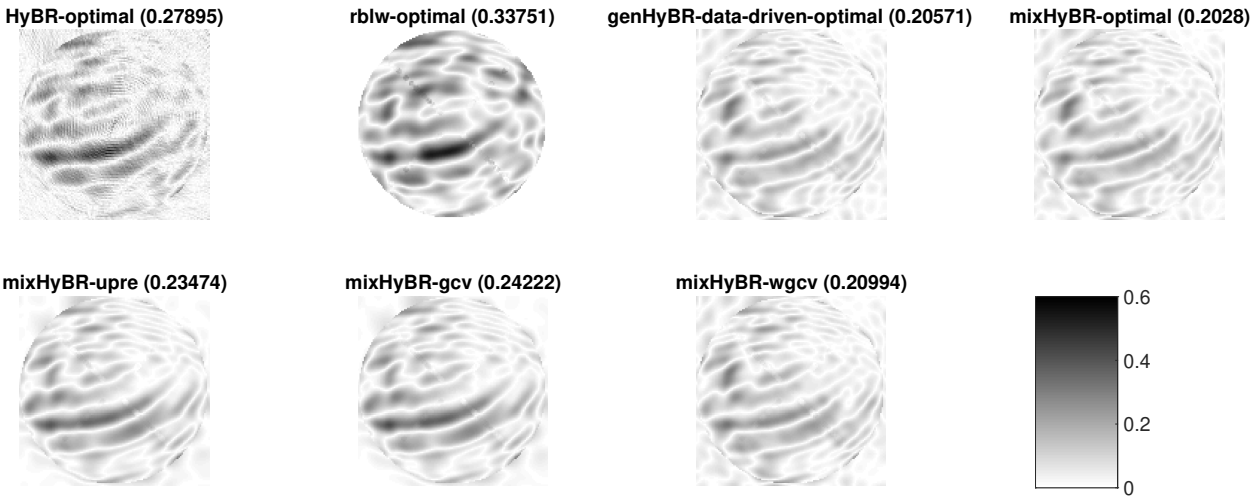


Figure 3.4: Absolute error images (in inverted colormap), with relative reconstruction error norms provided in the titles. The top row compares reconstructions using optimal regularization parameters, and the bottom row compares mixHyBR reconstructions with different parameter choice methods.

ative reconstruction error norms are provided in the titles. In the top row, we compare reconstructions at iteration 140 using the optimal regularization parameter. Absolute error

images in the bottom row correspond to mixHyBR reconstructions with automatic regularization parameter selection and correspond to the iteration determined by the stopping criteria. We notice that even with the optimal regularization parameter, the HyBR-optimal reconstruction suffers from the lack of sufficient prior information and the rblw-optimal reconstruction contains large errors due to the poor choice of $\gamma$ and disruptions due to freckles in the training data. Notice that the rblw-optimal reconstruction has a slight streak along the diagonal of the image, which is a result of the covariance matrix depending heavily on the training data. The mixHyBR and genHyBR-data-driven reconstructions have overall smaller absolute errors in the image. Moreover, we see that by including a combination of covariance matrices and allowing the observed data to inform the choice of $\gamma$, mixHyBR methods can reduce the deleterious effects from corruptions in the training data. For this example, all parameter selection methods combined with the stopping iteration performed reasonably well.

### 3.3.2 Seismic tomography example

In this experiment, we consider a linear inversion problem from crosswell tomography [2]. Crosswell tomography is used to image the seismic wave speed in some region of interest, given data collected from multiple source-receiver pairs. The sources send out a seismic wave, and the receivers measure the travel time taken by the seismic wave to hit the receiver. The goal of the inverse problem is to image the acoustic slowness (reciprocal wave velocity) of the medium in the domain. We consider an example from Continuous Active Source Seismic Monitoring (CASSM) [41], where the goal is to monitor the spatial development of a small scale injection of $CO_2$ into a high quality reservoir. We consider reconstruction at a single time point and investigate the impact of including mixed Gaussian priors on the reconstruction.

The inverse problem can be represented as (1.1) where the goal is to reconstruct the slowness $\mathbf{s} \in \mathbb{R}^{n\times 1}$ of the medium from the measured travel times $\mathbf{d} \in \mathbb{R}^{m\times 1}$ which are assumed to be corrupted by Gaussian white noise $\boldsymbol{\epsilon} \in \mathbb{R}^{m\times 1}$. In our problem setup, the true slowness field was discretized into $n = 188,356$ cells, where the slowness within each cell is assumed to be constant. The true image (normalized between 0 and 1) is of size $434 \times 434$ and was obtained from [40]. For the observations, there were $m_s = 20$ sources and $m_r = 50$ receivers, so a total of $m = m_r m_s$ measurements. Each row of the forward model matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ corresponds to a source-receiver pair. Since the wave travels along a straight line from source to receiver, only the cells lying on the straight line contribute to the non-zero entries. Hence, $\mathbf{A}$ is very sparse with $\mathcal{O}(\sqrt{mn})$ non-zero entries. The true image along with a schematic of the source-detector pairs are given in the left panel of Figure 3.5. The observations, which contain 1% noise, are provided in the right panel of Figure 3.5.

Next we investigate the impact of different choices of $\mathbf{Q}$ on the reconstruction. First, we

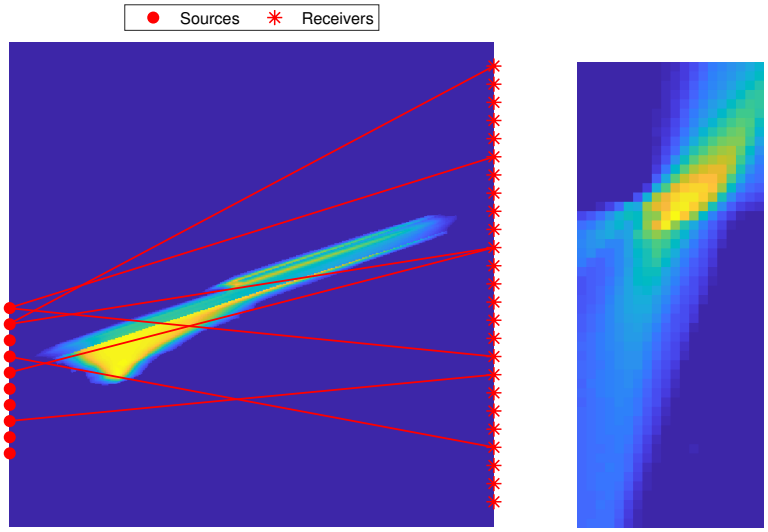


Figure 3.5: CASSM example. In the left panel, we provide the true slowness field image, along with some of the locations of the sources and the detectors. Seven of the source-receiver pairs are highlighted in the figure. In the right panel, we provide the observations corresponding to 20 sources and 50 receivers.

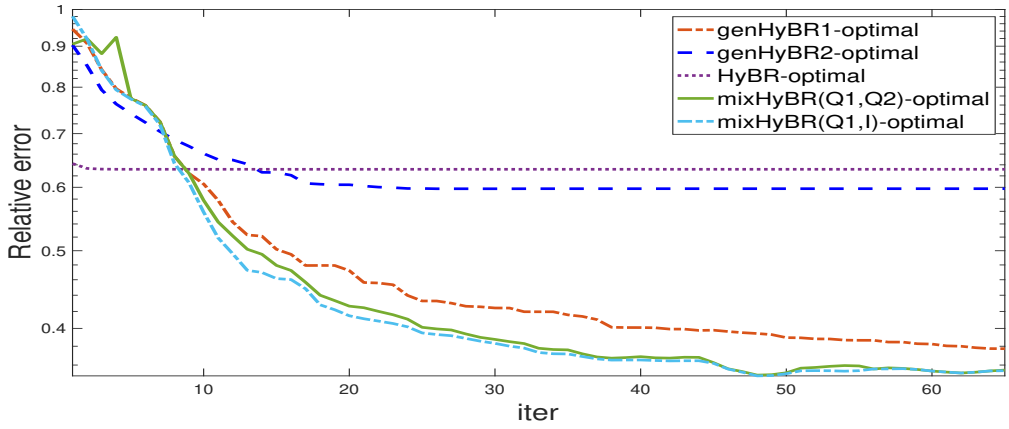


Figure 3.6: Comparison of relative reconstruction error norms for genHyBR and mixHyBR with optimal parameters $\gamma$ and $\ell$.

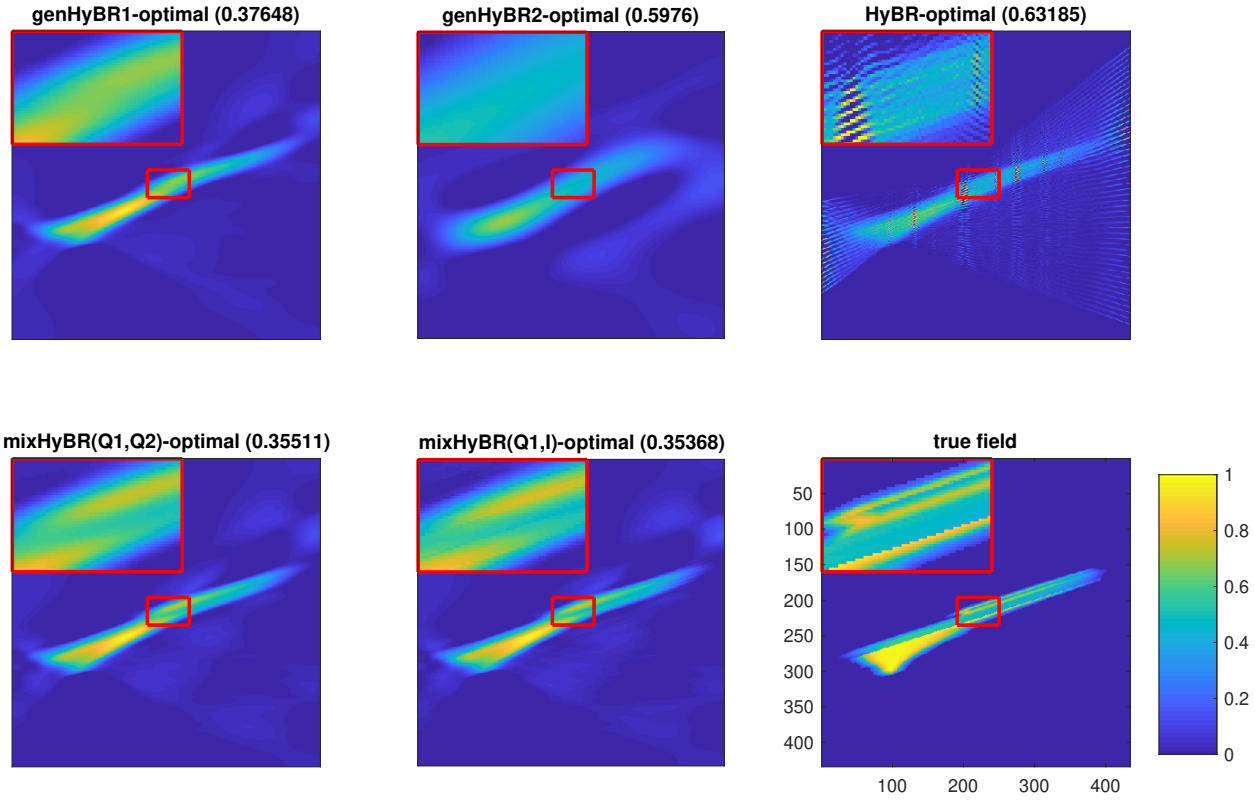


Figure 3.7: Reconstructions with zoomed subimages for CASSM example. All of the reconstructions use the optimal regularization parameter and relative reconstruction errors are provided in the titles.

consider the genHyBR method with three different prior covariance matrices $\mathbf{Q}_1, \mathbf{Q}_2$, and $\mathbf{Q}_3$ defined by a Matérn kernel with $\nu = 1$ and $\ell = 0.2$, a rational quadratic with $\nu = 2$ and $\ell = 0.1$, and the identity matrix $\mathbf{I}$, respectively. These approaches are denoted by 'genHyBR1', 'genHyBR2', and 'HyBR' respectively. Then we consider two mixHyBR approaches that include mixed Gaussian priors, where mixHyBR$(\mathbf{Q}_1, \mathbf{Q}_2)$ uses covariance matrix $\mathbf{Q} = \gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{Q}_2$ and mixHyBR$(\mathbf{Q}_1, \mathbf{I})$ uses covariance matrix $\mathbf{Q} = \gamma\mathbf{Q}_1 + (1-\gamma)\mathbf{I}$, where the mixing parameter $\gamma$ is selected during the reconstruction process. For the optimally selected regularization parameters, we provide in Figure 3.6 the relative reconstruction error norms per iteration.

We observe that if a good covariance matrix (in this case, $\mathbf{Q}_1$) is known in advance, stand-

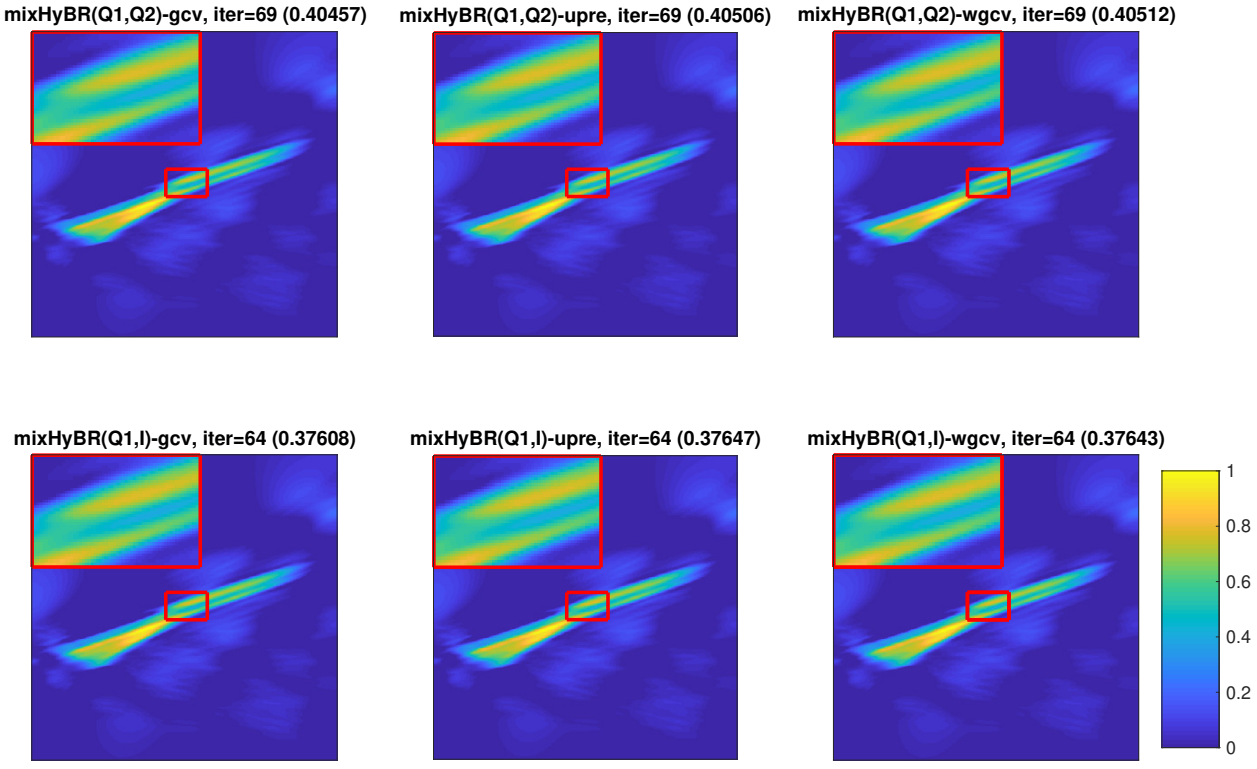


Figure 3.8: Reconstructions of mixHyBR$(\mathbf{Q}_1, \mathbf{Q}_2)$ (top row) and mixHyBR$(\mathbf{Q}_1, \mathbf{I})$ (bottom row) for different parameter choice methods. The automatically detected stopping iteration (iter) and corresponding relative reconstruction error norm are provided in the titles.

alone genHyBR can perform well and result in small relative reconstruction errors. Otherwise, the relative reconstruction errors may remain large, and multiple solves with different covariance matrices would be needed to determine a good prior. In this case, the mixHyBR approach can prove beneficial. The mixHyBR approaches produce reconstructions with overall smaller relative reconstruction errors than genHyBR with each covariance matrix alone. Image reconstructions, including a zoomed subregion, are provided in Figure 3.7. Notice that the mixed Gaussian priors are better able to resolve some details of the true image. While the reconstructions of genHyBR1 and genHyBR2 are too smooth, mixHyBR($\mathbf{Q}_1, \mathbf{Q}_2$) and mixHyBR($\mathbf{Q}_1, \mathbf{I}$) reconstructions reveal multiple layered regions that are present in the true field. Thus, incorporating mixed Gaussian priors can lead to improved reconstructions.

Next we investigate the performance of different regularization parameter selection methods within the mixHyBR methods. Reconstructions corresponding to GCV, WGCV, and UPRE parameter choices are provided in Figure 3.8, with stopping iterations and relative reconstruction errors provided in the titles. For the stopping criteria, we used a tolerance of $10^{-6}$ for the residual norm. Although the relative reconstruction error norms are larger than those for the optimal regularization parameter, which is likely due to the increased errors in the surrounding flat regions, the reconstructed images can all distinguish the two yellow layers.

# Chapter 4

# Hybrid Projection Methods for Large-scale Inverse Problems with Mean Estimation in Hierarchical Gaussian Prior

Satellites and ground based sensors observe greenhouse gas (GHG) concentrations in the atmosphere, and inverse problems must be solved to generate detailed maps of surface emissions using these observations of atmospheric mixing ratios [14, 78, 79, 81, 105, 109]. The goal of inverse modeling is to estimate spatiotemporal GHG fluxes at the Earth's surface using observations of these gases in the atmosphere. Similar to many other important scientific applications, this is a large-scale inverse problem and there are many computational challenges for solving the resulting optimization problems accurately and effectively. For example, atmospheric monitoring of carbon dioxide ($CO_2$) has increased the number of situ observation sites [54, 102], aircraft-based observations [77, 91], and $CO_2$ observing satellites [20, 46, 47, 82, 113]. Being able to handle these enormous datasets is one of the computational challenges for solving the inverse problem to track gas emissions. In a classical Bayesian inversion, various computational challenges arise for large-scale inverse problems, and in current approaches, it is assumed that a regularization parameter is known in ad-

vance. Furthermore, researchers assume a Gaussian prior distribution for use with Bayes' theorem, where the prior mean vector and covariance matrix are required. These parameters are usually assumed to be known and are estimated using expert knowledge. In this chapter, we argue that these parameters may be inappropriate choices for abnormal or extreme circumstances. We propose to use a hierarchical Gaussian model, and we develop optimization tools for reconstructing a desired parameter and its prior mean simultaneously and selecting a regularization parameter automatically.

## 4.1 Problem set up

We are interested in inverse problems of the form (1.1). Since ill-posedness occurs in inverse problems, small errors in the data can lead to huge errors in the reconstructed estimation $\mathbf{s}$. To estimate a stable reconstruction $\mathbf{s}$ in the inversion process, we treat $\mathbf{s}$ as a Gaussian random variable as described in (2.1). The reconstruction quality of (1.1) depends crucially on choosing appropriate hyperparameters that govern the prior and the noise distribution. For example, previous experiments in [39] for seismic tomography applications assumed the prior mean to be a zero vector or a constant vector in numerical experiments. In this chapter, $\mathbf{s}$ contains spatial or spatiotemporal parameters to be estimated. That means that using a simplified prior mean vector (e.g., where every entry has the same real value) is not necessarily appropriate. We seek an approach that incorporates some expert or physical knowledge. A better assumption is to use a spatial or temporal pattern in the prior mean. Geostatistical inverse modeling (GIM) has been a common approach to estimating spatiotemporal unknowns. Hence, a hierarchical Gaussian approach for modeling prior information that uses covariates (i.e., auxiliary variables) [80, 98] is considered. In this setting, we develop efficient Krylov subspace projection methods for solving inverse problems and quantifying

uncertainty with unknown mean. We represent the prior information in the form of the hierarchical model

$$\mathbf{s}|\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \lambda^{-2}\mathbf{Q}), \qquad \boldsymbol{\beta} \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \lambda_\beta^{-2}\mathbf{Q}_\beta). \tag{4.1}$$

Here $\mathbf{X} \in \mathbb{R}^{n\times p}$ is given and deterministic, $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is the prior covariance matrix that we assume known up to a constant, and $\lambda$ is a scaling parameter. We assume that matrix $\mathbf{Q}$ is defined by a covariance kernel, which may model spatial or spatiotemporal prior knowledge. The unknowns here are contained in the vector $\boldsymbol{\beta} \in \mathbb{R}^p$ and are assumed to follow a Gaussian distribution with given mean $\boldsymbol{\mu}_\beta \in \mathbb{R}^p$, covariance matrix $\mathbf{Q}_\beta \in \mathbb{R}^{p\times p}$, and $\lambda_\beta$ is a scaling parameter. We assume $\lambda_\beta = \gamma\lambda$ with a constant $\gamma > 0$ and $\gamma$ is predetermined. Previous works [80, 98] assume that $\boldsymbol{\beta}$ is uniform (i.e., $p(\boldsymbol{\beta}) \propto 1$). The novelty of (4.1) is treating $\boldsymbol{\beta}$ as a random variable of Gaussian distribution. Hence, this assumption allows more flexibility in the definition of $\boldsymbol{\beta}$.

## 4.2 Generalized hybrid approaches to solve the GIM and associated challenges for large datasets and mean estimation

Since the prior covariance matrix tends to be dense and its size follows the desired parameter vector, it is infeasible to compute and store the inverse of $\mathbf{Q}$. In this section, we avoid the inverse of the prior covariance matrix by exploiting a change of variables [38]. Thus, we can consider covariance matrices $\mathbf{Q}$ that are defined by some kernel function which is determined by a few kernel parameters [93]. We use a kernel based covariance matrix which means that we do not need to explicitly construct the full matrix $\mathbf{Q}$. Instead, we provide

functions to perform efficient matrix-vector multiplications with $\mathbf{Q}$. Likewise, $\mathbf{Q}_\beta$ can also be parameterized for efficient implementation.

### 4.2.1 Hierarchical Gaussian priors: Reformulation for mean estimation

Given the assumptions in (1.1) and (4.1), from Bayes' theorem the posterior probability density function can be written as

$$\pi(\mathbf{s},\boldsymbol{\beta}|\mathbf{d}) \propto \exp\left(-\frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|^2_{\mathbf{R}^{-1}} - \frac{\lambda^2}{2}\|\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\|^2_{\mathbf{Q}^{-1}} - \frac{\lambda^2_\beta}{2}\|\boldsymbol{\beta}-\boldsymbol{\mu}_\beta\|^2_{\mathbf{Q}_\beta^{-1}}\right), \tag{4.2}$$

where $\|\mathbf{x}\|^2_{\mathbf{M}} = \mathbf{x}^\top\mathbf{M}\mathbf{x}$ and $\mathbf{M}$ is a symmetric positive definite matrix. The maximum a posterior (MAP) estimate can be written as the solution of the optimization problem,

$$(\mathbf{s}_{\text{MAP}},\boldsymbol{\beta}_{\text{MAP}}) = \arg\min_{\mathbf{s},\boldsymbol{\beta}} f(\mathbf{s},\boldsymbol{\beta}) = \frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\|\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\|^2_{\mathbf{Q}^{-1}} + \frac{\lambda^2_\beta}{2}\|\boldsymbol{\beta}-\boldsymbol{\mu}_\beta\|^2_{\mathbf{Q}_\beta^{-1}} \tag{4.3}$$

If symmetric decompositions of the inverses of the covariance matrices are available,

$$\mathbf{R}^{-1} = \mathbf{L}_\mathbf{R}^\top\mathbf{L}_\mathbf{R}, \quad \mathbf{Q}^{-1} = \mathbf{L}_\mathbf{Q}^\top\mathbf{L}_\mathbf{Q}, \quad \text{and,} \quad \mathbf{Q}_\beta^{-1} = \mathbf{L}_\beta^\top\mathbf{L}_\beta,$$

then the optimization problem can be rewritten in least squares form as,

$$(\mathbf{s}_{\text{MAP}},\boldsymbol{\beta}_{\text{MAP}}) = \arg\min_{\mathbf{s},\boldsymbol{\beta}} f(\mathbf{s},\boldsymbol{\beta}) = \frac{1}{2}\left\|\begin{bmatrix}\mathbf{L}_\mathbf{R}\mathbf{A} & \mathbf{0}\\ \lambda\mathbf{L}_\mathbf{Q} & -\lambda\mathbf{L}_\mathbf{Q}\mathbf{X}\\ \mathbf{0} & \lambda_\beta\mathbf{L}_\beta\end{bmatrix}\begin{bmatrix}\mathbf{s}\\ \boldsymbol{\beta}\end{bmatrix} - \begin{bmatrix}\mathbf{L}_\mathbf{R}\mathbf{d}\\ \mathbf{0}\\ \lambda_\beta\mathbf{L}_\beta\boldsymbol{\beta}\end{bmatrix}\right\|^2_2.$$

Although $\mathbf{L}_\mathbf{R}$ and $\mathbf{L}_\beta$ are typically easy to compute, obtaining $\mathbf{L}_\mathbf{Q}$ is computationally in-

feasible for the problems of interest. Here, we propose a change of variables to apply the genHyBR method,

$$\mathbf{p} \leftarrow \mathbf{s} - \mathbf{X}\boldsymbol{\mu}_\beta, \quad \mathbf{q} \leftarrow \boldsymbol{\beta} - \boldsymbol{\mu}_\beta, \quad \mathbf{y} \leftarrow \mathbf{d} - \mathbf{A}\mathbf{X}\boldsymbol{\mu}_\beta, \tag{4.4}$$

Consider the composite vector $\widetilde{\mathbf{p}} = \begin{bmatrix} \mathbf{p} \\ \mathbf{q} \end{bmatrix} \in \mathbb{R}^{n+p}$ and a composite matrix $\widetilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m\times(n+p)}$. The optimization problem (4.3) can be written as follows,

$$\min_{\widetilde{\mathbf{p}}\in\mathbb{R}^{n+p}} \frac{1}{2}\|\widetilde{\mathbf{A}}\widetilde{\mathbf{p}} - \mathbf{y}\|^2_{\mathbf{R}^{-1}} + \frac{\lambda^2}{2}\|\widetilde{\mathbf{p}}\|^2_{\widetilde{\mathbf{Q}}^{-1}}. \tag{4.5}$$

where

$$\widetilde{\mathbf{Q}}^{-1} = \begin{bmatrix} \mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{X} \\ -\mathbf{X}^\top\mathbf{Q}^{-1} & \mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X} + \gamma^2\mathbf{Q}_\beta^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{I} \\ -\mathbf{X}^\top \end{bmatrix} \mathbf{Q}^{-1} \begin{bmatrix} \mathbf{I} & -\mathbf{X} \end{bmatrix} + \gamma^2 \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_\beta^{-1} \end{bmatrix}, \tag{4.6}$$

$$\widetilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} + \frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta\mathbf{X}^\top & \frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta \\ \frac{1}{\gamma^2}\mathbf{Q}_\beta\mathbf{X} & \frac{1}{\gamma^2}\mathbf{Q}_\beta \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \frac{1}{\gamma^2} \begin{bmatrix} \mathbf{X} \\ \mathbf{I} \end{bmatrix} \mathbf{Q}_\beta \begin{bmatrix} \mathbf{X}^\top & \mathbf{I} \end{bmatrix}. \tag{4.7}$$

Details of the derivations are provided in Appendix B.1 and B.2.

Note that we do not explicitly construct $\widetilde{\mathbf{Q}}$. Instead, we provide a function that performs an efficient matrix-vector multiplication with $\widetilde{\mathbf{Q}}$. Therefore, we can write the MAP estimate as the solution of (4.5) as the required form of the optimization problem in [38]. Then, we apply the generalized hybrid projection method and recover $\mathbf{s}$ and $\boldsymbol{\beta}$ from the solution of (4.5) by

$$\mathbf{s} = \mathbf{p} + \mathbf{X}\boldsymbol{\mu}_\beta \quad \text{and} \quad \boldsymbol{\beta} = \mathbf{q} + \boldsymbol{\mu}_\beta. \tag{4.8}$$

In particular, if $\boldsymbol{\mu}_\beta = \mathbf{0}$, then $\widetilde{\mathbf{p}} = \begin{bmatrix} \mathbf{s} \\ \boldsymbol{\beta} \end{bmatrix}$ and $\mathbf{y} = \mathbf{d}$. Likewise, the prior condition (2.1) can lead to a similar optimization problem as follow,

$$\min_{\mathbf{s}\in\mathbb{R}^n} \frac{1}{2}\|\mathbf{A}\mathbf{s} - \mathbf{d}\|^2_{\mathbf{R}^{-1}} - \frac{\lambda^2}{2}\|\mathbf{s} - \boldsymbol{\mu}\|^2_{\mathbf{Q}^{-1}} \tag{4.9}$$

where $\boldsymbol{\mu}$ is fixed.

### 4.2.2 Generalized hybrid projection methods for computing estimates and uncertainties

Given matrices $\widetilde{\mathbf{A}}$, $\mathbf{R}$, $\widetilde{\mathbf{Q}}$ and vector $\mathbf{y}$ from (4.5), the genGK bidiagonalization process generates matrices $\mathbf{B}_k$, $\mathbf{U}_{k+1}$, and $\mathbf{V}_k$ satisfying (2.12), (2.13), (2.14), and (2.15). Then, at the $k$th iteration,

$$\min_{\mathbf{x}_k\in\mathbb{R}^k} \frac{1}{2}\|\mathbf{B}_k\mathbf{x}_k - \beta_1\mathbf{e}_1\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{x}_k\|_2^2 \tag{4.10}$$

is the projected optimization problem. The regularization parameter $\lambda$ is selected by methods described in Section 2.3. The solution to (4.5) is recovered as $\widetilde{\mathbf{p}}_k = \widetilde{\mathbf{Q}}\mathbf{V}_k\mathbf{x}_k$.

The posterior covariance matrix is the inverse of the Hessian matrix of (4.3) with respect to $\mathbf{s}, \boldsymbol{\beta}$ given $\mathbf{d}$. Hence, it is given by

$$\mathbf{Q}_{\text{post}} := \begin{bmatrix} \lambda^2\mathbf{Q}^{-1} + \mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A} & -\lambda^2\mathbf{Q}^{-1}\mathbf{X} \\ -\lambda^2\mathbf{X}^\top\mathbf{Q}^{-1} & \lambda^2_\beta\mathbf{Q}_\beta^{-1} + \lambda^2\mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X} \end{bmatrix}^{-1} \tag{4.11}$$

Our goal is to approximate the posterior covariance of $\mathbf{s}$ and $\boldsymbol{\beta}$ given $\mathbf{d}$, but because its dimension is $(n+p)$-by-$(n+p)$, it is infeasible to construct explicitly. Hence, low-rank approximation methods provide a computationally reasonable alternative. Note that the

change of variables in (4.4) just shifts $\mathbf{s}$ and $\boldsymbol{\beta}$ into $\mathbf{p}$ and $\mathbf{q}$ respectively, and thus the posterior covariance matrix of $\widetilde{\mathbf{p}}$ given $\mathbf{y}$ is equivalent to (4.11). Therefore, from (4.6) and (4.11), we get the following expression for the posterior covariance matrix,

$$\mathbf{Q}_{\text{post}} = \left(\lambda^2 \widetilde{\mathbf{Q}}^{-1} + \widetilde{\mathbf{A}}^\top \mathbf{R}^{-1} \widetilde{\mathbf{A}}\right)^{-1}. \tag{4.12}$$

Now, we can approximate $\mathbf{Q}_{\text{post}}$ as described in Section 2.4 with matrices generated from the genGK bidiaognalization process with $\widetilde{\mathbf{Q}}$, $\widetilde{\mathbf{A}}$, and $\mathbf{R}$.

## 4.3 Numerical results

For the numerical experiments in this section, we use the following to denote the different considered iterative methods,

- `genHyBRs` refers to solving (1.1) with a Gaussian prior (2.1) via the optimization problem (4.9).
- `genHyBRmean` refers to solving (1.1) with the hierarchical Gaussian prior (4.1) via the optimization problem (4.5).

### 4.3.1 Experiment 1: Simulation Study

We demonstrate the potential of our proposed approach using a simulated example from the NOAA institute GHG monitoring network for the western US. This is a spatial reconstruction for a problem with small noise level. The main goal of this experiment is to show that `genHyBRmean` gives a better reconstruction than `genHyBRs`.

For simplicity we integrate over the time component and only consider a spatial reconstruction. For this example, the "true" emissions in $\mathbf{s}$ represent a resolution of 0.1° latitude by 0.1° longitude (see Fig. 4.1), matrix $\mathbf{A}$ is $1,215 \times 16,829$ and represents an atmospheric transport model. Using simulated observations with 4% Gaussian noise (i.e., $\frac{\|\boldsymbol{\epsilon}\|_2}{\|\mathbf{A}\mathbf{s}_{\text{true}}\|_2} = 0.04$), we compare reconstructions obtained by generalized LSQR (LSQR is a standard and widely used least-squares solver [86, 87]), `genHyBRs`, and `genHyBRmean`. We denote Generalized LSQR by `genHyBRs-none/genHyBRmean-none` because the generalized LSQR methods use a regularization parameter of 0 although the prior covariance matrix is not just a scaled identity matrix. `genHyBRs/genHyBRmean-opt` use gen-GK with the optimal regularization parameter. `genHyBRs/genHyBRmean-dp` use gen-GK with the DP and `genHyBRs/genHyBRmean-wgcv` use gen-GK with the WGCV method as decribed in Section 2.3. The UPRE method was also tested since it did not provide meaningful results for both `genHyBRs` and `genHyBRmean`, we omit these results. However, we remark that one observation from the UPRE function in the projected problem is that the residual term tends to dominate the UPRE function. Since the residual term rapidly increases as $\lambda$ increases, the UPRE function achieves its minimum value near zero which corresponds to no regularization.

In Fig. 4.2, the relative reconstruction errors are only computed on land (i.e., reconstructed values over the ocean are ignored) and are computed as $\frac{\|\mathbf{s}_k - \mathbf{s}_{\text{true}}\|}{\|\mathbf{s}_{\text{true}}\|}$, where $\mathbf{s}_{\text{true}}$ contains the true emissions and $\mathbf{s}_k$ is the reconstruction at the $k$th iteration. We observe that this problem is severely underdetermined, so the choices of prior and regularization parameter are very important. Here, $\mathbf{Q}$ is constructed using a Matérn kernel with $\nu = 2.5$ and $\ell = 0.1$ as described in [93]. For `genHyBRmean`, $\mathbf{X}$ is a vector of ones and $\mathbf{Q}_{\boldsymbol{\beta}} = 1$, and we simultaneously solve for the image and the mean constant. For `genHyBRmean` methods, three different $\gamma$ are tested and results are presented in Table 4.1. The reconstructions Fig. 4.1 and the relative errors Fig. 4.2 are computed for $\gamma = 1$. In addition to having smaller reconstruction errors

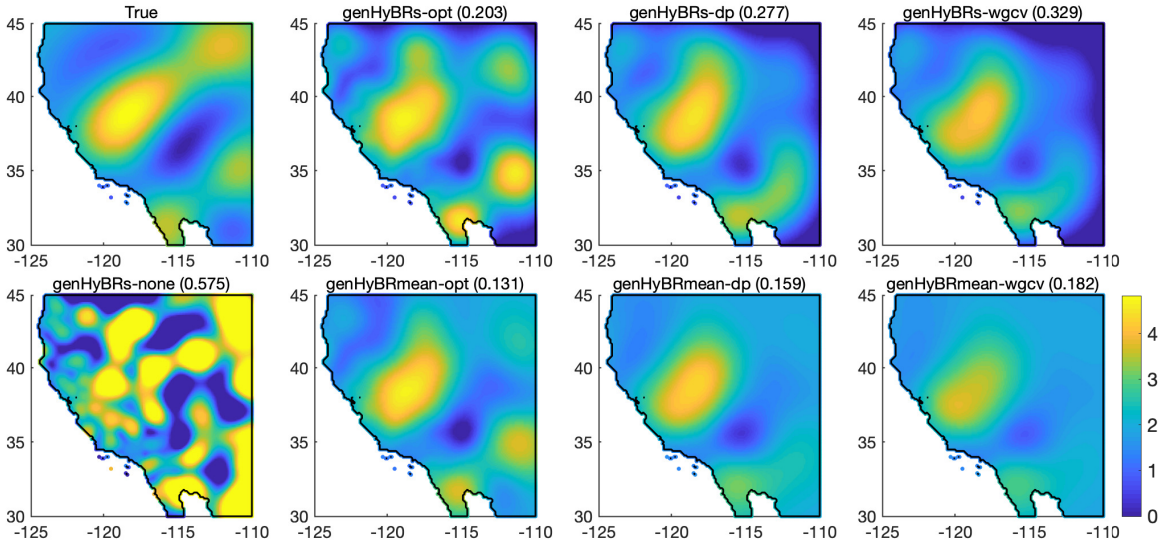


Figure 4.1: For each method, the relative errors are only on land without ocean.

than `genHyBRs-opt`, `genHyBRmean` reconstructions appear to have desirable nonnegative entries. Although this was not explicitly enforced, it can be attributed to the mean estimation which shifts the solution to the appropriate value. We observe that `genHyBRmean` tends to force over-smoothing on computed reconstructions, although the relative errors are smaller than `genHyBRs`.

We also provide an approximation of the solution uncertainties for $\beta$. $\beta_{\text{true}} = 3$ and the computed means are shown in Table 4.1. For small $\gamma$ where $\lambda_\beta = \gamma\lambda$, $\beta_{\text{MAP}}$ is close to 3. The approximation of $\sigma_{\text{post}}$ has no significant impact for different $\gamma$ values.

### 4.3.2 Experiment 2: 6 week case study

This is an example of an atmospheric reconstruction example where the goal is to reonstruct spatiotemporal $CO_2$ fluxes. This example is described in [80], and is a very large and challenging problem. The goal of this experiment is investigate the use of generalized hybrid

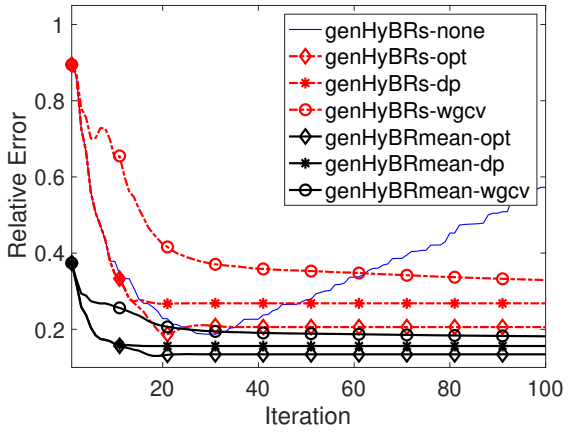


Figure 4.2: The plot shows the relative errors only on land without ocean for each iterative method.

| Methods($\gamma = \sqrt{0.1}$) | $\beta_{\text{MAP}}$ | $\sigma_\beta^{\text{post}}$ |
|---|---|---|
| `genHyBRmean-opt` | 2.8522 | 0.2325 |
| `genHyBRmean-dp` | 2.8634 | 0.0949 |
| `genHyBRmean-wgcv` | 2.8532 | 0.0471 |
| Methods($\gamma = 1$) | $\beta_{\text{MAP}}$ | $\sigma_\beta^{\text{post}}$ |
| `genHyBRmean-opt` | 2.7012 | 0.2358 |
| `genHyBRmean-dp` | 2.6816 | 0.0948 |
| `genHyBRmean-wgcv` | 2.6392 | 0.0466 |
| Methods($\gamma = \sqrt{10}$) | $\beta_{\text{MAP}}$ | $\sigma_\beta^{\text{post}}$ |
| `genHyBRmean-opt` | 1.7746 | 0.2417 |
| `genHyBRmean-dp` | 1.6932 | 0.0904 |
| `genHyBRmean-wgcv` | 1.5939 | 0.0488 |

Table 4.1: The table shows the computed $\beta$ and approximation of its posterior standard deviation for each iterative method.

methods for solving spatiotemporal $CO_2$ flux problems and to compare the results with automatic parameter selection to the results obtained using a direct method and and iterative method which has no selection methods for the regularization parameter. Here, we consider the use of the DP to select the regularization parameter, and we compare to problems with different noise levels.

We applied the hierarchical prior model (4.1) to genHyBR using the case study based on $CO_2$ observations from NASA's Orbiting Carbon Observatory 2 (OCO-2) satellite [47, 80]. We estimate 6 weeks of $CO_2$ fluxes over North America from late June through July 2015. There are $1.92 \times 10^4$ synthetic observations and the size of the unknown $CO_2$ fluxes is $1.06 \times 10^6$. Each flux represents a 3-hours temporal resolution and a $1^\circ \times 1^\circ$ latitude-longitude resolution. The synthetic observations were generated by the atmospheric transport model from NOAA's CarbonTracker-Lagrange program [66]. The matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{328} \in \mathbb{R}^{19,156 \times 3,222}$ were generated using the Weather Research and Forecasting (WRF) Stochastic Time-Inverted Lagrangian Transport Model (STILT) modeling system [73, 83]. Since the matrix $\mathbf{A}$ is con-

structed by stacking block matrices $\mathbf{A}_1, \ldots, \mathbf{A}_{328}$ horizontally as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \ldots & \mathbf{A}_{328} \end{bmatrix} \in \mathbb{R}^{19{,}156 \times 328 \cdot 3{,}222} \tag{4.13}$$

and it is too large to use as a single matrix, each matrix $\mathbf{A}_i$ is loaded when the matrix-vector multiplication is necessary. The matrix-vector multiplication operations with $\mathbf{A}$ and $\mathbf{A}^\top$ are implemented using object oriented programming.

The unknown $CO_2$ fluxes are contained in vector

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_{328} \end{bmatrix} \in \mathbb{R}^{328 \cdot 3222} \tag{4.14}$$

After obtaining observations or outputs from this atmospheric transport model, we added some noise to represent measurement error, as in (1.1).

For the hierarchical model (4.1), we use a matrix $\mathbf{X}_{\text{sub}}$ stacking vertically 41 times (for 6 weeks) to construct $\mathbf{X}$ such as

$$\mathbf{X} = \mathbf{1}_{41} \otimes \mathbf{X}_{\text{sub}} = \begin{bmatrix} \mathbf{X}_{\text{sub}} \\ \vdots \\ \mathbf{X}_{\text{sub}} \end{bmatrix} \in \mathbb{R}^{3222 \cdot 8 \cdot 41 \times 8} \tag{4.15}$$

where for one day,

$$\mathbf{X}_{\text{sub}} = \mathbf{I}_8 \otimes \mathbf{1}_{3222} = \begin{bmatrix} \mathbf{1}_{3222} & & \\ & \ddots & \\ & & \mathbf{1}_{3222} \end{bmatrix} \in \mathbb{R}^{3222 \cdot 8 \times 8} \tag{4.16}$$

with $\mathbf{1}_{3222} \in \mathbb{R}^{3222}$ is column of ones. We set the mean vector $\boldsymbol{\mu}_\beta$ to be zero. For the prior covariance matrix of unknown fluxes, $\mathbf{Q} = \mathbf{Q}_t \otimes \mathbf{Q}_s$ where $\mathbf{Q}_t$ represents the temporal covariance and $\mathbf{Q}_s$ represents the spatial covariance in the fluxes.

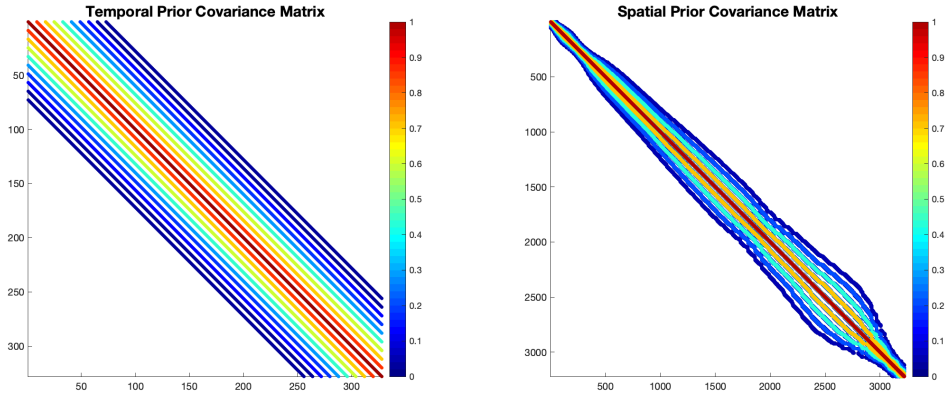


Figure 4.3: Prior covariance matrices for temporal and spatial priors.

These covariance matrices are provided in Figure 4.3, and they are defined by kernel functions

$$k_t(d_t;\theta_t) = \begin{cases} 1 - \frac{3}{2}\left(\frac{d_t}{\theta_t}\right) + \frac{1}{2}\left(\frac{d_t}{\theta_t}\right)^3 & \text{if } d_t \le \theta_t, \\ 0 & \text{if } d_t > \theta_t, \end{cases} \tag{4.17}$$

$$k_s(d_s;\theta_s) = \begin{cases} 1 - \frac{3}{2}\left(\frac{d_s}{\theta_s}\right) + \frac{1}{2}\left(\frac{d_s}{\theta_s}\right)^3 & \text{if } d_s \le \theta_s, \\ 0 & \text{if } d_s > \theta_s, \end{cases} \tag{4.18}$$

where $d_t$ is the day difference between two unknowns, $d_s$ is the spherical distance between two unknowns, and $\theta_t, \theta_s$ are kernel parameters. In this setting, we set $\theta_t = 9.854$ and $\theta_s = 555.42$. The covariance matrix $\mathbf{Q}_\beta$ is the identity matrix. The covariance matrix $\mathbf{R}$ is a diagonal matrix whose diagonal entries are all $\sigma^2$. In [80], $\sigma = 2$ was the standard deviation of the noise. Since $\sigma = 2$ leads to very large noise in the observations, in this experiment, we tested different $\sigma$ values as shown in Table 4.2. More specifically, for $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ the noise

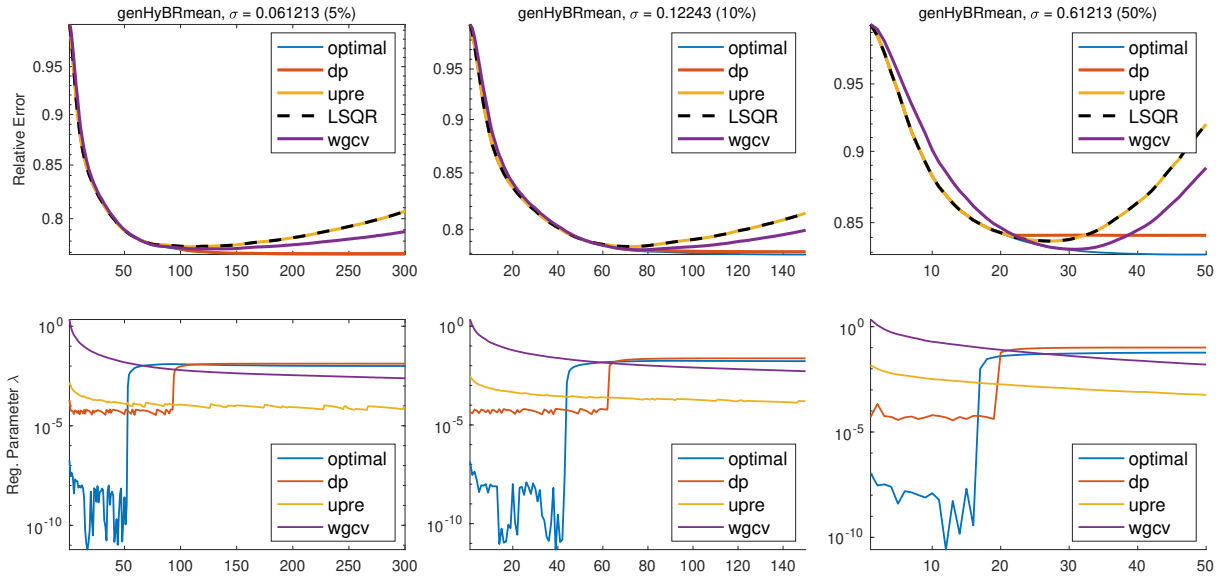


Figure 4.4: Error analysis of hybrid projection methods. Experiments of 5%, 10%, 50% noise level are corresponding to the first, second, and third columns respectively. The first row shows relative errors of `genHyBRmean` with different regularization parameter selection methods. The relative errors of Direct solutions are not included because they are over 100% relative error. The second row shows computed regularization parameter $\lambda$ at each iteration.

level (`nlevel`) of the observation corresponds to adding $\boldsymbol{\epsilon} = \sigma\mathbf{n}$ where $\sigma = \texttt{nlevel} \cdot \frac{\|\mathbf{z}\|_2}{\|\mathbf{n}\|_2}$.

| `nlevel` | $\sigma$ |
|---|---|
| 5% | 0.0612 |
| 10% | 0.1224 |
| 50% | 0.6121 |
| 163% | 2.0000 |

Table 4.2: Noise level and corresponding $\sigma$.

For each noise level, `genHyBRs` and `genHyBRmean` are tested. In the first row of Figure 4.4, the relative errors of hybrid methods are presented for different noise levels. `genHyBRmean-none`, `genHyBRmean-upre`, and `genHyBRmean-wgcv` show semi-convergence behavior (e.g., initial convergence and later divergence) while `genHyBRmean-dp` stabilizes and does not diverge from the optimal case. The behavior of the $\lambda$ values is provided in the second row of Figure

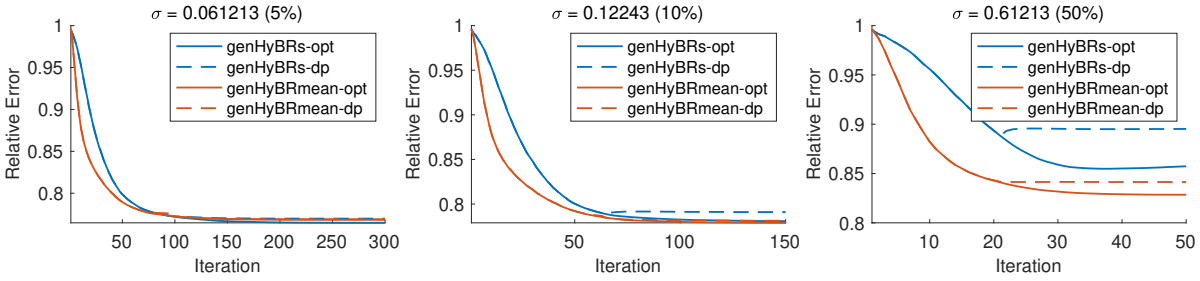


Figure 4.5: Compare relative errors of `genHyBRs` and `genHyBRmean` methods with 5%, 10%, and 50% noise levels.

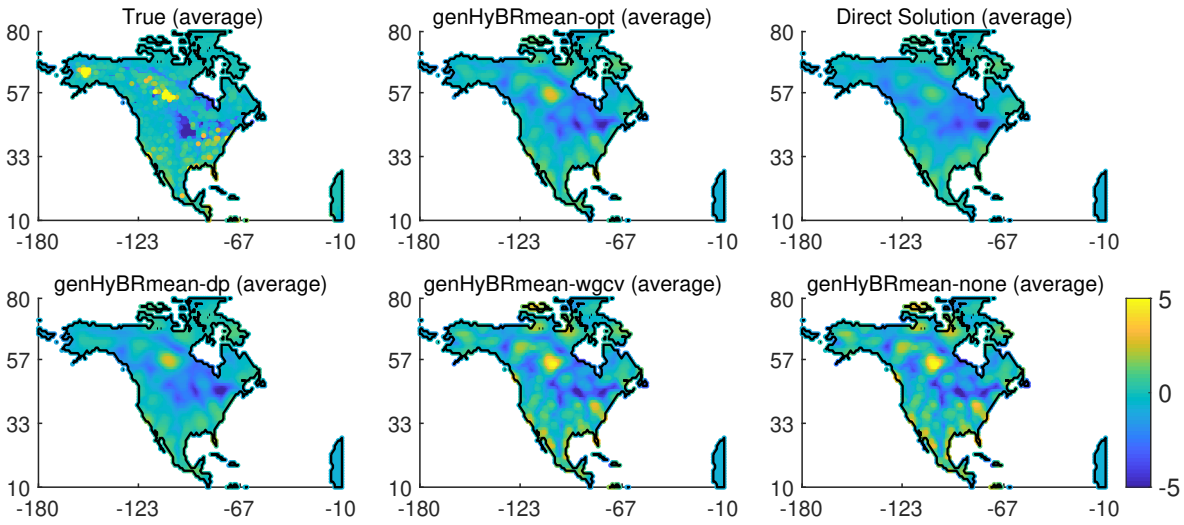


Figure 4.6: Reconstructions of average over 2-6 weeks. Since `genHyBRmean-upre` has a similar relative error with `genHyBRmean-none`, its reconstruction is omitted.

4.4. Corresponding to the relative errors, computed $\lambda$ values of `genHyBRmean-dp` are close to the optimal $\lambda$ values of `genHyBRmean-opt`. Since $\lambda$ values of `genHyBRmean-upre` are close to zero, the corresponding relative errors follow `genHyBRmean-none`. Figure 4.5 compares the relative errors between `genHyBRs` and `genHyBRmean`. For the converged cases, `genHyBRmean` attains smaller errors as the noise gets larger.

Figure 4.6 plots reconstructions of the true solution, the direct solution (i.e., obtained using

a current state-of-the art method), and the `genHyBRmean` solution. Since the direct method uses a regularization parameter that was tuned for $\sigma = 2$ as was done in previous work, the reconstructions are not very close to the true average map. However, the DP estimate of the average map is close to the `genHyBRmean-opt`. For `genHyBRmean-wgcv` and `genHyBRmean-none`, we provide the best reconstructions, but we can expect their estimations to be worse with additional iterations because of the semi-convergence behaviors that are showed in Figure 4.4.

In Figure 4.7, we provide approximations of the posterior standard deviation values for the average maps. Figure 4.8 shows the approximate posterior distribution of $\boldsymbol{\beta} \in \mathbb{R}^8$ with 95% credibility bounds.

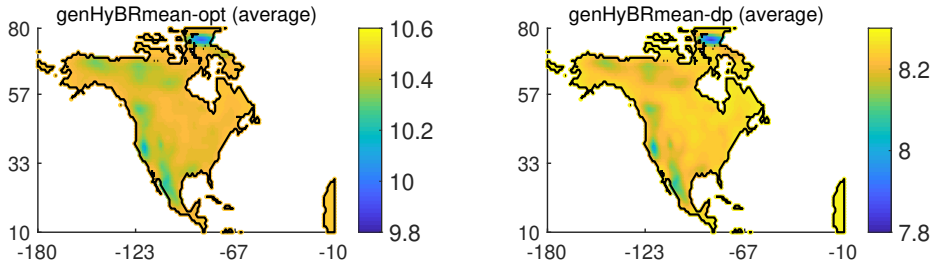


Figure 4.7: Posterior standard deviation of average over 2-6 weeks.

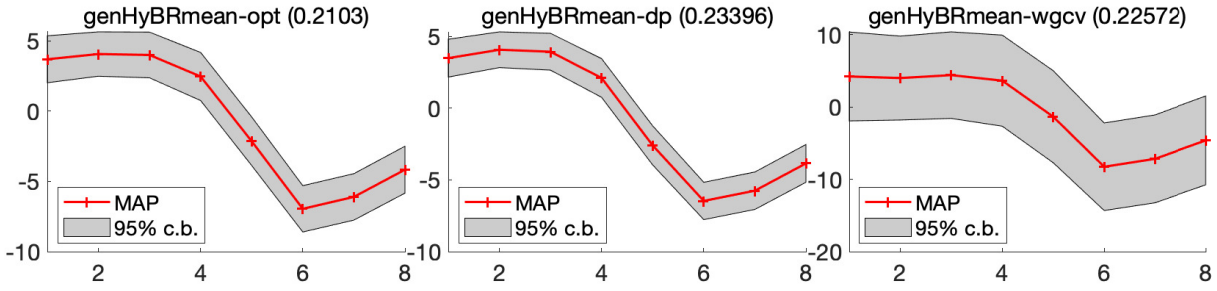


Figure 4.8: The elementwise 95% credibility bounds and MAPs computed from different regularization methods in 50% noise level. Each number in subtitle means the relative errors of $\boldsymbol{\beta}$.

# Chapter 5

# Computational Tools for Inversion and Uncertainty Estimation in Respirometry

The goal of this chapter is to develop practical mathematical and computational tools to advance reconstruction methodologies for the inverse problem of recovering signals in physiological systems from flow-through respirometry chambers. In many areas of biology and biomechanics, signals of interest cannot be measured directly, but instead must be estimated from indirect, noisy observations. For example, rates of oxygen consumption and $CO_2$ production are important for measuring energy expenditure associated with physiological phenomena and can promote understandings of energy regulation systems, but recovering such information generally requires indirect calorimetry using respiration chambers.

In flow-through respirometry systems, the goal is to determine the pattern of real instantaneous gas exchange of an animal that is put in a chamber. Air is pumped through the chamber and continuously mixes with the $CO_2$ and water vapor produced by the animal. Then, air that flows out of the chamber is brought to a gas analyzer that measures patterns of gas concentration. However, during this process the metabolic signals get distorted due to the washout kinetics. The recorded signal in the gas analyzer is a convolution of the true/instantaneous signal and the impulse response of the system which contains all characteristics

of the respirometry system. The instantaneous signals of interest can only be obtained by solving so-called "input estimation" or "inverse" problems [8, 43, 106]. This problem like any other deconvolution or input estimation problem is inherently ill-posed. Finding the instantaneous signal is particularly important if we study the synchrony between the metabolic signals and other physiological measurements such as locomotion, food or drug consumption, or circadian rhythms.

For problems where the impulse response function of the system is known, the inverse respirometry reconstruction problem can be formulated as a (1.1) that resembles the widely-studied problem of deconvolution. Some previous works that use Tikhonov regularization to solve the linear respirometry reconstruction problem include [43, 89, 90], among others. Since the forward model used in these regularized deconvolution methods is defined by the choice of the impulse response function, using an inaccurate impulse response function (e.g., one that is estimated experimentally) can result in significant degradation of the reconstruction accuracy. In flow through respirometry systems the pattern of the impulse response depends on the volume of the chamber, flow rate, size of the tubes between the chamber and gas analyzer, and even the size and location of the specimen in the chamber. Thus, it is important to consider methods that can either reconstruct the impulse response function or improve on a given impulse response function, while simultaneously reconstructing the desired signal. This problem is highly nonlinear and thus significantly more challenging to solve due to non-uniqueness of the solution. Indeed, joint reconstruction of the impulse response function and the physiological signal remains an open problem in the field of respirometry [23, 68].

In this chapter, we consider computational methods for both the linear and nonlinear respirometry reconstruction problems, with a particular emphasis on large-scale problems. We begin by describing the underlying mathematical model. Let $\mathbf{h} \in \mathbb{R}^n$ define the impulse

response function and let $\mathbf{s} \in \mathbb{R}^n$ contain the desired signal. Let $\mathbf{z} = \begin{bmatrix} \mathbf{s} \\ \mathbf{h} \end{bmatrix} \in \mathbb{R}^{2n}$, then the observed signal contained in $\mathbf{d} \in \mathbb{R}^n$ can be modeled as,

$$\mathbf{d} = f(\mathbf{z}) + \boldsymbol{\epsilon} \quad \text{with} \quad f(\mathbf{z}) = \mathbf{A}(\mathbf{h})\mathbf{s}, \tag{5.1}$$

where $\mathbf{A}(\cdot) : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ models the forward evolution process, and $\boldsymbol{\epsilon} \in \mathbb{R}^n$ represents noise or measurement errors. A common assumption is that the noise is independent and identically distributed from a Gaussian distribution with zero mean and variance $\sigma^{-2}$, i.e., $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^{-2}\mathbf{I})$. For a given $\mathbf{h}$, the respirometry forward model can be represented with matrix $\mathbf{A}(\mathbf{h})$, which is highly structured. Specific details regarding $\mathbf{h}$ and $\mathbf{A}$ will be provided (see Section 5.1). Given $\mathbf{d}$ and $\mathbf{A}(\cdot)$, the goal of the nonlinear (blind) respirometry problem is to reconstruct $\mathbf{z}$ (i.e., both $\mathbf{h}$ and $\mathbf{s}$). Oftentimes, (5.1) is referred to as a *separable* nonlinear inverse problem. Notice that if $\mathbf{h}$ is fixed, then we have a linear inverse problem.

There are many computational challenges to solving respirometry problems. First, due to ill-posedness, an appropriate choice of regularization should be incorporated for stable solution computation, and this goes hand-in-hand with the challenging task of selecting a suitable regularization parameter. More specifically, for classic variational regularization of the linear respirometry problem, solution approximations are obtained by solving optimization problems of the form,

$$\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s} - \mathbf{d}\|_2^2 + \frac{\lambda_p}{2}\Omega(\mathbf{s}) \tag{5.2}$$

where $\lambda_p > 0$ is a regularization parameter and $\Omega(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is a regularization functional determined by the choice of the prior. Previous studies on respirometry reconstruction employ standard Tikhonov regularization where $\Omega(\mathbf{s}) = \|\mathbf{s}\|_2^2$, and practitioners manually tune the regularization parameter $\lambda_p$. Selecting a suitable regularization parameter involves

finding a good balance between introducing bias in the solution and preserving fidelity to the system and the observed data. This can be an expensive and time consuming task that requires multiple solves for various parameter choices [43, 89, 90]. Second, iterative methods provide an efficient approach to handle very large problems (e.g., large signals with many unknown parameters), but preconditioning techniques are needed to accelerate convergence and these preconditioners need to be tailored to the structure of matrix $\mathbf{A}$. Third, it may be desirable to go beyond obtaining reconstructions to also provide uncertainty estimates for reconstructions, but this process often requires many expensive solves. The fourth, and most difficult, challenge is that methods need to be developed to handle nonlinearity in the problem (e.g., when the impulse response function contains errors or uncertainty). Due to difficulties of the nonlinear problem, previous respirometry studies do not formally consider this scenario. We describe various approaches to address these challenges.

We begin with a description of the mathematical set-up for the respirometry problem. We describe a Bayesian formulation of the linear respirometry problem and describe various tools for regularization and uncertainty quantification. We propose preconditioners for accelerating iterative methods. Then, we describe nonlinear optimization methods that can be used for nonlinear respirometry reconstruction. Numerical results for simulated and real respirometry data are provided to demonstrate the performance and potential of our proposed approaches.

## 5.1 Mathematical Problem Set-up

We begin with a mathematical description of the forward model underlying respirometry. In a continuous input estimation scenario, we assume that the system is linear and time-invariant, such that the output signal can be written as a convolution of the instantaneous

input signal and the impulse response function of the system. More precisely, the output of the system at time $t$ is given by

$$d(t) = \int_0^t h(t-\tau)s(\tau)d\tau$$

where $s(\tau)$ describes the state of the system at time $\tau$ and $h$ is the impulse response function. In a discrete formulation, we take observations at uniform time points $0 = t_0 < t_1 < \ldots < t_n < \infty$ denoted as

$$d_k = \sum_{i=0}^{k-1} h(t_k - t_i)s_i\delta t + \epsilon_k, \quad \text{for} \quad k = 1, \ldots, n$$

where $d_k$ and $s_k$ describe the output and input signals respectively at time $t_k$, $\delta t = t_{k+1} - t_k$ is the sampling interval, and $\epsilon_k \sim \mathcal{N}(0, \sigma^{-2})$ represent errors in the data. In matrix notation, we have the discrete respirometry problem,

$$\mathbf{d} = \mathbf{A}\mathbf{s} + \boldsymbol{\epsilon}, \tag{5.3}$$

where

$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \mathbf{s} = \begin{bmatrix} s_0 \\ s_1 \\ \vdots \\ s_{n-1} \end{bmatrix}, \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}, \text{ and } \mathbf{A} = \delta t \begin{bmatrix} h(\delta t) & 0 & \cdots & 0 \\ h(2\delta t) & h(\delta t) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ h(n\delta t) & \cdots & h(2\delta t) & h(\delta t) \end{bmatrix}.$$

Notice that if we let $\mathbf{h} = \delta t \begin{bmatrix} h(\delta t) \\ \vdots \\ h(n\delta t) \end{bmatrix} \in \mathbb{R}^n$ be the discretized impulse response function, then $\mathbf{A} = \mathbf{A}(\mathbf{h})$ is a lower-triangular Toeplitz matrix with $\mathbf{h}$ as the first column and we get a problem of the form (5.1).

For most respirometry problems, the impulse response function is not known in advance, but must be estimated experimentally. In practice, a $CO_2$ pulse is injected into an empty chamber for a short time (e.g., 0.5 seconds) and the normalized recorded output serves as the impulse response function. Due to various experimental errors and imperfections, this process may result in an imprecise estimate of the impulse response function. Nevertheless, this is the standard process used in practice. One could consider using blind deconvolution methods to solve for $\mathbf{h}$ and $\mathbf{s}$ simultaneously given $\mathbf{d}$, but this is a severely ill-posed problem where the main challenge is the existence of many local minimizers. Furthermore, the number of unknown variables doubles (i.e., $2n$ total unknowns in $\mathbf{s}$ and $\mathbf{h}$). We will assume that partial information about the impulse response function is available and consider the so-called *semi-blind deconvolution* problem. More specifically, we assume that the delay and the support of the impulse response function are known. That is, let $s \in \mathbb{Z}^+$ denote the support and $d \in \mathbb{Z}^+$ represent the delay, then we assume that $\mathbf{h}$ has the form

$$\mathbf{h} = [\ \underbrace{0\ \cdots\ 0}_{d}\ \bar{\mathbf{h}}^\top\ \underbrace{0\ \cdots\ 0}_{n-d-s}\ ]^\top \quad \text{where} \quad \bar{\mathbf{h}} = \delta t[\ h_{d+1}\ \cdots\ h_{d+s}\ ]^\top \in \mathbb{R}^s$$

has nonzero elements and $h_k = h(k\delta t)$. With these minor assumptions on the impulse response function, the nonlinear problem (5.1) reduces to smaller system given by,

$$\bar{\mathbf{d}} = \bar{\mathbf{A}}(\bar{\mathbf{h}})\bar{\mathbf{s}} + \bar{\boldsymbol{\epsilon}} \tag{5.4}$$

where $\mathbf{d} = [\ \underbrace{\bar{\bar{\mathbf{d}}}^\top}_{d}\ \bar{\mathbf{d}}^\top\ ]^\top$ and $\mathbf{s} = [\bar{\mathbf{s}}^\top\ \underbrace{\bar{\bar{\mathbf{s}}}}_{d}\ ]^\top \in \mathbb{R}^{n-d}$ with $\bar{\mathbf{d}}, \bar{\mathbf{s}}, \bar{\boldsymbol{\epsilon}} \in \mathbb{R}^{n-d}$, $\bar{\mathbf{A}}(\cdot) : \mathbb{R}^s \to \mathbb{R}^{(n-d)\times(n-d)}$ with $\bar{\mathbf{A}}(\bar{\mathbf{h}})$ being a lower-triangular Toeplitz matrix with $[\bar{\mathbf{h}}^\top\ \underbrace{0\ \cdots\ 0}_{n-d-s}\ ]^\top$ as the first column. Note that $\bar{\bar{\mathbf{d}}}$ and $\bar{\bar{\mathbf{s}}}$ represent the clipped elements of $\mathbf{d}$ and $\mathbf{s}$ respectively in Fig 5.1. The time response of a chamber is roughly about $V/F$, where $F$ is the air inflow rate and $V$ is the volume of the chamber [8, 89]. If there is no information about the support, we can assume it to be approximately 3 to 5 times the time response.

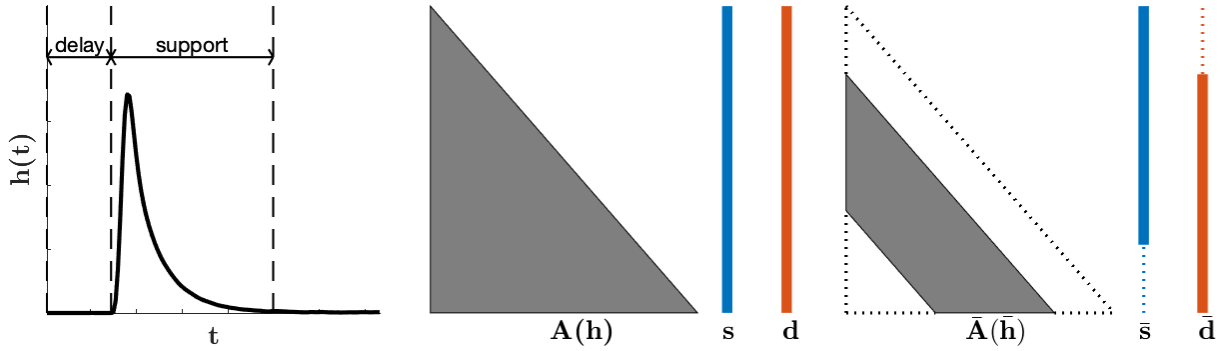


Figure 5.1: Illustration of the delay and support of the impulse response function $h(t)$ used in respirometry. The middle and right plots demonstrate the change in structure from the original system in (5.1) to the reduced system in (5.4) that occurs due to the inclusion of delay and support assumptions.

An example of the delay and support of an impulse response function used in respirometry is provided in Fig 5.1, along with an illustration of the impact on the resulting system due to the delay and support. Notice that since $\mathbf{s}$ is replaced by $\bar{\mathbf{s}}$ in the reduced system, the tail of $\mathbf{s}$ is not being reconstructed. However, this is not a significant loss since it is common practice to ignore the final points of the reconstruction even in the non-blind case. The physical reason is that the released $CO_2$ from the animal at the end of the experiment does not completely show up in our observed measurements since we have stopped recording before those $CO_2$ particles leave the chamber and reach the gas analyzer.

We close this section with a few remarks. First of all, regarding our assumption of knowl-

edge of the delay and support of the impulse response function, we can typically obtain good estimates of these values from the experimental impulse response function or from the respirometry problem set-up. Second, contrary to previous respirometry reconstruction methods, we do not assume any functional form for $\bar{\mathbf{h}}$. Third, we will describe various techniques to solve the reduced nonlinear system, and numerical results show that reconstructions are not sensitive to the choice of these parameters.

## 5.2 Respirometry with *Known* Impulse Response Function

In this section, we assume that $\mathbf{h}$ is fixed and focus on efficient computational tools for solving the linear respirometry reconstruction problem (5.3) and for performing subsequent uncertainty quantification. Since the linear reconstruction problem is ill-posed, some form of regularization must be included. We employ a Bayesian framework, which is a statistically robust way to include prior knowledge by treating $\mathbf{s}$ as a random variable. Furthermore, the Bayesian approach provides a natural framework for performing uncertainty quantification. Good overviews on Bayesian inverse problems, statistical inverse problems, and computational uncertainty quantification can be found here [7, 22, 67].

### Regularization and uncertainty quantification

Consider the stochastic extension of (5.3),

$$D = \mathbf{H}S + E$$

where $S$, $D$ and $E$ are random variables and $\mathbf{H}$ is deterministic. We assume that $S$ and $E$ are mutually independent and that the prior density function of $S$ is given by $\pi_{\text{prior}}(\mathbf{s})$ and the conditional density function of $D$ given $S$ is given by $\pi_{\text{like}}(\mathbf{d} \mid \mathbf{s})$. Using Bayes' Theorem, the posterior probability density function can be written as (2.2) assuming the marginal density function $\pi(\mathbf{d}) \neq 0$.

A key component of the Bayesian formulation is the choice of the prior distribution function $\pi_{\text{prior}}(\mathbf{s})$, which incorporates any knowledge about the solution $\mathbf{s}$ prior to data being collected. We consider two priors: a Gaussian prior and a Laplace prior. In both cases, we assume that the observation error can be modeled as $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^{-2}\mathbf{I})$, and thus the likelihood function can be written as

$$\pi_{\text{like}}(\mathbf{d}|\mathbf{s}) \;\propto\; \exp\left(-\frac{\sigma^2}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2\right).$$

We will see that a nice connection between the Bayesian and classical formulations for inverse problems is that various point estimators in the Bayesian framework coincide with classic regularized solutions that are obtained by solving optimization problems of the form (5.2).

**2-norm regularization.** Gaussian priors are commonly used, and these priors are defined by a known mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$ and known symmetric positive definite covariance matrix $\mathbf{Q}$, i.e., $\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}, \lambda^{-2}\mathbf{Q})$. This case is a particular case of (2.2). Hence, the posterior density function $\pi_{\text{post}}$ is a Gaussian distribution as (2.8) where $\mathbf{R} = \sigma^{-2}\mathbf{I}$.

Let $\mathbf{Q}^{-1} = \mathbf{L}_{\mathbf{Q}}^{\top}\mathbf{L}_{\mathbf{Q}}$ be a symmetric factorization (e.g., a Cholesky or eigenvalue decomposition), then the MAP estimate is the solution to the following optimization problem,

$$\mathbf{s}_{\text{MAP}} = \arg\min_{\mathbf{s}} \frac{\sigma^2}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{L}_{\mathbf{Q}}(\mathbf{s}-\boldsymbol{\mu})\|_2^2 \tag{5.5}$$

$$= \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 + \frac{\lambda_2^2}{2}\|\mathbf{L}_{\mathbf{Q}}(\mathbf{s}-\boldsymbol{\mu})\|_2^2 \tag{5.6}$$

which is commonly known as Tikhonov regularization where $\lambda_2 = \frac{\lambda}{\sigma}$. Thus, the Tikhonov regularized solution is the point estimate that corresponds to the maximum value of the posterior density function, or equivalently the minimizer of its negative log. Since the posterior density function is Gaussian, variance estimates for the solution can be obtained by computing the diagonal entries of $\mathbf{\Gamma}_{\text{post}}$. Furthermore, samples from the posterior can be obtained using efficient Krylov subspace methods [19] and approximation methods were described in Section 2.4.

Note that in the inverse problems community, $\mathbf{L_Q}$ is often referred to as the regularization matrix and is chosen to force smoothness of the desired solution. There are many choices for the regularization matrix $\mathbf{L_Q}$. In respirometry common choices for $\mathbf{L_Q}$ include the identity matrix $\mathbf{L_Q} = \mathbf{I}$ or a discretization of the derivative operator where $\mathbf{L_Q}$ is a lower triangular Toeplitz matrix with $[1 \ -2 \ 1 \ 0 \ \dots \ 0]^\top$ as the first column or $[1 \ -1 \ 0 \ \dots \ 0]^\top$ as the first column [89]. The parameter selection methods of 2-norm regularization were described in Section 2.3.

**1-norm regularization.** An alternative assumption to a Gaussian prior is a Laplace prior, where the signal is independent and identically Laplace distributed,

$$s_i \sim \text{Laplace}(0, \delta^{-1}), \quad i = 1, 2, \dots, n \tag{5.7}$$

where the probability density function for a Laplace distribution is given by $\pi(s) = \frac{\delta}{2}\exp(-\delta|s|)$ for $\delta > 0$. Thus, using the assumption of independence, the prior corresponding to assumption (5.7) can be written as

$$\pi_{\text{prior}}(\mathbf{s}) \ \propto \ \exp(-\delta\|\mathbf{s}\|_1),$$

where $\|\cdot\|_1$ is the 1-norm of a vector. The posterior density function $\pi_{\text{post}}$ is given by

$$\pi_{\text{post}}(\mathbf{s} \mid \mathbf{d}) \propto \exp\left(-\frac{\sigma^2}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 - \delta\|\mathbf{s}\|_1\right). \tag{5.8}$$

Notice that the posterior is no longer Gaussian; however, we can use various tools to explore the posterior. The MAP estimate corresponds to the mode of the posterior distribution and is given by

$$\begin{aligned}
\mathbf{x}_{\text{MAP}} &= \text{argmax}_{\mathbf{s}}\, \pi_{\text{post}}(\mathbf{s} \mid \mathbf{d}) && (5.9)\\
&= \arg\min_{\mathbf{s}} \frac{\sigma^2}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 + \delta\|\mathbf{s}\|_1 && (5.10)\\
&= \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 + \frac{\lambda_1}{2}\|\mathbf{s}\|_1, && (5.11)
\end{aligned}$$

which is an $\ell_1$ regularized problem (5.2) with $\Omega(\cdot) = \|\cdot\|_1$ and $\lambda_1 = \frac{2\delta}{\sigma^2}$.

It is common to use regularization terms of the form $\Omega(\mathbf{s}) = \|\mathbf{s}\|_1$ in signal and imaging processing, since these regularizers enforce sparsity in the desired parameters. The main computational difficulty with these regularizers is the absolute value, which has a discontinuous first derivative at zero, causing challenges for optimization algorithms. For small to medium size problems, it is well known that the problem can be reformulated as a quadratic programming problem, and standard optimization software packages can be used. However, the number of unknowns in the reformulated problem doubles, making this approach unrealistic for large-scale problems. A more computationally appealing approach is to solve $\ell_1$ regularized problems using the proximal gradient. That is, methods such as the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [11] use iterative techniques to solve

$$\mathbf{s}_{\lambda_1} = \arg\min_{\mathbf{s}} \frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2 + \frac{\lambda_1}{2}\|\mathbf{s}\|_1. \tag{5.12}$$

A summary of FISTA with a constant step size is provided in Algorithm 3.

**Algorithm 3** FISTA with constant stepsize

Choose $\lambda_1$.
Compute $L$, a Lipschitz constant of $\frac{1}{2}\|\mathbf{A}\mathbf{s}-\mathbf{d}\|_2^2$.
Set $\mathbf{y}_1 = \mathbf{s}_0 = \mathbf{0} \in \mathbb{R}^n$ and $t_1 = 1$.
**for** $k = 1, 2, \ldots$ **do**

$$\mathbf{s}_k = \arg\min_{\mathbf{s}} \frac{L}{2}\left\|\mathbf{s} - \left(\mathbf{y}_k - \frac{1}{L}\mathbf{A}^\top(\mathbf{A}\mathbf{y}_k - \mathbf{d})\right)\right\|_2^2 + \lambda_1\|\mathbf{s}\|_1,$$

$$t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2},$$

$$\mathbf{y}_{k+1} = \mathbf{s}_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(\mathbf{s}_k - \mathbf{s}_{k-1}).$$

**end for**

In addition to the choice of the regularization parameter $\lambda_1$ that must be selected in advance, the Lipschitz constant $L$ that depends on the maximum eigenvalue of $\mathbf{A}^\top\mathbf{A}$ must be estimated. It can be difficult to compute $L$ when $n$ is large, but an approach using backtracking was described in [11]. Similar to FISTA, the Sparse Reconstruction by Separable Approximation (SpaRSA) method [111] is an iterative method that can be used to solve (5.12), which uses a sequence of smooth approximations of the 1-norm. Although more general regularization terms can be included, SpaRSA requires more user-defined input parameters so we do not consider it here. Another class of methods for solving the $\ell_p$ regularized problem is based on flexible Krylov methods that use iterative techniques with flexible preconditioning within a hybrid framework to improve the solution subspace. Methods such as FLSQR-R can be used to solve $\ell_p$-regularized problems where $1 \leq p < 2$, see [34].

Various methods for solving the $\ell_1$-regularized problem can be used to approximate the MAP estimate, but subsequent uncertainty quantification for this case is significantly more challenging. Although the posterior (5.8) is not Gaussian, we can approximate the posterior with a Gaussian at the maximum a posterior (MAP) estimate using a linearization approach

[18]. Another approach to efficiently obtain samples from the posterior in this case is to use a change of variables or transformation to turn a non-Gaussian distribution into a Gaussian one, as described in [108]. More specifically, the transformation is defined by

$$\mathbf{s} = g(\mathbf{w}) := \begin{bmatrix} g_{1D}(w_1) & \cdots & g_{1D}(w_n) \end{bmatrix}^\top \tag{5.13}$$

where

$$g_{1D}(w) = \mathcal{L}^{-1}\mathcal{G}(w) \tag{5.14}$$

with $\mathcal{L}$ being the cumulative density function (cdf) of the Laplace distribution and $\mathcal{G}$ being the cdf of a Gaussian distribution. With this definition, $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and the transformation $\mathbf{w} = g^{-1}(\mathbf{s})$ generates

$$p(\mathbf{w}) = p(g(\mathbf{w}))|\mathbf{J}_g(\mathbf{w})| \tag{5.15}$$

where

$$\mathbf{J}_g(\mathbf{w}) = \text{diag}(g'_{1D}(w_1), \ldots, g'_{1D}(w_n)). \tag{5.16}$$

From these transformations and from (5.8), we obtain

$$p(\mathbf{w}|\mathbf{d}) \quad \propto \quad \exp\left( -\frac{1}{2} \left\| \begin{bmatrix} \mathbf{A}(g(\mathbf{w})) \\ \sqrt{\lambda_1}\mathbf{w} \end{bmatrix} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|_2^2 \right) \tag{5.17}$$

Hence, we can generate samples from $\mathbf{w}$ and transform these samples to get samples of $p(\mathbf{s}|\mathbf{d})$ via $\mathbf{s} = g(\mathbf{w})$. Although there are some known challenges with this approach, we found that it worked well for the respirometry reconstruction problem.

By following a Bayesian framework for inversion, we have established a natural framework not only for incorporating prior knowledge but also for quantifying solution uncertainties. In terms of software, IRTools [50] is a comprehensive package that contains many iterative

regularization routines for solving inverse problems along with various test problems. To the best of our knowledge there is no unified software package for performing UQ for inverse problems. We point the interested reader to the following book and associated codes [7].

### Accelerating Iterative Methods for Signal Reconstruction

In the previous section, we considered various regularization techniques for solving the linear respirometry reconstruction problem. Next, we focus on Tikhonov regularization, and we investigate efficient methods to accelerate the convergence of iterative methods when used to compute a solution,

$$\mathbf{s}_{\text{Tik}} = \arg\min_{\mathbf{s}} \frac{1}{2} \|\mathbf{A}\mathbf{s} - \mathbf{d}\|_2^2 + \frac{\lambda_2^2}{2} \|\mathbf{L}_{\mathbf{Q}}(\mathbf{s} - \boldsymbol{\mu})\|_2^2 \tag{5.18}$$

$$= \left(\mathbf{A}^\top \mathbf{A} + \lambda_2^2 \mathbf{Q}^{-1}\right)^{-1} (\mathbf{A}^\top \mathbf{d} + \lambda_2^2 \mathbf{Q}^{-1} \boldsymbol{\mu}), \tag{5.19}$$

where (5.19) comes from setting the gradient of the function in (5.18) equal to zero. For small problems, constructing the matrix and the solution in (5.19) is computationally feasible, and many of the previous works in respirometry reconstruction follow this approach. For example, Tikhonov methods described in [90] could be used here. However, more sophisticated iterative techniques should be used for large-scale problems.

Iterative methods, in particular Krylov subspace methods, are computationally attractive because each iteration only requires one matrix-vector-multiplication with $\mathbf{A}$ and perhaps $\mathbf{A}^\top$ [53, 96]. Thus, the matrix representing the respirometry forward model never needs to be constructed, but instead can be accessed via operations or function evaluations. However, it is widely known, especially in the numerical linear algebra community, that preconditioning is a very important tool for accelerating convergence and improving the robustness of Krylov methods [12]. The basic idea of preconditioning is to modify the problem by improving the

spectrum of the problem (so that eigenvalues or singular values of the preconditioned system are clustered around one and bounded away from zero), thereby accelerating the convergence of iterative methods.

For simplicity of presentation, we describe preconditioning techniques for the unregularized problem (i.e., $\lambda_2 = 0$) and focus on developing a good preconditioner for the respirometry matrix $\mathbf{A}$ that can exploit the special structure of these matrices. We first describe the general idea underlying preconditioning and then describe how to apply preconditioning to the regularized problem.

Assume that we have a preconditioner $\mathbf{M} \in \mathbb{R}^{n\times n}$ such that $\mathbf{M}^{-1} \approx \mathbf{A}^{-1}$ and solving systems involving $\mathbf{M}$ can be done easily and quickly. Then rather than solve (5.18), consider solving the right-preconditioned problem,

$$\min_{\mathbf{y}} \left\|\mathbf{A}\mathbf{M}^{-1}\mathbf{y} - \mathbf{b}\right\|_2^2 \quad \text{where} \quad \mathbf{y} = \mathbf{M}\mathbf{s},$$

or the left-preconditioned problem,

$$\min_{\mathbf{s}} \left\|\mathbf{M}^{-1}\mathbf{A}\mathbf{x} - \mathbf{M}^{-1}\mathbf{b}\right\|_2^2$$

using an iterative method such as the conjugate gradient for least-squares (CGLS) method [13]. Notice that each iteration requires one matrix-vector multiplication with $\mathbf{M}^{-1}\mathbf{A}$ and its transpose. Typical choices for $\mathbf{M}$ are based on incomplete matrix factorizations or multigrid methods [12]. However, for the respirometry problem, these approaches are not ideal for two main reasons. First, matrix $\mathbf{A}$ is large and construction of $\mathbf{A}$ is not possible. Instead, we access it via function evaluations. Second, $\mathbf{A}$ is severely ill-conditioned, so we would like to approximate a regularized pseudoinverse of $\mathbf{A}$ rather than $\mathbf{A}^{-1}$. Obtaining a good approximation of $\mathbf{A}^{-1}$ would result in very fast convergence to the undesired inverse solution.

For respirometry reconstruction problems, we propose various preconditioners that can be used to accelerate the convergence of iterative methods. Recall that $\mathbf{A}$ is a Toeplitz matrix with $\mathbf{h}$ as the first column. For large-scale problems, we have constructed an object class in MATLAB called `convMatrix.m`, where matrix-vector and matrix-transpose-vector operations with $\mathbf{A}$ are treated as function evaluations. In particular, `convMatrix` calls MATLAB's `conv` function to do convolution and then extracts the appropriate signal length.

Next, we describe how to exploit the Toeplitz structure of $\mathbf{A}$ to build a good preconditioner for the respirometry problem. Since circulant matrices provide good approximations to Toeplitz matrices [24, 25, 26, 58] and circulant matrices are diagonalized by the discrete Fourier transform, we propose to use a *circulant* matrix $\mathbf{M}$ such that the lower triangular part of $\mathbf{M}$ matches the lower triangular part of $\mathbf{A}$. Then since $\mathbf{M}$ can be diagonalized by the discrete Fourier transform, we can write

$$\mathbf{M} = \mathbf{F}^{*}\mathbf{\Theta}\mathbf{F} \tag{5.20}$$

where $\mathbf{F}$ represents the Fourier transform and $\mathbf{\Theta}$ is a diagonal matrix with eigenvalues computed as

```
theta = fft(h);
```

where $\mathbf{h}$ contains the impulse response function scaled by the sampling rate (i.e., this corresponds to the first column of $\mathbf{A}$). Thus, we can apply the preconditioner to any vector $\mathbf{y}$ as

$$\mathbf{M}^{-1}\mathbf{y} = \mathbf{F}^{*}\mathbf{\Theta}^{-1}\mathbf{F}\mathbf{y},$$

which corresponds to the following commands in MATLAB

```
ifft(fft(y)./theta);.
```

Notice that since $\mathbf{M}$ is likely ill-conditioned, $\mathbf{\Theta}$ has very small values on the diagonal which can result in erroneous computations. A small modification to the preconditioner can be done, where $\mathbf{\Theta}$ is replaced with a diagonal matrix $\widehat{\mathbf{\Theta}}$ with better spectral properties (i.e., removing small eigenvalues of $\mathbf{\Theta}$). For simplicity, we can use a TSVD-like preconditioner where $\mathbf{M} = \mathbf{F}^*\widehat{\mathbf{\Theta}}\mathbf{F}$ with diagonal entries of $\widehat{\mathbf{\Theta}}$ being

$$\hat{\theta}_i = \begin{cases} \theta_i, & \text{if } |\theta_i| \geq \tau \\ 1, & \text{else} \end{cases} \tag{5.21}$$

for some predetermined tolerance parameter $\tau$. Because the preconditioner inherently includes regularization, we avoid the danger of the preconditioner inadvertently magnifying the noise before a solution can be computed. Similar to the previous discussion about avoiding the construction of $\mathbf{A}$, we remark that construction of the preconditioner is also not advised. We have written an object class called `precMatrix.m` that can work with the preconditioner implicitly. The preconditioner for respirometry can be accessed using the MATLAB command: `M = precMatrix(h,tau);` where $\mathbf{h}$ contains the impulse response function without delay and $\tau$ is the tolerance parameter. We also describe an automatic approach to estimate $\tau$. Consider the approximate problem $\mathbf{Ms} = \mathbf{d}$. Since the diagonalization of $\mathbf{M}$ is computable, we can use the GCV method to efficiently compute a regularization parameter or truncation tolerance for TSVD [59]. This truncation tolerance can be used to define the preconditioner. Furthermore, we remark that $\mathbf{M}$ corresponds to convolution with the same impulse response function, where periodic boundary conditions are assumed. Thus, if $\mathbf{H}$ corresponds to periodic boundary conditions and $\tau$ is the smallest singular value, then $\mathbf{M} = \mathbf{A}$.

For a small respirometry example, we provide in Fig 5.2 the spectrum of $\mathbf{A}$ along with the spectrum of the preconditioned system $\mathbf{M}^{-1}\mathbf{A}$ for $\tau = 1.7 \times 10^{-3}, 1.7 \times 10^{-2}$, and $1.7 \times 10^{-1}$.

The GCV selected parameter for this example was $9.0 \times 10^{-2}$. Notice that the singular values for the preconditioned systems are clustered around 1. This is a very desirable property for the fast convergence of Krylov subspace methods [96]. However, the spectrum of the preconditioned system relies heavily on the choice of $\tau$. For small values of $\tau$, $\mathbf{M}$ clusters too many of the small singular values so that the preconditioned system will be very ill-posed. On the other hand, for larger values of $\tau$, only a few singular values are clustered so more iterations would be required. If $\tau$ is greater than or equal to the largest singular value of $\mathbf{A}$, then $\mathbf{M} = \mathbf{I}$ and we have no preconditioning.

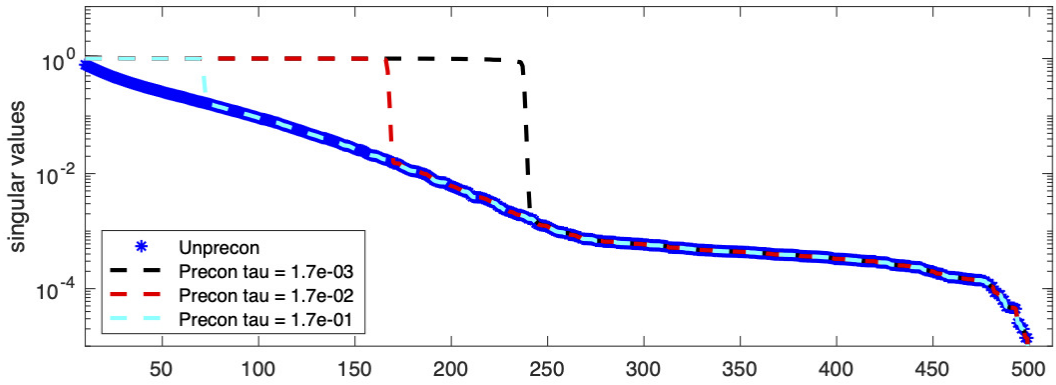


Figure 5.2: Spectrum of the unpreconditioned and preconditioned respirometry matrices for various choices of $\tau$. Note the desirable clustering of the larger eigenvalues, which results in fast convergence of iterative methods.

Thus far, we have focused on developing preconditioners that exploit the structure of $\mathbf{A}$ and described how to use these preconditioners to accelerate iterative methods. If one wishes to use these preconditioners for solving regularized problems (e.g.(5.18)), then a simple extension can be made. That is, one can solve preconditioned problem,

$$\min_{\mathbf{y}} \left\| \mathbf{A}\mathbf{M}^{-1}\mathbf{y} - \mathbf{d} \right\|_2^2 + \lambda_2^2 \left\| \mathbf{L}_\mathbf{Q}(\mathbf{M}^{-1}\mathbf{y} - \boldsymbol{\mu}) \right\|_2^2 \quad \text{where} \quad \mathbf{y} = \mathbf{M}\mathbf{s},$$

or

$$\min_{\mathbf{s}} \left\|\mathbf{M}^{-1}\mathbf{A}\mathbf{s} - \mathbf{M}^{-1}\mathbf{d}\right\|_2^2 + \lambda_2^2 \left\|\mathbf{L}_{\mathbf{Q}}(\mathbf{s} - \boldsymbol{\mu})\right\|_2^2 .$$

In general, proper preconditioning can be a very important, albeit delicate, task especially for inverse problems.

## 5.3 Respirometry with *Unknown* Impulse Response Function

Thus far, we have focused on the linear respirometry problem where the impulse response function is assumed known. However, this is not true in realistic experiments, where the impulse response function must be estimated. Given measured respirometry data, estimating both the impulse response function and the unknown signal simultaneously is a very challenging problem. Nevertheless, there are various reasons why we may want to consider a joint estimation approach. First, the estimated impulse response function which is obtained using a short burst of $CO_2$ in an empty chamber will likely contain errors. Second, although the respirometry systems is calibrated at construction, parameters may change over time and these changes are not accounted for without a full recalibration of the machine. Third, the impulse response depends on the size and location of the animal, which could change across experiments. For these and other reasons, we are interested in methods that can solve the nonlinear respirometry reconstruction problem. It is worth mentioning that we tried some off-the-shelf blind deconvolution methods such as MATLAB's deconvblind function, but found that reconstructions were very poor; thus motivating us to consider alternative approaches.

The impulse response function, which models the reaction of the system to a very short unit

impulse in a linear time-invariant system, is a key component of respirometry reconstruction. Conventional methods such as the Z-transform method described in [8] use an impulse response function defined by an exponential function, e.g., $h(t) = \alpha e^{-\beta t}$ where $\alpha$ and $\beta$ are parameters defined by the flow rate and chamber volume. In [89], the authors experimentally showed that for many chambers and flow rates, the impulse response has the form $h(t) = \alpha t^m e^{-\beta t}$ where $\alpha$, $m$, and $\beta$ are parameters of the system. Although the parameters for the impulse response function must be estimated, numerical experiments showed that this function performed better than the exponential function. For the methods described in this section, we do not enforce a functional form for the impulse response function. Instead we impose other less restrictive constraints on the impulse response function, and develop computational methods for nonlinear respirometry reconstruction, where both the signal $\mathbf{s}$ and the impulse response function $\mathbf{h}$ can be estimated simultaneously from the data. The goal is to solve *nonlinear* optimization problem,

$$\min_{\mathbf{s},\mathbf{h}} \|\mathbf{A}(\mathbf{h})\mathbf{s} - \mathbf{d}\|_2^2 + \lambda_p \Omega(\mathbf{s}) + \lambda_{\mathbf{h}} \|\mathbf{h}\|_2^2 \quad \text{s.t.} \quad \mathbf{h} \geq \mathbf{0} \quad \text{and} \quad \sum_{i=1}^{n} h_i \delta t = 1 \tag{5.22}$$

where $\lambda_{\mathbf{h}}$ is a regularization parameter for $\mathbf{h}$. Compared to problem (5.3), we have a nonlinear model represented by $\mathbf{A}(\mathbf{h})$, and we have various additional constraints on $\mathbf{h}$. These constraints include an additional Tikhonov regularization term for $\mathbf{h}$ to enforce smoothness, a nonnegativity constraint, and a mass preserving constraint to force the computed impulse response function to sum to 1. This last constraint corresponds to forcing the integral of the impulse response function to be 1 in the continuous framework.

Before we describe computational methods to solve nonlinear constrained optimization problem (5.22), we provide an example to illustrate why solving the nonlinear blind reconstruction problem is significantly more difficult. The main concern is the existence of multiple minimizers. That is, without additional constraints, there are multiple solution pairs $(\mathbf{s}, \mathbf{h})$ that

give small values of the data fit term in the objective function. The plots in Fig 5.3 show that under the convolution operation, two very different pairs $(\mathbf{s}, \mathbf{h})$ can result in nearly the same observation. Thus, methods for numerical optimization can easily get trapped in local minimizers. Including additional constraints can help with this problem. We observed that the choice of regularization for $\mathbf{s}$, i.e., the choice of $\Omega(\mathbf{s})$, was important. In particular, reconstructions obtained using Tikhonov regularization in the nonlinear framework resulted in significantly smaller residual error norms, which negatively impacted the convergence. However, using $\Omega(\mathbf{s}) = \|\mathbf{s}\|_1$ resulted in faster convergence and better reconstructions.

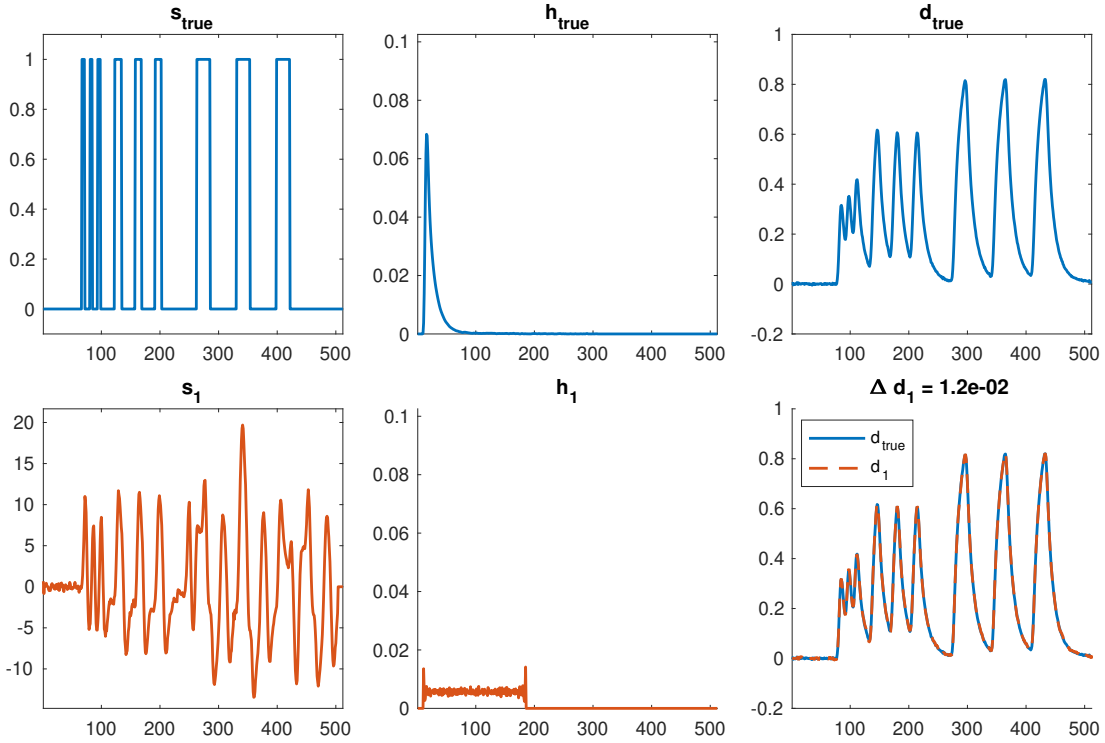


Figure 5.3: Illustration of the non-uniqueness problem in blind respirometry reconstruction. Both sets of parameters in $\mathbf{s}$ and $\mathbf{h}$ result in similar observed measurements in $\mathbf{d}$. The result in the second row corresponds to using Tikhonov regularization for $\mathbf{s}$ and solving $\mathbf{h}$ using alternating optimization.

Next we describe a computationally efficient method to solve (5.22). First, following the model described in (5.4), we assume that the delay and support of $\mathbf{h}$ are known and refor-

mulate problem (5.22) as

$$\min_{\bar{\mathbf{s}},\bar{\mathbf{h}}} \|\bar{\mathbf{A}}(\bar{\mathbf{h}})\bar{\mathbf{s}} - \bar{\mathbf{d}}\|_2^2 + \lambda_1\|\bar{\mathbf{s}}\|_1 + \lambda_{\bar{\mathbf{h}}}\|\bar{\mathbf{h}}\|_2^2 \quad \text{s.t.} \quad \bar{\mathbf{h}} \geq \mathbf{0} \quad \text{and} \quad \sum_{i=d+1}^{d+s} h_i \delta t = 1 \tag{5.23}$$

where $\lambda_1$, $\lambda_{\bar{\mathbf{h}}}$ are regularization parameters for $\bar{\mathbf{s}}, \bar{\mathbf{h}}$ respectively. For large scale problems, matrix-vector multiplications $\bar{\mathbf{A}}(\bar{\mathbf{h}})\bar{\mathbf{s}}$ are done via function evaluations so that $\bar{\mathbf{A}}$ is never constructed explicitly. Also note that $\bar{\mathbf{h}}$ and $\bar{\mathbf{s}}$ are exchangeable since

$$\bar{\mathbf{A}}(\bar{\mathbf{h}})\bar{\mathbf{s}} \;=\; \hat{\mathbf{A}}_s(\bar{\mathbf{s}})\bar{\mathbf{h}} \tag{5.24}$$

where $\hat{\mathbf{A}}_s(\bar{\mathbf{s}}) \in \mathbb{R}^{(n-d)\times s}$ contains the first $s$ columns of a lower-triangular Toeplitz matrix with $\bar{\mathbf{s}}$ as its first column. We will exploit this property in the described alternating optimization method.

Various nonlinear optimization methods can be used to solve problem (5.23) [62, 84]. A fully coupled approach would update all variables simultaneously, e.g., an inexact Newton method to solve for $\bar{\mathbf{z}} = \begin{bmatrix} \bar{\mathbf{s}} \\ \bar{\mathbf{h}} \end{bmatrix}$. The main caveats of this approach are that derivatives are required and convergence can be slow. On the other hand, an alternating approach can be used to exploit the separability of the parameters in $\bar{\mathbf{s}}$ and $\bar{\mathbf{h}}$. That is, we alternate between fixing $\bar{\mathbf{h}}$ and optimizing over $\bar{\mathbf{s}}$, and fixing $\bar{\mathbf{s}}$ and optimizing over $\bar{\mathbf{h}}$. An alternating optimization method to solve (5.23) is provided in Algorithm 4. Notice that a key computational benefit of the alternating optimization approach for this problem is that by exploiting property (5.24), each optimization problem corresponding to solving a linear inverse problem.

In summary, we reformulated the blind respirometry reconstruction problem as a constrained nonlinear optimization problem, where the additional constraints are modest and reasonable. We assume that the delay and the support of the impulse response function are known,

**Algorithm 4** Alternating Optimization for Blind Respirometry

choose initial $\bar{\mathbf{h}}_0$, tolerance $\text{tol}_h$ and $\text{tol}_r$
**for** $k = 0, 1, 2, \ldots$ **do**
$\bar{\mathbf{x}}_k = \arg\min_{\bar{\mathbf{s}}} \left\| \bar{\mathbf{A}}(\bar{\mathbf{h}}_k)\bar{\mathbf{s}} - \bar{\mathbf{d}} \right\|_2^2 + \lambda_1 \|\bar{\mathbf{s}}\|_1$
$\bar{\mathbf{h}}_{k+1} = \arg\min_{\bar{\mathbf{h}}} \left\| \hat{\mathbf{A}}_s(\bar{\mathbf{s}}_k)\bar{\mathbf{h}} - \bar{\mathbf{d}} \right\|_2^2 + \lambda_{\bar{\mathbf{h}}} \|\bar{\mathbf{h}}\|_2^2 \quad \text{s.t.} \quad \bar{\mathbf{h}} \geq \mathbf{0} \quad \text{and} \quad \sum_{i=d+1}^{d+s} h_i \delta t = 1$
$\bar{\mathbf{r}}_k = \bar{\mathbf{A}}(\bar{\mathbf{h}}_k)\bar{\mathbf{s}}_k - \bar{\mathbf{d}}$
**if** $\|\bar{\mathbf{h}}_{k+1} - \bar{\mathbf{h}}_k\|_2 < \text{tol}_h$ **or** $\|\bar{\mathbf{r}}_k - \bar{\mathbf{r}}_{k-1}\|_2 < \text{tol}_r$ **then**
**stop**
**end if**
**end for**

and we describe an alternating optimization method to estimate both the impulse response function and the instantaneous signal. In general, alternating optimization methods can be slow to converge but can have fast convergence if the initial guess is close to a minimizer. By exploiting structure in the problem, we have reduced the overall computational costs. We remark that for problems where the impulse response function can be parameterized using a few variables, a variable projection method may be used [35], but including additional constraints is not straightforward.

## 5.4 Numerical Results

In this section, we compare numerical optimization methods for different regularization functions and demonstrate the performance of the proposed preconditioners for the linear respirometry problem. We provide numerical results for uncertainty quantification for both Tikhonov and 1-norm regularizers. Then, we present results for a nonlinear respirometry reconstruction problem, where robustness of the proposed nonlinear optimization method is investigated. For the simulated dataset, we generate the measurements as in (5.1), where $\mathbf{s}_{\text{true}}$, $\mathbf{h}$ and $\mathbf{d}$ are provided in Fig 5.4 where the noise level is 0.5% and $n = 512$. We remark

that real metabolic signals usually have slowly varying patterns. However, some species exhibit discontinuous cycles of ventilation with periods of little to no $CO_2$ release [32, 88]. Here, to test the methods we choose $\mathbf{s}_{\text{true}}$ to be a series of rectangular pulses with various durations and frequencies. The rectangular pulses contain high frequency elements and recovering these signals is more challenging comparing to smooth patterns. For the linear respirometry results, we assume that we are given $\mathbf{h}$ and $\mathbf{d}$, and we seek reconstructions of $\mathbf{s}_{\text{true}}$. For the nonlinear respirometry results, we assume that we are given $\mathbf{b}$ as well as the delay and support of $\mathbf{h}$, and we seek reconstructions of $\mathbf{s}_{\text{true}}$ and $\mathbf{h}$. In addition to the simulated studies, we provide a case study for experimental validation on real data for both the linear and nonlinear respirometry problems.

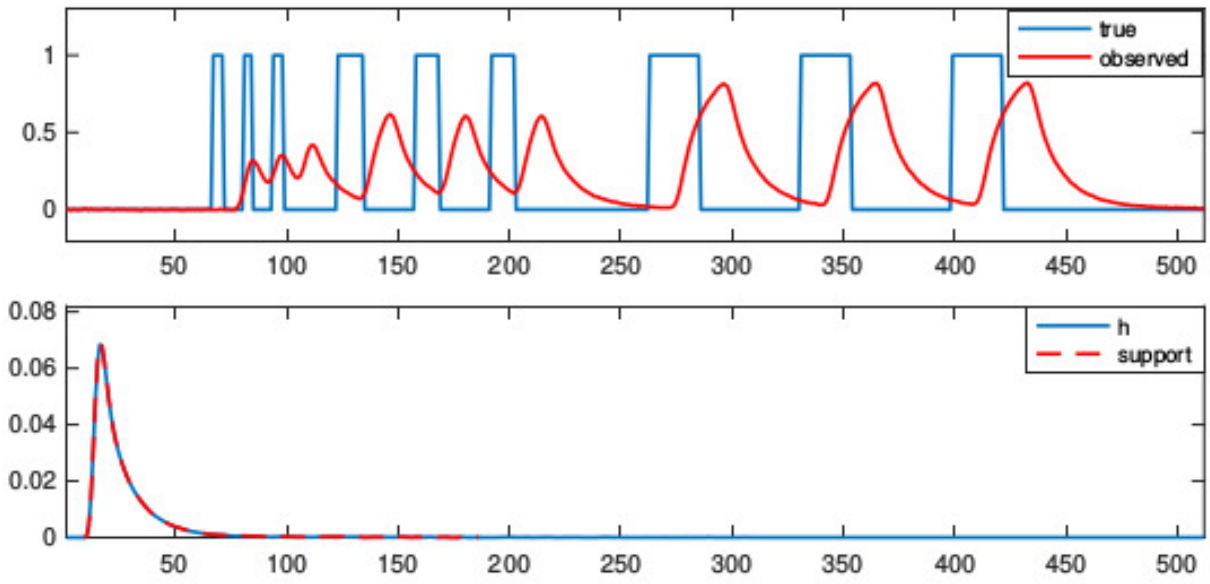


Figure 5.4: Simulated problem setup. The true signal and simulated observation with noise level 0.5% are provided in the top plot, and the impulse response function and its support are provided in the bottom plot.

### Linear Respirometry Reconstruction

For the linear respirometry problem, we investigate reconstructions using a 2-norm and a 1-norm regularization term. We consider two Tikhonov regularized solutions. Tikhonov-$\mathbf{Q}$ incorporates a regularization matrix $\mathbf{L}_\mathbf{Q}$, which is a lower triangular Toeplitz matrix with $[1 \ -1 \ 0 \ \dots \ 0]^\top$ as the first column [90], and HyBR-I corresponds to a hybrid iterative projection method with regularization matrix $\mathbf{Q} = \mathbf{I}$. For the 1-norm penalty, we investigate a flexible hybrid iterative method called FLSQR-R [34] and FISTA as described in Algorithm 3.

For the regularization parameter $\lambda$, we use the optimal regularization parameter for Tikhonov-$\mathbf{Q}$, which corresponds to minimizing the relative error between the reconstruction and the true signal. Both hybrid methods HyBR-I and FLSQR-R determine the regularization parameter automatically at each inner iteration using the weighted GCV method [34, 36]. For FISTA, the regularization parameter must be fixed in advance, and we set $\lambda_1 = 0.002$.

From the reconstructions in Fig 5.5, we observe that FLSQR-R and FISTA enforce sparsity in the reconstructions, and thus there are fewer artifacts. The FISTA reconstruction had the smallest relative reconstruction error norm among the considered methods, but this approach requires a good choice of the regularization parameter a priori, which requires time and careful tuning.

Next we provide credibility bounds for reconstructions of the linear respirometry problem. In Fig 5.6 we provide the Tikhonov-$\mathbf{Q}$ reconstruction from Fig 5.5 with $\lambda_2 = 0.0114$ along with the 95% credibility bounds. These 95% credibility bounds are computed from (2.8). Performing uncertainty quantification for Laplace priors (corresponding to the 1-norm) is a bit more difficult. We use the transformation described in [108] for Markov Chain Monte Carlo sampling. More specifically, we compute 1000 samples from the posterior distribution,

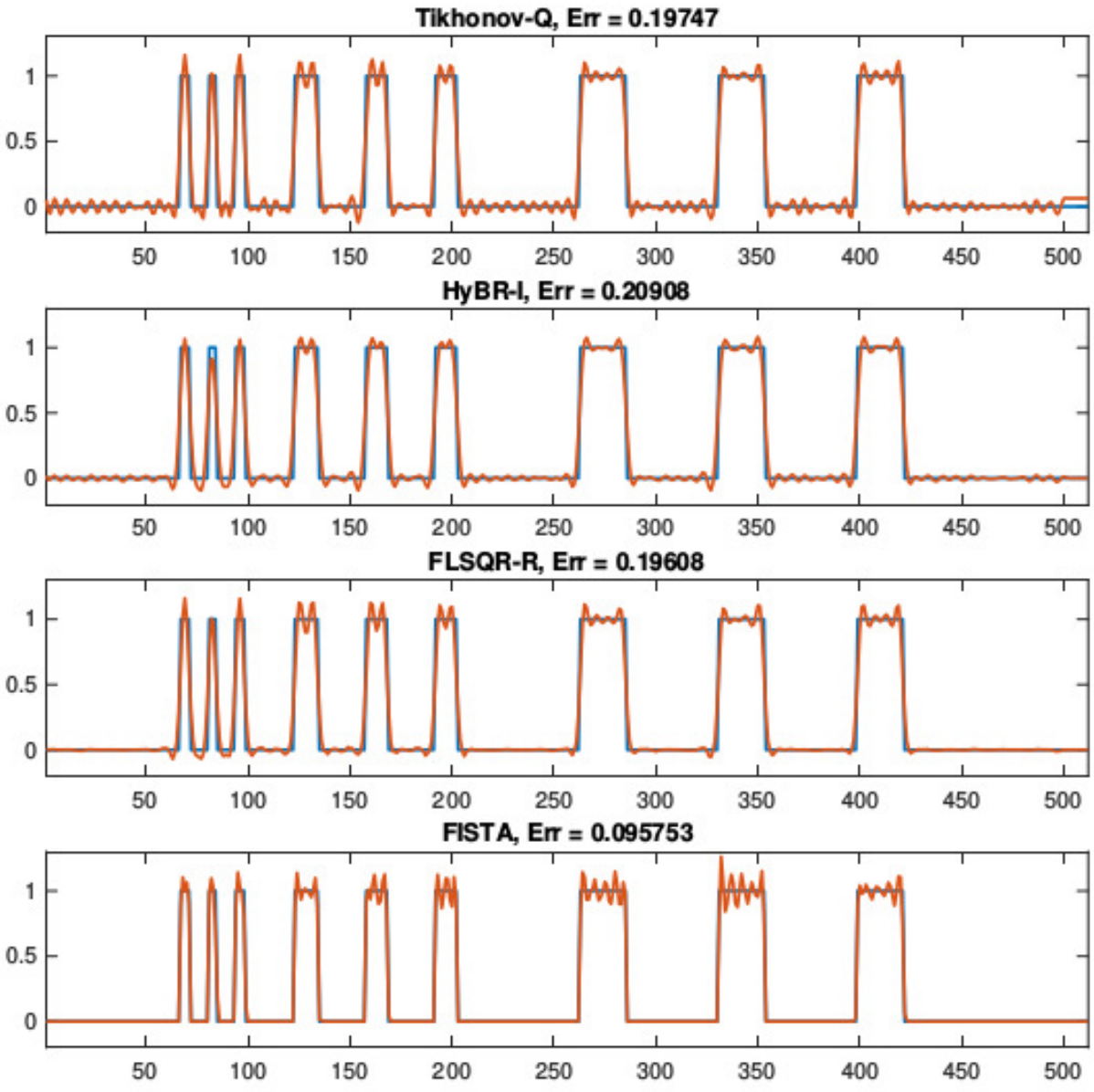


Figure 5.5: Reconstructions for linear respirometry reconstruction. Tikhonov-**Q** and HyBR-I reconstructions correspond to $\ell_2$ regularization, and FLSQR-R and FISTA reconstructions correspond to $\ell_1$ regularization. The true signal is provided in the blue line and the reconstructions are provided in red. Relative reconstruction error norms computed using the 2-norm are provided in the titles.

using the approach described in the Matlab code `OneDBlurHarr.m` from Section 6.4 of [7]. Due to computational difficulties for large-scale problems, we reduce the signal size to $n = 128$. As expected, we observe larger variances at the tail of the signal due to the delay in $\mathbf{h}$.

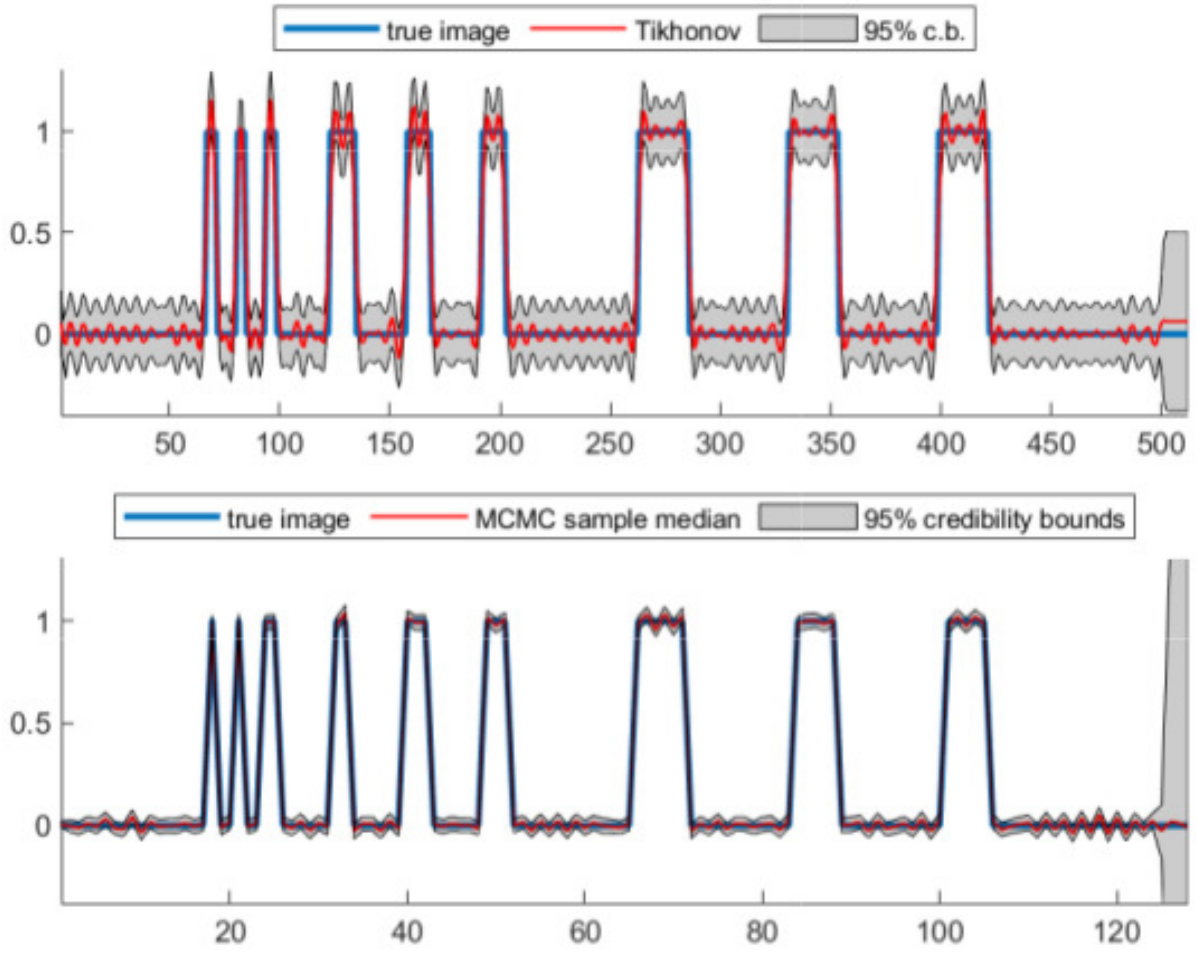


Figure 5.6: Uncertainty quantification for linear respirometry. The top plot contains the Tikhonov-$\mathbf{Q}$ solution with the 95% credibility bounds, and the bottom plot contains the sample median and 95% credibility bounds with 1000 samples corresponding to the Laplace prior.

For Tikhonov regularization, we investigate the proposed preconditioners. We provide relative reconstruction error norms per iteration in Fig 5.7 for two noise levels 0.5% and 1%. In practice the noise level is usually much lower. We show that both left and right precondi-

tioning result in faster convergence than unpreconditioned iterative methods. Furthermore although the preconditioned methods show semi-convergence behavior, whereby the error norms increase with later iterations, we can include appropriate regularization and compute a regularized solution. Thus, if one wishes to solve a large-scale nonlinear problem, preconditioned iterative methods can be used in an inner iteration to improve the overall efficiency of the algorithm.

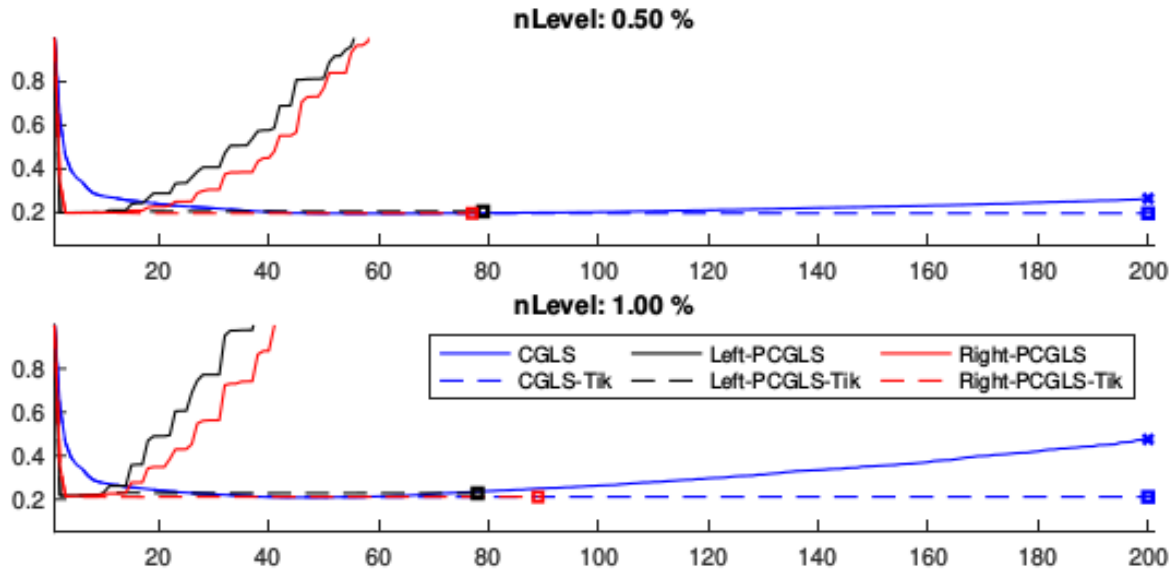


Figure 5.7: Results for preconditioned iterative methods. Relative reconstruction errors for preconditioned versus unpreconditioned iterative methods for Tikhonov regularization.

## Nonlinear Respirometry Reconstruction

For the blind respirometry reconstruction problem where we assume the impulse response function is unknown, we investigate the performance of the described alternating optimization approach. For Algorithm 4, we need an initial guess of the impulse response function. For this we take a uniform function on the support of the impulse response function (see

Fig 5.4) where the value is selected so that the impulse response function sums to 1. We found that our approaches can work for various choices of the initial guess of the impulse response function; however, it can be difficult to assess sensitivity for large-scale, nonlinear inverse problems. Our observation is motivated by our numerical experience in testing different initializations.

Recall that one of the nice features of the alternating optimization approach described in Algorithm 4 is that we split the main computational costs. Given an estimate of $\bar{\mathbf{h}}$, existing solvers can be used to compute a reconstruction $\bar{\mathbf{s}}$, and given an estimate of $\bar{\mathbf{s}}$, gradient-based constrained optimization methods can be used to efficiently estimate $\bar{\mathbf{h}}$. As we illustrated in Fig 5.3, the main challenge of the blind respirometry problem is the existence of multiple minimizers. In the experiments, we observe that using a 2-norm regularizer for $\bar{\mathbf{s}}$ resulted in overall smaller relative reconstruction error norms but much slower convergence. On the other hand, using a 1-norm regularizer for $\bar{\mathbf{s}}$ with FISTA was effective in avoiding the problem of getting stuck in undesirable local minimizers, especially at early iterations, but we needed to tune the regularization parameter. Here we use $\lambda_1 = 0.002$. For estimating $\bar{\mathbf{h}}$, we use MATLAB's `lsqlin` function to perform constrained optimization, where we enforce nonnegativity and summation to 1. For the choice of regularization parameter for $\bar{\mathbf{h}}$, we selected $\lambda_{\bar{\mathbf{h}}} = 0.01$. We tried a range of values from 0.2 to 0.005, and as expected, the reconstructed impulse response function was smoother for larger values of $\lambda_{\bar{\mathbf{h}}}$.

In Fig 5.8, we provide reconstructions of the impulse response function $\bar{\mathbf{h}}$ at various iterations of the alternating optimization method. Notice that the reconstructed function is shifted a few time units but is close to the support of the true impulse response function with small errors at the tails. Even though the initial guess is not close to true impulse response function, we obtain good reconstructions for the nonlinear respirometry problem.

After obtaining a good reconstruction of $\bar{\mathbf{h}}$, we test various reconstruction methods to com-

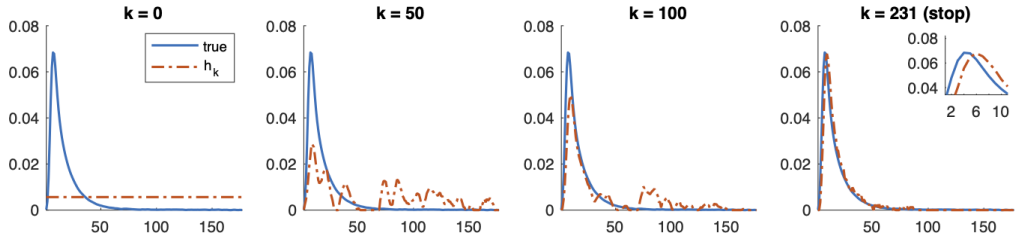


Figure 5.8: Reconstructed impulse response functions $\bar{\mathbf{h}}$ for the nonlinear respirometry problem at various iterations of the alternating optimization method. The subfigure in the right plot is a zoom of the peak of the reconstruction.

pute $\bar{\mathbf{s}}$. These results are provided in Fig 5.9 and show that FISTA and FLSQR-R are more sensitive than other methods when the reconstructed $\mathbf{h}$ has small errors at the tails. The regularization parameter was determined automatically in HyBR-I and FLSQR-R, while the optimal regularization parameter was used for Tikhonov-$\mathbf{Q}$. Since the reconstructed $\bar{\mathbf{h}}$ is shifted, the relative errors are not small. However, we can observe that the shape of the reconstructed $\bar{\mathbf{s}}$ is close to the true $\bar{\mathbf{s}}$. Next, we investigate robustness of the proposed algorithm to an inexact delay. For real experiments, it is hard to know the exact delay of the impulse response function, and thus we must estimate it. As shown in Fig 5.10, our nonlinear optimization approach can still reconstruct an impulse response function whose shape is similar to the true function.

## Experimental Validation

Finally, we test the described methods using real data from a flow-through respirometry chamber. First, we consider a linear reconstruction problem where we perfuse $CO_2$ with an arbitrary pattern into a respirometry chamber and record the output concentration of $CO_2$ with a gas analyzer. Then we applied the regularization methods (see Materials and Methods) to reconstruct the exact $CO_2$ injection pattern (i.e., the input signal) from the collected $CO_2$ observations. For this example, we use a fixed experimental impulse response

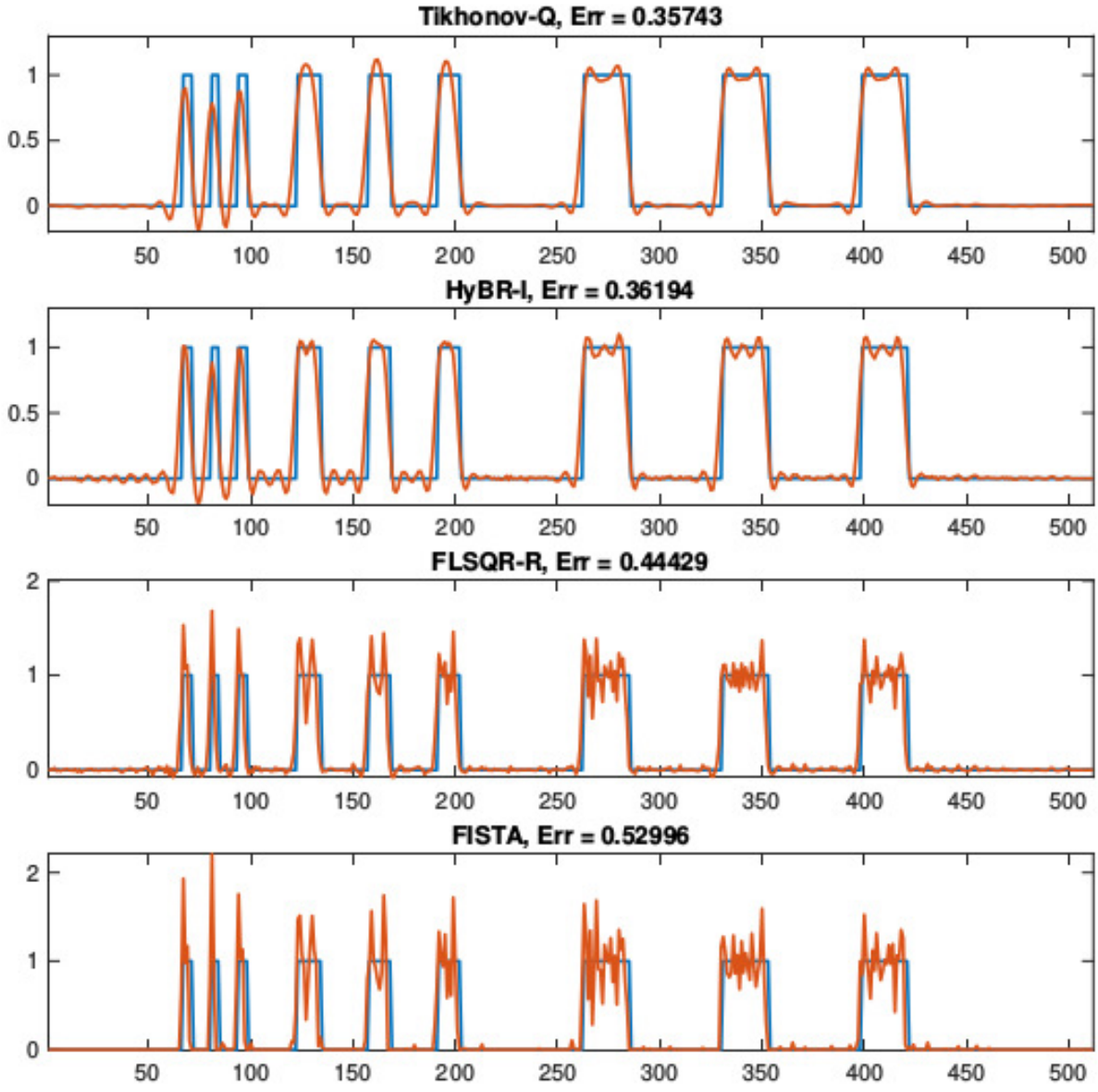


Figure 5.9: Reconstructed signals $\bar{\mathbf{s}}$ using different regularization techniques. All reconstructions correspond to $\bar{\mathbf{h}}$ in Fig 5.8.

function and an empty chamber. Then to demonstrate the effectiveness of the nonlinear reconstruction methods on a real biological system, we include an insect in the chamber and

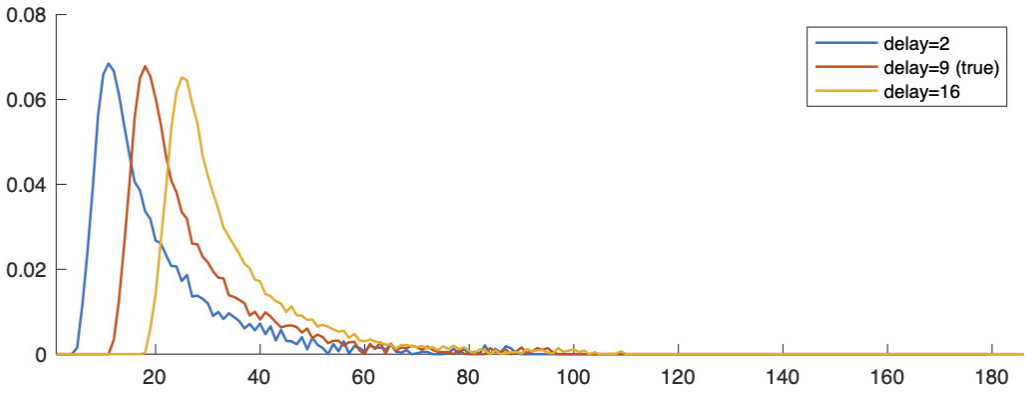


Figure 5.10: Investigation into the impact of selecting a different delay of the impulse response function in the nonlinear respirometry problem. Reconstructed impulse response functions $\bar{\mathbf{h}}$ are provided for different delays.

use Algorithm 4 to obtain a joint reconstruction of the impulse response function and $CO_2$ signal from the recorded $CO_2$ observations. We compare the reconstructed instantaneous signal with abdominal movements of the living organism. From these two experiments, we show the performance of the proposed method in experimental respirometry inverse problems.

### Linear case study

To validate the described methods, we designed an experimental setup to perfuse $CO_2$ with a controlled pattern into a 28 ml ($25 \times 25 \times 45$ mm$^3$) respirometry chamber. We used a high-speed valve (MHE2- MS1H-5/2-M7-K, Festo, NY, USA) to switch between dry air and $CO_2$ gas (100 ppm, balanced with N2) immediately before the chamber. The inlet flow rate into the chamber was 250 ml/min. The outlet of the chamber was connected to an infrared gas analyzer (LI 7000 Li-Cor, Nebraska, USA). To test the accuracy of the methods, $CO_2$ pulses with various width and frequencies were injected into the chamber and the concentration of the $CO_2$ in the outlet was recorded with a sampling rate of 10 Hz. To determine the impulse response of the respirometry chamber, a short pulse of $CO_2$ with the duration of

0.2s was injected into the chamber and the data were recorded for 5 minutes. The details of the experimental setup are described in Pendar et al [89, 90]. The output, observed $CO_2$ signal and the experimental impulse response are provided in Fig 5.11. The size of the input and observed signal is $13,413$. For the linear reconstruction problem, we evaluate the following methods: HyBR-I, FLSQR-R, and FISTA. For HyBR-I and FLSQR-R, the regularization parameter is computed automatically using weighted GCV, and for FISTA, we use $\lambda_1 = 0.002$. Reconstructions (including a zoomed image) are provided in Fig 5.12, along with the true signal and observation for comparison. Since the experimental impulse response function is well estimated in this case, all of the considered methods are able to nicely reconstruct the input signal. Notice that FISTA reconstructions are better able to resolve the peaks, especially when they are close together, as well as the flat regions (where no input is made).

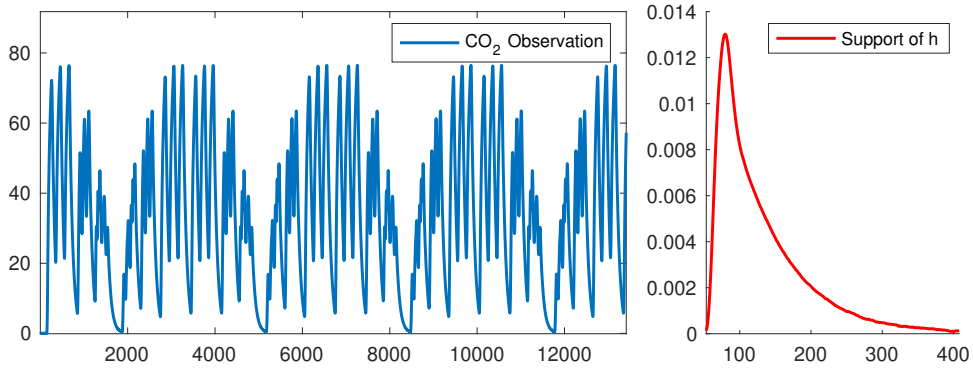


Figure 5.11: Experimental set-up for linear case study. $CO_2$ observations from manipulated input $CO_2$ signals to the empty chamber (left) and experimental impulse response function (right).

### Nonlinear case study: Abdominal pumping and $CO_2$ emission in a darkling beetle

Next, we include a breathing insect in the flow-through respirometry chamber and investigate the performance of the blind respirometry reconstruction methods for simultaneously

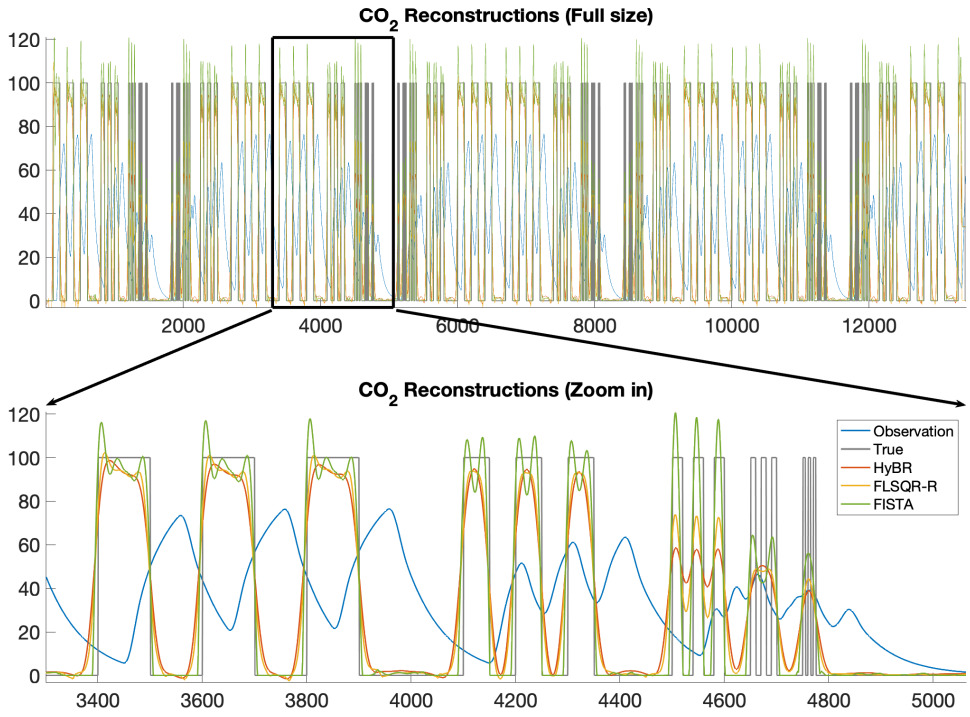


Figure 5.12: Experimental results for linear case study. Reconstruction of input $CO_2$ signal using different reconstruction methods. A zoomed plot is provided in the bottom plot.

estimating the impulse response function and the $CO_2$ instantaneous signal. A complicated network of tubes, called tracheae, run through an insect's body to deliver oxygen to the tissues and return $CO_2$ from the cells to the ambient air. The tube network is open to the outside air through valves which are called spiracles. Gas transport inside the tracheae occurs via diffusion and in larger insects via active ventilation, which is the result of compression of the tracheal tubes [110]. For larger and more active insects with higher metabolic rates, the diffusion is not sufficient to deliver enough oxygen to their tissues. They require an active ventilation to augment diffusive gas exchange. Active ventilation is known to be generated by abdominal pumping, a dorsoventral or anteroposterior compression of the abdomen [63, 64, 88]. However, some studies have shown that not all abdominal compressions are correlated

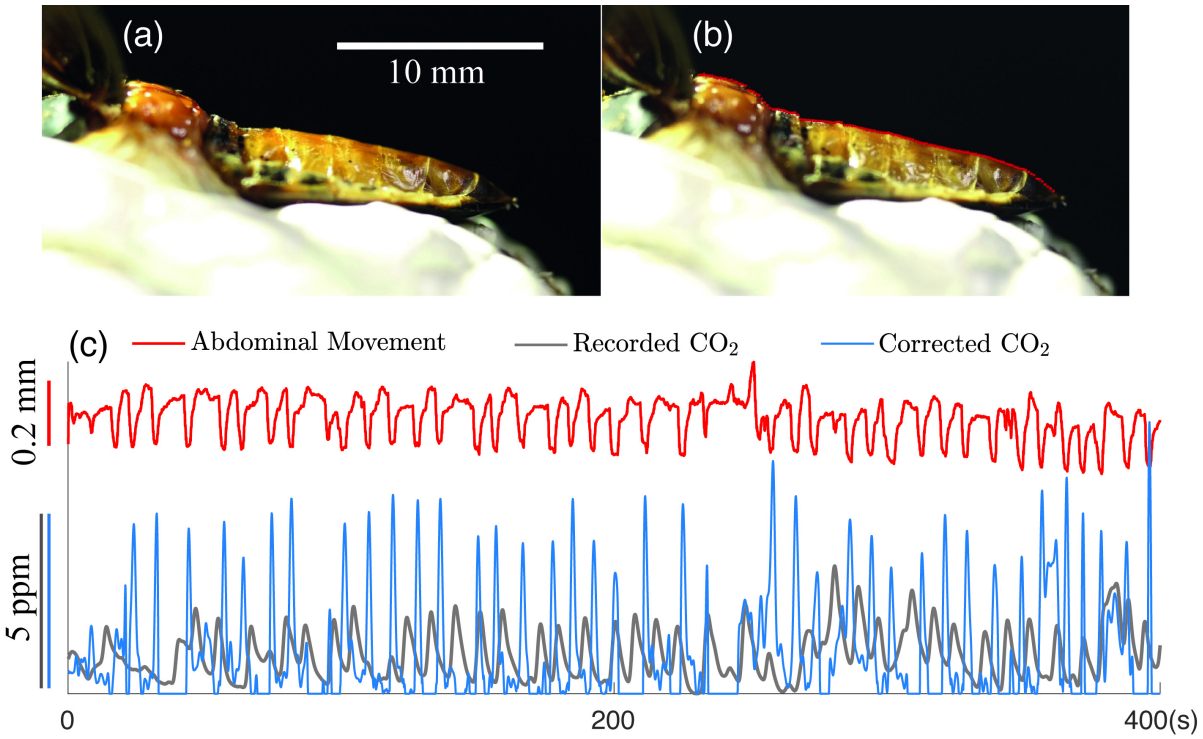


Figure 5.13: Experimental results for nonlinear case study. In this experiment, abdominal pumping and $CO_2$ emission of a breathing insect are recorded and synchronized. (a) A darkling beetle is breathing in the respirometry chamber. (b) Abdominal movements of a darkling beetle (red dots) are recorded. (c) The abdominal movement (red) signal samples. The recorded $CO_2$ emission samples (grey) from the chamber. The reconstructed $CO_2$ signals (blue) with Algorithm 4 from a flat initial guess of the impulse response function, delay of 21.3 sec, and support of 18.4 sec. The recovered $CO_2$ signal is concurrent with the abdominal movement of the insect.

with gas exchange, particularly in pupae and sub-adults [55, 88, 104]. In this study we used an adult tenebrionid beetle, Zophobas morio Fabricius, 1776 (Coleoptera: Tenebrionidae), to investigate the correlation between abdominal pumping and $CO_2$ emission. Before putting the beetle inside the respirometry chamber, the beetle was cold-anesthetized at 3°C. Then its legs, head, and antennae were secured using adhesive putty (Scotch adhesive putty, 3M, Minnesota, USA) to prevent body movements during the recording. To see the abdominal movement the elytra and the soft wings were pinned to the sides.

After putting the secured beetle shown in Fig 5.13 (a) inside the respirometry chamber and

letting the beetle rest for an hour, we recorded $CO_2$ emission and abdominal movement simultaneously. Recorded $CO_2$ observations can be found in Fig 5.13 (c). We also recorded the movement of the abdomen from the side with a video camera (NEX-VG10, Sony) at 30 frames per second. A flashing LED light was used to synchronize the video with the $CO_2$ data. To process the recorded video we used a custom MATLAB code to track 120 equally spaced points along the mid-tergites (see the red points in Fig 5.13 (b)) and considered the average displacement of these point as the dorso-ventral displacement of the abdomen.

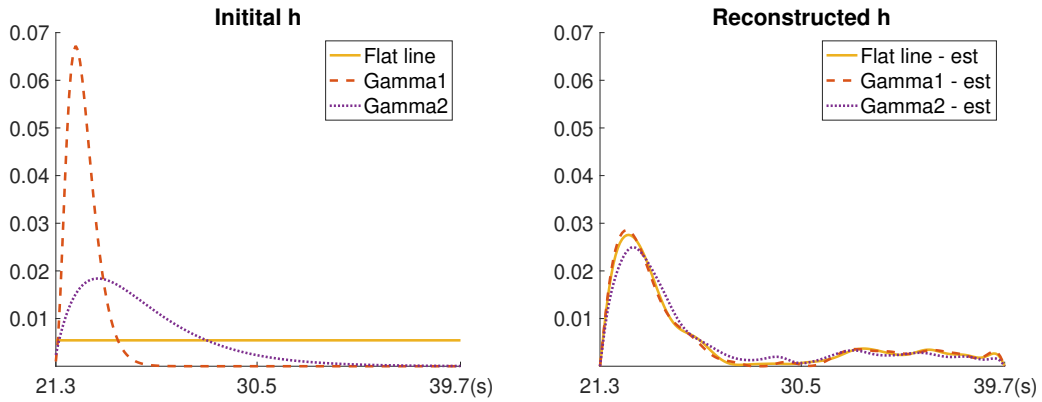


Figure 5.14: Impulse response function reconstruction for nonlinear case study. Initial guess for the unknown impulse response function (left) and reconstructed impulse response functions using a nonlinear respirometry reconstruction algorithm (right).

Then we used Algorithm 4 with the same regularization parameters used for the simulated experiments to simultaneously reconstruct the instantaneous $CO_2$ signal from the observations and the impulse response function. We tested different initial guesses for the impulse response function $\bar{\mathbf{h}}_0$, where the support is 18.4 sec and the delay is 21.3 sec. First, following the work in [90], we considered density functions of Gamma distributions (e.g., $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp^{-\beta x}$) for different choices of $\alpha$ and $\beta$. `Gamma1` corresponds to an initialization with $\alpha = 4$ and $\beta = 3$, and `Gamma2` corresponds to an initialization with $\alpha = 2$ and $\beta = 0.5$. The initializations of the impulse response functions are provided in Fig 5.14

To investigate the sensitivity of our approach with respect to the initialization of $\bar{\mathbf{h}}_0$, we

also considered an initial guesses where $\bar{\mathbf{h}}_0$ is a constant function. For each of the very different initial impulse response functions, the blind respirometry reconstruction method converges to an impulse response function with a similar shape and to similar reconstructed $CO_2$ signals. The reconstructed impulse response functions are provided in the right panel of Fig 5.14. Notice that all functions must satisfy two conditions: nonnegativity and the area under the curve over the support is 1. The reconstructed $CO_2$ signal (corresponding to an initialization of the flat line impulse response function) is provided in Fig 5.13 (c). Since this is a real data experiment, we do not have the true signal to compare to. Thus, we verify our results by comparing the reconstructed $CO_2$ signal to the recorded abdominal movement. In Fig 5.13 (c), we provide a superposition of signals in order to show a correlation between abdominal movement and $CO_2$ release. MATLAB code and data can be found at the website: *https://github.com/T-Cho-vt/respirometry*.

# Chapter 6

# Conclusions

This dissertation focuses on computational tools for solving inverse problems with numerical examples from various scientific applications. For linear inverse problems, we developed a hybrid projection method to handle mixed Gaussian priors and we described a problem reformulation to enable generalized hybrid projection methods for problems with a hierarchical Gaussian prior. Efficient preconditioning methods to accelerate iterative methods were introduced for linear problems where a given forward operator matrix is a lower triangular Toeplitz matrix. For nonlinear inverse problems, we developed computational tools for simultaneous estimation of the desired signal and the unknown impulse response function. We implemented different regularization methods to handle various prior distributions (or prior knowledge), and we considered efficient algorithms to solve the corresponding problems. Results for real-data experiments show the impact of this work in different areas ranging from geostatistical imaging to biology.

To summarize, in Chapter 3 we describe a hybrid iterative projection method, dubbed mixHyBR, that is based on an extension of the generalized Golub-Kahan bidiagonalization and that can be used for solving inverse problems (i.e., computing MAP estimates) with mixed Gaussian priors. The main advantage of this approach is that the mixing or blending parameter does not need to be known *a priori*, but rather can be estimated during the iterative process along with the regularization parameter. Various methods for selecting these parameters were considered and evaluated. Furthermore, mixHyBR methods can easily incorporate

data-driven priors where training data are used to define the prior covariance matrix itself (e.g., sample based priors) or to learn parameters for the covariance kernel function. Comparisons to widely-used shrinkage algorithms reveal that the mixed hybrid approaches are more robust under the presence of noise or freckles in the data and enable greater flexibility when selecting suitable priors. Numerical results from both spherical and seismic tomography show the potential of these methods.

In Chapter 4, we applied a hierarchical Gaussian priors to genHyBR and approximation of a posterior covariance matrix. This approach can estimate spatiotemporal unknowns and prior means simultaneously. Comparison to current existence methods for GHG tracking problems reduces computational cost by selecting regularization parameter automatically in the iterative process. Furthermore, the approximation of the posterior covariance matrix is almost free because we reused matrices generated from gen-GK process. Hence, large-scale inverse problems (e.g., the 6 week case study) are solved very efficiently with fast convergence.

In Chapter 5, we developed and investigated various computational tools for accurately and efficiently estimating the input signal, as well as the impulse response function, in any physiological system. By reformulating the linear respirometry problem in a Bayesian framework, we enabled tools for uncertainty quantification for both 2-norm and 1-norm regularization. Then, to accelerate the linear solve within iterative optimization methods (e.g., alternating optimization or Gauss-Newton methods), we developed preconditioners that are tailored to the respirometry forward model and demonstrated the excellent performance of these preconditioners for accelerating iterative reconstruction methods. Furthermore, by combining various constraints on both the impulse response function and the signal reconstruction, we developed sophisticated numerical optimization methods to tackle the very challenging problem of blind respirometry. Simulated and real-data results with a breathing insect

demonstrate that these methods can be used to extract high temporal information for the original signal. Overall, these improvements in input estimation have the potential to change the way physiologists view indirectly recorded data, most particularly for studies of gas exchange, and can change the interpretation of the underlying physiological processes. Future work will be to develop efficient methods for uncertainty quantification that can exploit the separable nonlinear structure of the respirometry problem. Furthermore, we addressed efficient computation of variance estimates, but sampling methods for both the Gaussian approximation and the fully nonlinear problem are still open problems, especially for very large problems.

In this disseration, we developed various extensions of generalized hybrid projection methods and blind deconvolution methods, and we have applied these techniques to a wide range of applications. However, there are still many interesting future works. For the mixHyBR method, one prior covariance matrix was learned from samples with a small number of parameters. We may extend this to non-kernel based covariance matrices. Furthermore, in this framework we can consider perturbations of kernels to see how the errors in the kernel function affect the reconstructions. For the hierarchical Gaussian prior assumptions, more investigations are needed on proper regularization parameter selection methods, especially for dataset with very high noise. Although DP worked well for selecting a regularization parameter, we were not able to get an appropriate regularization parameter using UPRE or GCV since the residual norm dominates the UPRE and the GCV functions. By analyzing the balance among terms in the UPRE and the GCV functions, we may be able to get reconstructions without the standard deviation of the white noise distribution. For the respirometry problems, future work remains to exploit parameterization of the impulse response function and to explore the posterior distribution for the nonlinear deconvolution problem.

# Bibliography


[1] NOAA Global Monitoring Division: CarbonTracker CT2017. 2019.

[2] S. Ambikasaran, J. Li, P. Kitanidis, and E. Darve. Large-scale stochastic linear inversion using hierarchical matrices. *Computational Geosciences*, 17:913–927, 2013.

[3] M. Arioli. Generalized Golub–Kahan bidiagonalization and stopping criteria. *SIAM Journal on Matrix Analysis and Applications*, 34(2):571–592, 2013.

[4] S. Arridge, P. Maass, O. Öktem, and C.-B. Schönlieb. Solving inverse problems using data-driven models. *Acta Numerica*, 28:1–174, 2019.

[5] M. Asch, M. Bocquet, and M. Nodet. *Data Assimilation: Methods, Algorithms, and Applications*, volume 11. SIAM, 2016.

[6] R. C. Aster, B. Borchers, and C. H. Thurber. Chapter four - Tikhonov regularization. In R. C. Aster, B. Borchers, and C. H. Thurber, editors, *Parameter Estimation and Inverse Problems (Third Edition)*, pages 93 – 134. Elsevier, third edition edition, 2019.

[7] J. M. Bardsley. *Computational Uncertainty Quantification for Inverse Problems*, volume 19. SIAM, 2018.

[8] G. A. Bartholomew, D. Vleck, and C. M. Vleck. Instantaneous measurements of oxygen consumption during pre-flight warm-up and post-flight cooling in sphingid and saturniid moths. *J. Exp. Biol.*, 90:17–32, 1981.

[9] F. S. V. Bazán and L. S. Borges. GKB-FP: An algorithm for large-scale discrete ill-posed problems. *BIT Numerical Mathematics*, 501:481–507, 2010.

[10] F. S. V. Bazán, L. S. Borges, and J. B. Francisco. On a generalization of Regińska's parameter choice rule and its numerical realization in large-scale multi-parameter Tikhonov regularization. *Appl. Math. Comput.*, 219(4):2100–2113, 2012.

[11] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.*, 2(1):183–202, 2009.

[12] M. Benzi. Preconditioning techniques for large linear systems: a survey. *Journal of computational Physics*, 182(2):418–477, 2002.

[13] Å. Björck. *Numerical methods for least squares problems*. SIAM, 1996.

[14] G. Brasseur and D. Jacob. *Modeling of Atmospheric Chemistry*. Cambridge University Press, 2017.

[15] R. D. Brown, J. M. Bardsley, and T. Cui. Semivariogram methods for modeling Whittle-Matérn priors in Bayesian inverse problems. *Inverse Problems*, 2020.

[16] T. A. Bubba, G. Kutyniok, M. Lassas, M. Maerz, W. Samek, S. Siltanen, and V. Srinivasan. Learning the invisible: A hybrid deep learning-shearlet framework for limited angle computed tomography. *Inverse Problems*, 2019.

[17] A. Buccini, M. Donatelli, and L. Reichel. Iterated Tikhonov regularization with a general penalty term. *Numerical Linear Algebra with Applications*, 24(4):e2089, 2017.

[18] T. Bui-Thanh, C. Burstedde, O. Ghattas, J. Martin, G. Stadler, and L. Wilcox. Extreme-scale uq for bayesian inverse problems governed by pdes. In *Proceedings of the international conference on high performance computing, networking, storage and analysis*, page 3. IEEE Computer Society Press, 2012.

[19] T. Bui-Thanh, O. Ghattas, J. Martin, and G. Stadler. A computational framework for infinite-dimensional Bayesian inverse problems Part i: The linearized case,

with application to global seismic inversion. *SIAM Journal on Scientific Computing*, 35(6):A2494–A2523, 2013.

[20] A. Butz, S. Guerlet, O. Hasekamp, D. Schepers, A. Galli, I. Aben, C. Frankenberg, J.-M. Hartmann, H. Tran, A. Kuze, G. Keppel-Aleks, G. Toon, D. Wunch, P. Wennberg, N. Deutscher, D. Griffith, R. Macatangay, J. Messerschmidt, J. Notholt, and T. Warneke. Toward accurate $CO_2$ and $CH_4$ observations from GOSAT. *Geophysical Research Letters*, 38(14), 2011.

[21] D. Calvetti and E. Somersalo. Prior conditioners for linear systems. *Inverse problems*, 21(4):1397, 2005.

[22] D. Calvetti and E. Somersalo. *An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*, volume 2. Springer Science & Business Media, 2007.

[23] D. U. Campos-Delgado, O. Gutierrez-Navarro, E. R. Arce-Santana, M. C. Skala, W. A. J., and J. A. Jo. Blind deconvolution estimation of fluorescence measurements through quadratic programming. *Journal of Biomedical Optics*, 20(7):075010, 2015.

[24] R. H. Chan, J. G. Nagy, and R. J. Plemmons. Circulant preconditioned Toeplitz least squares iterations. *SIAM Journal on Matrix Analysis and Applications*, 15(1):80–97, 1994.

[25] R. H. Chan and G. Strang. Toeplitz equations by conjugate gradients with circulant preconditioner. *SIAM Journal on Scientific and Statistical Computing*, 10(1):104–119, 1989.

[26] T. Chan and J. Olkin. Circulant preconditioners for Toeplitz-block matrices. *Numerical Algorithms*, 6(1):89–101, 1994.

[27] Y. Chen, A. Wiesel, Y. C. Eldar, and A. O. Hero. Shrinkage algorithms for MMSE covariance estimation. *IEEE Transactions on Signal Processing*, 58(10):5016–5029, 2010.

[28] Y. Chen, A. Wiesel, and A. O. Hero. Shrinkage estimation of high dimensional covariance matrices. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2937–2940. IEEE, 2009.

[29] T. Cho, J. Chung, and J. Jiang. Hybrid projection methods for large-scale inverse problems with mixed gaussian priors. *Inverse Problems*, 37(4):044002, 2021.

[30] T. Cho, J. Chung, S. Miller, and A. Saibaba. Hybrid projection method for large-scale inverse problems with mean estimation in hierarchical gaussian priors. *[in preperation]*, 2021.

[31] T. Cho, J. Chung, and H. Pendar. Computational tools for inversion and uncertainty estimation in respirometry. *Accepted to PLOS One*, 2021.

[32] S. L. Chown, S. Chown, S. Nicolson, et al. *Insect physiological ecology: mechanisms and patterns*. Oxford University Press, 2004.

[33] J. Chung, M. Chung, and D. P. O'Leary. Designing optimal spectral filters for inverse problems. *SIAM Journal on Scientific Computing*, 33(6):3132–3152, 2011.

[34] J. Chung and S. Gazzola. Flexible Krylov methods for $\ell_p$ regularization. *SIAM J. Sci. Comput.*, 41(5):149–171, 2018.

[35] J. Chung, E. Haber, and J. Nagy. Numerical methods for coupled super-resolution. *Inverse Problems*, 22:1261–1272, 2006.

[36] J. Chung, 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.

[37] J. Chung and K. Palmer. A hybrid LSMR algorithm for large-scale Tikhonov regularization. *SIAM J. Sci. Comput.*, 37(5):S562–S580, 2015.

[38] J. Chung and A. K. Saibaba. Generalized hybrid iterative methods for large-scale Bayesian inverse problems. *SIAM Journal on Scientific Computing*, 39(5):S24–S46, 2017.

[39] J. Chung, A. K. Saibaba, M. Brown, and E. Westman. Efficient generalized Golub–Kahan based methods for dynamic inverse problems. *Inverse Problems*, 34(2):024005, 2018.

[40] T. Daley and J. Ajo-Franklin. Frio 2 : Raw CASSM datasets. https://www.osti.gov/biblio/1490697-frio-raw-cassm-datasets. Accessed: 2019.

[41] T. Daley, J. Ajo-Franklin, and C. Doughty. Constraining the reservoir model of an injected $CO_2$ plume with crosswell CASSM at the Frio-II brine pilot. *International Journal of Greenhouse Gas Control*, 5:1022–1030, 2011.

[42] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart. Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization. *Mathematics of Computation*, 30(136):772–795, 1976.

[43] G. De Nicolao, G. Sparacino, and C. Cobelli. Nonparametric input estimation in physiological systems: problems, methods, and case studies. *Automatica*, 33(5):851–870, 1997.

[44] Y. Dong and C.-B. Schönlieb. Tomographic reconstruction with spatially varying parameter selection. *Inverse Problems*, 2020.

[45] D. L. Donoho. De-noising by soft-thresholding. *IEEE Transactions of Information Theory*, 41(3):613–627, 1995.

[46] A. Eldering, T. E. Taylor, C. W. O'Dell, and R. Pavlick. The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data. *Atmospheric Measurement Techniques*, 12(4):2341–2370, 2019.

[47] A. Eldering, P. O. Wennberg, D. Crisp, D. S. Schimel, M. R. Gunson, A. Chatterjee, J. Liu, F. M. Schwandner, Y. Sun, C. W. O'Dell, C. Frankenberg, T. Taylor, B. Fisher, G. B. Osterman, D. Wunch, J. Hakkarainen, J. Tamminen, and B. Weir. The orbiting carbon observatory-2 early science investigations of regional carbon dioxide fluxes. *Science*, 358(6360), 2017.

[48] H. P. Flath, L. C. Wilcox, V. Akcelik, J. Hill, B. van Bloemen Waanders, and O. Ghattas. Fast algorithms for bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian approximations. *SIAM Journal on Scientific Computing*, 33(1):407–432, 2011.

[49] A. Galka, O. Yamashita, and T. Ozaki. Spatiotemporal Kalman filtering: a new approach to solving dynamical inverse problems. In *International Symposium on Nonlinear Theory and its Applications*, 2004.

[50] S. Gazzola, P. C. Hansen, and J. G. Nagy. IR Tools: A MATLAB package of iterative regularization methods and large-scale test problems. *Numerical Algorithms*, 81(3):773–811, 2019.

[51] S. Gazzola and P. Novati. Multi-parameter Arnoldi-Tikhonov methods. *Electron. Trans. Numer. Anal.*, 40:452–475, 2013.

[52] G. H. Golub, M. Heath, and G. Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[53] G. H. Golub and C. F. Van Loan. *Matrix Computations, third edition*. Johns Hopkins University Press, 1996.

[54] S. M. Gourdji, K. L. Mueller, V. Yadav, D. N. Huntzinger, A. E. Andrews, M. Trudeau, G. Petron, T. Nehrkorn, J. Eluszkiewicz, J. Henderson, D. Wen, J. Lin, M. Fischer, C. Sweeney, and A. M. Michalak. North American $CO_2$ exchange: inter-comparison of modeled estimates with results from a fine-scale atmospheric inversion. *Biogeosciences*, 9(1):457–475, 2012.

[55] B. Groenewald, S. K. Hetz, S. L. Chown, and J. S. Terblanche. Respiratory dynamics of discontinuous gas exchange in the tracheal system of the desert locust, Schistocerca gregaria. *Journal of Experimental Biology*, 215(13):2301–2307, 2012.

[56] E. Haber and L. Tenorio. Learning regularization functionals: a supervised training approach. *Inverse Problems*, 19(3):611, 2003.

[57] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, pages 49–52, 1902.

[58] M. Hanke, J. Nagy, and R. Plemmons. Preconditioned iterative regularization. *Numerical linear algebra*, pages 141–163, 1992.

[59] P. C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, 2010.

[60] P. C. Hansen and J. S. Jorgensen. AIR Tools II: Algebraic Iterative Reconstruction Methods, Improved Implementation. *Numer. Algorithms*, 2018.

[61] P. C. Hansen, J. Nagy, and D. P. O'Leary. *Deblurring images: matrices, spectra, and filtering*. SIAM, Philadelphia, 2006.

[62] P. C. Hansen, V. Pereyra, and G. Scherer. *Least squares data fitting with applications*. JHU Press, 2013.

[63] J. F. Harrison. Ventilatory mechanism and control in grasshoppers. *American Zoologist*, 37(1):73–81, 1997.

[64] J. F. Harrison, J. S. Waters, A. J. Cease, J. M. VandenBrooks, V. Callier, C. J. Klok, K. Shaffer, and J. J. Socha. How locusts breathe. *Physiology*, 28(1):18–27, 2013.

[65] M. E. Hochstenbach and L. Reichel. An iterative method for Tikhonov regularization with a general linear regularization operator. *The Journal of Integral Equations and Applications*, pages 465–482, 2010.

[66] L. Hu, A. E. Andrews, K. W. Thoning, C. Sweeney, J. B. Miller, A. M. Michalak, E. Dlugokencky, P. P. Tans, Y. P. Shiga, M. Mountain, T. Nehrkorn, S. A. Montzka, K. McKain, J. Kofler, M. Trudeau, S. E. Michel, S. C. Biraud, M. L. Fischer, D. E. J. Worthy, B. H. Vaughn, J. W. C. White, V. Yadav, S. Basu, and I. R. van der Velde. Enhanced North American carbon uptake associated with El Niño. *Science Advances*, 5(6), 2019.

[67] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, New York, 2005.

[68] N. Kazemi. Automatic blind deconvolution with Toeplitz-structured sparse total least squares. *Geophysics*, 83(6):345–357, 2018.

[69] M. E. Kilmer and D. P. O'Leary. Choosing regularization parameters in iterative methods for ill-posed problems. *SIAM Journal on Matrix Analysis and Applications*, 22:1204–1221, 2001.

[70] P. Kuchment and L. Kunyansky. Mathematics of thermoacoustic tomography. *European Journal of Applied Mathematics*, 19(2):191–224, 2008.

[71] O. Ledoit and M. Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411, 2004.

[72] H. Li, J. Schwab, S. Antholzer, and M. Haltmeier. NETT: Solving inverse problems with deep neural networks. *Inverse Problems*, 2020.

[73] J. C. Lin, C. Gerbig, S. C. Wofsy, A. E. Andrews, B. C. Daube, K. J. Davis, and C. A. Grainger. A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research: Atmospheres*, 108(D16), 2003.

[74] C. J. Long, P. L. Purdon, S. Temereanca, N. U. Desai, M. S. Hämäläinen, and E. N. Brown. State-space solutions to the dynamic magnetoencephalography inverse problem using high performance computing. *The Annals of Applied Statistics*, 5(2B):1207, 2011.

[75] S. Lu and S. V. Pereverzev. Multi-parameter regularization and its numerical realization. *Numer. Math.*, 118(1):1–31, 2011.

[76] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos. Using deep neural networks for inverse problems in imaging: beyond analytical methods. *IEEE Signal Processing Magazine*, 35(1):20–36, 2018.

[77] T. Machida, H. Matsueda, Y. Sawa, Y. Nakagawa, K. Hirotani, N. Kondo, K. Goto, T. Nakazawa, K. Ishikawa, and T. Ogawa. Worldwide measurements of atmospheric $CO_2$ and other trace gas species using commercial airlines. *Journal of Atmospheric and Oceanic Technology*, 25(10):1744 – 1754, 01 Oct. 2008.

[78] A. M. Michalak, L. Bruhwiler, and P. P. Tans. A geostatistical approach to surface flux estimation of atmospheric trace gases. *Journal of Geophysical Research: Atmospheres*, 109(D14), 2004.

[79] S. M. Miller, A. M. Michalak, R. G. Detmers, O. P. Hasekamp, L. M. Bruhwiler, and S. Schwietzke. China's coal mine methane regulations have not curbed growing emissions. *Nature Communications*, 10(1):1–8, 2019.

[80] S. M. Miller, A. K. Saibaba, M. E. Trudeau, M. E. Mountain, and A. E. Andrews. Geostatistical inverse modeling with very large datasets: an example from the Orbiting Carbon Observatory 2 (OCO-2) satellite. *Geoscientific Model Development*, 13(3):1771–1785, 2020.

[81] S. M. Miller, S. C. Wofsy, A. M. Michalak, E. A. Kort, A. E. Andrews, S. C. Biraud, E. J. Dlugokencky, J. Eluszkiewicz, M. L. Fischer, G. Janssens-Maenhout, B. R. Miller, J. B. Miller, S. A. Montzka, T. Nehrkorn, and C. Sweeney. Anthropogenic emissions of methane in the United States. *Proceedings of the National Academy of Sciences*, 110(50):20018–20022, 2013.

[82] M. Nakajima, A. Kuze, and H. Suto. The current status of GOSAT and the concept of GOSAT-2. In R. Meynart, S. P. Neeck, and H. Shimoda, editors, *Sensors, Systems, and Next-Generation Satellites XVI*, volume 8533, pages 21 – 30. International Society for Optics and Photonics, SPIE, 2012.

[83] T. Nehrkorn, J. Eluszkiewicz, S. Wofsy, J. Lin, C. Gerbig, M. Longo, and S. Freitas. Coupled weather research and forecasting–stochastic time-inverted Lagrangian transport (WRF–STILT) model. *Meteorol Atmos Phys*, 107:51–64, 2010.

[84] J. Nocedal and S. Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

[85] D. P. O'Leary and J. A. Simmons. A bidiagonalization-regularization procedure for large scale discretizations of ill-posed problems. *SIAM Journal on Scientific and Statistical Computing*, 2:474–489, 1981.

[86] C. Paige and M. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71, 1982.

[87] C. C. Paige and M. A. Saunders. Algorithm 583, LSQR: Sparse linear equations and least-squares problems. *ACM Trans. Math. Soft.*, 8(2):195–209, 1982.

[88] H. Pendar, M. C. Kenny, and J. J. Socha. Tracheal compression in pupae of the beetle Zophobas morio. *Biology letters*, 11(6):20150259, 2015.

[89] H. Pendar and J. J. Socha. Estimation of instantaneous gas exchange in flow-through respirometry systems: A modern revision of Bartholomew's Z-transform method. *PLOS one*, 10(10):e0139508, 2015.

[90] H. Pendar, J. J. Socha, and J. Chung. Recovering signals in physiological systems with large datasets. *Biology open*, 5(8):1163–1174, 2016.

[91] A. Petzold, V. Thouret, C. Gerbig, A. Zahn, C. A. M. Brenninkmeijer, M. Gallagher, M. Hermann, M. Pontaud, H. Ziereis, D. Boulanger, J. Marshall, P. Nèdèlec, H. G. J. Smit, U. Friess, J.-M. Flaud, A. Wahner, J.-P. Cammas, A. Volz-Thomas, and I. TEAM. Global-scale atmosphere monitoring by in-service aircraft - current achievements and future prospects of the European Research Infrastructure IAGOS. *Tellus B: Chemical and Physical Meteorology*, 67(1):28452, 2015.

[92] C. E. Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71. Springer, 2003.

[93] C. E. Rasmussen and C. Williams. Gaussian processes for machine learning. *the MIT Press*, 2(3):4, 2006.

[94] L. Reichel, F. Sgallari, and Q. Ye. Tikhonov regularization based on generalized Krylov subspace methods. *Applied Numerical Mathematics*, 62(9):1215–1228, 2012.

[95] R. A. Renaut, S. Vatankhah, and V. E. Ardestani. Hybrid and iteratively reweighted regularization by unbiased predictive risk and weighted GCV for projected systems. *SIAM J. Sci. Comput.*, 39(2):B221–B243, 2017.

[96] Y. Saad. *Iterative methods for sparse linear systems*. SIAM, Philadelphia, 2003.

[97] A. K. Saibaba, J. Chung, and K. Petroske. Efficient Krylov subspace methods for uncertainty quantification in large Bayesian linear inverse problems. *Numerical Linear Algebra with Applications*, 27(5):e2325, 2020.

[98] A. K. Saibaba and P. K. Kitanidis. Fast computation of uncertainty quantification measures in the geostatistical approach to solve inverse problems. *Advances in Water Resources*, 82(0):124 – 138, 2015.

[99] J. Schäfer and K. Strimmer. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, 4(1), 2005.

[100] J. Schwab, S. Antholzer, and M. Haltmeier. Deep null space learning for inverse problems: Convergence analysis and rates. *Inverse Problems*, 2018.

[101] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, 2009.

[102] Y. P. Shiga, A. M. Michalak, Y. Fang, K. Schaefer, A. E. Andrews, D. H. Huntzinger, C. R. Schwalm, K. Thoning, and Y. Wei. Forests dominate the interannual variability of the North American carbon sink. *Environmental Research Letters*, 13(8):084015, aug 2018.

[103] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, and Y. Marzouk. Optimal low-

rank approximations of Bayesian linear inverse problems. *SIAM Journal on Scientific Computing*, 37(6):A2451–A2487, 2015.

[104] U. Tartes, A. Vanatoa, and A. Kuusik. The insect abdomenâ a heartbeat manager in insects? *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 133(3):611–623, 2002.

[105] A. J. Turner, D. J. Jacob, K. J. Wecht, J. D. Maasakkers, E. Lundgren, A. E. Andrews, S. C. Biraud, H. Boesch, K. W. Bowman, N. M. Deutscher, M. K. Dubey, D. W. T. Griffith, F. Hase, A. Kuze, J. Notholt, H. Ohyama, R. Parker, V. H. Payne, R. Sussmann, C. Sweeney, V. A. Velazco, T. Warneke, P. O. Wennberg, and D. Wunch. Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data. *Atmospheric Chemistry and Physics*, 15(12):7049–7069, 2015.

[106] J. D. Veldhuis, M. L. Carlson, and M. L. Johnson. The pituitary gland secretes in bursts: appraising the nature of glandular secretory impulses by simultaneous multiple-parameter deconvolution of plasma hormone concentrations. *Proceedings of the National Academy of Sciences*, 84(21):7686–7690, 1987.

[107] Z. Wang. Multi-parameter Tikhonov regularization and model function approach to the damped Morozov principle for choosing regularization parameters. *J. Comput. Appl. Math.*, 236(7):1815–1832, 2012.

[108] Z. Wang, J. M. Bardsley, A. Solonen, T. Cui, and Y. Marzouk. Bayesian inverse problems with $l_1$ priors: A randomize-then-optimize approach. *SIAM J. Sci. Comput.*, 39(5):140–166, 2017.

[109] K. J. Wecht, D. J. Jacob, C. Frankenberg, Z. Jiang, and D. R. Blake. Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY

satellite data. *Journal of Geophysical Research: Atmospheres*, 119(12):7741–7756, 2014.

[110] M. W. Westneat, O. Betz, R. W. Blob, K. Fezzaa, W. J. Cooper, and W.-K. Lee. Tracheal respiration in insects visualized with synchrotron x-ray imaging. *science*, 299(5606):558–560, 2003.

[111] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7), 2009.

[112] V. Yadav, A. M. Michalak, J. Ray, and Y. P. Shiga. A statistical approach for isolating fossil fuel emissions in atmospheric inverse problems. *Journal of Geophysical Research: Atmospheres*, 121(20):12–490, 2016.

[113] D. Yang, Y. Liu, and Z. Cai et al. First global carbon dioxide maps produces from TanSat measurements. *Adv. Atmos. Sci.*, 35:621–623, 2018.

[114] S. Yang, H. Chen, J. Li, and W. Wang. A spatially-variant morphological regularization method for inverse problems in image processing. In *2016 Chinese Control and Decision Conference (CCDC)*, pages 3789–3794. IEEE, 2016.

[115] Z. Yao, R. Roberts, and A. Tryggvason. Calculating resolution and covariance matrices for seismic tomography with the LSQR method. *Geophysical Journal International*, 138(3):886–894, 1999.

[116] J. Zhang and G. A. McMechan. Estimation of resolution and covariance for large matrix inversions. *Geophysical Journal International*, 121(2):409–426, 1995.

# Appendices

# Appendix A

# Appendix for Section 3

## A.1 Proof of Theorem 3.1

*Proof.* Based on (3.22) and (3.23),

$$\mathbf{y}_k(\lambda,\gamma) = \mathbf{C}_k(\gamma,\lambda)\begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix} \tag{A.1}$$

With $k = n$, by (2.12), (2.13), (2.14), and (A.9),

$$\begin{aligned}
\mathbf{s}_n &= \boldsymbol{\mu} + \mathbf{Q}\mathbf{V}_n\mathbf{y}_n \\
&= \boldsymbol{\mu} + \mathbf{Q}\mathbf{V}_n\mathbf{C}_n(\gamma,\lambda)\begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix} \\
&= \boldsymbol{\mu} + \mathbf{Q}\mathbf{V}_n(\mathbf{V}_n^\top\mathbf{Q}^\top\mathbf{A}^\top\mathbf{L}_\mathbf{R}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q}\mathbf{V}_n + \lambda^2\mathbf{V}_n^\top\mathbf{Q}\mathbf{V}_n)^{-1}\mathbf{V}_n^\top\mathbf{Q}\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{b} \\
&= \boldsymbol{\mu} + \mathbf{Q}(\mathbf{Q}^\top\mathbf{A}^\top\mathbf{L}_\mathbf{R}^\top\mathbf{L}_\mathbf{R}\mathbf{A}\mathbf{Q} + \lambda^2\mathbf{Q})^{-1}\mathbf{Q}\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{b} \\
&= \boldsymbol{\mu} + \mathbf{Q}(\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{A}\mathbf{Q} + \lambda^2\mathbf{I}_n)^{-1}\mathbf{A}^\top\mathbf{R}^{-1}\mathbf{b} \\
&= \mathbf{s}_{\text{MAP}}.
\end{aligned}$$

Therefore, the solution for (3.26) converges to the solution for (3.32) and the solution for (3.29) converges to the solution for (3.34) as $k$ increases. □

## A.2 Derivation of regularization parameter selection methods for mixHyBR

The derivations for both UPRE and GCV follow derivations for the Tikhonov case, see e.g. [7]. For UPRE for mixHyBR, the goal is to select parameters $\lambda$ and $\gamma$ that minimize the predictive risk,

$$\mathbb{E}\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 \tag{A.2}$$

where $\mathbf{y}_k^{(\gamma,\lambda)}$ is the solution to (3.17) with fixed $\gamma$ and $\lambda$, $\boldsymbol{\delta} = \begin{bmatrix} \beta\mathbf{e}_1 \\ \mathbf{0} \end{bmatrix}$, $\mathbf{y}_k^{(\gamma,\lambda)} = \mathbf{C}_k(\gamma,\lambda)\boldsymbol{\delta}$, and $\mathbf{D}_k(\gamma)$ and $\mathbf{C}_k(\gamma,\lambda)$ are defined in (3.23) and (3.24) respectively. For the projected problem (3.17), the vector $\boldsymbol{\delta}$ consists of a deterministic and a stochastic part as described in [95], so $\boldsymbol{\delta} = \mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda) + \boldsymbol{\eta}$ where $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and $\boldsymbol{\eta} \in \mathbb{R}^{2k+1}$. Then, (A.2) can be written as

$$\begin{aligned}
& \mathbb{E}\,\|\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)\boldsymbol{\delta} - \mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 \\
= \;& \mathbb{E}\,\|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda) + \mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)\boldsymbol{\eta}\|_2^2 \\
= \;& \|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 + \mathbb{E}\boldsymbol{\eta}^\top(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda))^2\boldsymbol{\eta} \\
= \;& \|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 + \sigma^2\mathrm{tr}((\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda))^2).
\end{aligned} \tag{A.3}$$

Next the expectation of the projected residual can be written as

$$\begin{aligned}
& \mathbb{E}\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \boldsymbol{\delta}\|_2^2 \\
= \; & \mathbb{E}\,\|\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)\boldsymbol{\delta} - \boldsymbol{\delta}\|_2^2 \\
= \; & \mathbb{E}\,\|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\,\boldsymbol{\delta}\|_2^2 \\
= \; & \mathbb{E}\,\|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\,\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda) + (\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\,\boldsymbol{\eta}\|_2^2 \qquad \text{(A.4)} \\
= \; & \|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\,\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 + \mathbb{E}(\boldsymbol{\eta}^\top\,(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})^2\,\boldsymbol{\eta}) \\
= \; & \|(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda) - \mathbf{I}_{2k+1})\,\mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 + \sigma^2\mathrm{tr}((\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda))^2) \\
& \quad -2\sigma^2\mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) + \sigma^2(2k+1).
\end{aligned}$$

By combining (A.3) and (A.4), we get

$$\begin{aligned}
\mathbb{E}\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \mathbf{D}_k(\gamma)\mathbf{y}_k(\gamma,\lambda)\|_2^2 &= \mathbb{E}\left\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \boldsymbol{\delta}\right\|_2^2 + 2\sigma^2\mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) - \sigma^2(2k+1) \\
&\approx \left\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \boldsymbol{\delta}\right\|_2^2 + 2\sigma^2\mathrm{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) - \sigma^2(2k+1).
\end{aligned} \qquad \text{(A.5)}$$

By dividing the function in (A.5) by $2k+1$, we get the following UPRE function at the $k$th iteration of the mixHyBR method,

$$\mathcal{U}_{\text{proj}}(\gamma,\lambda) = \frac{1}{2k+1}\|\mathbf{r}_k^{\text{proj}}(\gamma,\lambda)\|_2^2 + \frac{2\sigma^2}{2k+1}\mathrm{tr}(\mathbf{D}_k(\gamma),\mathbf{C}_k(\gamma,\lambda)) - \sigma^2 \qquad \text{(A.6)}$$

where $\mathbf{r}_k^{\text{proj}}(\gamma,\lambda) = \mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \boldsymbol{\delta}$.

The GCV method for selecting a regularization parameter is an approximation of a leave-one-out cross validation (LOOCV) method. However, for large-scale problems, the direct application of LOOCV and GCV for parameter selection is not feasible. Following the

derivation from [6], we get

$$V(\gamma,\lambda) = \frac{1}{2k+1}\sum_{i=1}^{2k+1}\left(\frac{[\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)}]_i - [\boldsymbol{\delta}]_i}{1-[\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)]_{ii}}\right)^2. \tag{A.7}$$

where $[\cdot]_i$ represents the $i$th component of a vector and $[\cdot]_{ij}$ represents the $i$th row and $j$th column component of a matrix. Next, since computing diagonal entries of $\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)$ can be expensive, the entries are replaced by the average of diagonal entries such that we get tje GCV function for the mixHyBR method,

$$G(\gamma,\lambda) = \frac{(2k+1)\|\mathbf{D}_k(\gamma)\mathbf{y}_k^{(\gamma,\lambda)} - \boldsymbol{\delta}\|_2^2}{(\mathrm{tr}(\mathbf{I}_{2k+1} - \mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)))^2}. \tag{A.8}$$

Notice that this is a scalar multiple of the GCV function (with constant $2k+1$) as given in (3.29).

## A.3 Proof of Lemma 3.2

*Proof.* For the projected residual for $\mathbf{x}_k$,

$$\begin{aligned}
\left\|\mathbf{r}_k^{\text{proj}}(\gamma,\lambda)\right\|_2^2 &= \left\|\mathbf{D}(\gamma)\mathbf{y}_k(\gamma,\lambda) - \begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix}\right\|_2^2 \\
&= \|(\gamma\mathbf{L_R}\mathbf{A}\mathbf{Q}_1\mathbf{V}_k + (1-\gamma)\mathbf{L_R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k)\mathbf{y}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2 \\
&= \left\|\begin{bmatrix}\mathbf{I}_{k+1} & \widetilde{\mathbf{U}}_{k+1}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\\ \mathbf{0} & \mathbf{R}_k\end{bmatrix}\begin{bmatrix}\gamma\mathbf{B}_k\\ (1-\gamma)\mathbf{I}_k\end{bmatrix}\mathbf{y}_k(\gamma,\lambda) - \begin{bmatrix}\beta_1\mathbf{e}_1\\ \mathbf{0}\end{bmatrix}\right\|_2^2 \\
&= \left\|\begin{bmatrix}\widetilde{\mathbf{U}}_{k+1} & \mathbf{L_R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k\end{bmatrix}\begin{bmatrix}\gamma\mathbf{B}_k\\ (1-\gamma)\mathbf{I}_k\end{bmatrix}\mathbf{y}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\right\|_2^2 \\
&= \left\|(\gamma\widetilde{\mathbf{U}}_{k+1}\mathbf{B}_k + (1-\gamma)\mathbf{L_R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k)\mathbf{y}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\right\|_2^2 \\
&= \|(\gamma\mathbf{L_R}\mathbf{A}\mathbf{Q}_1\mathbf{V}_k + (1-\gamma)\mathbf{L_R}\mathbf{A}\mathbf{Q}_2\mathbf{V}_k)\mathbf{y}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2 \\
&= \|\mathbf{L_R}\mathbf{A}\mathbf{Q}\mathbf{x}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2
\end{aligned}$$

and as $k = n$,

$$\|\mathbf{L_R}\mathbf{A}\mathbf{Q}\mathbf{x}_k(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2 \to \|\mathbf{L_R}\mathbf{A}\mathbf{Q}\mathbf{x}_n(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2 .$$

Since $\|\mathbf{L_R}\mathbf{A}\mathbf{Q}\mathbf{x}_n(\gamma,\lambda) - \mathbf{L_R}\mathbf{b}\|_2^2 = \left\|\mathbf{r}^{\text{full}}(\gamma,\lambda)\right\|_2^2$, when $k = n$ we have

$$\left\|\mathbf{r}_k^{\text{proj}}(\gamma,\lambda)\right\|_2^2 = \left\|\mathbf{r}^{\text{full}}(\gamma,\lambda)\right\|_2^2 .$$

For $k$th iteration in the projected problem (3.22),

$$\begin{aligned}
\mathbf{D}_k(\gamma)^\top \mathbf{D}_k(\gamma) &= \begin{bmatrix} \gamma \mathbf{B}_k + (1-\gamma)\widetilde{\mathbf{U}}_{k+1}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\ (1-\gamma)\mathbf{R}_k \end{bmatrix}^\top \begin{bmatrix} \gamma \mathbf{B}_k + (1-\gamma)\widetilde{\mathbf{U}}_{k+1}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\ (1-\gamma)\mathbf{R}_k \end{bmatrix} \\
&= \gamma^2 \mathbf{B}_k^\top \mathbf{B}_k + 2\gamma(1-\gamma)\mathbf{B}_k^\top \mathbf{U}_{k+1}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&\quad + (1-\gamma)^2 \mathbf{V}_k^\top \mathbf{Q}_2^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{U}_{k+1} \mathbf{U}_{k+1}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k + (1-\gamma)^2 \mathbf{R}_k^\top \mathbf{R}_k \\
&= \gamma^2 \mathbf{V}_k^\top \mathbf{Q}_1^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_1 \mathbf{V}_k + 2\gamma(1-\gamma)\mathbf{V}_k^\top \mathbf{Q}_1^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&\quad + (1-\gamma)^2 \mathbf{V}_k^\top \mathbf{Q}_2^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{U}_{k+1} \mathbf{U}_{k+1}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&\quad + (1-\gamma)^2 \mathbf{V}_k^\top \mathbf{Q}_2^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top (\mathbf{I}_{k+1} - \mathbf{L}_\mathbf{R} \mathbf{U}_{k+1} \mathbf{U}_{k+1}^\top \mathbf{L}_\mathbf{R}^\top) \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&= \gamma^2 \mathbf{V}_k^\top \mathbf{Q}_1^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_1 \mathbf{V}_k + 2\gamma(1-\gamma)\mathbf{V}_k^\top \mathbf{Q}_1^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&\quad + (1-\gamma)^2 \mathbf{V}_k^\top \mathbf{Q}_2^\top \mathbf{A}^\top \mathbf{L}_\mathbf{R}^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q}_2 \mathbf{V}_k \\
&= (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k
\end{aligned}$$

Therefore,

$$\begin{aligned}
\operatorname{tr}(\mathbf{D}_k(\gamma)\mathbf{C}_k(\gamma,\lambda)) &= \operatorname{tr}(\mathbf{D}_k(\gamma)(\mathbf{D}_k(\gamma)^\top \mathbf{D}_k(\gamma) + \lambda^2\gamma \mathbf{I}_k + \lambda^2(1-\gamma)\mathbf{V}_k^\top \mathbf{Q}_2 \mathbf{V}_k)^{-1} \mathbf{D}_k(\gamma)^\top) \\
&= \operatorname{tr}((\mathbf{D}_k(\gamma)^\top \mathbf{D}_k(\gamma) + \lambda^2\gamma \mathbf{I}_k + \lambda^2(1-\gamma)\mathbf{V}_k^\top \mathbf{Q}_2 \mathbf{V}_k)^{-1} \mathbf{D}_k(\gamma)^\top \mathbf{D}_k(\gamma)) \\
&= \operatorname{tr}(((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k + \lambda^2 \mathbf{V}_k^\top \mathbf{Q} \mathbf{V}_k)^{-1} (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)^\top (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)) \\
&= \operatorname{tr}((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k + \lambda^2 \mathbf{V}_k^\top \mathbf{Q} \mathbf{V}_k)^{-1} (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_k)^\top) \\
&\to \operatorname{tr}((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n)((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n)^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n + \lambda^2 \mathbf{V}_n^\top \mathbf{Q} \mathbf{V}_n)^{-1} (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n)^\top) \\
&= \operatorname{tr}((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n) \mathbf{V}_n^{-1} ((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q})^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} + \lambda^2 \mathbf{Q})^{-1} \mathbf{V}_n^{-\top} (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} \mathbf{V}_n)^\top) \\
&= \operatorname{tr}((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q})((\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q})^\top \mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q} + \lambda^2 \mathbf{Q})^{-1} (\mathbf{L}_\mathbf{R} \mathbf{A} \mathbf{Q})^\top) \\
&= \operatorname{tr}(A(\gamma,\lambda))
\end{aligned} \tag{A.9}$$

with the invertible $\mathbf{V}_n$ since $\mathbf{V}_n^\top \mathbf{Q}_1 \mathbf{V}_n = \mathbf{I}_n$ and $\mathbf{V}_n \in \mathbb{R}^{n \times n}$ is square matrix. □

# Appendix B

# Appendix for Section 4

## B.1 Reformulation to augmented variables

This is derivation from (4.3) to (4.9). For the data fit term,

$$\begin{aligned}
\frac{1}{2}\|\mathbf{As}-\mathbf{d}\|^2_{\mathbf{R}^{-1}} &= \frac{1}{2}\left\|\begin{bmatrix}\mathbf{A} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{s}\\ \boldsymbol{\beta}\end{bmatrix}-\mathbf{d}\right\|^2_{\mathbf{R}^{-1}}\\
&= \frac{1}{2}\left\|\begin{bmatrix}\mathbf{A} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}+\mathbf{X}\boldsymbol{\mu}_\beta\\ \mathbf{q}+\boldsymbol{\mu}_\beta\end{bmatrix}-\mathbf{d}\right\|^2_{\mathbf{R}^{-1}}\\
&= \frac{1}{2}\left\|\begin{bmatrix}\mathbf{A} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ \mathbf{q}\end{bmatrix}+\mathbf{AX}\boldsymbol{\mu}_\beta-\mathbf{d}\right\|^2_{\mathbf{R}^{-1}}\\
&= \frac{1}{2}\left\|\begin{bmatrix}\mathbf{A} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ \mathbf{q}\end{bmatrix}-\mathbf{y}\right\|^2_{\mathbf{R}^{-1}}
\end{aligned}$$

and for the regularization terms in (4.3), and we have

$$\begin{aligned}
\frac{\lambda^2}{2}\|\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\|^2_{\mathbf{Q}^{-1}}+\frac{\lambda_\beta^2}{2}\|\boldsymbol{\beta}-\boldsymbol{\mu}_\beta\|^2_{\mathbf{Q}_\beta^{-1}} &= \frac{\lambda^2}{2}\|\mathbf{p}-\mathbf{X}\mathbf{q}\|^2_{\mathbf{Q}^{-1}}+\frac{\lambda_\beta^2}{2}\|\mathbf{q}\|^2_{\mathbf{Q}_\beta^{-1}} \\
&= \frac{1}{2}(\lambda^2\mathbf{p}^\top\mathbf{Q}^{-1}\mathbf{p}-2\lambda^2\mathbf{p}^\top\mathbf{Q}^{-1}\mathbf{X}\mathbf{q}+\mathbf{q}^\top(\lambda^2\mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X}+\lambda_\beta^2\mathbf{Q}_\beta^{-1})\mathbf{q}) \\
&= \frac{1}{2}(\lambda^2\mathbf{p}^\top\mathbf{Q}^{-1}\mathbf{p}-2\lambda^2\mathbf{p}^\top\mathbf{Q}^{-1}\mathbf{X}\mathbf{q}+\mathbf{q}^\top(\lambda^2\mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X}+(\gamma\lambda)^2\mathbf{Q}_\beta^{-1})\mathbf{q}) \\
&= \frac{1}{2}\begin{bmatrix}\mathbf{p}^\top & \mathbf{q}^\top\end{bmatrix}\begin{bmatrix}\lambda^2\mathbf{Q}^{-1} & -\lambda^2\mathbf{Q}^{-1}\mathbf{X} \\ -\lambda^2\mathbf{X}^\top\mathbf{Q}^{-1} & \lambda^2\mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X}+(\gamma\lambda)^2\mathbf{Q}_\beta^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{p}\\ \mathbf{q}\end{bmatrix} \\
&= \frac{\lambda^2}{2}\left\|\begin{bmatrix}\mathbf{p}\\ \mathbf{q}\end{bmatrix}\right\|^2_{\widetilde{\mathbf{Q}}^{-1}}.
\end{aligned}$$

## B.2 Derivation of the augmented prior covariance matrix

This is derivation of (4.7). Since we need $\widetilde{\mathbf{Q}}$ and not $\widetilde{\mathbf{Q}}^{-1}$ in genHyBR, we use the formula for the inverse of $2\times 2$ block matrix,

$$\begin{bmatrix}\mathbf{E} & \mathbf{F}\\ \mathbf{G} & \mathbf{H}\end{bmatrix}^{-1}=\begin{bmatrix}(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1} & -(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1}\mathbf{F}\mathbf{H}^{-1}\\ -\mathbf{H}^{-1}\mathbf{G}(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1} & \mathbf{H}^{-1}\mathbf{G}(\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G})^{-1}\mathbf{F}\mathbf{H}^{-1}+\mathbf{H}^{-1}\end{bmatrix}$$

if $\mathbf{H}$ and $\mathbf{E}-\mathbf{F}\mathbf{H}^{-1}\mathbf{G}$ are invertible. From this formula, we get

$$\begin{bmatrix}\mathbf{Q}+\frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta\mathbf{X}^\top & \frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta\\ \frac{1}{\gamma^2}\mathbf{Q}_\beta\mathbf{X}^\top & \frac{1}{\gamma^2}\mathbf{Q}_\beta\end{bmatrix}^{-1}=\begin{bmatrix}\mathbf{Q}^{-1} & -\mathbf{Q}^{-1}\mathbf{X}\\ -\mathbf{X}^\top\mathbf{Q}^{-1} & \mathbf{X}^\top\mathbf{Q}^{-1}\mathbf{X}+\gamma^2\mathbf{Q}_\beta^{-1}\end{bmatrix}$$

so that

$$\widetilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} + \frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta\mathbf{X}^\top & \frac{1}{\gamma^2}\mathbf{X}\mathbf{Q}_\beta \\ \frac{1}{\gamma^2}\mathbf{Q}_\beta\mathbf{X} & \frac{1}{\gamma^2}\mathbf{Q}_\beta \end{bmatrix} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \frac{1}{\gamma^2}\begin{bmatrix} \mathbf{X} \\ \mathbf{I} \end{bmatrix} \mathbf{Q}_\beta \begin{bmatrix} \mathbf{X}^\top & \mathbf{I} \end{bmatrix}.$$