# Gradient-, Ensemble-, and Adjoint-Free Data-Driven Parameter Estimation

by

Ankit Goel

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in The University of Michigan
2019

Doctoral Committee:

Professor Dennis S. Bernstein, Chair
Professor Karthik Duraisamy
Assistant Professor Alex Gorodetsky
Professor Aaron Ridley

Ankit Goel

ankgoel@umich.edu

ORCID ID: 0000-0002-4146-6275

# ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Dennis Bernstein for giving me the opportunity to work under his guidance during last several years. This dissertation would not have been possible without his infallible support, encouragement, and numerous memorable discussions. Also, I would like to thank my committee members Professor Karthik Duraisamy, Professor Aaron Ridley, and Professor Alex Gorodetsky for giving me challenging projects and providing helpful comments and suggestions. I am extremely grateful for the opportunity to study at the University of Michigan, and learn from some of the best professors.

I would like to also acknowledge all my friends and colleagues in Ann Arbor: Sneha Sanjeevini, Ambarish Desai, Nitin Sharma, Pradyumna Srinivasan, Albert Thomas, Sushant Choudhary, Pradeep Kumar, Ahmad Ansari, Ming-Jui (Ray) Yu, Antai Xie, Khaled Aljanaideh, Yousaf Rahman, Frant Sobolic, Syed Aseem Ul Islam, Adam Bruce, Angeline G. Burrell, Pedro Di Donato, Robert Zidek, Sweewarman Balachandran, Gianluca Kapiris, Christian Spiess, who have made my time in the graduate school very enjoyable.

Finally, I would like to thank my parents and my brother for their unconditional love, support, and encouragement. I am very grateful for the family I have.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# LIST OF TABLES

**Table**

# ABSTRACT

In many applications, models of physical systems have known structure but unknown parameters. By viewing the unknown parameters as constant states, nonlinear estimation methods such as the extended Kalman filter, unscented Kalman filter, and ensemble Kalman filter can be used to estimate the states of the augmented system, thereby providing estimates of the parameters along with the dynamic states. These methods tend to be computationally expensive due to the need for Jacobians, ensembles, or adjoints, especially when the models are high-dimensional.

This dissertation presents retrospective cost parameter estimation (RCPE), which does not require gradients, ensembles, or adjoints. Rather, RCPE estimates unknown parameters from a single trajectory, and requires updating an adaptive integrator gain for each unknown parameter. RCPE is applicable to parameter estimation in linear and nonlinear models, where the parameterization may be either affine or nonaffine.

The main contribution of this work is to show that the parameter estimates may be permuted in an arbitrary way, and thus a permutation is needed to correctly associate each parameter estimate with the corresponding unknown parameter. RCPE is illustrated through several numerical examples including the Burgers equation and the Global Ionosphere Thermosphere Model (GITM), where the goal is to estimate representational parameters such as eddy diffusion coefficient and thermal conductivity coefficients using measurements of atmospheric variables such as total electron content, density, temperatures etc.

The next part of the dissertation focuses on forgetting in the context of recursive least squares (RLS) algorithm. It is a well-known fact that classical RLS with for-

getting diverges in the cases where the excitation is not persistent. In this work, an information-driven directional forgetting technique is proposed, which constrains the forgetting to directions in which new information is available, thereby allowing RLS to operate without divergence during periods of loss of persistency.

In the last part of this dissertation, retrospective cost adaptive control (RCAC) is extended to the problem of control allocation in overactuated systems. In particular, it is shown that the applied control input lies in the range of the target model used in RCAC, thereby providing a simple technique to constrain the control input to a desired subspace. Finally, RCAC is extended to asymptotically enforce output constraint by formulating the problem as a problem of following conflicting commands, and is used to prevent a scramjet combustor from unstarting using pressure measurements.

# CHAPTER 1

# Introduction

## 1.1 Motivation and Purpose

In many applications of science and engineering, models of physical systems have known structure but unknown parameters. In such models, often called gray-box models, the structure of the function describing the evolution of the state of the system is known, but the values of the parameters in the function may be unknown. These parameters might be embedded in the model in such a way that direct calculation is not possible due to lack of measurements. Such parameters are called *inaccessible*, which means that they relate unmeasured signals, and thus cannot be determined by regression. Furthermore, these parameters might represent the cumulative effect of a complex phenomena, and thus might not have a true measurable value.

To illustrate this problem, consider the problem of estimating the spring stiffness $k_2$ in the mass-spring system shown in Figure 1.1. The system is modeled by

$$m_1\ddot{q}_1 + (k_1 + k_2)q_1 - k_2q_2 = 0, \tag{1.1}$$

$$m_2\ddot{q}_2 - k_2q_1 + k_2q_2 = F, \tag{1.2}$$

where $q_1$ is the position of the mass $m_1$, $q_2$ is the position of the mass $m_2$, and $k_1, k_2$ are the stiffness of the first and second spring, and $F$ is the force applied to $m_2$. Note

that the relaxed length of the springs are assumed to be zero.

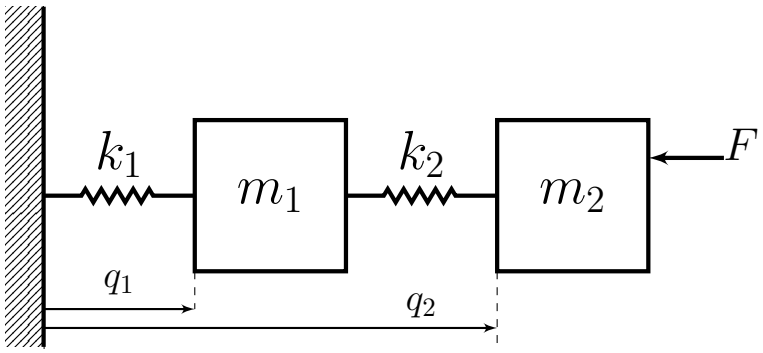



Figure 1.1: Mass-spring system with unknown spring stiffness

If appropriate measurements are available, the unknown parameters can be estimated using regression. For example, if $m_2$ is known and the acceleration $\ddot{q}_2$, the positions $q_1$ and $q_2$, and the force $F$ are measured, then

$$k_2 = \frac{F - m_2\ddot{q}_2}{q_2 - q_1}. \tag{1.3}$$

Similarly, if the reaction force $f_{k_2/m_1}$ applied to $k_2$ by $m_1$ and the reaction force $f_{k_2/m_2}$ applied to $k_2$ by $m_2$ are measured, then

$$k_2 = \frac{f_{k_2/m_1} - f_{k_2/m_2}}{q_2 - q_1}. \tag{1.4}$$

However, such measurements are not typically available.

Another approach uses the dynamic model of the system along with the measured signals to estimate the unknown parameter. Let the system be modeled by

$$y_k = f(\phi_k, \mu), \tag{1.5}$$

where $\phi_k$ contains the measured input to the system and the measured output of the system, and $\mu \in \mathbb{R}^{l_\mu}$ is the unknown parameter parameterizing the map $f$. The problem is to estimate $\mu$ using measurements $y_k$ and $\phi_k$.

The most common approach is to use least-squares formulation to estimate the unknown parameter. In this case, the estimate $\hat{\mu}_k$ of the unknown parameter $\mu$ at step $k$ is given by the minimizer of

$$J(\hat{\mu}) \triangleq \sum_{i=0}^{k} (y_i - \hat{y}_i)^{\mathrm{T}} (y_i - \hat{y}_i), \tag{1.6}$$

where $\hat{y}_i$ is the computed output of the estimation model given by

$$\hat{y}_i = f(\phi_i, \hat{\mu}). \tag{1.7}$$

for all $i \in \{0, 1, \cdots, k\}$.

If the model (1.5) is linearly parameterized by $\mu$, that is, the output $y_k$ can be written as $y_k = \phi_k \mu$, then the minimizer of (1.6) has a closed-form solution [1]. In the case where $f$ is nonlinearly parameterized by $\mu$, using analytical expression of $f$, [2] uses Taylor series expansion to iteratively minimize $J(\hat{\mu})$, [3] uses gradient-descent method to minimize $J(\hat{\mu})$, while [4] optimally interpolates the solution given by the Taylor series expansion and the gradient-descent method. The severity of nonlinearity in function $f$ and the order of the system (1.5) render minimization of $J(\hat{\mu})$ using anylytical methods difficult. In addition, these methods are plagued by slow convergence or divergence of the estimates. Note that these methods require an analytical expression of $f$ to compute the gradients. Furthermore, the computational cost increases with $k$.

In the case where analytical expression of $f$ is not available, the gradient of the cost function (1.6) can be computed by finite differences. Note that this method uses $f$ as a black-box model. However, this method becomes computationally expensive as the size of the vector $\mu$ and the number of samples $k$ increase. Alternatively, the gradient of the cost function (1.6) can be computed by variational methods [5–7], which require an adjoint formulation of the dynamics. Adjoint based methods tend

to be computationally expensive due to the need for multiple iterations of the forward model and backward adjoint. To use the adjoint method to compute the gradient, note that

$$\frac{\mathrm{d}J}{\mathrm{d}\mu} = \sum_{i=0}^{k} \frac{\partial J}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial \mu} \tag{1.8}$$

For $i \in \{0, 1, \cdots, k\}$, define $g_i(\hat{y}_i, \mu) \stackrel{\triangle}{=} \hat{y}_i - f(\phi_i, \mu)$. Note that $g_i(y_i, \mu) = 0$. Thus,

$$\frac{\partial g_i}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial \mu} + \frac{\partial g_i}{\partial \mu} = 0. \tag{1.9}$$

It follows from (1.8) and (1.9) that

$$\frac{\mathrm{d}J}{\mathrm{d}\mu} = \sum_{i=0}^{k} \lambda_i^{\mathrm{T}} \frac{\partial g_i}{\partial \mu}. \tag{1.10}$$

where $\lambda_i$ is obtained by solving

$$\frac{\partial g_i}{\partial \hat{y}_i}^{\mathrm{T}} \lambda_i = -\frac{\partial J}{\partial \hat{y}_i}^{\mathrm{T}}. \tag{1.11}$$

Note that (1.11) is called the *adjoint equation*. Various formulations based on adjoint method to estimate unknown parameter in a dynamical system modeled using state-space representation are discussed in [8–12].

As a special case of this problem, a linear system may have uncertain entries in its state space representation. For this problem, a two-step procedure is used in [13], where a black-box model is first constructed based on the input-output data, and a similarity transformation is used to recover the unknown parameters. In [14], a sequential convex relaxation method is used to estimate unknown entries in the matrices of a state space realization.

The measurements of $u_k$ and $y_k$ may be corrupted by noise. In presence of noise,

the unknown parameter $\mu$ is treated as a random variable in a probabilistic framework. Consequently, the statistical characteristics of $\hat{\mu}_k$ depend on the the statistical characteristics of the noise. In such cases, the parameter estimation problem has a Bayesian interpretation. Using Bayes' rule, the conditional probability distribution of the unknown parameter given the measurements is formulated and maximized in order to obtain the parameter estimate $\hat{\mu}_k$ [5, 15–17]. In [18], parameters in the special case of a linear system are estimated using expectation maximization algorithm under the assumption of Gaussian noise.

State estimation techniques provide another framework for estimating the unknown parameters. In this framework, (1.5) is written in the state-space form as

$$x_{k+1} = f(x_k, u_k, \mu), \tag{1.12}$$

$$y_k = g(x_k, u_k, \mu), \tag{1.13}$$

where $x_k \in \mathbb{R}^{l_x}$ is the state of the system at step $k$, $u_k \in \mathbb{R}^{l_u}$ is the input to the system, $y_k \in \mathbb{R}^{l_y}$ is the measured output of the system, and $\mu \in \mathbb{R}^{l_\mu}$ is the unknown parameter parameterizing the dynamics map $f$ and $g$. Usually, the state $x_k$ of the system (1.12) is not available to estimate $\mu$. By viewing the unknown parameters as constant states, and augmenting the original states with the constant states, state estimation techniques can be used to estimate the states of the augmented system, thereby providing estimates of the parameters along with the dynamic states [19]. However, the parameter states multiply the dynamic states, thus the resulting estimation dynamics are nonlinear irrespective of whether the "original" dynamics (1.12), (1.13) are linear or nonlinear. Consequently, nonlinear state estimation techniques such as the extended Kalman filter (EKF), unscented Kalman filter (UKF), and ensemble Kalman filter (EnKF) can be applied to these problems [19–25]. For example, to estimate $k_2$ in (1.1), (1.2) using Kalman filter based estimation techniques, the

estimation model is constructed as

$$\dot{\hat{x}} = f(\hat{x}, u), \tag{1.14}$$

where

$$\hat{x} \triangleq \begin{bmatrix} \hat{q}_1 & \dot{\hat{q}}_1 & \hat{q}_2 & \dot{\hat{q}}_2 & \hat{k}_2 \end{bmatrix}^{\mathrm{T}}, \tag{1.15}$$

$$f(x,u) = \begin{bmatrix} x_2 \\ \dfrac{1}{m_1}(x_5 x_3 - k_1 x_1 - x_5 x_1) \\ x_4 \\ \dfrac{1}{m_2}(u - x_5 x_3 + x_5 x_1) \\ 0 \end{bmatrix}. \tag{1.16}$$

Note that the estimation model is nonlinear.

In the problem of state estimation in a linear system, Kalman filter minimizes the covariance of the state estimate error to compute the optimal state estimate given the correct knowledge of process noise and measurement noise characteristics. However, Kalman filter based techniques are robust to noise characteristics [26]. In the noise-free case, Kalman filter based techniques can be applied to estimate the augmented state where the covariance variables can be used to tune the convergence rate of the estimator.

The application of EKF to estimate the unknown parameters requires the computation of Jacobian of the augmented dynamics at each time step. Note that this requires access to the full state of the estimation model to compute the Jacobian at each time step. The application of EnKF or UKF requires an ensemble of models. In particular, UKF is based on an esnsemble of $2N + 1$ models, where $N = n + p$, $n$ is number of dynamic states and $p$ is the number of unknown parameters. The total number of states that must be propagated at each iteration is thus $2N^2 + N = O(N^2)$.

Consequently, for a system with $n$ dynamic states and $p$ unknown parameters, it follows that $N = n + p$, and thus the total number of states that must be propagated at each iteration is $(2N + 1)N = [2(n + p) + 1](n + p)$. Note that UKF also requires access to the full state of the estimation model to construct the ensemble.

The purpose of this dissertation is to present a gradient free, ensemble free, and adjoint free data driven parameter estimation technique. In particular, this dissertation presents retrospective cost parameter estimation (RCPE) algorithm for estimating multiple unknown parameters in linear and nonlinear systems with affine or nonaffine parameterizations.

Like UKF but unlike EKF, RCPE does not require a Jacobian of the dynamics in order to update the parameter estimates. However, unlike UKF, RCPE does not require an ensemble of models. In contrast to UKF, RCPE requires the propagation of only a single copy of the "original" system dynamics, so that the number of states that must be propagated at each iteration is simply $n$. In addition, unlike EKF and UKF, RCPE does not require access to the states of the estimation model. For both UKF and RCPE, this model need only be given as an executable simulation; explicit knowledge of the equations and source code underlying the simulation is not required. Finally, unlike variational methods, RCPE does not require an adjoint model. However, the price paid for not requiring an explicit model or an ensemble of models is the need within RCPE to select a permutation matrix that correctly associates each parameter estimate with the corresponding unknown parameter.

The retrospective cost parameter estimation algorithm is a variation of retrospective cost model refinement (RCMR) developed in [27–29] and is based on retrospective cost adaptive control (RCAC) [30]. RCMR was developed to estimate unknown parameters in an affinely parameterized dynamical system. To estimate $\mu$ using RCMR, (1.12) is written as

$$x_{k+1} = f(x_k, u_k) + \sum_{i=1}^{l_\mu} v_{i,k}, \tag{1.17}$$

$$v_{i,k} = \mu_i w_{i,k}, \tag{1.18}$$

$$w_{i,k} = h_i(x_k, u_k). \tag{1.19}$$

Note that $\mu_i$ appears as a static feedback in the dynamics given by (1.17). RCMR minimizes a retrospective cost function based on the measurements of $v_{i,k}$ and $w_{i,k}$ whose minimizer provides the parameter estimate $\hat{\mu}_k$. Although RCMR does not use $h_i(x_k, u_k)$ to compute the estimate $\hat{\mu}_k$, RCMR does require that the function $h_i(x_k, u_k)$ be available to compute $w_{i,k}$. These requirements prohibits RCMR to be applicable to problems where the dynamics (1.12) is nonlinearly parameterized or the dynamics (1.12) are so complicated that construction of the functions $h_i(x_k, u_k)$ is cumbersome.

On the other hand, RCPE is applicable to parameter estimation in linear and nonlinear models, where the parameterization may be either affine or nonaffine. In order to update the parameter estimate, RCPE uses an error signal given by the difference between the output of the system and the output of the estimation model. The parameter update is obtained by minimizing a retrospective cost function whose minimizer provides an update of the gains of an integrator. The output of the adaptive integrator consists of the parameter pre-estimates, whose absolute values are the parameter estimates. However, the parameter estimates may be permuted in an unknown way, and thus a permutation is needed to correctly associate each parameter estimate with the corresponding unknown parameter.

## 1.2 Contributions

The major contributions of the dissertation are listed below.

1. Development of RCPE, which is a gradient-, ensemble-, and adjoint-free data driven parameter estimation algorithm applicable to affinely or nonaffinely parameterized systems.

2. Analysis of the retrospective cost in RCPE to show that the parameter estimate is constrained to the subspace defined by the filter coefficients used to define the retrospective cost.

3. Systematic demonstration of RCPE on low-order systems and high-dimensional systems such as the Burgers equation and the global ionosphere thermosphere model to show the effect of the ordering of the filter coefficients and the necessity of the permutation matrix.

4. Formulation of the biquadratic retrospective cost to simultaneously optimize the adaptive integrator gains and the filter coefficients in RCPE.

5. Analysis of asymptotic convergence of recursive least squares algorithm using discrete-time Lyapunov theory and asymptotic bounds on RLS variables under persistent excitation.

6. Development of the information-driven directional forgetting scheme in recursive least squares algorithm to constrain forgetting to the information subspace in the case of lack of persistent excitation.

7. Extension of RCAC to constrain the input to a desired input subspace and application to the problem of control allocation in wide systems.

8. Extension of RCAC to enforce output constrains asymptotically and application to a 2D scramjet combustor to prevent unstart.

## 1.3 Dissertation Outline

This disseration is organized as follows.

### Chapter 2 Summary

Chapter 2 presents the parameter estimation problem in dynamical systems. First, The parameter estimator, consisting of an adaptive integrator, permutation, and a nonlinear transformation to the first orthant is presented. Then, the retrospective cost parameter estimation (RCPE) algorithm is presented. Next, it is shown in Theorem 2.1 that the parameter estimates produced by RCPE are constrained to lie in a subspace defined by the filter coefficients that define the retrospective cost. Then, the need of permutation is shown through several numerical examples. Finally, RCPE is applied to high-dimensional nonlinear systems.

### Chapter 3 Summary

Chapter 3 extends RCPE by simultaneously optimizing the the filter coefficients that define the retrospective cost. It is shown that that the optimization problem is biquadratic but nonconvex. An alternating convex search is used to converge to a local minimizer. The extended algorithm is demonstrated on a low-dimensional system, and finally, it is used to estimate eddy diffusion coefficient (EDC) in global ionosphere-thermosphere model (GITM) by using measurements of total electron content (TEC).

### Chapter 4 Summary

Chapter 4 presents a novel directional forgetting algorithm to prevent estimator divergence under lack of persistency in the context of recursive least squares. Various results on the effect of forgetting with and without the persistence of excitation are

presented, and it is shown that some singular values of the covariance matrix diverge in the case where excitation is not persistent. Finally, a matrix forgetting scheme is proposed which constrains forgetting to the directions receiving new information.

### Chapter 5 Summary

Chapter 5 presents the control allocation problem. In the context of the retrospective cost adaptive control (RCAC), it is shown that the the control input lies in the range of the target model. An extension of Theorem 2.1 that includes control penalty in the retrospective cost is presented. Finally, numerical examples demonstrate control allocation in wide plant using RCAC.

### Chapter 6 Summary

Chapter 6 extends RCAC to enforce auxiliary output constraints. First, it is shown that in the case of conflicting commands, RCAC trades off output error based on the choice of the target model. Next, the problem of output constraints is formulated as a command-following problem with conflicting commands. Although the constraint is violated, it is shown that the asymptotic magnitude of the constraint violation error can be arbitrarily reduced by tuning the target model in RCAC. Finally, the extended RCAC algorithm is used to prevent unstart in a two-dimensional scramjet combustor model.

Finally, the thesis is concluded in Chapter 7.

## 1.4 Publications

The following is the list of publications relevant to the research presented in this dissertation.

### 1.4.1 Journal Articles

- A. Goel, and D. S. Bernstein, "Gradient-, Ensemble-, and Adjoint-Free Data-Driven Parameter Estimation", Accepted for publication in *AIAA Journal of Guidance, Control, and Dynamics*.

- A. Goel, and D. S. Bernstein, " Recursive Least Squares with Directional Forgetting", submitted to *IEEE Control System Magazine.*

- A. Goel, K. Duraisamy, and D. S. Bernstein, "Retrospective Cost Adaptive Control of Unstart in a Model Scramjet Combustor", *AIAA Journal*, vol. 56, no. 3, pp. 1085–1096, 2018.

- A. G. Burrell, A. Goel, A. J. Ridley, and D. S. Bernstein, "Correction of the Photoelectron Heating Efficiency Within the Global Ionosphere-Thermosphere Model Using Retrospective Cost Model Refinement", *Journal of Atmospheric Solar-Terrestrial Physics*, Vol. 124, pp. 30–38, 2015.

### 1.4.2 Peer–reviewed Conference Papers

- A. Goel and D. S. Bernstein, "Data-Driven Parameter Estimation for Models with Nonlinear Parameter Dependence," in Proceedings of *Conference on Decision and Control*, pp. 1470–1475, Miami, FL, Dec 2018.

- A. Goel , K. Duraisamy, and D. S. Bernstein, "Output-Constrained Adaptive Control for Unstart Prevention in a 2D Scramjet Combustor ," *AIAA Scitech 2019 Forum* , San Diego, CA, Jan 2019.

- A. Goel, A. Ansari and D. S. Bernstein, "Adaptive Squaring-Based Control Allocation for Wide Systems with Application to Lateral Flight Control," in Proceedings of *Conference on Decision and Control*, pp. 5140–5145, Miami, FL, Dec 2018.

- A. Goel and D. S. Bernstein, "A Targeted Forgetting Factor for Recursive Least Squares," in Proceedings of *Conference on Decision and Control*, pp. 3899–3903, Miami, FL, Dec 2018.

- A. Goel and D. S. Bernstein, "Parameter Estimation for Nonlinearly Parameterized Gray-Box Models," in Proceedings of *American Control Conference*, pp. 5280–5285, Milwaukee, WI, June 2018.

- A. Goel, A. Ridley, and D. S. Bernstein, "Estimation of the Eddy Diffusion Coefficient Using Total Electron Content Data," in Proceedings of *American Control Conference*, pp. 3298–3303, Milwaukee, WI, June 2018.

- A. Goel, K. Duraisamy and D. S. Bernstein, "Parameter estimation in the Burgers equation using retrospective-cost model refinement," in Proceedings of *American Control Conference*, pp. 6983–6988, Boston, MA, July 2016.

# CHAPTER 2

# Gradient-, Ensemble-, and Adjoint-Free Data-Driven Parameter Estimation

## 2.1 Introduction

In many applications, models of physical systems have known structure but unknown parameters. By viewing the unknown parameters as constant states, nonlinear estimation methods can be used to estimate the states of the augmented system, thereby providing estimates of the parameters along with the dynamic states [19]. The extended Kalman filter (EKF), unscented Kalman filter (UKF), and ensemble Kalman filter (EnKF) can be applied to these problems [19–25]. An alternative approach to parameter estimation is variational methods [6, 7, 31], which require an adjoint formulation of the dynamics, These methods tend to be computationally expensive due to the need for multiple iterations of the forward model and backward adjoint.

As a special case of this problem, a linear system may have uncertain entries in its state space representation. Since the parameter states multiply the dynamic states, the resulting estimation dynamics are nonlinear despite the fact that the "original" dynamics are linear. For this problem, a two-step procedure is used in [13], where a black-box model is first constructed based on the input-output data, and a similarity

transformation is used to recover the unknown parameters. In [14], a sequential convex relaxation method is used to estimate unknown entries in the matrices of a state space realization.

The present chapter focuses on retrospective cost parameter estimation (RCPE), which is a variation of retrospective cost model refinement (RCMR) developed in [27–29] and is based on retrospective cost adaptive control [30]. RCPE is applicable to parameter estimation in linear and nonlinear models, where the parameterization may be either affine or nonaffine. In order to update the parameter estimate, RCPE uses an error signal given by the difference between the output of the physical system and the output of the estimation model. The parameter update is obtained by minimizing a retrospective cost function whose minimizer provides an update of the gains of an integrator. The output of the adaptive integrator consists of the parameter pre-estimates, whose absolute values are the parameter estimates. However, the parameter estimates may be permuted in an unknown way, and thus a permutation is needed to correctly associate each parameter estimate with the corresponding unknown parameter.

Like UKF but unlike EKF, RCPE does not require a Jacobian of the dynamics in order to update the parameter estimates. However, unlike UKF, RCPE does not require an ensemble of models. In particular, for parameter estimation, the unscented Kalman filter (UKF) is based on an ensemble of $2N+1$ models, where $N = n+p$, $n$ is number of dynamic states and $p$ is the number of unknown parameters. The total number of states that must be propagated at each iteration is thus $2N^2+N = O(N^2)$. Consequently, for a system with $n$ dynamic states and $p$ unknown parameters, it follows that $N = n+p$, and thus the total number of states that must be propagated at each iteration is $(2N+1)N = [2(n+p)+1](n+p)$. In contrast to UKF, RCPE requires the propagation of only a single copy of the "original" system dynamics, so that the number of states that must be propagated at each iteration is simply $n$. For both

UKF and RCPE, this model need only be given as an executable simulation; explicit knowledge of the equations and source code underlying the simulation is not required. However, the price paid for not requiring an explicit model or an ensemble of models is the need within RCPE to select a permutation matrix that correctly associates each parameter estimate with the corresponding unknown parameter. Finally, unlike variational methods, RCPE does not require an adjoint model.

The contribution of the present chapter is to present, analyze, and demonstrate the RCPE algorithm for estimating multiple unknown parameters in linear and nonlinear systems with affine or nonaffine parameterizations. RCPE is shown to be applicable without explicit knowledge of the system equations, and thus is implementable using only an executable simulation. The chapter analyzes the effect of the filter coefficients in determining the search directions leading to the parameter estimates. Most importantly, this chapter demonstrates the need for the permutation matrix in problems with multiple unknown parameters. Finally, a numerical example with 101 dynamic states and two parameter states shows that the computation required by RCPE (202 propagated states) is substantially less than the computation required by UKF (21,321 propagated states).

The chapter is structured as follows. Section 2.2 describes the parameter-estimation problem. Section 2.3 describes the RCPE algorithm. Next, section 2.4 analyzes the effect of the user-defined filter in RCPE on the performance of the parameter estimator. Sections 2.5-2.9 present several numerical examples (summarized in Table 2.1) demonstrating the application of RCPE and its features.

| Example | System | Parameterization | $l_\mu$ | $l_y$ | Objective |
|---|---|---|---|---|---|
| 2.1 | Linear | Affine | 1 | 1 | Effect of $x(0)$ and $u$ on the choice of $N_1$ |
| 2.2 | Linear | Nonaffine | 1 | 1 | Effect of noise and comparison with UKF |
| 2.3 | Linear | Nonaffine | 2 | 2 | Effect of sparse $G_\mathrm{f}$ |
| 2.4 | Linear | Affine | 2 | 1 | Choice of $N_1, N_2$, and $\mathcal{O}_\mathrm{p}$ |
| 2.5 | Nonlinear | Nonaffine | 2 | 1 | Choice of $\mathcal{O}_\mathrm{p}$ with fixed $N_1, N_2$ |
| 2.6 | Linear | Affine | 3 | 1 | Choice of $N_1, N_2, N_3$ with fixed $\mathcal{O}_\mathrm{p}$ |
| 2.7 | Nonlinear | Affine | 3 | 1 | Choice of $\mathcal{O}_\mathrm{p}$ with fixed $N_1, N_2, N_3$ |
| 2.9 | Nonlinear | Affine | 2 | 1 | High-dimensional application |

Table 2.1: Summary of the numerical examples.

## 2.2 Parameter-Estimation Problem

Consider the discrete-time system

$$x(k+1) = f(x(k), u(k), \mu) + w_1(k), \tag{2.1}$$

$$y(k) = h(x(k), u(k), \mu) + w_2(k), \tag{2.2}$$

where $x(k) \in \mathbb{R}^{l_x}$ is the state, $u(k) \in \mathbb{R}^{l_u}$ is the measured input, $y(k) \in \mathbb{R}^{l_y}$ is the measured output, $w_1(k) \in \mathbb{R}^{l_x}$ is the process noise, $w_2(k) \in \mathbb{R}^{l_y}$ is the measurement noise, and $\mu = [\mu_1 \ \cdots \ \mu_{l_\mu}]^\mathrm{T} \in \mathcal{M} \subseteq \mathbb{R}^{l_\mu}$ is the *true parameter*, which is unknown. The

set $\mathcal{M}$ is assumed to be known and satisfy $\mathcal{M} \subseteq [0, \infty)^{l_\mu}$, that is, $\mathcal{M}$ is contained in the nonnegative orthant. If $\mathcal{M}$ does not satisfy this condition, then it may be possible to replace $\mathcal{M}$ by $\mathcal{M}' \stackrel{\triangle}{=} \overline{\mu} + \mathcal{M}$ and $\mu$ by $\mu - \overline{\mu}$ in (2.1), (2.2), where $\overline{\mu} \in \mathbb{R}^{l_\mu}$ shifts $\mathcal{M}$ such that $\mathcal{M}'$ is contained in the nonnegative orthant. With this transformation, which can always be done if $\mathcal{M}$ is bounded, it can be assumed that $\mu$ is an element of the nonnegative orthant. The system (2.1), (2.2) is viewed as the *truth model* of a physical system.

Based on (2.1), (2.2), the *estimation model* is constructed as

$$\hat{x}(k+1) = f(\hat{x}(k), u(k), \hat{\mu}(k)), \tag{2.3}$$

$$\hat{y}(k) = h(\hat{x}(k), u(k), \hat{\mu}(k)), \tag{2.4}$$

where $\hat{x}(k)$ is the computed state, $\hat{y}(k)$ is the computed output of (2.3), (2.4), and $\hat{\mu}(k)$ is the *parameter estimate*. It is assumed that $f$ and $h$ are known, and thus they can be used to construct (2.3), (2.4). Since $w_1(k)$ and $w_2(k)$ are unknown, they do not appear in (2.3), (2.4). Since $\mu$ is unknown, it is replaced by $\hat{\mu}(k)$ in (2.3), (2.4). The objective is to construct $\hat{\mu}(k)$ based on the *output error* $z(k) \in \mathbb{R}^{l_y}$ defined by

$$z(k) \stackrel{\triangle}{=} \hat{y}(k) - y(k). \tag{2.5}$$

The ability to estimate $\mu$ is based on the assumption that (2.1), (2.2) is structurally identifiable [32–34] and the data are sufficiently persistent [35, 36].

Since measurements of only $y$ are available, the state $x$ is unknown, and thus $x(0)$ is unknown. For all examples in this chapter, the initial state of the estimation model (2.3), (2.4) is chosen to be zero to reflect the absence of additional modeling information. However, the initial state of (2.1), (2.2) is unknown and nonzero.

**Definition 2.1.** The system (2.1), (2.2) is *affinely parameterized* if there exist func-

tions $f_0, f_1, \ldots, f_{l_\mu}$ and $h_0, h_1, \ldots, h_{l_\mu}$ such that

$$f(x,u,\mu) = f_0(x,u) + \sum_{i=1}^{l_\mu} \mu_i f_i(x,u), \tag{2.6}$$

$$h(x,u,\mu) = h_0(x,u) + \sum_{i=1}^{l_\mu} \mu_i h_i(x,u). \tag{2.7}$$

Otherwise, (2.1), (2.2) is *nonaffinely parameterized*.

A specialization of (2.1), (2.2) is given by the linear discrete-time system

$$x(k+1) = A(\mu)x(k) + B(\mu)u(k) + w_1(k), \tag{2.8}$$

$$y(k) = C(\mu)x(k) + D(\mu)u(k) + w_2(k). \tag{2.9}$$

In this case, the estimation model (2.3), (2.4) becomes

$$\hat{x}(k+1) = A(\hat{\mu}(k))\hat{x}(k) + B(\hat{\mu}(k))u(k), \tag{2.10}$$

$$\hat{y}(k) = C(\hat{\mu}(k))\hat{x}(k) + D(\hat{\mu}(k))u(k). \tag{2.11}$$

**Definition 2.2.** The linear system (2.8), (2.9) is *affinely parameterized* if there exist constant matrices $A_0, A_1, \ldots, A_{l_\mu} \in \mathbb{R}^{l_x \times l_x}$, $B_0, B_1, \ldots, B_{l_\mu} \in \mathbb{R}^{l_x \times l_u}$, $C_0, C_1, \ldots, C_{l_\mu} \in \mathbb{R}^{l_y \times l_x}$, and $D_0, D_1, \ldots, D_{l_\mu} \in \mathbb{R}^{l_y \times l_u}$ such that

$$A(\mu) = A_0 + \sum_{i=1}^{l_\mu} \mu_i A_i, \quad B(\mu) = B_0 + \sum_{i=1}^{l_\mu} \mu_i B_i, \tag{2.12}$$

$$C(\mu) = C_0 + \sum_{i=1}^{l_\mu} \mu_i C_i, \quad D(\mu) = D_0 + \sum_{i=1}^{l_\mu} \mu_i D_i. \tag{2.13}$$

Otherwise, (2.8), (2.9) is *nonaffinely parameterized*.

## 2.3 Retrospective Cost Parameter Estimation

This section presents retrospective cost parameter estimation (RCPE). RCPE uses the estimation model (2.3), (2.4) along with a parameter estimator to construct $\hat{\mu}(k)$. The parameter estimator constructs $\hat{\mu}(k)$ by minimizing a cost function based on the output error $z$.

### 2.3.1 Parameter Estimator

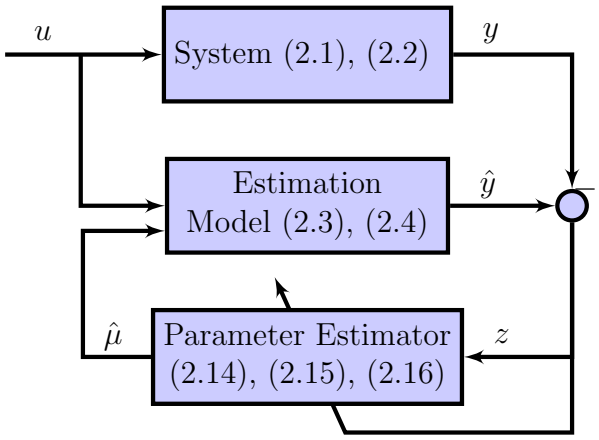



Figure 2.1: Retrospective cost parameter estimation.

The *parameter estimator* consists of an adaptive integrator and an output nonlinearity. In particular, the *parameter pre-estimate* $\nu$ is given by

$$\nu(k) = R(k)\phi(k), \tag{2.14}$$

where the *integrator state* $\phi(k) \in \mathbb{R}^{l_y}$ is updated by

$$\phi(k) = \phi(k-1) + z(k-1). \tag{2.15}$$

The *adaptive integrator gain* $R(k) \in \mathbb{R}^{l_\mu \times l_y}$ is updated by RCPE as described later in this section. Since $\nu(k)$ is not necessarily an element of the nonnegative orthant, an

output nonlinearity is used to transform $\nu(k)$. In particular, the parameter estimate $\hat{\mu}(k)$ is given by

$$\hat{\mu}(k) = \mathcal{O}_{\mathrm{p}}|\nu(k)|, \tag{2.16}$$

where the absolute value is applied componentwise. The matrix $\mathcal{O}_{\mathrm{p}}$ is explained below. The parameter estimator, which consists of (2.14), (2.15), (2.16), is represented in Figure 2.1. Since $z(k) \to 0$ is a necessary condition for $\phi$ to converge, the integrator (2.15) allows $z$ to converge to zero while $\phi$ converges to a finite value. Consequently, the parameter pre-estimate $\nu$ given by (2.14) can converge to a nonzero value, which, in turn, allows the parameter estimate $\hat{\mu}$, given by (2.16), to converge to $\mu$.

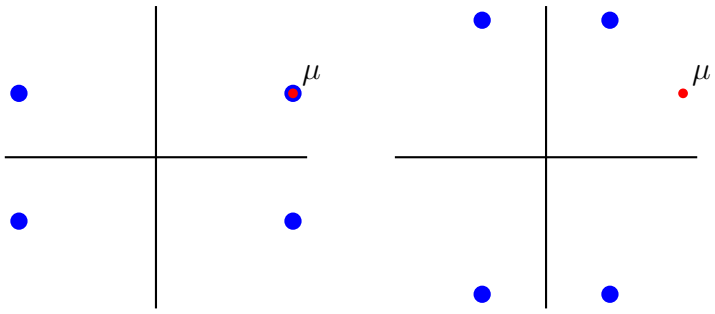
Let the $l_\mu$-tuple $\mathrm{p} = (i_1, \ldots, i_{l_\mu})$ denote a permutation of $(1, \ldots, l_\mu)$. Then the matrix $\mathcal{O}_{\mathrm{p}} \in \mathbb{R}^{l_\mu \times l_\mu}$ maps $(1, \ldots, l_\mu)$ to $(i_1, \ldots, i_{l_\mu})$. Since $\mathcal{O}_{\mathrm{p}}$ is a permutation matrix, each of its rows and columns contains exactly one "1" and the remaining entries are all zero. Specifically, row $j$ of $\mathcal{O}_{\mathrm{p}}$ is row $i_j$ of the identity matrix $I_{l_\mu}$. Now, define the set

$$\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} \stackrel{\triangle}{=} \{s \in \mathbb{R}^{l_\mu} \colon \mathcal{O}_{\mathrm{p}}|s| = \mu\}, \tag{2.17}$$

whose elements are the vectors that are mapped to $\mu$ by the componentwise absolute value and the permutation $O_{\mathrm{p}}$. For illustration, Figure 2.2(a) shows the elements of $\mathcal{S}_{\mathcal{O}_{12}}$, and Figure 2.2(b) shows the elements of $\mathcal{S}_{\mathcal{O}_{21}}$.

To facilitate the subsequent development, note that the parameter pre-estimate (2.14) can be rewritten as

$$\nu(k) = \Phi(k)\theta(k), \tag{2.18}$$




(a) p = 12. (b) p = 21.

Figure 2.2: The set $\mathcal{S}_{\mathcal{O}_\mathrm{p}}$ for $l_\mu = 2$ consists of the blue dots; $\mu$ is shown in red.

where the regressor matrix $\Phi(k)$ is defined by

$$\Phi(k) \triangleq I_{l_\mu} \otimes \phi^{\mathrm{T}}(k) \in \mathbb{R}^{l_\mu \times l_\theta}, \tag{2.19}$$

and the coefficient vector $\theta(k)$ is defined by

$$\theta(k) \triangleq \mathrm{vec}\ R(k) \in \mathbb{R}^{l_\theta}, \tag{2.20}$$

where $l_\theta \triangleq l_\mu l_y$, "$\otimes$" is the Kronecker product, and "vec" is the column-stacking operator. Note that $\theta(k)$ is an alternative representation of the adaptive integrator gain $R(k)$.

### 2.3.2 Retrospective Cost Optimization

The *retrospective error variable* is defined by

$$\hat{z}(k, \hat{\theta}) \triangleq z(k) + G_{\mathrm{f}}(\mathbf{q})[\Phi(k)\hat{\theta} - \nu(k)], \tag{2.21}$$

where $\mathbf{q}$ is the forward-shift operator and $\hat{\theta} \in \mathbb{R}^{l_\theta}$ is determined by optimization to obtain the updated coefficient vector $\theta(k+1)$. The filter $G_{\mathrm{f}}$ has the form

$$G_{\mathrm{f}}(\mathbf{q}) = \sum_{i=1}^{n_{\mathrm{f}}} \frac{1}{\mathbf{q}^i} N_i, \tag{2.22}$$

where $N_1, \ldots, N_{n_{\mathrm{f}}} \in \mathbb{R}^{l_y \times l_\mu}$ are the *filter coefficients*. Note that $G_{\mathrm{f}}$ is an $l_y \times l_\mu$ finite impulse response filter. The retrospective error variable (2.21) can thus be rewritten as

$$\hat{z}(k, \hat{\theta}) = z(k) + N\overline{\Phi}(k)\hat{\theta} - N\overline{V}(k), \tag{2.23}$$

where

$$N \triangleq [N_1 \ \cdots \ N_{n_{\mathrm{f}}}] \in \mathbb{R}^{l_y \times n_{\mathrm{f}} l_\mu}, \tag{2.24}$$

$$\overline{\Phi}(k) \triangleq \begin{bmatrix} \Phi(k-1) \\ \vdots \\ \Phi(k-n_{\mathrm{f}}) \end{bmatrix} \in \mathbb{R}^{l_\mu n_{\mathrm{f}} \times l_\theta}, \quad \overline{V}(k) \triangleq \begin{bmatrix} \nu(k-1) \\ \vdots \\ \nu(k-n_{\mathrm{f}}) \end{bmatrix} \in \mathbb{R}^{l_\mu n_{\mathrm{f}}}. \tag{2.25}$$

The retrospective cost function is defined by

$$J(k, \hat{\theta}) \triangleq \sum_{i=1}^{k} \lambda^{k-i} \hat{z}(i, \hat{\theta})^{\mathrm{T}} \hat{z}(i, \hat{\theta}) + \lambda^k \hat{\theta}^{\mathrm{T}} R_\theta \hat{\theta}, \tag{2.26}$$

where $R_\theta \in \mathbb{R}^{l_\theta \times l_\theta}$ is positive definite and $\lambda \in (0, 1]$ is the forgetting factor. The following result uses recursive least squares (RLS) to minimize (2.26).

**Proposition 2.1.** *Let $P(0) = R_\theta^{-1}$, $\theta(0) = 0$, and $\lambda \in (0, 1]$. For all $k \geq 1$, denote the minimizer of the retrospective cost function* (2.26) *by*

$$\theta(k+1) = \operatorname*{argmin}_{\hat{\theta} \in \mathbb{R}^n} J(k, \hat{\theta}). \tag{2.27}$$

*Then, for all $k \geq 1$, $\theta(k+1)$ is given by*

$$P(k+1) = \lambda^{-1}[P(k) - P(k)\overline{\Phi}(k)^{\mathrm{T}} N^{\mathrm{T}} \Gamma(k)^{-1} N \overline{\Phi}(k) P(k)], \tag{2.28}$$

$$\theta(k+1) = \theta(k) - P(k+1)\overline{\Phi}(k)^{\mathrm{T}} N^{\mathrm{T}} [N\overline{\Phi}(k)\theta(k) + z(k) - N\overline{V}(k)], \tag{2.29}$$

*where*

$$\Gamma(k) \stackrel{\triangle}{=} \lambda I_{l_y} + N\overline{\Phi}(k) P(k) \overline{\Phi}(k)^{\mathrm{T}} N^{\mathrm{T}}. \tag{2.30}$$

Furthermore, the parameter estimate at step $k+1$ is given by

$$\hat{\mu}(k+1) = \mathcal{O}_{\mathrm{p}}|\nu(k+1)| = \mathcal{O}_{\mathrm{p}}|\Phi(k+1)\theta(k+1)|. \tag{2.31}$$

Since $\theta(0) = 0$, it follows that $\nu(0) = 0$ and thus $\hat{\mu}(0) = 0$.

## 2.4 Analysis of RCPE

This section analyzes the role of the filter $G_{\mathrm{f}}$ in the update of the parameter pre-estimate $\nu$. In particular, it is shown that the filter coefficients determine the subspace of $\mathbb{R}^{l_\mu}$ that contains $\nu$.

### 2.4.1 The filter $G_{\mathrm{f}}$

To analyze the role of $G_{\mathrm{f}}$, the cost function (2.26) is rewritten as

$$J(k, \hat{\theta}) = \hat{\theta}^{\mathrm{T}} A_\theta(k) \hat{\theta} + 2 b_\theta(k)^{\mathrm{T}} \hat{\theta} + c_\theta(k), \tag{2.32}$$

where

$$A_\theta(k) \triangleq \sum_{i=1}^{k} \lambda^{k-i} \overline{\Phi}(i)^{\mathrm{T}} N^{\mathrm{T}} N \overline{\Phi}(i) + \lambda^k R_\theta, \tag{2.33}$$

$$b_\theta(k) \triangleq \sum_{i=1}^{k} \lambda^{k-i} \overline{\Phi}(i)^{\mathrm{T}} N^{\mathrm{T}} (z(i) - N\overline{V}(i)), \tag{2.34}$$

$$c_\theta(k) \triangleq \sum_{i=1}^{k} \lambda^{k-i} (z(i) - N\overline{V}(i))^{\mathrm{T}} (z(i) - N\overline{V}(i)). \tag{2.35}$$

At step $k$, the batch least squares minimizer $\theta(k+1)$ of (2.26) is given by

$$\theta(k+1) = -A_\theta(k)^{-1} b_\theta(k), \tag{2.36}$$

which is equal to $\theta(k+1)$ given by (2.29).

The following result shows that the parameter pre-estimate $\nu(k)$, and thus the estimate $\hat{\mu}(k)$, is constrained to lie in a subspace determined by the coefficients of $G_{\mathrm{f}}$.

**Theorem 2.1.** *Let $\beta > 0$, $R_\theta = \beta I_{l_\theta}$, $\nu(k)$ be given by (2.18), $\Phi(k)$ be given by (2.19), $N, \overline{\Phi}(k), \overline{V}(k)$ be given by (2.24), (2.25), and $\theta(k+1)$ be given by (2.36). Then, for all $k \geq 1$,*

$$\begin{aligned} \nu(k+1) = &-\frac{1}{\beta}[N_1^{\mathrm{T}} \ \cdots \ N_{n_{\mathrm{f}}}^{\mathrm{T}}] \cdot \\ &\sum_{i=1}^{k} \lambda^{-i} \Psi(k,i)[z(i) + N\overline{\Phi}(i)\theta(k+1) - N\overline{V}(i)] \\ &\in \mathcal{R}([N_1^{\mathrm{T}} \ \cdots \ N_{n_{\mathrm{f}}}^{\mathrm{T}}]), \end{aligned} \tag{2.37}$$

*where*

$$\Psi(k,i) \triangleq \begin{bmatrix} \phi(k+1)^{\mathrm{T}} \phi(i-1) \otimes I_{l_y} \\ \vdots \\ \phi(k+1)^{\mathrm{T}} \phi(i-n_{\mathrm{f}}) \otimes I_{l_y} \end{bmatrix}. \tag{2.38}$$

*Proof.* Note that

$$\begin{aligned}
\Phi(k+1)A_\theta(k)\theta(k+1) &= \sum_{i=1}^{k} \lambda^{k-i}\Phi(k+1)\left(\overline{\Phi}(i)^{\mathrm{T}}N^{\mathrm{T}}N\overline{\Phi}(i)\right)\theta(k+1)+ \\
&\qquad \lambda^k\beta\Phi(k+1)\theta(k+1) \\
&= \sum_{i=1}^{k}\left(\lambda^{k-i}\sum_{j=1}^{n_{\mathrm{f}}}(I_{l_u}\otimes\phi(k+1)^{\mathrm{T}})(I_{l_u}\otimes\phi(i-j))N_j^{\mathrm{T}}\right)\cdot \\
&\qquad N\overline{\Phi}(i)\theta(k+1)+\lambda^k\beta\Phi(k+1)\theta(k+1) \\
&= \sum_{i=1}^{k}\left(\lambda^{k-i}\sum_{j=1}^{n_{\mathrm{f}}}N_j^{\mathrm{T}}\phi(k+1)^{\mathrm{T}}\phi(i-j)\right)N\overline{\Phi}(i)\theta(k+1)+ \\
&\qquad \lambda^k\beta\Phi(k+1)\theta(k+1) \\
&= [N_1^{\mathrm{T}}\ \cdots\ N_{n_{\mathrm{f}}}^{\mathrm{T}}]\sum_{i=1}^{k}\lambda^{k-i}\Psi(k,i)N\overline{\Phi}(i)\theta(k+1)+ \\
&\qquad \lambda^k\beta\Phi(k+1)\theta(k+1) \qquad (2.39)
\end{aligned}$$

and

$$\begin{aligned}
\Phi(k+1)b_\theta(k) &= \Phi(k+1)\sum_{i=1}^{k}\lambda^{k-i}\overline{\Phi}(i)^{\mathrm{T}}N^{\mathrm{T}}\left(z(i)-N\overline{V}(i)\right) \\
&= [N_1^{\mathrm{T}}\ \cdots\ N_{n_{\mathrm{f}}}^{\mathrm{T}}]\sum_{i=1}^{k}\lambda^{k-i}\Psi(k,i)\left(z(i)-N\overline{V}(i)\right). \qquad (2.40)
\end{aligned}$$

Writing (2.36) as $A_\theta(k)\theta(k+1) = -b_\theta(k)$, multiplying by $\Phi(k+1)$, and using (2.39) and (2.40) yields (2.37). □

It follows from Lemma 2.1 that the parameter pre-estimate $\nu$ is constrained to lie in the subspace of $\mathbb{R}^{l_\mu}$ spanned by the coefficients of the filter used by RCPE. In addition to the subspace constraint, the numerical examples in sections 2.5–2.8 show that the feasible region is determined by the choice of the filter coefficients. The *feasible region* is the set of parameter pre-estimates in $\mathbb{R}^{l_\mu}$ that are asymptotically

reachable by the estimator. Consequently, the permutation matrix $\mathcal{O}_{\mathrm{p}}$ must be chosen such that at least one element of $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}}$, defined in (2.17), lies in the feasible region.

In view of Lemma 2.1, for all examples in this chapter where $l_z = 1$, $n_{\mathrm{f}}$ is set to be equal to $l_\mu$, and each filter coefficient is chosen to be an element of $\{e_1, e_2, \ldots, e_{l_\mu}\}$, where $e_i$ is the $i^{\mathrm{th}}$ row of the identity matrix $I_{l_\mu}$. For $l_z > 1$, the filter coefficients must be selected such that $\mu \in \mathcal{R}([N_1^{\mathrm{T}} \ \cdots \ N_{n_{\mathrm{f}}}^{\mathrm{T}}])$.

## 2.5 Examples with $l_\mu = l_y = 1$

In this section, RCPE is used to estimate one parameter in an affinely and non-affinely parameterized linear systems.

**Example 2.1.** *Affinely parameterized linear dynamics with one unknown parameter in the dynamics matrix. This example shows the effect of $u$ , $x(0)$ and $N_1$ on the feasible region.* Consider the linear system (2.8), (2.9), where

$$A(\mu) = \begin{bmatrix} \mu & 0.2 \\ 0.1 & 0.6 \end{bmatrix}, \quad B = \begin{bmatrix} 0.9 \\ 0.3 \end{bmatrix}, \quad C = \begin{bmatrix} 1.1 & 0.5 \end{bmatrix}, \tag{2.41}$$

and $\mu = 0.3$. The initial state is $x(0) = [10 \ \ 10]^{\mathrm{T}}$, the input is

$$u(k) = 2 + \sum_{j=1}^{15} \sin\left(\frac{2\pi j}{100} k\right), \tag{2.42}$$

$N_1 = 1$, $\lambda = 1$, and $R_\theta = 10^6$. Furthermore, $\mathcal{O}_1 = 1$, and thus $\mathcal{S}_1 = \{\pm\mu\}$. Figure 2.3 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $-\mu$, and thus, by (2.16), $\hat{\mu}(k) = |\nu(k)|$ converges to $\mu$. Unless stated otherwise, the abscissa of all plots denotes the iteration step.

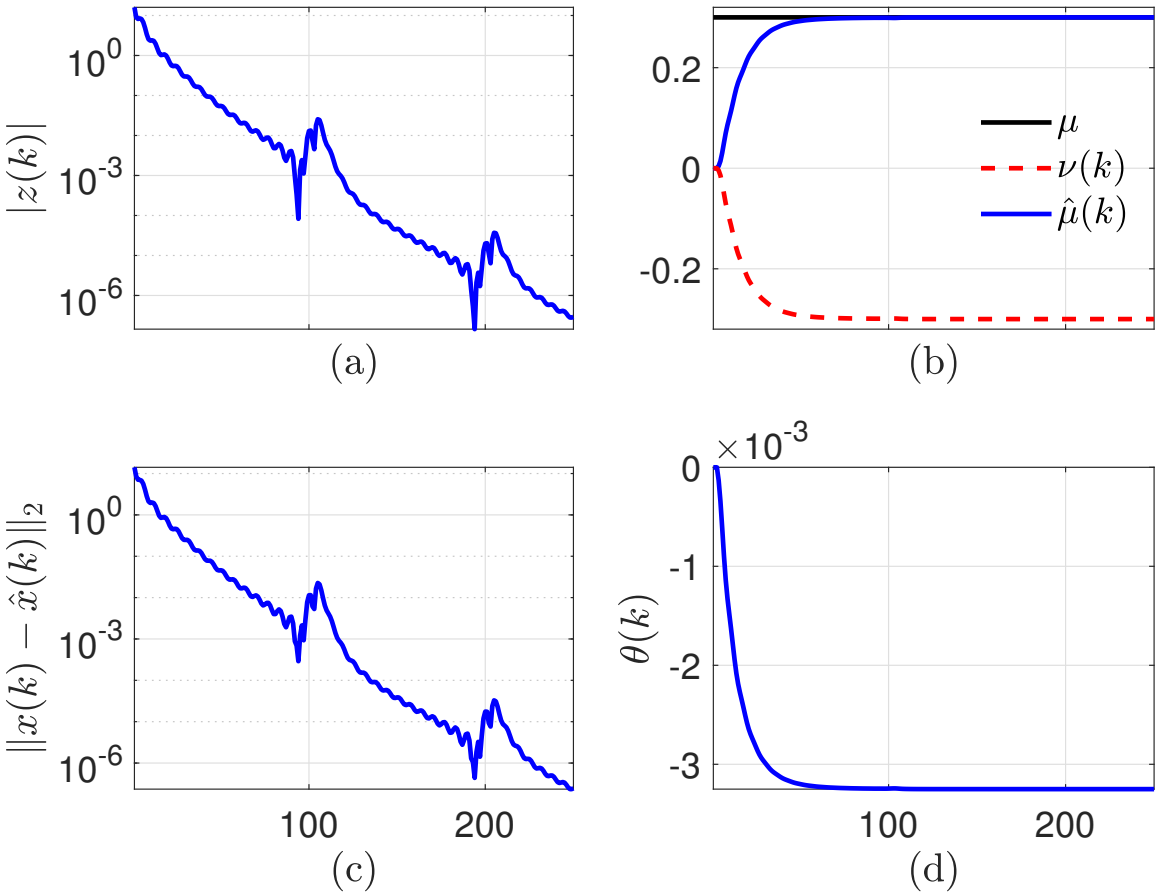
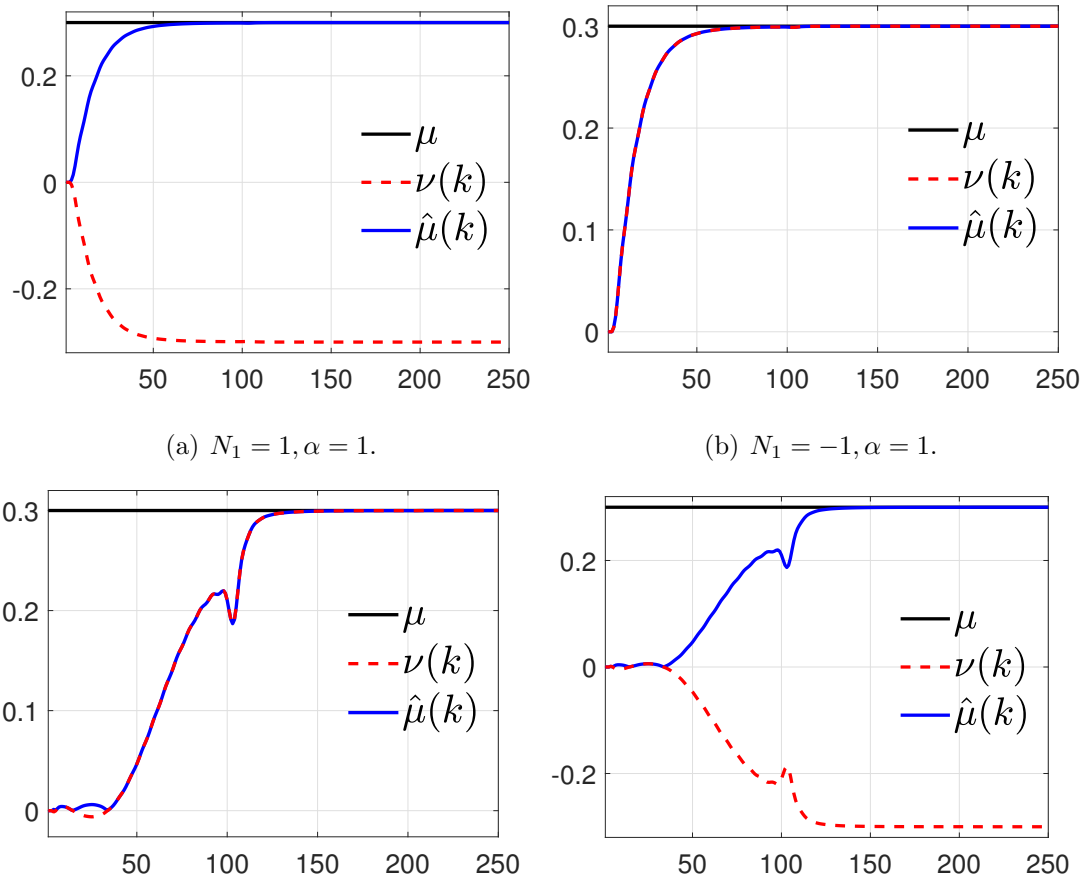



Figure 2.3: Example 2.1. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficient.

In order to investigate the effect of $u$ and $N_1$ on the feasible region, $\mu$ is estimated with the input $\alpha u$, where $u$ is given by (2.42), $\alpha = \pm 1$, and $N_1 = \pm 1$. For all four cases, Figure 2.4 shows the true parameter, parameter pre-estimate, and parameter estimate. Note that, for a given input $u$ and filter coefficient $N_1$, $\nu(k)$ converges to either $\mu$ or $-\mu$, and thus $\hat{\mu}(k)$ converges to $\mu$.




(a) $N_1 = 1, \alpha = 1$.

(b) $N_1 = -1, \alpha = 1$.

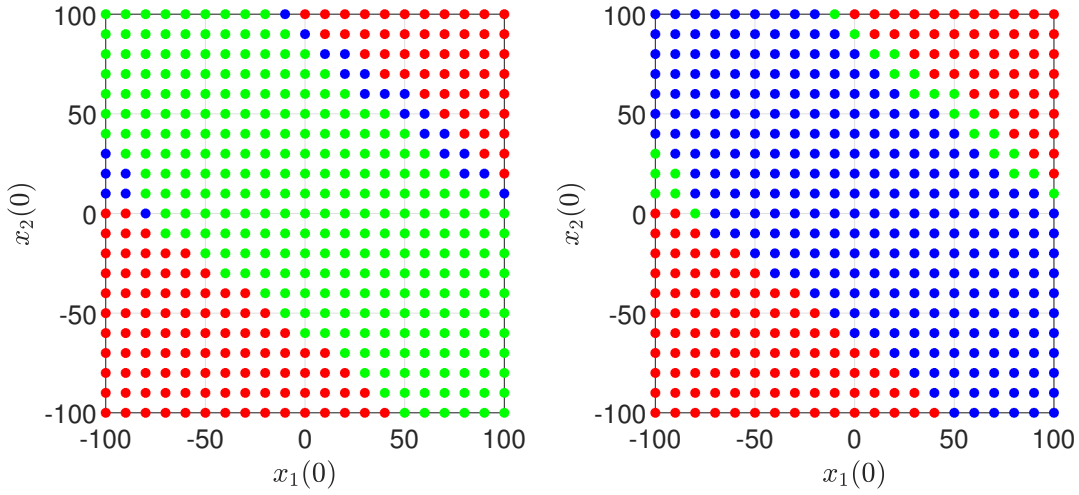
(c) $N_1 = 1, \alpha = -1$.

(d) $N_1 = -1, \alpha = -1$.

Figure 2.4: Example 2.1. Parameter pre-estimate $\nu(k)$ and the estimate $\hat{\mu}(k)$ for various choices of the filter coefficient $N_1$ and the input $u$ determined by the parameter $\alpha$.

Next, to investigate the effect of the initial conditions of (2.8), (2.9) on the performance of RCPE, $\mu$ is estimated with $x_1(0)$ and $x_2(0)$ varied from $-100$ to 100. The input $u(k)$ is given by (2.42) in all cases. Each point in Figure 2.5 (a),(b) corresponds to an initial condition of (2.8), (2.9), where green indicates that $\nu(k)$ converges to $\mu$, blue indicates that $\nu(k)$ converges to $-\mu$, and red indicates that $\nu(k)$ diverges. Note that all cases are obtained by running RCPE under the same values of $\lambda$ and $R_\theta$; however, the set of initial conditions $x(0)$ for which $\nu(k)$ converges can be expanded by varying these parameters. Figures 2.4 and 2.5 suggest that, for the given input

$u$ and filter coefficient $N_1$, the feasible region is either $(-\infty, 0]$ or $[0, \infty)$, and thus cannot be determined a priori. Consequently, (2.16) ensures that there exists $s \in \mathcal{S}_{\mathcal{O}_p}$ in the feasible region such that, $\nu(k)$ converges to $s$ , and thus $\hat{\mu}(k)$ converges to $\mu$. $\diamond$




(a) $N_1 = -1$ (b) $N_1 = 1$

Figure 2.5: Example 2.1. Convergence of the parameter estimate for a grid of initial conditions of (2.8), (2.9).

**Example 2.2.** *Nonaffinely parameterized linear dynamics with one unknown parameter in the dynamics matrix. This example investigates the effect of noise, and compares the performance of RCPE with UKF.* Consider the linear system (2.8), (2.9) , where

$$A(\mu) = \begin{bmatrix} \sin \mu & \cos \dfrac{\mu}{3} \\ \dfrac{e^{-\mu}}{3} & \dfrac{0.5}{1.1 + \mu^2} \end{bmatrix}, \tag{2.43}$$

$\mu = 0.3$, and $B$ and $C$ are given by (2.41). The initial state is $x(0) = [10 \quad 10]^{\mathrm{T}}$, $u(k)$ is given by (2.42), $N_1 = 1$, $\lambda = 1$, and $R_\theta = 10^6$. Furthermore, $\mathcal{O}_1 = 1$, and thus $\mathcal{S}_1 = \{\pm\mu\}$. Figure 2.6 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $-\mu$.

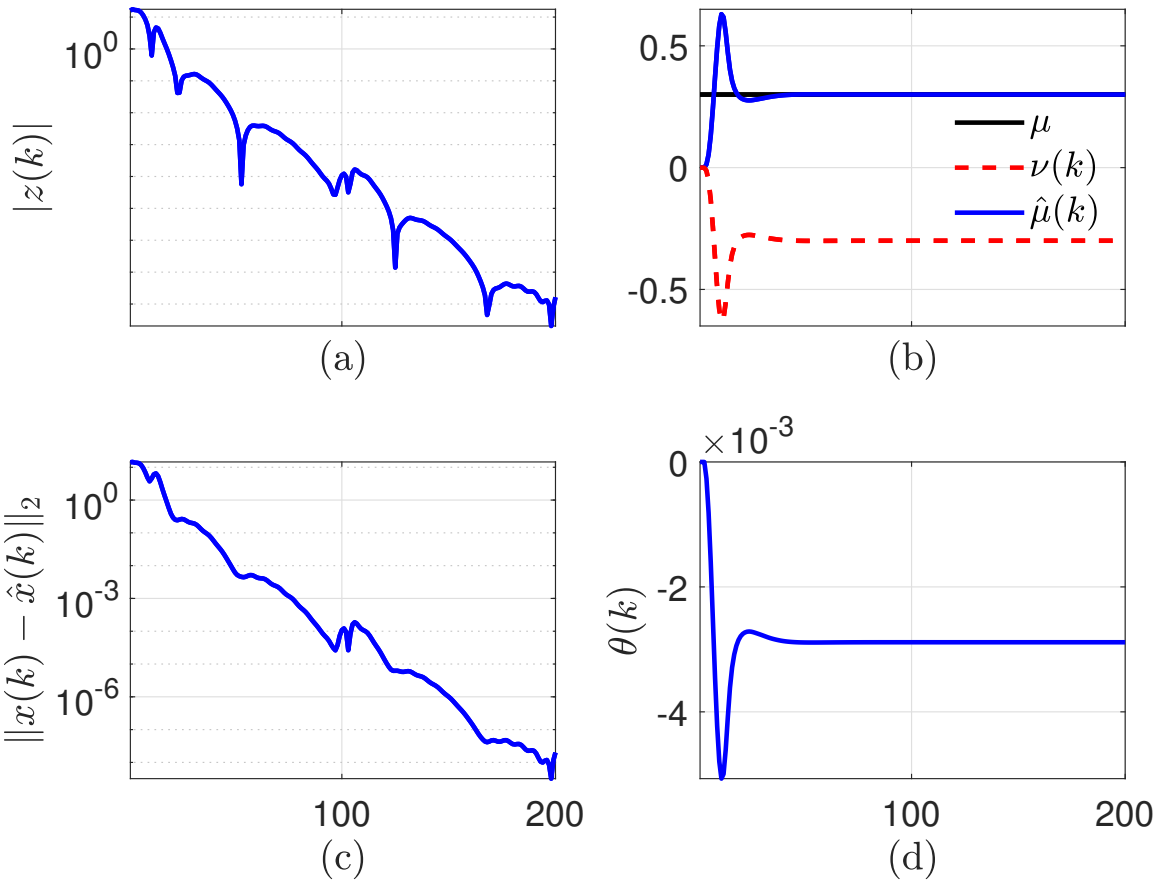



Figure 2.6: Example 2.2. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficient.

Next, $\mu$ is estimated using UKF with $\alpha = 1.01$, $P(0) = \text{diag}(10^3, 10^3, 1), Q = 10^{-10} I_2$, and $R = 10^{-10}$, where $\alpha$ affects the distribution of sigma points, and $P, Q, R$ are the augmented state, process and measurement covariance matrices. Figure 2.7 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. Note that the UKF gain $K(k)$ does not converge.

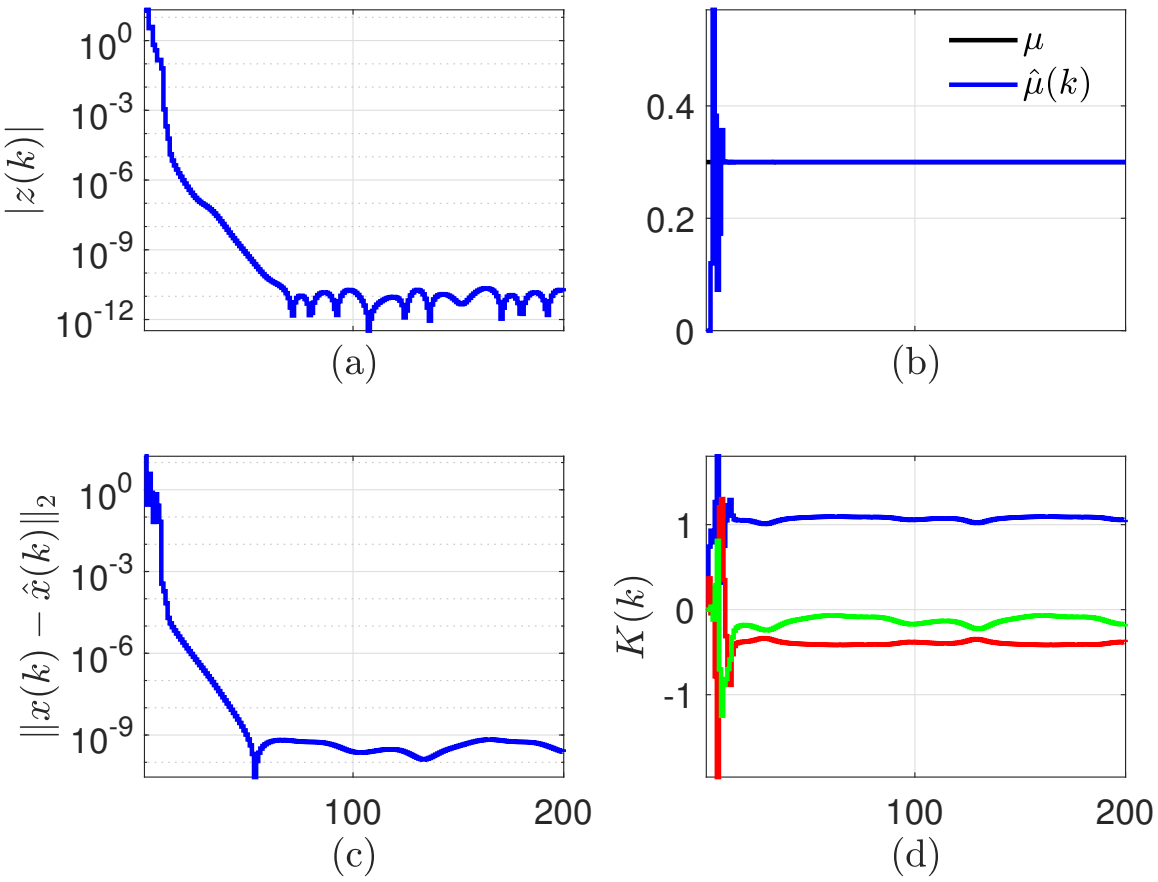
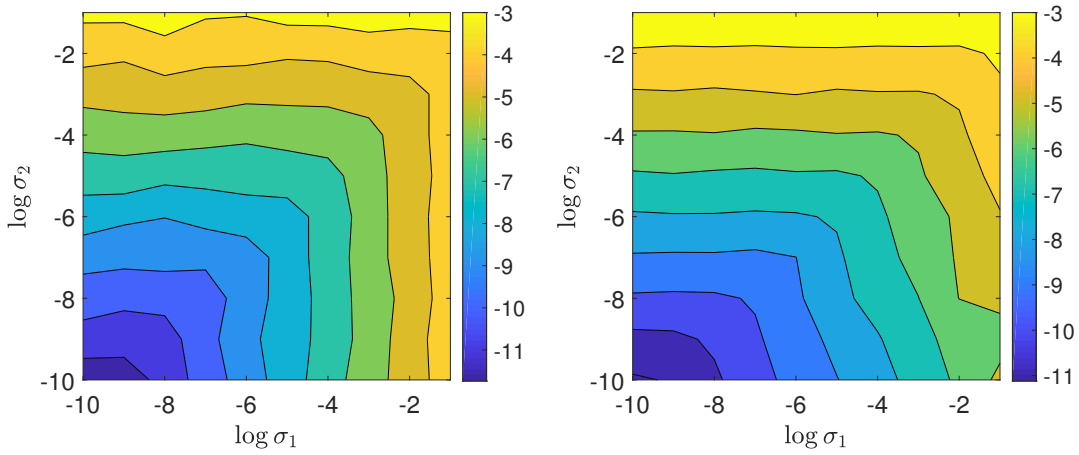



Figure 2.7: Example 2.2. UKF-based parameter estimation. (a) output error, (b) true parameter and parameter estimate, (c) state-estimate error, (d) components of the UKF based estimator gain.

Next, to compare the accuracy of RCPE and UKF in the presence of noise, $\mu$ is estimated with process noise $w_1 \sim \mathcal{N}(0, \sigma_1^2 I_2)$ and measurement noise $w_2 \sim \mathcal{N}(0, \sigma_2^2)$. For RCPE, $N_1 = 1$, $\lambda = 1$, and $R_\theta = 10^6$; for UKF, $\alpha = 1.01$, $P(0) = \mathrm{diag}(10^3, 10^3, 1), Q = \sigma_1 I_2$, and $R = \sigma_2$. For a range of values of $\sigma_1$ and $\sigma_2$, Figure 2.8 shows

$$\varepsilon_\mu \triangleq \frac{1}{100}\sqrt{\sum_{i=901}^{1000}(\hat{\mu}(i)-\mu)^2}. \tag{2.44}$$

Note that, unlike UKF, RCPE uses no knowledge of the noise statistics $Q$ and $R$ to compute $\hat{\mu}$. $\diamond$




(a) RCPE (b) UKF

Figure 2.8: Example 2.2. Effect of noise on the estimation accuracy of (a) RCPE and (b) UKF; the color scale denotes values of $\log \varepsilon_\mu$ as a function of $\sigma_1$ and $\sigma_2$.

## 2.6 Example with $l_\mu = l_y = 2$

In this section, RCPE is used to estimate two unknown parameters in linear systems that are affinely and nonaffinely parameterized with two measurements.

**Example 2.3.** *Nonaffinely parameterized linear dynamics with two measurements and two unknown parameters. This example shows how RCPE can be implemented with a sparse $R(k)$.* Consider the linear system (2.8), (2.9), where

$$A(\mu) = \begin{bmatrix} \sin \mu_1 & \cos \frac{\mu_1}{3} \\ \dfrac{e^{-\mu_1}}{3} & \dfrac{0.5}{1.1 + \mu_1^2} \end{bmatrix}, \; B(\mu) = \begin{bmatrix} \log(1 + \mu_2^2) \\ 1 + \sin \mu_2 \end{bmatrix}, \; C(\mu) = \begin{bmatrix} \mu_2 & 4\mu_1^2 \\ \sin \mu_1 & 2\mu_2 \end{bmatrix}, \tag{2.45}$$

and $\mu = [\mu_1 \quad \mu_2]^{\mathrm{T}} = [0.4 \quad 0.2]^{\mathrm{T}}$. The initial state is $x(0) = [10 \quad 10]^{\mathrm{T}}$, $u(k)$ is given by (2.42), $N_1 = I_2$, $\lambda = 1$, and $R_\theta = 10^8 I_4$. Furthermore, p = 12, and thus $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} = \{[\pm\mu_1 \ \pm\mu_2]^{\mathrm{T}}\}$. Note that $R(k) \in \mathbb{R}^{2\times 2}$, and thus the estimates of $\mu_1$ and $\mu_2$ are determined by both $z_1$ and $z_2$. Figure 2.9 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $-\mu$.

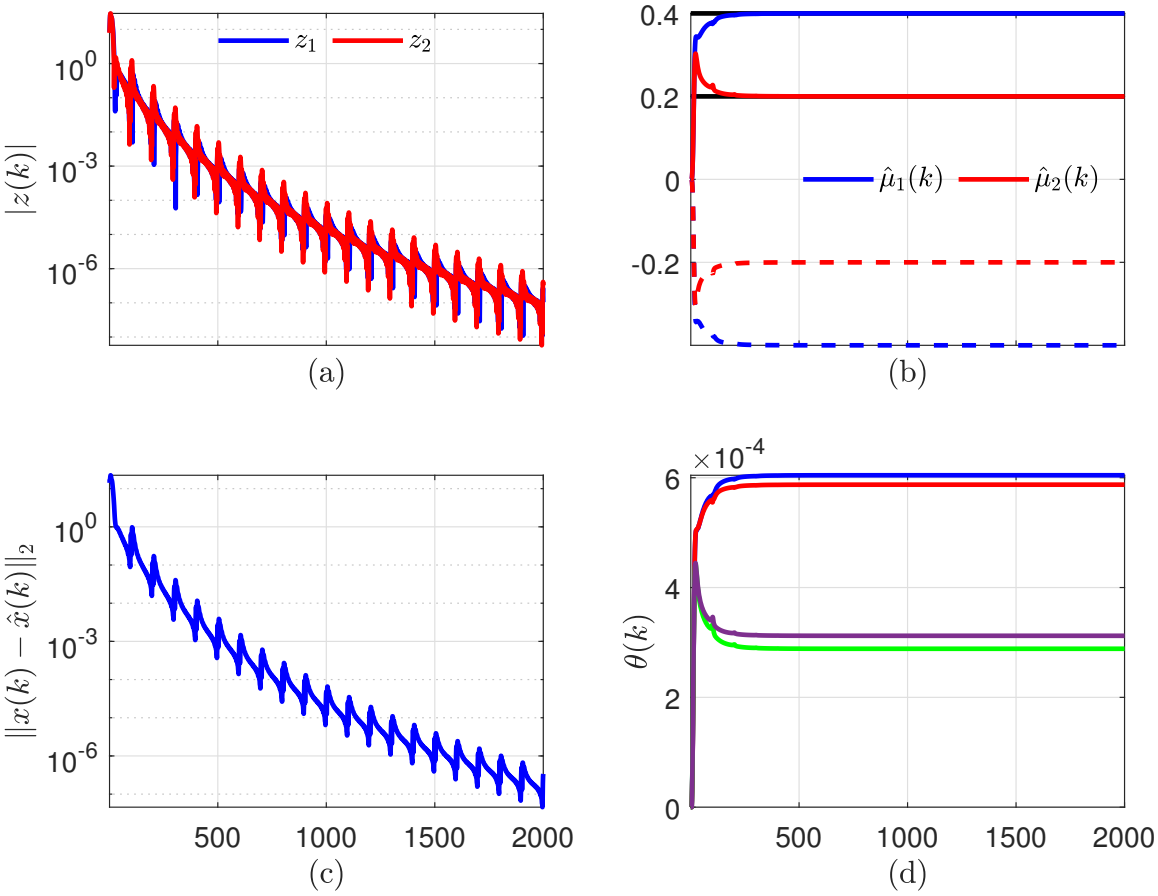



Figure 2.9: Example 2.3. (a) output error, (b) parameter estimates, (c) state-estimation error, (d) parameter estimator coefficients.

Next, the adaptive integrator (2.14) is constrained such that

$$\nu(k) = \begin{bmatrix} R_{11}(k) & 0 \\ 0 & R_{22}(k) \end{bmatrix} \phi(k). \tag{2.46}$$

Since $R(k)$ is sparse, it follows that $\mu_1$ is determined by $z_1$ only, and $\mu_2$ is determined by $z_2$ only. Furthermore, p = 12, $N_1 = I_2$, $\lambda = 1$, and $R_\theta = 10^8$. Figures 2.10 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. ◇

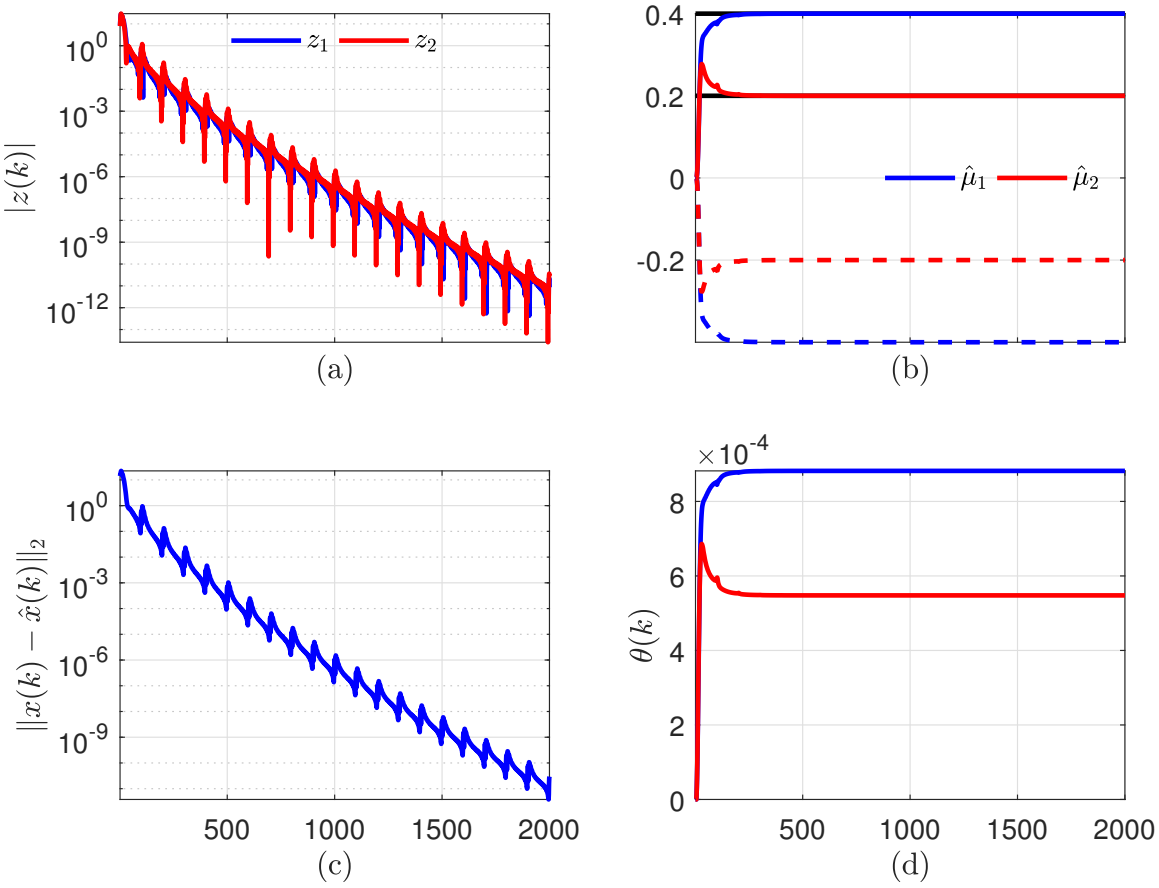



Figure 2.10: Example 2.3. Sparse integrator. (a) output error, (b) parameter estimates, (c) state-estimation error, (d) parameter estimator coefficients.

## 2.7 Examples with $l_\mu = 2$ and $l_y = 1$

In this section, RCPE is used to estimate two unknown parameters in affinely and nonaffinely parameterized systems with one measurement. These examples show that the feasible region is determined by the choice and ordering of the filter coefficients.

**Example 2.4.** *Affinely parameterized linear dynamics with two unknown parameters in the dynamics matrix. This example investigates the effect of $N_1, N_2$, and $\mathcal{O}_{\mathrm{p}}$ on the feasible region.* Consider the linear system (2.8), (2.9), where

$$A(\mu) = \begin{bmatrix} \mu_1 & \mu_2 \\ 0.1 & 0.6 \end{bmatrix}, \tag{2.47}$$

the input and output matrices are given by (2.41), and $\mu = [\mu_1 \ \ \mu_2]^{\mathrm{T}} = [0.3 \ \ 0.2]^{\mathrm{T}}$. The initial state is $x(0) = [10 \ \ 10]^{\mathrm{T}}$, $u(k)$ is given by (2.42), $N_1 = e_1$, $N_2 = e_2$, $\lambda = 0.999$, and $R_\theta = 10^6 I_2$. Furthermore, $\mathrm{p} = 12$, and thus $\mathcal{S}_{\mathcal{O}_{12}} = \{[\pm\mu_1 \ \ \pm\mu_2]^{\mathrm{T}}\}$. Figures 2.11 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $-\mu$.

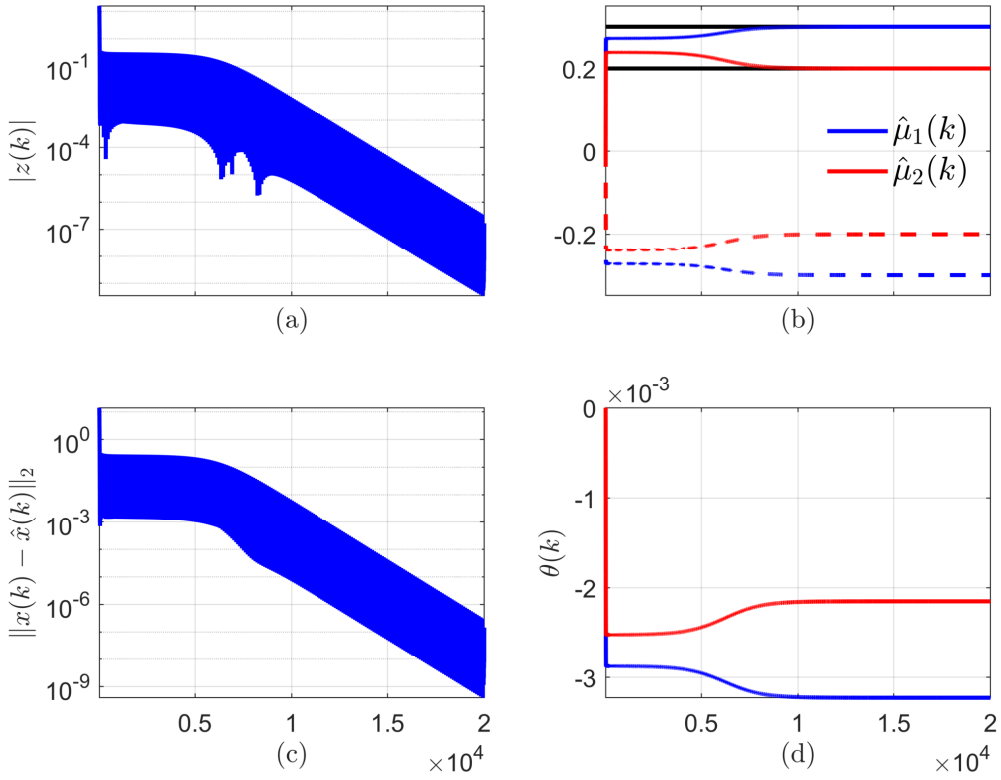



Figure 2.11: Example 2.4. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficients.

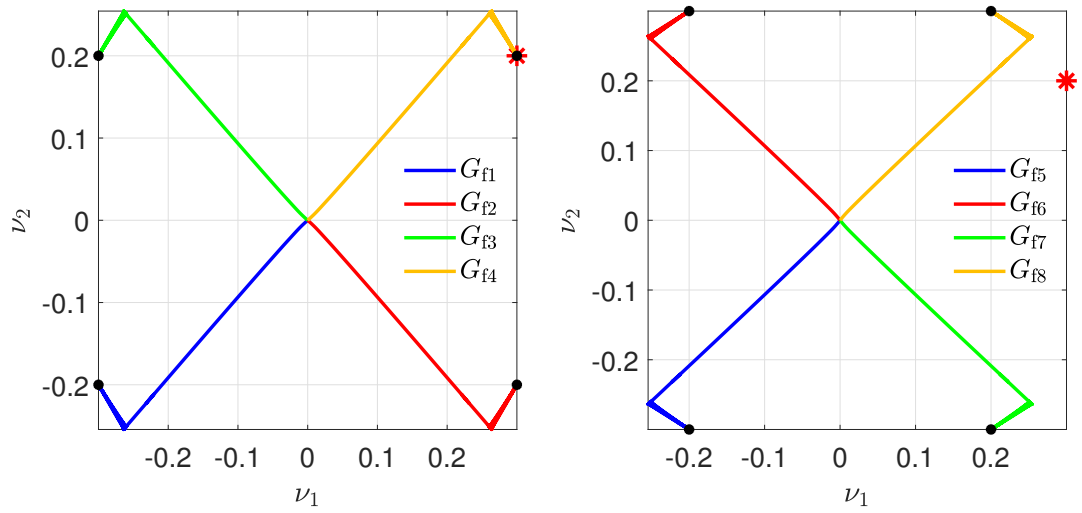
Next, the effect of the choice of $\mathcal{O}_\mathrm{p}$ and $N_1$ and $N_2$ is investigated. For $l_\mu = 2$, there are two choices of $\mathcal{O}_\mathrm{p}$ and two ways to order the filter coefficients $e_1$ and $e_2$. Further, for each ordering, there are four ways to allocate signs. Table 2.2 shows all such filter choices. Figure 2.12(a) shows $\nu(k)$ for $\mathrm{p} = 12$, and Figure 2.12(b) shows $\nu(k)$ for $\mathrm{p} = 21$, where the corresponding filter coefficients are given in Table 2.2. Note that, for a fixed ordering of the filter coefficients, there is exactly one permutation matrix $\mathcal{O}_\mathrm{p}$ such that the parameter pre-estimate $\nu$ converges to an element of $\mathcal{S}_{\mathcal{O}_\mathrm{p}}$. Conversely, for a fixed permutation matrix $\mathcal{O}_\mathrm{p}$, there is exactly one ordering of the

filter coefficients such that the parameter pre-estimate $\nu$ converges to an element of $\mathcal{S}_{\mathcal{O}_\mathrm{p}}$. $\diamond$

| | $N_1$ | $N_2$ |
|---|---|---|
| $G_{\mathrm{f}1}$ | $e_1$ | $e_2$ |
| $G_{\mathrm{f}2}$ | $-e_1$ | $e_2$ |
| $G_{\mathrm{f}3}$ | $e_1$ | $-e_2$ |
| $G_{\mathrm{f}4}$ | $-e_1$ | $-e_2$ |

| | $N_1$ | $N_2$ |
|---|---|---|
| $G_{\mathrm{f}5}$ | $e_2$ | $e_1$ |
| $G_{\mathrm{f}6}$ | $-e_2$ | $e_1$ |
| $G_{\mathrm{f}7}$ | $e_2$ | $-e_1$ |
| $G_{\mathrm{f}8}$ | $-e_2$ | $-e_1$ |

Table 2.2: Filter coefficients for Example 2.4.




(a) $\mathrm{p} = 12$. (b) $\mathrm{p} = 21$.

Figure 2.12: Example 2.4. Parameter pre-estimate $\nu(k)$ for various choices of $G_\mathrm{f}$ given in Table 2.2. The true parameter $\mu$ is shown in red.

**Example 2.5.** *Nonaffinely parameterized nonlinear dynamics with two unknown parameters. This example investigates the effect of $N_1, N_2$, and $\mathcal{O}_\mathrm{p}$ on the feasible region.*

Consider the (3,3) type nonlinear system [37, p. 183]

$$x(k+1) = \begin{bmatrix} x_2(k) \\ \dfrac{1 + 0.8x_2(k) + x_1(k)}{1 + \mu_1 x_2(k) + \mu_2 x_1(k)} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k), \tag{2.48}$$

$$y(k) = x_1(k), \tag{2.49}$$

where $\mu = [\mu_1 \;\; \mu_2]^{\mathrm{T}} = [0.6 \;\; 1.1]^{\mathrm{T}}$. The initial state is $x(0) = [10 \;\; 10]^{\mathrm{T}}$, $u(k)$ is given by

$$u(k) = 2 + \sum_{j=1}^{5} \frac{1}{j} \sin\left(\frac{2\pi j}{100}k + j^2\right), \tag{2.50}$$

$N_1 = e_2$, $N_2 = e_1$, $\lambda = 0.999$, and $R_\theta = 10^{12} I_2$. Furthermore, p = 21, and thus $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} = \{[\pm\mu_2 \; \pm\mu_1]^{\mathrm{T}}\}$. Figure 2.13 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $\mathcal{O}_{\mathrm{p}}^{-1}\mu$. Analogous results shown in Figure 2.12 are obtained for other choices of the filter coefficients and permutation matrix $\mathcal{O}_{\mathrm{p}}$.

Figure 2.14(a) shows $\nu(k)$ for p = 21, and Figure 2.14(b) shows $\nu(k)$ for p = 12, where the corresponding filter coefficients are given in Table 2.2. Note that, for a fixed ordering of the filter coefficients, there is exactly one permutation matrix $\mathcal{O}_{\mathrm{p}}$ such that the parameter pre-estimate $\nu$ converges to an element of $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}}$. Conversely, for a fixed permutation matrix $\mathcal{O}_{\mathrm{p}}$, there is exactly one ordering of the filter coefficients such that the parameter pre-estimate $\nu$ converges to an element of $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}}$. $\diamond$

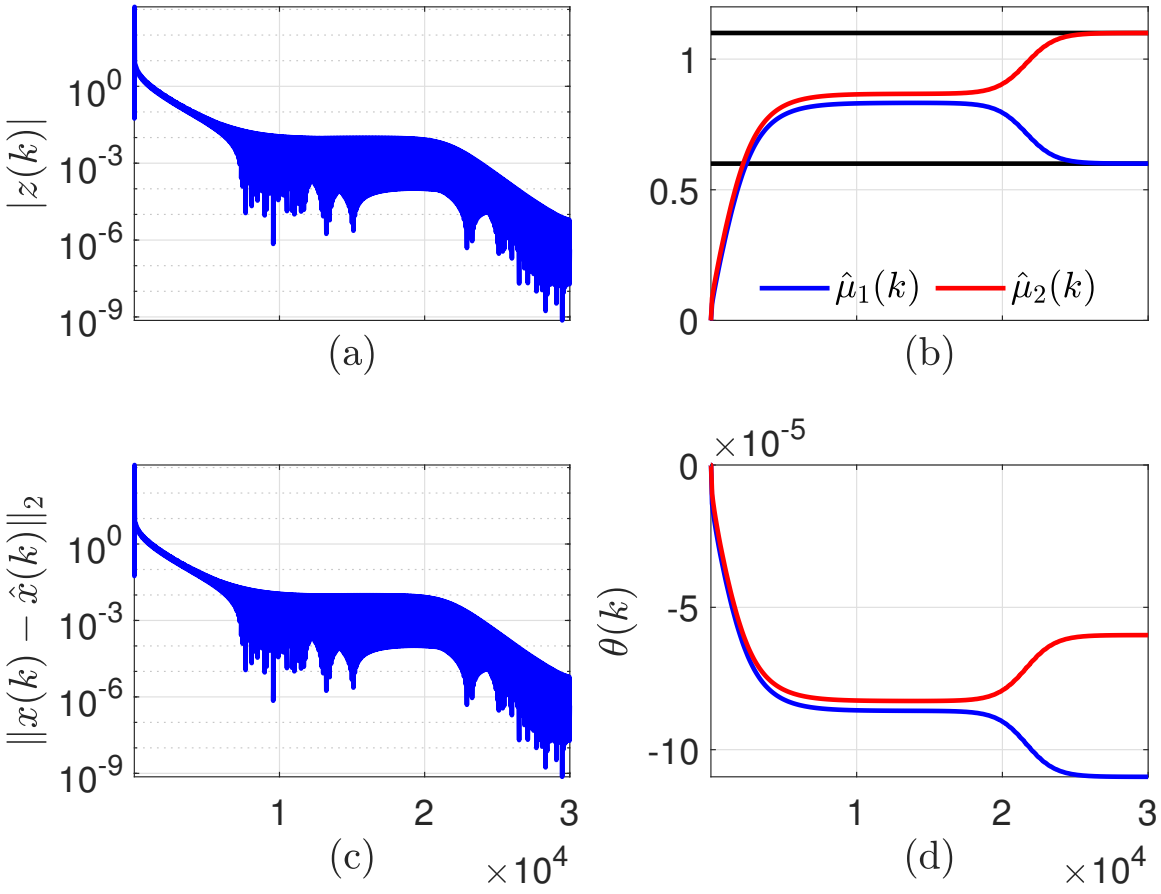
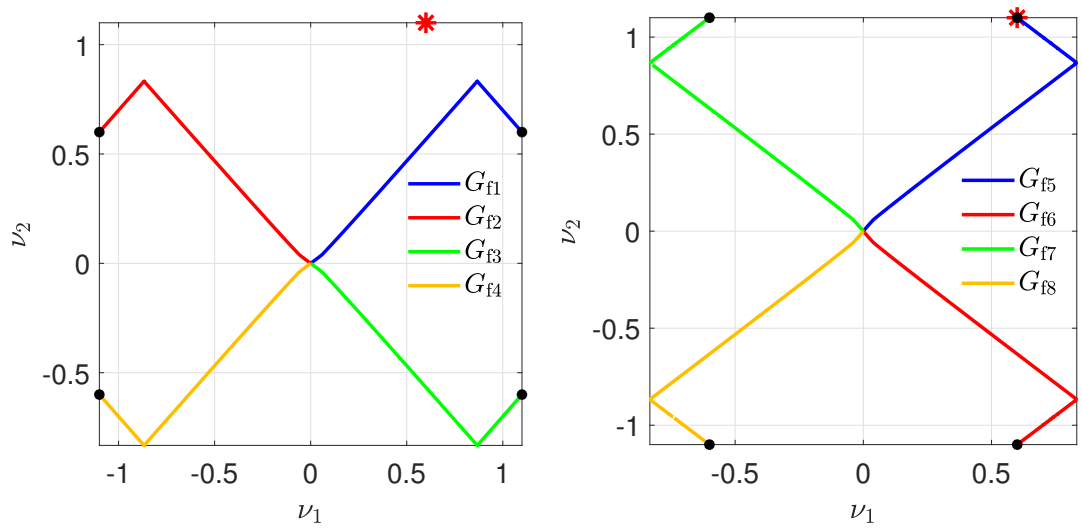



Figure 2.13: Example 2.5. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficients.




(a) p = 21.

(b) p = 12.

Figure 2.14: Example 2.5. Parameter pre-estimate $\nu(k)$ for various choices of $G_{\mathrm{f}}$ given in Table 2.2. The true parameter $\mu$ is shown in red.

## 2.8 Examples with $l_\mu = 3$ and $l_y = 1$

In this section, RCPE is used to estimate three unknown parameters in affinely and nonaffinely parameterized systems with one measurement.

**Example 2.6.** *Affinely parameterized linear dynamics with three unknown parameters in the dynamics matrix. This example investigates the effect of* $N_1, N_2, N_3$, *and* $\mathcal{O}_{\mathrm{p}}$ *on the feasible region.* Consider the linear system (2.8), (2.9), where

$$A(\mu) = \begin{bmatrix} \mu_1 & \mu_2 \\ 0.1 & \mu_3 \end{bmatrix}, \tag{2.51}$$

the input and output matrices are given by (2.41), and $\mu = [\mu_1 \quad \mu_2 \quad \mu_3]^{\mathrm{T}} = [0.3 \quad 0.2 \quad 0.6]^{\mathrm{T}}$. The initial state is $x(0) = [10 \quad 10 \quad 10]^{\mathrm{T}}$, $u(k)$ is given by (2.42), $N_1 = e_3$, $N_2 = e_2$, $N_3 = e_1$, $\lambda = 0.999$, and $R_\theta = 10^6 I_2$. Furthermore, p = 123, and thus $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} = \{[\pm\mu_1 \quad \pm\mu_2 \quad \pm\mu_3]^{\mathrm{T}}\}$. Figures 2.15 shows the output error, true parame-

ter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. (b) shows that $\nu(k)$ converges to $-\mu$.

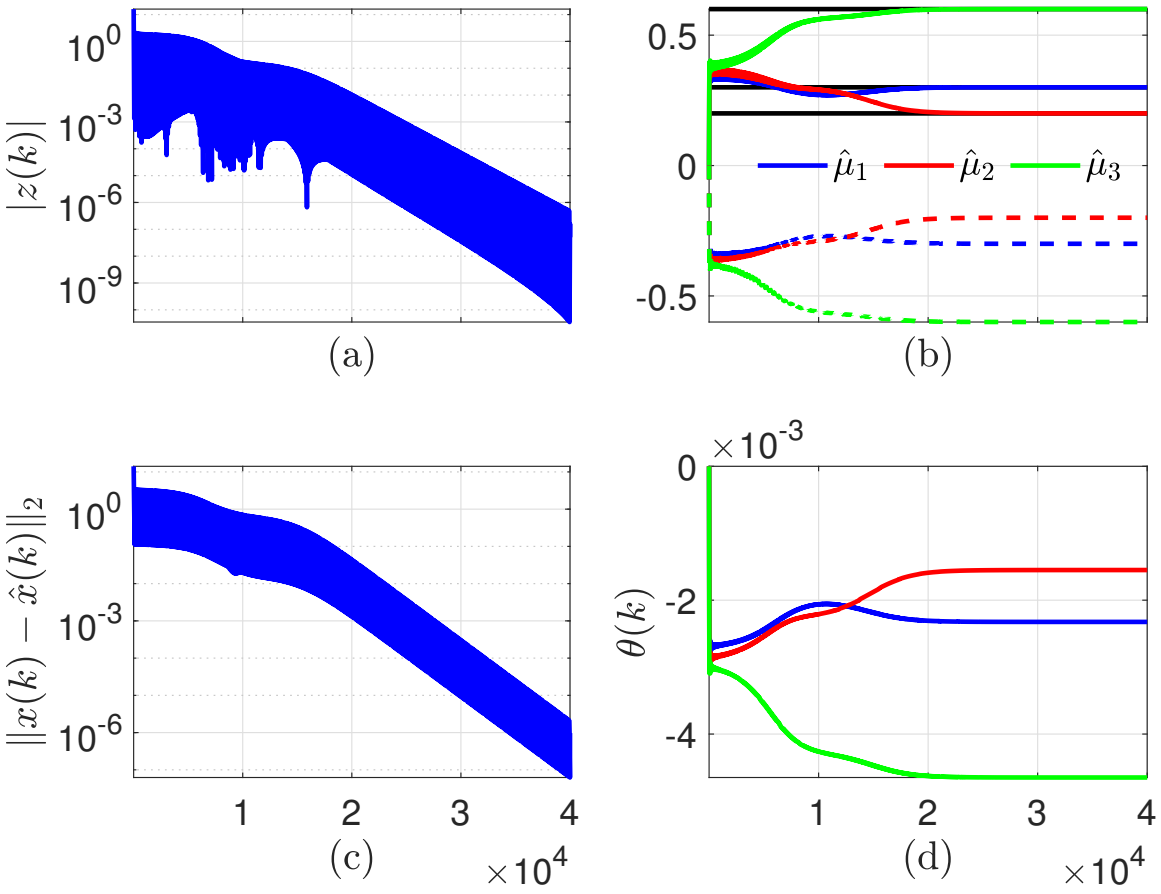
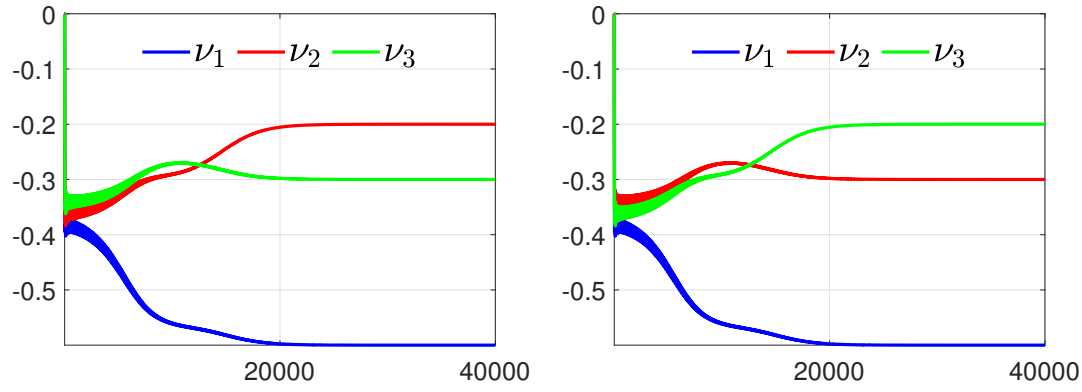



Figure 2.15: Example 2.6. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficients.

Next, the effect of the choice of $\mathcal{O}_\mathrm{p}$, $N_1, N_2$, and $N_3$ is investigated. For $l_\mu = 3$, there are six ways of ordering $e_1$, $e_2$, and $e_3$ and six choices of $\mathcal{O}_\mathrm{p}$. Figure 2.16 shows $\nu(k)$ for each ordering of the filter coefficients, where the corresponding $\mathcal{O}_\mathrm{p}$ is given in Table 2.3. Note that, in each case, $|\nu|$ converges to $\mathcal{O}_\mathrm{p}^{-1}\mu$.

| Case | $N_1$ | $N_2$ | $N_3$ | p |
|---|---|---|---|---|
| 1 | $e_1$ | $e_2$ | $e_3$ | 321 |
| 2 | $e_1$ | $e_3$ | $e_2$ | 231 |
| 3 | $e_2$ | $e_1$ | $e_3$ | 312 |
| 4 | $e_2$ | $e_3$ | $e_1$ | 132 |
| 5 | $e_3$ | $e_2$ | $e_1$ | 123 |
| 6 | $e_3$ | $e_1$ | $e_2$ | 213 |

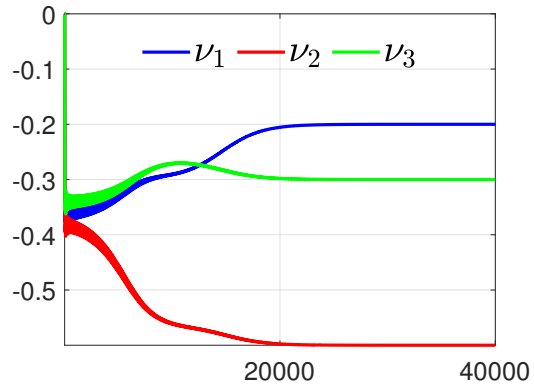
Table 2.3: Filter coefficients and $\mathcal{O}_{\mathrm{p}}$ for Example 2.6.



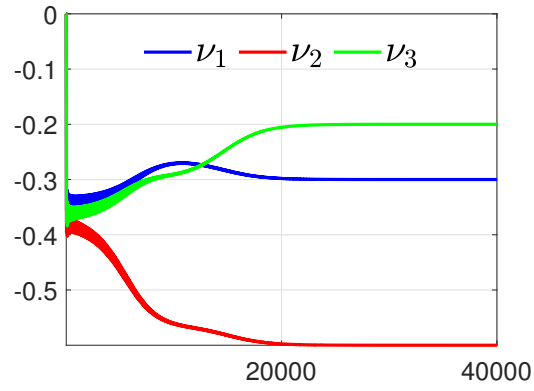

(a) Case 1



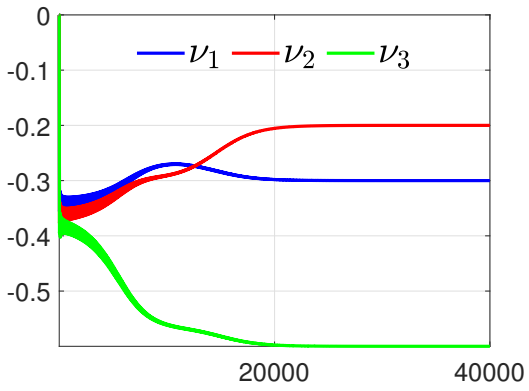

(b) Case 2



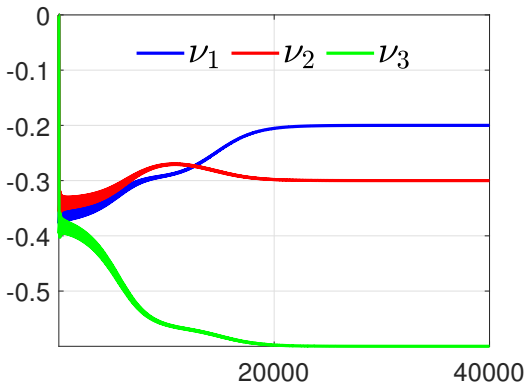

(c) Case 3




(d) Case 4




(e) Case 5




(f) Case 6

Figure 2.16: Example 2.6. Parameter pre-estimate $\nu(k)$ for various choices of $G_{\mathrm{f}}$ and $\mathcal{O}_{\mathrm{p}}$ given in Table 2.3.

Next, for each choice of the ordering of $e_1, e_2$, and $e_3$, the effect of the choice of the attitude of each filter coefficient is investigated. In particular, for $i \in \{1, 2, 3\}$,

each filter coefficient could be $\pm e_i$. Thus, the filter if given by

$$G_{\mathrm{f}i}(\mathbf{q}) = \frac{s_1 N_1}{\mathbf{q}} + \frac{s_2 N_2}{\mathbf{q}^2} + \frac{s_3 N_3}{\mathbf{q}^3}, \tag{2.52}$$

where $s_1, s_2$, and $s_3$ are given in 2.4. Figure 2.17 shows $\nu(k)$ for each choice of the filter coefficients with the corresponding $\mathcal{O}_{\mathrm{p}}$ given in Table 2.3. Note that, for a choice of filter coefficient ordering with the corresponding $\mathcal{O}_{\mathrm{p}}$, each choice of the attitude of the filter coefficients corresponds to exactly one element in $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}}$.

| Filter | $s_1$ | $s_2$ | $s_3$ |
|---|---|---|---|
| $G_{\mathrm{f1}}(\mathbf{q})$ | 1 | 1 | 1 |
| $G_{\mathrm{f2}}(\mathbf{q})$ | $-1$ | 1 | 1 |
| $G_{\mathrm{f3}}(\mathbf{q})$ | 1 | $-1$ | 1 |
| $G_{\mathrm{f4}}(\mathbf{q})$ | 1 | 1 | $-1$ |
| $G_{\mathrm{f5}}(\mathbf{q})$ | $-1$ | $-1$ | 1 |
| $G_{\mathrm{f6}}(\mathbf{q})$ | 1 | $-1$ | $-1$ |
| $G_{\mathrm{f7}}(\mathbf{q})$ | $-1$ | 1 | $-1$ |
| $G_{\mathrm{f8}}(\mathbf{q})$ | $-1$ | $-1$ | $-1$ |

Table 2.4: Filter coefficient sign for Example 2.6.

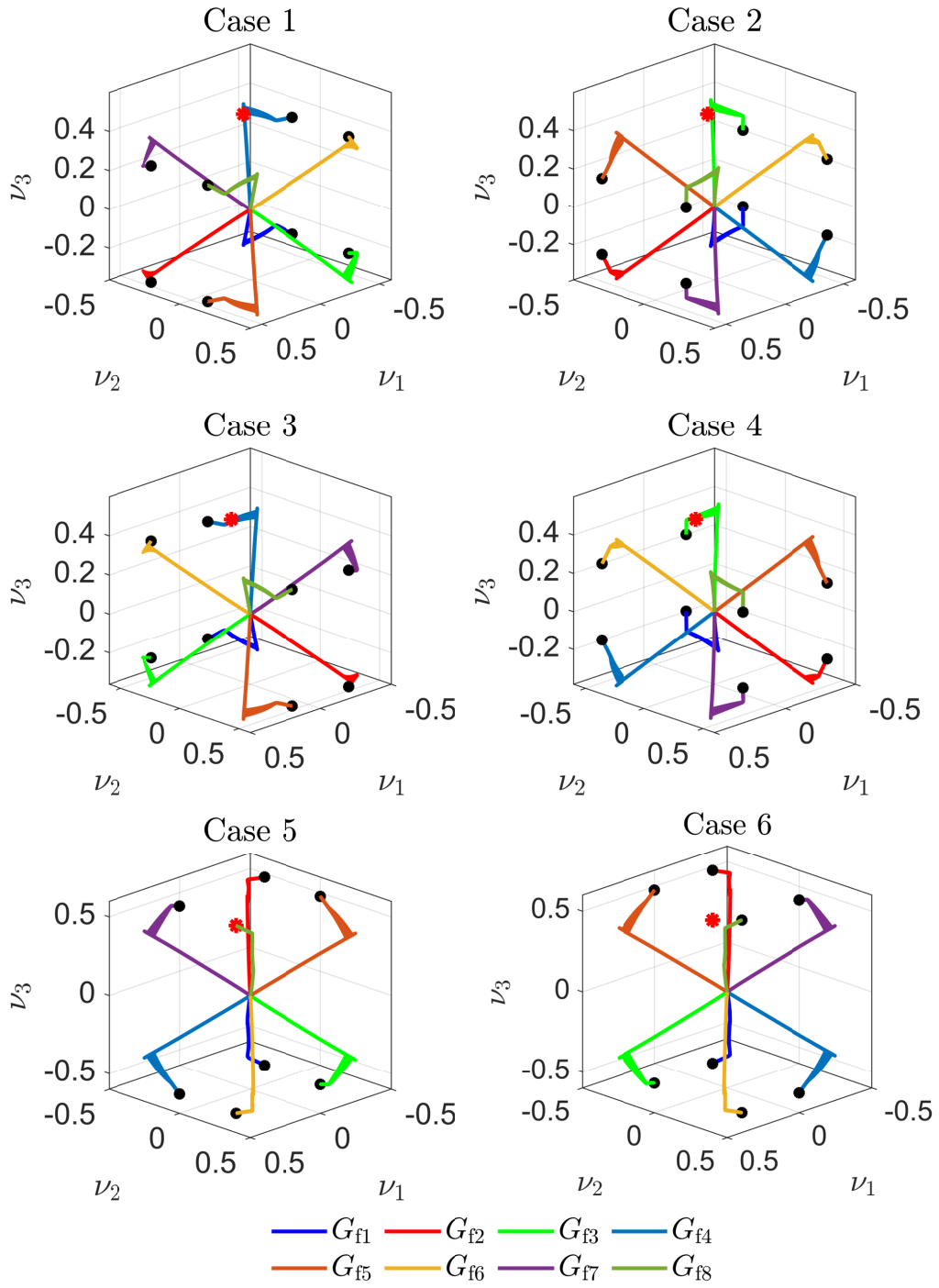



Figure 2.17: Example 2.6. Parameter pre-estimate $\nu(k)$ for various choices of $G_f$ and $\mathcal{O}_p$ given in (2.52), where $N_1, N_2$, and $N_3$ are given in Table 2.3 and $s_1, s_2$, and $s_3$ are given by Table 2.4.

$\diamond$

**Example 2.7.** *Affinely parameterized nonlinear dynamics with three unknown parameters. This example investigates the effect of $N_1, N_2, N_3$, and $\mathcal{O}_{\mathrm{p}}$ on the feasible region.* Consider the (3,3) type nonlinear system [37, p. 183]

$$x(k+1) = \begin{bmatrix} x_2(k) \\ \dfrac{\mu_1 + \mu_2 x_2(k) + \mu_3 x_1(k)}{1 + 0.6x_2(k) + 1.1x_1(k)} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k), \tag{2.53}$$

$$y(k) = x_1(k), \tag{2.54}$$

where $\mu = [\mu_1 \;\; \mu_2 \;\; \mu_3]^{\mathrm{T}} = [0.5 \;\; 0.8 \;\; 1.0]^{\mathrm{T}}$. The initial state is $x(0) = [10 \;\; 10 \;\; 10]^{\mathrm{T}}$, $u(k)$ is given by (2.42), $N_1 = e_1$, $N_2 = e_2$, $N_3 = e_3$, $\lambda = 0.9999$, and $R_\theta = 10^6 I_2$. Furthermore, p = 213, and thus $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} = \{[\pm\mu_2 \;\; \pm\mu_1 \;\; \pm\mu_3]^{\mathrm{T}}\}$. Figures 2.18 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient. Analogous results shown in Figure 2.16 are obtained for other choices of the filter coefficients and the permutation matrix $\mathcal{O}_{\mathrm{p}}$.

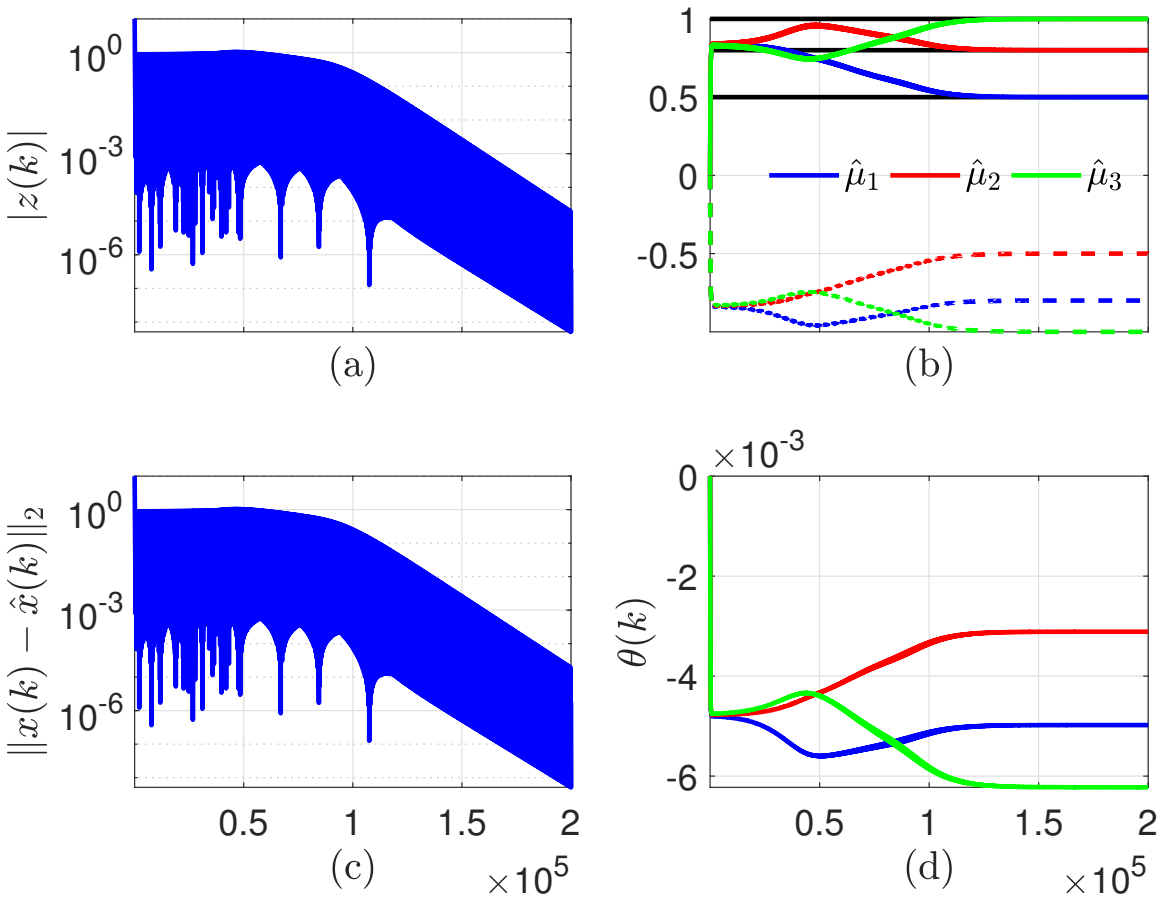



Figure 2.18: Example 2.7. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficients.

Figure 2.19 shows the output error for all six permutations. For clarity, a subset of the data is shown. Note that the output error diverges for all permutation but p = 213. Thus, diverging output error can be used to rule out the incorrect permutations.

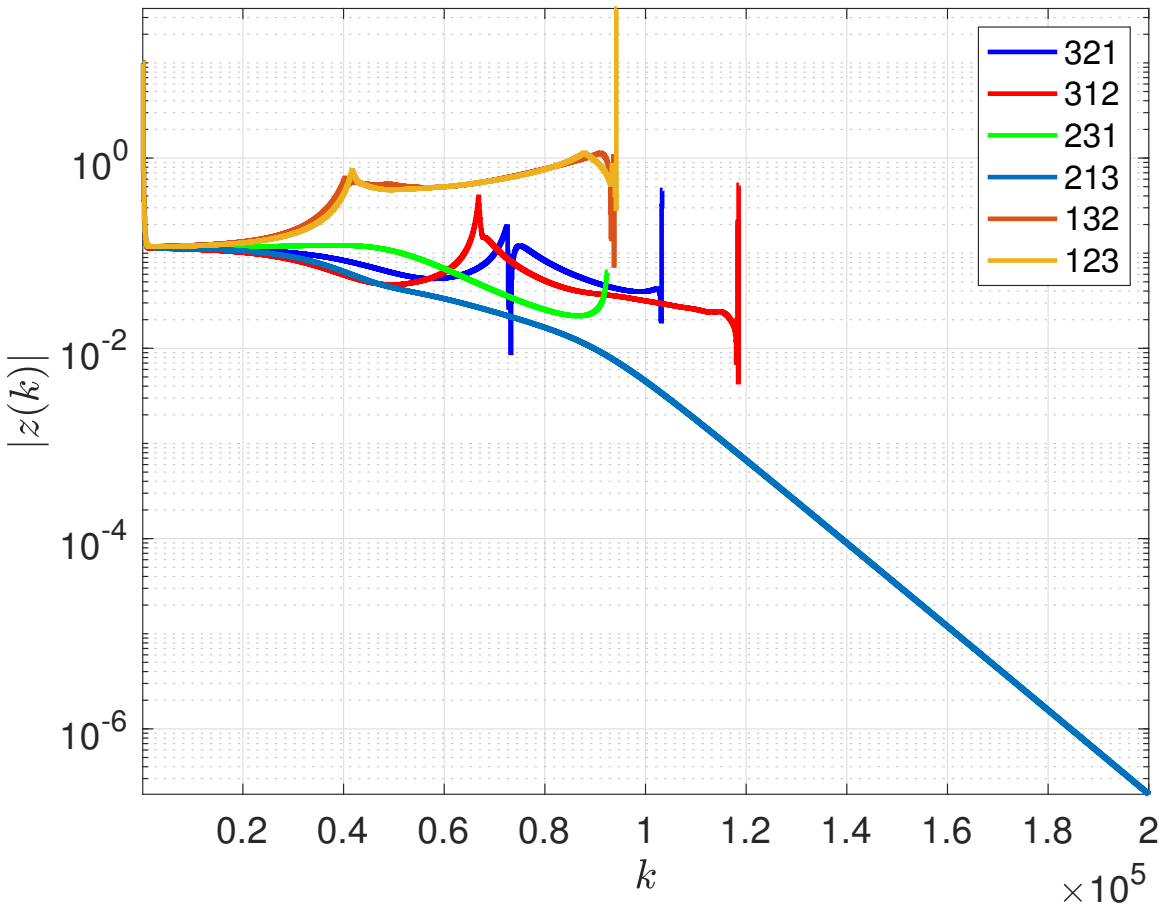



Figure 2.19: Example 2.7. Output error for all six permutations. [For clarity, only a subset of the data is plotted.] For five of the six permutations, the parameter error diverges. However, the correct permutation 213 yields convergence to the true parameters.

Next, for each choice of the ordering of $e_1, e_2$, and $e_3$, the effect of the choice of the attitude of each filter coefficient is investigated. Figure 2.20 shows $\nu(k)$ for each choice of the filter coefficients with the corresponding $\mathcal{O}_\mathrm{p}$ given in Table 2.5. Note that, for a choice of filter coefficient ordering with the corresponding $\mathcal{O}_\mathrm{p}$, each choice of the attitude of the filter coefficients corresponds to exactly one element in $\mathcal{S}_{\mathcal{O}_\mathrm{p}}$.

| Case | $N_1$ | $N_2$ | $N_3$ | p |
|---|---|---|---|---|
| 1 | $e_1$ | $e_2$ | $e_3$ | 213 |
| 2 | $e_1$ | $e_3$ | $e_2$ | 312 |
| 3 | $e_2$ | $e_1$ | $e_3$ | 123 |
| 4 | $e_2$ | $e_3$ | $e_1$ | 321 |
| 5 | $e_3$ | $e_2$ | $e_1$ | 231 |
| 6 | $e_3$ | $e_1$ | $e_2$ | 132 |

Table 2.5: Filter coefficients and $\mathcal{O}_{\mathrm{p}}$ for Example 2.7.

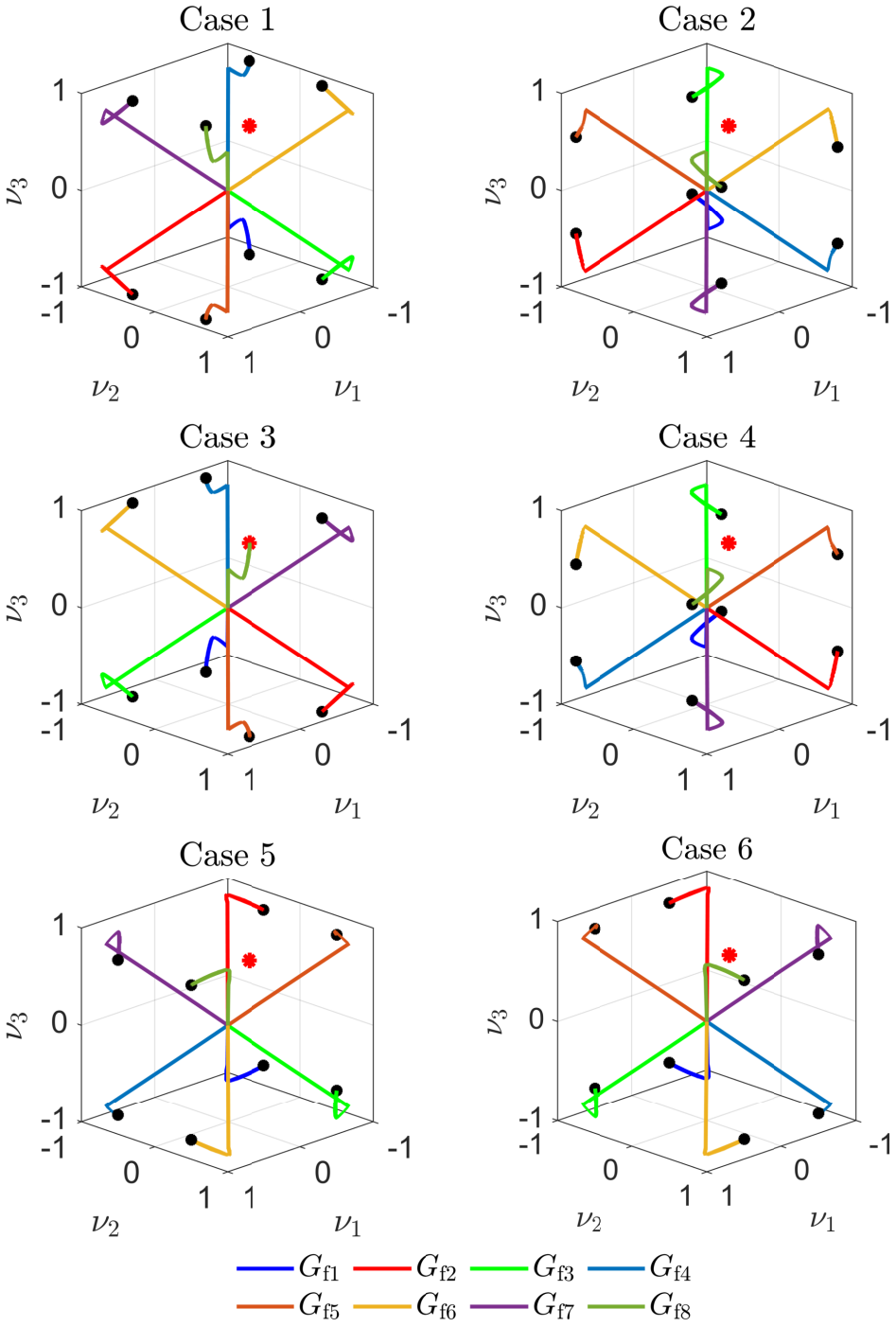



Figure 2.20: Example 2.7. Parameter pre-estimate $\nu(k)$ for various choices of $G_f$ and $\mathcal{O}_p$ given in (2.52), where $N_1, N_2$, and $N_3$ are given in Table 2.5 and $s_1, s_2$, and $s_3$ are given by Table 2.4.

$\diamond$

## 2.9 Parameter Estimation in the Generalized Burgers Equation

In this section, we consider the generalized one-dimensional viscous Burgers equation [38]

$$\frac{\partial u}{\partial t} + \mu_1 \frac{\partial}{\partial x}\frac{u^2}{2} = \frac{\partial}{\partial x}\left(\mu_2 \frac{\partial u}{\partial x}\right), \tag{2.55}$$

where $u(x,t)$ is a function of space and time with domain $[0,1]\times[0,\infty)$, $\mu_1 > 0$ is the convective constant and $\mu_2 > 0$ is the viscosity. Note that there is no external input to this system and $u$ is used to denote the solution of this partial differential equation. The initial condition is $u(x,0) = 0$ for all $x \in [0,1]$, and the boundary conditions are $u(0,t) = 0$ and $u(1,t) = \sin(5t) + 0.25\sin(10t)$ for all $t \geq 0$. The objective is to estimate the unknown parameter $\mu \stackrel{\triangle}{=} [\mu_1 \;\; \mu_1]^{\mathrm{T}}$ using measurements of $u$ at a single location.

The Burgers equation (2.55) is discretized using a forward Euler approximation for the time derivative, a second-order-accurate upwind method for the convective term, and a second-order-accurate central difference scheme for the viscous term. The spatial domain $[0,1]$ is discretized using $N$ equally spaced grid points; thus $\Delta x \stackrel{\triangle}{=} \frac{1}{N-1}$. The time step $\Delta t$ is chosen to satisfy the CFL condition, that is,

$$\Delta t < \frac{C_{\max}\Delta x}{|\max(u)|}, \tag{2.56}$$

where the Courant number $C_{\max}$ depends on the discretization scheme [39]. Finally, the discrete variable $u_j(k) \stackrel{\triangle}{=} u((j-1)\Delta x, k\Delta t)$ is defined on the grid points $j \in$

$\{1, \ldots, N\}$ for all time steps $k \geq 0$. Hence, at each grid point, $j \in \{3, \ldots, N-1\}$,

$$u_j(k+1) = u_j(k) - \mu_1 \frac{\Delta t}{2\Delta x}(1.5u_j(k)^2 - 2u_{j-1}(k)^2 + 0.5u_{j-2}(k)^2) + \mu_2 \frac{\Delta t}{\Delta x^2}(u_{j+1}(k) - 2u_j(k) + u_{j-1}(k)). \tag{2.57}$$

For all $k \geq 0$, the discretized boundary conditions are

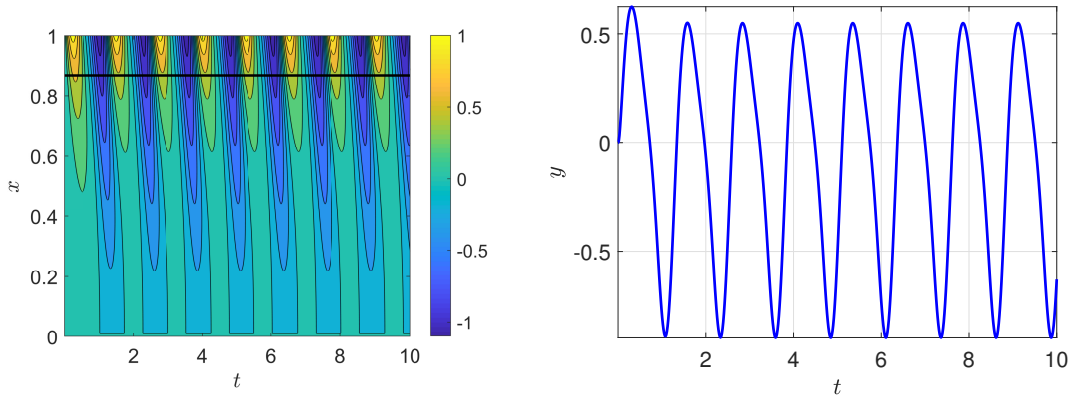
$$u_1(k) = u_2(k) = 0, \quad u_N(k) = \sin(5\Delta t k) + 0.25\sin(10\Delta t k), \tag{2.58}$$

and, for all $j \in \{3, \ldots, N-1\}$, the initial condition is

$$u_j(0) = 0. \tag{2.59}$$

In this example, $\mu_1 = 1.4, \mu_2 = 0.3$, $C_{\max} = 0.25$, $N = 100$, and $\Delta t = 10^{-4}$ s. Figure 2.21(a) shows the numerical solution of (2.57) with the boundary conditions (2.58) and initial conditions (2.59), where the solid black line shows the measurement location. Figure 2.21(b) shows the measurement $y(k) \stackrel{\triangle}{=} u_{87}(k) = u(0.87, k\Delta t)$.

In order to start the estimation model, nonzero values of $\hat{\mu}_1(0)$ and $\hat{\mu}_2(0)$ are needed. A simple way to ensure this is to replace $\mu$ by $\hat{\mu}(k) = \overline{\mu} + \mathcal{O}_{\mathrm{p}}\nu(k)$, where $\overline{\mu} = [\overline{\mu}_1 \;\; \overline{\mu}_1]^{\mathrm{T}} = [1 \;\; 0.01]^{\mathrm{T}}$, so that $\hat{\mu}(0) \neq 0$ . Furthermore, $N_1 = e_1$, $N_2 = e_2$, $\lambda = 0.9999$, and $R_\theta = 10^6 I_2$. Let $\mathrm{p} = 21$ so that $\mathcal{S}_{\mathcal{O}_{\mathrm{p}}} = \{[\pm(\mu_2 - \overline{\mu}_2) \;\; \pm(\mu_1 - \overline{\mu}_2)]^{\mathrm{T}}\}$. Figure 2.22 shows the output error, true parameter, parameter pre-estimate, parameter estimate, state-estimate error, and estimator coefficient.




(a) $u(x,t)$ (b) $y(t)$

Figure 2.21: Simulation of the generalized Burgers equation with the discretization (2.57).

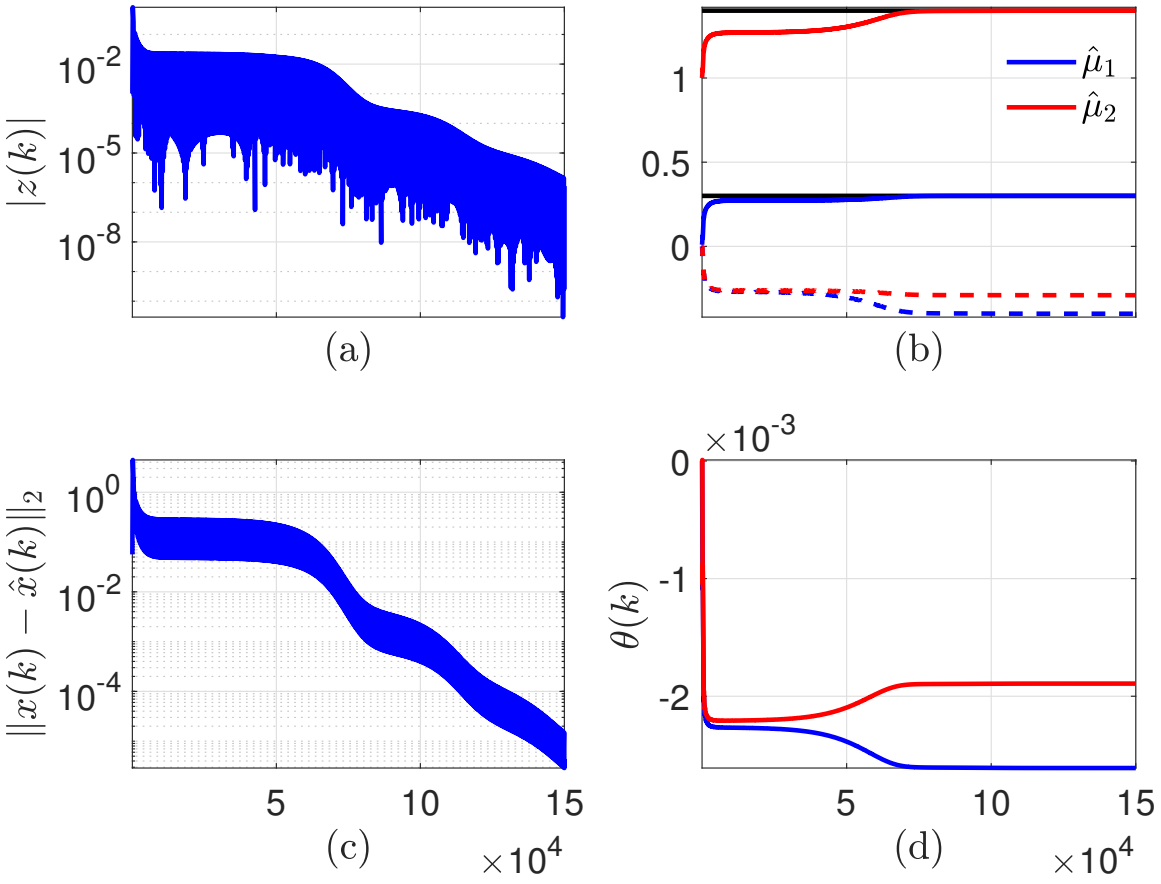



Figure 2.22: Generalized Burgers equation. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) state-estimate error, (d) parameter estimator coefficients.

## 2.10 Thermal Conductivity Estimation using Density Measurements

Finally, we consider the problem of estimating thermal conductivity using density measurements. In particular, we estimate the thermal conductivity coefficients $\kappa_{O_2}$ and $\kappa_O$ using simulated measurements of global maximum and minimum density altitudes between 100 km and 500 km. The temperature at one level of the atmosphere

depends on the vertical structure of the temperature through the equation

$$\frac{\delta T}{\delta t} \propto \frac{\delta}{\delta z}\lambda\frac{\delta T}{\delta z}, \tag{2.60}$$

where $T$ is the temperature, $t$ is time, $z$ is altitude, and $\lambda$ is approximated as

$$\lambda = \frac{\kappa_{O_2} N_{O_2} + \kappa_O N_O}{N_{O_2} + N_O} T^S, \tag{2.61}$$

where the parameters $\kappa_{O_2}$ and $\kappa_O$ depend on the density of atomic ($N_O$) and molecular ($N_{O_2}$) oxygen, and $S$ is independent of both.

To illustrate the application of RCPE to estimate thermal conductivity coefficients $\kappa_{O_2}$ and $\kappa_O$, we generate measurements from a simulation of Global Ionosphere Thermosphere Model (GITM), described in detail in Section 3.2, with thermal conductivity coefficients $\kappa_{O_2} = 0.00036$ and $\kappa_O = 0.00056$. In particular, we record minimum and maximum values of the density between 100 and 500 km.

In the estimation model, the thermal conductivity coefficients estimates are given by

$$\begin{bmatrix} \hat{\kappa}_{O_2}(k) \\ \hat{\kappa}_O(k) \end{bmatrix} = \begin{bmatrix} \overline{\kappa}_{O_2} \\ \overline{\kappa}_O \end{bmatrix} + \mathcal{O}_p|\nu(k)|, \tag{2.62}$$

where $\overline{\kappa}_{O_2} = 0.00031$ and $\overline{\kappa}_O = 0.00030$ are the initial guesses of thermal conductivity coefficients. In RCPE, we set $N_1 = I_2$, $\lambda = 0.999$, $R_\theta = 10^8 I_2$, and p = 12. Figure 2.23 shows the output error, true parameter, parameter pre-estimate, parameter estimate, estimated and measured densities, and the estimator coefficient.

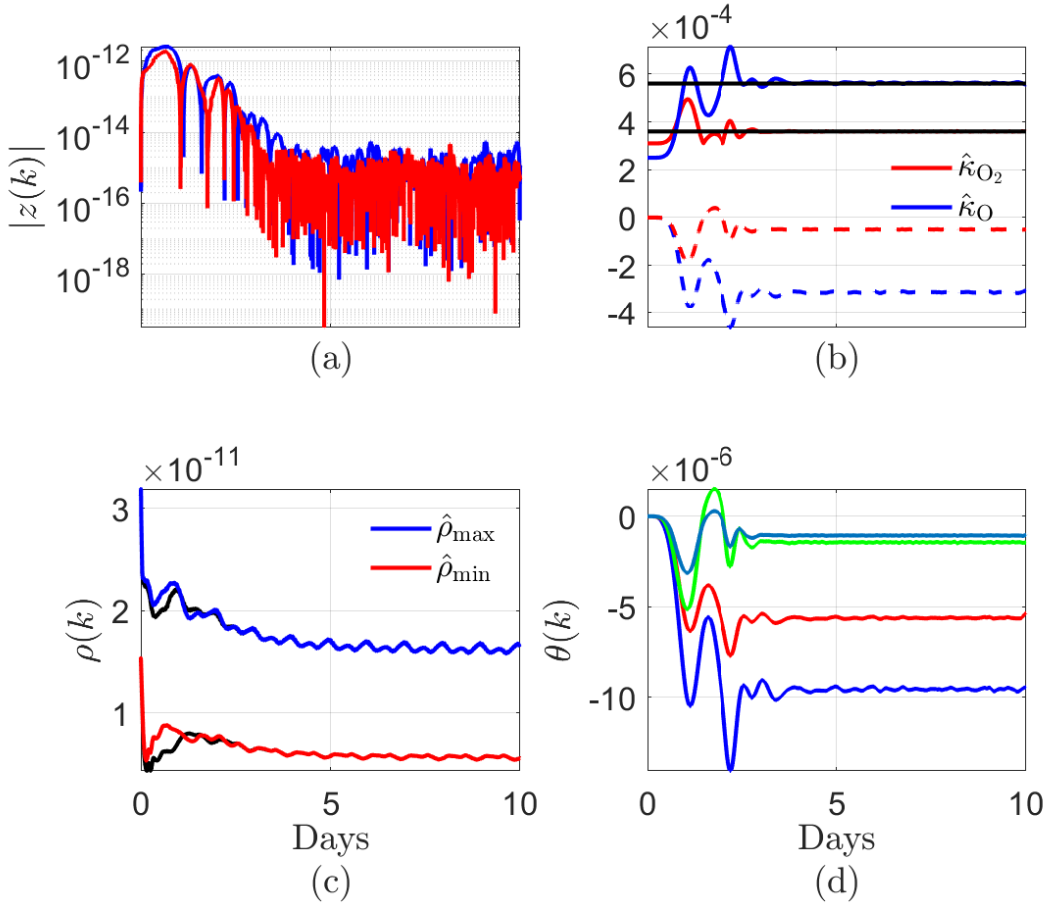



Figure 2.23: Estimation of thermal conductivity coefficients. (a) output error, (b) true parameter, parameter pre-estimate, and parameter estimate, (c) measured and estimated densities, (d) parameter estimator coefficients.

Next, thermal conductivity coefficients are estimated with various values of initial guesses with same RCPE tuning parameters. Figure 2.24 shows the thermal conductivity coefficients estimates for various values of initial guesses.

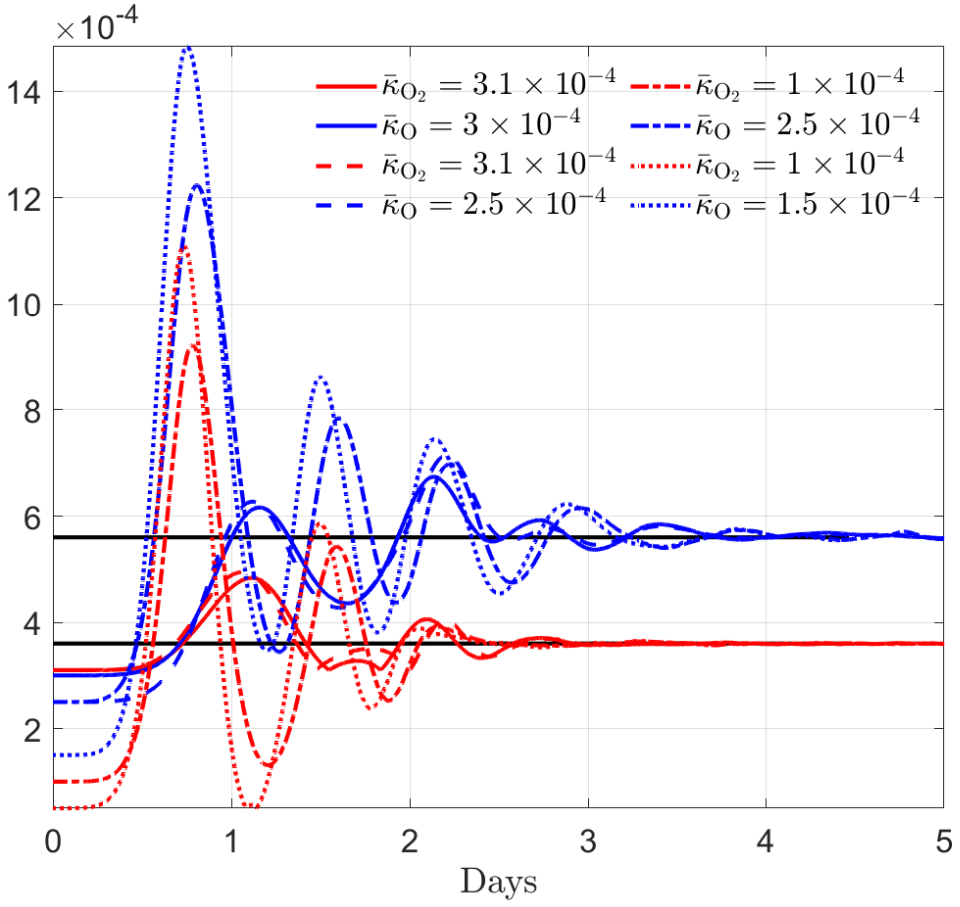



Figure 2.24: Effect of the initial parameter guess on the estimates of thermal conductivity coefficients.

Note that RCPE does not need to propagate an ensemble of GITM state to estimate the unknown parameters. In contrast, UKF would need to propagate a five member ensemble in GITM at each step to estimate the unknown the thermal conductivity coefficients.

## 2.11 Conclusions

This chapter presented RCPE, which is an iterative, data-driven technique for estimating unknown parameters in linear and nonlinear dynamical systems. Unlike the extended Kalman filter, RCPE is gradient-free; unlike the unscented Kalman filter, RCPE is ensemble-free; and unlike variational methods, RCPE is adjoint-free. It is shown that to estimate multiple unknown parameters, a permutation matrix is required to correctly associate each parameter estimate with the corresponding unknown parameter. The need to select the permutation matrix is the price paid for not requiring gradient information, an ensemble of models, or an adjoint model. The potential usefulness of RCPE was demonstrated by application to the Burgers equation and GITM. In the Burgers equation, UKF requires an ensemble of 207 models each with 101 states, requiring a total of 21,321 state updates. In contrast, RCPE required one model and testing of two permutations, which requires a total of 202 state updates (101 for each permutation). In GITM which has $\sim 10^6$ states, in the case where UKF is applied to update the unknown parameters only, a five member ensemble needs to be propagated at each step to estimate the unknown the thermal conductivity coefficients, increasing the computational cost by at least five times. In contrast, RCPE required a single propagation of GITM states and testing of two permutations.

In the next chapter, RCPE is extended to the case where the filter coefficients that define the retrospective cost are simultaneously optimized along with the parameter estimator coefficient. The extended RCPE is applied to the problem of estimating eddy diffusion coefficient in GITM using the measurements of total electron content.

# CHAPTER 3

# Estimation of the Eddy Diffusion Coefficient Using Total Electron Content Data

## 3.1 Introduction

Computational models of high-dimensional complex physical systems often have unknown parameters. These parameters might be unknown due to lack of a physical measurement equipment. For example, measurements of friction force to estimate friction coefficients might not be available. In addition, the unknown parameters might be used as a proxy for the cumulative effect of a complex phenomenon, and thus are representational nature, that is, they may not have a true value. For example, the subgrid stresses in large eddy simulations are modeled using constitutive relationships using unknown coefficients.

The present chapter focuses on estimation of such uncertain parameters arising in a model of the upper atmosphere, which extends from about 60 km to 1000 km. The basis of this model is the Navier-Stokes equations coupled with electrodynamics equations. The source terms driving the ionosphere, such as the solar flux and the Earth's magnetic field, as well as coupling effects such as viscosity and diffusion, are constructed statistically [40] or empirically to match the predicted output of the model with measurements from various ground stations and satellites in orbit.

It is well known that turbulent diffusion affects energy deposition and transport of chemical species in the upper atmosphere. Eddy diffusion coefficient (EDC) is used to model this mixing process. Specifically, EDC models turbulent mixing in the upper atmosphere [41]. According to mass continuity and the momentum equations, the altitude profile of the neutral constituents changes from full mixing at lower altitudes (below 150 km), where turbulent mixing prevails, to molecular diffusion at higher altitudes (above 150 km). The value of EDC represents the intensity of the turbulent mixing, which is a key factor in determining the free-electron density in the ionosphere, usually measured by total electron content (TEC). However, no measurement device can directly measure EDC, and no first-principles physics model is available to determine its value. Nevertheless, models of the upper atmosphere routinely simulate the turbulent mixing at various altitudes by using shape functions parameterized by EDC [42].

The estimates of EDC obtained in the present chapter are based on measurements of the total electron content (TEC). TEC, defined as the total number of electrons integrated along a vertical column of one meter squared cross section, is a widely used quantity to describe the ionosphere. TEC is measured in TEC unit (TECU), where 1 TECU = $10^{16}$ electrons/$m^2$. The free electrons in the ionosphere cause delay in the propagation of radio waves in the atmosphere. Ionospheric irregularities also cause random amplitude and phase fluctuations in the signals [43]. TEC measurements are thus used to correct positioning errors and improve the accuracy of the Global Navigation Satellite Systems [44]. Consequently, the physics of TEC are widely studied [45–48]. TEC is routinely used to estimate the state of the ionosphere and thermosphere [49–51]. The use of TEC measurements to estimate EDC is a novel element of the present chapter.

In this chapter, we use retrospective cost parameter estimation (RCPE) to the problem of estimating EDC. RCPE was applied to atmospheric models in [27, 52–54].

The present chapter extends RCPE by simultaneously optimizing the filter $G_{\mathrm{f}}$, which is used to define the retrospective cost function [55, 56]. Within the context of adaptive control, $G_{\mathrm{f}}$ is chosen to capture knowledge of the leading sign, relative degree, and nonminimum-phase zeros. For parameter estimation within the context of a high-order nonlinear model such as the atmospheric model considered in this chapter, there are no clear guidelines for constructing $G_{\mathrm{f}}$. Consequently, the present chapter updates $G_{\mathrm{f}}$ online through a combined optimization procedure. A related optimization technique within the context of adaptive control was considered in [57, 58].

The optimization problem involving both the unknown parameters and the filter $G_{\mathrm{f}}$ turns out to be biquadratic. Although this biquadratic function is strictly convex in each variable separately, it is highly nonconvex as a joint function of its arguments. Unfortunately, numerical algorithms that converge globally to the global minimizer of biquadratic functions are not available [59], and this poses a technical challenge within the context of parameter estimation for large-scale physics models. In the present chapter we apply an alternating convex search method that is guaranteed to converge to a local minimizer.

The chapter is organized as follows. In Section 3.2, we describe GITM and the role played by EDC. In Section 3.3, we summarize the RCPE algorithm. In section 3.4, an algorithm to optimize the biquadratic retrospective cost function is presented. A low-dimensional system illustrating the application of RCPE is presented in Section 3.5. Estimation of EDC using TEC is presented in 3.6. Finally, in Section 3.7, we summarize the results of the chapter and discuss future directions.

## 3.2 Global Ionosphere-Thermosphere Model

GITM is a computational code that models the thermosphere and the ionosphere of the Earth as well as that of various planets and moons by solving coupled continuity,

momentum, and energy equations [60]. By propagating the governing equations, GITM computes neutral, ion, and electron temperatures, neutral-wind and plasma velocities, and mass and number densities of neutrals, ions, and electrons. GITM uses a uniform grid in latitude with width $\frac{2\pi}{n_{\text{lat}}}$ rad, where $n_{\text{lat}}$ is the number of grid points. In longitude and altitude, GITM uses a stretched grid to account for temperature and density variations.

GITM is implemented in parallel, where the computational domain (the atmosphere from 100 km to 600 km) is divided into blocks. Ghost cells border the physical blocks to exchange information. GITM can be run in one-dimensional mode, where horizontal transport is ignored, or in global three-dimensional mode. Furthermore, GITM can be run at either a constant or a variable time step, which is calculated by GITM based on the physical state and the user-defined CFL number in order to maintain numerical stability. To initialize GITM, neutral and ion densities and temperatures for a chosen time are set using the Mass Spectrometer and Incoherent Scatter radar (MSIS) model [61] and International Reference Ionosphere (IRI) [62].

The model inputs for GITM are 10.7 cm solar radio flux (F10.7 index), hemispheric power index (HPI), interplanetary magnetic field (IMF), solar wind plasma (SWP), and solar irradiance, all of which are read from a text file containing the time of measurements and the measured values. These signals are available from various terrestrial sensor platforms.

## 3.3 Retrospective Cost Parameter Estimation

Consider the parameter estimator where the parameter pre-estimate $\nu$ is given by

$$\mu(k) = \sum_{i=1}^{l_{\text{w}}} P_i \nu(k-i) + \sum_{i=1}^{l_{\text{w}}} Q_i z(k-i) + R(k)\gamma(k), \tag{3.1}$$

where $l_{\mathrm{w}}$ is the data-window length, and the integrator state $\gamma(k) \in \mathbb{R}^{l_y}$ is updated by

$$\gamma(k) = \gamma(k-1) + z(k-1), \tag{3.2}$$

and $P_i(k) \in \mathbb{R}^{l_\mu \times l_\mu}$, $Q_i(k), R(k) \in \mathbb{R}^{l_\mu \times l_y}$ are the coefficient matrices, which are updated by the RCPE algorithm. Finally, the parameter estimate $\hat{\mu}(k)$ is given by

$$\hat{\mu}(k) = \mathcal{O}_{\mathrm{p}}|\nu(k)|, \tag{3.3}$$

where $\mathcal{O}_{\mathrm{p}}$ is a permutation matrix described in Section 3.3.

We rewrite (3.1) as

$$\nu(k) = \Phi(k)\theta(k), \tag{3.4}$$

where the regressor matrix $\Phi(k)$ is defined by

$$\Phi(k) \triangleq I_{l_\mu} \otimes \phi^{\mathrm{T}}(k) \in \mathbb{R}^{l_\mu \times l_\theta},$$

where

$$\phi(k) \triangleq \begin{bmatrix} \nu(k-1) \\ \vdots \\ \nu(k-l_{\mathrm{w}}) \\ z(k-1) \\ \vdots \\ z(k-l_{\mathrm{w}}) \\ g(k) \end{bmatrix}, \tag{3.5}$$

$$\theta(k) \triangleq \operatorname{vec} \left[ \begin{array}{cccc} P_1(k) \cdots P_{l_{\mathrm{w}}}(k) & Q_1(k) \cdots Q_{l_{\mathrm{w}}}(k) & R(k) \end{array} \right] \in \mathbb{R}^{l_\theta}, \tag{3.6}$$

$l_\theta \triangleq l_\mu^2 l_{\mathrm{w}} + l_\mu l_y (l_{\mathrm{w}}+1)$, "$\otimes$" is the Kronecker product, and "vec" is the column-stacking operator.

### 3.3.1 Retrospective Cost Optimization

The *retrospective error variable* is defined by

$$\hat{z}(k, \hat{\theta}, \hat{N}) = z(k) + \hat{G}_{\mathrm{f}}(\mathbf{q})(\Phi(k)\hat{\theta} - \nu(k)), \tag{3.7}$$

where $\hat{\theta} \in \mathbb{R}^{l_\theta}$ contains the parameter estimator coefficients to be optimized, and $\hat{G}_{\mathrm{f}}$ is an FIR filter of order $n_{\mathrm{f}}$ to be optimized, given by

$$\hat{G}_{\mathrm{f}}(\mathbf{q}) = \sum_{i=1}^{n_{\mathrm{f}}} \frac{1}{\mathbf{q}^i} \hat{N}_i, \tag{3.8}$$

where $\hat{N}_i \in \mathbb{R}^{l_y \times l_\mu}$ are the filter coefficients to be optimized. The retrospective error variable (3.7) can thus be written as

$$\hat{z}(k, \hat{\theta}, \hat{N}) = z(k) + \hat{N}\overline{\Phi}(k)\hat{\theta} - \hat{N}\overline{V}(k), \tag{3.9}$$

where

$$\hat{N} \triangleq \left[ \begin{array}{ccc} \hat{N}_1 & \cdots & \hat{N}_{n_{\mathrm{f}}} \end{array} \right] \in \mathbb{R}^{l_y \times n_{\mathrm{f}} l_\mu}, \tag{3.10}$$

and $\overline{\Phi}(k)$ and $\overline{V}(k)$ are defined by (2.25).

Using the retrospective error variable $\hat{z}(k, \hat{\theta}, \hat{N})$, the retrospective cost function is

defined by

$$J(k,\hat{\theta},\hat{N}) \triangleq \sum_{i=1}^{k} \lambda^{k-i} \hat{z}(k,\hat{\theta},\hat{N})^{\mathrm{T}} R_z \hat{z}(k,\hat{\theta},\hat{N}) + \lambda^k \hat{\theta}^{\mathrm{T}} R_\theta \hat{\theta}, \tag{3.11}$$

where $R_z \in \mathbb{R}^{l_y \times l_y}$ and $R_\theta \in \mathbb{R}^{l_\theta \times l_\theta}$ are positive definite, and $\lambda \in (0,1]$ is the forgetting factor. Note that the retrospective cost $J(k,\hat{\theta},\hat{N})$ is a biquadratic function, that is, $J(k,\hat{\theta},\hat{N})$ is a quadratic function of $\hat{\theta}$ for fixed $\hat{N}$, and a quadratic function of $\hat{N}$ for fixed $\hat{\theta}$. We use alternating convex search (ACS) described in [59] to optimize the retrospective cost with respect to the parameter estimator coefficients and the filter. The parameter pre-estimate is given by

$$\nu(k+1) = \Phi(k+1)\theta(k+1), \tag{3.12}$$

where $\theta(k+1)$ is the minimizer of $J(k,\hat{\theta},\hat{N})$ at the $k^{\text{th}}$ step. Finally, the parameter estimate at step $k+1$ is given by

$$\hat{\mu}(k+1) = \mathcal{O}_{\mathrm{p}}|\nu(k+1)| = \mathcal{O}_{\mathrm{p}}|\Phi(k+1)\theta(k+1)|. \tag{3.13}$$

## 3.4 Biquadratic Retrospective Cost Optimization

In this section we show that $J(k,\hat{\theta},\hat{N})$ defined by (3.11) is biquadratic as a joint function of the arguments $\hat{\theta}$ and $\hat{N}$ and strictly convex in $\hat{\theta}$ and $\hat{N}$ separately. This property suggests an optimization algorithm in which $J(k,\hat{\theta},\hat{N})$ is minimized alternately with respect to $\hat{\theta}$ and $\hat{N}$

### 3.4.1 Retrospective cost as a quadratic function of $\hat{\theta}$

We write the retrospective cost function (3.11) as

$$J(k,\hat{\theta},\hat{N}) = \hat{\theta}^{\mathrm{T}} A_\theta(k,\hat{N})\hat{\theta} + \hat{\theta}^{\mathrm{T}} b_\theta(k,\hat{N}) + b_\theta(k,\hat{N})^{\mathrm{T}}\hat{\theta} + c_\theta(k,\hat{N}), \tag{3.14}$$

where

$$A_\theta(k,\hat{N}) \triangleq \sum_{i=1}^{k} \lambda^{k-i}\overline{\Phi}(i)^{\mathrm{T}}\hat{N}^{\mathrm{T}}\hat{N}\overline{\Phi}(i) + \lambda^k R_\theta \in \mathbb{R}^{l_\theta \times l_\theta}, \tag{3.15}$$

$$b_\theta(k,\hat{N}) \triangleq \sum_{i=1}^{k} \lambda^{k-i}\big(\hat{N}\overline{\Phi}(i)\big)^{\mathrm{T}}\big(z(i) - \hat{N}\overline{V}(i)\big) \in \mathbb{R}^{l_\theta}, \tag{3.16}$$

$$c_\theta(k,\hat{N}) \triangleq \sum_{i=1}^{k} \lambda^{k-i}[\overline{V}(i)^{\mathrm{T}}\hat{N}^{\mathrm{T}}\hat{N}\overline{V}(i) + z(i)^{\mathrm{T}}z(i) + \overline{V}(i)^{\mathrm{T}}\hat{N}^{\mathrm{T}}z(i)] \in \mathbb{R}. \tag{3.17}$$

For fixed $\hat{N}$, the global minimizer $\theta^*(k,\hat{N})$ of (3.11) is given by

$$\theta^*(k,\hat{N}) = -A_\theta(k,\hat{N})^{-1} b_\theta(k,\hat{N}). \tag{3.18}$$

### 3.4.2 Retrospective cost as a quadratic function of $\hat{N}$

We write the retrospective cost function (3.11) as

$$J(k,\hat{\theta},\hat{N}) \triangleq \operatorname{tr}\big(\hat{N}A_N(k,\hat{\theta})\hat{N}^{\mathrm{T}} + \hat{N}B_N(k,\hat{\theta})^{\mathrm{T}} + B_N(k,\hat{\theta})\hat{N}^{\mathrm{T}} + C_N(k,\hat{\theta})\big), \tag{3.19}$$

where

$$A_N(k,\hat{\theta}) \triangleq \sum_{i=1}^{k} \lambda^{k-i}\big(\overline{\Phi}(i)\hat{\theta} - \overline{V}(i)\big)\big(\overline{\Phi}(i)\hat{\theta} - \overline{V}(i)\big)^{\mathrm{T}} \in \mathbb{R}^{n_{\mathrm{f}}l_\mu \times n_{\mathrm{f}}l_\mu}, \tag{3.20}$$

$$B_N(k,\hat{\theta}) \triangleq \sum_{i=1}^{k} \lambda^{k-i} z(i)\big(\overline{\Phi}(i)\hat{\theta} - \overline{V}(i)\big)^{\mathrm{T}} \in \mathbb{R}^{l_y \times n_{\mathrm{f}}l_\mu}, \tag{3.21}$$

$$C_N(k,\hat{\theta}) \triangleq \sum_{i=1}^{k} \lambda^{k-i} z(i)z(i)^{\mathrm{T}} \in \mathbb{R}^{l_y \times l_y}. \tag{3.22}$$

For fixed $\hat{\theta}$, the global minimizer $N^*(k,\hat{\theta})$ of (3.11) is given by

$$N^*(k,\hat{\theta}) \triangleq -B_N(k,\hat{\theta})A_N(k,\hat{\theta})^{-\mathrm{T}} \in \mathbb{R}^{l_y \times n_{\mathrm{f}}l_\mu}. \tag{3.23}$$

It follows from Theorem 2.1 that the pre-estimate $\nu(k)$ is constrained to lie in a subspace defined by the coefficients of $G_{\mathrm{f}}(\mathbf{q})$. In particular, for all $k \geq 1$, $\nu(k) \in \mathcal{R}([N_1^{\mathrm{T}} \ \cdots \ N_{n_{\mathrm{f}}}^{\mathrm{T}}])$. Consequently, the filter order $n_{\mathrm{f}}$ should be chosen such that

$$\mathcal{R}([N_1^{\mathrm{T}} \quad \cdots \quad N_{n_{\mathrm{f}}}^{\mathrm{T}}]) = \mathbb{R}^{l_\mu}. \tag{3.24}$$

For example, to estimate $l_\mu$ unknown parameters with one measurement, $n_{\mathrm{f}}$ should be at least $l_\mu$.

#### 3.4.2.1 Sparse Filter

The following result, given in [63], allows the construction of sparse filter coefficients.

**Proposition 3.1.** Let $N$ be a sparse matrix with $p$ nonzero entries. Let $q$ be the number of rows of $N$ where the $p$ nonzero entries appear; and let $r$ be the number of rows of $N$ where the $p$ nonzero entries appear. Then $N$ can be written as $E_1\overline{N}E_2$, where $E_1$ and $E_2$ are matrices containing zero and one, and $\overline{N}$ contains the $p$ nonzero

entries, if and only if $p = qr$.

In the case where $l_y = 1$, we set $n_{\mathrm{f}} = l_\mu$, and constrain the filter coefficients $N_i$ to lie along the canonical basis vectors. In particular, it follows from Proposition 3.1 that $N$ can be written as $N = nE_2$. For example, let $l_\mu = 2$ and $\hat{N} = [\hat{n}_1 e_2 \;\; \hat{n}_2 e_1]$. Then,

$$\hat{n} = \begin{bmatrix} \hat{n}_1 & \hat{n}_2 \end{bmatrix}, \tag{3.25}$$

$$E_2 = \begin{bmatrix} e_2 & 0 \\ 0 & e_1 \end{bmatrix}. \tag{3.26}$$

In cases where it is not possible to write $N = E_1 \overline{N} E_2$, it is possible to set certain components of the filter coefficients as constants, so that Proposition 3.1 can be used. In such cases, $N$ can be written as

$$N = E_1 n E_2 + E_0, \tag{3.27}$$

where $E_0$ contains the constants. For example,

$$N = \begin{bmatrix} n_1 & 0 & n_2 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \tag{3.28}$$

$$= E_1 \begin{bmatrix} n_1 & n_2 \end{bmatrix} E_2 + E_0, \tag{3.29}$$

where

$$E_0 \triangleq \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad E_1 \triangleq \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad E_2 \triangleq \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.30}$$

Finally, with sparse filter, the retrospective cost function (3.11) can be written as

$$J(k,\hat{\theta},\hat{n}) \triangleq \text{tr}\ \left(E_1\hat{n}\overline{A}_N(k,\hat{\theta})\hat{n}^{\mathrm{T}}E_1^{\mathrm{T}} + E_1\hat{n}\overline{B}_N(k,\hat{\theta})^{\mathrm{T}} + \overline{B}_N(k,\hat{\theta})\hat{n}^{\mathrm{T}}E_1^{\mathrm{T}} + \overline{C}_N(k,\hat{\theta})\right), \tag{3.31}$$

where

$$\overline{A}_N(k,\hat{\theta}) = E_2A_N(k,\hat{\theta})E_2^{\mathrm{T}}, \tag{3.32}$$

$$\overline{B}_N(k,\hat{\theta}) = E_2B_N(k,\hat{\theta}) + E_2A_N(k,\hat{\theta})E_0^{\mathrm{T}}, \tag{3.33}$$

$$\overline{C}_N(k,\hat{\theta}) = \overline{C}_N(k,\hat{\theta}) + E_0A_N(k,\hat{\theta})E_0^{\mathrm{T}} + E_0B_N(k,\hat{\theta}) + B_N(k,\hat{\theta})^{\mathrm{T}}E_0^{\mathrm{T}} \tag{3.34}$$

where $A_N(k,\hat{\theta})$ is given by (3.20), $B_N(k,\hat{\theta})$ is given by (3.21), and $C_N(k,\hat{\theta})$ is given by (3.22). For fixed $\hat{\theta}$, the global minimizer $n^*(k,\hat{\theta})$ of (3.31) is given by

$$n^*(k,\hat{\theta}) \triangleq -(E_1^{\mathrm{T}}E_1)^{-1}E_1^{\mathrm{T}}\overline{B}_N(k,\hat{\theta})\overline{A}_N(k,\hat{\theta})^{-\mathrm{T}}. \tag{3.35}$$

### 3.4.3 Alternating Convex Search algorithm

The ACS algorithm consists of using (3.23) and (3.18) alternately to converge to a stationary point of (3.11). At step $k$, ACS consists of the following rules:

1. Set $i = 0$ and choose nonzero $N_0$.
2. Use (3.18) with $N_i$ to compute $\theta_{i+1}$.
3. Use (3.23) with $\theta_{i+1}$ to compute $N_{i+1}$.
4. Compute $J_{i+1}(k,\theta_{i+1},N_{i+1})$ using either (3.14) or (3.19).
5. For $i > 2$, if $J_{i+1} - J_i \leq \varepsilon$, then stop, where $\varepsilon > 0$ is the user-defined stopping criteria, and set $\theta(k+1) = \theta_{i+1}$ and $N(k+1) = N_{i+1}$. Otherwise, replace $i$ by $i+1$ and go to 2).

Finally, the parameter pre-estimate is given by

$$\nu(k+1) = \Phi(k+1)\theta(k). \tag{3.36}$$

## 3.5 Illustrative example

In this section, RCPE is used to estimate an unknown scalar parameter $\mu$ that nonaffinely parameterizes a linear system realization. Consider the LTI physical system model

$$x(k+1) = A(\mu)x(k) + B(\mu)u(k) + w_1(k), \tag{3.37}$$

$$y(k) = C(\mu)x(k) + w_2(k), \tag{3.38}$$

where

$$A(\mu) = \begin{bmatrix} e^{-\mu} & 1-\mu \\ \mu^2 & \log(1+\mu^2) \end{bmatrix}, \tag{3.39}$$

$$B(\mu) = \begin{bmatrix} \sin\mu \\ 1+\cos\mu \end{bmatrix}, \tag{3.40}$$

$$C(\mu) = \begin{bmatrix} 1+\mu & \mu^2 \end{bmatrix}. \tag{3.41}$$

The true value of $\mu$ is 0.8. The estimation model is thus

$$\hat{x}(k+1) = A(\mu(k))\hat{x}(k) + B(\mu(k))u(k), \tag{3.42}$$

$$\hat{y}(k) = E(\mu(k))\hat{x}(k), \tag{3.43}$$

where $\mu(k)$ is the output of the parameter estimator (3.1), (3.3) updated by RCPE.

The measurement $y(k)$ is generated using the input $u(k) = 2 + \sin\left(\frac{2\pi}{40}k\right) +$

$\sin\left(\frac{2\pi}{80}k - 0.3\right) + \sin\left(\frac{2\pi}{160}k - 0.5\right)$, the initial state $x(0) = [10\ 10]^{\mathrm{T}}$, each component of the process noise $w_1$ is $\mathcal{N}(0, 10^{-6})$, and the measurement noise $w_2$ is $\mathcal{N}(0, 10^{-6})$. To reflect the absence of additional information, the initial state $\hat{x}(0)$ of the estimation model and the initial estimate $\mu(0)$ of the unknown parameter $\mu$ are set to zero. We use RCPE to estimate the unknown parameter $\mu$ in the linear system (3.37), (3.38) with the nonlinear parameter dependence (3.39)-(3.41). We set $l_{\mathrm{w}} = 1$, $n_{\mathrm{f}} = 1$, $R_z = 1$, and $R_\theta = 10^6 I_{l_\theta}$. At each step, ACS is initialized with $N_0 = -1$. Figure 3.1 shows the estimate $\mu(k)$ of $\mu$.

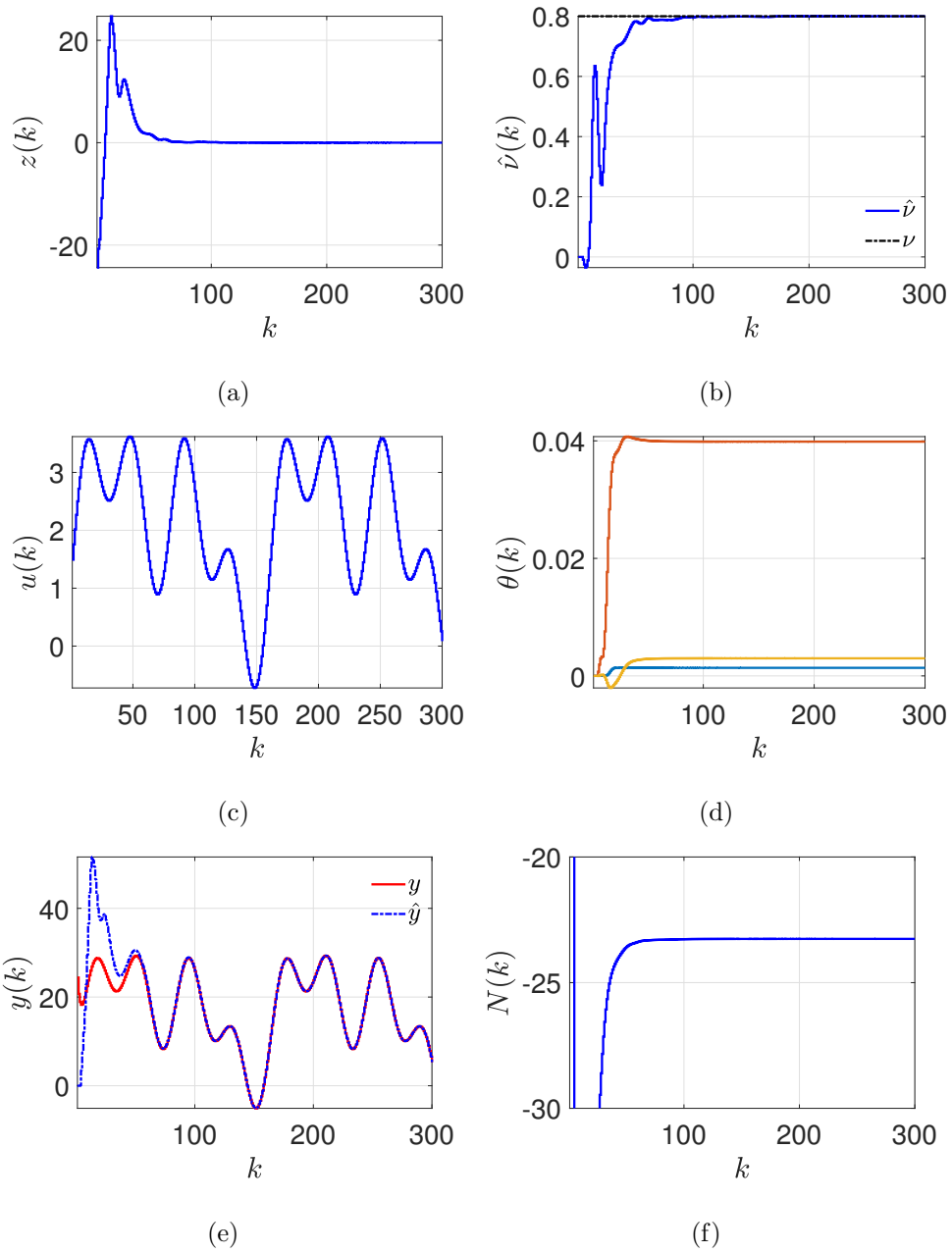



Figure 3.1: RCPE estimate of the unknown parameter $\mu$ in the linear system (3.37), (3.38) with the nonlinear parameter dependence (3.39)-(3.41). (a) shows the performance $z$ on a linear scale, (b) shows the parameter estimate $\mu$, (c) shows the measured input $u$ to the system, (d) shows the adapted coefficients $\theta$ of the parameter estimator, (e) shows the measured output and the output of the estimation model, and (f) shows the adapted coefficient $N(k)$ of the filter $G_{\mathrm{f}}$.

Next, we investigate the effect of the ACS filter initialization $N_0$ and the weight $R_\theta$ on the performance of RCPE. We set $l_{\mathrm{w}} = 1$, $n_{\mathrm{f}} = 1$, and $R_z = 1$. Figure 3.2 shows the estimate $\mu(k)$ of $\mu$ and the filter coefficient $N(k)$ for various initializations and weights. Note that RCPE successfully estimates the unknown parameter $\mu$ for ACS filter initialization choices ranging several orders of magnitude, thus indicating that ACS filter initialization choice is not critical.

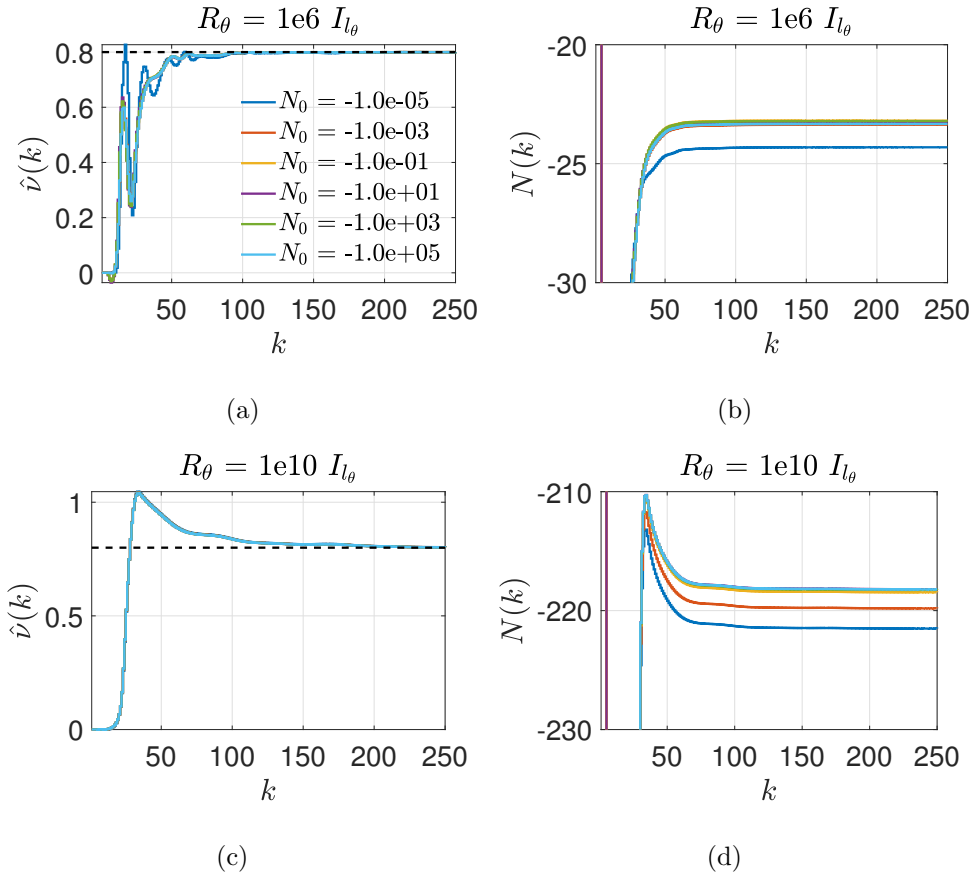



Figure 3.2: Effect of the ACS filter initialization $N_0$ and the weight $R_\theta$ on the performance of RCPE. (a) and (c) show the parameter estimate $\mu$ for various values of the filter initialization $N_0$ and the weight $R_\theta$. (b) and (d) show the filter coefficient $N$ for various values of the filter initialization $N_0$ and the weight $R_\theta$.

## 3.6 Estimation of EDC using TEC GITM

Finally, we consider the problem of estimating the EDC in the global ionosphere thermosphere model (GITM) using measurements of TEC at a fixed ground station on Earth.

To generate the measurements of TEC, we simulate the upper atmosphere of Earth using GITM for the period starting at 00:00:00, 21-Nov-2002 and ending at 00:00:00, 8-Dec-2002 with EDC = 1750. TEC is computed at every minute at a fictitious ground station located at 1 deg North, 45 deg East. The initial state of the upper atmosphere, comprising neutral and ion densities and temperature, is set using MSIS and IRI for the chosen start time. The inputs to GITM, such as F10.7 index, IMF data, SWP data, and HPI data, are read from text files.

To estimate the unknown EDC, we compute the TEC at every minute at the ground location 1 deg North, 45 deg East for the period starting from 00:00:00, 21-Nov-2002 to 00:00:00, 8-Dec-2002. The initial state of the upper atmosphere, comprising neutral and ion densities and temperature, is set using MSIS and IRI for the chosen start time. GITM is run for one simulation day, that is from 00:00:00, 21-Nov-2002 to 00:00:00, 22-Nov-2002 with EDC = 1500. At the start of the second simulation day, RCPE is switched on. Note that the delayed starting of RCPE ensures that, at the instant RCPE starts, the state of the atmosphere updated by the estimation GITM model is the not same as the GITM model used to generate the TEC measurements.

We set $l_{\mathrm{w}} = 2$, $n_{\mathrm{f}} = 1$, $R_z = 1$, and $R_\theta = 10^{-1} I_{l_\theta}$, and use RCPE to update the estimate of EDC at every minute. Figure 3.3 shows the estimate of EDC using RCPE. Figure 3.4 shows the measured and computed TEC.

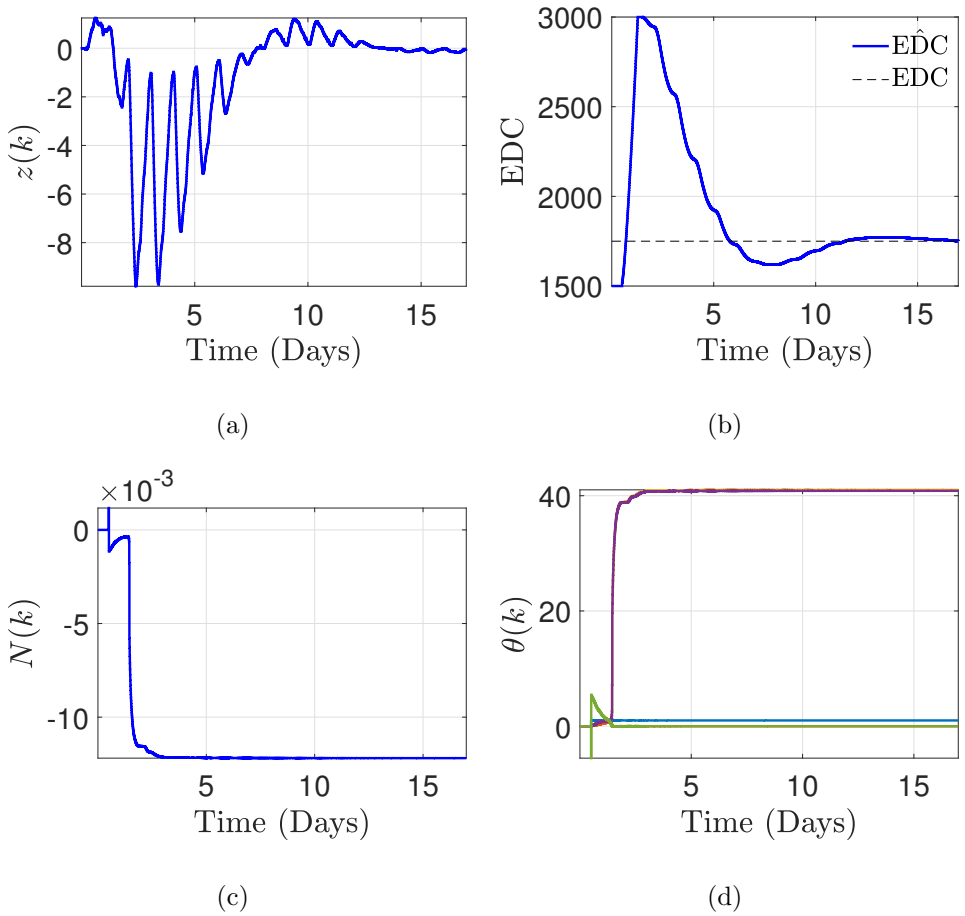



Figure 3.3: RCPE estimate of the unknown EDC in GITM. (a) shows the performance $z$ on a linear scale, (b) shows the EDC estimate, (c) shows the adapted coefficient $N(k)$ of the filter $G_{\mathrm{f}}$, and (d) shows the adapted coefficients $\theta$ of the parameter estimator.

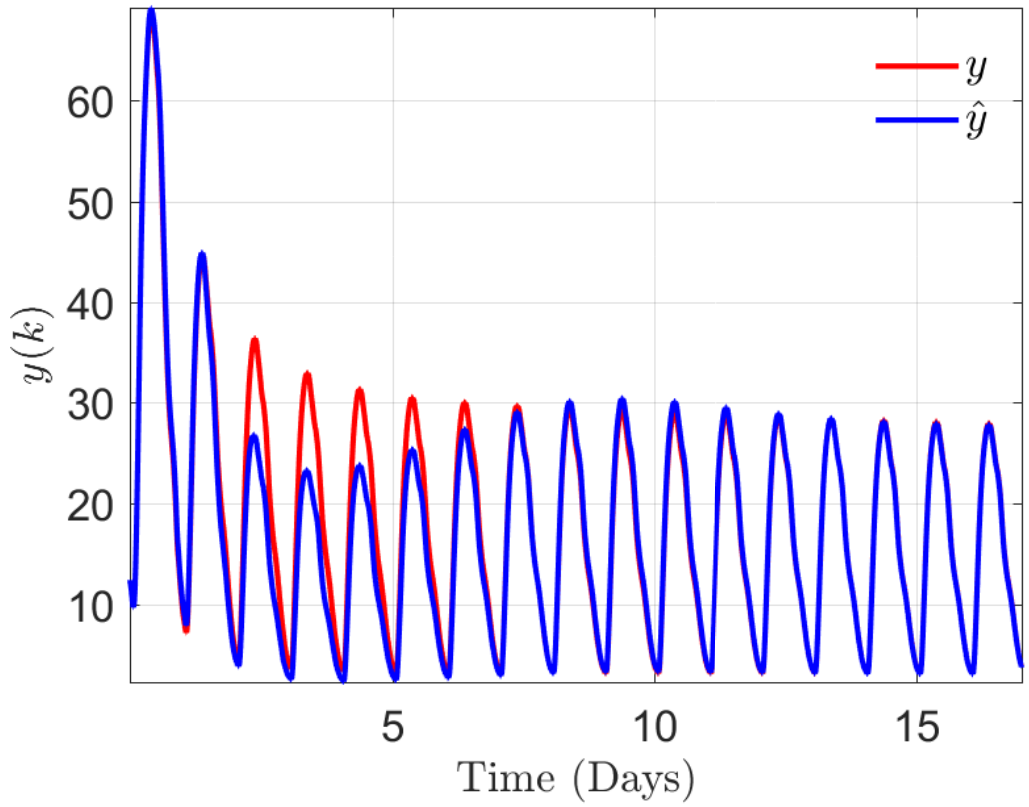



Figure 3.4: Measured and computed TEC at the fictitious ground station located at 1 deg North, 45 deg East. $y$ denotes the TEC measurements generated using GITM with a constant EDC= 1750, and $\hat{y}$ denotes the TEC computed by GITM where EDC is updated at every minute by RCPE.

Note that RCPE does not use GITM to update the parameter estimator. Instead, GITM is used as a black box model, although the EDC estimate is injected into GITM as a gray box model. In fact, the internal parameter dependence of GITM on EDC is extremely complicated; fortunately, there is no need to explicitly characterize this dependence.

## 3.7 Conclusions

This chapter presented an extension of retrospective cost parameter estimation (RCPE) by concurrently optimizing the filter $G_{\mathrm{f}}$ and the parameter estimator. This technique was used to estimate an unknown parameter in a large-scale model of a physical system, namely, the eddy diffusion coefficient in the global ionosphere-thermosphere model.

Analysis of RCPE focused on the biquadratic nature of the retrospective cost function. Alternating convex search algorithm was used to optimize the biquadratic retrospective cost function.

# CHAPTER 4

# Recursive Least Squares with Information-Driven Directional Forgetting

Within the context of system identification, persistent excitation is crucial for constructing estimators that converge to the unknown parameters. This convergence depends on two things, namely, identifiability [64], which is ability to distinguish distinct parameters, and persistent excitation, which refers to the spectral content of the signals needed to ensure convergence of the parameter estimates to the true parameter values [35, 65, 66]. Roughly speaking, the level of persistency must be commensurate with the number of unknown parameters. For example, a harmonic input has two-dimensional persistency and thus can be used to identify two parameters, whereas white noise is sufficiently persistent for identifying an arbitrary number of parameters. Within the context of adaptive control, persistent excitation is needed to avoid bursting [67]; recent research has focused on relaxing these requirements [68–70].

Under persistent excitation, a key issue in practice is the speed of convergence, especially under changing conditions. For example, the parameters of a system may change abruptly, and the goal is to ensure fast convergence to the modified parameters. The speed of convergence depends on the ability to forget past parameters and incorporate new information. Although somewhat counterintuitive, the ability

to accommodate new information depends on the ability to forget. The ability to forget is thus crucial to the ability to learn. This paradox is widely recognized, and effective forgetting is of intense interest in machine learning [71].

This chapter considers forgetting within the context of recursive least squares (RLS). In the classical RLS formulation, a constant forgetting factor $\lambda \in (0, 1]$ can be set by the user. However, it often occurs in practice that if $\lambda$ is too small—perhaps as small as 0.99—then RLS may diverge. This phenomenon has been extensively studied and various remedies have been studied in the form of variable-rate forgetting [72–78].

The present chapter considers a technique that complements variable-rate forgetting, namely, *directional forgetting*. Directional forgetting has been widely studied within the context of recursive least squares [79–85]. In the absence of persistent excitation, new information is confined to a limited number of directions. The goal of directional forgetting is thus to determine these directions and thereby constrain forgetting to the directions in which new information is available. This technique allows RLS to operate without divergence during periods of loss of persistency.

The chapter is organized as follows. Section 4.1 presents the classical recursive least squares algorithm with constant scalar forgetting factor. Section 4.2 defines persistency of excitation and presents results on bounds of the estimator variables. Section 4.3 connects persistency of excitation of a sequence to the condition number of the information matrix. Section 4.4 presents Lyapunov analysis of the parameter error. Section 4.5 shows the effect of lack of persistent excitation on the estimator. Section 4.6 introduces the information subspace, and Section 4.7 introduces the matrix forgetting that constrains forgetting to the information directions receiving new information. Finally, the chapter concludes in Section 4.8.

Table 4.1 summarizes the results and examples in this chapter.

| | |
|---|---|
| Definition 4.1 | Persistently exciting regressor |
| Definition 4.2 | Uniformly Lyapunov stable equilibrium |
| Definition 4.3 | Uniformly globally aymptotically stable equilibrium |
| Definition 4.4 | Uniformly globally geometrically stable equilibrium |
| Theorem 4.1 | Recursive least squares (RLS) |
| Theorem 4.2 | Lyapunov stability theorem |
| Theorem 4.3 | Lyapunov analysis of RLS for $\lambda \in (0, 1)$ |
| Theorem 4.4 | Stability analysis of RLS for $\lambda \in (0, 1]$ based on $\theta_k$ |
| Proposition 4.1 | Convergence of $z_k$ and $\theta_k$ |
| Proposition 4.2 | Data-dependent subspace constraint on $\theta_k$ |
| Proposition 4.3 | Bounds on $P_k$ for $\lambda = 1$ |
| Proposition 4.4 | Bounds on $P_k$ for $\lambda \in (0, 1)$ |
| Proposition 4.5 | Converse of Proposition 4.4 |
| Proposition 4.6 | Persistent excitation and $\mathcal{A}_k$ |
| Example 4.1 | Necessary, but not sufficient condition for persistent excitation |
| Example 4.2 | Persistent excitation and bounds on $P_k^{-1}$ |
| Example 4.3 | Lack of persistent excitation and bounds on $P_k^{-1}$ |
| Example 4.4 | Using $\kappa(P_k)$ to determine whether $(\phi_k)_{k=0}^{\infty}$ is persistently exciting |
| Example 4.5 | Effect of $\lambda$ on the rate of convergence of $\theta_k$ |
| Example 4.6 | Lack of persistent excitation in scalar estimation |
| Example 4.7 | Subspace constrained regressor |
| Example 4.8 | Effect of lack of persistent excitation on $\theta_k$ |
| Example 4.9 | Lack of persistent excitation and the information subspace |
| Example 4.10 | Information-driven forgetting for a regressor lacking persistent excitation |
| Example 4.11 | Effect of Information-driven forgetting on $\theta_k$ |

Table 4.1: Summary of definitions, results, and examples in this chapter.

## 4.1 Recursive Least Squares

Consider the model

$$y_k = \phi_k \theta, \tag{4.1}$$

where, for all $k \geq 0$, $y_k \in \mathbb{R}^p$ is the measurement, $\phi_k \in \mathbb{R}^{p \times n}$ is the regressor matrix, and $\theta \in \mathbb{R}^n$ is the vector of unknown parameters. The goal is to estimate $\theta$ as new data becomes available. One approach to this problem is to minimize the quadratic cost function

$$J_k(\hat{\theta}) \triangleq \sum_{i=0}^{k} \lambda^{k-i}(y_i - \phi_i \hat{\theta})^{\mathrm{T}}(y_i - \phi_i \hat{\theta}) + \lambda^{k+1}(\hat{\theta} - \theta_0)^{\mathrm{T}} R(\hat{\theta} - \theta_0), \tag{4.2}$$

where $\lambda \in (0, 1]$ is the *forgetting factor*, $R \in \mathbb{R}^{n \times n}$ is positive definite, and $\theta_0 \in \mathbb{R}^n$ is the initial estimate of $\theta$. The forgetting factor applies higher weighting to more recent data, which is useful for estimating time-varying parameters. The following result is *recursive least squares*.

**Theorem 4.1.** *For all $k \geq 0$, let $\phi_k \in \mathbb{R}^{p \times n}$ and $y_k \in \mathbb{R}^p$, let $R \in \mathbb{R}^{n \times n}$ be positive definite, and let $P_0 = R^{-1}$, $\theta_0 \in \mathbb{R}^n$, and $\lambda \in (0, 1]$. Furthermore, for all $k \geq 0$, denote the minimizer of (4.2) by*

$$\theta_{k+1} = \operatorname*{argmin}_{\hat{\theta} \in \mathbb{R}^n} J_k(\hat{\theta}). \tag{4.3}$$

*Then, for all $k \geq 0$, $\theta_{k+1}$ is given by*

$$P_{k+1} = \frac{1}{\lambda} P_k - \frac{1}{\lambda} P_k \phi_k^{\mathrm{T}} \left( \lambda I_p + \phi_k P_k \phi_k^{\mathrm{T}} \right)^{-1} \phi_k P_k, \tag{4.4}$$

$$\theta_{k+1} = \theta_k + P_{k+1} \phi_k^{\mathrm{T}} (y_k - \phi_k \theta_k). \tag{4.5}$$

*Furthermore, for all $k \geq 0$, $P_k$ is positive definite and*

$$P_{k+1}^{-1} = \lambda P_k^{-1} + \phi_k^{\mathrm{T}} \phi_k. \tag{4.6}$$

*Proof.* See [86]. □

Let $k \geq 0$. By defining the *parameter error*

$$\tilde{\theta}_k \triangleq \theta_k - \theta, \tag{4.7}$$

it follows that

$$\phi_i \theta_k - y_i = \phi_i \tilde{\theta}_k, \tag{4.8}$$

and

$$J_k(\theta_{k+1}) = \sum_{i=0}^{k} \lambda^{k-i} \tilde{\theta}_{k+1}^{\mathrm{T}} \phi_i^{\mathrm{T}} \phi_i \tilde{\theta}_{k+1} + \lambda^{k+1} (\tilde{\theta}_{k+1} - \tilde{\theta}_0)^{\mathrm{T}} R (\tilde{\theta}_{k+1} - \tilde{\theta}_0). \tag{4.9}$$

Furthermore, it follows from (4.5) and (4.7) that $\tilde{\theta}_k$ satisfies

$$\tilde{\theta}_{k+1} = (I_n - P_{k+1} \phi_k^{\mathrm{T}} \phi_k) \tilde{\theta}_k \tag{4.10}$$

$$= \lambda P_{k+1} P_k^{-1} \tilde{\theta}_k. \tag{4.11}$$

Note that (4.11) implies that, for all $k, l \geq 0$,

$$\tilde{\theta}_k = \lambda^{k-l} P_k P_l^{-1} \tilde{\theta}_l. \tag{4.12}$$

The following lemma is used in the proof of Proposition 4.1.

**Lemma 4.1.1.** *Let $A \in \mathbb{R}^{n\times n}$ be positive semidefinite, and let $\lambda > 0$. Then,*

$$I_n - A(\lambda I_n + A)^{-1} > 0. \tag{4.13}$$

*Proof.* Write $A = SDS^{\mathrm{T}}$, where $D = \mathrm{diag}(d_1, \ldots, d_n)$ is diagonal and $S$ is unitary. For all $i \in \{1, \ldots, n\}$, $d_i \geq 0$, and thus $\frac{d_i}{\lambda + d_i} < 1$. Thus,

$$D(\lambda I_n + D)^{-1} = \mathrm{diag}\left(\tfrac{d_1}{\lambda + d_1}, \ldots, \tfrac{d_n}{\lambda + d_n}\right) < I_n. \tag{4.14}$$

Pre-multiplying and post-multiplying (4.14) by $S$ and $S^{\mathrm{T}}$, respectively, yields (4.13). □ □

The following result shows that the estimate $\theta_k$ of $\theta$ converges.

**Proposition 4.1.** *For all $k \geq 0$, let $\phi_k \in \mathbb{R}^{p\times n}$ and $y_k \in \mathbb{R}^p$, let $R \in \mathbb{R}^{n\times n}$ be positive definite, and let $P_0 = R^{-1}$, $\theta_0 \in \mathbb{R}^n$, and $\lambda \in (0, 1]$. Let $P_k$ and $\theta_k$ be given by* (4.4) *and* (4.5)*, respectively, and define the predicted error $z_k \stackrel{\triangle}{=} \phi_k\theta_k - y_k$. Then,*

$$\lim_{k\to\infty} z_k = 0 \tag{4.15}$$

*and*

$$\lim_{k\to\infty} (\theta_{k+1} - \theta_k) = 0. \tag{4.16}$$

*Proof.* Note that $z_k = \phi_k\tilde{\theta}_k$, and, for all $k \geq 0$, define $V_k \stackrel{\triangle}{=} \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} \tilde{\theta}_k$. Note that, for

all $k \geq 0$ and $\tilde{\theta}_k \in \mathbb{R}^n$, $V_k \geq 0$. Furthermore, for all $k \geq 0$,

$$\begin{aligned}
V_{k+1} - V_k &= \tilde{\theta}_{k+1}^{\mathrm{T}} P_{k+1}^{-1} \tilde{\theta}_{k+1} - \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} \tilde{\theta}_k \\
&= \lambda^2 \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} P_{k+1} P_k^{-1} \tilde{\theta}_k - \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} \tilde{\theta}_k \\
&= \left( \lambda \tilde{\theta}_{k+1}^{\mathrm{T}} - \tilde{\theta}_k^{\mathrm{T}} \right) P_k^{-1} \tilde{\theta}_k \\
&= \left( -(1-\lambda) \tilde{\theta}_k^{\mathrm{T}} - \lambda \tilde{\theta}_k^{\mathrm{T}} \phi_k^{\mathrm{T}} \phi_k P_{k+1} \right) P_k^{-1} \tilde{\theta}_k \\
&= -(1-\lambda) \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} \tilde{\theta}_k - \lambda \tilde{\theta}_k^{\mathrm{T}} \phi_k^{\mathrm{T}} \phi_k P_{k+1} P_k^{-1} \tilde{\theta}_k \\
&= -(1-\lambda) \tilde{\theta}_k^{\mathrm{T}} P_k^{-1} \tilde{\theta}_k - \tilde{\theta}_k^{\mathrm{T}} \phi_k^{\mathrm{T}} \left( I_p - \phi_k P_k \phi_k^{\mathrm{T}} (\lambda I_p + \phi_k P_k \phi_k^{\mathrm{T}})^{-1} \right) \phi_k \tilde{\theta}_k \\
&= - \left( (1-\lambda) V_k + z_k^{\mathrm{T}} \left( I_p - \phi_k P_k \phi_k^{\mathrm{T}} (\lambda I_p + \phi_k P_k \phi_k^{\mathrm{T}})^{-1} \right) z_k \right) \\
&\leq 0.
\end{aligned}$$

Note that, since $(V_k)_{k=1}^{\infty}$ is a nonnegative, nonincreasing sequence, it converges to a nonnegative number. Hence $\lim_{k \to \infty} (V_{k+1} - V_k) = 0$, which implies that

$$\lim_{k \to \infty} \left( (1-\lambda) V_k + z_k^{\mathrm{T}} R_k z_k \right) = 0,$$

where $R_k \stackrel{\triangle}{=} I_p - \phi_k P_k \phi_k^{\mathrm{T}} (\lambda I_p + \phi_k P_k \phi_k^{\mathrm{T}})^{-1}$. Lemma 4.1.1 implies that $R_k$ is positive definite. Since $V_k \geq 0$, it follows that $\lim_{k \to \infty} z_k = 0$. Finally, (4.16) follows from (4.5) and (4.15). □

The following lemma is used in the proof of Proposition 4.2.

**Lemma 4.1.2.** *Let $X \in \mathbb{R}^{p \times n}$ and $y \in \mathbb{R}^{1 \times n}$, and let $W \in \mathbb{R}^{p \times p}$ be positive definite. Then,*

$$(I_n + X^{\mathrm{T}} W X)^{-1} y^{\mathrm{T}} \in \mathcal{R}([X^{\mathrm{T}} \;\; y^{\mathrm{T}}]). \tag{4.17}$$

*Proof.* Note that

$$\begin{aligned}
y^{\mathrm{T}} &\in \mathcal{R}([X^{\mathrm{T}} \;\; y^{\mathrm{T}}]) \\
&= \mathcal{R}[X^{\mathrm{T}} \;\; y^{\mathrm{T}} + X^{\mathrm{T}}WXy^{\mathrm{T}}] \\
&= \mathcal{R}\left( [X^{\mathrm{T}} \;\; (I_n + X^{\mathrm{T}}WX)y^{\mathrm{T}}] \begin{bmatrix} I_p + WXX^{\mathrm{T}} & 0 \\ 0 & 1 \end{bmatrix} \right) \\
&= \mathcal{R}([X^{\mathrm{T}}(I_p + WXX^{\mathrm{T}}) \;\; (I_n + X^{\mathrm{T}}WX)y^{\mathrm{T}}]) \\
&= \mathcal{R}([(I_n + X^{\mathrm{T}}WX)X^{\mathrm{T}} \;\; (I_n + X^{\mathrm{T}}WX)y^{\mathrm{T}}]) \\
&= (I_n + X^{\mathrm{T}}WX)\mathcal{R}([X^{\mathrm{T}} \;\; y^{\mathrm{T}}]),
\end{aligned}$$

which implies (4.17). □

The following result shows that the estimate $\theta_k$ of $\theta$ is constrained to a data-dependent subspace.

**Proposition 4.2.** *For all $k \geq 0$, let $\phi_k \in \mathbb{R}^{p\times n}$ and $y_k \in \mathbb{R}^p$, let $R \in \mathbb{R}^{n\times n}$ be positive definite, let $\theta_0 \in \mathbb{R}^n$, let $\lambda \in (0, 1]$, and define $\theta_{k+1}$ by (4.3). Then, $\theta_{k+1}$ satisfies*

$$\left( \sum_{i=0}^{k} \lambda^{k-i} \phi_i^{\mathrm{T}} \phi_i + \lambda^{k+1} R \right) \theta_{k+1} = \sum_{i=0}^{k} \lambda^{k-i} \phi_i^{\mathrm{T}} y_i + \lambda^{k+1} R\theta_0. \tag{4.18}$$

*Furthermore,*

$$\theta_{k+1} \in \mathcal{R}(\Phi_k^{\mathrm{T}}\Phi_k + R^{-1}\Phi_k^{\mathrm{T}}\Phi_k R^{-1} + \theta_0\theta_0^{\mathrm{T}}), \tag{4.19}$$

*where*

$$\Phi_k \triangleq [\phi_0^{\mathrm{T}} \;\; \cdots \;\; \phi_k^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{(k+1)p\times n}. \tag{4.20}$$

*Proof.* Note that

$$J_k(\hat{\theta}) = \hat{\theta}^{\mathrm{T}} A_k \hat{\theta} + \hat{\theta}^{\mathrm{T}} b_k + c_k,$$

where

$$\begin{aligned} A_k &\triangleq \left( \sum_{i=0}^{k} \lambda^{k-i} \phi_i^{\mathrm{T}} \phi_i + \lambda^{k+1} R \right), \\ b_k &\triangleq \sum_{i=0}^{k} -\lambda^{k-i} \phi_i^{\mathrm{T}} y_i - \lambda^{k+1} R \theta_0, \\ c_k &\triangleq \sum_{i=0}^{k} \lambda^{k-i} y_i^{\mathrm{T}} y_i + \lambda^{k+1} \theta_0^{\mathrm{T}} R \theta_0. \end{aligned}$$

Since $A_k$ is positive definite, it follows from Lemma 1 in [86] that the minimizer $\theta_{k+1}$ of $J_k$ satisfies (4.18).

Next, define $W_k \triangleq \operatorname{diag}(\lambda^{-1} I_p, \ldots, \lambda^{-1-k} I_p) \in \mathbb{R}^{(k+1)p \times (k+1)p}$. Using (4.18) and Lemma 4.1.2, it follows that

$$\begin{aligned} \theta_{k+1} &= \left( I_n + \Phi_k^{\mathrm{T}} W_k \Phi_k \right)^{-1} \left( \sum_{i=0}^{k} \lambda^{-i-1} R^{-1} \phi_i^{\mathrm{T}} y_i + \theta_0 \right) \\ &= \sum_{i=0}^{k} \left( I_n + \Phi_k^{\mathrm{T}} W_k \Phi_k \right)^{-1} \lambda^{-i-1} R^{-1} \phi_i^{\mathrm{T}} y_i + \left( I_n + \Phi_k^{\mathrm{T}} W_k \Phi_k \right)^{-1} \theta_0 \\ &\in \sum_{i=0}^{k} \mathcal{R}([\Phi_k^{\mathrm{T}} \quad R^{-1} \phi_i^{\mathrm{T}}]) + \mathcal{R}([\Phi_k^{\mathrm{T}} \quad \theta_0]) \\ &= \mathcal{R}([\Phi_k^{\mathrm{T}} \quad R^{-1} \Phi_k^{\mathrm{T}} \quad \theta_0]) \\ &= \mathcal{R}(\Phi_k^{\mathrm{T}} \Phi_k + R^{-1} \Phi_k^{\mathrm{T}} \Phi_k R^{-1} + \theta_0 \theta_0^{\mathrm{T}}). \end{aligned}$$

□

Table 4.2 summarizes the various expressions for the RLS variables.

| Variable | Equation | Reference |
|---|---|---|
| $P_k$ | • $P_{k+1} = \frac{1}{\lambda}P_k - \frac{1}{\lambda}P_k\phi_k^{\mathrm{T}}\left(\lambda I_p + \phi_k P_k \phi_k^{\mathrm{T}}\right)^{-1}\phi_k P_k$ | (4.4) |
| | • $P_{k+1}^{-1} = \lambda P_k^{-1} + \phi_k^{\mathrm{T}}\phi_k$ | (4.6) |
| | • $P_{k+1}^{-1} = \lambda^{k+1}P_0^{-1} + \sum_{i=0}^{k}\lambda^{k-i}\phi_i^{\mathrm{T}}\phi_i$ | (4.6) |
| $\theta_k$ | • $\theta_{k+1} = \theta_k + P_{k+1}\phi_k^{\mathrm{T}}(y_k - \phi_k\theta_k)$ | (4.5) |
| | • $\theta_{k+1} = \theta_k + P_k\phi_k^{\mathrm{T}}\left(\lambda I_p + \phi_k P_k\phi_k^{\mathrm{T}}\right)^{-1}(y_k - \phi_k\theta_k)$ | [86] |
| | • $\theta_{k+1} = P_{k+1}\left(\sum_{i=0}^{k}\lambda^{k-i}\phi_i^{\mathrm{T}}y_i + \lambda^{k+1}P_0^{-1}\theta_0\right)$ | (4.18) |
| $\tilde{\theta}_k$ | • $\tilde{\theta}_k = \theta_k - \theta$ | (4.7) |
| | • $\tilde{\theta}_{k+1} = (I_n - P_{k+1}\phi_k^{\mathrm{T}}\phi_k)\tilde{\theta}_k$ | (4.10) |
| | • $\tilde{\theta}_{k+1} = \lambda P_{k+1}P_k^{-1}\tilde{\theta}_k$ | (4.11) |
| | • $\tilde{\theta}_k = \lambda^{k-l}P_kP_l^{-1}\tilde{\theta}_l$ | (4.12) |

Table 4.2: Various expressions for RLS variables.

## 4.2 Persistent Excitation and Forgetting

In this section, we define persistent excitation of the regressor sequence and investigate the effect of persistent excitation and forgetting on $P_k$. For all $j \geq 0$ and $k \geq j$, define

$$F_{j,k} \triangleq \sum_{i=j}^{k}\phi_i^{\mathrm{T}}\phi_i. \tag{4.21}$$

**Definition 4.1.** The sequence $(\phi_k)_{k=0}^{\infty} \subset \mathbb{R}^{p\times n}$ is *persistently exciting* if there exist $N \geq n/p$ and $\alpha, \beta \in (0,\infty)$ such that, for all $j \geq 0$,

$$\alpha I_n \leq F_{j,j+N} \leq \beta I_n. \tag{4.22}$$

Suppose that $(\phi_k)_{k=0}^{\infty}$ is persistently exciting and (4.22) is satisfied for $N, \alpha, \beta$. Then, it can be seen that, with suitable values of $\alpha$ and $\beta$, (4.22) is satisfied for all larger values of $N$. For example, if $N$ is replaced by $2N$, then (4.22) is satisfied with $\alpha$ replaced by $2\alpha$ and $\beta$ replaced by $2\beta$.

**Lemma 4.2.1.** *Let $\lambda = 1$. For all $k \geq 1$, define $P_k$ as in Theorem 4.1. Then,*

$$P_k^{-1} = F_{0,k} + P_0^{-1}. \tag{4.23}$$

The following result shows that, if $(\phi_k)_{k=0}^{\infty}$ is persistently exciting and $\lambda = 1$, then $P_k$ converges to zero.

**Proposition 4.3.** *Assume that $(\phi_k)_{k=0}^{\infty} \in \mathbb{R}^{p\times n}$ is persistently exciting, let $N, \alpha, \beta$ be given by Definition 4.1, let $R \in \mathbb{R}^{n\times n}$ be positive definite, define $P_0 \triangleq R^{-1}$, let $\lambda = 1$, and, for all $k \geq 0$, let $P_k$ be given by (4.4). Then, for all $k \geq N+1$,*

$$\lfloor \tfrac{k}{N+1} \rfloor \alpha I_n + P_0^{-1} \leq P_k^{-1} \leq \lceil \tfrac{k}{N+1} \rceil \beta I_n + P_0^{-1}. \tag{4.24}$$

*Furthermore,*

$$\lim_{k\to\infty} P_k = 0. \tag{4.25}$$

*Proof.* First, note that, for all $k \geq 0$,

$$\begin{aligned} F_{0,k} &= \sum_{i=1}^{\lfloor \frac{k}{N+1} \rfloor} F_{(i-1)(N+1),i(N+1)-1} + F_{\lfloor \frac{k}{N+1} \rfloor (N+1),k} \\ &\leq \sum_{i=1}^{\lceil \frac{k}{N+1} \rceil} F_{(i-1)(N+1),i(N+1)-1}, \end{aligned}$$

and thus (4.22) implies that

$$\begin{aligned}\lfloor \tfrac{k}{N+1} \rfloor \alpha I_n &\leq \sum_{i=1}^{\lfloor \frac{k}{N+1} \rfloor} F_{(i-1)(N+1),i(N+1)-1} \\ &\leq \sum_{i=1}^{\lceil \frac{k}{N+1} \rceil} F_{(i-1)(N+1),i(N+1)-1} \\ &\leq \lceil \tfrac{k}{N+1} \rceil \beta I_n. \end{aligned} \tag{4.26}$$

It follows from Lemma 4.2.1 and (4.26) that, for all $k \geq N+1$,

$$\begin{aligned}\lfloor \tfrac{k}{N+1} \rfloor \alpha I_n + P_0^{-1} &\leq F_{0,\lfloor \frac{k}{N+1} \rfloor (N+1)-1} + P_0^{-1} \\ &\leq F_{0,k} + P_0^{-1} \\ &= P_k^{-1} \\ &\leq F_{0,\lceil \frac{k}{N+1} \rceil (N+1)-1} + P_0^{-1} \\ &\leq \lceil \tfrac{k}{N+1} \rceil \beta I_n + P_0^{-1}. \end{aligned}$$

Finally, it follows from (4.24) that $\lim_{k\to\infty} P_k = 0$. □

Example 4.1 shows that $\lim_{k\to\infty} P_k = 0$ is a necessary condition $(\phi_k)_{k=0}^{\infty}$ to be persistently exciting, but not a sufficient condition.

**Example 4.1. Necessary, but not sufficient condition for persistent excitation.** For all $k \geq 0$, let $\phi_k = \frac{1}{\sqrt{k+1}}$. Let $\lambda = 1$. For all $N \geq 1$, note that $F_{j,j+N} \leq \frac{N+1}{j+1}$, and thus there does not exist $\alpha$ satisfying (4.22). Hence, $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. However, it follows from (4.6) that, for all $k \geq 0$,

$$P_k^{-1} = \sum_{i=0}^{k} \frac{1}{i+1} + P_0^{-1}. \tag{4.27}$$

Thus, $\lim_{k\to\infty} P_k = 0$. ◇

The following result given in [87] shows that, if $(\phi_k)_{k=0}^{\infty}$ is persistently exciting and $\lambda \in (0,1)$, then $P_k$ is bounded.

**Proposition 4.4.** *Assume that $(\phi_k)_{k=0}^{\infty} \in \mathbb{R}^{p\times n}$ is persistently exciting, let $N, \alpha, \beta$ be given by Definition 4.1, let $R \in \mathbb{R}^{n\times n}$ be positive definite, define $P_0 \triangleq R^{-1}$, let $\lambda \in (0,1)$, and, for all $k \geq 0$, let $P_k$ be given by (4.4). Then, for all $k \geq N+1$,*

$$\frac{\lambda^N(1-\lambda)\alpha}{1-\lambda^{N+1}} I_n \leq P_k^{-1} \leq \frac{\beta}{1-\lambda^{N+1}} I_n + P_N^{-1}. \tag{4.28}$$

*Proof.* It follows from (4.6) that, for all $i \geq 0$, $\lambda P_i^{-1} \leq P_{i+1}^{-1}$ and $\phi_i^{\mathrm{T}}\phi_i \leq P_{i+1}^{-1}$, and thus, for all $i, j \geq 0$, $\lambda^j P_i^{-1} \leq P_{i+j}^{-1}$. Hence, for all $k \geq N+1$,

$$\begin{aligned}
\alpha I_n &\leq \sum_{i=k-N-1}^{k-1} \phi_i^{\mathrm{T}}\phi_i \\
&\leq \sum_{i=k-N}^{k} P_i^{-1} \\
&\leq (\lambda^{-N} + \cdots + 1) P_k^{-1} \\
&= \frac{1-\lambda^{N+1}}{\lambda^N(1-\lambda)} P_k^{-1},
\end{aligned}$$

which proves the first inequality in (4.28). To prove the second inequality in (4.28),

note that, for all $k \geq N+1$,

$$\begin{aligned} P_k^{-1} &\leq \frac{1-\lambda}{1-\lambda^{N+1}} \sum_{i=k-1}^{k+N-1} P_{i+1}^{-1} \\ &\leq \frac{1-\lambda}{1-\lambda^{N+1}} \left( \lambda \sum_{i=k-1}^{k+N-1} P_i^{-1} + \beta I_n \right) \\ &\leq \frac{1-\lambda}{1-\lambda^{N+1}} \left( \lambda^k \sum_{i=0}^{N} P_i^{-1} + \frac{1-\lambda^k}{1-\lambda} \beta I_n \right) \\ &\leq \lambda^{k-N} P_N^{-1} + \frac{(1-\lambda^k)\beta}{1-\lambda^{N+1}} I_n. \\ &\leq P_N^{-1} + \frac{\beta}{1-\lambda^{N+1}} I_n. \end{aligned}$$

□

The next result, which is an immediate consequence of (4.6), is a converse of Proposition 4.4.

**Proposition 4.5.** *Define $\phi_k$, $y_k$, $R$, and $P_0$ as in Theorem 4.1. Let $\lambda \in (0,1)$, and let $P_k$ be given by (4.4). Furthermore, assume there exist $\overline{\alpha}, \overline{\beta} \in (0,\infty)$ such that, for all $k \geq 0$, $\overline{\alpha} I_n \leq P_k^{-1} \leq \overline{\beta} I_n$. Let $N \geq \frac{\lambda\overline{\beta}-\overline{\alpha}}{(1-\lambda)\overline{\alpha}}$. Then, for all $j \geq 0$,*

$$[(1+(1-\lambda)N)\overline{\alpha} - \lambda\overline{\beta}] I_n \leq \sum_{i=j}^{j+N} \phi_i^{\mathrm{T}} \phi_i \leq \frac{1-\lambda^{N+1}}{\lambda^N(1-\lambda)} \overline{\beta} I_n. \tag{4.29}$$

*Consequently, $(\phi_k)_{k=0}^{\infty}$ is persistently exciting.*

*Proof.* Note that, for all $j \geq 0$,

$$\begin{aligned}
[(1+(1-\lambda)N)\overline{\alpha}-\lambda\overline{\beta}]I_n &= \overline{\alpha}I_n + (1-\lambda)N\overline{\alpha}I_n - \overline{\beta}I_n \\
&\leq P_{j+N+1}^{-1} + (1-\lambda)\sum_{i=j+1}^{j+N} P_i^{-1} - \lambda P_j^{-1} \\
&= \sum_{i=j}^{j+N}(P_{i+1}^{-1} - \lambda P_i^{-1}) \\
&= \sum_{i=j}^{j+N} \phi_i^{\mathrm{T}}\phi_i,
\end{aligned}$$

which proves the first inequality in (4.29). To prove the second inequality in (4.29), note that (4.6) implies that, for all $i \geq 0$, $\lambda P_i^{-1} \leq P_{i+1}^{-1}$ and $\phi_i^{\mathrm{T}}\phi_i \leq P_{i+1}^{-1}$, and thus, for all $i, j \geq 0$, $\lambda^j P_i^{-1} \leq P_{i+j}^{-1}$. Hence, for all $j \geq 0$,

$$\begin{aligned}
\sum_{i=j}^{j+N} \phi_i^{\mathrm{T}}\phi_i &\leq \sum_{i=j}^{j+N} P_{i+1}^{-1} \\
&\leq (\lambda^{-N} + \cdots + 1)P_{j+N+1}^{-1} \\
&\leq \frac{1-\lambda^{N+1}}{\lambda^N(1-\lambda)}\overline{\beta}I_n.
\end{aligned}$$

Finally, it follows from Definition 4.1 with $N \geq \frac{\lambda\overline{\beta}-\overline{\alpha}}{(1-\lambda)\overline{\alpha}}$, $\alpha = (1+(1-\lambda)N)\overline{\alpha} - \lambda\overline{\beta}$, and $\beta = \frac{1-\lambda^{N+1}}{\lambda^N(1-\lambda)}\overline{\beta}$, that $(\phi_k)_{k=0}^{\infty}$ is persistently exciting. □

Note that the proof of Proposition 4.5 shows that the lower bound in Definition 4.1 requires that $N$ satisfy $N \geq \frac{\lambda\overline{\beta}-\overline{\alpha}}{(1-\lambda)\overline{\alpha}}$. However, the upper bound in Definition 4.1 is satisfied for all $N \geq 1$.

**Example 4.2. Persistent excitation and bounds on $P_k^{-1}$.** Let $\phi_k = [u_k \;\; u_{k-1}]$, where $u_k$ is the periodic signal

$$u_k = \sin\frac{2\pi k}{17} + \sin\frac{2\pi k}{23} + \sin\frac{2\pi k}{53}. \tag{4.30}$$

For $j = 1, \ldots, 5000$, Figures 4.1a) and b) show the singular values of $F_{j,j+N}$ for $N = 2$ and $N = 10$. For each value of $N$, the constants $\alpha, \beta$ are chosen to satisfy (4.22). Letting $\lambda = 0.99$, Figure 4.1c) shows the singular values of $P_k^{-1}$ with the corresponding upper and lower bounds given by (4.28) for $N = 2$ and $N = 10$. Since $u_k$ is periodic, it follows that, for all $j \geq 0$, the lower and upper bounds (4.22) on $F_{j,j+N}$ are satisfied. Hence, $(\phi_k)_{k=0}^{\infty}$ is persistently exciting. Finally, note that $\alpha$ and $\beta$ are larger for $N = 10$ than for $N = 2$, as expected. ◇

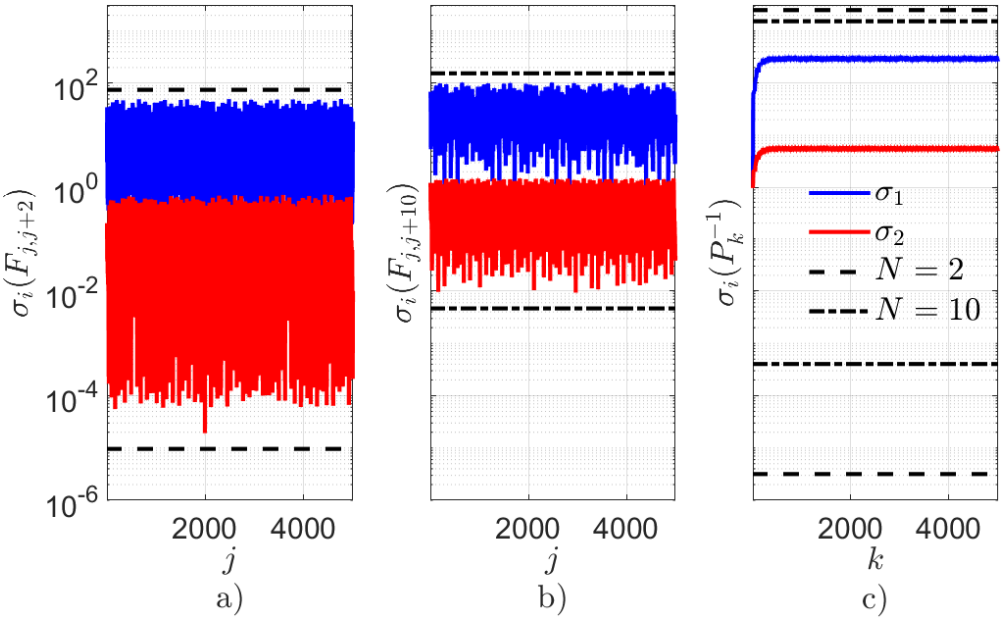



Figure 4.1: Example 4.2. Persistent excitation and bounds on $P_k^{-1}$. a) and b) show the singular values of $F_{j,j+N}$ for $N = 2$ and $N = 10$, where $\alpha$ and $\beta$ are chosen to satisfy (4.22). c) shows the singular values of $P_k^{-1}$, with corresponding bounds given by (4.28) for $\lambda = 0.99$.

**Example 4.3. Lack of persistent excitation and bounds on $P_k^{-1}$.** Let $\phi_k = [u_k \;\; u_{k-1}]$, where $u_k$ is given by (4.30) for all $k < 2500$ and $u_k = 1$ for all $k \geq 2500$. For $j = 1, \ldots, 5000$, Figure 4.2a) shows the singular values of $F_{j,j+2}$. Note that the

smaller singular value of $F_{j,j+2}$ reaches zero in machine precision, and thus there does not exist $\alpha > 0$ satisfying (4.22). Hence, $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. Figures 4.2b) and c) show the singular values of $P_k^{-1}$ for $\lambda = 1$ and $\lambda = 0.9$, respectively. Note that, if $\lambda = 1$, then one of the singular values of $P_k^{-1}$ diverges, whereas, if $\lambda \in (0, 1)$, then one of singular values of $P_k^{-1}$ converges to zero. ◇

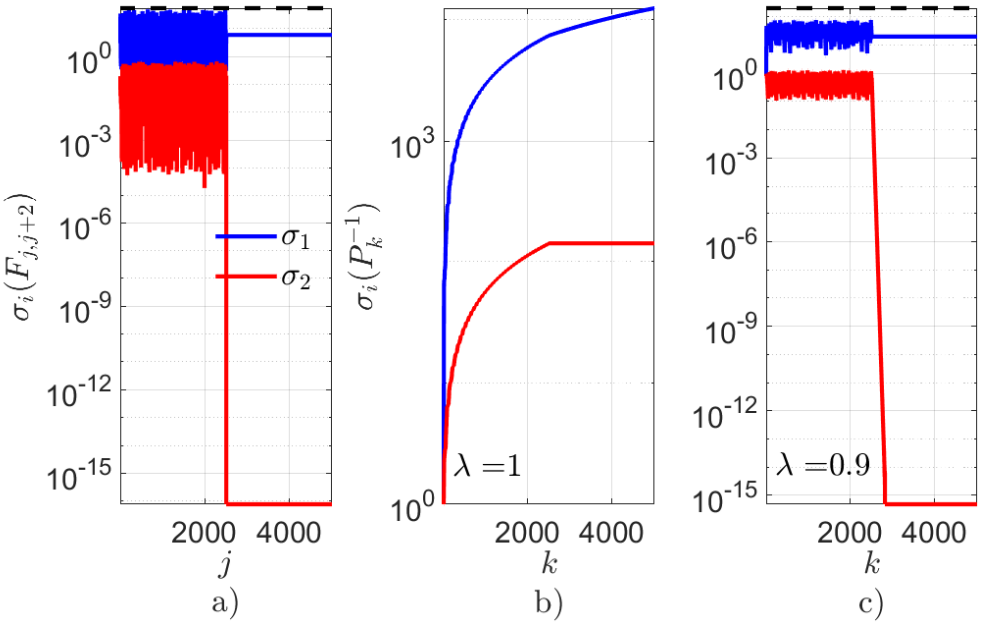



Figure 4.2: Example 4.3. Lack of persistent excitation and bounds on $P_k^{-1}$. a) shows the singular values of $F_{j,j+2}$. Note that the smaller singular value of $F_{j,j+2}$ reaches zero in machine precision, and thus that $\alpha > 0$ satisfying (4.22) does not exist. Hence, $\phi_k$ is not persistently exciting. The upper bound $\beta$ shown by the dashed line is chosen to satisfy (4.22). b) and c) show the singular values of $P_k^{-1}$ for $\lambda = 1$ and $\lambda = 0.9$, respectively. Note that, if $\lambda = 1$, then one of the singular values of $P_k^{-1}$ diverges, whereas, if $\lambda \in (0, 1)$, then one of singular values of $P_k^{-1}$ converges to zero.

Table 4.3 summarizes the results in this section.

| Excitation \ $\lambda$ | $\lambda = 1$ | $\lambda \in (0, 1)$ |
|---|---|---|
| Persistent | • $P_k$ converges to zero<br>• Proposition 4.3<br>• Example 4.2 | • $P_k$ remains bounded<br>• Propositions 4.4, 4.5<br>• Example 4.2 |
| Not Persistent | • Some singular values of $P_k$ remain bounded<br>• Other singular values of $P_k$ converge to zero<br>• Example 4.3 | • Some singular values of $P_k$ diverge<br>• Other singular values of $P_k$ remain bounded<br>• Example 4.3 |

Table 4.3: Behavior of $P_k$ under persistent and not persistent excitation.

## 4.3 Persistent Excitation and the Condition Number

For nonsingular $A \in \mathbb{R}^{n\times n}$, the condition number of $A$ is defined by

$$\kappa(A) \triangleq \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)}, \tag{4.31}$$

For $B \in \mathbb{R}^{n\times m}$, let $\|B\|$ denotes the maximum singular value of $B$. If $A$ is positive definite, then

$$\|A^{-1}\|^{-1} I_n = \sigma_{\min}(A) I_n \leq A \leq \sigma_{\max}(A) I_n = \|A\| I_n. \tag{4.32}$$

Therefore, if $\alpha, \beta \in (0, \infty)$ satisfy $\alpha \leq \sigma_{\min}(A)$ and $\sigma_{\max}(A) \leq \beta$, then $\kappa(A) \leq \dfrac{\beta}{\alpha}$. Thus, if $\lambda = 1$ and $(\phi_k)_{k=0}^{\infty}$ is persistently exciting with $N, \alpha, \beta$ given by Definition 4.1, then it follows from (4.24) that

$$\kappa(P_k) \leq \frac{\beta}{\alpha}. \tag{4.33}$$

Similarly, if $\lambda \in (0,1)$ and $(\phi_k)_{k=0}^{\infty}$ is persistently exciting with $N, \alpha, \beta$ given by Definition 4.1, then it follows from (4.28) that

$$\kappa(P_k) \leq \frac{\beta + (1-\lambda^{N+1})\|P_N^{-1}\|}{\lambda^N(1-\lambda)\alpha}. \tag{4.34}$$

However, as shown by Example 4.3, if $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting, then there may not exist $\alpha > 0$ satisfying (4.22), and thus $\kappa(P_k)$ cannot be bounded. Hence $\kappa(P_k)$ can be used to determine whether or not the $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, where a bounded condition number implies that $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, whereas a diverging condition number implies that $\phi_k$ is not persistently exciting, as illustrated by the following example. [88] provides a recursive algorithm for computing $\kappa(P_k)$.

**Example 4.4. Using the condition number of $P_k$ to determine whether $(\phi_k)_{k=0}^{\infty}$ is persistently exciting.** Consider the 5th-order IIR system

$$y_k = \frac{0.68\mathbf{q}^4 - 0.16\mathbf{q}^3 - 0.12\mathbf{q}^2 - 0.18\mathbf{q} + 0.09}{\mathbf{q}^5 - \mathbf{q}^4 + 0.41\mathbf{q}^3 - 0.17\mathbf{q}^2 - 0.03\mathbf{q} + 0.01}u_k, \tag{4.35}$$

where $\mathbf{q}$ is the forward-shift operator, and $u_k$ is given by (4.30). To apply RLS, let $\theta$ consist of the coefficients in (4.35) and

$$\phi_k = [u_{k-1} \ \cdots \ u_{k-5} \ y_{k-1} \ \cdots \ y_{k-5}], \tag{4.36}$$

Note that the regressor (4.36) corresponds to a 5th-order IIR parameterization of a linear input-output model. Let $P_0 = I_{10}$. Figure 4.3a) shows the singular values of $F_{j,j+20}$, where the singular values of $F_{j,j+20}$ close to machine precision ($\sim 10^{-15}$) are essentially zero. Thus, Definition 4.1 implies that $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting.

Figures 4.3b) and c) show the singular values and the condition number of $P_k$ for $\lambda = 1$. Note that the six singular values of $P_k$ decrease due to the presence of three harmonics in $u_k$. Note that $\kappa(P_k)$ is diverging due to the lack of persistent excitation.

Figures 4.3d) and e) show the singular values and the condition number of $P_k$ for $\lambda = 0.99$. Note that the six singular values of $P_k$ remain bounded due to the presence of three harmonics in $u_k$. However, $P_k$ becomes ill-conditioned due to the lack of persistent excitation. Consequently, numerical computation with $P_k$ becomes erroneous once $P_k$ becomes ill-conditioned. $\diamond$

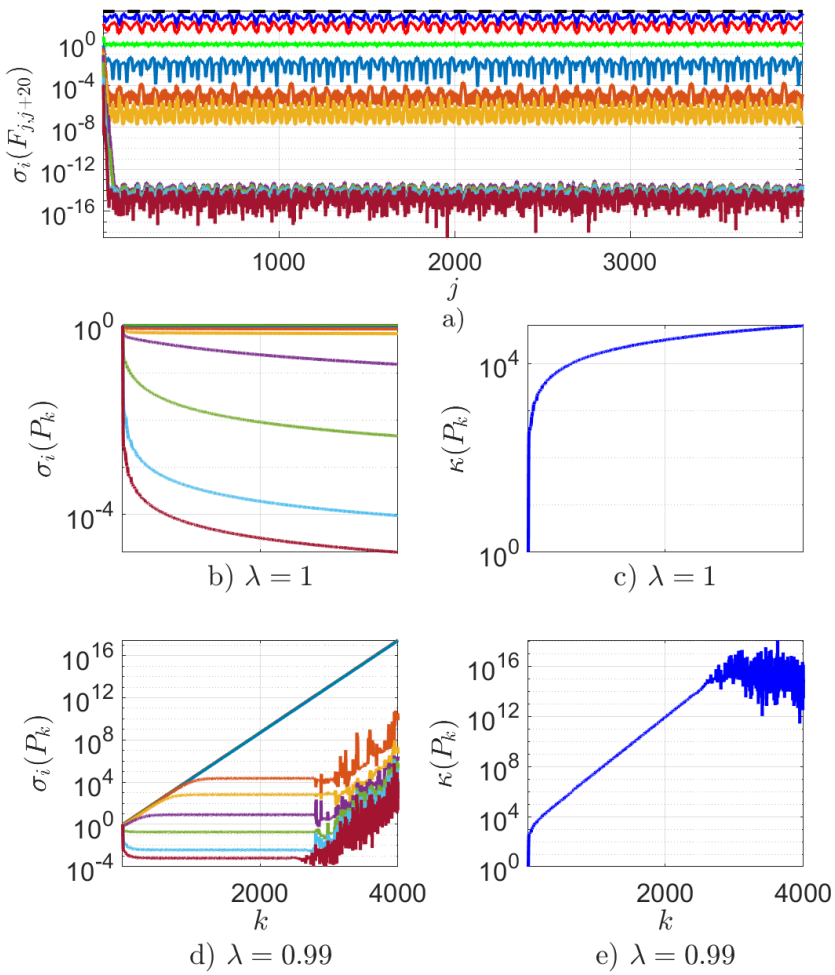



Figure 4.3: Example 4.4. Using the condition number of $P_k$ to evaluate persistency. a) shows the singular values of $F_{j,j+20}$, where the singular values of $F_{j,j+20}$ close to machine precision ($\sim 10^{-15}$) are essentially zero, thus implying that $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. b) and c) shows the singular values and the condition number of $P_k$ for $\lambda = 1$. Note that the six singular values of $P_k$ decrease due to the presence of three harmonics in $u_k$. d) and e) shows the singular values and the condition number of $P_k$ for $\lambda = 0.99$. Note that the six singular values of $P_k$ remain bounded due to the presence of three harmonics in $u_k$. However, $P_k$ becomes ill-conditioned due to the lack of persistent excitation.

Example 4.4 shows that if $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting and $\lambda = 1$, then some singular values of $P_k$ converge while the rest decrease to zero. On the other hand, if $(\phi_k)_{k=0}^{\infty}$ is not persistent and $\lambda \in (0, 1)$, then some singular values of $P_k$ diverge while the rest remain bounded. In particular, let $(\phi_k)_{k=0}^{\infty}$ be not persistent such that, for $\lambda = 1$, $n_1$ singular values of $P_k$ converge to zero, while $n - n_1$ remain bounded. Then, for $\lambda \in (0, 1)$, $n_1$ singular values remain bounded, while $n - n_1$ singular values diverge.

## 4.4 Lyapunov Analysis of the Parameter Error

Let $k \geq 0$, and consider the system

$$x_{k+1} = f(k, x_k), \tag{4.37}$$

where $x_k \in \mathbb{R}^n$, $f\colon \{0, 1, 2, \ldots\} \times \mathbb{R}^n \to \mathbb{R}^n$ is continuous, and, for all $k \geq 0$, $f(k, 0) = 0$.

**Definition 4.2.** The zero solution $x_k \equiv 0$ of (4.37) is *uniformly Lyapunov stable* if, for all $\varepsilon > 0$, there exists $\delta \in (0, \varepsilon]$ such that, for all $k_0 \geq 0$ and all $x_{k_0} \in \mathbb{R}^n$ such that $\|x_{k_0}\| < \delta$, it follows that, for all $k \geq k_0$, $\|x_k\| < \varepsilon$.

**Definition 4.3.** The zero solution $x_k \equiv 0$ of (4.37) is *uniformly globally aymptotically stable* if it is uniformly Lyapunov stable, and for all $x_0 \in \mathbb{R}^n$, if follows that $\lim_{k\to\infty} x_k = 0$.

**Definition 4.4.** The zero solution $x_k \equiv 0$ of (4.37) is *uniformly globally geometrically stable* if there exist $\alpha_0 > 0$ and $\beta_0 > 1$ such that, for all $x_0 \in \mathbb{R}^n$, it follows that, for all $k \geq 0$,

$$\|x_k\| \leq \alpha_0 \|x_0\| \beta_0^{-k}. \tag{4.38}$$

Note that, if the zero solution $x_k \equiv 0$ of (4.37) is uniformly globally geometrically stable, then it is uniformly globally aymptotically stable as well as uniformly Lyapunov stable.

The following result is a specialization of Theorem 13.11 given in [89, pp. 784, 785].

**Theorem 4.2.** *Consider* (4.37)*, and assume there exist a continuous function* $V : \mathbb{R}^n \to \mathbb{R}$ *and* $\alpha_1, \beta_1 > 0$ *such that, for all* $k \geq 0$*,*

$$\alpha_1 \|x\|^2 \leq V(k, x) \leq \beta_1 \|x\|^2, \tag{4.39}$$

$$V(k+1, f(k, x)) - V(k, x) \leq 0. \tag{4.40}$$

*Then the zero solution* $x_k \equiv 0$ *of* (4.37) *is uniformly Lyapunov stable. Furthermore, assume there exists* $\gamma_1 > 0$ *such that, for all* $k \geq 0$ *and* $x \in \mathbb{R}^n$*,*

$$V(k+1, f(k, x)) - V(k, x) \leq -\gamma_1 \|x\|^2. \tag{4.41}$$

*Then the zero solution* $x_k \equiv 0$ *of* (4.37) *is uniformly globally geometrically stable.*

The following result uses Lyapunov theory to show that, if $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, then the RLS estimate $\theta_k$ with $\lambda \in (0, 1)$ converges to $\theta$ in the sense of Definition 4.4. A related result is given in [87].

**Theorem 4.3.** *Assume that* $(\phi_k)_{k=0}^{\infty}$ *is persistently exciting, let* $N, \alpha, \beta$ *be given by Definition 4.1, let* $R \in \mathbb{R}^{n \times n}$ *be positive definite, define* $P_0 \stackrel{\triangle}{=} R^{-1}$*, let* $\lambda \in (0, 1]$*, and, for all* $k \geq 0$*, let* $P_k$ *be given by* (4.4)*. Then the zero solution of* (4.10) *is Lyapunov stable. In addition, if* $\lambda \in (0, 1)$*, then the zero solution of* (4.10) *is uniformly globally geometrically stable.*

*Proof.* Define the Lyapunov candidate

$$V(k,x) \stackrel{\triangle}{=} x^{\mathrm{T}} P_k^{-1} x,$$

where $x \in \mathbb{R}^n$. It follows from Proposition 4.4 that, for all $k \geq N+1$,

$$\begin{aligned}\frac{\lambda^N(1-\lambda)\alpha}{1-\lambda^{N+1}}\|x\|^2 \leq V(k,x) &\leq \frac{\beta}{1-\lambda^{N+1}}\|x\|^2 + x^{\mathrm{T}} P_N^{-1} x \\ &\leq \left(\frac{\beta}{1-\lambda^{N+1}} + \|P_N^{-1}\|\right)\|x\|^2,\end{aligned}$$

which confirms (4.39) for all $\lambda \in (0,1]$ with $\alpha_1 = \frac{\lambda^N(1-\lambda)\alpha}{1-\lambda^{N+1}}$, and $\beta_1 = \frac{\beta}{1-\lambda^{N+1}} + \|P_N^{-1}\|$. Next, defining

$$f(k,x) \stackrel{\triangle}{=} (I_n - P_{k+1}\phi_k^{\mathrm{T}}\phi_k)x,$$

it follows that

$$\begin{aligned}V(k+1, f(k,x)) - V(k,x) &= f(k,x)^{\mathrm{T}} P_{k+1}^{-1} f(k,x) - x^{\mathrm{T}} P_k^{-1} x \\ &= x^{\mathrm{T}}[(I_n - \phi_k^{\mathrm{T}}\phi_k P_{k+1})P_{k+1}^{-1}(I_n - P_{k+1}\phi_k^{\mathrm{T}}\phi_k) - P_k^{-1}]x \\ &= x^{\mathrm{T}}[(P_{k+1}^{-1} - \phi_k^{\mathrm{T}}\phi_k)(I_n - P_{k+1}\phi_k^{\mathrm{T}}\phi_k) - P_k^{-1}]x \\ &= x^{\mathrm{T}}[(\lambda-1)P_k^{-1} - \lambda P_k^{-1} P_{k+1}\phi_k^{\mathrm{T}}\phi_k]x \\ &= x^{\mathrm{T}}[(\lambda-1)P_k^{-1} - \lambda\phi_k^{\mathrm{T}}(\lambda I_p - \phi_k P_k \phi_k^{\mathrm{T}})^{-1}\phi_k]x. \\ &\leq 0,\end{aligned}$$

which confirms (4.40). It thus follows from Theorem 4.2 that the zero solution of (4.10) is uniformly Lyapunov stable.

Furthermore, if $\lambda \in (0,1)$, then

$$V(k+1, f(k,x)) - V(k,x) \leq (\lambda - 1)x^{\mathrm{T}} P_k^{-1} x$$
$$\leq (\lambda - 1)\left(\frac{\beta}{1-\lambda^{N+1}} + \|P_N^{-1}\|\right)\|x\|^2,$$

which confirms (4.41) with $\gamma_1 = (1-\lambda)(\frac{\beta}{1-\lambda^{N+1}} + \|P_N^{-1}\|)$. It thus follows from Theorem 4.2 that the zero solution of (4.10) is uniformly globally geometrically stable. □

The following result provides an alternative proof of Theorem 4.3 that does not depend on Theorem 4.2. In addition, this result considers the case $\lambda = 1$, where the RLS estimate $\theta_k$ converges to $\theta$ in the sense of Definition 4.3.

**Theorem 4.4.** *Assume that $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, let $N, \alpha, \beta$ be given by Definition 4.1, let $R \in \mathbb{R}^{n\times n}$ be positive definite, define $P_0 \triangleq R^{-1}$, let $\lambda \in (0,1]$, and, for all $k \geq 0$, let $P_k$ be given by (4.4). Then the zero solution of (4.10) is uniformly globally asymptotically stable. Furthermore, if $\lambda \in (0,1)$, then the zero solution of (4.10) is uniformly globally geometrically stable.*

*Proof.* Let $k_0 \geq 0$ and $\tilde{\theta}_{k_0} \in \mathbb{R}^n$. Then, it follows from (4.12) that, for all $k \geq k_0$,

$$\|\tilde{\theta}_k\| = \lambda^{k-k_0}\|P_k P_{k_0}^{-1}\tilde{\theta}_{k_0}\|$$
$$\leq \|P_k P_{k_0}^{-1}\tilde{\theta}_{k_0}\|$$
$$\leq \|P_k\|\|P_{k_0}^{-1}\|\|\tilde{\theta}_{k_0}\|. \tag{4.42}$$

First, consider the case where $\lambda = 1$. Let $\delta > 0$, and suppose that $\tilde{\theta}_{k_0} \in \mathbb{R}^n$ satisfies $\|\tilde{\theta}_{k_0}\| \leq \delta$. Now note that it follows from (4.6) with $\lambda = 1$ that $P_k \leq P_{k-1} \leq \cdots \leq P_{k_0}$. Hence, for all $k \geq k_0$, $\|P_k\| \leq \|P_{k_0}\|$, and thus $\|P_k\|\|P_{k_0}^{-1}\| = \|P_k\|/\|P_{k_0}\| \leq 1$. It thus follows from (4.42) that, for all $k \geq k_0$, $\|\tilde{\theta}_k\| \leq \|\tilde{\theta}_{k_0}\| < \delta$.It thus follows from Definition 4.2 with $\varepsilon = \delta$ that the zero solution of (4.10) is uniformly Lyapunov stable.

Next, let $\tilde{\theta}_0 \in \mathbb{R}^n$. Then, Proposition 4.3 implies that

$$\lim_{k\to\infty} \tilde{\theta}_k = \lim_{k\to\infty} P_k P_0^{-1} \tilde{\theta}_0 = 0.$$

It thus follows from Definition 4.3 that the zero solution of (4.10) is uniformly globally asymptotically stable.

Next, consider the case where $\lambda \in (0,1)$. Let $k_0 \geq 0$ and $\delta > 0$, and let $\tilde{\theta}_{k_0} \in \mathbb{R}^n$ satisfy $\|\tilde{\theta}_{k_0}\| \leq \delta$. It follows from Proposition 4.4 and (4.42) that, for all $k \geq \max(N+1, k_0)$,

$$\|\tilde{\theta}_k\| \leq \varepsilon,$$

where

$$\varepsilon \triangleq \frac{\beta + (1-\lambda^{N+1})\|P_N^{-1}\|}{\lambda^N(1-\lambda)\alpha}\delta.$$

It thus follows from Definition 4.2 that the zero solution of (4.10) is uniformly Lyapunov stable.

Next, let $\tilde{\theta}_0 \in \mathbb{R}^n$. Then, it follows from (4.12) and Proposition 4.4 that, for all $\tilde{\theta}_0 \in \mathbb{R}^n$, and $k \geq N+1$,

$$\|\tilde{\theta}_k\| \leq \alpha_0 \|\tilde{\theta}_0\| \beta_0^{-k},$$

where $\beta_0 \triangleq 1/\lambda$ and

$$\alpha_0 \triangleq \frac{1-\lambda^{N+1}}{\lambda^N(1-\lambda)\alpha}\|P_0^{-1}\|.$$

Thus, it follows from Definition 4.4 that the zero solution of (4.10) is uniformly globally geometrically stable, and thus uniformly globally asymptotically stable. □

The following result shows that persistent excitation produces an infinite sequence of matrices whose product converges to zero.

**Proposition 4.6.** *Let $P_0 \in \mathbb{R}^{n\times n}$ be positive definite, let $\lambda \in (0,1]$, and, for all $k \geq 0$, let $P_k$ be given by (4.4). Then, for all $k \geq 0$, all of the eigenvalues of $P_{k+1}\phi_k^{\mathrm{T}}\phi_k$ are contained in $[0,1]$. If, in addition, $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, then*

$$\lim_{k\to\infty} \mathcal{A}_k = 0, \tag{4.43}$$

*where*

$$\mathcal{A}_k \triangleq (I_n - P_{k+1}\phi_k^{\mathrm{T}}\phi_k)\cdots(I_n - P_1\phi_0^{\mathrm{T}}\phi_0). \tag{4.44}$$

*Proof.* It follows from (4.6) that, for all $k \geq 0$, $\phi_k^{\mathrm{T}}\phi_k \leq P_{k+1}^{-1}$, and thus, for all $k \geq 0$, $P_{k+1}^{1/2}\phi_k^{\mathrm{T}}\phi_k P_{k+1}^{1/2} \leq I_n$. Hence, for all $k \geq 0$,

$$0 \leq \lambda_{\max}(P_{k+1}\phi_k^{\mathrm{T}}\phi_k) = \lambda_{\max}(P_{k+1}^{1/2}\phi_k^{\mathrm{T}}\phi_k P_{k+1}^{1/2}) \leq 1.$$

To prove (4.43), suppose that $(\phi_k)_{k=0}^{\infty}$ is persistently exciting and, for all $i \in \{1,\ldots,n\}$, define $\theta_0 \triangleq e_i + \theta$, where $e_i$ is the $i$th column of $I_n$. Note that, for all $i \in \{1,\ldots,n\}$, $\tilde{\theta}_0 \triangleq \theta_0 - \theta = e_i$. Then, (4.12) implies that, for all $k \geq 0$ and all $i \in \{1,\ldots,n\}$,

$$\tilde{\theta}_{k+1} = \mathcal{A}_k e_i = \lambda^{k+1} P_{k+1} P_0^{-1} e_i.$$

It follows from Proposition 4.4 that, for all $i \in \{1,\ldots,n\}$, $\tilde{\theta}_k$ converges to 0. Hence, for all $i \in \{1,\ldots,n\}$, the $i$th column of $\mathcal{A}_k$ converges to zero as $k \to \infty$, which implies (4.43). □

It follows from Proposition 4.4 that, if $(\phi_k)_{k=0}^{\infty}$ is persistently exciting, then, for all $\lambda \in (0,1]$, $\tilde{\theta}_k$ converges to zero. In addition, if $\lambda \in (0,1)$, then $\tilde{\theta}_k$ converges to zero

geometrically, that is, the rate of convergence of $\|\tilde{\theta}_k\|$ is $O(\lambda^k)$. However, in the case $\lambda = 1$, as shown in [87] and the next example, $\tilde{\theta}_k$ converges to zero as $O(1/k)$, and thus the convergence is not geometric.

**Example 4.5. Effect of $\lambda$ on the rate of convergence of $\theta_k$.** Consider the 3rd-order FIR system

$$y_k = \frac{\mathbf{q}^2 + 0.8\mathbf{q} + 0.5}{\mathbf{q}^3} u_k, \tag{4.45}$$

where $\mathbf{q}$ is the forward-shift operator. To apply RLS, let $\theta = [1\ 0.8\ 0.5]$, $\theta_0 = 0$, and $\phi_k = [u_{k-1}\ u_{k-2}\ u_{k-3}]$, where the input $u_k$ is zero-mean Gaussian white noise with unity standard deviation. Figures 4.4a) - f) show the parameter error norm $\|\tilde{\theta}_k\|$ for several values of $P_0$ and $\lambda$. Note that the convergence rate of $\|\tilde{\theta}_k\|$ is $O(1/k)$ for $\lambda = 1$ and geometric for all $\lambda \in (0, 1)$. Figures 4.4g), h), and i) show the condition number of the corresponding $P_k$ for several values of $P_0$ and $\lambda$. Note that, as $\lambda$ is decreased, the convergence rate of $\theta_k$ increases; however, the condition number of $P_k$ degrades, and the effect of $P_0$ is reduced. ◇

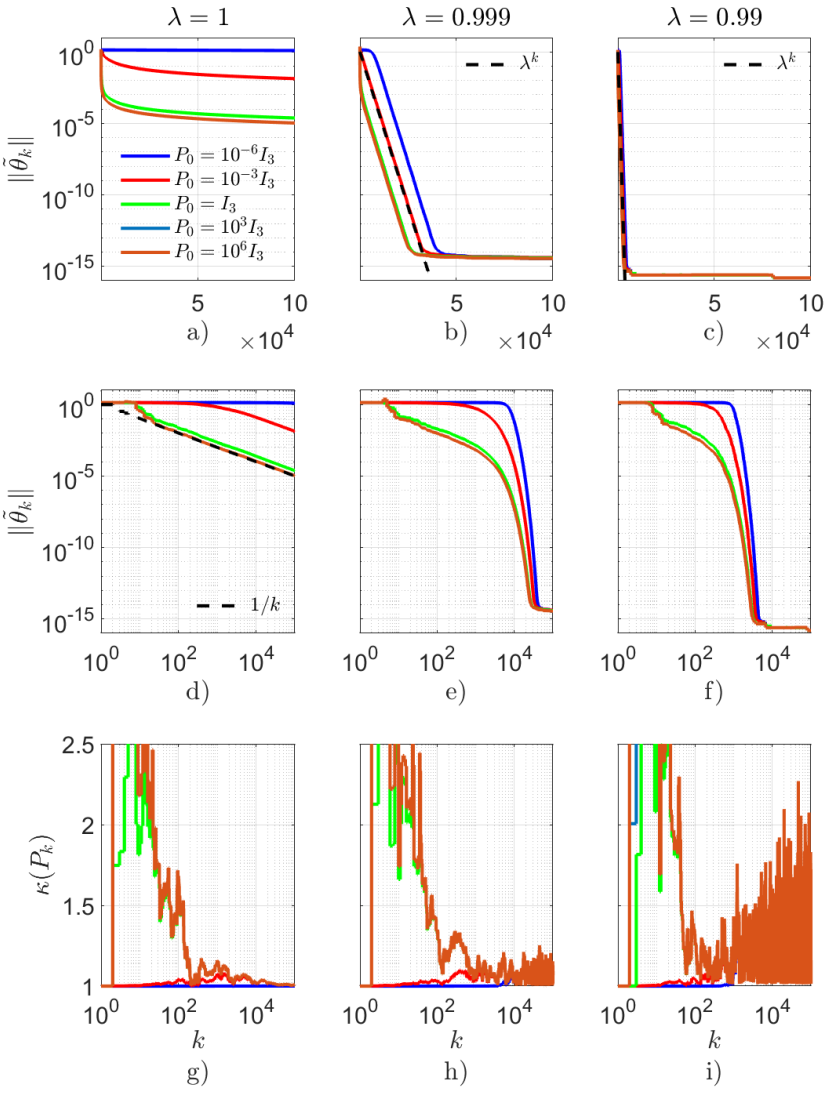



Figure 4.4: Example 4.5. Effect of $\lambda$ on the rate of convergence of $\theta_k$. a)-f) show the parameter error norm $\|\tilde{\theta}_k\|$ for several values of $P_0$ and $\lambda$. Note that the slope of $-1$ between $\log \|\tilde{\theta}_k\|$ and $\log k$ in d) is consistent with the fact that the rate of convergence of $\|\tilde{\theta}_k\|$ is $O(1/k)$ for $\lambda = 1$. Similarly, the slope of $\log \lambda$ between $\log \|\tilde{\theta}_k\|$ and $k$ in b) and c) is consistent with the fact that the rate of convergence of $\|\tilde{\theta}_k\|$ is $O(\lambda^k)$ for $\lambda \in (0, 1)$. g), h), and i) show the condition number of the corresponding $P_k$ for several values of $P_0$ and $\lambda$. Note that, as $\lambda$ is decreased, the convergence rate of $\theta_k$ increases; however, the condition number of $P_k$ degrades, and the effect of $P_0$ is reduced.

## 4.5 Lack of Persistent Excitation

In this section, we investigate the effect of lack of persistent excitation on $P_k$ and $\theta_k$ using numerical examples. Recall that, as shown in Example 4.3 and Example 4.4, if $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting and $\lambda = 1$, then some singular values of $P_k$ converge to zero, while the rest remain bounded. On the other hand, if $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting and $\lambda \in (0, 1)$, then some singular values of $P_k$ remain bounded, while the rest diverge. Furthermore, it follows from Proposition 4.1 that the predicted error $z_k$ and the parameter error $\tilde{\theta}_k$ converge to zero, irrespective of whether $(\phi_k)_{k=0}^{\infty}$ is persistent or not.

**Example 4.6. Lack of persistent excitation in scalar estimation**. Let $n = 1$, so that (4.4), (4.5) are given by

$$P_{k+1} = \frac{P_k}{\lambda + P_k \phi_k^2}, \tag{4.46}$$

$$\tilde{\theta}_{k+1} = \frac{\lambda \tilde{\theta}_k}{\lambda + P_k \phi_k^2}. \tag{4.47}$$

Now, let $k_0 \geq 0$ and assume that, for all $k \geq k_0$, $\phi_k = 0$. Therefore, for all $j \geq 0$ and $N \geq 1$, $F_{j,j+N}$ cannot be lower bounded as in (4.22), and thus $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. For $\lambda = 1$, $P_k$ and $\tilde{\theta}_k$ converge in $k_0$ steps to $\overline{P}$ and $\overline{\tilde{\theta}}$, respectively, where $\overline{P} \neq 0$ and, if $\theta_0 \neq \theta$, then $\overline{\tilde{\theta}} \neq 0$. However, for all $\lambda \in (0, 1)$, $P_k$ diverges geometrically and $\tilde{\theta}_k$ converges in $k_0$ steps.

Note that, for all $\lambda \in (0, 1]$, since $\phi_k = 0$ for all $k \geq k_0$, it follows from (4.46) and (4.47) that, for all $k \geq k_0$, the minimum of (4.2) is achieved in a finite number of steps. Consequently, RLS provides no further refinement of the estimate $\theta_k$ of $\theta$, and thus $\overline{\tilde{\theta}} \neq 0$ implies that $\theta_k$ does not converge to $\theta$.

Next, assume that, for all $k \geq 0$, $\phi_k = \overline{\phi}$, where $\overline{\phi} \neq 0$. Then Definition 4.1 is satisfied with $N = 1$, $\alpha = \overline{\phi}^2$, and $\beta = 3\overline{\phi}^2$, and thus $(\phi_k)_{k=0}^{\infty}$ is persistently

exciting. If $\lambda = 1$, then both $P_k$ and $\tilde{\theta}_k$ converge to zero. However, if $\lambda \in (0, 1)$, then $P_k$ converges to $\frac{1-\lambda}{\overline{\phi}^2}$ and $\tilde{\theta}_k$ converges geometrically to zero. Table 4.4 shows the asymptotic behavior of $\tilde{\theta}_k$ and $P_k$ for various cases. $\diamond$

| Excitation \ $\lambda$ | $\lambda = 1$ | $\lambda \in (0, 1)$ |
|---|---|---|
| Not persistently exciting | $\tilde{\theta}_k \to \overline{\tilde{\theta}}, P_k \to \overline{P}$ | $\tilde{\theta}_k \to \overline{\tilde{\theta}}, P_k$ diverges |
| Persistently exciting | $\tilde{\theta}_k \to 0, P_k \to 0$ | $\tilde{\theta}_k \to 0, P_k \to \frac{1-\lambda}{\overline{\phi}^2}$ |

Table 4.4: Asymptotic behavior of RLS in Example 4.6. In the case of persistent excitation with $\lambda < 1$, the convergence of $\tilde{\theta}_k$ is geometric.

**Example 4.7. Subspace constrained regressor.** Consider (4.1), where $\phi_k = (\sin \frac{2\pi k}{100})[1\ 1]$ and $\theta = [0.4\ 1.4]^{\mathrm{T}}$. To estimate $\theta$ using RLS, let $P_0 = I_2$ and $\theta_0 = 0$. Figures 4.5 and 4.6 show the estimate $\theta_k$ of $\theta$ with $\lambda = 1$ and $\lambda = 0.99$. Note that all regressors $\phi_k$ lie along the same one-dimensional subspace, and thus, $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. It follows from (4.19) that the estimate $\theta_k$ of $\theta$ lies in this subspace.

For $\lambda = 1$, note that one of the singular value decreases to zero, while the other singular value remains bounded. Note that $\tilde{\theta}_k$ converges along the singular vector corresponding to the bounded singular value. For $\lambda = 0.99$, one of the singular value remains bounded, while the other singular value diverges. Note that $\tilde{\theta}_k$ converges along the singular vector corresponding to the diverging singular value. $\diamond$

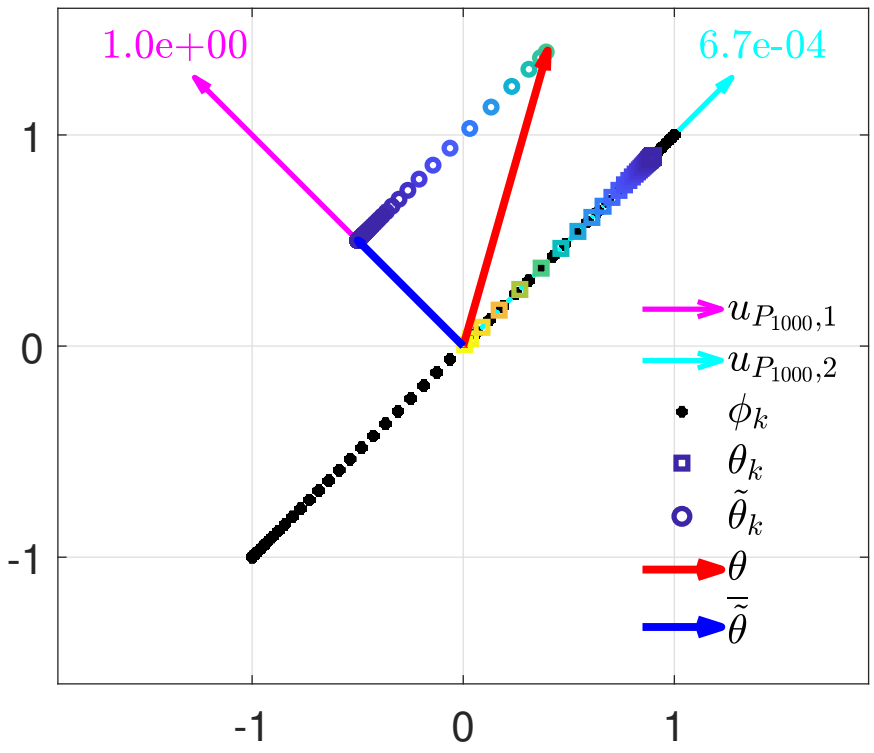



Figure 4.5: Example 4.7. Subspace constrained regressor. The first component of each vector is plotted along the horizontal axis, and the second component is plotted along the vertical axis. The singular values $\sigma_i(P_{1000})$ are shown with the corresponding singular vector $u_{P_{1000},i}$. All regressors $\phi_k$ lie along the same one-dimensional subspace, and thus, $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. Consequently, each estimate $\theta_k$ of $\theta$ lies in this subspace. The color gradient from yellow to blue of $\theta_k$ and $\tilde{\theta}_k$ shows the evolution from $k = 1$ to $k = 1000$. In a), the singular value corresponding to the cyan singular vector decreases to zero, while the singular value corresponding to the magenta singular vector remains bounded. Note that $\tilde{\theta}_k$ converges along the singular vector corresponding to the bounded singular value.

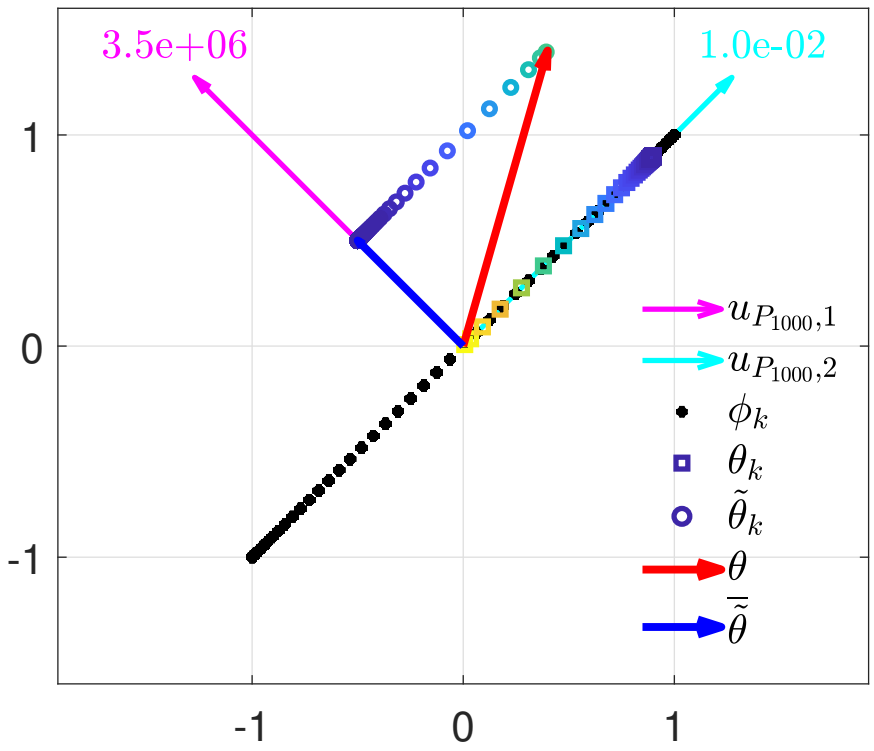



Figure 4.6: Example 4.7. Subspace constrained regressor. The first component of each vector is plotted along the horizontal axis, and the second component is plotted along the vertical axis. The singular values $\sigma_i(P_{1000})$ are shown with the corresponding singular vector $u_{P_{1000},i}$. All regressors $\phi_k$ lie along the same one-dimensional subspace, and thus, $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. Consequently, each estimate $\theta_k$ of $\theta$ lies in this subspace. The color gradient from yellow to blue of $\theta_k$ and $\tilde{\theta}_k$ shows the evolution from $k = 1$ to $k = 1000$. The singular value corresponding to the cyan singular vector remains bounded, while the singular value corresponding to the magenta singular vector diverges. Note that $\tilde{\theta}_k$ converges along the singular vector corresponding to the diverging singular value.

**Example 4.8. Lack of persistent excitation and finite-precision arithmetic.** Consider the problem of fitting a 5th-order IIR model to the measured input-output data from the system (4.35), where the input $u_k$ is given by (4.30). Note that $\phi_k$

is given by (4.36), and is not persistently exciting as shown in Example 4.4. Let $P_0 = I_{10}$, $\theta_0 = 0$, and $\lambda = 0.999$. Figure 4.7a) shows the predicted error $z_k$, and b) shows the norm of the parameter error $\tilde{\theta}_k$. Note that the $\tilde{\theta}_k$ does not converge to 0. c) shows the singular values of $P_k$, and d) shows the condition number of $P_k$. Note that six singular values of $P_k$ remain bounded due to the presence of three harmonics in the regresssor. Due to finite-precision arithmetic, computation becomes erroneous as $P_k$ becomes numerically ill-conditioned, and thus, the estimate $\theta_k$ diverges. $\diamond$

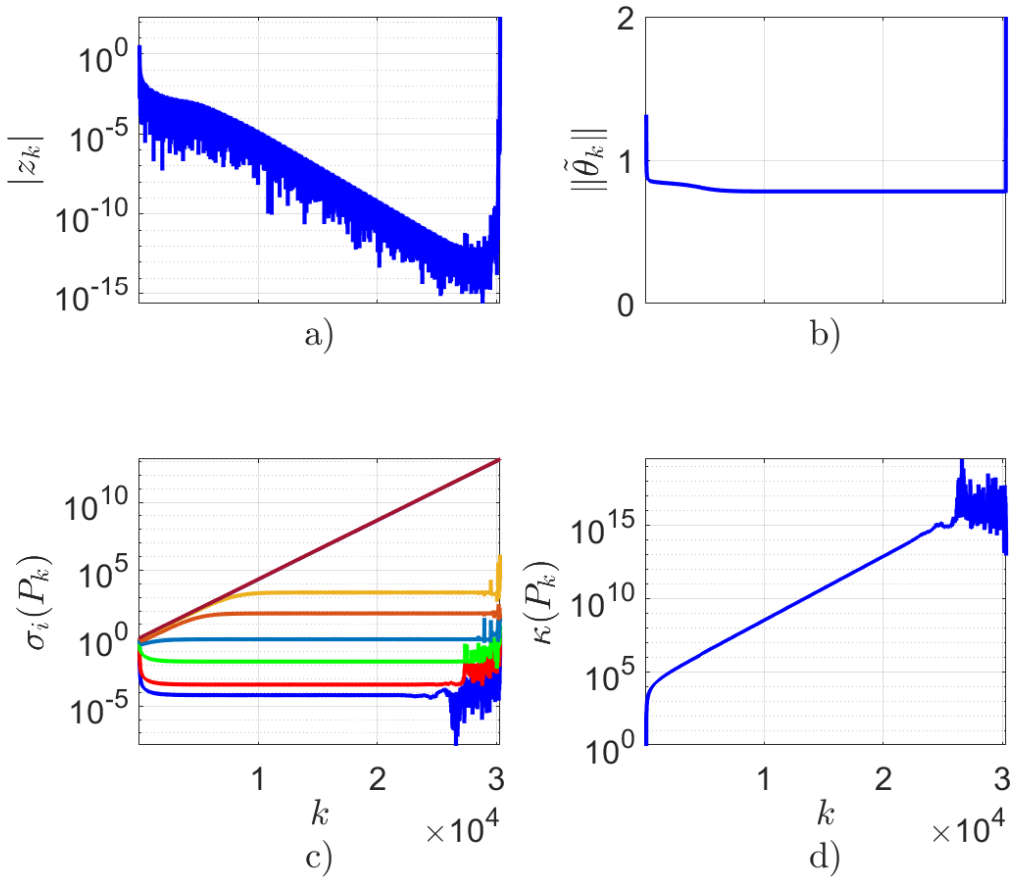



Figure 4.7: Example 4.8. Effect of lack of persistent excitation on $\theta_k$. a) shows the predicted error $z_k$, b) shows the norm of the parameter error $\tilde{\theta}_k$, c) shows the singular values of $P_k$, and d) shows the condition number of $P_k$. Note that six singular values of $P_k$ remain bounded due to the presence of three harmonics in the regresssor. Due to finite-precision arithmetic, computation becomes erroneous as $P_k$ becomes numerically ill-conditioned, and thus, the estimate $\theta_k$ diverges.

The numerical examples in this section show that, if $\lambda \in (0, 1]$ and $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting, then $\tilde{\theta}_k$ does not necessarily converge to zero. Furthermore, if $\lambda \in (0, 1)$ and $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting, then some singular values of $P_k$ di-

verge, and $\theta_k$ diverges due to finite-precision arithmetic when $P_k$ becomes numerically ill-conditioned.

## 4.6 Information Subspace

Using the singular value decomposition, (4.6) can be written as

$$P_{k+1}^{-1} = \lambda U_k \Sigma_k U_k^{\mathrm{T}} + U_k \psi_k^{\mathrm{T}} \psi_k U_k^{\mathrm{T}}, \tag{4.48}$$

where $U_k \in \mathbb{R}^{n \times n}$ is an orthonormal matrix whose columns are the singular vectors of $P_k^{-1}$, $\Sigma_k \in \mathbb{R}^{n \times n}$ is a diagonal matrix whose diagonal entries are the corresponding singular values, and

$$\psi_k \triangleq \phi_k U_k. \tag{4.49}$$

The columns of $U_k$ are the *information directions* at step $k$, and each row of $\psi_k$ is the projection of the corresponding row of $\phi_k$ onto the information directions. The magnitude of column of $\psi_k$ thus indicates the *information content* present in $\phi_k$ along the corresponding information direction. The smallest subspace that is spanned by a subset of the information directions and that contain all rows of $\phi_k$ is the *information subspace* $\mathcal{I}_k$ at step $k$. Figure 4.8 shows an illustrative example of the information subspace.

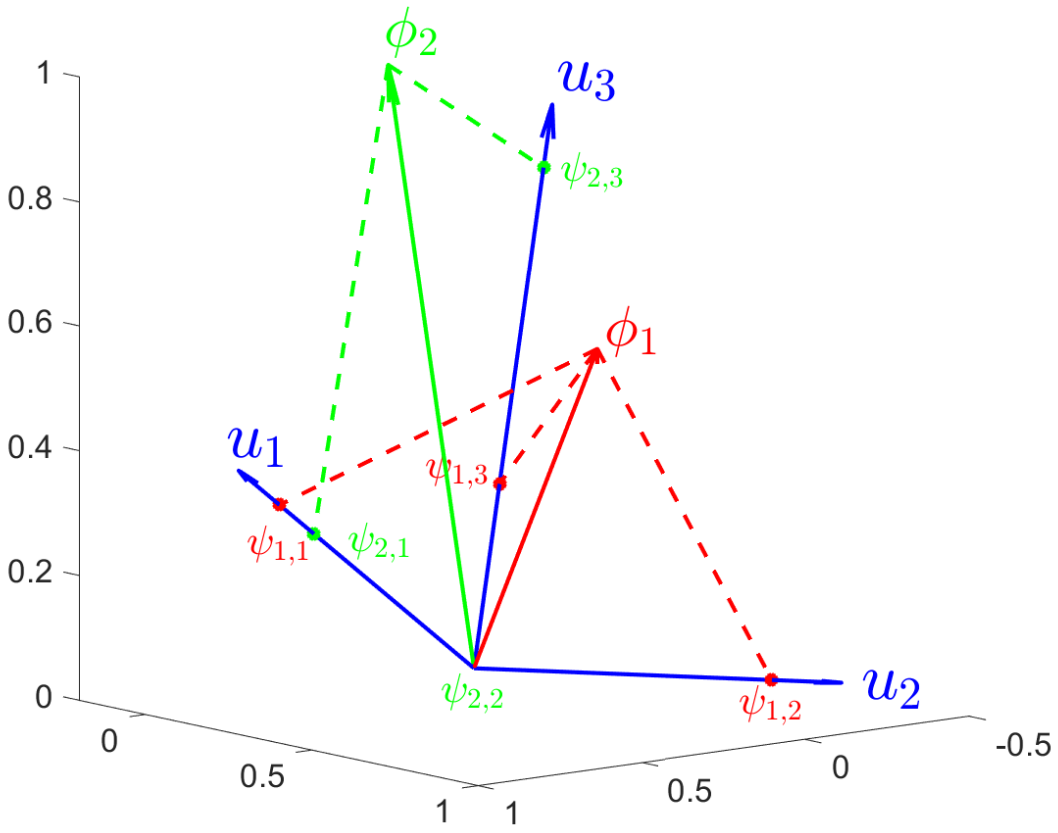



Figure 4.8: Illustrative example of the information subspace. Let $u_1, u_2$, and $u_3$ be the information directions (shown in blue). The regressor $\phi_1$ (shown in red) has new information along all three information directions, as shown by the nonzero values $\psi_{1,1}$, $\psi_{1,2}$, and $\psi_{1,3}$; the information subspace is thus $\mathcal{R}([u_1\ u_2\ u_3])$. On the other hand, the regressor $\phi_2$ (shown in green) has new information only along $u_1$ and $u_3$, as shown by the nonzero values $\psi_{2,1}$ and $\psi_{2,3}$; the information subspace is thus $\mathcal{R}([u_1\ u_3])$.

Now, consider the case where

$$\psi_k = \begin{bmatrix} \psi_{k,1} & 0_{p\times(n-n_1)} \end{bmatrix}, \tag{4.50}$$

where $\psi_{k,1} \in \mathbb{R}^{p\times n_1}$, which implies that $\phi_k$ provides new information along the first $n_1$ columns of $U_k$; these directions constitute the information subspace. It thus follows from (4.48) and (4.50) that $P_{k+1}^{-1}$ is given by

$$P_{k+1}^{-1} = U_k \begin{bmatrix} \lambda\Sigma_{k,1} + \psi_{k,1}^{\mathrm{T}}\psi_{k,1} & 0 \\ 0 & \lambda\Sigma_{k,2} \end{bmatrix} U_k^{\mathrm{T}}, \tag{4.51}$$

where $\Sigma_{k,1} \in \mathbb{R}^{n_1\times n_1}$ is the diagonal matrix whose diagonal entries are the first $n_1$ singular values of $P_k^{-1}$, and $\Sigma_{k,2}$ is the diagonal matrix whose diagonal entries are the remaining $n - n_1$ singular values of $P_k^{-1}$. In particular, writing

$$U_k = \begin{bmatrix} U_{k,1} & U_{k,2} \end{bmatrix}, \tag{4.52}$$

where $U_{k,1} \in \mathbb{R}^{n\times n_1}$ contains the first $n_1$ columns of $U_k$, and $U_{k,2} \in \mathbb{R}^{n\times n-n_1}$ contains the remaining $n - n_1$ columns of $U_k$, it follows that

$$P_{k+1}^{-1} = \begin{bmatrix} U_{k+1,1} & U_{k+1,2} \end{bmatrix} \begin{bmatrix} \Sigma_{k+1,1} & 0 \\ 0 & \Sigma_{k+1,2} \end{bmatrix} \begin{bmatrix} U_{k+1,1}^{\mathrm{T}} \\ U_{k+1,2}^{\mathrm{T}} \end{bmatrix}, \tag{4.53}$$

where

$$U_{k+1,1} = U_{k,1}V_k, \tag{4.54}$$

$$\Sigma_{k+1,1} = D_k, \tag{4.55}$$

$$U_{k+1,2} = U_{k,2}, \tag{4.56}$$

$$\Sigma_{k+1,2} = \lambda\Sigma_{k,2}, \tag{4.57}$$

where $V_k \in \mathbb{R}^{n_1 \times n_1}$ contains the singular vectors of $\lambda\Sigma_{k,1} + \psi_{k,1}^{\mathrm{T}}\psi_{k,1}$ and $D_k \in \mathbb{R}^{n_1 \times n_1}$ is the diagonal matrix containing the corresponding singular values. It follows from (4.56), (4.57) that if, for all $k \geq 0$, $\psi_k$ is given by (4.50) and $\lambda \in (0, 1)$, then the last $n - n_1$ singular vectors of $P_k^{-1}$ do not change and the corresponding singular values of $P_k^{-1}$ decrease to zero geometrically. It thus follows from Proposition 4.4 that $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. Furthermore, since $P_k$ and $P_k^{-1}$ have the same singular vectors and the singular values of $P_k$ are the inverse of the singular values of $P_k^{-1}$, the last $n - n_1$ singular values of $P_k$ diverge.

The following example shows that, if $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting, then the information subspace is a proper subset of $\mathbb{R}^n$ and the singular values of $P_k^{-1}$ corresponding to the singular vectors in the orthogonal complement of the information subspace converge to zero.

**Example 4.9. Lack of persistent excitation and the information subspace .** Consider the regressor $\phi_k$ given by (4.36) used in Example 4.4. Recall that $(\phi_k)_{k=0}^{\infty}$ is not persistently exciting. Let $P_0 = I_{10}$. Figures 4.9a), b), and c) show the information content $|\psi_{k,i}|$ for several values of $\lambda$. Note that the information subspace is six dimensional due to the presence of three harmonics in $u_k$ as shown by six relatively large components of $\psi_k$. Figures 4.9d), e), and f) show the singular values of $P_k^{-1}$ for several values of $\lambda$. Note that, for $\lambda < 1$, the singular values that correspond to the singular vectors not in the information subspace convergeto zero in machine precision. $\diamond$

## 4.7 Information-Driven Forgetting

Motivated by the fact that some singular values of $P_k^{-1}$ converge to zero if $\phi_k$ is not persistently exciting, as shown in Examples 4.3, 4.4, 4.6, 4.7, and 4.8, we propose the following modification. We modify (4.6) by replacing the scalar forgetting factor

$\lambda$ by a data-dependent forgetting matrix $\Lambda_k$. To do this, we redefine $P_{k+1}^{-1}$ as

$$P_{k+1}^{-1} = \Lambda_k P_k^{-1} \Lambda_k + \phi_k^{\mathrm{T}} \phi_k, \tag{4.58}$$

where $\Lambda_k$ is a positive-definite (and thus symmetric) matrix constructed below. Note that, for all $k \geq 0$, $P_{k+1}^{-1}$ given by (4.58) is positive-definite. Using the singular value decomposition, (4.58) can be written as

$$P_{k+1}^{-1} = \Lambda_k U_k \Sigma_k U_k^{\mathrm{T}} \Lambda_k + U_k \psi_k^{\mathrm{T}} \psi_k U_k^{\mathrm{T}}, \tag{4.59}$$

where $U_k$, $\Sigma_k$, and $\psi_k$ are as defined in the previous section.

The objective is to apply forgetting to only those singular values of $P_k^{-1}$ that correspond to the singular vectors in the information subspace, that is, forgetting is restricted to the subspace of $P_k^{-1}$ where sufficient new information is provided by $\phi_k$. Specifically, forgetting is applied to those information directions where the information content is greater than $\varepsilon > 0$, where $\varepsilon$ should be selected to be larger than the noise to signal ratio or larger than the machine zero, if no noise is present. To do so, we write (4.59) as

$$P_{k+1}^{-1} = U_k \overline{\Lambda}_k \Sigma_k \overline{\Lambda}_k U_k^{\mathrm{T}} + U_k \psi_k^{\mathrm{T}} \psi_k U_k^{\mathrm{T}}, \tag{4.60}$$

where $\overline{\Lambda}_k$ is a diagonal matrix whose diagonal entries are either $\sqrt{\lambda}$ or 1. In particular,

$$\overline{\Lambda}_k(i,i) \triangleq \begin{cases} \sqrt{\lambda}, & \|\psi_{k,i}\| > \varepsilon, \\ 1, & \text{otherwise}, \end{cases} \tag{4.61}$$

where $\psi_{k,i}$ is the $i$th column of $\psi_k$ and $\lambda \in (0,1]$. Next, it follows from (4.59), (4.60)

that

$$\Lambda_k = U_k \overline{\Lambda}_k U_k^{\mathrm{T}}, \tag{4.62}$$

which is positive definite.

Using $U_k$ obtained by the singular value decomposition of $P_k$, the information content $\psi_k$ given by (4.49), $\overline{\Lambda}_k$ is given by (4.61), and applying the matrix inversion lemma to (4.58), the recursive update of $P_k$ is given by

$$\Lambda_k^{-1} = U_k \overline{\Lambda}_k^{-1} U_k^{\mathrm{T}}, \tag{4.63}$$

$$\overline{P}_k = \Lambda_k^{-1} P_k \Lambda_k^{-1}, , \tag{4.64}$$

$$P_{k+1} = \overline{P}_k - \overline{P}_k \phi_k (I_p + \phi_k^{\mathrm{T}} \overline{P}_k \phi_k)^{-1} \phi_k^{\mathrm{T}} \overline{P}_k. \tag{4.65}$$

Next, we show that the forgetting scheme (4.63) in (4.65) prevents the singular values of $P_k$ from diverging. Consider the case where, for all $k$,

$$\psi_k = \left[ \begin{array}{cc} \psi_{k,1} & 0 \end{array} \right], \tag{4.66}$$

where $\psi_{k,1} \in \mathbb{R}^{p \times n_1}$, that is, the information subspace is spanned by the first columns of $U_k$. It thus follows from (4.60) and (4.66) that $P_{k+1}^{-1}$ is given by

$$P_{k+1}^{-1} = U_k \left[ \begin{array}{cc} \lambda \Sigma_{k,1} + \psi_{k,1}^{\mathrm{T}} \psi_{k,1} & 0 \\ 0 & \Sigma_{k,2} \end{array} \right] U_k^{\mathrm{T}}. \tag{4.67}$$

It follows from the $(2,2)$ block of (4.67) that the last $n - n_1$ information directions and the corresponding singular values are not affected by $\phi_k$. Furthermore, if $n_1 = n$, that is, new information is present in $\phi_k$ along every information direction, then forgetting is applied to all of the singular values of $P_k^{-1}$, and thus information-driven directional forgetting specializes to spatially uniform forgetting, that is, RLS with the update for

$P_k$ given by (4.6).

**Example 4.10. Information-driven forgetting for a regressor lacking persistent excitation.** Reconsider Example 4.9. Let $P_0 = I_{10}$, and $P_k^{-1}$ be given by (4.65), where $\varepsilon = 10^{-8}$. Figures 4.10a) and b) show the information content $|\psi_{k,i}|$ for several values of $\lambda$. Note that the information subspace is six dimensional due to the presence of three harmonics in $u_k$ as shown by six relatively large components of $\psi_k$. Figures 4.10c) and d) show the singular values of $P_k^{-1}$ for several values of $\lambda$. Note that the singular values that correspond to the singular vectors not in the information subspace do not convergeto zero. ◇

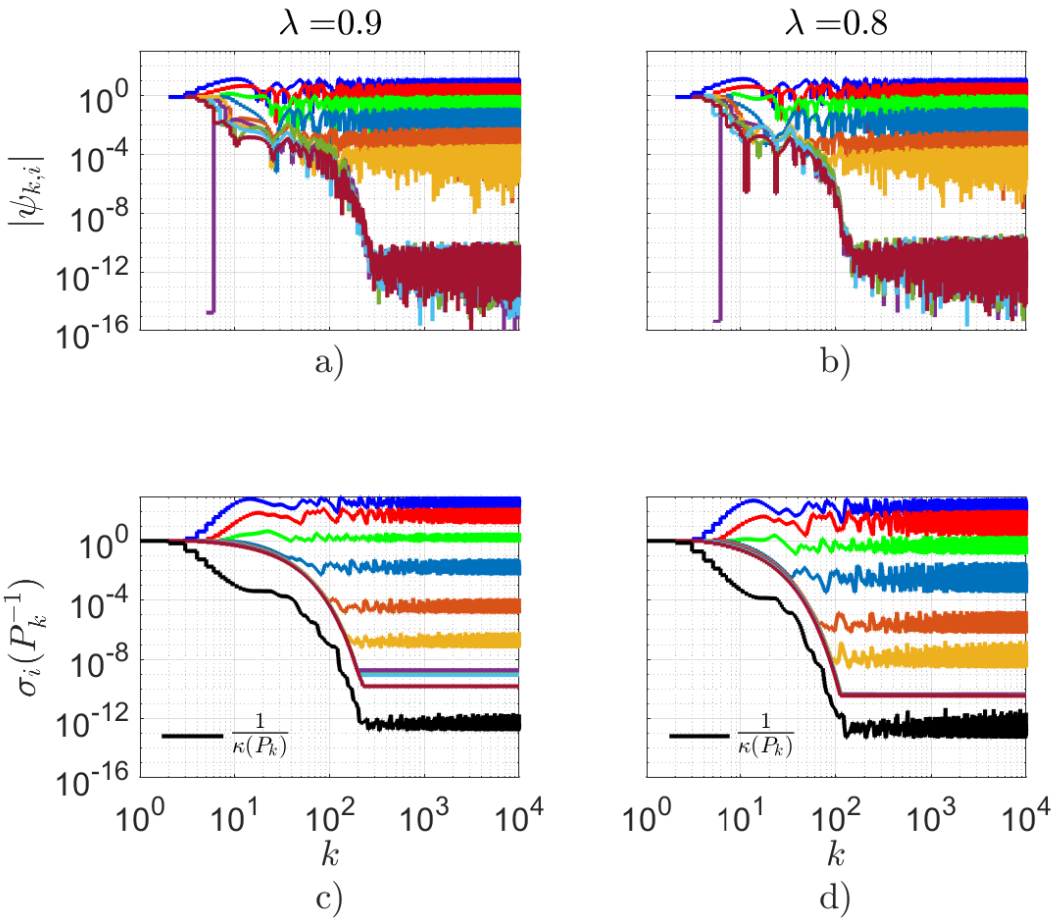



Figure 4.10: Example 4.10. Information-driven forgetting for a regressor lacking persistent excitation. a) and b) show the information content $\|\psi_k\|$ for $\lambda = 0.9$ and $\lambda = 0.8$. c) and d) show the singular values of $P_k^{-1}$ for $\lambda = 0.9$ and $\lambda = 0.8$. The inverse of the condition number of $P_k$ is shown in black. Note that, for $\lambda < 1$, the singular values that correspond to the singular vectors not in the information subspace do not converge to zero.

**Example 4.11. Effect of information-driven forgetting on $\theta_k$.** Reconsider Example 4.8. Let $P_0 = I_{10}$, and $P_k^{-1}$ be given by (4.65), where $\varepsilon = 10^{-8}$. Figure 4.11a) shows the predicted error $z_k$, and b) shows the norm of the parameter error

$\tilde{\theta}_k$. Note that the $\tilde{\theta}_k$ does not converge to 0. c) shows the singular values of $P_k$, and d) shows the condition number of $P_k$. Note that, unlike spatially uniform forgetting, all singular values of $P_k$ remain bounded, and $\theta_k$ does not diverge. ◇

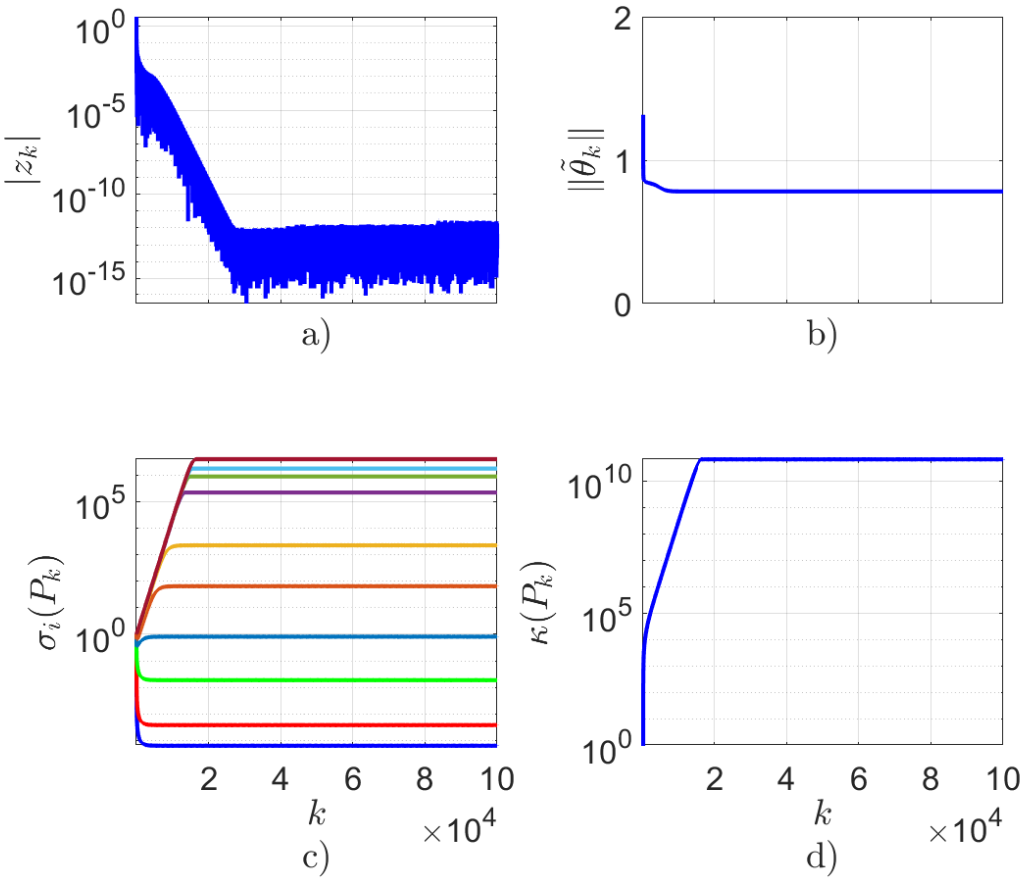



Figure 4.11: Example 4.11. Effect of Information-driven forgetting on $\theta_k$. a) shows the predicted error $z_k$, b) shows the norm of the parameter error $\tilde{\theta}_k$, c) shows the singular values of $P_k$, and d) shows the condition number of $P_k$. Note that all singular values of $P_k$ remain bounded.

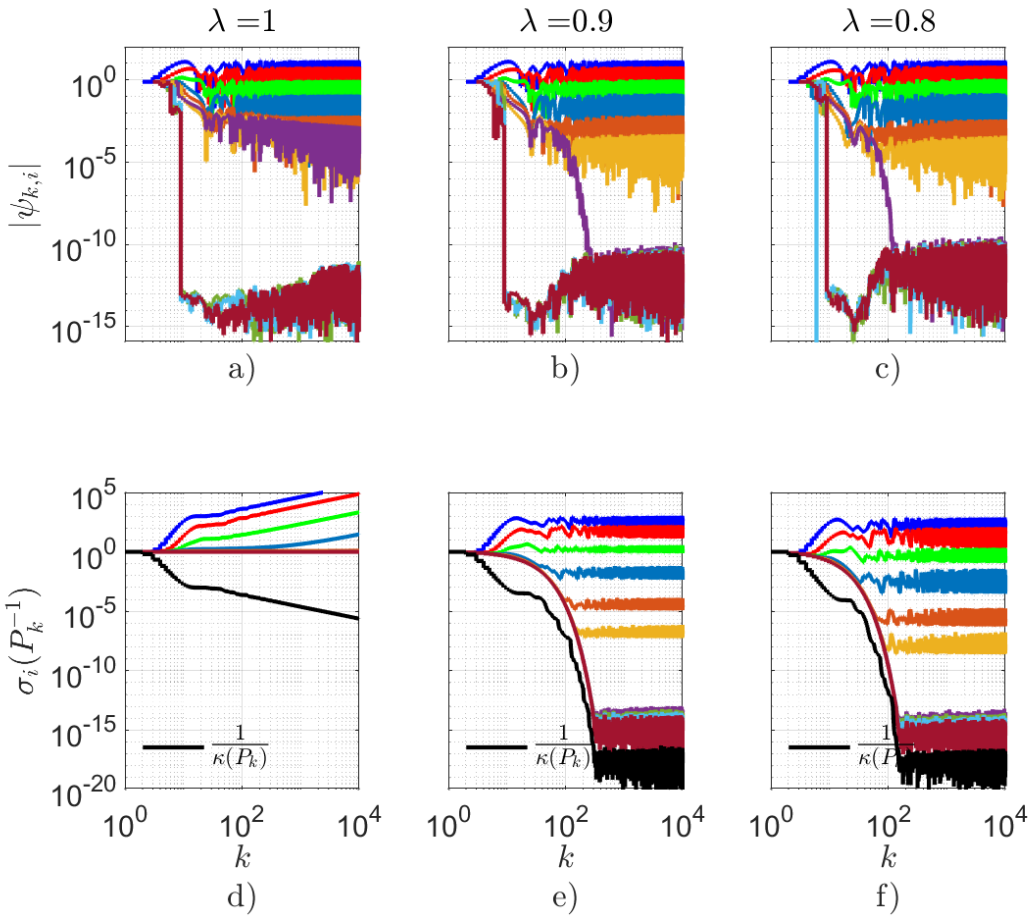



Figure 4.9: Example 4.9. Relation between $P_k$ and the information content $\psi_k$. a), b), and c) show the information content $\psi_{k,i}$ for several values of $\lambda$. Note that, in each case, the information subspace is six dimensional due to the presence of three harmonics in $u_k$. d), e), and (f) show the singular values of $P_k^{-1}$ for several values of $\lambda$. The inverse of the condition number of $P_k$ is shown in black. Note that, for $\lambda < 1$, the singular values of $P_k^{-1}$ corresponding to the singular vectors in the orthogonal complement of the information subspace converge to zero.

## 4.8 Conclusions

This chapter presented various results on uniform and directional forgetting within the context of RLS. It was shown that, in the case of persistent excitation without forgetting, the parameter estimates converge asymptotically, whereas, with forgetting, the parameter estimates converge geometrically. Numerical examples were presented to illustrate this behavior.

In the case where the excitation is not persistent and forgetting is used, it was shown that forgetting is enforced in all information directions, whether or not new information is present in along all information directions. Consequently, the parameter estimates converge, but not necessarily to their true values; furthermore, the matrix $P_k$ diverges leading to numerical instability. This phenomenon was traced to the divergence of the singular values of $P_k$ corresponding to singular vectors orthogonal to the information subspace.

In order to address this problem, a data-dependent forgetting matrix was constructed to restrict forgetting to the information subspace. Numerical examples presented show that this directional forgetting technique prevents $P_k$ from diverging under lack of persistent excitation.

# CHAPTER 5

# Adaptive Squaring-Based Control Allocation for Wide Systems

## 5.1 Introduction

Redundant actuators provide the opportunity to allocate control effort to account for saturation and other input constraints. This is especially true in wide systems, where the number of input channels is greater than the number of outputs. The present chapter considers the control allocation problem within the context of adaptive control.

Many control applications possess the mixed blessing of more actuators than controlled outputs. For these applications, the required control inputs can be realized by combinations of redundant actuators, thus providing reliability. At the same time, however, the presence of actuators with similar effect on the system demands that these actuators be used optimally with respect to magnitude and rate saturation, energy usage, and other limitations. This is the *control allocation problem* [90, 91].

Control allocation is widely studied within the context of aerospace vehicles [92–95], where the moments obtainable from control surfaces are limited by there size and depend on the current the aircraft speed. A classical case of actuator redundancy in lateral flight control concerns the use of ailerons and rudders to perform turning

maneuvers. Separate use of these actuators for turning yields poor performance, due to adverse aileron yaw and adverse rudder roll. Consequently, control allocation can ensure that the ailerons and rudder are used in a desirable combination to perform coordinated turns.

In the present chapter, we consider control allocation within the context of retrospective cost adaptive control (RCAC) [96]. RCAC is a direct digital control technique that is applicable to stabilization, command following, and disturbance rejection. A key feature of RCAC discussed in [96] is the role of the filter $G_{\mathrm{f}}$ in defining the retrospective cost variable. As shown in [96], $G_{\mathrm{f}}$ serves as a target model for the intercalated transfer function from the virtual external control perturbation to the performance variable. Modeling information concerning the leading numerator sign, relative degree, and nonminimum-phase (NMP) zeros is used to construct $G_{\mathrm{f}}$.

The present chapter extends the development of [96] by focusing on wide (that is, overactuated) systems for the purpose of control allocation. In particular, we show that the applied control input lies in the range of the target model. For example, if the plant is MISO and the target model $G_{\mathrm{f}}$ is chosen to be $G_{\mathrm{f}}(\mathbf{q}) = N_1 \frac{1}{\mathbf{q}}$, where $\mathbf{q}$ is the forward shift operator and $N_1$ is a row vector, then the control input vector is constrained to the direction $N_1^{\mathrm{T}}$. For example, if the plant has two inputs and one output and $N_1$ is chosen to be $N_1 = [3 \;\; 6]$, then, for all time steps $k$, $u_1(k)/u_2(k) = 1/2$. Consequently, the target model provides a simple and convenient technique for control allocation in overactuated systems.

The contents of the chapter are as follows. Section 5.2 states the control allocation problem. Next, Section 5.3 reviews RCAC, and Theorem 5.1 demonstrates how the target model constrains the allowable directions of the control input. Section 5.4 presents several examples illustrating control allocation for RCAC. The chapter ends with some conclusions and directions for future research.

## 5.2 Control allocation problem

Consider the MIMO discrete-time plant

$$x(k+1) = f(x(k), u(k)) + D_1 w(k), \tag{5.1}$$

$$y(k) = g(x(k)) + D_2 w(k), \tag{5.2}$$

$$z(k) = Ex(k) + E_0 w(k), \tag{5.3}$$

where $x(k) \in \mathbb{R}^{l_x}$ is the state, $y(k) \in \mathbb{R}^{l_y}$ is the measurement, $u(k) \in \mathbb{R}^{l_u}$ is the control signal, $w(k) \in \mathbb{R}^{l_w}$ is the exogenous signal, and $z(k) \in \mathbb{R}^{l_z}$ is the performance variable. The components of $w$ can represent either command signals to be followed, external disturbances to be rejected, or both. Figure 5.1 shows a block diagram representation of (5.1)–5.3.

The goal is to develop an adaptive output-feedback controller that produces the control effort in a desired ratio and minimizes $z$ in the presence of the exogenous signal $w$ with limited modeling information about (5.1)–(5.3) .

In this chapter, we consider overactuated systems, that is, $l_u > l_y$. In overactuated systems, the control input required to generate a desired output is not unique. For example, consider step command following for a two-input, one-output LTI plant. At steady state,

$$y_\infty = C(I - A)^{-1} B u_\infty, \tag{5.4}$$

where $y_\infty \in \mathbb{R}$ is the asymptotic output of the plant, and $u_\infty \in \mathbb{R}^2$ is the asymptotic control input. Since $C(I - A)^{-1}B$ is a wide matrix, it follows that the asymptotic control input $u_\infty$ is not unique. The goal of the control allocation problem is thus to achieve the specified performance objectives while confining the control input vector $u$ to a chosen subspace that constrains the direction of the allowable control inputs.

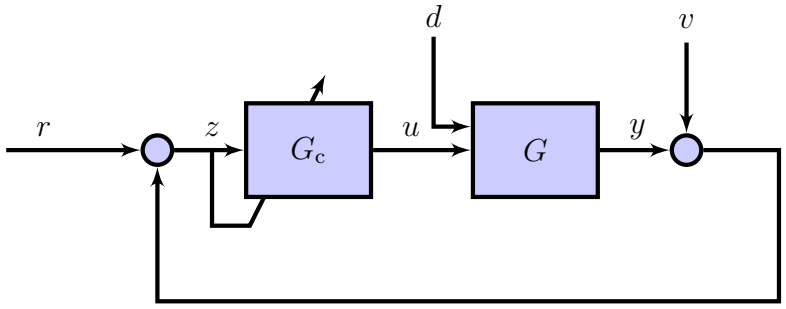



Figure 5.1: Block diagram representation of the adaptive control allocation problem with the adaptive controller $G_c$ and plant $G$. The goal is to design the controller $G_c$ that produces the control effort in a desired ratio such that the plant output $y$ follows the reference command $r$ in presence of process noise $d$ and measurement noise $v$.

## 5.3 RCAC algorithm

We consider a dynamic compensator represented by an ARMA model.The control $u(k)$ is thus given by

$$u(k) = \sum_{i=1}^{n_c} P_i(k)u(k-i) + \sum_{i=1}^{n_c} Q_i(k)z(k-i), \tag{5.5}$$

where the coefficient matrices $P_i(k) \in \mathbb{R}^{l_u \times l_u}$, $Q_i(k), R(k) \in \mathbb{R}^{l_u \times l_z}$ are updated by the RCAC algorithm.

We rewrite (5.5) as

$$u(k) = \Phi(k)\theta(k), \tag{5.6}$$

where the regressor matrix $\Phi(k)$ is defined by

$$\Phi(k) \triangleq I_{l_u} \otimes \phi^{\mathrm{T}}(k) \in \mathbb{R}^{l_u \times l_\theta}, \tag{5.7}$$

where

$$\phi(k) \triangleq [\hat{\mu}(k-1)^{\mathrm{T}} \ \cdots \ \hat{\mu}(k-n_{\mathrm{c}})^{\mathrm{T}} \ z(k-1)^{\mathrm{T}} \ \cdots \ z(k-n_{\mathrm{c}})^{\mathrm{T}}]^{\mathrm{T}},$$

$$\theta(k) \triangleq \mathrm{vec} \left[ \begin{array}{c} P_1(k) \cdots P_{n_{\mathrm{c}}}(k) \ Q_1(k) \cdots Q_{n_{\mathrm{c}}}(k) \end{array} \right] \in \mathbb{R}^{l_\theta},$$

$l_\theta \triangleq l_u^2 n_{\mathrm{c}} + l_u l_y n_{\mathrm{c}}$, "$\otimes$" is the Kronecker product, and "vec" is the column-stacking operator.

### 5.3.1 Retrospective Performance Variable

We define the retrospective performance variable

$$\hat{z}(k) = z(k) + G_{\mathrm{f}}(\mathbf{q})(\Phi(k)\hat{\theta} - u(k)), \tag{5.8}$$

where $\mathbf{q}$ is the forward-shift operator, $\hat{\theta} \in \mathbb{R}^{l_\theta}$ contains the controller coefficients to be optimized,

$$G_{\mathrm{f}}(\mathbf{q}) = \sum_{i=1}^{n_{\mathrm{f}}} N_i \mathbf{q}^{-i}, \tag{5.9}$$

and, for all $i = 1, \ldots, n_{\mathrm{f}}$, $N_i \in \mathbb{R}^{l_y \times l_u}$. $G_{\mathrm{f}}$ is an FIR filter of order $n_{\mathrm{f}}$ whose choice is discussed below. We rewrite (5.8) as

$$\hat{z}(k) = z(k) + N\Phi_{\mathrm{b}}(k)\hat{\theta} - NU_{\mathrm{b}}(k), \tag{5.10}$$

where

$$N \triangleq \begin{bmatrix} N_1 & \cdots & N_{n_\mathrm{f}} \end{bmatrix} \in \mathbb{R}^{l_y \times n_\mathrm{f} l_u},$$

$$\Phi_\mathrm{b}(k) \triangleq \begin{bmatrix} \Phi(k-1) \\ \vdots \\ \Phi(k-n_\mathrm{f}) \end{bmatrix} \in \mathbb{R}^{l_u n_\mathrm{f} \times l_\theta},$$

$$U_\mathrm{b}(k) \triangleq \begin{bmatrix} u(k-1) \\ \vdots \\ u(k-n_\mathrm{f}) \end{bmatrix} \in \mathbb{R}^{l_u n_\mathrm{f}}.$$

The vector $\hat{\theta}$, which contains the controller coefficients, is determined by minimizing the retrospective cost function, as described next.

### 5.3.2 Retrospective Cost Function

Using the retrospective performance variable $\hat{z}(k)$, we define the retrospective cost function

$$J(k,\hat{\theta}) \triangleq \sum_{i=1}^{k-1} \lambda^{k-i} \big( \hat{z}(i)^\mathrm{T} R_z \hat{z}(i) + \hat{\theta}^\mathrm{T} \Phi_\mathrm{b}(i)^\mathrm{T} N^\mathrm{T} R_\mathrm{f} N \Phi(i)_\mathrm{b} \hat{\theta} + \hat{\theta}^\mathrm{T} \Phi(i)^\mathrm{T} R_u \Phi(i) \hat{\theta} \big) + \lambda^k \hat{\theta}^\mathrm{T} R_\theta \hat{\theta}, \tag{5.11}$$

where $R_z, R_\mathrm{f}, R_u$, and $R_\theta$ are positive-definite matrices, and $\lambda \leq 1$ is the forgetting factor. The following result uses recursive least squares (RLS) to minimize (5.11).

**Proposition 5.1.** *Let $P(0) = R_\theta^{-1}$ and $\theta(0) = 0$. Then, for all $k \geq 1$, the retrospec-*

*tive cost function* (5.11) *has a unique global minimizer* $\theta(k)$*, which is given by*

$$\theta(k) = \theta(k-1) - P(k) \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1) \\ \Phi(k-1) \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} R_z & 0 \\ 0 & R_u \end{bmatrix} \cdot \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1)\theta(k-1) + z(k-1) - NU_{\mathrm{b}}(k-1) \\ \Phi(k-1)\theta(k-1) \end{bmatrix}, \tag{5.12}$$

$$P(k) = \lambda^{-1}P(k-1) - \lambda^{-1}P(k-1) \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1) \\ \Phi(k-1) \end{bmatrix}^{\mathrm{T}} \cdot \Gamma(k)^{-1} \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1) \\ \Phi(k-1) \end{bmatrix} P(k-1), \tag{5.13}$$

*where*

$$\Gamma(k) \triangleq \lambda \begin{bmatrix} R_z & 0 \\ 0 & R_u \end{bmatrix}^{-1} + \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1) \\ \Phi(k-1) \end{bmatrix} P(k-1) \cdot \begin{bmatrix} N\Phi_{\mathrm{b}}(k-1) \\ \Phi(k-1) \end{bmatrix}^{\mathrm{T}}. \tag{5.14}$$

*Furthermore, the control input at step* $k$ *is given by*

$$u(k) = \Phi(k)\theta(k). \tag{5.15}$$

### 5.3.3 The Target Model $G_{\mathrm{f}}$

The following result shows that the control $u(k)$ is constrained to the subspace spanned by the transposes of the coefficients of $G_{\mathrm{f}}$.

**Theorem 5.1.** *Let* $R_\theta = \beta I_{l_\theta}$*,* $R_u = \gamma I_{l_u}$ *and let* $\theta(k)$ *be given by* (5.15)*. Let*

$\Phi \triangleq I_{l_u} \otimes \phi^{\mathrm{T}}$, *where* $\phi \in \mathbb{R}^{l_\phi}$, *and* $l_\phi \triangleq l_\theta / l_u$. *Then, for all* $k \geq 1$,

$$\Phi\theta(k) \in \mathcal{R}\left(\begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix}\right). \tag{5.16}$$

*Proof.* Note that

$$\begin{aligned}
\Phi A_\theta(k)\theta(k) &= \sum_{i=1}^{k-1} \lambda^{k-i} \Phi \left(\Phi_{\mathrm{b}}(i)^{\mathrm{T}} N^{\mathrm{T}} (R_z + R_{\mathrm{f}}) N \Phi_{\mathrm{b}}(i) + \Phi(i)^{\mathrm{T}} \gamma \Phi(i)\right) \theta(k) + \\
&\quad \lambda^k \beta \Phi \theta(k) \\
&= \sum_{i=1}^{k-1} \left( \lambda^{k-i} \sum_{j=1}^{n_{\mathrm{f}}} (I_{l_u} \otimes \phi^{\mathrm{T}})(I_{l_u} \otimes \phi(i-j)) N_j^{\mathrm{T}} \right) (R_z + R_{\mathrm{f}}) N \Phi_{\mathrm{b}}(i) \theta(k) + \\
&\quad \lambda^{k-i} (I_{l_u} \otimes \phi^{\mathrm{T}})(I_{l_u} \otimes \phi(i)) \gamma \Phi(i) \theta(k) + \lambda^k \beta \Phi \theta(k) \\
&= \sum_{i=1}^{k-1} \left( \lambda^{k-i} \sum_{j=1}^{n_{\mathrm{f}}} N_j^{\mathrm{T}} \phi^{\mathrm{T}} \phi(i-j) \right) (R_z + R_{\mathrm{f}}) N \Phi_{\mathrm{b}}(i) \theta(k) + \\
&\quad \lambda^{k-i} \gamma \phi^{\mathrm{T}} \phi(i) \Phi(i) \theta(k) + \lambda^k \beta \Phi \theta(k) \\
&= \begin{bmatrix} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{bmatrix} \sum_{i=1}^{k-1} \lambda^{k-i} \Xi(i) (R_z + R_{\mathrm{f}}) N \Phi_{\mathrm{b}}(i) \theta(k) + \bar{\Phi}(k) \theta(k),
\end{aligned}$$

where

$$\begin{aligned}
\Xi(i) &\triangleq \begin{bmatrix} \phi^{\mathrm{T}} \phi(i-1) \otimes I_{l_z} \\ \vdots \\ \phi^{\mathrm{T}} \phi(i-n_{\mathrm{f}}) \otimes I_{l_z} \end{bmatrix}, \\
\bar{\Phi}(k) &\triangleq I_{l_u} \otimes \bar{\phi}(k)^{\mathrm{T}}, \\
\bar{\phi}(k) &\triangleq M(k) \phi, \\
M(k) &\triangleq \sum_{i=1}^{k-1} \lambda^{k-i} \gamma \phi(i) \phi(i)^{\mathrm{T}} + \lambda^k \beta I_{l_\phi}.
\end{aligned}$$

Note that $M(k)$ is positive-definite. Further,

$$\begin{aligned}
&\Phi b_\theta(k) \\
&= \Phi \sum_{i=1}^{k-1} \lambda^{k-i} \Phi_{\mathrm{b}}(i)^{\mathrm{T}} N^{\mathrm{T}} R_z \left( z(i) - N U_{\mathrm{b}}(i) \right) \\
&= \left[ \begin{array}{ccc} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{array} \right] \sum_{i=1}^{k-1} \lambda^{k-i} \Xi(i) R_z \left( z(i) - N U_{\mathrm{b}}(i) \right).
\end{aligned}$$

Using $\Phi A_\theta(k)\theta(k) = -\Phi b_\theta(k)$ yields

$$\begin{aligned}
\bar{\Phi}(k)\theta(k) = &- \left[ \begin{array}{ccc} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{array} \right] \sum_{i=1}^{k-1} \lambda^{k-i} \Xi(i) \\
&\cdot \left( R_z z(i) + (R_z + R_{\mathrm{f}}) N \Phi_{\mathrm{b}}(i) \theta(k) - R_z N U_{\mathrm{b}}(i) \right) \\
&\in \mathcal{R} \left( \left[ \begin{array}{ccc} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{array} \right] \right).
\end{aligned}$$

For $\phi \in \mathbb{R}^{l_\phi}$, there exists a unique $\tilde{\phi} \in \mathbb{R}^{l_\phi}$ such that $\phi = M(k)\tilde{\phi}$. Consequently, $\Phi = I_{l_u} \otimes (M(k)\tilde{\phi})^{\mathrm{T}}$. It follows that $\Phi\theta(k) \in \mathcal{R} \left( \left[ \begin{array}{ccc} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{array} \right] \right)$. □

The following proposition follows from Theorem 5.1.

**Proposition 5.2.** *Let $R_\theta = \beta I_{l_\theta}$, $R_u = \gamma I_{l_u}$ and let $\theta(k)$ be given by* (5.15)*. Then, for all $k \geq 1$,*

$$u(k) = \Phi(k)\theta(k) \in \mathcal{R} \left( \left[ \begin{array}{ccc} N_1^{\mathrm{T}} & \cdots & N_{n_{\mathrm{f}}}^{\mathrm{T}} \end{array} \right] \right). \tag{5.17}$$

## 5.4 Illustrative Examples

In this section, we present several examples demonstrating control allocation based on RCAC.

### 5.4.1 Colocated actuators

Consider the two-input, one-output discrete-time system

$$x(k+1) = Ax(k) + Bu(k), \tag{5.18}$$

$$y(k) = Cx(k), \tag{5.19}$$

where

$$A = \begin{bmatrix} -0.2650 & -0.2225 & -0.0938 & -0.1211 \\ -0.0727 & 0.3296 & 0.0594 & -0.3238 \\ 0.2661 & -0.2316 & 0.1396 & 0.2074 \\ -0.2016 & -0.1591 & 0.8285 & 0.1352 \end{bmatrix}, \tag{5.20}$$

$$B = \begin{bmatrix} 0.8752 & 0.8752 \\ 0.3179 & 0.3179 \\ 0.2732 & 0.2732 \\ 0.6765 & 0.6765 \end{bmatrix}, \tag{5.21}$$

$$C = \begin{bmatrix} 0.8700 & 0.2437 & 0.8429 & 0.5577 \end{bmatrix}. \tag{5.22}$$

The identical columns of $B$ indicate that the two actuators are colocated. Let the command $r$ be a unit step. We set

$$G_{\mathrm{f}}(\mathbf{q}) = N_1\mathbf{q}^{-1} = \begin{bmatrix} 0.5 & 1 \end{bmatrix}\mathbf{q}^{-1}, \tag{5.23}$$

$n_{\mathrm{c}} = 2$, and $R_\theta = 10^{-3}I_{l_\theta}$. With this choice of $G_{\mathrm{f}}$, the control input $u(k)$ produced by RCAC is such that $u_2(k) = 2u_1(k)$ for all $k$. Figure 5.2 shows the closed-loop response of the system.

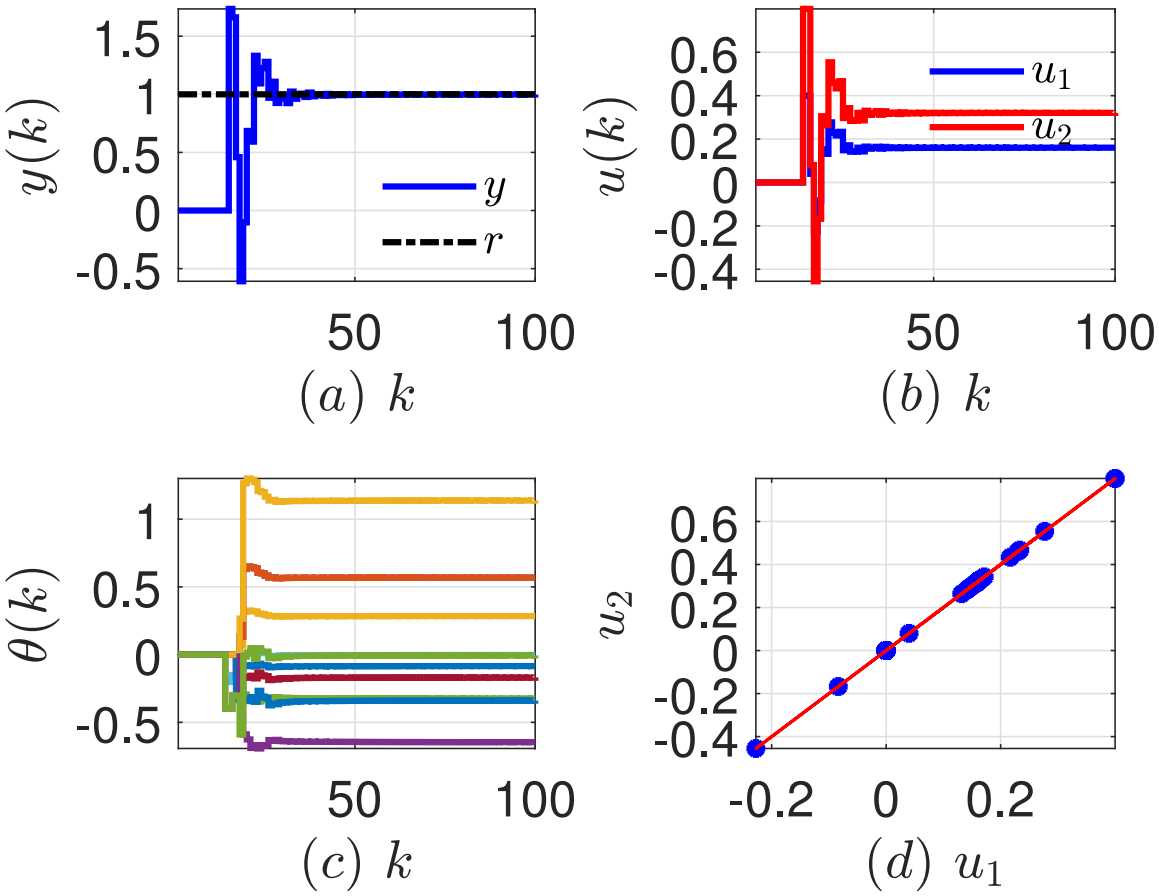



Figure 5.2: Example 5.4.1. Control allocation for setpoint command following. (a) shows the response $y$, (b) shows the control signal $u$ generated by RCAC, (c) shows the controller coefficients $\theta$ adapted by RCAC, and (d) shows the control signals. In (d), the direction of $N_1^{\mathrm{T}}$ is shown in red.

### 5.4.2 Independent actuators

We reconsider the plant in Example 5.4.1 with the modified input matrix

$$B = \begin{bmatrix} 0.8752 & 0.0712 \\ 0.3179 & 0.1966 \\ 0.2732 & 0.5291 \\ 0.6765 & 0.1718 \end{bmatrix}. \tag{5.24}$$

Let the command $r$ be a unit step. We set

$$G_{\mathrm{f}}(\mathbf{q}) = N_1 \mathbf{q}^{-1} = \begin{bmatrix} 0.5 & 1 \end{bmatrix} \mathbf{q}^{-1}, \tag{5.25}$$

$n_{\mathrm{c}} = 2$, and $R_\theta = 10^{-3} I_{l_\theta}$. With this choice of $G_{\mathrm{f}}$, the control input produced by RCAC is in the direction $N_1^{\mathrm{T}}$. Figure 5.3 shows the closed-loop response of the system.

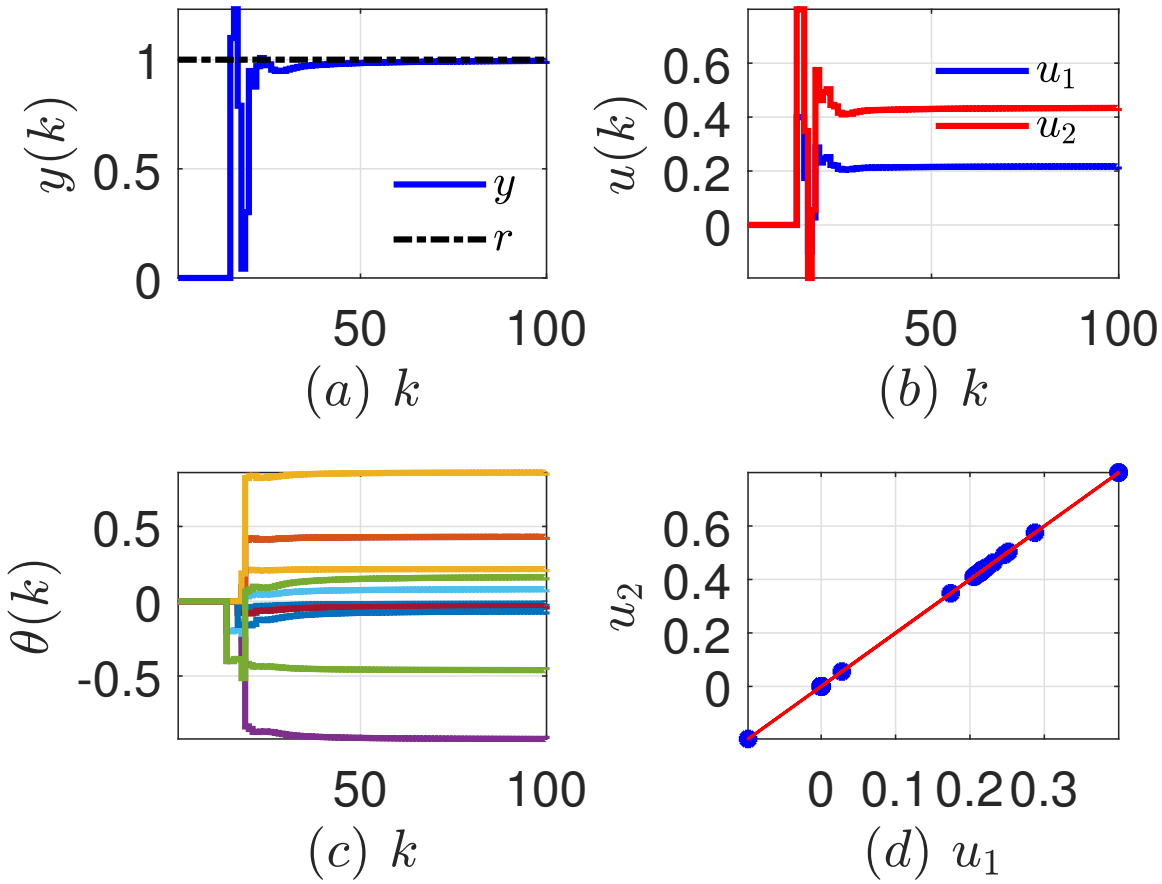



Figure 5.3: Example 5.4.2. Control allocation for setpoint command following. (a) shows the response $y$, (b) shows the control signal $u$ generated by RCAC, (c) shows the controller coefficients $\theta$ adapted by RCAC, and (d) shows the control signals. In (d), the direction of $N_1^{\mathrm{T}}$ is shown in red.

### 5.4.3 Uncontrollable channels

We reconsider the plant in Example 5.4.1 with the modified input matrix

$$B = \begin{bmatrix} b_1 & b_2 \end{bmatrix} = \begin{bmatrix} 0.8826 & 0.0984 \\ 0.5405 & 0.2135 \\ 0.0808 & 0.5128 \\ 0.3543 & 0.2059 \end{bmatrix}. \tag{5.26}$$

Note that $(A, b_1)$ and $(A, b_2)$ each have one uncontrollable mode, but $(A, B)$ is controllable. Hence, $u_1$ and $u_2$ must work together to achieve the control objective. Let the command $r$ be a unit step. We set

$$G_{\mathrm{f}}(\mathbf{q}) = N_1\mathbf{q}^{-1} = \begin{bmatrix} 0.5 & 1 \end{bmatrix} \mathbf{q}^{-1}, \tag{5.27}$$

$n_{\mathrm{c}} = 2$, and $R_\theta = 10^{-3}I_{l_\theta}$. With this choice of $G_{\mathrm{f}}$, the control input produced by RCAC is in the direction $N_1^{\mathrm{T}}$. Figure 5.4 shows the closed-loop response of the system.

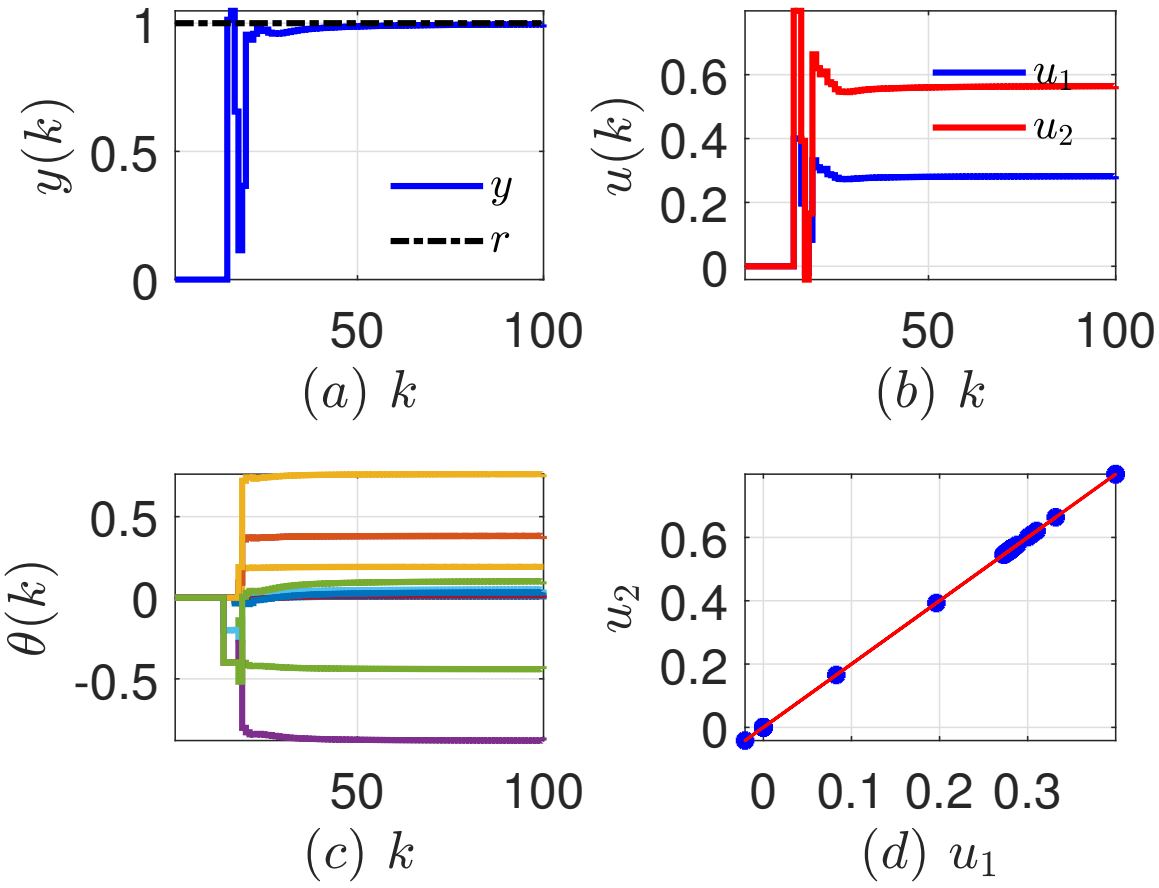



Figure 5.4: Example 5.4.3. Control allocation for setpoint command following. (a) shows the response $y$, (b) shows the control signal $u$ generated by RCAC, (c) shows the controller coefficients $\theta$ adapted by RCAC, and (d) shows the control signals. In (d), the direction of $N_1^{\mathrm{T}}$ is shown in red.

## 5.5 Conclusions

This chapter presented a novel approach to control allocation for wide systems within the context of retrospective cost adaptive control (RCAC). The approach is based on the fact that the control input is constrained to the range of the transpose of the target model. This technique was applied to illustrative examples.

The next chapter extends RCAC to asymptotically enforce auxiliary output con-

straints by formulating the auxiliary output constraints as conflicting commands.

# CHAPTER 6

# Output-Constrained Adaptive Control for Unstart Prevention in a 2D Scramjet Combustor

## 6.1 Introduction

A supersonic combustion ramjet (scramjet) engine is an air-breathing propulsion system that compresses and ignites a supersonic stream of air and generates thrust by expelling high-energy gases through a nozzle without any moving parts. Scramjet engines have high potential for high-speed transport, access to space, and payload delivery. Although mechanically simple, the flow inside a scramjet engine, comprised of pockets of subsonic regions and multiple shock structures, is extremely complex, and is highly sensitive to external disturbances such as inlet mach number and angle of attack of the leading ramp.

A major challenge in the operation of a scramjet is the problem of preventing the scramjet from unstarting. The flow inside the scramjet can thermally choke when sufficient heat is added in the combustor, resulting in a normal shock that travels upstream, thereby destroying the shock structure and the supersonic flow. Consequently, the thrust generated by the scramjet drops precipitously. Since it is extremely difficult to re-establish stable supersonic flow structure quickly enough once the engine has unstarted completely, unstart is usually fatal to the engine operation.

In addition to excessive heat addition, the scramjet engine can unstart because of flow and structural perturbations.

In [97], unstart was prevented by constraining the hypersonic vehicle to remain within a predefined flight envelope. In [98], a second-order linear, time-invariant model relating the ramp angle and the shock train location was identified. A PID controller was developed based on the identified model, and was applied to the high-order model to shift the location of the shock train to prevent unstart. In [99], a particle filter was used to estimate the state of a scramjet described by a one-dimensional model, and a full-state-feedback control law was suggested to prevent the scramjet from unstarting. As a result, the control altered the cycle length as well as the burst length of the fuel flow based on the predicted future location of the shock front computed using the filtered state. In [100], retrospective cost adaptive control (RCAC) was used for unstart prevention in a 2D scramjet model with a single pressure measurement near the combustor, however, the closed-loop response was oscillatory. In the present work, RCAC is used to prevent unstart by enforcing auxiliary output constraints.

RCAC is a direct discrete-time adaptive control algorithm that is applicable to stabilization, command following, and disturbance rejection. RCAC minimizes a convex cost function based on a performance variable, such as the thrust error, to optimize the controller coefficients based on past control inputs and past measurements. The retrospective cost minimized by RCAC is based on the retrospective performance, which is defined using the target model $G_{\mathrm{f}}$. The target model $G_{\mathrm{f}}$ reflects the qualitative response of the performance variable to the control input, which is usually captured by the key dynamics features of the system to be controlled such as its relative degree, sign of the leading Markov parameter, and nonminimum-phase (NMP) zeros, if any. As shown in [30], RCAC adapts the control law in order to match the intercalated transfer function to the target model $G_{\mathrm{f}}$.

RCAC was applied to the problem of possibly conflicting commands in an aircraft to follow trim commands in [101]. The contribution of the present chapter is to show that output constraints can be asymptotically enforced by casting them as additional conflicting commands; these commands are conflicting in the sense that the tradeoff necessitates a constraint violation. As in the case of conflicting commands, the asymptotic response of the controlled system may not be able follow all of the commands simultaneously, and thus RCAC automatically trades off the output error and the constraint violation error. Although the constraint is violated, it is shown that the asymptotic magnitude of the constraint violation can be reduced by varying the tuning parameters in RCAC. By defining a conservative constraint, the required constraint can be satisfied.

The chapter is organized as follows: in Section 6.2, the scramjet model used in this chapter is presented; unstart and its characterization based on open-loop simulations of the scramjet model is presented in Section 6.3. The control architecture for the setpoint command following is presented in Section 6.4. In Section 6.5, RCAC is extended to the problem of following conflicting setpoint commands. In Section 6.6, RCAC is extended to the problem of preventing the scramjet from unstarting by enforcing an auxiliary output constraint. Finally, the chapter is concluded in Section 6.7.

## 6.2 Scramjet Model

The scramjet engine considered in this work is based on the Hyshot-II geometry [102] shown in Figure 6.1. The flow is modeled by coupled nonlinear partial differential equations based on the assumption of inviscid flow along with a heat release model.

The two-dimensional CFD model is obtained by discretizing the Euler equations

$$\frac{\partial}{\partial t}\begin{bmatrix}\rho\\ \rho u\\ \rho v\\ e\end{bmatrix}+\frac{\partial}{\partial x}\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ u(e+p)\end{bmatrix}+\frac{\partial}{\partial y}\begin{bmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ v(e+p)\end{bmatrix}=\begin{bmatrix}0\\ 0\\ 0\\ \dot{Q}\end{bmatrix}, \tag{6.1}$$

where $u, v$ are the longitudinal and lateral velocity components, $\rho$ is the fluid density, and $p$ is the pressure. The total energy per unit volume

$$e=\frac{p}{\gamma-1}+\frac{1}{2}\rho(u^2+v^2), \tag{6.2}$$

where $\gamma$ is the ratio of specific heats.

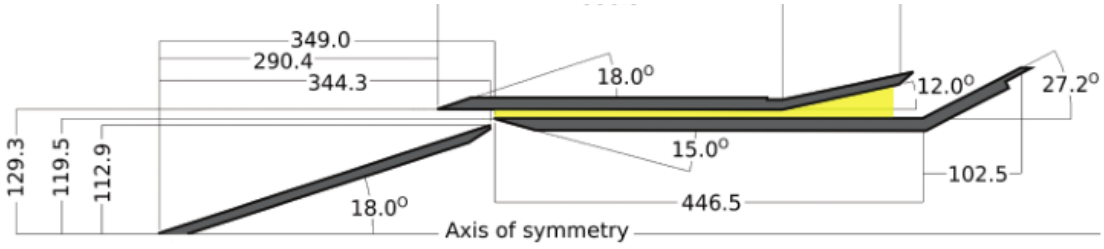



Figure 6.1: Geometry of the Hyshot-II scramjet engine [103]. The solution domain is shown in yellow. Note that all lengths are shown in millimeters.

The discretized governing equations are solved over the isolator, combustor and the nozzle section of the engine (shown in yellow in Figure 6.1). Note that the model has $O(10^5)$ degrees of freedom. The flow is assumed to be deflected by two shocks from the vehicle's forebody and cowl at the entrance of the isolator section. All numerical computations are performed using the unstructured mesh compressible flow solver Joe [103–105].

The combustion process is modeled by the heat-release model given in [106]. In

particular, the source term $Q$ is modeled by

$$\dot{Q} = \phi f_{\mathrm{st}} H_{\mathrm{f}} \dot{m}_{\mathrm{air}} \eta(x/L_{\mathrm{c}}), \tag{6.3}$$

$$\eta(x/L_{\mathrm{c}}) = 1 - e^{-(C_{\mathrm{c}}(x-x_{\mathrm{c}})/L_{\mathrm{c}})^{D_{\mathrm{c}}}}, \tag{6.4}$$

where $x_{\mathrm{c}} < x < x_{\mathrm{c}} + L_{\mathrm{c}}$, $\phi$ is the fuel-air equivalence ratio, $\eta$ is the heat-release distribution function, and the numerical values of various parameters is given in Table 6.1. The thrust generated by the scramjet is computed by integrating the pressure along the flow path, and is normalized and nondimensionalized by the free-stream dynamic pressure and unit area. Finally, the heat addition, and thus the thrust is controlled by varying the fuel-air equivalence ratio $\phi$

| Symbol | Definition | Nominal Value |
|---|---|---|
| $f_{\mathrm{st}}$ | Stoichiometric fuel/air ratio | 0.028 |
| $H_{\mathrm{f}}$ | Fuel heating value ($H_2$) | 120 MJ/kg |
| $x_{\mathrm{c}}$ | Combustion ignition position | 0.408 m |
| $L_{\mathrm{c}}$ | Combustor length | 0.368 m |
| $K_{\mathrm{c}}$ | Fraction of completed combustion | 0.95 |
| $D_{\mathrm{c}}$ | Shape parameter | 0.75 |
| $C_{\mathrm{c}}$ | Shape parameter | $-\log(1-K_{\mathrm{c}})^{1/D_{\mathrm{c}}}$ |

Table 6.1: Summary of parameters used in the heat-release model (6.3), (6.4).

## 6.3 Unstart and Unstart detection

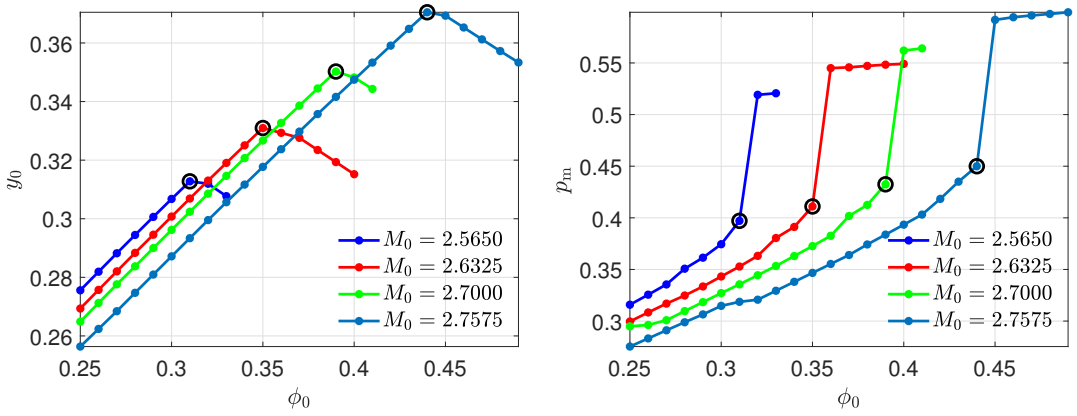
A major challenge in the operation of a scramjet is the need to prevent the scramjet from unstarting. When sufficient heat is released in the combustor, usually by means of too much fuel flow, the flow can thermally choke, resulting in a normal shock that travels upstream, thereby destroying the shock structure and the supersonic

flow. Consequently, the thrust generated by the scramjet drops precipitously. This is usually fatal to the engine operation since it is extremely difficult to establish a stable supersonic flow structure quickly enough once the engine has unstarted. In addition to excessive heat addition, a scramjet engine can unstart because of flow and structural perturbations. Because of unstart, there is an upper limit on the thrust that a scramjet engine can produce at a given operating condition. The maximum asymptotic thrust that can be generated by the scramjet at a constant inlet Mach number $M_0$ is called the *critical thrust* at $M_0$, and the constant equivalence ratio $\phi_0$ that corresponds to the critical thrust at $M_0$ is called the *critical equivalence ratio* at $M_0$.

To characterize the onset of unstart, we simulate the scramjet without feedback control for various fixed values of the equivalence ratio $\phi_0$ at various constant inlet Mach numbers $M_0$ for 3000 time steps. At each $M_0$, if the applied equivalence ratio is less than the critical equivalence ratio, the scramjet reaches a stable equilibrium state and generates a constant thrust $y_0$ at the end of the simulation. If, however, the equivalence ratio is greater than the critical equivalence ratio, the state of the scramjet starts diverging as it attempts to establish a stable subsonic flow inside the scramjet, leading to a sharp loss of thrust. Figure 6.2(a) shows the asymptotic thrust $y_0$ for various values of constant equivalence ratios $\phi_0$ at various constant inlet Mach numbers $M_0$.

In [103], a functional defined based on the pressure profile on the upper wall of the combustor is used to detect unstart. In [100], a single pressure measurement is used to detect and prevent unstart. Since the shock structure shifts longitudinally as the equivalence ratio is varied, the measured pressure can change sharply as the shock moves across the pressure sensor. In the present work, the mean $p_{\mathrm{m}}$ of the pressure along the bottom wall near the combustor is used to detect unstart. Specifically, 20 pressure sensors located uniformly between $x = 0.279$ m and $x = 0.294$ m are used

to define the *pressure metric* $p_{\mathrm{m}}$. This sensor arrangement corresponds to pressure measurements from 20 grid points in the simulation. Figure 6.2(b) shows the mean pressure $p_{\mathrm{m}}$ for various values of constant equivalence ratios $\phi_0$ at various constant inlet Mach numbers $M_0$. Note that the mean pressure metric $p_{\mathrm{m}}$ increases abruptly once the scramjet unstarts due to the presence of a normal shock wave. Hence, the value of the metric $p_{\mathrm{m}}$ can be used to constrain the control algorithm to prevent the scramjet from unstarting.




(a) Steady state Thrust

(b) Pressure Metric

Figure 6.2: Critical thrust $y_0$ and the pressure metric $p_{\mathrm{m}}$ in the normal operating and unstarting scramjet. The critical thrust at each value of $M_0$ is shown with a black circle.

## 6.4 Adaptive Setpoint Command Following

In this work, RCAC is used to optimize an infinite impulse response (IIR) controller to follow setpoint thrust commands. The notation used for RCAC variables is given in [100]. The scramjet is assumed to be operating in a steady state with the equivalence ratio $\phi_0$ at the constant inlet Mach number $M_0$. The thrust error, defined as $z(k) \triangleq r(k) - y(k)$, where $r(k)$ is the commanded value of the thrust and $y(k)$ is the generated thrust, is used to drive and optimize the controller $G_{\mathrm{c}}$. The equivalence ratio applied to the scramjet is $\phi(k) = \phi_0 + u_{\mathrm{c}}(k)$, where $u_{\mathrm{c}}(k)$ is the output of $G_{\mathrm{c}}$.

The control system architecture is shown in Figure 6.3.

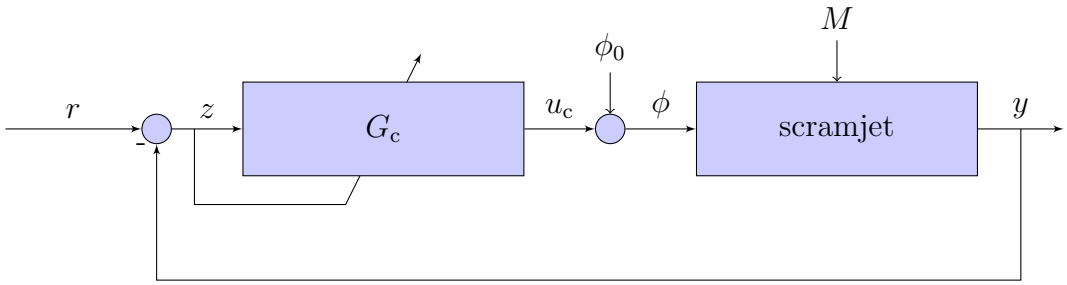



Figure 6.3: Scramjet command-following control system architecture. The controller $G_c$ is updated by RCAC.

In [100], a feedback/feedforward controller structure is optimized using RCAC, whereas, in this chapter, a feedback controller with an embedded integrator is optimized. The embedded integrator ensures that $u_c(k)$ converges to the required constant value as $k \to \infty$ and $z(k) \to 0$.

## 6.5 Adaptive Command Following with Conflicting Commands

Consider the SISO system

$$x(k+1) = f(x(k), u(k)), \tag{6.5}$$

$$y(k) = g(x(k), u(k)), \tag{6.6}$$

where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}$ is the input, and $y(k) \in \mathbb{R}$ is the measured output. The objective is to construct a control law such that $y(k) - r(k) \to 0$ as $k \to \infty$, where $r(k)$ is the command.

As a first step in developing a strategy for enforcing constraints, we consider the response of RCAC in the case of conflicting commands. In particular, we define the

performance variable

$$z(k) \stackrel{\triangle}{=} \begin{bmatrix} y(k) - r_1(k) \\ \vdots \\ y(k) - r_q(k) \end{bmatrix} \in \mathbb{R}^q, \tag{6.7}$$

where $r_1(k), \ldots, r_q(k) \in \mathbb{R}$. If $r_1, \ldots, r_q$ are different signals, then the commands are *conflicting*, that is, the output $y(k)$ is simultaneously requested to follow different commands. It is shown below that conflicting commands provide a technique for asymptotically enforcing constraints.

Let

$$N_1 = \begin{bmatrix} N_{1,1} \\ \vdots \\ N_{1,q} \end{bmatrix}, \quad R_z = \operatorname{diag}(R_{z,1}, \ldots, R_{z,q}), \tag{6.8}$$

where $R_{z,i} > 0$. Numerical examples show that, if $\lim_{k\to\infty} z(k)$ exists, then

$$\lim_{k\to\infty} \sum_{j=1}^{q} N_{1,j} R_{z,j} z_j(k) = 0. \tag{6.9}$$

Consequently,

$$\lim_{k\to\infty} z(k) \in \mathcal{N}(N_1^{\mathrm{T}} R_z), \tag{6.10}$$

and thus,

$$\lim_{k\to\infty} y(k) = \frac{\sum_{j=1}^{q} R_{z,j} N_{1,j} r_j}{\sum_{j=1}^{q} R_{z,j} N_{1,j}}. \tag{6.11}$$

Thus, by choosing $N_1$, and $R_z$, (6.11) can be used to specify the tradeoff between several conflicting commands. $\diamond$

**Example 6.1.** Consider the spring-mass-damper system

$$m\ddot{x} + c\dot{x} + kx = u, \tag{6.12}$$

where $x$ is the position of the particle with mass $m = 1$, the damping coefficient is $c = 5$, the spring stiffness is $k = 5$, and $u$ is the applied input. Let the measured output be $y(k) = x(kT_\mathrm{s})$, where $T_\mathrm{s} = 0.01$ sec is the sampling time.

The conflicting commands for the two-component performance variable (6.7) are chosen to be the setpoints $r_1(k) = 1$ and $r_2(k) = 2$. RCAC is implemented with $N_1 = [1\ 1]^\mathrm{T}$, $R_z = I_2$, $R_\theta = 10^2 I_{l_\theta}$, and $\lambda = 0.999$. Figure 6.4 shows the closed-loop response of the spring-mass-damper system with conflicting commands, (a) shows the measured output $y(k)$, (b) shows the control $u(k)$, (c) shows the output error $z(k)$, and (d) shows the coefficients of the controller optimized by RCAC.

Next, RCAC is implemented with $N_1 = [1\ N_{1,2}]^\mathrm{T}$, $R_z = \mathrm{diag}(1, R_{z,2})$, $R_\theta = 10^2 I_{l_\theta}$, and $\lambda = 0.999$, where $N_{1,2}$ and $R_{z,2}$) are varied from 0.1 to 10. For each choice of $N_{1,2}$ and $R_{z,2}$, Figure 6.5 shows the asymptotic output $y(k)$ the end of the simulation. Note that the asymptotic output satisfies (6.11). $\diamond$

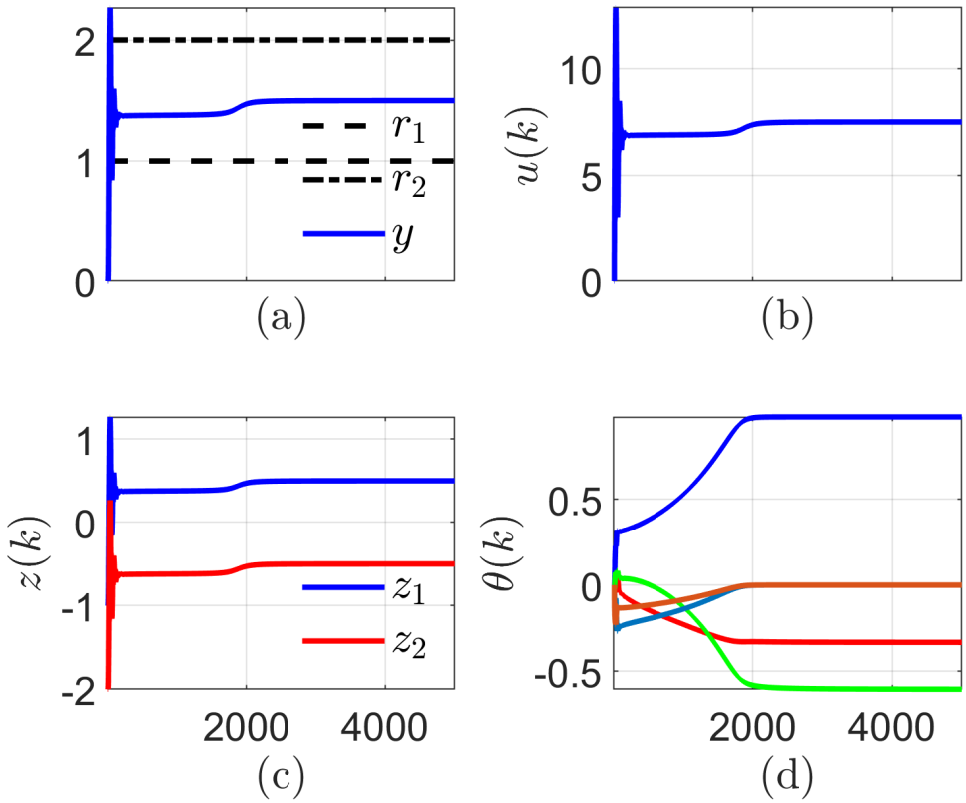



Figure 6.4: Example 6.1. Closed-loop response of the spring-mass-damper system with conflicting commands. (a) shows the commands and the output, (b) shows the equivalence ration $\phi$, (c) shows the output error $z$, and (d) shows the controller coefficients $\theta$. Note that the output converges to the mean of the conflicting commands.

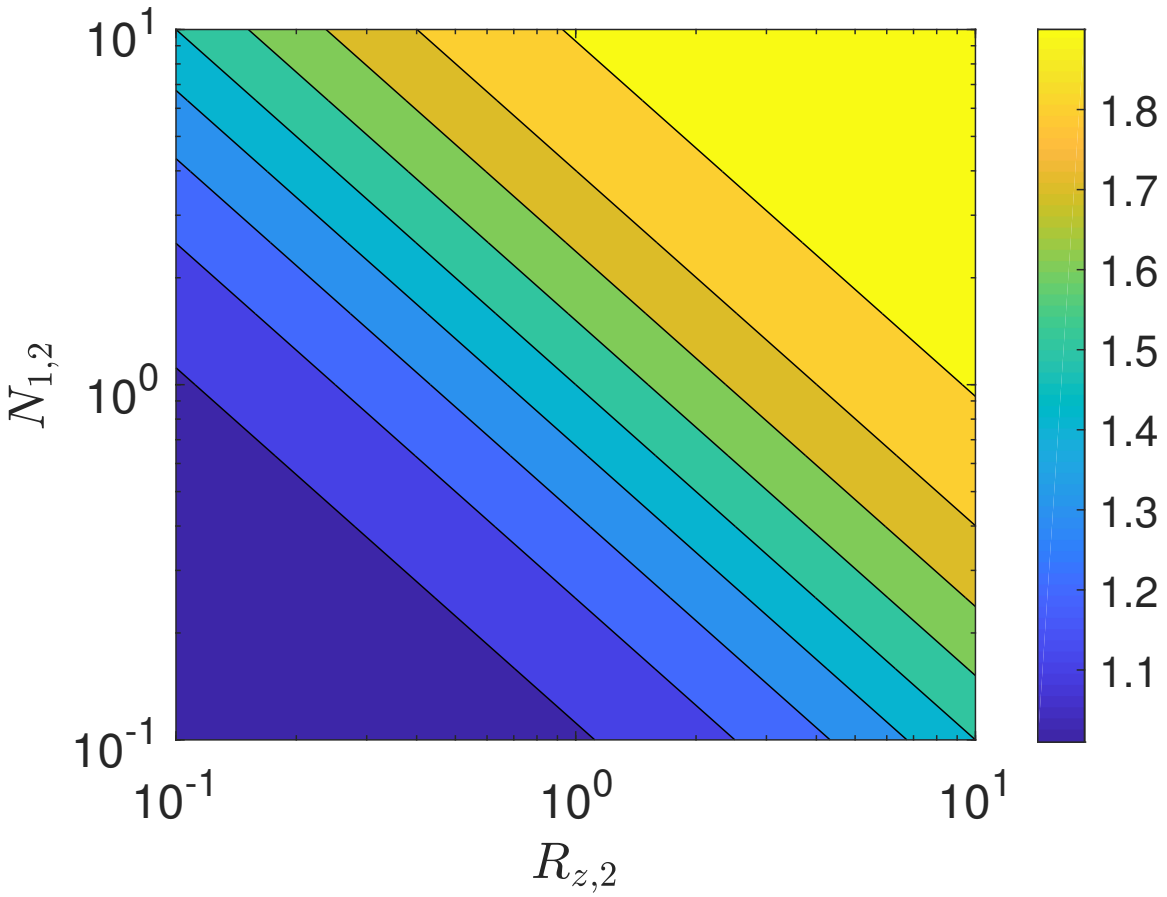



Figure 6.5: Example 6.1. Asymptotic output of the spring-mass-damper system with conflicting commands for various filter and performance weight choices. Note that the asymptotic output satisfies (6.11).

Next, the closed-loop response of the scramjet to conflicting commands is considered. The conflicting commands are assumed to be less than the critical thrust at the operating conditions. The scramjet is commanded to follow the step commands $r_1(k) = 0.24$ and $r_2(k) = 0.29$ for all $k > 0$. At $k = 0$, the scramjet is assumed to be operating in steady state with $\phi = 0.2$ at the inlet Mach number $M = 2.7$. RCAC is implemented with $N_1 = [0.1\ 0.1]^{\mathrm{T}}$, $\lambda = 0.9999$, $R_\theta = 10^7 I_8$, and $R_z = I_2$. Figure 6.6 shows the closed-loop response of the scramjet to the conflicting commands. Note that the asymptotic value of the generated thrust is in the middle of the commanded

values since $N_{1,1}R_{z,11} = N_{1,2}R_{z,22}$. Figure 6.7(a) shows the generated thrust with various filter choices. Note that the assymptotic value of the generated thrust can be arbitrarily chosen by varying the entries of $N_1$ and $R_z$. Figure 6.7(b) shows the correspondng values of equivalence ratios.

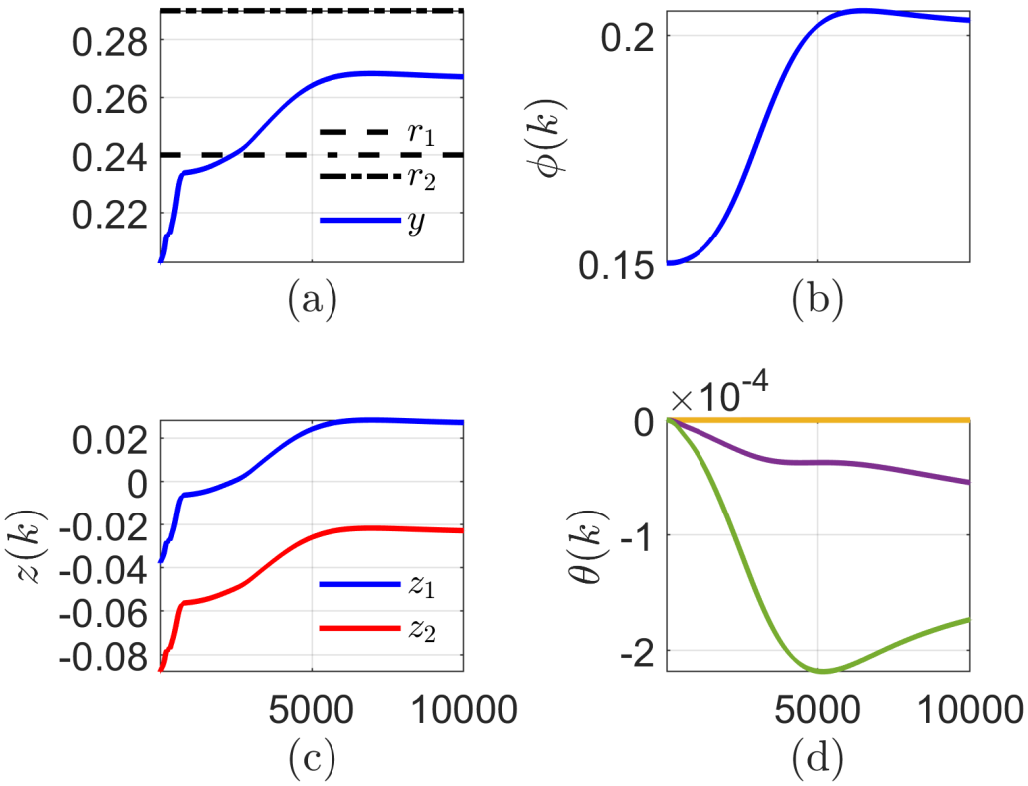
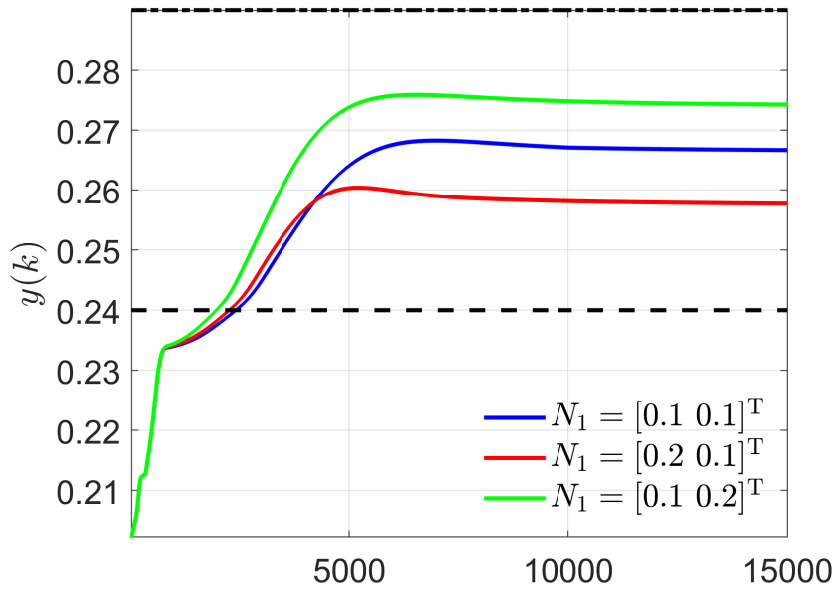
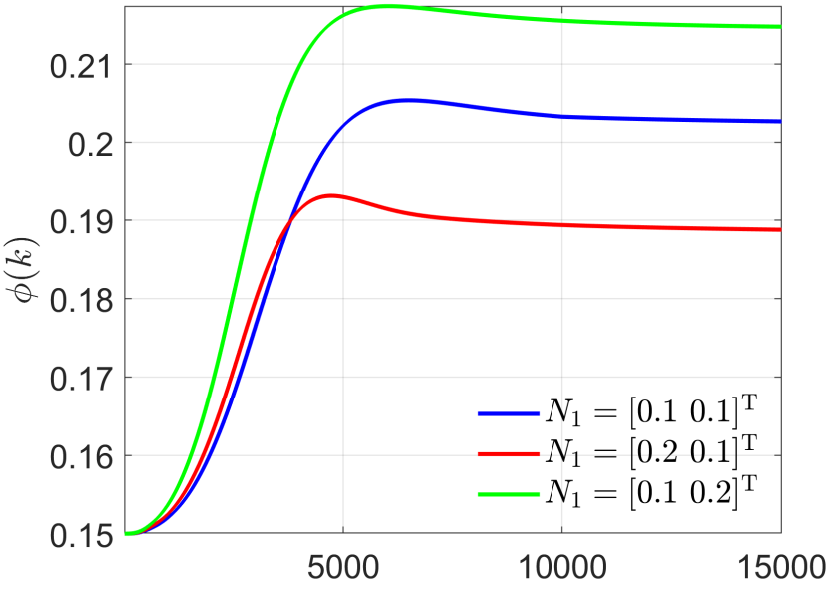



Figure 6.6: Closed-loop response of the scramjet to conflicting commands. (a) shows the commands and the output, (b) shows the equivalence ration $\phi$, (c) shows the output error $z$, and (d) shows the controller coefficients $\theta$.




(a) Thrust




(b) Equivalence ratio

Figure 6.7: Closed-loop response of the scramjet to conflicting commands for various filter choices. (a) shows the commands and the output, and (b) shows the equivalence ration $\phi$.

## 6.6 Adaptive Command Following with Auxiliary Output Constraints

Consider the system (6.5), (6.6), with auxiliary outputs

$$y_{\rm c}(k) = h_{\rm c}(x(k), u(k)), \tag{6.13}$$

where $y_{\rm c}(k) \in \mathbb{R}^q$ is an auxiliary measured output to be constrained. Let $r(k)$ be the reference command. The objective is to construct a control law such that $y(k) - r(k) \to 0$ as $k \to \infty$ while enforcing constraints of the form $y_{\rm c}(k) \in \mathcal{Y}_{\rm c}$ for all $k \geq 0$, where $\mathcal{Y}_{\rm c} \subset \mathbb{R}^q$. For all $i = 1, \ldots, q$, the constraint on $y_{{\rm c}i}(k)$ is assumed to be given by

$$\bar{y}_{{\rm c}i}^{\rm l} \leq y_{{\rm c}i}(k) \leq \bar{y}_{{\rm c}i}^{\rm u}, \tag{6.14}$$

where $\bar{y}_{{\rm c}i}^{\rm l}, \bar{y}_{{\rm c}i}^{\rm u}$ are the lower and the upper *constraint bounds*.

Now, consider the performance variable

$$z(k) \triangleq \begin{bmatrix} R_1 d(y_{\rm c1}(k), \bar{y}_{\rm c1}^{\rm l}, \bar{y}_{\rm c1}^{\rm u}) \\ \vdots \\ R_q d(y_{{\rm c}q}(k), \bar{y}_{{\rm c}q}^{\rm l}, \bar{y}_{{\rm c}q}^{\rm u}) \\ y(k) - r(k) \end{bmatrix}, \tag{6.15}$$

where $R_i \geq 0$ and the two-sided *constraint violation metric* is defined by

$$d(y_{{\rm c}i}(k), \bar{y}_{{\rm c}i}^{\rm l}, \bar{y}_{{\rm c}i}^{\rm u}) \triangleq \begin{cases} y_{{\rm c}i}(k) - \bar{y}_{{\rm c}i}^{\rm u}, & y_{{\rm c}i}(k) > \bar{y}_{{\rm c}i}^{\rm u}, \\ 0, & \bar{y}_{{\rm c}i}^{\rm l} \leq y_{{\rm c}i}(k) \leq \bar{y}_{{\rm c}i}^{\rm u}, \\ y_{{\rm c}i}(k) - \bar{y}_{{\rm c}i}^{\rm l}, & y_{{\rm c}i}(k) < \bar{y}_{{\rm c}i}^{\rm l}. \end{cases} \tag{6.16}$$

Note that, if $y(k) = r(k)$ and $\bar{y}_{ci}^{\mathrm{l}} \leq y_{ci}(k) \leq \bar{y}_{ci}^{\mathrm{u}}$ hold asymptotically, then the $i^{\mathrm{th}}$ constraint is said to be *satisfied*. Otherwise, the constraint is *violated*.

Let $N_1$ and $R_z$ be given by (6.8). Numerical examples show that, if $\lim_{k\to\infty} z(k)$ exists, then

$$\lim_{k\to\infty} \sum_{j=1}^{q+1} N_{1,j} R_{z,j} z_j(k) = 0. \tag{6.17}$$

Consequently,

$$\lim_{k\to\infty} z(k) \in \mathcal{N}(N_1^{\mathrm{T}} R_z). \tag{6.18}$$

Thus, by choosing $N_1$ and $R_z$, (6.17) can be used to specify the tradeoff between the output error and the constraint violation.

**Example 6.2.** Consider the linear time-invariant system

$$x(k+1) = Ax(k) + Bu(k), \tag{6.19}$$

$$y(k) = Cx(k), \tag{6.20}$$

where

$$A \triangleq \begin{bmatrix} 0.41 & 0.17 & 0.65 \\ 0.18 & 0.34 & 0.29 \\ 0.14 & 0.77 & 0.02 \end{bmatrix}, \quad B \triangleq \begin{bmatrix} 0.19 \\ 0.24 \\ 0.56 \end{bmatrix}, \quad C \triangleq [0.43 \;\; 0.61 \;\; 0.04]. \tag{6.21}$$

Let $r = 1$. At steady state, the equilibrium state $x_{\mathrm{eq}}$ is given by

$$x_{\mathrm{eq}} = (I - A)^{-1} B (C(I - A)^{-1} B)^{-1} r = [1.18 \;\; 0.75 \;\; 0.84]^{\mathrm{T}}, \tag{6.22}$$

so that $Cx_{\mathrm{eq}} = r = 1$. It follows that $x(k) \to x_{\mathrm{eq}}$ if $y(k) \to r$ as $k \to \infty$.

Consider the auxiliary output

$$y_{\mathrm{c}}(k) = x_3(k), \tag{6.23}$$

and consider the constraint given by $\bar{y}_{\mathrm{c}}^{\mathrm{u}} = 0.5$. Thus, if $r = 1$, it follows from (6.22) that the constraint is violated in the case where $y(k) \to r$ as $k \to \infty$. RCAC is implemented with $N_1 = [5 \;\; 1]^{\mathrm{T}}$, $R_z = \mathrm{diag}(10, 1)$, $R_\theta = 10^2 I_{l_\theta}$, $R_1 = 2$, and $\lambda = 0.999$. It follows from (6.17) that

$$\lim_{k\to\infty} 100(y_{\mathrm{c}}(k) - \bar{y}_{\mathrm{c}}^{\mathrm{u}}) + y(k) - r = 0. \tag{6.24}$$

Figure 6.8 shows the states and the constrained output of the system (6.19), (6.20). Note that, asymptotically, the output error $z_2(k) = -0.4$, and the constraint violation metric $y_{\mathrm{c}}(k) - \bar{y}_{\mathrm{c}}^{\mathrm{u}} = 0.004 = -0.01 z_2(k)$ satisfy (6.24). $\diamond$

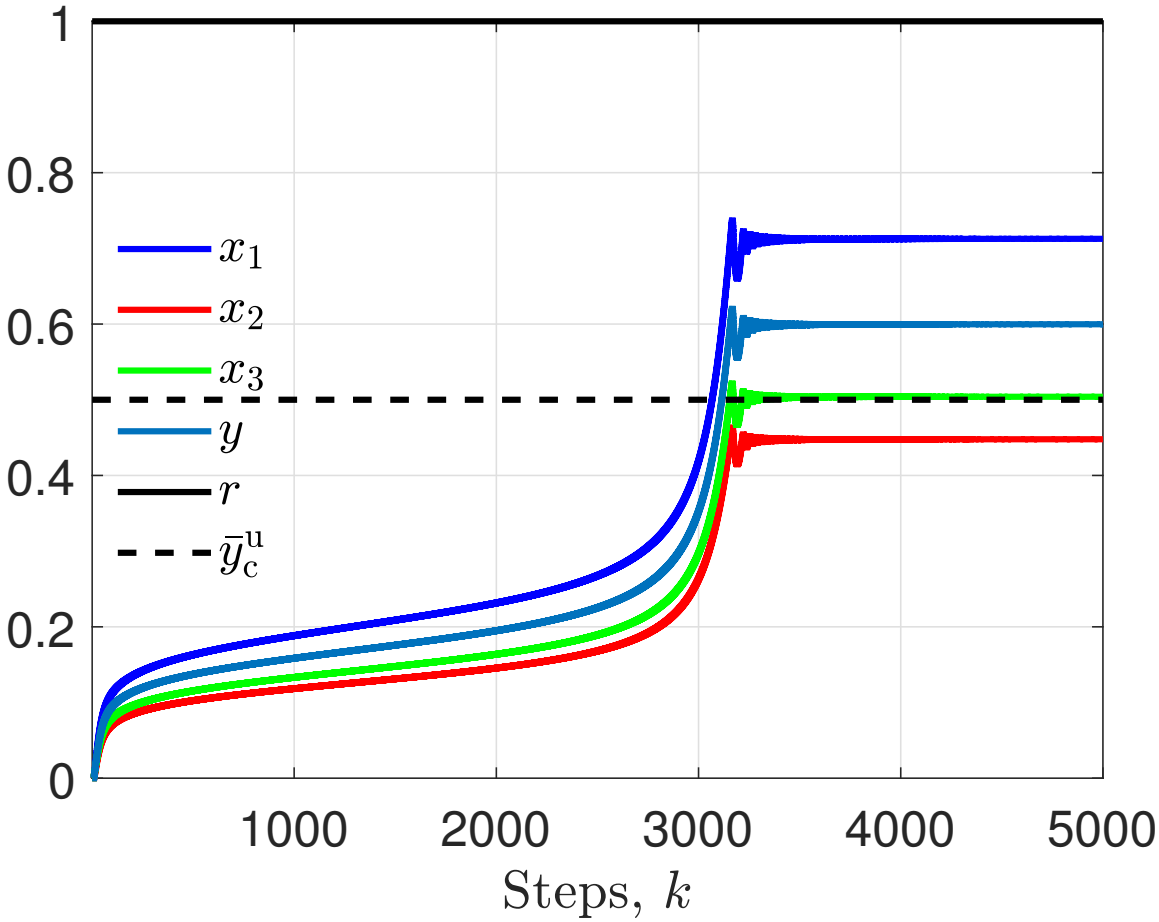



Figure 6.8: Example 6.2. Closed-loop response of the system (6.19), (6.20) with auxiliary output constraints. Note that the constraint is violated if the command is followed. Thus, asymptotically, the output error $y(k)-r(k)$, and the constraint violation metric $y_c(k) - \bar{y}_c^u$ satisfy (6.24).

Next, the closed-loop response of the scramjet with the auxiliary output constraint is considered, where the auxiliary output is the pressure metric $p_m$ defined in Section 6.3. Specifically, the constraint is given by

$$p_m(k) \leq 0.4. \tag{6.25}$$

Note that $p_m = 0.4325$ at the critical thrust $M_0 = 2.7$. The scramjet is commanded

to follow the step command

$$r(k) = \begin{cases} 0.32, & k < 8000, \\ 0.355, & k \geq 8000. \end{cases} \tag{6.26}$$

At $k = 0$, the scramjet is assumed to be operating in steady state with $\phi_0 = 0.3$ at the inlet Mach number $M = 2.7$. RCAC is implemented with $N_1 = [1\ 1]^{\mathrm{T}}$, $R_z = I_2$, , $R_\theta = 10^7 I_8$, $R_1 = 1$, and $\lambda = 0.9999$.

For $k < 8000$, the commanded thrust is less than the critical thrust at the operating conditions, and thus the scramjet does not unstart and reaches a steady state. Note that $p_{\mathrm{m}} < 0.4$ for $k < 9000$, and thus the constraint is satisfied. For $k \geq 8000$, the commanded thrust is greater than the critical thrust at the operating conditions, thus the scramjet begins to unstart as the controller increases the equivalence ratio to increase the generated thrust. However, the constraint eventually becomes active, the scramjet converges to a steady state, and the output error and the constraint violation metric converge to constant values based on the choice of $N_1$, $R_z$, and $R_1$.

Figure 6.9 shows the closed-loop response of the scramjet with auxiliary output constraint, (a) shows the measured output $y(k)$, (b) shows the equivalence ratio $\phi(k)$, (c) shows the constraint violation metric $z_1(k)$ and the output error $z_2(k)$, and (d) shows the coefficients of the controller optimized by RCAC.

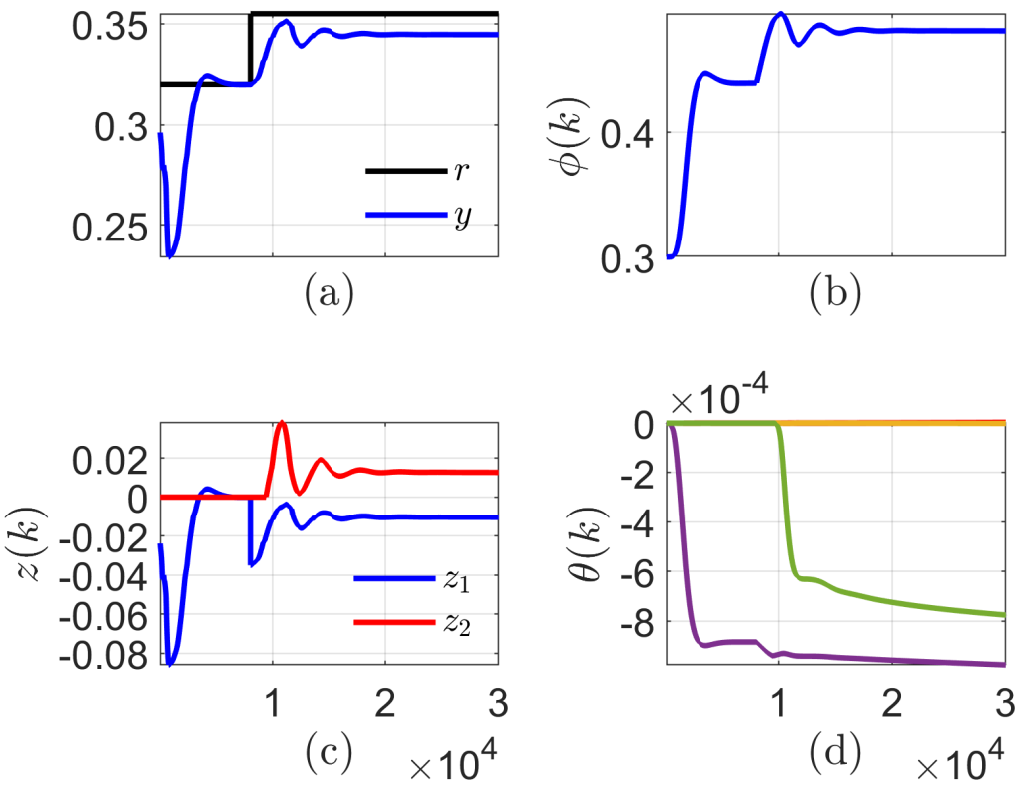



Figure 6.9: Closed-loop response of the scramjet with the auxiliary output constraint. (a) shows the commands and the output, (b) shows the equivalence ration $\phi$, (c) shows the output error $z$, and (d) shows the controller coefficients $\theta$.

## 6.7 Conclusions

In this chapter, retrospective cost adaptive control (RCAC) with auxiliary output constraints was used to prevent a 2D scramjet combustor from unstarting. As a first step towards enforcing output constraints, RCAC was applied to the problem of following conflicting setpoint commands. The asymptotic response of the system to the conflicting commands was shown to be directly related to the filter choice

and the performance weights in RCAC. Finally, RCAC was extended to implement output constraints by defining the constraint violation metric and augmenting the performance variable. The asymptotic response of the system and the asymptotic constraint violation metric were shown to be directly related to the filter choice and the performance weights in RCAC.

RCAC does not prevent the constraint from being violated. Instead, RCAC balances the several conflicting objectives by trading the cost between the output error and the constraint violation metric. This tradeoff can be designed by the choice of the filter choice and the performance weights in RCAC.

RCAC was implemented to follow conflicting commands and enforce auxiliary output constraints in low-order linear systems to numerically explore the asymptotic response. Finally, RCAC was applied to prevent the scramjet from unstarting by using several pressure measurements as the auxiliary output.

The upper bound of the pressure metric used to prevent unstart used in this work is based on open-loop simulations of the scramjet, which are chosen to be sufficiently far from unstart.

# CHAPTER 7

# Conclusions and Future Work

This dissertation first presented retrospective cost parameter estimation algorithm, which is an iterative, data-driven technique for estimating unknown parameters in linear and nonlinear dynamical systems. Unlike the extended Kalman filter, RCPE is gradient-free; unlike the unscented Kalman filter, RCPE is ensemble-free; and unlike variational methods, RCPE is adjoint-free.

The parameter estimator system consisted of an adaptive integrator with an output nonlinearity and a permutation. It was shown that to estimate multiple unknown parameters, a permutation of the estimated paramters is required to correctly associate each parameter estimate with the corresponding unknown parameter. The need to select the permutation matrix is the price paid for not requiring gradient information, an ensemble of models, or an adjoint model. In particular, it was shown that for a choice filter coefficients that define the retrospective cost, there is exactly one correct choice of permutation matrix. In fact, the parameter estimate diverges with all other choices of the permutation matrix. In addition, it was shown that the RCPE estimate is constrained to lie in a subspace defined by filter coefficients that define the retrospective cost.

The potential usefulness of RCPE was demonstrated by application to the problem of estimating two unknown parameters in the Burgers equation, for which UKF

requires an ensemble of 207 models each with 101 states, requiring a total of 21,321 state updates. In contrast, RCPE required one model and testing of two permutations, which requires a total of 202 state updates (101 for each permutation).

Next, it was shown that RCPE can implemented with simultaneous optimization of the filter coefficients. The resulting optimization problem is biquadratic and non-convex, however it is convex in the filter coefficients and the parameter estimator coefficients separately. Thus, an alternating convex search algorithm was used to optimize the retrospective cost function. The extended RCPE is applied to the problem of estimating eddy diffusion coefficient in global ionosphere-thermosphere model using the measurements of total electron content.

Several key issues remain unexplored in RCPE. First, the number of permutation matrices that must be tested is potentially $l_\mu!$, which is prohibitive when $l_\mu$ is large. Fortunately, divergence of the parameter estimates can be used to rule out incorrect permutation matrices. An online technique based on the performance metric may be developed to determine the correct permutation matrices. Next, RCPE is based on RLS with a constant, scalar forgetting factor, and the performance of RCPE tends to be sensitive to the choice of $\lambda$. Variable-rate forgetting technique can be implemented within RCPE to optimize the retrospective cost. Next, all of the examples considered in this dissertation involve asymptotically stable dynamics. Preliminary results suggest that RCPE is effective for systems that are Lyapunov stable; however, in this case, the performance of RCPE is sensitive to the weighting parameters. Better understanding of this case is needed to mitigate the sensitivity.

Finally, this dissertation did not consider the problem of uncertainty propagation in parameter estimation. It maybe possible to compute the parameter error covariance in RCPE by using ensemble methods or analytical expressions. Furthermore, methods such as unscented transformation, quadrature formulas maybe used to propagate uncertainty separately. Finally, the independent propagation of state error and

parameter error covariance may be compared with Monte-Carlo based methods to propagate uncertainty.

Next, this dissertation focused on forgetting in the context of recursive least squares algorithm. Note that the minimizer of the retrospective cost in RCPE is given by the RLS algorithm. It was shown that in the case where the excitation is not persistent and forgetting factor is less than one, singular values of $P_k$ corresponding to the singular vectors orthogonal to the information subspace diverge. A data-dependent forgetting matrix scheme was developed to restrict forgetting to the information subspace, and prevent $P_k$ from diverging under lack of persistent excitation.

Next, the dissertation focused on the problem of control allocation in wide systems using RCAC. The approach is based on the fact that the control input is constrained to the range of the transpose of the target model. This technique was demonstrated using numerical examples. The illustrative examples that were considered were minimum phase, and thus it was not necessary to account for the presence of NMP zeros. Although NMP zeros are nongeneric in nonsquare plants, they arise generically in square systems (with SISO plants as the special case $1 \times 1$). It is thus of interest to consider control allocation for NMP square systems. The challenge is thus to enforce control allocation by choosing the filter coefficients to also account for the presence of NMP zeros and simultaneously allowing the constraining the subspace defined by the filter coefficients.

Finally, the dissertation presented a technique to asymptotically enforce auxiliary output constraints in RCAC. In particular, the problem of enforcing auxiliary output constraints was formulated as a problem of following conflicting commands. The asymptotic response of the system to the conflicting commands was shown to be directly related to the filter choice and the performance weights in RCAC. Finally, RCAC was extended to implement output constraints by defining the constraint vio-

lation metric and augmenting the performance variable. The asymptotic response of the system and the asymptotic constraint violation metric were shown to be directly related to the filter choice and the performance weights in RCAC.

It was shown that RCAC does not prevent the constraint from being violated. Instead, RCAC balances the several conflicting objectives by trading the cost between the output error and the constraint violation metric. This tradeoff can be designed by the choice of the filter choice and the performance weights in RCAC.

The constraint violation metric was defined using a non-smooth nonlinear function. Future research should focus on using smooth nonlinear functions to construct the constraint violation metric to improve the transient behaviour of the closed loop system.

# BIBLIOGRAPHY

# BIBLIOGRAPHY


[1] L. Ljung, *System Identification: Theory for the User*, 2nd ed., ser. Prentice Hall Information and System Sciences Series. Prentice Hall, 1 1999.

[2] D. Morrision, "Methods for non-linear least squares problems and convergene proofs, tracking programs and orbit determination," in *Proceedings of the Jet Propulsion Laboratory Seminar*, 1960, pp. 1–9.

[3] G. Booth, G. Box, M. Muller, and T. Peterson, "Forecasting by generalized regression methods, nonlinear estimation,(princeton-ibm)," *Inter-national Business Machines Corporation, New York*, 1959.

[4] D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 2005. [Online]. Available: http://www.siam.org/journals/ojsa.php

[5] T. S. Jaakkola and M. I. Jordan, "Bayesian Paramater Estimation via Variational Methods," *Statistics and Computing*, vol. 10, no. 10, pp. 25–37, 2000. [Online]. Available: http://link.springer.com/10.1023/A:1008932416310

[6] R. L. Raffard, K. Amonlirdviman, J. D. Axelrod, and C. J. Tomlin, "An adjoint-based parameter identification algorithm applied to planar cell polarity signaling," *IEEE Transactions on Automatic Control*, vol. 53, pp. 109–121, 2008.

[7] M. Eknes and G. Evensen, "Parameter estimation solving a weak constraint variational formulation for an Ekman model," *Journal of Geophysical Research: Oceans*, vol. 102, pp. 12 479–12 491, 1997.

[8] R. M. Errico, "What is an adjoint model?" *Bulletin of the American Meteorological Society*, vol. 78, no. 11, pp. 2577–2592, 1997.

[9] I. M. Navon, "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography," *Dynamics of atmospheres and oceans*, vol. 27, no. 1-4, pp. 55–79, 1998.

[10] D. Kondrashov, C. Sun, and M. Ghil, "Data assimilation for a coupled ocean–atmosphere model. part ii: Parameter estimation," *Monthly Weather Review*, vol. 136, no. 12, pp. 5062–5076, 2008.

[11] Y. Zhu and I. M. Navon, "Impact of parameter estimation on the performance of the fsu global spectral model using its full-physics adjoint," *Monthly Weather Review*, vol. 127, no. 7, pp. 1497–1517, 1999.

[12] Z. Yang and J. M. Hamrick, "Variational inverse parameter estimation in a cohesive sediment transport model: An adjoint approach," *Journal of Geophysical Research: Oceans*, vol. 108, no. C2, 2003.

[13] G. Mercere, O. Prot, and J. A. Ramos, "Identification of parameterized gray-box state-space systems: From a black-box linear time-invariant representation to a structured one," *IEEE Transactions on Automatic Control*, vol. 59, pp. 2873–2885, 2014.

[14] C. Yu, L. Ljung, and M. Verhaegen, "Identification of structured state-space models," *Automatica*, vol. 90, pp. 54–61, 4 2018.

[15] A. Gelman, H. S. Stern, J. B. Carlin, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. Chapman and Hall/CRC, 2013.

[16] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013, vol. 3.

[17] H. White, "Maximum likelihood estimation of misspecified models," *Econometrica: Journal of the Econometric Society*, pp. 1–25, 1982.

[18] Z. Ghahramani and G. E. Hinton, "Parameter estimation for linear dynamical systems," Technical Report CRG-TR-96-2, University of Totronto, Dept. of Computer Science, Tech. Rep., 1996.

[19] R. Van der Merwe and E. A. Wan, "The square-root unscented Kalman filter for state and parameter-estimation," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6. IEEE, 2001, pp. 3461–3464.

[20] L. Ljung, "Asymptotic behavior of the Extended Kalman filter as a parameter estimator for linear systems," *IEEE Transactions on Automatic Control*, vol. 24, pp. 36–50, 2 1979.

[21] G. L. Plett, "Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs - Part 3. State and parameter estimation," *Journal of Power Sources*, vol. 134, pp. 277–292, 2004.

[22] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *Adaptive Systems for Signal Processing, Communications, and Control Symposium*. IEEE, 2000, pp. 153–158.

[23] G. Evensen, "The ensemble Kalman filter for combined state and parameter estimation," *IEEE Control Systems Magazine*, vol. 29, pp. 83–104, 6 2009.

[24] H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser, "Dual state and parameter estimation of hydrological models using ensemble Kalman filter," *Advances in Water Resources*, vol. 28, pp. 135–147, 2005.

[25] R. Madankan, P. Singla, T. Singh, and P. D. Scott, "Polynomial-chaos-based Bayesian approach for state and parameter estimations," *Journal of Guidance, Control, and Dynamics*, vol. 36, pp. 1058–1074, 2013.

[26] J. L. Anderson, "An ensemble adjustment kalman filter for data assimilation," *Monthly weather review*, vol. 129, no. 12, pp. 2884–2903, 2001.

[27] A. M. D'Amato, A. J. Ridley, and D. S. Bernstein, "Retrospective-cost-based adaptive model refinement for the ionosphere and thermosphere," *Statistical Analysis and Data Mining*, vol. 4, pp. 446–458, 8 2011.

[28] X. Zhou, D. S. Bernstein, J. L. Stein, and T. Ersal, "Battery state of health monitoring by estimation of side reaction current density via retrospective-cost subsystem identification," *Journal of Dynamic Systems, Measurement, and Control*, vol. 139, p. 091007, 2017.

[29] A. Goel, K. Duraisamy, and D. S. Bernstein, "Parameter estimation in the burgers equation using retrospective-cost model refinement," in *Proc. Amer. Contr. Conf.*, Boston, MA, July 2016, pp. 6983–6988.

[30] Y. Rahman, A. Xie, and D. S. Bernstein, "Retrospective Cost Adaptive Control: Pole Placement, Frequency Response, and Connections with LQG Control," *IEEE Control System Magazine*, vol. 37, pp. 28–69, 10 2017.

[31] T. S. Jaakkola and M. I. Jordan, "Bayesian parameter estimation via Variational methods," *Statistics and Computing*, vol. 10, pp. 25–37, 2000.

[32] R. Bellman and K. J. Åström, "On structural identifiability," *Mathematical biosciences*, vol. 7, no. 3-4, pp. 329–339, 1970.

[33] M. Grewal and K. Glover, "Identifiability of linear and nonlinear dynamical systems," *IEEE Trans. Autom. Contr.*, vol. 21, pp. 833–837, 12 1976.

[34] S. Stanhope, J. E. Rubin, and D. Swigon, "Identifiability of linear and linear-in-parameters dynamical systems from a single trajectory," *SIAM Journal on Applied Dynamical Systems*, vol. 13, no. 4, pp. 1792–1815, 2014.

[35] I. M. Y. Mareels, R. R. Bitmead, M. Gevers, C. R. Johnson, R. L. Kosut, and M. A. Poubelle, "How exciting can a signal really be?" *Systems & Control Letters*, vol. 8, pp. 197–204, 1987.

[36] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[37] M. R. S. M. R. S. Kulenovic and G. E. Ladas, *Dynamics of second order rational difference equations : with open problems and conjectures.* Chapman & Hall/CRC, 2002.

[38] D. T. Blackstock, "Generalized burgers equation for plane waves," *The Journal of the Acoustical Society of America*, vol. 77, no. 6, pp. 2050–2053, June 1985.

[39] R. Courant, K. Friedrichs, and H. Lewy, "On the partial difference equations of mathematical physics," *IBM Journal of Research and Development*, vol. 11, pp. 215–234, march 1967.

[40] K. F. Tapping, "The 10.7 cm solar radio flux (f10.7)," *Space Weather*, vol. 11, no. 7, pp. 394–406, 2013.

[41] L. Qian, A. G. Burns, S. C. Solomon, and W. Wang, "Annual/semiannual variation of the ionosphere," *Geophys. Res. Lett.*, vol. 40, no. 10, pp. 1928–1933, 2013.

[42] T. Shimazaki, "Effective eddy diffusion coefficient and atmospheric composition in the lower thermosphere," *J. Atmos. Terr. Physics*, vol. 33, no. 9, pp. 1383–1401, 1971.

[43] P. M. Kintner, B. M. Ledvina, and E. R. de Paula, "Gps and ionospheric scintillations," *Space Weather*, vol. 5, no. 9, 2007, s09003. [Online]. Available: http://dx.doi.org/10.1029/2006SW000260

[44] J. A. Klobuchar, "Ionospheric time-delay algorithm for single-frequency gps users," *IEEE Transactions on aerospace and electronic systems*, no. 3, pp. 325–331, 1987.

[45] L. Liu, W. Wan, B. Ning, and M.-L. Zhang, "Climatology of the mean total electron content derived from gps global ionospheric maps," *J. Geophys. Res.: Space Phys.*, vol. 114, no. A6, 2009, a06308. [Online]. Available: http://dx.doi.org/10.1029/2009JA014244

[46] N. Jakowski, C. Mayer, M. M. Hoque, and V. Wilken, "Total electron content models and their use in ionosphere monitoring," *Radio Sci.*, vol. 46, no. 6, 2011, rS0D18. [Online]. Available: http://dx.doi.org/10.1029/2010RS004620

[47] J. L. Lean, R. R. Meier, J. M. Picone, and J. T. Emmert, "Ionospheric total electron content: Global and hemispheric climatology," *J. Geophys. Res.: Space Phys.*, vol. 116, no. A10, 2011, a10318. [Online]. Available: http://dx.doi.org/10.1029/2011JA016567

[48] L. Ciraolo and P. Spalla, "Comparison of ionospheric total electron content from the navy navigation satellite system and the gps," *Radio Sci.*, vol. 32, no. 3, pp. 1071–1080, 1997. [Online]. Available: http://dx.doi.org/10.1029/97RS00425

[49] B. Tsurutani, A. Mannucci, B. Iijima, M. A. Abdu, J. H. A. Sobral, W. Gonzalez, F. Guarnieri, T. Tsuda, A. Saito, K. Yumoto, B. Fejer, T. J. Fuller-Rowell, J. Kozyra, J. C. Foster, A. Coster, and V. M. Vasyliunas, "Global dayside ionospheric uplift and enhancement associated with interplanetary electric fields," *J. Geophys. Res.: Space Phys.*, vol. 109, no. A8, 2004, a08302. [Online]. Available: http://dx.doi.org/10.1029/2003JA010342

[50] J. Liu, B. Zhao, and L. Liu, "Time delay and duration of ionospheric total electron content responses to geomagnetic disturbances," *Annal. Geophys.*, vol. 28, no. 3, pp. 795–805, 2010. [Online]. Available: https://www.ann-geophys.net/28/795/2010/

[51] E. G. Thomas, J. B. H. Baker, J. M. Ruohoniemi, A. J. Coster, and S.-R. Zhang, "The geomagnetic storm time response of gps total electron content in the north american sector," *J. Geophys. Res.: Space Phys.*, vol. 121, no. 2, pp. 1744–1759, 2016, 2015JA022182. [Online]. Available: http://dx.doi.org/10.1002/2015JA022182

[52] A. G. Burrell, A. Goel, A. J. Ridley, and D. S. Bernstein, "Correction of the photoelectron heating efficiency within the global ionosphere-thermosphere model using retrospective cost model refinement," *J. Atmos. Solar-Terr. Phys.*, vol. 124, pp. 30–38, 2015.

[53] A. M. Morozov, A. A. Ali, A. M. D'Amato, A. J. Ridley, S. L. Kukreja, and D. S. Bernstein, "Retrospective-cost-based model refinement for system emulation and subsystem identification," in *Proc. Conf. Dec. Contr.*, Orlando, FL, December 2011, pp. 2142–2147.

[54] A. A. Ali, A. Goel, A. J. Ridley, and D. S. Bernstein, "Retrospective-cost-based adaptive input and state estimation for the ionosphere–thermosphere," *J. Aero. Inf. Sys.*, vol. 12, no. 12, pp. 767–783, 2015.

[55] Y. Rahman, A. Xie, and D. S. Bernstein, "Retrospective cost adaptive control: Pole placement, frequency response, and connections with lqg control," *IEEE Contr. Sys. Mag.*, vol. 37, 2017, october.

[56] ——, "A Tutorial and Overview of Retrospective-Cost-Based Adaptive Control," *Proc. Amer. Contr. Conf.*, pp. 3386–3409, July 2016.

[57] F. Sobolic and D. S. Bernstein, "Adaptive control with concurrent closed-loop identification of time-varying nonminimum-phase zeros," in *Proc. Conf. Dec. Contr.*, Las Vegas, NV, December 2016, pp. 371–376.

[58] F. M. Sobolic, A. Goel, and D. S. Bernstein, "Retrospective cost adaptive control using concurrent controller optimization and target-model identification," in *Proc. Amer. Contr. Conf.*, 2016, pp. 3416–3421.

[59] J. Gorski, F. Pfeuffer, and K. Klamroth, "Biconvex sets and optimization with biconvex functions: a survey and extensions," *Math. Meth. Oper. Res.*, vol. 66, no. 3, pp. 373–407, 2007.

[60] A. Ridley, Y. Deng, and G. Toth, "The global ionosphere–thermosphere model," *J. Atmos. Solar-Terr. Phys.*, vol. 68, no. 8, pp. 839–864, 2006.

[61] A. E. Hedin, "Extension of the msis thermosphere model into the middle and lower atmosphere," *J. Geophys. Res.: Space Phys.*, vol. 96, no. A2, pp. 1159–1172, 1991.

[62] D. Bilitza, "International reference ionosphere 2000," *Radio Sci.*, vol. 36, no. 2, pp. 261–275, 2001.

[63] M.-J. Yu and D. S. Bernstein, "Retrospective cost subsystem estimation and smoothing for linear systems with structured uncertainty," *Journal of Aerospace Information Systems*, pp. 1–19, October 2018.

[64] M. Grewal and K. Glover, "Identifiability of linear and nonlinear dynamical systems," *IEEE Trans. Autom. Contr.*, vol. 21, pp. 833–837, 12 1976.

[65] I. Y. Mareels and M. Gevers, "Persistence of excitation criteria," in *Proc. Conf. Dec. Contr*, 1986, pp. 1933–1935. [Online]. Available: http://ieeexplore.ieee.org/document/4049134/

[66] I. M. Mareels and M. Gevers, "Persistency of excitation criteria for linear, multivariable, time-varying systems," *Mathematics of Control, Signals, and Systems*, vol. 1, no. 3, pp. 203–226, 1988. [Online]. Available: http://link.springer.com/10.1007/BF02551284

[67] B. D. O. Anderson, "Adaptive systems, lack of persistency of excitation and bursting phenomena," *Automatica*, vol. 21, no. 3, pp. 247–258, 1985.

[68] G. Chowdhary and E. Johnson, "Concurrent learning for convergence in adaptive control without persistency of excitation," in *Proc. Conf. Dec. Contr.*, 2010, pp. 3674–3679.

[69] G. Chowdhary, M. Mühlegg, and E. Johnson, "Exponential parameter and tracking error convergence guarantees for adaptive controllers without persistency of excitation," *Int. J. Contr.*, vol. 87, no. 8, pp. 1583–1603, 2014.

[70] S. Aranovskiy, A. Bobtsov, R. Ortega, and A. Pyrkin, "Performance enhancement of parameter estimators via dynamic regressor extension and mixing," *IEEE Trans. Autom. Contr.*, vol. 62, no. 7, pp. 3546–3550, 2017.

[71] P. Panda, J. M. Allred, S. Ramanathan, and K. Roy, "Learning to forget with adaptive synaptic plasticity in spiking neural networks," *J. Emerg. Selec. Top. Circ. Syst.*, vol. 8, no. 1, pp. 51–64, 2018.

[72] T. R. Fortescue, L. S. Kershenbaum, and B. E. Ydstie, "Implementation of Self-tuning Regulators with Variable Forgetting Factors," *Automatica*, vol. 17, no. 6, pp. 831–835, 1981.

[73] C. Paleologu, J. Benesty, and C. Silviu, "A Robust Variable Forgetting Factor Recursive Least-Squares Algorithm for System Identification," *IEEE Sig. Proc. Lett.*, vol. 15, pp. 597–600, 2008.

[74] S.-H. Leung and C. F. So, "Gradient-Based Variable Forgetting Factor RLS Algorithm in Time-Varying Environments," *IEEE Trans. Sig. Proc.*, vol. 53, no. 8, pp. 3141–3150, 2005.

[75] S. Song, J.-S. Lim, S. J. Baek, and K.-M. Sung, "Gauss Newton variable forgetting factor recursive least squares for time varying parameter tracking," *Electron. Lett.*, vol. 36, no. 11, pp. 988–990, 2000.

[76] D. J. Park, B. E. Jun, and K. J. H., "Fast tracking RLS algorithm using novel variable forgetting factor with unity zone," *Electron. Lett.*, vol. 27, no. 23, pp. 2150–2151, 1991.

[77] A. A. Ali, J. B. Hoagg, M. Mossberg, and D. S. Bernstein, "On the stability and convergence of a sliding-window variable-regularization recursive-least-squares algorithm," *Int. J. Adapt. Control Signal Process.*, vol. 30, pp. 715–735, 2016.

[78] R. M. Canetti and M. D. España, "Convergence analysis of the least-squares identification algorithm with a variable forgetting factor for time-varying linear systems," *Automatica*, vol. 25, no. 4, pp. 609–612, 1989.

[79] R. Kulhavỳ, "Restricted exponential forgetting in real-time identification," *Automatica*, vol. 23, no. 5, pp. 589–600, 1987.

[80] R. Kulhavỳ and M. Kárnỳ, "Tracking of slowly varying parameters by directional forgetting," *IFAC Proc.*, vol. 17, no. 2, pp. 687–692, 1984.

[81] G. Kreisselmeier, "Stabilized least-squares type adaptive identifiers," *IEEE Trans. Autom. Contr.*, vol. 35, no. 3, pp. 306–310, 1990.

[82] L. Cao and H. Schwartz, "Directional forgetting algorithm based on the decomposition of the information matrix," *Automatica*, vol. 36, no. 11, pp. 1725–1731, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0005109800000935

[83] G. Kubin, "Stabilization of the RLS algorithm in the absence of persistent excitation," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing*, 1988, pp. 1369–1372.

[84] S. Bittanti, P. Bolzern, and M. Campi, "Convergence and exponential convergence of identification algorithms with directional forgetting factor," *Automatica*, vol. 26, no. 5, pp. 929–932, 1990.

[85] ——, "Exponential convergence of a modified directional forgetting identification algorithm," *Systems & Control Letters*, vol. 14, no. 2, pp. 131–137, 1990.

[86] S. A. U. Islam and D. S. Bernstein, "Recursive Least Squares for Real-Time Implementation," *Contr. Sys. Mag.*, vol. 39, June 2019.

[87] R. M. Johnstone, C. R. Johnson, R. R. Bitmead, and B. D. O. Anderson, "Exponential convergence of recursive least squares with exponential forgetting factor," in *Proc. Conf. Dec. Contr.*, 1982, pp. 994–997.

[88] J. Benesty and T. Gänsler, "New Insights into the RLS Algorithm," *Eurasip Jour. App. Sig. Proc.*, no. 3, pp. 331–339, 2004. [Online]. Available: https://link.springer.com/content/pdf/10.1155%2FS1110865704310188.pdf

[89] W. M. Haddad and V. Chellaboina, *Nonlinear Dynamical Systems and Control: A Lyapunov-based Approach*. Princeton University Press, 2008. [Online]. Available: https://press.princeton.edu/titles/8700.html

[90] O. Härkegård and S. T. Glad, "Resolving actuator redundancy–optimal control vs. control allocation," *Automatica*, vol. 41, pp. 137–144, 2005.

[91] T. A. Johansen and T. I. Fossen, "Control allocationA survey," *Automatica*, vol. 49, pp. 1087–1103, 2013.

[92] W. C. Durham, "Constrained Control Allocation," *AIAA J. Guid. Contr. Dyn.*, vol. 16, pp. 717–725, 1993.

[93] D. F. Enns, "Control Allocation Approaches," in *Proc. Guid. Nav. Contr. Conf.*, Boston, August 1998, AIAA-98-4109.

[94] J. J. Burken, P. Lu, Z. Wu, and C. Bahm, "Two Reconfigurable Flight-Control Design Methods: Robust Servomechanism and Control Allocation," *AIAA J. Guid. Contr. Dyn.*, vol. 24, pp. 482–493, 2001.

[95] M. Bodson, "Evaluation of Optimization Methods for Control Allocation," *AIAA J. Guid. Contr. Dyn.*, vol. 25, pp. 703–711, 2002.

[96] Y. Rahman, A. Xie, and D. S. Bernstein, "Retrospective Cost Adaptive Control: Pole Placement, Frequency Response, and Connections with LQG Control," *IEEE Contr. Sys. Mag.*, vol. 37, pp. 28–69, October 2017.

[97] M. Brocanelli, Y. Gunbatar, A. Serrani, and M. A. Bolender, "Robust Control for Unstart Recovery in Hypersonic Vehicles," in *AIAA Guidance, Navigation, and Control Conference*, ser. Guidance, Navigation, and Control and Co-located Conferences. Reston, Virigina: American Institute of Aeronautics and Astronautics, 8 2012.

[98] J. R. Hutzel, *Scramjet Isolator Modeling and Control*. Air Force Institute of Technology: ProQuest Dissertations Publishing, 2011, uMI number 3481728.

[99] N. West, G. Papanicolaou, P. Glynn, and G. Iaccarino, "A Numerical Study of Filtering and Control for Scramjet Engine Flow," in *20th AIAA Computational Fluid Dynamics Conference*, ser. Fluid Dynamics and Co-located Conferences. American Institute of Aeronautics and Astronautics, 6 2011.

[100] A. Goel, K. Duraisamy, and D. S. Bernstein, "Retrospective Cost Adaptive Control of Unstart in a Model Scramjet Combustor," *AIAA Journal*, vol. 56, no. 3, pp. 1085–1096, 2018.

[101] M.-J. Yu, E. M. Atkins, I. Kolmanovsky, and D. Bernstein, "Trim-Commanded Adaptive Control for Waypoint-Defined Trajectory Following," in *AIAA Guidance, Navigation, and Control (GNC) Conference*. Reston, Virginia: American Institute of Aeronautics and Astronautics, aug 2013.

[102] M. K. Smart, N. E. Hass, and A. Paull, "Flight Data Analysis of the HyShot 2 Scramjet Flight Experiment," *AIAA Journal*, vol. 44, no. 10, pp. 2366–2375, 10 2006.

[103] Q. Wang, K. Duraisamy, J. J. Alonso, and G. Iaccarino, "Risk assessment of scramjet unstart using adjoint-based sampling methods," *AIAA journal*, vol. 50, no. 3, pp. 581–592, 2012.

[104] R. Pecnik, V. E. Terrapon, F. Ham, G. Iaccarino, and H. Pitsch, "Reynolds-averaged Navier-Stokes simulations of the HyShot II scramjet," *AIAA journal*, vol. 50, no. 8, pp. 1717–1732, 2012.

[105] K. Duraisamy and J. Alonso, "Adjoint based techniques for uncertainty quantification in turbulent flows with combustion," in *42nd AIAA Fluid Dynamics Conference and Exhibit*, 2012, p. 2711.

[106] C. Doolan and R. Boyce, "A Quasi-One-Dimensional Mixing and Combustion Code for Trajectory Optimisation and Design Studies," in *AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, no. AIAA 2008-2603, Dayton, Ohio, 2008, pp. 898–907.