The observational network analysis for atmospheric inverse modelling extended by emission rates

Dissertation
zur Erlangung des akademischen Grades Doktor der Naturwissenschaften
an der Fakultät für Mathematik und Naturwissenschaften
der Bergischen Universität Wuppertal

vorgelegt von
Xueran Wu
aus Beijing

Wuppertal, 2016

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20160414-142859-8
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20160414-142859-8]

Contents

Acknowledgements

This dissertation has been written during my HITEC fellowship at IEK-8 of Forschungszentrum Jülich. I gratefully acknowledge Prof. Andreas Wahner, the director of IEK-8 and Dr. Bärbel Köster, the managing director for granting me this valuable opportunity to study and work in Germany. It provides the opportunity to broaden my horizon on scientific research and experience a different culture.

At the same time, I sincerely thank my supervisors, Prof. Dr. Birgit Jacob and PD Dr. Hendrik Elbern, for their patient guidance and continuous support of my Ph.D project. Their dedication and enthusiasm for scientific research inspire me. Their wide knowledge enlightens me. Their constant encouragement is always the great strength keeping me going when I met difficulties. They provided me with so many valuable suggestions, helpful advice and important guidance during the last three years.

Besides, I express my thanks to the colleagues in the functional analysis group of University of Wuppertal and in Rhenish Institute for Environmental Research at the University of Cologne for the creation of positive working atmosphere. It was a great time for me to work with you.

Especially I thank my husband Long Teng for the proofreading of this thesis and his full support of me. I will never forget those frank discussions about the academic research and happy time in our life.

Finally, I would like to dedicate this thesis to my mother Chunyue Liu and father Peijing Wu. Thanks for giving me the unconditional support, encouragement and endless love in my lifetime.

List of Figures

List of Tables

Chapter 1

Introduction

Air quality and climate change are influenced by the fluxes of green house gases, reactive gas emissions and aerosols in the atmosphere (Intergovernmental panel on climate change, fifth assessment report). The temporal evolution of chemistry in the atmosphere is usually modelled by atmospheric chemistry transport models. The ability to quantify variables, yet hardly observable emission rates is a key problem to be solved for the analysis of atmospheric systems, and typically addressed by elaborate and costly field campaigns or permanently operational observation networks. Especially for chemistry transport or greenhouse gas models with high dependence on the emissions in the troposphere, the optimization of the initial state is no longer the only issue. This renders initial-value-only optimization by traditional data assimilation methods as insufficient. The observation of fluxes can be achieved by eddy covariance measurements, mounted on special towers. By logistic and cost reasons, only a very sparse network is globally available. The lack of ability to observe clearly surface emission fluxes directly with necessary accuracy is a major roadblock, hampering the progress in predictive skills of climate and atmospheric chemistry models.

In order to get the better estimates from the model with limited observations, efforts of optimization have been made including the emission rates by spatio-temporal data assimilation. A meanwhile classical task is greenhouse gas inversion, aiming at the estimation of carbon dioxide, methane, and nitrous oxide, from which a rich set of literature emerged. For example, in case of CO_2, Peters et al. [82] devised an ensemble data assimilation approach, approximating the covariance matrix without need to use an adjoint model version. Singular value decomposition (SVD) can help identifying the priorities of observations by detecting the fastest growing uncertainties. Singular vector analysis based on SVD was firstly introduced to numerical weather prediction by Lorenz [67]. In [12], [13] and [15], the singular vector analysis for high-dimensional meteorological models was shown to be feasible to determine the direction with the strong influence of observations. Daescu [26] exploited the error covariance sensitivity analysis, to finally assess the data impact on analysis and forecasts. Kang and Xu [57] applied a four-dimensional variation (4D-Var) system to Burgers' equation to optimize sensor deployment by maximizing observability using a gradient projection approach. Sandu et al. [86] determined the dominant model singular vectors to identify regions of maximal error growth, which are then candidate locations for optimized sensor placement. With focus

on efficiency, Daescu and Navon [27] presented a method, resting on only one additional adjoint model integration for measurement network optimization. Cioaca and Sandu [18] introduced a general framework to optimize a set of parameters controlling the 4D-Var data assimilation system, which includes means to identify erroneous data, observation accuracy and location. In a related paper Cioaca and Sandu [17] quantified the observation impact in terms of reduction of uncertainties of shallow water model state and other parameters. The first full chemical implementation of the 4D-variational method for reactive atmospheric chemistry initial values was introduced in [34]. Further, Elbern et al. [36] introduced the strong constraint of the diurnal profile shape of emission rates such that their amplitudes and initial values are the only uncertainty to be optimized, and then implemented it by 4D-Var inversion. This strong constraint approach is reasonable because the diurnal evolution sequence of emissions is typically much better known than the absolute amount of daily emissions. Moreover, several data assimilation strategies were designed to adjust ozone initial conditions and emission rates separately or jointly in [93]. Bocquet et al. introduced a straightforward extension of the iterative ensemble Kalman smoother in [10].

Furthermore, the choices of observation locations and control locations, which may be mutually dual problems, are of great importance for improving the estimation and designing control systems in various practical problems. Many researchers have focused on the study of finding the optimal locations of control hardware and observation instruments and different criteria of optimizing control locations were established, such as maximization of observability and controllability [53], [79], or minimizing the linear quadratic (LQ) regulator cost [77]. Geromel [42] successfully reformulated the LQ cost function into a convex optimization problem by mapping the locations of controller into zero-one vectors and expressed the solution of the classic LQ problem in terms of a Riccati equation. Morris [75] optimized controller locations of time-invariant systems on an infinite-time horizon in Hilbert spaces by solving an algebraic Riccati equation and showed the convergence of optimal controller locations of a sequence of approximated finite-dimensional systems. Further, the H_∞-optimal actuator location problem of time-invariant systems on an infinite-time horizon was considered in [58]. Besides, an algorithm [29] for the linear quadratic optimal problem of controller locations based on the convexity shown in [42] are introduced. Bensoussan [7] studied the optimal problem of n sensor locations with filtering on a finite-time interval for time-invariant systems.

However, in atmospheric chemistry, the better estimations of both the initial state and emission rates are not always sustained, based on appropriate observational network configurations. It may hamper the optimization by unbalanced weights between the initial state and emission rates, which can, in practice, even result in degraded simulations beyond the time interval with available observations. The ability to evaluate the suitability of an observational network to control chemical states and emission rates is a key qualification, which needs to be addressed. At the same time, it is also important to find the optimal locations of observation to improve the estimations of specified uncertain variables. The objective of this work is therefore the development of a method for a quantitative evaluation of the efficiency and sensitivity of observational networks and to study the optimal problem of control and observation locations. This thesis is orga-

nized as follows. In Chapter 2, we review the current popular data assimilation method, four-dimensional variation method, Kalman filter and smoother and their ensemble implementation. In Chapter 3 we present the current approach of the optimization of initial values and emission rates by 4D-Var and establish the dynamic model for emission rates with the constraint of their diurnal profiles and introduce the atmospheric transport model extended by emission rates.

In Chapter 4 based on the Kalman smoother, a quantitative assessment method on the efficiency of observation configurations is theoretically developed by the singular value decomposition in order to evaluate and balance the potential improvements of initial values and emission rates associated within the entire data assimilation window. Further, the ensemble based approach is derived to guarantee the feasibility of the approach. An elementary example based on a 3D advection-diffusion equation is given to illustrate this method. Here the sensitive parameters to specific observation networks can be identified and targeted by determining the directions and strength of maximum perturbation within a finite-time interval. Besides, we apply the singular vector analysis of observation networks to determine the apportionments of different emission sources.

In Chapter 5, starting with partial differential equations, we consider the optimal problem of control locations for time-varying systems on a finite-time horizon in Hilbert spaces. The existence of the optimal locations based on the linear-quadratic control for both deterministic and stochastic systems on a finite-time horizon is studied. In order to provide the feasibility to solve the optimal problems on infinite-dimensional systems in practice, we develop the conditions to guarantee the convergence of the minimal costs and optimal control location of a (sub)sequence of approximations in finite-dimensional space of the original time-varying system. In Chapter 6 the optimal location of observations for improving the estimation of the state at the final time, based on the Kalman filter, is considered as the dual problem to the LQ optimal problem of the control locations. In addition, the existence and convergence of optimal locations of observations for improving the estimation of the initial state, based on the Kalman smoother is discussed. The results obtained are applied to a linear advection-diffusion model extended by emission rates.

In Chapter 7 we summarize the results and contributions developed and discussed in this thesis and mention the outlooks of the future research on the uncertainties qualification and optimal problem of control and observation locations.

Chapter 2

Overview of Data Assimilation Approaches

The temporal evolution of chemistry in the atmosphere is usually modelled by atmospheric chemistry transport models. These model can be used to predict the future evolution of atmospheric chemical compounds driven by initial values. Unfortunately, the initial values always contain inaccuracies and uncertainties. In this case, a sequence of observations can be incorporated into the model to "correct" the initial values. However, in practice, the error-equipped observations have the insufficient spacial and temporal density. They usually produce the ill-posed assimilation problem and fail to engender the picture depicting the true chemical evolution of atmosphere. Data assimilation aims at providing the most possibly accurate estimates of model states by incorporating the prior information and observations. In this chapter, for the future convenience, we review the most popular data assimilation approaches in current. A great amount of literature about the approaches of data assimilation is available, we mainly refer to [37], [55], [56], [61], [68] and [69] in this chapter.

2.1 Four-dimensional variational data assimilation

Four dimensional variational technique is an assimilation algorithm to estimate variables by minimizing the difference between model states and observations over a given data assimilation window [61]. The objective function of minimization is defined according to the maximum likelihood criteria. It combines the model and observation information to estimate the uncertain parameters and propagates the information both forward and backward in time via the adjoint of models and discretization algorithms.

In order to describe the dynamic system, we first define transition matrices of dynamic systems.

Definition 2.1.1. *$M(t,s)$ is called the transition matrix for any time pair (t,s) if it satisfies*

1. *$M(t,t) = I$, I is the identity matrix,*
2. *$M(t,r)M(r,s) = M(t,s), \quad s \leqslant r \leqslant t,$*

3. $M(s,t) = M^{-1}(t,s)$.

We now consider the discrete-time dynamic system within a discrete-time interval $[t_0, \cdots, t_N]$

$$x(t_{i+1}) = M(t_{i+1}, t_i)x(t_i) + \varepsilon(t_i)$$

with the observation system

$$y(t_i) = H(t_i)x(t_i) + \nu(t_i),$$

where $x(t_i) \in \mathbb{R}^n$ is the state variable at time t_i, $y(t_i) \in \mathbb{R}^m$ is the observation vector at time t_i, $M(t_{i+1}, t_i)$, $i \in [0, \cdots, N-1]$ is the transition matrix, $\varepsilon(t_i)$ and $\nu(t_i)$ are random variables of Gaussian distributions with zero mean and the following covariance matrices

$$\begin{aligned}\text{cov}[\varepsilon(t_i), \varepsilon(t_j)] &= Q(t_i)\delta(t_i - t_j),\\ \text{cov}[\nu(t_i), \nu(t_j)] &= R(t_i)\delta(t_i - t_j),\end{aligned}$$

where δ is the Dirac delta function. We denote $\varepsilon(t_i) \sim \mathcal{N}(0, Q(t_i))$ and $\nu \sim \mathcal{N}(0, R(t_i))$.

The prior estimate of the model state $x(t)$ are usually assumed to be known and can be obtained from previous analysis, denoted by $x_b(t)$. The covariance of error of prior estimates is denoted by

$$P(t_0) = \mathbb{E}[(x(t_0) - x_b(t_0))(x(t_0) - x_b(t_0))^\top].$$

In order to evaluate the inaccuracy of initial estimate, the model and observations, the objective function is given by

$$\begin{aligned}J(x(t_0)) = &\frac{1}{2}(x(t_0) - x_b(t_0))^\top P^{-1}(t_0)(x(t_0) - x_b(t_0))\\ &+\frac{1}{2}\sum_{i=0}^{N}(y(t_i) - H(t_i)x(t_i))^\top R^{-1}(t_i)(y(t_i) - H(t_i)x(t_i))\\ &+\frac{1}{2}\sum_{i=0}^{N-1}(x(t_{i+1}) - M(t_{i+1}, t_i)x(t_i))^\top Q^{-1}(t_i)(x(t_{i+1}) - M(t_{i+1})x(t)).\end{aligned} \tag{2.1}$$

The minimization of (2.1) is termed as the weak constraint four-dimensional variational assimilation, see [89].

If we ignore the model error and consider the dynamic model

$$x(t_{i+1}) = M(t_{i+1}, t_i)x(t_i),$$

the corresponding cost function turns to be

$$\begin{aligned}J(x(t_0)) = &\frac{1}{2}(x(t_0) - x_b(t_0))^\top P^{-1}(t_0)(x(t_0) - x_b(t_0))\\ &+\frac{1}{2}\sum_{i=0}^{N}(y(t_i) - H(t_i)x(t_i))^\top R^{-1}(t_i)(y(t_i) - H(t_i)x(t_i)).\end{aligned} \tag{2.2}$$

The minimization of (2.2) is termed as the strong constraint four-dimensional variational assimilation [89]. In this thesis, we mainly focus on problems based on the perfect models. Thus, we abbreviate the strong constraint four-dimensional variational assimilation by four-dimensional variational assimilation or 4D-Var.

In order to find out the minimal solution of (2.2), we calculate its gradient with respect to $x(t_0)$ and obtain

$$\frac{\partial J(x(t_0))}{\partial x(t_0)} = P^{-1}(t_0)(x(t_0) - x_b(t_0)) + \sum_{i=0}^{N} H^{\top}(t_i)M^{\top}(t_i, t_0)R^{-1}(t_i)(H(t_i)x(t_i) - y(t_i)). \tag{2.3}$$

The Hessian matrix of (2.2) is given by

$$\frac{\partial^2 J(x(t_0))}{(\partial x(t_0))^2} = P^{-1}(t_0) + \sum_{i=0}^{N} H^{\top}(t_i)M^{\top}(t_i, t_0)R^{-1}(t_i)H(t_i)M^{\top}(t_i, t_0). \tag{2.4}$$

It is clear that the Hessian matrix (2.4) is always positive-definite. It indicates that if there exists $\hat{x}(t_0)$ such that

$$\frac{\partial J(x(t_0))}{\partial x(t_0)} = 0,$$

$\hat{x}(t_0)$ is the minimum of (2.2).

The gradient (2.3) can be calculated firstly by propagating the model forward within the time interval $[t_0, \cdots, t_N]$ and then by the backward integration to the initial time via the adjoint model. It gives an access to obtain an numerically accurate minimal solution of the cost function.

2.2 Kalman filter and smoother in finite-dimensional spaces

Combining the information from the evolution of models with a sequence of observations associated with models, the Kalman filter and smoother, refer to [55],[56], [68] and [69], are recursive estimators to provide the best linear unbiased estimates (BLUE) of the unknown variables of models and the statistical description of the uncertainties based on the sequence of observations over time with inaccuracies. Here $\hat{x}$, the best linear unbiased estimate of a variable x, is of properties that

$$\mathbb{E}(\hat{x}) = \mathbb{E}(x)$$

and $\mathrm{cov}(x - \hat{x})$ is minimized by the certain norm. In fact, since 1960's, the Kalman filter and smoother have been widely applied in many fields including in meteorology, to produce optimal linear estimations of states and parameters. In the following, we summarize the main expressions of the Kalman filter and smoother for both continuous-time and discrete-time systems.

2.2.1 KF and KS for continuous-time systems

In this section we consider the continuous-time system within the time interval $[t_0, t_N]$,

$$\begin{aligned} x(t) &= M(t, t_0)x(t_0) + \varepsilon(t), \quad t_0 \leqslant t \leqslant t_N \\ y(t) &= H(t)x(t) + \nu(t), \end{aligned}$$

where $x(t) \in \mathbb{R}^n$ is the state variable at time t, $y(t) \in \mathbb{R}^m$ is observation vector at time t, $M(t, t_0)$ is the transition matrix, $\varepsilon(t)$, $\varepsilon(s)$, $\nu(t)$ and $\nu(s)$, $t \neq s$ are independent and $\varepsilon(t) \sim \mathcal{N}(0, Q(t))$ and $\nu(t) \sim \mathcal{N}(0, R(t))$.

In the context of the continuous-time Kalman filter and smoother, we denote the estimation of $x(t)$ based on $Y_\tau^c = \{y(t_o), t_o \in [t_0, \tau]\}$ by $\hat{x}(t|\tau)$, In addition, it can be found in several references, for example, [16], [39] and [68], that for any time t, τ, the BLUEs of $x(t)$ based on Y_τ can be generally written as the conditional expectation of $x(t)$ based on Y_τ^c, which is denoted by

$$\hat{x}(t|\tau) = \mathbb{E}[x(t)|Y_\tau^c].$$

For the continuous-time system above, the Kalman filter, also called the Kalman-Bucy filter was introduced by [56]. It aims at finding out the BLUE of $x(t)$ based on $Y_t^c = \{y(t_o), t_o \in [t_0, t]\}$, the observations until time t.

Concerning with the Kalman-Bucy filter, we term $\hat{x}(t|t)$ as the analysis estimate of the state $x(t)$ and $x(t|\tau)$, $\tau \leqslant t$ as the forecast estimate of the state. Correspondingly, $P(t|t)$ and $P(t|\tau)$ are the analysis and forecast covariance matrices of $\hat{x}(t|t)$ and $\hat{x}(t|\tau)$ respectively. According to [56], we summarize the main results of Kalman-Bucy filters in the integral form as follows:
(1) Analysis step:

$$\begin{aligned} K(t) &= P(t|t)H^\top(t)R^{-1}(t), \\ \hat{x}(t|t) &= \hat{x}(t|\tau) + \int_\tau^t M(t,s)K(s)(y(s) - H(s)\hat{x}(s|s))ds, \\ P(t|t) &= M(t, t_0)P(t_0|t_{-1})M^\top(t, t_0) \\ &\quad + \int_{t_0}^t M(t,s)[Q(s) - P(s|s)H^\top(s)R^{-1}(s)H(s)P(s|s)]M^\top(t,s)ds. \end{aligned} \tag{2.5}$$

(2) Forecasting step:

$$\begin{aligned} \hat{x}(t|\tau) &= M(t,\tau)\hat{x}(\tau|\tau), \\ P(t|\tau) &= M(t,\tau)P(\tau|\tau)M^\top(t,\tau) + Q(t). \end{aligned}$$

Differing from the filtering problem, the objective of the smoothing problem is to obtain the BLUE of $x(t)$ based on $Y_\tau^c = \{y(t_o), t_o \in [t_0, \tau], \tau \geqslant t\}$. It contains more observation information such that the estimation of states can be further improved. Various kinds of smoothers are exploited to solve the realistic problems, where three classes of smoothers have been widely used.

1. Fixed-interval smoother: With a fixed continuous-time interval $[t_0, t_N]$, it utilizes the observations within the entire time interval to provide the BLUE of $x(t)$, denoted by $\mathbb{E}[x(t)|Y_{t_N}^c]$, $t \in [t_0, t_N]$. There are several algorithms developed for the fixed-interval smoother. For example, Mayne [70] derived the fixed-interval smoother by combining the BLUE of the state by the Kalman filter with the optimal estimates from the future observations. Rauch, Tung and Striebel (RTS) smoother [83] is developed by combining the backward filter with the smoothing step into one recursive process. In this thesis we consider the data assimilation window as the fixed time interval and develop the approaches and theorems based on the fixed-interval smoother.

2. Fixed-point smoother: It is usually applied to estimate the state at a specific time $\tilde{t}$ by the observations within a certain time interval $[t_0, t_N]$. The estimator can be also represented by $\mathbb{E}[x(\tilde{t})|Y_{t_N}^c]$. Compared with the fixed-interval smoother, if $t = \tilde{t}$, the estimates of both smoothers are equivalent to each other.

3. Fixed-lag smoother: This smoother is designed to seek the BLUE of the state at time t through the observations from the initial time to the time $t + T$. Here T is a constant. Generally, the estimator of $x(t)$ of the fixed-lag smoother can be denoted by $\mathbb{E}[x(t)|Y_{t+T}^c]$.

2.2.2 KF and KS for discrete-time systems

In this section we consider the discrete-time system:

$$\begin{aligned} x(t_{k+1}) &= M(t_{k+1}, t_k)x(t_k) + \varepsilon(t_k), \\ y(t_k) &= H(t_k)x(t_k) + \nu(t_k), \end{aligned} \tag{2.6}$$

where $x(\cdot) \in \mathbb{R}^n$ is the state variable, $y(t_k) \in \mathbb{R}^{m(t_k)}$ is the observation vector, the model error $\varepsilon(t_k)$ and the observation error $\nu(t_k)$, $k = 1, \cdots, N$ follow Gaussian distributions with zero mean and the covariance matrices $Q(t_k)$ and $R(t_k)$ respectively.

For the discrete-time Kalman filter, we term $x(t_k|t_k)$ as the analysis estimate and $x(t_k|t_{k-1})$ as the forecasting estimate. Besides, $P(t_k|t_k)$ and $P(t_k|t_{k-1})$ are the corresponding analysis and forecasting covariance matrices. For the future convenience, the main results of the discrete-time Kalman filter are summarized as follows [55]:
(1) Analysis step:

$$\begin{aligned} &K(t_k) = P(t_k|t_{k-1})H^\top(t_k)(H(t_k)P(t_k|t_{k-1})H^\top(t_k) + R(t_k))^{-1}; \\ &\hat{x}(t_k|t_k) = \hat{x}(t_k|t_{k-1}) + K(t_k)(y(t_k) - H(t_k)\hat{x}(t_k|t_{k-1})); \\ &P(t_k|t_k) = (I - K(t_k)H(t_k))P(t_k|t_{k-1}); \end{aligned} \tag{2.7}$$

(2) Forecasting step:

$$\begin{aligned} &\hat{x}(t_{k+1}|t_k) = M(t_{k+1}, t_k)\hat{x}(t_k|t_k); \\ &P(t_{k+1}|t_k) = M(t_{k+1}, t_k)P(t_k|t_k)M^\top(t_{k+1}, t_k) + Q(t_k). \end{aligned} \tag{2.8}$$

Defining $Y_\tau = \{y(t_o), t_o \in [t_0, \cdots, \tau]\}$ for all time τ, similar with the continuous-time case, the widely used discrete-time smoothers are generally summarized as follows:

1. Fixed-interval smoother: With a fixed discrete-time interval $[t_0, t_1, \cdots, t_N]$, the BLUE of $x(t_i)$ from the fixed-interval smoother is given by $\mathbb{E}[x(t_i)|Y_{t_N}]$, $t_i \in [t_0, \cdots, t_N]$.

2. Fixed-point smoother: The optimal estimate of the state at a specific time $\tilde{t}$ using the observations within a certain discrete-time interval $[t_0, \cdots, t_N]$ is represented by $\mathbb{E}[x(\tilde{t})|Y_{t_N}]$.

3. Fixed-lag smoother: Assuming T is a positive integral, the estimator of $x(t_i)$ of the fixed-lag smoother can be denoted by $\mathbb{E}[x(t_i)|Y_{t_{i+T}}]$.

It is clear that as the 4D-Var approach, the Kalman filter and smoother calculate the best linear unbiased estimate of the state vector by a series of observations over time. In addition, the Kalman filter and smoother update the variance of the BLUE of the state vector, which gives us an access to evaluate the error between the estimate and the true value of the state.

2.2.3 Ensemble Kalman filter and smoother

In practice, the standard Kalman filter and smoother cannot be directly applied to transport models due to their computational complexity. The ensemble Kalman filter (EnKF) and smoother (EnKS), as the Monte Carlo implementations originating from the Kalman filter and smoother, are designed for problems with a large number of control variables. EnKF and EnKS have been important tools in the field of data assimilation [37]. In this section we briefly introduce EnKF and EnKS according to [37].

For the discrete-time system (2.6) with $\varepsilon(t_i) = 0$, we denote the ensemble samples of $\hat{x}(t_i|t_{i-1})$ and $\hat{x}(t_i|t_i)$, $i = 1, \cdots, N$ respectively by

$$\begin{aligned} X(t_i|t_{i-1}) &= (\hat{x}_1(t_i|t_{i-1}), \hat{x}_2(t_i|t_{i-1}), \cdots, \hat{x}_q(t_i|t_{i-1})), \\ X(t_i|t_i) &= (\hat{x}_1(t_i|t_i), \hat{x}_2(t_i|t_i), \cdots, \hat{x}_q(t_i|t_i)), \end{aligned} \tag{2.9}$$

where q is the number of ensemble members.

Correspondingly, their ensemble means are

$$\begin{aligned} \bar{x}(t_i|t_{i-1}) &= \frac{1}{q}\sum_{k=1}^{q} \hat{x}_k(t_i|t_{i-1}) = \frac{1}{q} X(t_i|t_{i-1}) \mathbb{1}_{q\times 1}, \\ \bar{x}(t_i|t_i) &= \frac{1}{q}\sum_{k=1}^{q} \hat{x}_k(t_i|t_i) = \frac{1}{q} X(t_i|t_i) \mathbb{1}_{q\times 1}, \end{aligned}$$

where $\mathbb{1}_{i\times j}$ is a $i \times j$ matrix of which each element is equal to 1.

Note the ensemble perturbation matrix consists of the perturbation of each sampling by

$$\tilde{X}(t_i|t_{i-1}) = X(t_i|t_{i-1}) - \frac{1}{q}X(t_i|t_{i-1})\mathbb{1}_{q\times q},$$
$$\tilde{X}(t_i|t_i) = X(t_i|t_i) - \frac{1}{q}X(t_i|t_i)\mathbb{1}_{q\times q}.$$

Thus, the ensemble covariance matrices are given by

$$\begin{aligned}\bar{P}(t_i|t_{i-1}) &= \frac{1}{q-1}\tilde{X}(t_i|t_{i-1})\tilde{X}^\top(t_i|t_{i-1}),\\ \bar{P}(t_i|t_i) &= \frac{1}{q-1}\tilde{X}(t_i|t_i)\tilde{X}^\top(t_i|t_i).\end{aligned} \tag{2.10}$$

In addition, we define the ensemble observations as

$$\hat{y}_k(t_i) = y(t_i) + \nu_k(t_i), \quad k = 1, \cdots, q,\ i = 1, \cdots, N$$

where

$$\bar{\nu}(t_i) = \frac{1}{q}\sum_{k=1}^{q}\nu_k(t_i) = 0, \quad \bar{R}(t_i) = \frac{1}{q-1}\sum_{k=1}^{q}\nu_k(t_i)\nu_k^\top(t_i).$$

Further, we denote $Y_{\mathsf{en}}(t_i) = (\hat{y}_1(t_i), \cdots, \hat{y}_q(t_i))$ and

$$\bar{\mathcal{R}}^{-1} = \begin{pmatrix} \bar{R}^{-1}(t_0) & & & \\ & \bar{R}^{-1}(t_1) & & \\ & & \ddots & \\ & & & \bar{R}^{-1}(t_N) \end{pmatrix}.$$

We denoting the ensemble Kalman gain matrix by

$$\bar{K}(t_i) = \bar{P}(t_i|t_{i-1})H^\top(t_i)(H(t_i)\bar{P}(t_i|t_{i-1})H^\top(t_i) + \bar{R}(t_i))^{-1}. \tag{2.11}$$

It is worth noting that the inverse of the matrix in (2.11) is not always guaranteed. However, we can use the pseudo inverse of matrix to replace it.

Then, the analysis scheme of the ensemble Kalman filter has the consistent form with the standard Kalman filter as

$$\hat{x}_k(t_i|t_i) = \hat{x}_k(t_i|t_{i-1}) - \bar{K}(t_i)(\hat{y}_k(t_i) - H(t_i)\hat{x}_k(t_i|t_{i-1})), \quad k = 1, \cdots, q.$$

It allows that

$$\bar{x}_k(t_i|t_i) = \bar{x}_k(t_i|t_{i-1}) - \bar{K}(t_i)(\bar{y}_k(t_i) - H(t_i)\bar{x}_k(t_i|t_{i-1})), \quad k = 1, \cdots, q.$$

The ensemble analysis covariance is given by

$$\bar{P}(t_i|t_i) = (I - \bar{K}(t_i)H(t_i))\bar{P}(t_i|t_{i-1}).$$

The ensemble state evolves according to the model as

$$\hat{x}(t_{i+1}|t_i) = M(t_{i+1}, t_i)\hat{x}(t_i|t_i)$$

with the ensemble covariance matrix

$$\bar{P}(t_{i+1}|t_i) = M(t_{i+1}, t_i)\bar{P}(t_i|t_i)M^{\top}(t_{i+1}, t_i).$$

The ensemble Kalman smoother can be considered as the extension of the ensemble Kalman filter with the information propagating backward in time. We define

$$\tilde{Y}_f(t_i) = H(t_i)\tilde{X}(t_i|t_{i-1}), \quad S_y(t_i|t_{i-1}) = \tilde{Y}_f(t_i)\tilde{Y}_f^{\top}(t_i) + (q-1)\bar{R}^{-1}(t_i)$$

and

$$F_{\text{en}}(t_i) = I + \tilde{Y}_f^{\top}(t_i)S_y^{-1}(t_i|t_{i-1})(Y_{\text{en}}(t_i) - H(t_i)X(t_i|t_{i-1})),$$

According to [37], for a fixed time interval $[t_0, \cdots, t_N]$, the optimal estimate from ensemble Kalman smoother using the ensemble Kalman filter as a prior is given by

$$X(t_i|t_N) = X(t_i|t_i)\Pi_{j=i+1}^{m} F_{\text{en}}(t_i). \tag{2.12}$$

Chapter 3

Approaches to Optimizing Initial Values and Emission Rates

The evolution of chemical compounds in the troposphere is described by several physical processes and jointly effected by various chemical parameters. However, most of chemical parameters are not known precisely and hardly to be observed. In predictive geophysical model systems, uncertain initial values and emission rates jointly influence the temporal evolution of the system and play the equally important roles in improving the predictive skill.

In this chapter we firstly describe the original atmospheric transport model with emission rates and briefly review the current approach to optimize initial values and emission rates by 4D-Var [36]. Then we establish the dynamic model for emission rates in a novel way and extend the atmospheric transport with emission rates by reconstructing the state vector such that the original states and emission rates are included dynamically. Finally based on the novel atmospheric transport model extended with emission rates, we show how the initial-value-only and emission-rate-only optimization work and prove the joint optimization of initial values and emission rates can provide same or better estimates of initial values and emission rates than the initial-value-only optimization and the emission-rate-only optimization.

3.1 Current approach to optimizing initial values and emission rates by 4D-Var

We usually describe the chemical tendency in the atmosphere, propagating forward in time by the atmospheric transport model

$$\frac{dc}{dt} = \mathcal{A}(c) + e(t), \tag{3.1}$$

where $\mathcal{A}$ is a nonlinear model operator, $c(t)$ and $e(t)$ are the state vector of chemical constituents and emission rates at time t, respectively.

The prior estimate of the state vector of concentrations $c(t)$ is given and denoted by $c_b(t)$, termed as the background state. The prior estimate of emission rates, usually taken

from emission inventories, is denoted by $e_b(t)$. The incremental forms of the state vector and emission rates are given by

$$\delta c(t) = c(t) - c_b(t), \quad \delta e(t) = e(t) - e_b(t). \tag{3.2}$$

Let $\mathbf{A}$ be the tangent linear operator of $\mathcal{A}$. The evolution of the perturbations of $c(t)$ and $e(t)$ follows the tangent linear model with $\mathbf{A}$ as

$$\frac{d\delta c}{dt} = \mathbf{A}\delta c + \delta e(t). \tag{3.3}$$

With the discretization of the tangent linear model in space, we denote the finite-dimensional approximation of $\mathbf{A}$ by $\mathbf{A}_n$ and then obtain the transition matrix or resolvent generated by $\mathbf{A}_n$, denoted by $M_c(\cdot,\cdot)$, n is the dimension of the partial phase space of concentration. In order to simplify the notation and without loss of generality, we still denote the discretized state vector and emission rates as $\delta c(t)$, $\delta e(t)$ and assume they have the same dimension, namely $\delta c(t) \in \mathbb{R}^n$, $\delta e(t) \in \mathbb{R}^n$. Obviously, $M_c(\cdot,\cdot) \in \mathbb{R}^{n\times n}$. It is straightforward to obtain the linear solution of (3.3) discretized in space and continuous in time as

$$\delta c(t) = M_c(t,t_0)\delta c(t_0) + \int_{t_0}^{t} M_c(t,s)\delta e(s)ds. \tag{3.4}$$

In addition, let $y(t)$ be the observation vector of $c(t)$ and define

$$\delta y(t) = y(t) - \mathcal{H}(t)(c_b(t)), \tag{3.5}$$

where $\mathcal{H}(t)$ is a nonlinear forward observation operator mapping the model space to the observation space. We linearize and discretize the nonlinear operator $\mathcal{H}$ as H and then obtain the observation system

$$\delta y(t) = H(t)\delta c(t) + \nu(t), \tag{3.6}$$

where $\delta y(t) \in \mathbb{R}^{m(t)}$, $m(t)$ the dimension of the phase space of observation configurations at time t. $\nu(t)$ is the observation error at time t following the Gaussian distribution which has zero mean and covariance matrix $R(t) \in \mathbb{R}^{m(t)\times m(t)}$.

In this chapter we work on the linear model (3.4) with the observation system (3.6).

As the initial value of the state vector and emission rates play the equally important roles in improving the accuracy of estimations [36], the 4D-Var approach introduced in Section 2.1 is only feasible to the initial-value-only optimization rather than the optimization of both the initial state and emission rates. Elbern et al. [36] regulated the emission rate by preserving their diurnal profiles such that the total amounts of emitted species can be controlled by a emission factor f such that $e(t) = fe_b(t)$. Here f is time invariant and location dependent. Then the joint optimization is presented as follows. Firstly, the constant emission factor is transformed by the logarithm as

$$\delta u = \ln e(t) - \ln e_b(t) = \ln f.$$

We denote the variance of δu by P^u and combine the initial value and the emission rate into one vector, denoted by

$$\delta z = (\delta c^\top(t_0), \delta u^\top)^\top.$$

Denoting the covariance of $\delta c(t_0)$ by $P^c(t_0)$, we assume that $\delta c(t_0)$ and δu are uncorrelated since, on one hand, it is already rather numerically costly to deal with $P^c(t_0)$ itself. On the other hand, the correlation between $\delta c(t_0)$ and δu varies in different scenarios and is hard to be formulated uniformly. Then the objective function of 4D-Var is defined as

$$\begin{aligned} J(\delta z) =& \frac{1}{2}(\delta c^\top(t_0), \delta u^\top) \begin{pmatrix} P^c(t_0) & 0 \\ 0 & P^u \end{pmatrix}^{-1} \begin{pmatrix} \delta c(t_0) \\ \delta u \end{pmatrix} \\ &+ \frac{1}{2}\sum_{i=0}^{N}(y(t_i) - H(t_i)x(t_i))^\top R^{-1}(t_i)(y(t_i) - H(t_i)x(t_i)). \end{aligned} \tag{3.7}$$

In order to search for the minimal solution of (3.7), we calculate its equilibrium by its gradient with respect to δz

$$\begin{aligned} \frac{\partial J(\delta z)}{\partial \delta z} =& (P^c(t_0))^{-1}\delta c(t_0) + (P^u)^{-1}\delta u \\ &+ \sum_{i=0}^{N} H^\top(t_i) M_c^\top(t_i, t_0) R^{-1}(t_i)(H(t_i)\delta c(t_i) - y(t_i)). \end{aligned}$$

Due to the uncorrelated assumption of $\delta c(t_0)$ and δu, if we only consider optimizing the emission rates, by assuming $\delta c(t_0) = 0$, the objective function (3.7) can be simply rewritten as

$$\begin{aligned} J(\delta u) =& \frac{1}{2}\delta u^\top (P^u)^{-1}\delta u \\ &+ \frac{1}{2}\sum_{i=0}^{N}(y(t_i) - H(t_i)x(t_i))^\top R^{-1}(t_i)(y(t_i) - H(t_i)x(t_i)). \end{aligned}$$

3.2 Novel approach to optimizing initial values and emission rates by KS

The 4D-Var approach summarized in the last section gives us an access to improve the estimations of the initial value and emission factors. However, we hardly gain the statistic information of the accuracy of the estimates by 4D-Var. Especially for emission rates, in the previous method, we assume δu follows the Gaussian distribution or to say f is log-normal distributed, rather than giving the Gaussian assumption directly on the perturbed emission rates themselves.

Reviewing the main formulas of the discrete-time Kalman filter and smoother in Section 2.2.2, we can easily find that it is only feasible to apply the Kalman filter and smoother into the linear model (3.4) with observations (3.6) within a given time interval to optimize the initial value of the concentration. As to the emission-rate optimization or

the joint optimization of initial value and emission rates, the Kalman filter and smoother is not feasible to the current model.

Hence, in this section we extend the atmospheric transport model with emission rates in a novel way such that the Kalman filter and smoother can be directly applied into the new model to optimize both the initial value and emission rates. At the same time we are able to obtain the analysis covariance to evaluate the improvements of estimations, which is a foundation of the work in the next chapters.

3.2.1 Atmospheric transport model extended by emission rates

Generally we still use the model (3.4) to formulate the evolution of chemical compounds in atmosphere. As mentioned before, it has been shown in [36] that the diurnal profiles of emission rates are better known than the amplitude of emission rates. Thus, the diurnal profiles of emission rates can be taken as constraints such that the amplitudes of emission rates become the optimized parameters. Thus we firstly formulate the background evolution of emission rates from time s to t into the dynamic form within a given data assimilation window $[t_0, t_N]$

$$e_b(t) = M_e(t,s)e_b(s), \quad t_0 \leqslant s \leqslant t \leqslant t_N, \tag{3.8}$$

where $e_b(\cdot)$ is a n-dimensional vector of which the i^{th} element is denoted by $e_b^i(\cdot)$ and $M_e(t,s)$ is the diagonal matrix defined as

$$M_e(t,s) = \begin{pmatrix} \frac{e_b^1(t)}{e_b^1(s)} & & & \\ & \frac{e_b^2(t)}{e_b^2(s)} & & \\ & & \ddots & \\ & & & \frac{e_b^n(t)}{e_b^n(s)} \end{pmatrix},$$

where

$$\frac{e_b^i(t)}{e_b^i(s)} \in \mathbb{R}, \quad i = 1, \cdots, n.$$

We establish the dynamic model of emission rates by forcing emission rates to follow the background evolution of emission rates as

$$\delta e(t) = M_e(t,s)\delta e(s), \quad t_0 \leqslant s \leqslant t \leqslant t_N. \tag{3.9}$$

We have stated in Section 2.2 that the estimate of the variable x via the fix-interval Kalman smoother can be generally expressed as the conditional expectation based on the observations in the whole time interval, denoted by $\mathbb{E}[x|\{y(t_o), t_o \in [t_0, t_N]\}]$. With the dynamic model (3.9), the estimate of $e(t)$ by Kalman smoother on $[t_0, t_N]$ follows the linear property of the conditional expectation,

$$\begin{aligned} &\mathbb{E}[e(t)|\{y(t_o), t_o \in [t_0, t_N]\}] = \mathbb{E}[M_e(t,s)e(s)|\{y(t_o), t_o \in [t_0, t_N]\}] \\ =&M_e(t,s)\mathbb{E}[e(s)|\{y(t_o), t_o \in [t_0, t_N]\}]. \end{aligned} \tag{3.10}$$

It implies that the BLUEs of emission rates with the dynamic model (3.9) by Kalman smoother preserve the proportioned diurnal profiles of their backgrounds.

We rewrite (3.4) as

$$\delta c(t) = M_c(t, t_0)\delta c(t_0) + \int_{t_0}^{t} M_c(t, s)M_e(s, t_0)\delta e(t_0)ds. \tag{3.11}$$

Combining (3.4) with (3.9), we obtain the transport model extended with emission rates

$$\begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} = \begin{pmatrix} M_c(t, t_0) & \int_{t_0}^{t} M_c(t, s)M_e(s, t_0)ds \\ 0 & M_e(t, t_0) \end{pmatrix} \begin{pmatrix} \delta c(t_0) \\ \delta e(t_0) \end{pmatrix}. \tag{3.12}$$

Typically, there is no direct observation for emissions, apart from the flux tower observations used for carbon dioxide, which are not considered here. Therefore, we reformulate the observation mapping as

$$\delta y(t) = (H(t), 0_{n\times n}) \begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} + \nu(t), \tag{3.13}$$

where $0_{n\times n}$ is a $n \times n$ matrix with zero elements.

Now we see that both concentrations and emission rates are included into the state vector of the homogeneous dynamic model (3.12). It allows us to apply the Kalman smoother within a fixed time interval $[t_0, t_N]$ to optimize both parameters.

3.2.2 Joint optimization of initial values and emission rates

In order to study how the Kalman filter and smoother work on the atmospheric transport model extended by emission rates in details and practice, in this section we consider the discrete-time model extended by emission rate. Firstly, we rewrite the model (3.1) discretized in time as

$$\delta c(t_{k+1}) = M_c(t_{k+1}, c_k)\delta c(t_k) + B(t_k)\delta e(t_k),$$

then we can formulate the extended model (3.12) discretized in time as

$$\begin{pmatrix} \delta c(t_{k+1}) \\ \delta e(t_{k+1}) \end{pmatrix} = \begin{pmatrix} M_c(t_{k+1}, t_k) & B(t_k) \\ 0 & M_e(t_{k+1}, t_k) \end{pmatrix} \begin{pmatrix} \delta c(t_k) \\ \delta e(t_k) \end{pmatrix}. \tag{3.14}$$

Correspondingly, we discretize the observations system (3.13) as

$$\delta y(t_k) = (H(t_k), 0) \begin{pmatrix} \delta c(t_k) \\ \delta e(t_k) \end{pmatrix} + \nu(t_k), \quad \nu_k \sim \mathcal{N}(0, R_k).$$

In the following, we formulate the transition matrix of (3.14) as

$$M(t_{k+1}, t_k) = \begin{pmatrix} M_c(t_{k+1}, t_k) & B(t_k) \\ 0 & M_e(t_{k+1}, t_k) \end{pmatrix}.$$

and the initial state of the model (3.14) by

$$\begin{pmatrix} \delta\hat{c}(t_0|t_{-1}) \\ \delta\hat{e}(t_0|t_{-1}) \end{pmatrix}.$$

The initial covariance is given by

$$P(t_0|t_{-1}) = \begin{pmatrix} P^c(t_0|t_{-1}) & 0 \\ 0 & P^e(t_0|t_{-1}) \end{pmatrix}.$$

For the extended model (3.14), the analysis and forecasting covariance matrices have the following block forms

$$P(t_k|t_k) = \begin{pmatrix} P^c(t_k|t_k) & P^{ce}(t_k|t_k) \\ P^{ec}(t_k|t_k) & P^e(t_k|t_k) \end{pmatrix},$$
$$P(t_k|t_{k-1}) = \begin{pmatrix} P^c(t_k|t_{k-1}) & P^{ce}(t_k|t_{k-1}) \\ P^{ec}(t_k|t_{k-1}) & P^e(t_k|t_{k-1}) \end{pmatrix},$$

where $P^c(t_k|\cdot)$ and $P^e(t_k|\cdot)$ are the covariance matrices of $\delta\hat{c}(t_k|\cdot)$ and $\delta\hat{e}(t_k|\cdot)$. $P^{ce}(t_k|\cdot)$ is the covariance matrix between $\delta\hat{c}(t_k|\cdot)$ and $\delta\hat{e}(t_k|\cdot)$. Besides, $P^{ce}(t_k|\cdot) = (P^{ec}(t_k|\cdot))^\top$. According to Section 2.2, we summarize the main steps of Kalman filter and smoother as follows.

Analysis step of KF for the joint optimization:

1. Gain matrix:

$$\begin{aligned} K(t_k) &= \begin{pmatrix} K^c(t_k) \\ K^e(t_k) \end{pmatrix} \\ &= P(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix} [(H(t_k), 0) P(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix} + R(t_k)]^{-1} \\ &= \begin{pmatrix} P^c(t_k|t_{k-1}) H^\top(t_k) [H(t_k) P^c(t_k|t_{k-1}) H^\top(t_k) + R(t_k)]^{-1} \\ P^{ec}(t_k|t_{k-1}) H^\top(t_k) [H(t_k) P^c(t_k|t_{k-1}) H^\top(t_k) + R(t_k)]^{-1} \end{pmatrix}. \end{aligned}$$

2. Estimate of the extended state vector:

$$\begin{pmatrix} \delta\hat{c}(t_k|t_k) \\ \delta\hat{e}(t_k|t_k) \end{pmatrix} = \begin{pmatrix} \delta\hat{c}(t_k|t_{k-1}) + K^c(t_k)(y(t_k) - H(t_k)\delta\hat{c}(t_k|t_{k-1})) \\ \delta\hat{e}(t_k|t_{k-1}) + K^e(t_k)(y(t_k) - H(t_k)\delta\hat{c}(t_k|t_{k-1})) \end{pmatrix}.$$

3. Analysis covariance matrix:

$$P(t_k|t_k) = \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} K^c(t_k) \\ K^e(t_k) \end{pmatrix} (H(t_k), 0) \right] P(t_k|t_{k-1}). \tag{3.15}$$

Forecasting step of KF for the joint optimization:

1. Estimate of the extended state vector:

$$\begin{aligned} \begin{pmatrix} \delta\hat{c}(t_{k+1}|t_k) \\ \delta\hat{e}(t_{k+1}|t_k) \end{pmatrix} &= M(t_{k+1}, t_k) \begin{pmatrix} \delta\hat{c}(t_k|t_k) \\ \delta\hat{e}(t_k|t_k) \end{pmatrix} \\ &= \begin{pmatrix} M_c(t_{k+1}, t_k)\delta\hat{c}(t_k|t_k) + B(t_k)\delta\hat{e}(t_k|t_k) \\ M_e(t_{k+1}, t_k)\delta\hat{e}(t_k|t_k) \end{pmatrix}. \end{aligned}$$

2. Forecasting covariance matrix:

$$P(t_{k+1}|t_k) = M(t_{k+1}, t_k)P(t_k|t_k)M^\top(t_{k+1}, t_k).$$

Estimate results by KS:
1. Estimate of the extended state vector at the initial time:

$$\begin{pmatrix} \delta\hat{c}(t_0|t_N) \\ \delta\hat{e}(t_0|t_N) \end{pmatrix} = M^{-1}(t_N, t_0) \begin{pmatrix} \delta\hat{c}(t_N|t_N) \\ \delta\hat{e}(t_N|t_N) \end{pmatrix}.$$

2. Covariance matrix of the initial extended state:

$$P(t_0|t_N) = M^{-1}(t_N, t_0)P(t_N|t_N)M^{-\top}(t_N, t_0),$$

where $M(t_N, t_0) = \Pi_{i=0}^{N-1} M(t_{i+1}, t_i)$.

3.2.3 Initial-value-only optimization

In this section we consider the initial value as the only parameter to be optimized. In another word, emission rates are viewed as the input of the model. This indicates we only need to apply Kalman filter and smoother into the original transport model (3.14) with $\delta e(t_k) = 0$. In order to distinguish the notations with the joint optimization, we denote the state of (3.14) by $c_I(t_k)$ in the case of initial-value-only optimization. For the given initial priori estimate $\hat{c}_I(t_0|t_{-1})$, we have

$$\delta c_I(t_{k+1}) = M_c(t_{k+1}, t_k)\delta c_I(t_k), \tag{3.16}$$

where $\delta c_I(t_k) = c_I(t_k) - c_b(t_k)$. Correspondingly, the observation system can be rewritten as

$$\delta y(t_k) = H(t_k)\delta c_I(t_k) + \nu(t_k). \tag{3.17}$$

The analysis and forecasting covariance matrices of concentrations of the initial-value-only optimization are denoted by

$$\begin{aligned} P_I^c(t_k|t_k) &= \mathbb{E}[(\delta\hat{c}_I(t_k|t_k) - \delta c_I(t_k))(\delta\hat{c}_I(t_k|t_{k-1}) - \delta c_I(t_k))^\top], \\ P_I^c(t_k|t_{k-1}) &= \mathbb{E}[(\delta\hat{c}_I(t_k|t_{k-1}) - \delta c_I(t_k))(\delta\hat{c}_I(t_k|t_{k-1}) - \delta c_I(t_k))^\top], \end{aligned} \tag{3.18}$$

where $\delta\hat{c}_I(t_k|t_k) = c_I(t_k) - \hat{c}_I(t_k|t_k)$ and $\delta\hat{c}_I(t_k|t_{k-1}) = c_I(t_k) - \hat{c}_I(t_k|t_{k-1})$.

We summarize the main steps of the Kalman filter and smoother for initial-value-only optimization as

Analysis step of KF of only initial value optimization:
1. Gain matrix:

$$K_I^c(t_k) = P_I^c(t_k|t_{k-1})H^\top(t_k)[H(t_k)P_I^c(t_k|t_{k-1})H^\top(t_k) + R(t_k)]^{-1}.$$

2. Estimate of the initial value:

$$\delta\hat{c}_I(t_k|t_k) = \delta\hat{c}_I(t_k|t_{k-1}) + K_I^c(t_k)(\delta y(t_k) - H(t_k)\delta\hat{c}_I(t_k|t_{k-1})).$$

3. Analysis covariance of the concentrations:

$$P_I^c(t_k|t_k) = (I - K_I^c(t_k)H(t_k))P_I^c(t_k|t_{k-1}).$$

Forecasting step of KF of only initial value optimization:
1. Forecasting estimate of the concentration

$$\delta\hat{c}_I(t_{k+1}|t_k) = M_c(t_{k+1}, t_k)\delta\hat{c}_I(t_k|t_k).$$

2. Forecasting covariance of the concentrations:

$$P_I^c(t_{k+1}|t_k) = M_c(t_{k+1}, t_k)P_I^c(t_k|t_k)M_c^\top(t_{k+1}, t_k).$$

Estimation results of KS for initial-value-only optimization:
1. Estimate of the initial state:

$$\delta\hat{c}_I(t_0|t_{-1}) = M_c^{-1}(t_N, t_0)\delta\hat{c}_I(t_N|t_N).$$

2. Covariance matrix of the initial estimate:

$$P_I^c(t_0|t_N) = M_c^{-1}(t_N, t_0)P_I(t_N|t_N)M_c^{-\top}(t_N, t_0).$$

Preserving the diurnal profiles of emission rates, one of our objectives in this chapter is to study how the extended model influences on the joint optimization and compare it with the optimization of the initial state or emission rates. Hence, in order to formulate the block form of the initial-value-only optimization extended by emission rates, we notate $e_I(t_k)$ as the state vector of emission rates at t_k and $\delta e_I(t_{k+1}) = e_I(t_k) - e_b(t_k)$. We rewrite (3.16), (3.17) and the corresponding results from Kalman filter and smoother into the following block form:

$$\begin{pmatrix} \delta c_I(t_{k+1}) \\ \delta e_I(t_{k+1}) \end{pmatrix} = \begin{pmatrix} M_c(t_{k+1}, t_k) & B(t_k) \\ 0 & M_e(t_{k+1}, t_k) \end{pmatrix} \begin{pmatrix} \delta c_I(t_k) \\ \delta e_I(t_k) \end{pmatrix}, \tag{3.19}$$

$$\delta y(t_k) = (H(t_k), 0) \begin{pmatrix} \delta c_I(t_k) \\ \delta e_I(t_k) \end{pmatrix} + \nu(t_k), \quad \nu(t_k) \sim \mathcal{N}(0, R(t_k)). \tag{3.20}$$

The initial priori estimate of the extended state and its covariance are respectively given by

$$\begin{pmatrix} \delta\hat{c}_I(t_0|t_{-1}) \\ 0 \end{pmatrix} \quad \text{and} \quad P_I(t_0|t_{-1}) = \begin{pmatrix} P_I^c(t_0|t_{-1}) & 0 \\ 0 & 0 \end{pmatrix}. \tag{3.21}$$

With the initial condition (3.21), we denote the analysis and forecasting estimates of the emission rate based on the model (3.19) with the observation system (3.20) by $\delta\hat{e}_I(t_k|t_k)$ and $\delta\hat{e}_I(t_k|t_{k-1})$. Correspondingly, the analysis and forecasting covariance matrices of $(\delta c_I^\top(t_k), \delta e_I^\top(t_k))^\top$ are denoted by

$$P_I(t_k|t_k) = \begin{pmatrix} P_I^c(t_k|t_k) & P_I^{ce}(t_k|t_k) \\ P_I^{ec}(t_k|t_k) & P_I^e(t_k|t_k) \end{pmatrix},$$

$$P_I(t_k|t_{k-1}) = \begin{pmatrix} P_I^c(t_k|t_{k-1}) & P_I^{ce}(t_k|t_{k-1}) \\ P_I^{ec}(t_k|t_{k-1}) & P_I^e(t_k|t_{k-1}) \end{pmatrix}.$$

We rewrite the main results of Kalman filter and smoother for the initial-value-only optimization in the following block form.

Equivalent analysis step of KF:

1. Gain Matrix:

$$\begin{aligned} K_{I}(t_k) &= \begin{pmatrix} K_{I}^{c}(t_k) \\ K_{I}^{e}(t_k) \end{pmatrix} \\ &= \begin{pmatrix} P_{I}^{c}(t_k|t_{k-1})H^{\top}(t_k)[H(t_k)P_{I}^{c}(t_k|t_{k-1})H^{\top}(t_k)+R(t_k)]^{-1} \\ 0 \end{pmatrix}. \end{aligned}$$

2. Estimate of the state:

$$\begin{pmatrix} \delta\hat{c}_{I}(t_k|t_k) \\ \delta\hat{e}_{I}(t_k|t_k) \end{pmatrix} = \begin{pmatrix} \delta\hat{c}_{I}(t_k|t_{k-1}) + K_{I}^{c}(t_k)(\delta y(t_k) - H(t_k)\delta\hat{c}_{I}(t_k|t_{k-1})) \\ 0 \end{pmatrix}.$$

3. Analysis covariance matrix:

$$\begin{aligned} P_{I}(t_k|t_k) &= \begin{pmatrix} P_{I}^{c}(t_k|t_k) & 0 \\ 0 & 0 \end{pmatrix} \\ &= \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} K_{I}^{c}(t_k) \\ 0 \end{pmatrix}(H(t_k),\ 0)\right]\begin{pmatrix} P_{I}^{c}(t_k|t_{k-1}) & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} (I - K_{I}^{c}(t_k)H(t_k))P_{I}^{c}(t_k|t_{k-1}) & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \tag{3.22}$$

Equivalent forecasting step of KF:

1. Forecasting estimate of the state:

$$\begin{pmatrix} \delta\hat{c}_{I}(t_{k+1}|t_k) \\ \delta\hat{e}_{I}(t_{k+1}|t_k) \end{pmatrix} = \begin{pmatrix} M_c(t_{k+1},t_k)\delta\hat{c}_{I}(t_k|t_k) \\ 0 \end{pmatrix}.$$

2. Forecasting covariance matrix:

$$P_{I}(t_{k+1}|t_k) = \begin{pmatrix} M_c(t_{k+1},t_k)P_{I}^{c}(t_k|t_k)M_{c}^{\top}(t_{k+1},t_k) & 0 \\ 0 & 0 \end{pmatrix}.$$

Equivalent estimation results of KS:

1. Estimate of the initial state:

$$\begin{pmatrix} \delta\hat{c}_{I}(t_0|t_N) \\ \delta\hat{e}_{I}(t_0|t_N) \end{pmatrix} = \begin{pmatrix} M_c(t_N,t_0)^{-1}\delta\hat{c}_{I}(t_N|t_N) \\ 0 \end{pmatrix}.$$

2. Covariance matrix of the initial estimate:

$$P_{I}(t_0|t_N) = \begin{pmatrix} M_{c}^{-1}(t_N,t_0)P_{I}^{c}(t_N|t_N)M_{c}^{-\top}(t_N,t_0) & 0 \\ 0 & 0 \end{pmatrix}.$$

3.2.4 Emission-rate-only optimization

In this section we consider the emission rates as the only parameters to be optimized and state the main results of Kalman filter and smoother for this emission-rate-only optimization. Similar to the initial-value-only optimization case, we denote the concentration and emission rate at t_k of the emission-rate-only optimization case by $c_E(t_k)$, $e_E(t_k)$ and define $\delta c_E(t_k) = c_E(t_k) - c_b(t_k)$, $\delta e_E(t_k) = e_E(t_k) - e_b(t_k)$. Then we consider the following model

$$\begin{pmatrix} \delta c_E(t_{k+1}) \\ \delta e_E(t_{k+1}) \end{pmatrix} = \begin{pmatrix} M_c(t_{k+1}, t_k) & B(t_k) \\ 0 & M_e(t_{k+1}, t_k) \end{pmatrix} \begin{pmatrix} \delta c_E(t_k) \\ \delta e_E(t_k) \end{pmatrix}, \tag{3.23}$$

$$\delta y(t_k) = (H(t_k), 0) \begin{pmatrix} \delta c_E(t_k) \\ \delta e_E(t_k) \end{pmatrix} + \nu(t_k), \quad \nu(t_k) \sim \mathcal{N}(0, R(t_k)) \tag{3.24}$$

with the initial prior estimate and covariance matrix

$$\begin{pmatrix} 0 \\ \delta \hat{e}_E(t_0|t_{-1}) \end{pmatrix} \quad \text{and} \quad P_E(t_0|t_{-1}) = \begin{pmatrix} 0 & 0 \\ 0 & P_E^e(t_0|t_{-1}) \end{pmatrix}.$$

Further, we denote the analysis and forecasting estimates of concentrations and emission rates by $\delta \hat{c}_E(t_k|t_k)$, $\delta \hat{c}_E(t_k|t_{k-1})$ and $\delta \hat{e}_E(t_k|t_k)$, $\delta \hat{e}_E(t_k|t_{k-1})$. Similarly we assume the analysis and forecasting covariance matrices of $(\delta c_E^\top(t_k), \delta e_E^\top(t_k))^\top$ to be

$$P_E(t_k|t_k) = \begin{pmatrix} P_E^c(t_k|t_k) & P_E^{ce}(t_k|t_k) \\ P_E^{ec}(t_k|t_k) & P_E^e(t_k|t_k) \end{pmatrix},$$

$$P_E(t_k|t_{k-1}) = \begin{pmatrix} P_E^c(t_k|t_{k-1}) & P_E^{ce}(t_k|t_{k-1}) \\ P_E^{ec}(t_k|t_{k-1}) & P_E^e(t_k|t_{k-1}) \end{pmatrix}.$$

Now we summarize the main steps of Kalman filter and smoother for the emission-rate-only optimization as follows.

Analysis step of KF:

1. Gain matrix:

$$\begin{aligned} K_E(t_k) &= \begin{pmatrix} K_E^c(t_k) \\ K_E^e(t_k) \end{pmatrix} \\ &= \begin{pmatrix} P_E^c(t_k|t_{k-1}) H^\top(t_k) [H(t_k) P_E^c(t_k|t_{k-1}) H^\top(t_k) + R(t_k)]^{-1} \\ P_E^{ec}(t_k|t_{k-1}) H^\top(t_k) [H(t_k) P_E^c(t_k|t_{k-1}) H^\top(t_k) + R(t_k)]^{-1} \end{pmatrix}. \end{aligned}$$

2. Estimate of the state:

$$\begin{pmatrix} \delta \hat{c}_E(t_k|t_k) \\ \delta \hat{e}_E(t_k|t_k) \end{pmatrix} = \begin{pmatrix} \delta \hat{c}_E(t_k|t_{k-1}) + K_E^c(t_k)(y(t_k) - H(t_k)\delta \hat{c}(t_k|t_{k-1})) \\ \delta \hat{e}_E(t_k|t_{k-1}) + K_E^e(t_k)(y(t_k) - H(t_k)\delta \hat{c}(t_k|t_{k-1})) \end{pmatrix}.$$

3. Analysis covariance matrix:

$$\begin{aligned} &P_E(t_k|t_k) \\ &= \left[\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} K_E^c(t_k) \\ K_E^e(t_k) \end{pmatrix} (H(t_k),\ 0) \right] \begin{pmatrix} P_E^c(t_k|t_{k-1}) & P_E^{ce}(t_k|t_{k-1}) \\ P_E^{ec}(t_k|t_{k-1}) & P_E^e(t_k|t_{k-1}) \end{pmatrix}. \end{aligned}$$

Forecasting step by KF:
1. Forecasting estimates of states:

$$\begin{aligned}\begin{pmatrix} \delta\hat{c}_{E}(t_{k+1}|t_k) \\ \delta\hat{e}_{E}(t_{k+1}|t_k) \end{pmatrix} &= M(t_N,t_0)\begin{pmatrix} \delta\hat{c}_{E}(t_k|t_k) \\ \delta\hat{e}_{E}(t_k|t_k) \end{pmatrix} \\ &= \begin{pmatrix} M_c(t_{k+1},t_k)\delta\hat{c}_{E}(t_k|t_k) + B(t_k)\delta\hat{e}_{E}(t_k|t_k) \\ M_e(t_{k+1},t_k)\delta\hat{e}_{E}(t_k|t_k) \end{pmatrix}.\end{aligned}$$

2. Forecasting covariance matrix:

$$P_{E}(t_{k+1}|t_k) = M(t_N,t_0)P_{E}(t_k|t_k)M^{\top}(t_N,t_0).$$

General solution of KS at initial time:
1. Estimate of the initial state:

$$\begin{pmatrix} \delta\hat{c}_{E}(t_0|t_N) \\ \delta\hat{e}_{E}(t_0|t_N) \end{pmatrix} = M^{-1}(t_N,t_0)\begin{pmatrix} \delta\hat{c}_{E}(t_N|t_N) \\ \delta\hat{e}_{E}(t_N|t_N) \end{pmatrix}.$$

2. Covariance matrix of the initial estimate:

$$P_{E}(t_0|t_N) = M^{-1}(t_N,t_0)P_{E}(t_N|t_N)M^{-\top}(t_N,t_0).$$

3.2.5 Comparison

In this section we compare the estimation results of the joint optimization with the estimation results of the initial-value-only optimization and the emission-rate-only optimization respectively. Taking the diurnal profiles of emission rates as constraints, it will be shown that the estimates of the joint optimization based on our atmospheric transport model extended with emissions are better or at least as good as the estimates of the initial-value-only optimization and the emission-rate-only optimization in various situations.

Joint optimization versus only initial value optimization

In Section 3.2.3, we assume $\delta e_{I}(t) = 0$ by ignoring the error of the background of emission rates. However, it is impossible that emission rates are perfectly known in practice. There are more or less some inaccuracies of the background knowledge of emission rates. This implies $\delta e(t) \neq 0$. Taking the diurnal profiles of emissions as constraints and the error of emission rates into account, there is a difference between $\delta e_{I}(t_k)$ and $\delta e(t_k)$ given by $-\delta e(t_k)$. Thus, the true analysis and forecasting covariances for the initial-value-only optimization based on the model (3.19) can be uniformly represented as:

$$\begin{aligned}P_{TI}(t_k|\cdot) =& \begin{pmatrix} P^{c}_{TI}(t_k|\cdot) & P^{ce}_{TI}(t_k|\cdot) \\ P^{ec}_{TI}(t_k|\cdot) & P^{e}_{TI}(t_k|\cdot) \end{pmatrix} \\ =& \mathbb{E}\left[\begin{pmatrix} \delta\hat{c}_{I}(t_k|\cdot) - \delta c(t_k) \\ -\delta e(t_k) \end{pmatrix}\left((\delta\hat{c}_{I}(t_k|\cdot) - \delta c(t_k))^{\top}, -\delta e^{\top}(t_k)\right)\right],\end{aligned}$$

where the dot "$\cdot$" could be any time step t_i, $i = 0, \cdots N$.

In particular, we denote the true initial prior covariance as

$$P_{TI}(t_0|t_{-1}) = \begin{pmatrix} P^c_{TI}(t_0|t_{-1}) & 0 \\ 0 & P^e_{TI}(t_0|t_{-1}) \end{pmatrix}.$$

Since $P_{TI}(t_0|t_{-1})$ is also the first guess from the previous knowledge, we can always assume that

$$P_{TI}(t_0|t_{-1}) = P(t_0|t_{-1}). \tag{3.25}$$

Defining the covariance of the *difference* of the model (3.14) and (3.19) by

$$\begin{aligned} D_I(t_k) :=& \begin{pmatrix} D^c_I(t_k) & D^{ce}_I(t_k) \\ D^{ec}_I(t_k) & D^e_I(t_k) \end{pmatrix} \\ =&\mathbb{E}\left[\begin{pmatrix} \delta c_I(t_k) - \delta c(t_k) \\ -\delta e(t_k) \end{pmatrix} (\delta c_I(t_k) - \delta c(t_k))^\top, -\delta e^\top(t_k))\right], \end{aligned}$$

we have

$$\begin{aligned} & P_{TI}(t_k|\cdot) \\ =&\mathbb{E}\left[\begin{pmatrix} (\delta\hat{c}_I(t_k|\cdot) - \delta c_I(t_k)) + (\delta c_I(t_k) - \delta c(t_k)) \\ -\delta e(t_k) \end{pmatrix}\right. \\ &\left.\cdot\begin{pmatrix} (\delta\hat{c}_I(t_k|\cdot) - \delta c_I(t_k)) + (\delta c_I(t_k) - \delta c(t_k)) \\ -\delta e(t_k) \end{pmatrix}^\top\right] \\ =&\mathbb{E}\left[\begin{pmatrix} (\delta\hat{c}_I(t_k|\cdot) - \delta c_I(t_k)) \\ 0 \end{pmatrix} ((\delta\hat{c}_I(t_k|\cdot) - \delta c_I(t_k))^\top, 0^\top)\right] \\ &+\mathbb{E}\left[\begin{pmatrix} \delta c_I(t_k) - \delta c(t_k) \\ -\delta e(t_k) \end{pmatrix} (\delta c_I(t_k) - \delta c(t_k))^\top, -\delta e^\top(t_k))\right] \\ =&P_I(t_k|\cdot) + D_I(t_k). \end{aligned}$$

Before we compare the estimations of the joint optimization with the estimations of the initial-value-only optimization, we introduce a notation and a lemma.

We assume that $P_1, P_2 \in \mathbb{R}^{n\times n}$ are two (semi-) positive definite matrices. If for any $x \neq 0_{n\times 1} \in \mathbb{R}^n$, $x^\top(P_1 - P_2)x < (\leqslant)0$ holds, then we denote $P_1 \prec (\preccurlyeq)P_2$. Now we have the following lemma.

Lemma 3.2.1. *Let $R \in \mathbb{R}^{m\times m}$ be positive-definite matrix. P_1 and $P_2 \in \mathbb{R}^{n\times n}$ be semi-positive definite matrices. $H \in \mathbb{R}^{m\times n}$ is any matrix. If $P_1 \succcurlyeq P_2$, then*

$$P_1H^\top(HP_1H^\top + R)^{-1}HP_1 \succcurlyeq P_2H^\top(HP_2H^\top + R)^{-1}HP_2. \tag{3.26}$$

Proof. (1) We firstly consider the case that $P_1 \succ P_2 \succ 0$. According to the matrix inversion lemma, also known as Woodbury formula [99], in order to prove

$$P_1H^\top(HP_1H^\top + R)^{-1}HP_1 \succcurlyeq P_2H^\top(HP_2H^\top + R)^{-1}HP_2,$$

it is equivalent to prove

$$P_1 - (P_1^{-1} + H^\top R^{-1} H)^{-1} \succcurlyeq P_2 - (P_2^{-1} + H^\top R^{-1} H)^{-1}.$$

Thus, for any semi-positive definite matrix A, if we define

$$f_A(P) = P - (P^{-1} + A)^{-1}, \tag{3.27}$$

we need to prove that $f_A(P)$ is monotone increasing function with P in the sense of the quadratic form. Defining

$$g_A(P) := f_A(P^{-1}) = P^{-1} - (P + A)^{-1},$$

then we will equivalently prove that $g_A(P)$ is monotone decreasing function with P in the sense of quadratic form.

Since $P_1 \succ P_2 \succ 0$, we define

$$P_1 = P_2 + \delta P, \quad \delta P \succ 0.$$

If we assume that $g_{\delta P}(P)$ is not monotone decreasing with P in the sense of quadratic form, then for any $n \in \mathbb{N}$, there exists at least one vector $x_n \neq 0_{n\times 1} \in \mathbb{R}^n$ such that

$$\frac{x_n^\top}{\|x_n\|}((nP_1)^{-1} - ((nP_1) + \delta P)^{-1})\frac{x_n}{\|x_n\|} > \frac{x_n^\top}{\|x_n\|}(P_2^{-1} - (P_2 + \delta P)^{-1})\frac{x_n}{\|x_n\|}, \tag{3.28}$$

where $\|\cdot\|$ is the Euclidean norm of vectors.

On one hand, for any $n \in \mathbb{N}$, according to (3.28), we have

$$\begin{aligned}
\frac{x_n^\top}{\|x_n\|}(P_2^{-1} - P_1^{-1})\frac{x_n}{\|x_n\|} =& \frac{x_n^\top}{\|x_n\|}(P_2^{-1} - (P_2 + \delta P)^{-1})\frac{x_n}{\|x_n\|} \\
<& \frac{x_n^\top}{\|x_n\|}((nP_1)^{-1} - (nP_1 + \delta P)^{-1})\frac{x_n}{\|x_n\|} \\
<& \frac{x_n^\top}{\|x_n\|}((nP_1)^{-1} - ((n+1)P_1)^{-1})\frac{x_n}{\|x_n\|} \\
=& \frac{x_n^\top}{\|x_n\|}\frac{1}{n(n+1)}P_1^{-1}\frac{x_n}{\|x_n\|}.
\end{aligned}$$

Since $\{\frac{x_n}{\|x_n\|}_{n\in\mathbb{N}}\}$ is bounded, then there is subsequence $\{\frac{x_{n_k}}{\|x_{n_k}\|}\}_{k\in\mathbb{N}}$ converging to x ($\|x\| \neq 0$) when $k \to \infty$. Therefore, we have

$$x^\top(P_2^{-1} - P_1^{-1})x \leqslant 0, \quad n \to \infty. \tag{3.29}$$

On the other hand,

$$\begin{aligned}
P_2^{-1} - P_1^{-1} &= P_2^{-1} - (P_2 + \delta P)^{-1} \\
&= P_2^{-1}\delta P^{\frac{1}{2}}(I + \delta P^{\frac{1}{2}} P_2^{-1} \delta P^{\frac{1}{2}})^{-1}\delta P^{\frac{1}{2}} P_2^{-1} \succ 0,
\end{aligned} \tag{3.30}$$

which is conflict with (3.29). Therefore, we can conclude that (3.26) holds for $P_1 \succ P_2 \succ 0$.

(2) Now we consider the case $P_1 \succcurlyeq P_2 \succcurlyeq 0$, which implies P_1 and P_2 are probably not invertible. Thus, from (1) above, for any $\varepsilon > 0$ and the constant $N > 1$, we have

$$\begin{aligned}&(P_1 + N\varepsilon I)H^\top (H(P_1 + N\varepsilon I)H^\top + R)^{-1}H(P_1 + N\varepsilon I)\\ \succcurlyeq&(P_2 + \varepsilon I)H^\top (H(P_2 + \varepsilon I)H^\top + R)^{-1}H(P_2 + \varepsilon I).\end{aligned}$$

Since

$$\begin{aligned}&(P_1 + N\varepsilon I)H^\top (H(P_1 + N\varepsilon I)H^\top + R)^{-1}H(P_1 + N\varepsilon I),\\ &(P_2 + \varepsilon I)H^\top (H(P_2 + \varepsilon I)H^\top + R)^{-1}H(P_2 + \varepsilon I)\end{aligned}$$

are continuous in ε, then let $\varepsilon \to 0$, we can conclude that (3.26) still holds for the assumption $P_1 \succcurlyeq P_2 \succcurlyeq 0$. □

Then under the assumption $P(t_0|t_{-1}) = P_{TI}(t_0|t_{-1})$, we have

Theorem 3.2.2. *We assume that* $\delta c(t_0)$ *in* (3.14) *and* $\delta c_I(t_0)$ *in* (3.19) *have the same priori estimate* $\delta\hat{c}(t_0|t_{-1})$ *, then*

$$\begin{aligned}&P(t_k|t_k) \preccurlyeq P_{TI}(t_k|t_k), \quad P^c(t_k|t_k) \preccurlyeq P^c_{TI}(t_k|t_k)\\ &P^e(t_k|t_k) \preccurlyeq P^e_{TI}(t_k|t_k), \quad k \in 0, \cdots, N.\end{aligned}$$

Further, $P^c(t_0|t_N) \preccurlyeq P^c_{TI}(t_0|t_N)$ *and* $P^e(t_0|t_N) \preccurlyeq P^e_{TI}(t_0|t_N)$.

Proof. We firstly assume

$$P(t_k|t_{k-1}) \preccurlyeq P_{TI}(t_k|t_{k-1}). \tag{3.31}$$

According to (3.15) and the matrix inverse lemma [99],

$$\begin{aligned}&P(t_k|t_k)\\ =&\left[P^{-1}(t_k|t_{k-1}) + \begin{pmatrix} H^\top(t_k)R^{-1}(t_k)H(t_k) & 0\\ 0 & 0\end{pmatrix}\right]^{-1}\\ \preccurlyeq&\left[P^{-1}_{TI}(t_k|t_{k-1}) + \begin{pmatrix} H^\top(t_k)R^{-1}(t_k)H(t_k) & 0\\ 0 & 0\end{pmatrix}\right]^{-1}\\ =&\begin{pmatrix} P^c_I(t_k|t_{k-1}) + D^c_I(t_k) & D^{ce}_I(t_k)\\ D^{ec}_I(t_k) & D^e_I(t_k)\end{pmatrix}\\ &-\begin{pmatrix} P^c_I(t_k|t_{k-1}) + D^c_I(t_k) & D^{ce}_I(t_k)\\ D^{ec}_I(t_k) & D^e_I(t_k)\end{pmatrix}\begin{pmatrix} H^\top(t_k)\\ 0\end{pmatrix}\\ &\cdot\left[(H(t_k),0)\begin{pmatrix} P^c_I(t_k|t_{k-1}) + D^c_I(t_k) & D^{ce}_I(t_k)\\ D^{ec}_I(t_k) & D^e_I(t_k)\end{pmatrix}\begin{pmatrix} H^\top(t_k)\\ 0\end{pmatrix} + R(t_k)\right]^{-1}\\ &\cdot(H(t_k),0)\begin{pmatrix} P^c_I(t_k|t_{k-1}) + D^c_I(t_k) & D^{ce}_I(t_k)\\ D^{ec}_I(t_k) & D^e_I(t_k)\end{pmatrix}.\end{aligned}$$

From Lemma 3.2.1, we have

$$\begin{aligned}
&P(t_k|t_k)\\
\preccurlyeq&\begin{pmatrix} P_I^c(t_k|t_{k-1})+D_I^c(t_k) & D_I^{ce}(t_k)\\ D_I^{ec}(t_k) & D_I^e(t_k)\end{pmatrix}-\begin{pmatrix} P_I^c(t_k|t_{k-1}) & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix} H^\top(t_k)\\ 0\end{pmatrix}\\
&\cdot[(H(t_k),0)\begin{pmatrix} P_I^c(t_k|t_{k-1}) & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix} H^\top(t_k)\\ 0\end{pmatrix}+R(t_k)]^{-1}\\
&\cdot(H(t_k),0)\begin{pmatrix} P_I^c(t_k|t_{k-1}) & 0\\ 0 & 0\end{pmatrix}\\
=&P_I(t_k|t_k)+D_I(t_k)\\
=&P_{TI}(t_k|t_k).
\end{aligned}$$

Since $P(t_0|t_{-1})=P_{TI}(t_0|t_{-1})$ satisfies (3.31), we obtain

$$P(t_k|t_k)\preccurlyeq P_{TI}(t_k|t_k),\quad k\in\mathbb{N}.$$

Moreover, since

$$P^c(t_k|t_k)=(I,0)P(t_k|t_k)\begin{pmatrix} I\\ 0\end{pmatrix},\quad P_{TI}^c(t_k|t_k)=(I,0)P_{TI}(t_k|t_k)\begin{pmatrix} I\\ 0\end{pmatrix},$$

we obtain $P^c(t_k|t_k)\preccurlyeq P_{TI}^c(t_k|t_k)$. In the similar way, we get $P^e(t_k|t_k)\preccurlyeq P_{TI}^e(t_k|t_k)$.

Further, since

$$D_I(t_k)=M(t_k,t_0)D_I(t_0),$$

then we have

$$\begin{aligned}
P^c(t_0|t_N)=&(I,0)M^{-1}(t_N,t_0)P(t_N|t_N)M^{-\top}(t_N,t_0)\begin{pmatrix} I\\ 0\end{pmatrix}\\
\preccurlyeq&(I,0)M^{-1}(t_N,t_0)P_{TI}(t_0|t_N)M^{-\top}(t_N,t_0)\begin{pmatrix} I\\ 0\end{pmatrix}\\
=&P_{TI}^c(t_0|t_N).
\end{aligned}$$

Similarly, we obtain $P^e(t_0|t_N)\preccurlyeq P_{TI}^e(t_0|t_N)$. □

Joint optimization versus emission-rate-only optimization

In Section 3.2.4, as to the emission-rate-only optimization, we assumed $\delta\hat{c}_E(t_0|t_{-1})=0$. In order to compare the joint optimization and emission-rate-only optimization, the true analysis and forecasting covariances for the case only emission rates can be uniformly represented as:

$$\begin{aligned}
&P_{TE}(t_k|\cdot):=\begin{pmatrix} P_{TE}^c(t_k|\cdot) & P_{TE}^{ce}(t_k|\cdot)\\ P_{TE}^{ec}(t_k|\cdot) & P_{TE}^e(t_k|\cdot)\end{pmatrix}\\
=&\mathbb{E}\left[\begin{pmatrix} \delta\hat{c}_E(t_k|\cdot)-\delta c(t_k)\\ \delta\hat{e}_E(t_k|\cdot)-\delta e(t_k)\end{pmatrix}((\delta\hat{c}_E(t_k|\cdot)-\delta c(t_k))^\top,(\delta\hat{e}_E(t_k|\cdot)-\delta e(t_k))^\top\right].
\end{aligned}$$

where the dot "$\cdot$" could be any time step t_i, $i = 0, \cdots, N$.

In particular, we denote the initial prior covariance including the prior error of the concentrations by

$$P_{TE}(t_0|t_{-1}) = \begin{pmatrix} P^c_{TE}(t_0|t_{-1}) & 0 \\ 0 & P^e_{TE}(t_0|t_{-1}) \end{pmatrix}.$$

Since $P_{TE}(t_0|t_{-1})$ is also the first guess from the previous information, we can also assume that

$$P_{TE}(t_0|t_{-1}) = P(t_0|t_{-1}). \tag{3.32}$$

Defining the covariance of the *difference* between the model (3.14) and (3.23)

$$\begin{aligned} D_E(t_k) :=& \begin{pmatrix} D^c_E(t_k) & D^{ce}_E(t_k) \\ D^{ec}_E(t_k) & D^e_E(t_k) \end{pmatrix} \\ =& \mathbb{E}\left[\begin{pmatrix} \delta c_E(t_k) - \delta c(t_k) \\ \delta e_E(t_k) - \delta e(t_k) \end{pmatrix} (\delta c_E(t_k) - \delta c(t_k))^\top, (\delta e_E(t_k) - \delta e(t_k))^\top \right], \end{aligned}$$

then

$$\begin{aligned} & P_{TE}(t_k|\cdot) \\ =& \mathbb{E}\left[\begin{pmatrix} (\delta \hat{c}_E(t_k|\cdot) - \delta c_E(t_k)) + (\delta c_E(t_k) - \delta c(t_k)) \\ (\delta \hat{e}_E(t_k|\cdot) - \delta e_E(t_k)) + (\delta e_E(t_k) - \delta e(t_k)) \end{pmatrix} \right. \\ & \left. \cdot \begin{pmatrix} (\delta \hat{c}_E(t_k|\cdot) - \delta c_E(t_k)) + (\delta c_E(t_k) - \delta c(t_k)) \\ (\delta \hat{e}_E(t_k|\cdot) - \delta e_E(t_k)) + (\delta e_E(t_k) - \delta e(t_k)) \end{pmatrix}^\top \right] \\ =& P_E(t_k|\cdot) + D_E(t_k). \end{aligned}$$

Under the initial assumption (3.32), similar to Theorem 3.2.2, we have

Theorem 3.2.3. *We assume $P^e(t_0|t_{-1}) = P^e_E(t_0|t_{-1})$, then $P^c(t_k|t_k) \preccurlyeq P^c_{TE}(t_k|t_k)$, $P^e(t_k|t_k) \preccurlyeq P^e_{TE}(t_k|t_k)$ and $P(t_k|t_k) \preccurlyeq P_{TE}(t_k|t_k)$ hold for any $k \in \mathbb{N}$. Further, $P^c(t_0|t_N) \preccurlyeq P^c_{TE}(t_0|t_N)$ and $P^e(t_0|t_N) \preccurlyeq P^e_{TE}(t_0|t_N)$.*

Proof. We firstly assume

$$P(t_k|t_{k-1}) \preccurlyeq P_{TE}(t_k|t_{k-1}). \tag{3.33}$$

According to (3.15) and the matrix inverse lemma, we have

$$\begin{aligned} & P(t_k|t_k) \\ =& \left[P^{-1}(t_k|t_{k-1}) + \begin{pmatrix} H^\top(t_k) R^{-1}(t_k) H(t_k) & 0 \\ 0 & 0 \end{pmatrix} \right]^{-1} \\ \preccurlyeq& \left[P^{-1}_{TE}(t_k|t_{k-1}) + \begin{pmatrix} H^\top(t_k) R^{-1}(t_k) H(t_k) & 0 \\ 0 & 0 \end{pmatrix} \right]^{-1} \\ =& P_{TE}(t_k|t_{k-1}) - P_{TE}(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix} \\ & \cdot [(H(t_k), 0) P_{TE}(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix} + R(t_k)]^{-1} (H(t_k), 0) P_{TE}(t_k|t_{k-1}). \end{aligned}$$

By Lemma 3.2.1, we have

$$\begin{aligned}
&P(t_k|t_k)\\
\preccurlyeq &(P_E(t_k|t_{k-1}) + D_E(t_k)) - P_E(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix}\\
&\cdot [(H(t_k), 0) P_E(t_k|t_{k-1}) \begin{pmatrix} H^\top(t_k) \\ 0 \end{pmatrix} + R(t_k)]^{-1} (H(t_k), 0) P_E(t_k|t_{k-1})\\
=&P_E(t_k|t_{k-1}) + D_E(t_k)\\
=&P_{TE}(t_k|t_{k-1}).
\end{aligned}$$

Since $P(t_0|t_{-1}) = P_{TE}(t_0|t_{-1})$, which satisfies (3.33), we can conclude that

$$P(t_k|t_k) \preccurlyeq P_{TE}(t_k|t_k), \quad k \in \mathbb{N}.$$

Further, similar with Theorem 3.2.2, we obtain

$$\begin{aligned}
P^c(t_k|t_k) \preccurlyeq P^c_{TE}(t_k|t_k), \quad P^e(t_k|t_k) \preccurlyeq P^e_{TE}(t_k|t_k),\\
P^c(t_0|t_N) \preccurlyeq P^c_{TE}(t_0|t_N), \quad P^e(t_0|t_N) \preccurlyeq P^e_{TE}(t_0|t_N).
\end{aligned}$$

□

Now we can conclude that the joint optimization of the initial value and emission rates based on the atmospheric transport model extended by emissions can provide the same or better estimates for both the initial value and emission rates than the initial-value-only optimization and emission-rate-only optimization.

3.2.6 Application to EnKF and EnKS

In this section we give a basic example by applying the ensemble Kalman filter and smoother into a one-dimensional transport model extended by emission rates. It illustrates that the daily profile of the emission rate can be preserved only by the ensemble Kalman smoother rather than the ensemble Kalman filter.

We consider the following one-dimensional transport equation with the periodic boundary condition on the domain $[0, 14]$,

$$\frac{\partial \delta c}{\partial t} + v \frac{\partial \delta c}{\partial x} = e(t), \tag{3.34}$$

where δc, δe are the perturbations of the concentration, emission rate respectively, $v = 0.5$ represents the wind speed. Within the spacial domain $[0, 14]$ and data assimilation window (DAW) $[0, 23.5]$, applying the Lax-wendroff scheme, we discretize the transport equation (3.34) in space with $\triangle x = 0.5$ and in time with $\triangle t = 0.5$. Thus, the dimension of the state space is $N_x = 30$ and there are $T = 48$ time steps within the entire data assimilation window.

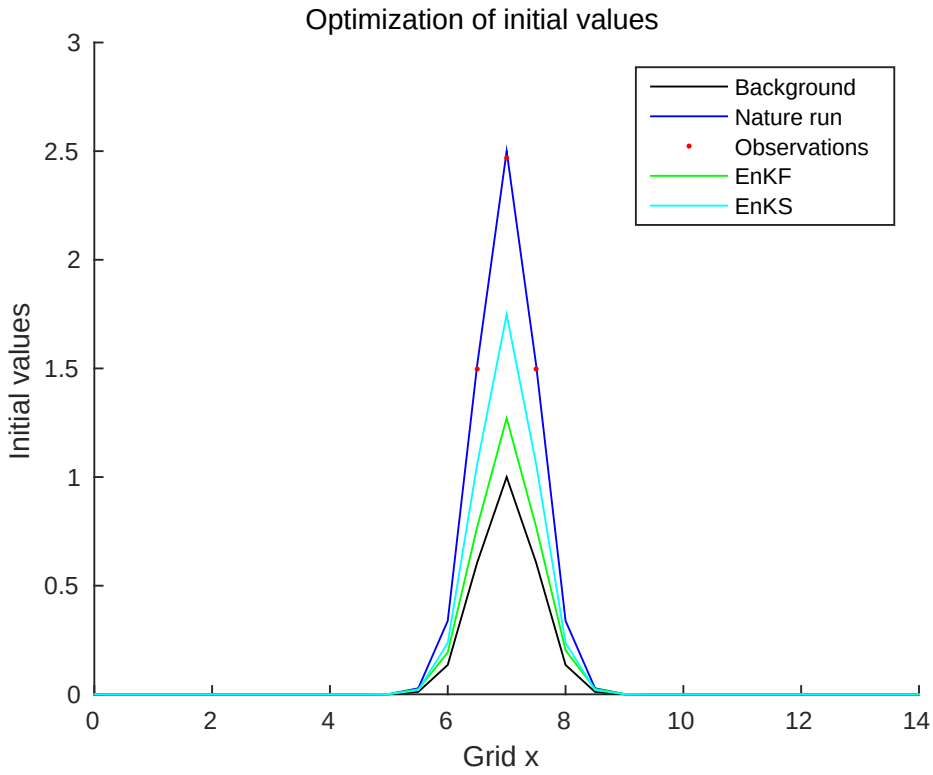

Figure 3.1: BLUEs of the initial value by EnKF and EnKS.

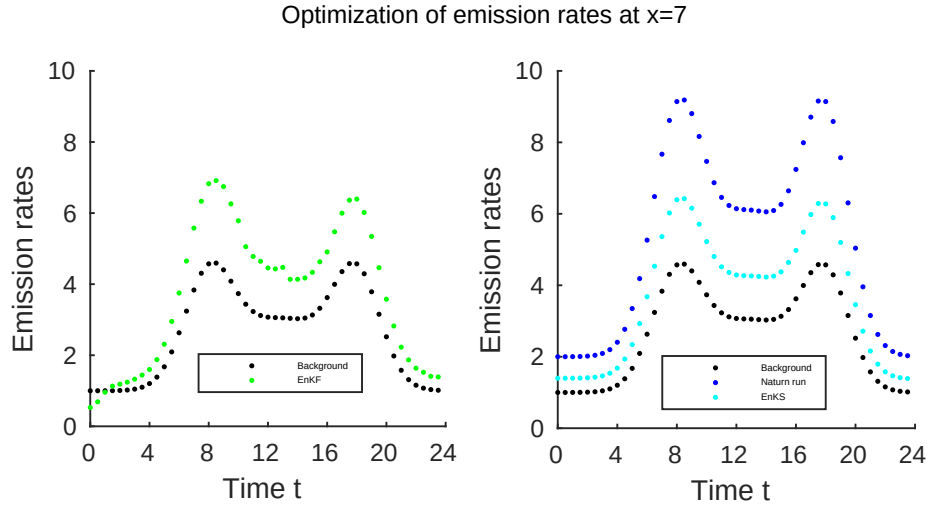

Figure 3.2: BLUEs of the emission rate at $x = 7$ within the entire data assimilation window by EnKF and EnKS.

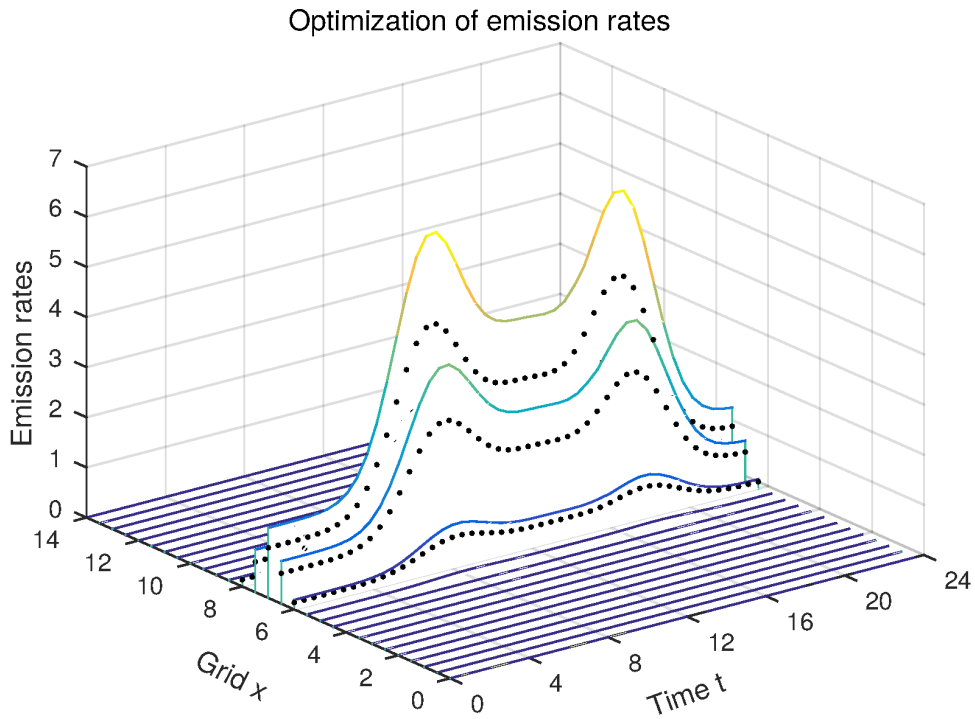

Figure 3.3: BLUEs of the emission rates at all grid points within the entire data assimilation window by EnKF and EnKS. The dots are the diurnal profiles of the background emission rates. The lines are BLUEs of the emission rates.

With the same temporal and spacial discretization of $c(t)$, we write the background information of emission rate into a vector

$$e_b(t_n) = (e_b^1(t_n), \cdots, e_b^{N_x}(t_n))^\top, \quad n = 1, \cdots, T,$$

which is shown in Fig. 3.2 and Fig. 3.3. Then we extract the transition matrix of the dynamic model of the emission rate from the background of the emission rate

$$M_e(t_{n+1}, t_n) = \begin{pmatrix} \frac{e_b^1(t_{n+1})}{e_b^1(t_n)} & & & \\ & \frac{e_b^2(t_{n+1})}{e_b^2(t_n)} & & \\ & & \ddots & \\ & & & \frac{e_b^{15}(t_{n+1})}{e_b^{15}(t_n)} \end{pmatrix}, \quad n = 0, \cdots, T-1.$$

Further, we establish the discrete dynamic model of the perturbation of the emission rate according to (4.36)

$$\delta e(t_{n+1}) = M_e(t_{n+1}, t_n)\delta e(t_n), \quad n = 1, \cdots, T-1.$$

Besides, we generate three observations of the concentration at $x = 6.5, 7$ and 7.5 at each time step. The observation errors follow the Gaussian distributions of zero mean and diagonal covariance matrix with the diagonal $0.2c_n(t)$, where $c_n(t)$ represents the nature run of the concentrations. The plot convection of observations in Fig. 3.1 is given by dots.

Fig. 3.1 shows that the optimal estimates of the initial values by the data assimilation procedure based on the ensemble Kalman filter and smoother. Obviously, the BLUEs of the initial values based on the EnKS are better than the BLUEs based on the EnKF since the estimates from the EnKS are closer to the given nature run $c_n(t)$.

Fig. 3.2 shows that the diurnal profiles of the background, nature run and the estimates of the emission rate at $x = 7$. Since in general

$$\mathbb{E}(e(t_{n+1})|Y_{t_{n+1}}) \neq M(t_{n+1}, t_n)\mathbb{E}(e(t_n)|Y_{t_n}),$$

the left panel of Fig. 3.2 indicates that the optimization of the emission rates from the ensemble Kalman filter may not preserve the same diurnal profile as the background emission rates. From the right panel of Fig. 3.2 we can find that the optimal estimates of emission rates from EnKS follow the same evolution of the background emission rates, which verifies (3.10). Similarly, Fig. 3.3 shows the BLUEs of the emission rates at all grid points within the entire data assimilation window.

Chapter 4

Efficiency and Sensitivity Analysis of Observational Networks

In previous sections, we have reviewed the current common-used data assimilation approaches and introduced our novel atmospheric transport model extended by emission rates, aiming at obtaining the better estimation of the model state with limited observations. As mentioned before, the better estimations of both the initial state and emission rates are probably not achieved by certain observational network configurations. It is worth to address the qualification problem in order to evaluate the capacity of an observational network to control and effect chemical states and emission rates quantitatively in advance of data assimilation procedure.

In fact, in atmospheric chemistry, studies about the importance of observations are still sparse. Khattatov et al. [54] firstly analysed the uncertainty of a chemical compositions. Liao et al. [65] focused on the optimal placement of observation locations of the chemical transport model. Starting with a given sensor network, Singh et al. [90] introduced theoretical metrics to quantify the value of measurements to reduce the analysis error in the frame of ensemble runs. For accidental releases, Abida and Bocquet [1] sought to reconstruct the plume of emitted compounds by sequentially optimizing observation locations for mobile monitor platforms. However, singular vector analysis and other methods for atmospheric chemistry with emissions are different since emissions play a similarly important role in forecast accuracy with initial values. Goris and Elbern [47] recently used the singular vector decomposition to determine the sensitivity of the chemical composition to emissions and initial values for a variety of chemical scenarios and integration length. This methodology has been generalized for the 3-dimensional EURAD-IM (European Air pollution Dispersion-Inverse Model) and applied to a field campaign with airship borne measurements in Goris and Elbern [48]. While that paper describes an approach to optimize an atmospheric chemistry observation network, both in terms of individual compounds to be observed with preference and their location, the assessment of the information potential of an established and mainly fixed observation network, like for example the AIRNow Air Quality Monitor Maps (`http://www.airnow.gov/index.cfm?action=airnow.pointmaps`) in the United State or from the European Environment Agency (`http://www.eea.europa.eu/data-and-maps/`

explore-interactive-maps#c5=&c0=5&b_start=0) needs a different algorithmic approach. Most measurement devices monitor concentrations hourly or half-hourly. In practice, the deployment of in situ observations follows manly legal requirements as manifested in official regulations to monitor concentration threshold violations for public healthcare with emphasis on populated area. This is in stark contrast to observation network design principles of weathers services, which nearly exclusively aspire to comply with data assimilation requirements to optimise initial values for predictions. With the growing importance of earth system modelling and its combination with measurements, existing observation networks need to be validated for forecasting purposes. For this, attention should be paid to the fact that in atmospheric chemistry emission rates are also candidate parameters for optimisation, as they are typically both, insufficiently well known and of high impact on the simulation. Moreover, network assessment results are dependent on meteorological conditions, most notably wind direction and vertical exchange.

Using the Kalman smoother as the required data assimilation method we introduce a novel approach to identify the efficiency and sensitivity of the observation networks for controlling linear tangent diffusion models. This chapter is organized as follows. In Section 4.1, based on the Kalman smoother, we derive the theoretical approach to determine the efficiency of observation networks for both discrete-time and continuous-time systems. In Section 4.2, we develop the ensemble approach to evaluate the efficiency of observation configurations. A 3D advection-diffusion equation is extended by the dynamic model of the emission rate and several elementary experiments are given to verify the approaches. In Section 4.3, we present the approach to identify the sensitivity of observations by determining the directions of maximum perturbation growth to the initial perturbation and focus on the relationship between the efficiency and sensitivity analysis of observation networks. In Section 4.4, based on the atmospheric transport model extended by emission rates, we apply the sensitivity analysis of observation networks into the emission source apportionment problem in order to distinguish the different emission sources and determine their apportionments.

4.1 Efficiency analysis of observational networks

In the case that the estimation of both the initial state and emission rates can be improved significantly, we say that the corresponding observation configurations are efficient for both the initial state and emission rates. Otherwise, the observation configurations are only efficient to initial state or emission rates, or even to none in case of undue sparseness of measurements. However, it is usually difficult to foresee the efficiency of observation configurations. The lack of the knowledge of the efficiency of observations may lead us to give the poor initial guesses, imbalanced results and wasted computational resources. In this section we will introduce the theoretical approach to determine the efficiency of observations by the Kalman smoother within a finite-time interval.

4.1.1 Efficiency analysis for discrete-time systems

We generalize the atmospheric transport model extended by emission rates (3.12) and the observation system (3.13) into the following discrete-time linear system on the time interval $[t_0, t_1, \cdots, t_N]$

$$x(t_{k+1}) = M(t_{k+1}, t_k)x(t_k) + \varepsilon(t_k), \tag{4.1}$$
$$y(t_k) = H(t_k)x(t_k) + \nu(t_k), \tag{4.2}$$

where $x(\cdot) \in \mathbb{R}^n$ is the state variable, $y(t_k) \in \mathbb{R}^{m(t_k)}$ is the observation vector at time t_k, the model error $\varepsilon(t_k)$ and the observation error $\nu(t_k)$, $k = 1, \cdots, N$ follow Gaussian distributions with zero means and the model error covariance matrix $Q(t_k)$ and observation error $R(t_k)$, respectively.

We denote the first guess of initial variance by $P(t_0|t_{-1})$ and assume that $P(t_0|t_{-1})$ and $R(t_k)$ to be symmetric and positive definite matrices. Then applying the matrix inverse lemma [99] into (2.7), we have

$$P^{-1}(t_k|t_k) = P^{-1}(t_k|t_{k-1}) + H^\top(t_k)R^{-1}(t_k)H(t_k). \tag{4.3}$$

Assuming the model errors $\varepsilon(t_i)$, $i = 1, \cdots, N$ and consequently also the model error covariance matrix $Q(t_i)$, are negligible, we obtain

$$P^{-1}(t_{k+1}|t_k) = M^{-\top}(t_{k+1}, t_k)P^{-1}(t_k|t_k)M^{-1}(t_{k+1}, t_k). \tag{4.4}$$

Hence, by the deduction based on (4.3) and (4.4), we have

$$\begin{aligned}
& P^{-1}(t_{k+1}|t_k) \\
= \; & M^{-\top}(t_{k+1}, t_k)P^{-1}(t_k|t_{k-1})M^{-1}(t_{k+1}, t_k) \\
& + M^{-\top}(t_{k+1}, t_k)H^\top(t_k)R^{-1}(t_k)H(t_k)M^{-1}(t_{k+1}, t_k) \\
= \; & M^{-\top}(t_{k+1}, t_{k-1})P^{-1}(t_{k-1}|t_{k-2})M^{-1}(t_{k+1}, t_{k-1}) \\
& + M^{-\top}(t_{k+1}, t_{k-1})H^\top(t_{k-1})R^{-1}(t_{k-1})H(t_{k-1})M^{-1}(t_{k+1}, t_{k-1}) \\
& + M^{-\top}(t_{k+1}, t_k)H^\top(t_k)R^{-1}(t_k)H(t_k)M^{-1}(t_{k+1}, t_k) \\
= \; & M^{-\top}(t_{k+1}, t_0)P^{-1}(t_0|t_{-1})M^{-1}(t_{k+1}, t_0) \\
& + \sum_{i=0}^{k} M^{-\top}(t_{k+1}, t_i)H^\top(t_i)R^{-1}(t_i)H(t_i)M^{-1}(t_{k+1}, t_i).
\end{aligned}$$

Thus, the covariance of the estimate of the initial state by the fixed-interval Kalman smoother [69] is given by

$$\begin{aligned}
& P^{-1}(t_0|t_k) \\
= \; & \mathbb{E}[(x(t_0) - \hat{x}(t_0|t_k))(x(t_0) - \hat{x}(t_0|t_k))^\top] \\
= \; & \mathbb{E}[M^{-1}(t_{k+1}, t_0)(x(t_{k+1}) - \hat{x}(t_{k+1}|t_k))(x(t_0) - \hat{x}(t_{k+1}|t_k))^\top M^{-\top}(t_{k+1}, t_0)]^{-1} \\
= \; & M^\top(t_{k+1}, t_0)P^{-1}(t_{k+1}|t_k)M(t_{k+1}, t_0) \\
= \; & P^{-1}(t_0|t_{-1}) + \sum_{i=0}^{k} M^\top(t_i, t_0)H^\top(t_i)R^{-1}(t_i)H(t_i)M(t_i, t_0).
\end{aligned} \tag{4.5}$$

In particular, for the case $k = N$, which indicates to take the observations in the entire time interval into account, we have

$$P^{-1}(t_0|t_N) = P^{-1}(t_0|t_{-1}) + \sum_{i=0}^{N} M^\top(t_i, t_0) H^\top(t_i) R^{-1}(t_i) H(t_i) M(t_i, t_0), \quad (4.6)$$

which is the inverse of the analysis error covariance matrix and the optimal Hessian of the underlying cost function of 4D-Var (2.4). It implies the equivalence between Kalman smoother and 4D-Var method for linear models and guarantees that the following approach in this paper is available for the data assimilation based on 4D-Var method.

It is clear that (4.6) comprises the information of the initial condition, model evolution, observation configurations and errors over the entire time interval $[t_0, \cdots, t_N]$. At the same time, it is independent of any specific data and states unknown before fulfilling the data assimilation procedure, apart from the reference model evolution $M(\cdot, \cdot)$ needed for the linearisation, as well as the observation operator $H(\cdot)$. In fact, if we define

$$\mathcal{G} = \begin{pmatrix} H(t_0)M(t_0, t_0) \\ H(t_1)M(t_1, t_0) \\ \vdots \\ H(t_N)M(t_N, t_0) \end{pmatrix}, \ \mathcal{R}^{-1} = \begin{pmatrix} R^{-1}(t_0) & & & \\ & R^{-1}(t_1) & & \\ & & \ddots & \\ & & & R^{-1}(t_N) \end{pmatrix}, \quad (4.7)$$

we can rewrite (4.6) as

$$P^{-1}(t_0|t_N) = P^{-1}(t_0|t_{-1}) + \mathcal{G}^\top \mathcal{R}^{-1} \mathcal{G}, \quad (4.8)$$

where $\mathcal{G}^\top \mathcal{R}^{-1} \mathcal{G}$ is the observability Gramian with respect to $\mathcal{R}^{-1}$ in control theory [11], [103]. It represents the observable capacity of a model.

Though (4.8) meets the demand to represent the estimate covariance by all available information before starting the data assimilation procedure, it cannot be applied directly to evaluate the potential improvement of the estimate by the Kalman smoother due to the lack of clear statistical significance. Thus, aspiring a means to compare efficiencies with respect to initial values and emission rates in a scaled way, we define *the relative improvement covariance* as

$$\begin{aligned} & P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\ =& I - P^{-\frac{1}{2}}(t_0|t_{-1})P(t_0|t_N)P^{-\frac{1}{2}}(t_0|t_{-1}). \end{aligned} \quad (4.9)$$

The relative improvement covariance (4.9) is a normalized matrix of the difference between the initial variance $P(t_0|t_{-1})$ and the covariance matrix $P(t_0|t_N)$ from the Kalman smoother. Especially, $P^{-\frac{1}{2}}(t_0|t_{-1})P(t_0|t_N)P^{-\frac{1}{2}}(t_0|t_{-1})$ can be understood as the covariance matrix from the fixed-interval Kalman smoother normalized by the initial variance. Thus, the symmetry of (4.9) guarantees the relative improvement covariance matrix to be nonnegative-definite. In fact,

$$0 \preccurlyeq P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \prec I, \quad (4.10)$$

where the left equality holds for the situation that there is no observation within the entire time interval $[t_0, \cdots, t_N]$. Further, (4.10) implies that singular values of the relative improvement covariance are bounded by n since their sum is less than n, the trace of I.

Since $P(t_0|t_N)$ is unknown prior to the data assimilation procedure, we use (4.8) to rewrite the relative improvement covariance as

$$\begin{aligned}
\tilde{P} &:= P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\
&= P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - (P^{-1}(t_0|t_{-1}) + \mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G})^{-1})P^{-\frac{1}{2}}(t_0|t_{-1}) \\
&= I - P^{-\frac{1}{2}}(t_0|t_{-1})(P^{-1}(t_0|t_{-1}) + \mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G})^{-1}P^{-\frac{1}{2}}(t_0|t_{-1}) \\
&= I - (I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1}))^{-1}. \qquad (4.11)
\end{aligned}$$

It is worth noting that in (4.11),

$$I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1})$$

is always invertible even if the observation Gramian $\mathcal{G}^\top\mathcal{G}$ is not full-rank. Thus, the relative improvement covariance is well-defined for all models with invertible initial covariance and observation systems with invertible error covariances within $[t_0, \cdots, t_N]$. Due to the high computational cost of (4.11), we apply the singular value decomposition to

$$P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top \mathcal{R}^{-\frac{1}{2}} = VSU^\top,$$

where V and U are unitary matrices consisting of the left and right singular vectors, while S is the rectangular diagonal matrix consisting of the singular values.

The normalized covariance $\tilde{P}$ can be simplified as

$$\begin{aligned}
\tilde{P} &= I - (I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1}))^{-1} \\
&= I - (I + VSS^\top V^\top)^{-1} \\
&= VV^\top - (VV^\top + VSS^\top V^\top)^{-1} \\
&= VV^\top - (V(I + SS^\top)V^\top)^{-1} \\
&= V(I - (I + SS^\top)^{-1})V^\top \\
&= \sum_{i=1}^{r} \frac{s_i^2}{1 + s_i^2} v_i v_i^\top, \qquad (4.12)
\end{aligned}$$

where r is the rank of (4.11) and v_i is the i^{th} left singular vector in V related to the singular value s_i, which is the i^{th} element on the diagonal of S.

The relative improvement of each element in the state vector x is given by the corresponding value in the diagonal of the relative improvement covariance, and remains to be specified. Denoting the relative improvement of j^{th} element in $x(t_0)$ by $\tilde{P}_j$, from (4.12) we have

$$\tilde{P}_j = \sum_{i=1}^{r} \frac{s_i^2}{1 + s_i^2} (v_{ij})^2,$$

where v_{ij} is the j^{th} element of v_i.

In order to get a deeper insight into the capacity of the observation networks to improve the estimation of all model states, we consider the sum of the diagonal entries of the relative improvement matrix as the evaluation of the total relative improvement of the model. Thus the 1-norm, also named as the nuclear norm, is appropriately taken as the metric, which is defined as

$$\|A\|_1 = \text{tr}(\sqrt{A^\top A}),$$

where A is any matrix and $\text{tr}(\cdot)$ denotes the trace of the matrix.

From (4.12), we obtain

$$\|\tilde{P}\|_1 = \text{tr}(\tilde{P}) = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2},$$

which is called the *total improvement value*. According to [85], it can be also considered as the degree of signal freedom.

As mentioned before, $\|\tilde{P}\|_1 < \|I\|_1 = n$. Here n is considered as the total relative improvement if the system is fully observed. Thus, if we consider the ratio

$$\tilde{p} = \frac{\|\tilde{P}\|_1}{\|I\|_1} = \frac{\|\tilde{P}\|_1}{n} \in [0,1), \tag{4.13}$$

the percentage of the total improvement of the model is obtained, which is called the *relative improvement degree*.

4.1.2 Efficiency analysis of the atmospheric transport model extended by emission rates

For the atmospheric transport model extended with emissions, without loss of generality, we simply assume the original state $c \in \mathbb{R}^n$ and emission rates $e \in \mathbb{R}^n$. We divide (4.11) into the following block matrix according to the dimension of c and e

$$\tilde{P} = \begin{pmatrix} \tilde{P}^c & \tilde{P}^{ce} \\ \tilde{P}^{ec} & \tilde{P}^e \end{pmatrix} = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} \begin{pmatrix} v_i^c \\ v_i^e \end{pmatrix} (v_i^{c\top}, v_i^{e\top}) \in \mathbb{R}^{2n\times 2n},$$

where P^c is the relative improvement covariance of the state $c(t_0)$, $\tilde{P}^e$ is the relative improvement covariance of the emission rates $e(t_0)$, $\tilde{P}^{ce} = (\tilde{P}^{ec})^\top$ is the relative improvement covariance between $c(t_0)$ and $e(t_0)$ and $(v_i^{c\top}, v_i^{e\top})^\top = v_i$.

It is easy to see that

$$\tilde{P}^c = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} v_i^c v_i^{c\top}, \quad \tilde{P}^e = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} v_i^e v_i^{e\top}.$$

Further, the relative improvements of j^{th} element in $c(t_0)$ and $e(t_0)$ are given by

$$\tilde{P}_j^c = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} (v_{ij}^c)^2, \quad \tilde{P}_j^e = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} (v_{ij}^e)^2,$$

where v_{ij}^c and v_{ij}^e are the j^{th} elements of v_i^c and v_i^e respectively.

Moreover, the total improvement values of concentration and emission rates are

$$\|\tilde{P}^c\|_1 = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} \mathrm{tr}(v_i^c v_i^{c^\top}), \quad \|\tilde{P}^e\|_1 = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2} \mathrm{tr}(v_i^e v_i^{e^\top}).$$

It is worth noticing that

$$\tilde{P}^c = (P^c(t_0|t_{-1}))^{-\frac{1}{2}}(P^c(t_0|t_{-1}) - P^c(t_0|t_N))(P^c(t_0|t_{-1}))^{-\frac{1}{2}},$$
$$\tilde{P}^e = (P^e(t_0|t_{-1}))^{-\frac{1}{2}}(P^e(t_0|t_{-1}) - P^e(t_0|t_N))(P^e(t_0|t_{-1}))^{-\frac{1}{2}}$$

if and only if there is no prior correlation between the initial concentration and emission rates. In this case $P^{ce}(t_0|t_{-1}) = 0_{n\times n}$, the corresponding relative improvement degrees of concentration and emission rates are defined as

$$\tilde{p}^c = \frac{\|\tilde{P}^c\|_1}{n}, \quad \tilde{p}^e = \frac{\|\tilde{P}^e\|_1}{n}.$$

From (4.13), it is obvious that $\tilde{p}^c \in [0,1)$ and $\tilde{p}^e \in [0,1)$ show the percentages of the relative improvements of concentration and emission rates, respectively. However, efficient observation networks probably lead both of them to be close to 1 such that

$$\frac{\|\tilde{P}^c\|_1}{n} + \frac{\|\tilde{P}^e\|_1}{n} > 1.$$

It indicates the normalization of $\tilde{P}$ is only with respect to the extended covariance matrix rather than specified to the state c and emission rates e. The relative improvement degree cannot serve our objective to distinguish the observability of concentration and emission rates and balance them quantitatively. However, by observing the block form of $\tilde{P}$, we have

$$\|\tilde{P}^c\|_1 + \|\tilde{P}^e\|_1 = \|\tilde{P}\|_1.$$

Thus, in order to compare the improvements of the concentration and emission rates, we define *relative improvement ratios* for concentrations or emission rates as

$$\tilde{p}^c = \frac{\|\tilde{P}^c\|_1}{\|\tilde{P}\|_1}, \quad \tilde{p}^e = \frac{\|\tilde{P}^e\|_1}{\|\tilde{P}\|_1}, \quad \tilde{p}^e + \tilde{p}^c \equiv 1.$$

If the total improvement value or relative improvement degree of the model is almost zero, an improvement cannot be expected. In contrast, $\{\tilde{P}_j^c\}_{j=1}^n$ and $\{\tilde{P}_j^e\}_{j=1}^n$, which show the improvement of each parameter j of concentrations and emission rates respectively, can help us determining which parameters can be optimized by the existing observation configurations. Furthermore, comparing $\tilde{p}^c$ with $\tilde{p}^e$, we can conclude that the estimate of that one with the larger relative improvement ratio can be improved more efficiently by the existing observation configurations than the other. In other words, if $\tilde{p}^c > \tilde{p}^e$, the existing observation configurations are more efficient to the initial values of

concentrations. Conversely, if $\tilde{p}^c < \tilde{p}^e$, the observation configurations can help improving the estimate of emission rates more. According to $\tilde{p}^c$ and $\tilde{p}^e$, the "weights" between the concentrations and emission rates can be quantitatively identified. In a data assimilation context, where observations are in a weighted relation to the background, the BLUE favours those parameters with higher observation efficiency.

The special case that $\tilde{p}^e$ is very close to zero implies that observation network is nearly "blind" for emission rate optimization.

4.1.3 Efficiency analysis for continuous-time systems

In this section we generalize the efficiency analysis of discrete-time systems into continuous-time systems in order to provide the possibility of wider applications of this approach. We consider the continuous-time system

$$x(t) = M(t, t_0)x(t_0) + \varepsilon(t), \tag{4.14}$$
$$y(t) = H(t)x(t) + \nu(t), \tag{4.15}$$

where $x(t) \in \mathbb{R}^n$ is the state variable, $y(t) \in \mathbb{R}^{m(t)}$ is observation vector, the model error $\varepsilon(t)$ and the observation error $\nu(t)$, $t \in [t_0, t_N]$ follow Gaussian distributions with zero mean, while $Q(t)$ and $R(t)$ are their covariance matrices respectively.

As in Section 4.1.1, we ignore the model error. We assume

$$M_K(t, t_0) := M(t, t_0) - \int_{t_0}^{t} M(t, s)K(s)H(s)M_K(s, t_0)ds,$$

where $K(t) = P(t|t)H(t)R^{-1}(t)$. According to (2.5), we calculate that

$$P(t|t) = M_K(t, t_0)P(t_0|t_{-1})M_K^\top(t, t_0) + \int_{t_0}^{t} M_K(t, s)K(s)R(s)K^\top(s)M_K^\top(t, s)ds.$$

On one hand,

$$\begin{aligned}
& M_K(t, t_0)P(t_0|t_{-1})M_K^\top(t, t_0) \\
= \; & M(t, t_0)P(t_0|t_{-1})M_K^\top(t, t_0) \\
& - \int_{t_0}^{t} M(t, s)K(s)H(s)M_K(s, t_0)P(t_0|t_{-1})M_K^\top(t, t_0)ds \\
= \; & M(t, t_0)P(t_0|t_{-1})M_K^\top(t, t_0) - \int_{t_0}^{t} M(t, s)K(s)H(s)P(s|s)M_K^\top(t, s)ds \\
& + \int_{t_0}^{t}\int_{t_0}^{s} M(t, s)K(s)H(s)M_K(s, \eta)K(\eta)R(\eta)K^\top(\eta)M_K^\top(t, \eta)d\eta ds.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\int_{t_0}^{t} M_K(t,s)K(s)R(s)K^\top(s)M_K^\top(t,s)ds \\
=&\int_{t_0}^{t}[M(t,s)-\int_{s}^{t} M(t,\eta)K(\eta)H(\eta)M_K(\eta,s)d\eta]K(s)R(s)K^\top(s)M_K^\top(t,s)ds \\
=&\int_{t_0}^{t} M(t,s)K(s)R(s)K^\top(s)M_K^\top(t,s)ds \\
&-\int_{t_0}^{t}\int_{t_0}^{\eta} M(t,\eta)K(\eta)H(\eta)M_K(\eta,s)K(s)R(s)K^\top(s)M_K^\top(t,s)dsd\eta \\
=&\int_{t_0}^{t} M(t,s)K(s)R(s)K^\top(s)M_K^\top(t,s)ds \\
&-\int_{t_0}^{t}\int_{t_0}^{s} M(t,s)K(s)H(s)M_K(s,\eta)K(\eta)R(\eta)K^\top(\eta)M_K^\top(t,\eta)d\eta ds.
\end{aligned}$$

Therefore, $P(t|t) = M(t,t_0)P(t_0|t_{-1})M_K^\top(t,t_0)$.

Since

$$\begin{aligned}
&M^{-1}(t,t_0) = M_K^{-1}(t,t_0)M_K(t,t_0)M^{-1}(t,t_0) \\
=\;& M_K^{-1}(t,t_0)[M(t,t_0)-\int_{t_0}^{t} M_K(t,s)K(s)H(s)M(s,t_0)ds]M^{-1}(t,t_0) \\
=\;& M_K^{-1}(t,t_0)-\int_{t_0}^{t} M_K^{-1}(s,t_0)L(s)H(s)M(t,s)ds,
\end{aligned}$$

we obtain

$$M_K^{-1}(t,t_0) = M^{-1}(t,t_0)+\int_{t_0}^{t} M_K^{-1}(s,t_0)K(s)H(s)M^{-1}(t,s)ds.$$

Hence,

$$\begin{aligned}
&P^{-1}(t_0|t) \\
=&[M^{-1}(t,t_0)P(t|t)M^{-\top}(t,t_0)]^{-1} \\
=&[P(t_0|t_{-1})M_K^\top(t,t_0)M^{-\top}(t,t_0)]^{-1} \\
=&M^\top(t,t_0)[M^{-\top}(t,t_0)+\int_{t_0}^{t} M^{-\top}(t,s)H^\top(s)K^\top(s)M_K^{-\top}(s,t_0)ds]P^{-1}(t_0|t_{-1}) \\
=&P^{-1}(t_0|t_{-1})+\int_{t_0}^{t} M^\top(s,t_0)H^\top(s)R^{-1}(s)H(s)M(s,t_0)ds.
\end{aligned}$$

Let $t = t_N$ and define the observability mapping $\mathcal{G}:\mathbb{R}^n \to L^2([t_0,t_N];\mathbb{R}^m)$ as

$$\mathcal{G}f := H(\cdot)M(\cdot,t_0)f, \quad f\in\mathbb{R}^n,$$

its adjoint operator $\mathcal{G}^*$ is

$$\mathcal{G}^* f = -\int_{t_N}^{t_0} M^\top(s,t_0)H^\top(s)f(s)ds, \quad f \in L^2([t_0,t_N];\mathbb{R}^m).$$

Further, we define $\mathcal{R}^{-1} : \mathcal{L}^2([t_0,t_N];\mathbb{R}^m) \to L^2([t_0,t_N];\mathbb{R}^m)$,

$$\mathcal{R}^{-1} f := R^{-1}(\cdot)f(\cdot), \quad f \in \mathcal{L}^2([t_0,t_N];\mathbb{R}^m).$$

Thus,

$$P^{-1}(t_0|t_N) = P^{-1}(t_0|t_{-1}) + \mathcal{G}^*\mathcal{R}^{-1}\mathcal{G}, \tag{4.16}$$

where $\mathcal{G}^*\mathcal{R}^{-1}\mathcal{G}$ is the observability Gramian of continuous-time systems, see [11] and [103]. By the singular value decomposition,

$$P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^*\mathcal{R}^{-\frac{1}{2}} = VSU^\top,$$

we obtain

$$\begin{aligned} & P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}) \\ = \;& I - (I + P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^*\mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1}))^{-1} \\ = \;& V(I - (I + SS^\top)^{-1})V^\top, \end{aligned} \tag{4.17}$$

where V and U are unitary matrices consisting of the left and right singular vectors, S is the rectangular diagonal matrix consisting of the singular values.

It is clear now that (4.16) and (4.17) has the same pattern with (4.6) and (4.12). Thus, following the similar steps as in Section 4.1, we also choose the 1-norm as the metric. If we denote

$$\tilde{P} = P^{-\frac{1}{2}}(t_0|t_{-1})(P(t_0|t_{-1}) - P(t_0|t_N))P^{-\frac{1}{2}}(t_0|t_{-1}), \tag{4.18}$$

then $\tilde{P}_j$, the relative improvement of j^{th} element in $x(t_0)$ of the continuous system (4.14) is also given by

$$\tilde{P}_j = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2}(v_{ij})^2,$$

where v_{ij} is the j^{th} element of v_i.

Correspondingly, the total improvement value and the relative improvement degree of the continuous system are respectively given by

$$\|\tilde{P}\|_1 = \text{tr}(\tilde{P}) = \sum_{i=1}^{r} \frac{s_i^2}{1+s_i^2}, \quad \tilde{p} = \frac{\|\tilde{P}\|_1}{\|I\|_1} = \frac{\|\tilde{P}\|_1}{n} \in [0,1).$$

4.2 The ensemble approach for the efficiency analysis

In this section we develop the ensemble approach to analyse the efficiency of observational networks. It provides the feasibility to apply this approach for high dimensional problems in practice.

4.2.1 The ensemble approach for discrete-time systems

Based on the ensemble Kalman filter and smoother introduced in Section 2.2.3, we further introduce some notations. For the discrete-time system (4.1), we define the ensemble observation configurations in the entire time interval as

$$y_k^f = \mathcal{G}\hat{x}_k(t_0|t_{-1}), \quad k = 1, \cdots, q.$$

And the ensemble mean and the forecast error covariance matrix of the ensemble observation configurations are given by

$$\bar{y}^f = \frac{1}{q}\sum_{k=1}^{q} y_k^f, \quad \bar{P}_{yy}^f = \frac{1}{q-1}\sum_{k=1}^{q}(\hat{y}_k^f - \bar{y}^f)(\hat{y}_k^f - \bar{y}^f)^\top = \mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^\top.$$

Similarly, we denote the ensemble covariance between the initial states and the forecasting observations by

$$\bar{P}_{xy}^f = \frac{1}{q-1}\sum_{k=1}^{q}(\hat{x}_k(t_0|t_{-1}) - \bar{x}(t_0|t_{-1}))(\hat{y}_k^f - \bar{y}^f)^\top = \bar{P}(t_0|t_{-1})\mathcal{G}^\top.$$

It is shown by Evensen [37] that the ensemble forecasting and analysis covariances have the same form as the covariances in the standard Kalman filter. It indicates that (2.7) and (2.8) are also true for $\bar{P}(t_i|t_i)$ and $\bar{P}(t_i|t_{i-1})$. However, the ensemble size q is usually less than the dimension of the model n in the real world. It causes (4.3), (4.4) and further (4.9) to be infeasible since the initial ensemble covariance $\bar{P}(t_0|t_{-1})$ is not invertible. In this case, the pseudo inverse is a widely used alternative of the inverse of a matrix, due to its best fitness and uniqueness. We denote the pseudo inverse of a matrix A by $A^\dagger$. Then concerning about the initial ensemble covariance

$$\bar{P}(t_0|t_{-1}) = \frac{1}{q-1}\tilde{X}(t_0|t_{-1})\tilde{X}^\top(t_0|t_{-1}),$$

we apply the singular value decomposition to

$$\frac{1}{\sqrt{q-1}}\tilde{X}(t_0|t_{-1}) = V_0 S_0 U_0^\top,$$

where $V_0 \in \mathbb{R}^{n\times n}$ and $U_0 \in \mathbb{R}^{q\times q}$ consist of the left and right singular vectors respectively, and $S_0 \in \mathbb{R}^{n\times q}$ is a rectangular diagonal matrix with singular values $\{s_{0i}|s_{0i} \geqslant 0\}_{i=1}^q$ on its diagonal. Thus,

$$\bar{P}(t_0|t_{-1}) = V_0 S_0 U_0^\top U_0 S_0^\top V_0^\top = V_0 S_0 S_0^\top V_0^\top = V_0 \hat{S}_0^2 V_0^\top,$$

where $\hat{S}_0^2 = S_0 S_0^\top \in \mathbb{R}^{n\times n}$ is a block diagonal matrix with the diagonal

$$(s_{01}^2, \cdots, s_{0r_0}^2, 0_{1\times(n-r_0)}),$$

where r_0 is the rank of S_0. Hence,

$$\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) = V_0\hat{S}_0^{\dagger}V_0^{\top},$$

where $\hat{S}_0^{\dagger}$ is the pseudo inverse of $\hat{S}_0$ with the diagonal

$$(\frac{1}{s_{01}}, \cdots, \frac{1}{s_{0r_0}}, 0_{1\times(n-r_0)}).$$

Analogy to (4.9), we define the *ensemble relative improvement covariance* as

$$\tilde{P} = \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}).$$

Likewise, corresponding to (4.2), we present the observation system in the entire time interval as

$$y = \mathcal{G}x(t_0) + \nu,$$

where $y = (y^{\top}(t_0), \cdots, y^{\top}(t_N))^{\top}$, $\nu = (\nu^{\top}(t_0), \cdots, \nu^{\top}(t_N))^{\top}$ and $\mathcal{G}$ as the observation configuration for $x(t_0)$. As an analogy to the ensemble case of (2.7), we obtain

$$\begin{aligned}
& \bar{P}(t_0|t_N) \\
= \;& \bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_{-1})\mathcal{G}^{\top}(\mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^{\top} + \mathcal{R})^{-1}\mathcal{G}\bar{P}(t_0|t_{-1}) \\
= \;& \bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_{-1})\mathcal{G}^{\top}\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^{\top}\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}\mathcal{G}\bar{P}(t_0|t_{-1}) \\
= \;& \bar{P}(t_0|t_{-1}) - \bar{P}^f_{xy}\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\bar{P}^f_{yy}\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}(\bar{P}^f_{xy})^{\top}. \qquad (4.19)
\end{aligned}$$

Further,

$$\begin{aligned}
& \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\
=& \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}^f_{xy}\mathcal{R}^{-\frac{1}{2}}(I + \mathcal{R}^{-\frac{1}{2}}\bar{P}^f_{yy}\mathcal{R}^{-\frac{1}{2}})^{-1}\mathcal{R}^{-\frac{1}{2}}(\bar{P}^f_{xy})^{\top}\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}). \qquad (4.20)
\end{aligned}$$

Let $\sum_{i=0}^{N} m(t_i) = m$ be the number of observations. To proceed with (4.20), we apply again the singular value decomposition into,

$$\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}^f_{xy}\mathcal{R}^{-\frac{1}{2}} = VSU^{\top} \in \mathbb{R}^{n\times m}, \qquad (4.21)$$

where $U \in \mathbb{R}^{m\times m}$ consists of the eigenvectors of $\mathcal{R}^{-\frac{1}{2}}\mathcal{G}\bar{P}(t_0|t_{-1})\mathcal{G}^{\top}\mathcal{R}^{-\frac{1}{2}}$, $V \in \mathbb{R}^{n\times n}$ consists of the eigenvectors of $\bar{P}^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^{\top}\mathcal{R}^{-1}\mathcal{G}\bar{P}^{\frac{1}{2}}(t_0|t_{-1})$, $S \in \mathbb{R}^{n\times m}$ consists of the singular values on its diagonal.

We denote the rank of (4.21) by r. The *ensemble relative improvement covariance* can be rewritten as

$$\begin{aligned}
& \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\
= \;& VS^{\top}U^{\top}(UU^{\top} + U(SS^{\top})U^{\top})^{-1}USV^{\top} \\
= \;& VS^{\top}(I + S^{\top}S)^{-1}SV^{\top} \\
= \;& \sum_{i=1}^{r}\frac{s_i^2}{1+s_i^2}v_iv_i^{\top} \qquad (4.22)
\end{aligned}$$

and its diagonal elements show *ensemble relative improvements* of the corresponding states.

We observe that (4.22) and (4.12) have a similar form. By virtue of

$$\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}^f_{xy}\bar{\mathcal{R}}^{-\frac{1}{2}} = \bar{P}^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top\bar{\mathcal{R}}^{-\frac{1}{2}}, \tag{4.23}$$

the final results of (4.12) and (4.22) are equivalent. However, compared with

$$P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{-\frac{1}{2}},$$

the ensemble expression $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}^f_{xy}\bar{\mathcal{R}}^{-\frac{1}{2}}$ processes the absolute benefit, since in the calculation of $\bar{P}^f_{xy}$, we do not need the explicit form of $\mathcal{G}$. It allows us to code it line by line such that our approach is much more computationally efficient.

Since $\bar{P}(t_0|t_{-1})$ is not full rank for most cases,

$$\begin{aligned}
& \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})(\bar{P}(t_0|t_{-1}) - \bar{P}(t_0|t_N))\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\
= \;& \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_{-1})\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) - \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\
= \;& V_0\hat{S}_0^\dagger V_0^\top(V_0\hat{S}_0^2V_0^\top)V_0\hat{S}_0^\dagger V_0^\top - \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}) \\
= \;& V_0I_{r_0}V_0^\top - \bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1}),
\end{aligned}$$

where I_{r_0} is the diagonal matrix with the diagonal $(\mathbb{1}_{1\times r_0}, 0_{1\times(n-r_0)})$.

It is clear from (4.19) that $\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})\bar{P}(t_0|t_N)\bar{P}^{\dagger\frac{1}{2}}(t_0|t_{-1})$ is still nonnegative definite and $0_{n\times n} \preccurlyeq \tilde{P} \prec I_{r_0}$. Thus, the *ensemble relative improvement degree* is defined by

$$\tilde{p} = \frac{\|\tilde{P}\|_1}{\|I_{r_0}\|_1} = \frac{\|\tilde{P}\|_1}{r_0} \in [0, 1). \tag{4.24}$$

For the distinction of the improvements for concentrations and emission rates, the *ensemble relative ratios* remain

$$\tilde{p}^c = \frac{\|\tilde{P}^c\|_1}{\|\tilde{P}\|_1}, \quad \tilde{p}^e = \frac{\|\tilde{P}^e\|_1}{\|\tilde{P}\|_1}.$$

If we further consider the nonlinear dynamic model, we can renew the definition of the forecasting observation configurations as

$$y_k^f = \mathcal{G}(\hat{x}_k(t_0|t_{-1})), \quad k = 1, \cdots, q,$$

such that it can follow the nonlinear model, where $\mathcal{G}$ is a nonlinear operator.

Correspondingly, its ensemble mean and covariance are given by

$$\bar{y}^f = \frac{1}{q}\sum_{k=1}^q y_k^f, \quad \bar{P}^f_{yy} = \frac{1}{q-1}\sum_{k=1}^q(\hat{y}_k^f - \bar{y}^f)(\hat{y}_k^f - \bar{y}^f)^\top. \tag{4.25}$$

In addition,

$$\bar{P}^f_{xy} = \frac{1}{q-1}\sum_{k=1}^q(\hat{x}_k(t_0|t_{-1}) - \bar{x}(t_0|t_{-1}))(\hat{y}_k^f - \bar{y}^f)^\top. \tag{4.26}$$

It easily found that (4.25) and (4.26) can be substituted into (4.19) in order to determine the efficiency of observational networks of nonlinear models by solving the singular value decomposition of (4.21).

4.2.2 Example for the efficiency analysis

Consider a linear advection-diffusion model with Dirichlet horizontal boundary condition and Neumann boundary condition in the vertical direction on the domain $[0, 14]\times[0, 14]\times[0, 4]$,

$$\frac{\partial \delta c}{\partial t} = -v_x\frac{\partial \delta c}{\partial x} - v_y\frac{\partial \delta c}{\partial y} + \frac{\partial}{\partial z}(K(z)\frac{\partial \delta c}{\partial z}) + \delta e - \delta d, \tag{4.27}$$

where δc, δe and δd are the perturbations of the concentration, the emission rate and deposition rate of a species respectively. For vertical diffusion, $K(z)$ is a differentiable function of height z.

Assume $\triangle t = 0.5$, the numerical solution is based on the symmetric operator splitting technique [102] with the following operator sequence

$$\delta c(t + \triangle t) = T_x T_y D_z A D_z T_y T_x \delta c(t), \tag{4.28}$$

where T_x and T_y are transport operators in horizontal directions x and y, D_z is the diffusion operator in vertical direction z. The parameters of emission and deposition rates are included in A. The Lax-Wendroff algorithm is chosen as the discretization method for horizontal advection with $\triangle x = \triangle y = 1$. The vertical diffusion is discretized with $\triangle z = 1$ by Crank-Nicolson scheme with the Thomas algorithm [50] as the solver. The number of the grid points is $N_g = 1125$.

With the same temporal and spacial discretization of the concentration, the background knowledge of the emission rate is given by $e_b(t_n, i, j, l)$, where $n = 1, \cdots, N$ and $\{(i, j, l), i, j \in \{0, \cdots, 14\}, l \in \{0, \cdots, 4\}\}$. We rearrange $e_b(t_n, i, j, l)$ into one vector $e_b(t_n) = (e_b^1(t_n), \cdots, e_b^{N_g}(t_n))^\top$ and establish the discrete dynamic model of the emission rate according to (4.36)

$$\delta e(t_{n+1}) = M_e(t_{n+1}, t_n)\delta e(t_n), \quad n = 1, \cdots, N,$$

where $M_e(t_{n+1}, t_n)$ is a diagonal matrix of which i^{th} element on the diagonal is given by $\frac{e_b^i(t_{n+1})}{e_b^i(t_n)}$.

For expository reasons we assume δd be a constant over time and the only one fixed observation configuration is time-invariant. It indicates that the observation operator mapping the state space to the observation space is a $1 \times 2N_g$ time-invariant matrix.

In addition, we produce $q = 500$ (the ensemble numbers) samplings for the initial concentration and emission rate respectively by pseudo independent random numbers and make the states correlated by the moving average technique. In the following, we present three different tests, aiming to demonstrate the roles of variable winds, emissions, and vertical diffusion.

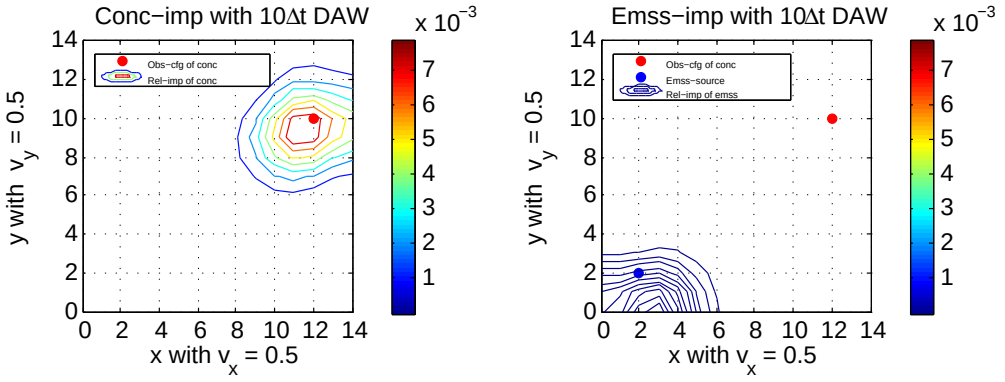

Figure 4.1: Advection test with $10\triangle t$ DAW and southwesterly wind. Isopleths of ensemble relative improvements of the concentration and emission rate are shown in the left and right figure panels respectively. The point located at $(12, 10, 0)$ named as'Obs-cfg of conc' shows the invariant observation configuration. The point located at $(2, 2, 0)$ named as 'Emss-source' is the source of the emission rate.

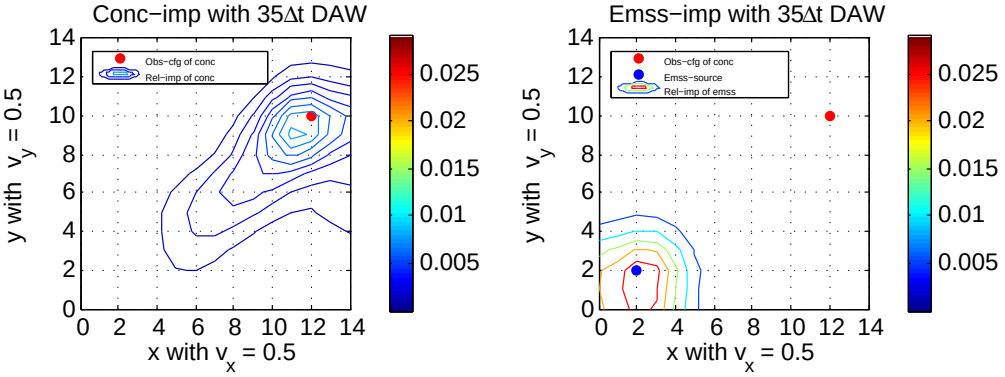

Figure 4.2: Advection test with $35\triangle t$ DAW and southwesterly wind. Plotting conventions are as in Fig. 4.1.

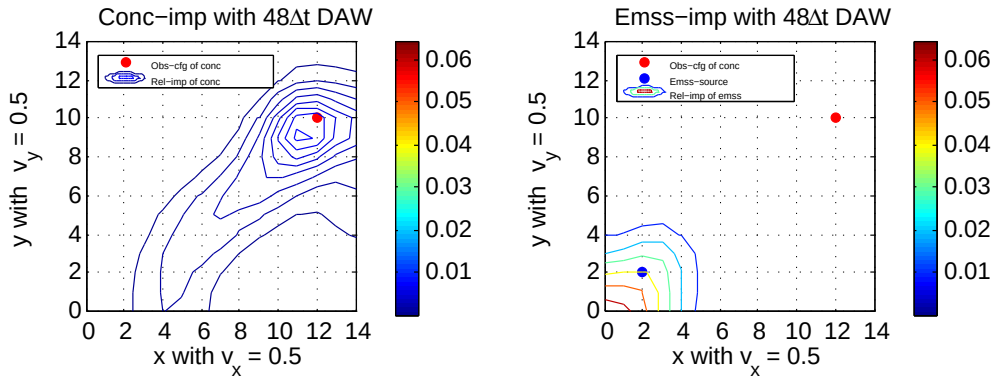

Figure 4.3: Advection test with $48\triangle t$ DAW and southwesterly wind. Plotting conventions are as in Fig. 4.1.

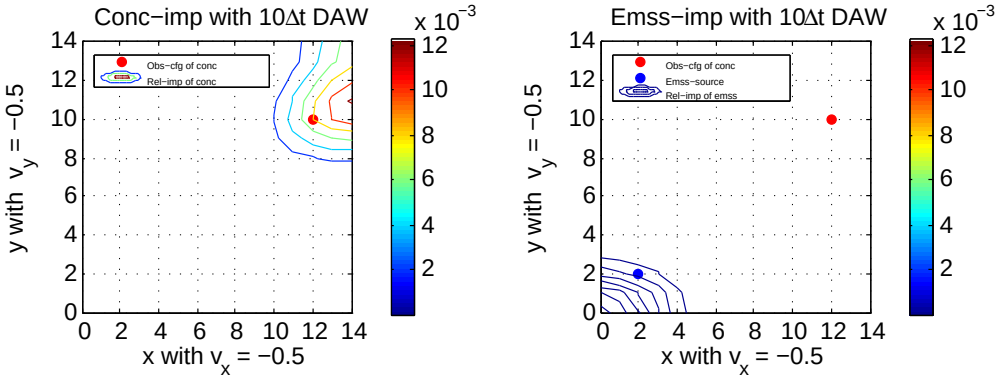

Figure 4.4: Advection test with $10\triangle t$ DAW and northeasterly wind. Plotting conventions are as in Fig. 4.1.

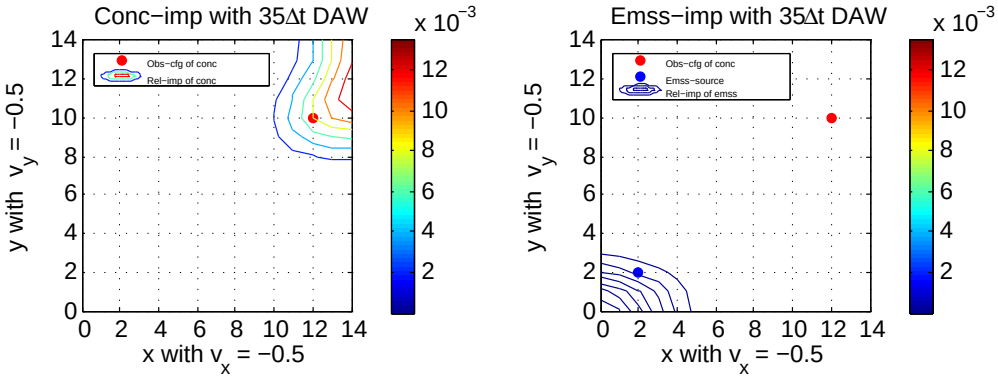

Figure 4.5: Advection test with $35\triangle t$ DAW and northeasterly wind. Plotting conventions are as in Fig. 4.1.

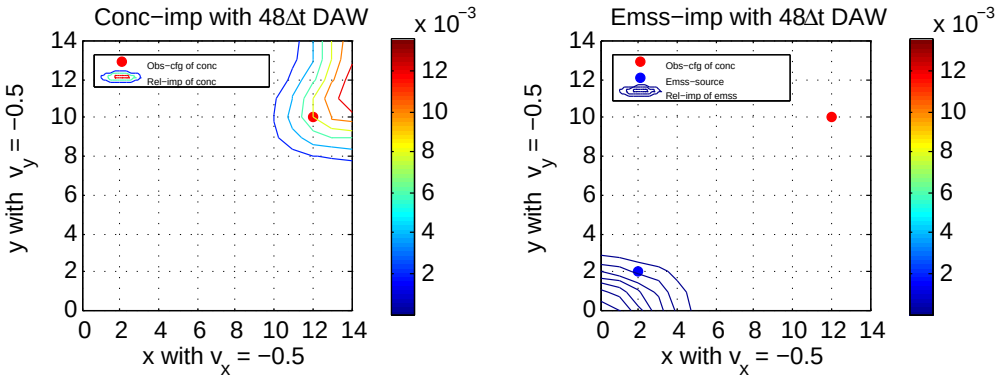

Figure 4.6: Advection test with $48\triangle t$ DAW and northeasterly wind. Plotting conventions are as in Fig. 4.1.

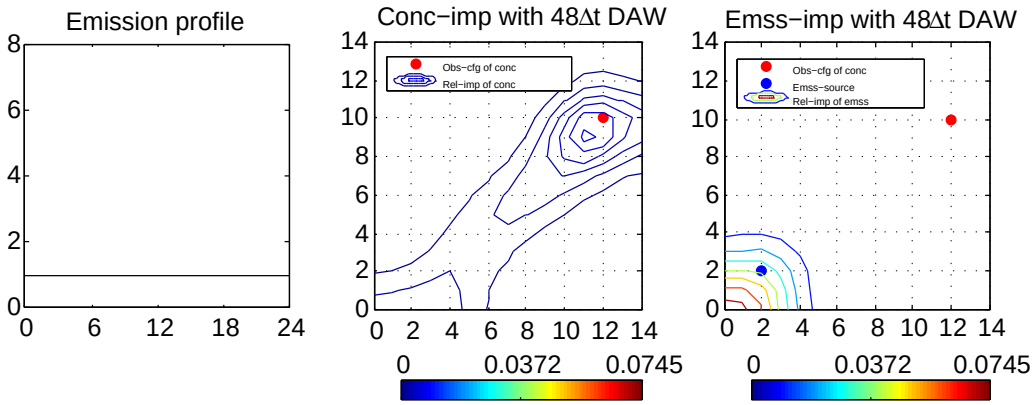

Figure 4.7: Emission signal test (weak) with $48\triangle t$ DAW and southwesterly wind ($v_x = 1$ and $v_y = 1$). Plotting conventions are as in Fig. 4.1.

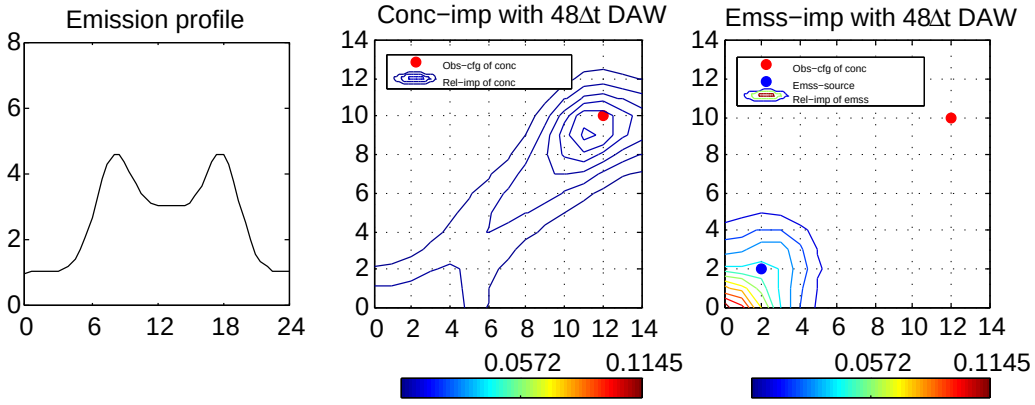

Figure 4.8: Emission signal test (strong) with $48\triangle t$ DAW and southwesterly wind ($v_x = 1$ and $v_y = 1$). Plotting conventions are as in Fig. 4.1.

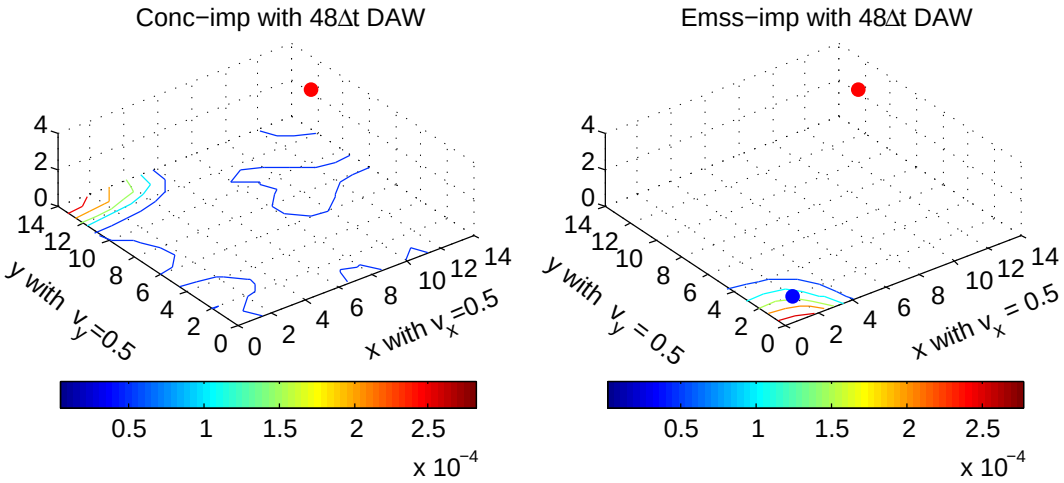

Figure 4.9: Diffusion test (weak) with $35\triangle t$ DAW and southwesterly wind. Plotting conventions are as in Fig. 4.1.

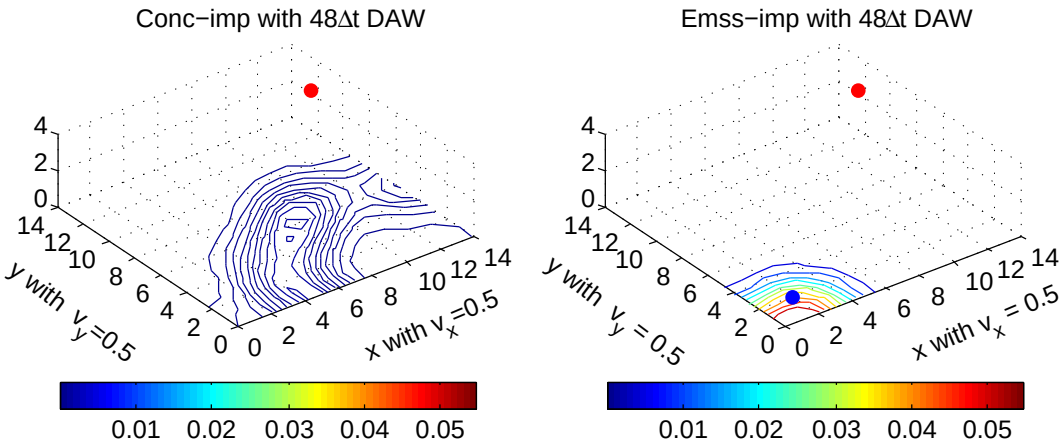

Figure 4.10: Diffusion test (strong) with $35\triangle t$ DAW and southwesterly wind. Plotting conventions are as in Fig. 4.1.

Fig.	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8	4.9	4.10
$\tilde{p}^c$	0.9920	0.4548	0.2848	0.9937	0.9939	0.9939	0.2248	0.1808	0.9928	0.1905
$\tilde{p}^e$	0.0080	0.5452	0.7152	0.0063	0.0061	0.0061	0.7752	0.8192	0.0072	0.8095

Table 4.1: Ensemble relative ratios of the initial value and emission rate at the lowest layer.

	$\tilde{P}^c_{low}$	$\tilde{P}^e_{low}$
Fig. 4.3	0.2767	0.8851
Fig. 4.9	0.0102	0.0030
Fig. 4.10	0.0500	0.7892

Table 4.2: The total improvement values of the initial value and emission rate at the lowest layer.

Advection test: The objective of the elementary advection test (Fig. 4.1 to Fig. 4.6) is to identify the most improvable parameters with different wind direction and data assimilation window (DAW) in advance of data assimilation procedure. Focusing on the advection effects, we assume the model with a weak diffusion process ($K(z) = 0.5e^{-z^2}$).

In Fig. 4.1 to Fig. 4.3 we assume southwesterly winds and the potential data assimilation windows are $10\triangle t$, $35\triangle t$ and $48\triangle t$ respectively. The ensemble relative improvements of the initial concentration are shown in the left panels of Fig. 4.1 to Fig. 4.3. We can find that the horizontal fields at lowest layer ($z = 0$) where the estimates of the concentration probably improved are enlarged with the extension of data assimilation windows since more and more grid points of the concentration are correlated with longer data assimilation windows.

The right panels of Fig. 4.1 to Fig. 4.3 show the ensemble relative improvements of the emission rate at each grid point with $z = 0$. From Fig. 4.1, we can observe that the ensemble relative improvements of the emission rate are smaller than the case of initial value in the influenced area. It indicates that the observations cannot detect the emission rate within $10\triangle t$ data assimilation window. Thus, in this case initial values alone can be optimized. It is shown in the right panels of Fig. 4.2 and Fig. 4.3 that the emission rate plays a more and more important role on the influence of observations. In this two cases, we consider both the concentration and emission rate as optimized parameters. At the same time, the quantitative balances between the concentration and emission rate are provided in Table 4.1.

Fig. 4.4 to Fig. 4.6 also show the ensemble relative improvements of the concentration and emission rate under the same assumptions as Fig. 4.1 to Fig. 4.3 respectively, except that the northeasterly wind is assumed. Clearly, with the northeasterly wind, whatever the duration of the assimilation window is, the emission is not detectable and improvable by that particular observation configuration. This hypothesis is demonstrated by our method. The quantitative balances are exposed in Table 4.1.

Emission signal test: The purpose of emission signal test (Fig. 4.7 and Fig. 4.8) is

to assess the efficiency of observation configurations to the emission rates evolved with different diurnal profiles. We have the same assumptions as Fig. 4.3 except the wind speed in Fig. 4.7 and Fig. 4.8 is accelerated such that the profiles of the emission rate is better detectable as to observations. The only distinction between the situations in Fig. 4.7 and Fig. 4.8 is the pronounced diurnal cycle background profile of the emission rate during the assimilation window $48\triangle t$, schematically simulating a rush hour induced source. Since the profiles of emission rates are correlated with the emitted amount of that species during the data assimilation window, it is clearly shown in Table 4.1 that the distinct variation of the emission rate during the data assimilation window acts to level $\tilde{p}^c$ and $\tilde{p}^e$, and thus helps to improve the estimates of source.

Diffusion test: The diffusion test (Fig. 4.9 and Fig. 4.10) aims to test the approach via comparing the ensemble relative improvements of the concentration and the emission rate at the layer $z = 0$, imposing both a weak diffusion process and a strong diffusion process. We assume the observation configuration at each time step is located at $(12, 10, 4)$ in the diffusion test. The diffusion coefficients are $K(z) = 0.5e^{-z^2}$ in Fig. 4.9 and $K(z) = 0.5e^{-z^2} + 1$ in Fig. 4.10. Besides, Fig. 4.9 and Fig. 4.10 preserve the same assumptions with Fig. 4.3.

It is obviously seen from Fig. 4.3 and Fig. 4.9 that the different observation locations strongly influence on the distribution of the relative improvements of the concentration. Table 1 shows that with the same diffusion coefficient the total improvement value of the concentration in the lowest layer in Fig. 4.3 is definitely larger than the one in Fig. 4.9. Moreover, it can be seen from Table 1 that the observation configuration at the top layer is not efficient to the emission rate with such weak diffusion within $48\triangle t$ data assimilation window .

Comparing Fig. 4.9 with Fig. 4.10, we can find that the ensemble relative improvements of concentration and emission rate increase with the stronger diffusion process. The increasing efficiency of the observation configuration with the stronger diffusion is also verified by the total improvement values of the concentration and emission rate for Fig. 4.9 and Fig. 4.10 in Table 4.2. The balances for Fig. 4.9 and Fig. 4.10 are shown in Table 4.1. The significant difference of the "weight" of emission rate in Table 4.1 implies that the observation configuration cannot detect the emission at the lowest layer with such a weak diffusion in Fig. 4.9. At the same time, with the stronger in Fig. 4.10 both the concentration and emission rate should be considered as optimized parameters with the corresponding "weights".

4.3 Sensitivity analysis of observational networks

The discussions about the observation efficiency above aim to evaluate and balance the probable estimating improvements of initial values and emission rates in advance of the execution of data assimilations. In this section, we will introduce the singular vector approach to identify the sensitive directions of observation networks to initial values and emission rates and show the association between the efficiency and sensitivity of observation networks.

4.3.1 Sensitivity analysis for discrete-time systems

Denoting

$$\delta x(t_0) = x(t_0) - \hat{x}(t_0) \in \mathbb{R}^{n\times 1},$$

where $\hat{x}(t_0)$ is any estimate of $x(t_0)$ and consider the discrete-time linear system in $[t_0, \cdots, t_N]$, we consider

$$\delta x(t_{k+1}) = M(t_{k+1}, t_k)\delta x(t_k)$$

with the observation configuration

$$\delta y_c(t_k) = H(t_k)\delta x(t_k), \quad \delta y_c(t_k) \in \mathbb{R}^{m(t)\times 1}, \tag{4.29}$$

where $M(t_{k+1}, t_k)$ is the transition matrix and $H(t_k)$ is a matrix mapping model states into the observation space. Then we define the magnitude of the perturbation of the initial state by the norm in the state space with respect to a positive definite matrix W_0

$$\|\delta x(t_0)\|^2_{W_0} = \langle \delta x(t_0), W_0 \delta x(t_0)\rangle.$$

Similarly, we define the magnitude of the related observations perturbation in the time interval $[t_0, \cdots, t_N]$ by the norm with respect to a sequence of positive definite matrices $\{W(t_k)\}_{k=1}^N$

$$\|\delta y_c\|^2_{\{W(t_k)\}} = \sum_{k=0}^{N} \langle \delta y_c(t_k), W(t_k)\delta y_c(t_k)\rangle,$$

where

$$\delta y_c = \begin{pmatrix} \delta y_c(t_0) \\ \delta y_c(t_1) \\ \vdots \\ \delta y_c(t_N) \end{pmatrix}.$$

In order to find the direction of observation configuration which can minimize the perturbation of the initial state, the ratio

$$\frac{\|\delta x(t_0)\|^2_{W_0}}{\|\delta y_c\|^2_{\{W(t_k)\}}}, \quad \delta y \neq 0_{m\times 1}.$$

should be minimized. It is equivalent to maximize the ratio between the magnitude of observation perturbation and the initial perturbation

$$\frac{\|\delta y_c\|^2_{\{W(t_k)\}}}{\|\delta x(t_0)\|^2_{W_0}}, \quad \delta x(t_0) \neq 0_{n\times 1}.$$

Thus, we define the measure the perturbation growth as

$$
\begin{aligned}
g^2 &= \frac{\|\delta y_c\|^2_{\{W(t_k)\}}}{\|\delta x(t_0)\|^2_{W_0}} \\
&= \sum_{k=0}^{N} \frac{\langle \delta y_c(t_k), W(t_k)\delta y_c(t_k)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle} \\
&= \sum_{k=0}^{N} \frac{\langle H(t_k)\delta x(t_k), W(t_k)H(t_k)\delta x(t_k)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle} \\
&= \sum_{k=0}^{N} \frac{\langle \delta x(t_k), H^\top(t_k)W(t_k)H(t_k)\delta x(t_k)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle} \\
&= \sum_{k=0}^{N} \frac{\langle \delta x(t_0), M^\top(t_k,t_0)H^\top(t_k)W(t_k)H(t_k)M(t_k,t_0)\delta x(t_0)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle} \\
&= \frac{\langle \delta x(t_0), \sum_{k=0}^{N} M^\top(t_k,t_0)H^\top(t_k)W(t_k)H(t_k)M(t_k,t_0)\delta x(t_0)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle} \\
&= \frac{\langle \delta x(t_0), \mathcal{G}^\top \mathcal{W}\mathcal{G}\delta x(t_0)\rangle}{\langle \delta x(t_0), W_0 \delta x(t_0)\rangle}, \quad \delta x(t_0) \neq 0,
\end{aligned} \tag{4.30}
$$

where $\mathcal{G}$ has the same definition with (4.7) and

$$
\mathcal{W} = \begin{pmatrix} W(t_0) & & \\ & \ddots & \\ & & W(t_N) \end{pmatrix}.
$$

According to Liao and Sandu [65], singular vectors refer to the directions of the error growth in a descend sequence with respect to the decreasing singular values. Hence, in order to search the maximal directions of g^2, we need to find out the solutions of the singular value problem:

$$
\begin{aligned}
W_0^{-\frac{1}{2}} \mathcal{G}^\top \mathcal{W}\mathcal{G} W_0^{-\frac{1}{2}} v_k &= s_k^2 v_k, \\
W^{\frac{1}{2}} \mathcal{G} W_0^{-1} \mathcal{G}^\top W^{\frac{1}{2}} u_k &= s_k^2 u_k,
\end{aligned}
$$

where $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_n \geqslant 0$ are singular values, $\{v_k\}_{i=1}^n$ and $\{u_k\}_{i=1}^n$ are the corresponding orthogonal singular vectors. Then,

$$
\max_{\delta x(t_0)\neq 0} g^2 = s_1^2.
$$

Especially, if the perturbation norms are provided by the choice $W_0 = P^{-1}(t_0|t_{-1})$ and $\mathcal{W} = \mathcal{R}^{-1}$,

$$
g^2 = \frac{\langle \delta x(t_0), \mathcal{G}^\top \mathcal{R}^{-1}\mathcal{G}\delta x(t_0)\rangle}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle}, \quad \delta x(t_0) \neq 0.
$$

We need to search the directions of

$$\begin{aligned} &P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1})v_k = s_k^2 v_k; \\ &\mathcal{R}^{-\frac{1}{2}}\mathcal{G}P(t_0|t_{-1})\mathcal{G}^\top\mathcal{R}^{-\frac{1}{2}}u_k = s_k^2 u_k, \quad k = 1, \cdots, n. \end{aligned} \tag{4.31}$$

The singular value s_k shows the amplification of the impact of the initial state to the observation configurations during the entire time interval. The associated singular vector in the state space v_k is the direction of k$^{\text{th}}$-fast growth of the perturbation of observations evolved from the initial perturbation. We compare the sensitivity analysis with the efficiency analysis in Section 4.1. Since

$$\frac{s_k^2}{1+s_k^2}$$

are decreasing with the decrease of s_k, $k = 1, \cdots, n$, it is clear that v_k is also the k$^{\text{th}}$ direction which maximizes the relative improvement of estimates based on the Kalman smoother. It indicates that most efficient directions of observation networks are the same with the most sensitive directions of the observation networks to the initial perturbations. Besides, the leading singular value s_1 is related to the operator norm of $\tilde{P}$

$$\|\tilde{P}\| = \max_{\|x\|=1} \|\tilde{P}x\| = \frac{s_1^2}{1+s_1^2},$$

which implies the upper boundedness of the relative improvement covariance.

From the analysis above we can find that the sensitivity analysis does not provide us the information of covariances of estimates directly. However, it gives us an access to approximate and target the sensitive parameters or areas with the certain metric of the leading singular vectors weighted by the corresponding singular values.

4.3.2 Sensitivity analysis of the atmospheric transport model extended by emission rates

Due to the homogeneity of the atmospheric transport model extended by emissions, the sensitivity analysis can be easily applied into the extended atmospheric transport model by dividing singular vectors into the block form according to the dimensions of the initial state and emissions. The corresponding block parts of different singular vectors indicate the different sensitive directions of the initial state and emissions.

We consider the same example in Section 4.2.2. Fig. 4.11 exhibits in its upper row panels the singular values of Fig. 4.1 to Fig. 4.3. We approximate the sensitivities of the initial concentrations by the first five leading singular vectors weighted by the associated singular values in the nuclear norm and show the results in the three panels in the second row. It is clearly visible that the sensitive area can be well targeted by only few singular vectors, although the sensitivity analysis cannot provide the quantitative solutions with a clear statistical significance as the efficiency analysis of observation networks.

Finally, similar to Fig. 4.11, the singular values of Fig. 4.9 and Fig. 4.10 and the approximating targeting results of sensitive parameters are shown in Fig. 4.12.

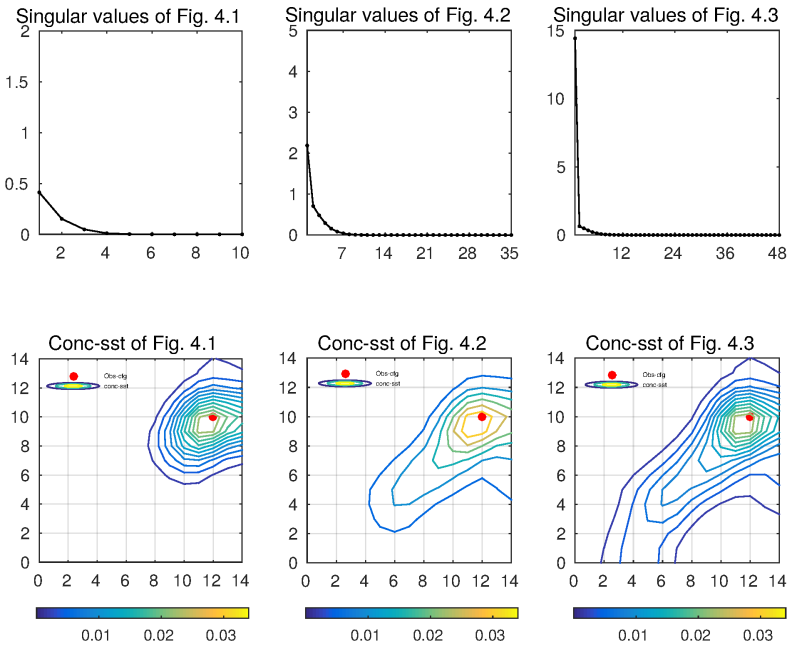

Figure 4.11: Singular values of Fig. 4.1 to Fig. 4.3 and sensitivities of initial states approximated by 5 leading singular values.

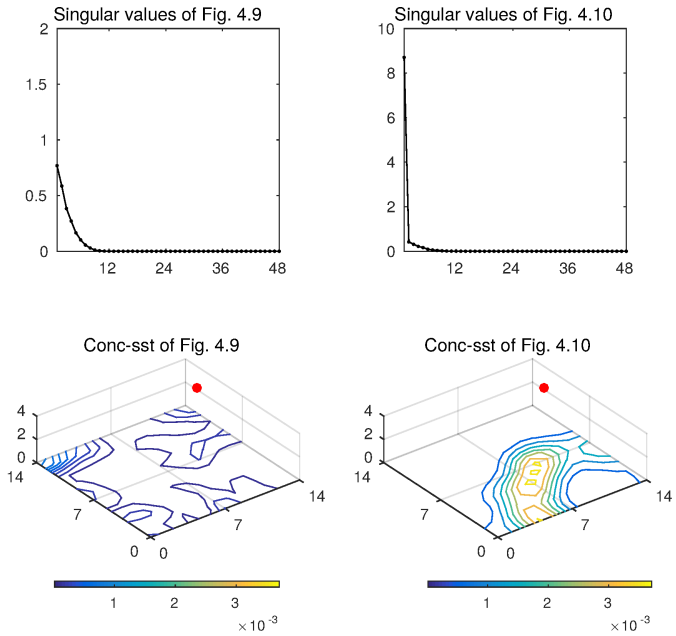

Figure 4.12: Singular values of Fig. 4.9 and Fig. 4.10 and sensitivities of initial states approximated by 5 leading singular values.

4.3.3 Sensitivity analysis for continuous-time systems

We generalize the sensitivity analysis in Section 4.3.1 into the following continuous-time system

$$\delta x(t) = M(t, t_0)\delta x(t_0),$$

with the corresponding forecast perturbation of observations evolving from $\delta x(t_0)$

$$\delta y(t) = H(t)\delta x(t).$$

To be brief, we assume $W_0 = P^{-1}(t_0|t_{-1})$ and $\mathcal{W}(t) = R^{-1}(t)$ and define the magnitude of the perturbation of the initial state and observations respectively by

$$\begin{aligned}
\|\delta x(t_0)\|^2_{P^{-1}(t_0|t_{-1})} &= \langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle, \\
\|\delta y\|^2_{\{R^{-1}(t)\}} &= \int_{t_0}^{t_N} \langle \delta y(t), R^{-1}(t)\delta y(t)\rangle dt.
\end{aligned}$$

Thus, the perturbation growth for continuous-time system can be measured by

$$\begin{aligned}
g^2 &= \frac{\|\delta y\|^2_{\{R^{-1}(t))\}}}{\|\delta x(t_0)\|^2_{P^{-1}(t_0|t_{-1})}} \\
&= \frac{\int_{t_0}^{t_N} \langle H(t)\delta x(t), R^{-1}(t)H(t)\delta x(t)\rangle dt}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle} \\
&= \frac{\int_{t_0}^{t_N} \langle H(t)M(t,t_0)\delta x(t_0), R^{-1}(t)H(t)M(t,t_0)\delta x(t_0)\rangle dt}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle} \\
&= \frac{\langle \delta x(t_0), \int_{t_0}^{t_N} M^\top(t,t_0)H^\top(t)R^{-1}(t)H(t)M(t,t_0)\delta x(t_0)dt\rangle}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle} \\
&= \frac{\langle \delta x(t_0), \mathcal{G}^*\mathcal{R}^{-1}\mathcal{G}\delta x(t_0)\rangle}{\langle \delta x(t_0), P^{-1}(t_0|t_{-1})\delta x(t_0)\rangle}, \quad \delta x(t_0) \neq 0,
\end{aligned} \tag{4.32}$$

where $\mathcal{G}$ and $\mathcal{R}^{-1}$ are defined in (4.7).

To find the directions maximizing the ratio, we need to find the solutions of the singular value problem:

$$\begin{aligned}
P^{\frac{1}{2}}(t_0|t_{-1})\mathcal{G}^*\mathcal{R}^{-1}\mathcal{G}P^{\frac{1}{2}}(t_0|t_{-1})v_k &= s_k^2 v_k, \\
\mathcal{R}^{-\frac{1}{2}}\mathcal{G}P(t_0|t_{-1})\mathcal{G}^*\mathcal{R}^{-\frac{1}{2}}u_k &= s_k^2 u_k,
\end{aligned} \tag{4.33}$$

where $s_1 \geqslant s_2 \geqslant \cdots \geqslant s_n \geqslant 0$ are singular values, $\{v_k\}_{i=1}^n$ and $\{u_k\}_{i=1}^n$ are orthogonal singular vectors.

Compared (4.33) with (4.17), similar analysis and conclusions as Section 4.3 can be extended to continuous-time systems.

Until now, approaches to determining the efficiency and sensitivity of observation configurations for discrete-time and continuous-time systems have been established. It

can be found that some special operators are usually applied in order to deal with the specific questions in atmospheric chemistry. For example, in order to consider the efficiency and sensitivity of observations in some certain locations, the local projection operator introduced by Buizza et al. [13] can be applied into approaches in Section 4.2 and Section 4.3.

Let L be a $0-1$ diagonal matrix defined as

$$L_{ii} = \begin{cases} 1, & l_i \in L_a, \\ 0, & \text{otherwise.} \end{cases}$$

where L_a is a fixed area and l_i is the coordinate of i^{th} grid point.

To test the efficiency and sensitivity of observation configurations in a special area, by rearranging the observations y according to the locations, $\mathcal{G}$ in (4.7) should be defined as

$$\mathcal{G} = \begin{pmatrix} LH(t_0)M(t_0,t_0) \\ LH(t_1)M(t_1,t_0) \\ \vdots \\ LH(t_N)M(t_N,t_0) \end{pmatrix}.$$

If $LH(\cdot)$ is considered as the observation mapping, it is straightforward to apply the approaches stated above to analyse the efficiency and sensitivity of specialized observational networks.

4.4 Emission source apportionments to observation networks by singular vector analysis

In this section we apply the sensitivity analysis in Section 4.3 to solve the emission rate apportionment problem. As mentioned before, emission rates are almost unobservable in practice and we can extract the regular diurnal profiles of different-type emissions for atmospheric models but have poor knowledge about their amplitude. Take CO_2 for example, the two peaked rush hour emission profile over the day is applicable for traffic emissions. Base load operated power plants emit continuously. Biogenic sources and sinks also have different profiles: CO_2 from photosynthesis during daylight the sinking peaks at noon. CO_2 source from plant and soil respiration fairly continuous over 24 hours.

Hence, it is highly desired in practice to investigate, to what extend these different diurnal source and sink shapes of one certain emitted species can be taken for source apportionments. By author's knowledge, the most existing papers concerning the source apportionment problem, such as [25] and [64], usually identify the emission source by the application of principal component analysis and multiple linear regression based on the observation data. Independent with the observation data, we study the emission source apportionment problem with the aid of the dynamic model extended by emission rates, which has been introduced in Section 3.2.1. It benefits us that the concentration and emission rates can be jointly considered with the time evolution in the entire time interval. By means of singular vector analysis, we can determine the emission source apportionments

of one species under different observation configurations. If we take the data into account, the above emission source apportionments can be considered as the result to the expectation of observations, in the statistical sense.

We establish the atmospheric transport models extended by emissions from different sources in Section 4.4.1 and review the theoretical foundation based on singular value decomposition for both the original model and its ensemble case in Section 4.4.2. By investigating the covariance between the normalized initial perturbation and coefficients of the observation perturbation under the basis of observation space, we develop a sequence of indexes in order to quantitatively study the contributions or namely sensitivities of concentrations and emissions. In Section 4.4.3, we give an elementary example to verify the analysis above and study how the different profiles, locations of emissions and boundary conditions effect the apportionments of the concentration and emissions. Finally, in Section 4.4.4 we discuss the relationship of contributions of different states between observation configurations to highly dimensional observation space and observation configurations to one dimensional observation space.

4.4.1 Model description

We describe the chemical tendency equation including various emission rates, propagating forward in time, by the following atmospheric transport model

$$\frac{dc}{dt} = \mathcal{A}(c) + \mathbf{B}e(t),$$

where $\mathcal{A}$ is a nonlinear model operator, $\mathbf{B}$ is a linear operator, $c(t)$ is the state vector of chemical constituents and $e(t)$ is the emission rate for one species from different sources at time t.

A prior estimate of the state vector of concentrations $c(t)$ and the emission rate $e(t)$ are still denoted by $c_b(t)$ and $e_b(t)$, respectively.

Let $\mathbf{A}$ be the tangent linear operator of $\mathcal{A}$, the evolution of the perturbation of states $c(t)$ and $e_i(t)$ follows the tangent linear model with $\mathbf{A}$ as

$$\frac{d\delta c}{dt} = \mathbf{A}\delta c + \mathbf{B}\delta e(t), \tag{4.34}$$

where $\delta c(t)$ is the perturbation evolving from the perturbation of initial state of chemical state $\delta c(t_0) = c(t_0) - c_b(t_0)$ and emission rates $\delta e(t) = e(t) - e_b(t)$.

After discretizing the tangent linear model in space, let $M_c(\cdot,\cdot)$ be the evolution operator or resolvent generated by A_n, the approximation of $\mathbf{A}$ in $n-$dimensional space. It is straightforward to obtain the linear solution of (4.34) with continuous time as

$$\delta c(t) = M_c(t,t_0)\delta c(t_0) + \int_{t_0}^{t} M_c(t,s)B\delta e(s)ds, \tag{4.35}$$

where $\delta c(t) \in \mathbb{R}^n$. $M_c(\cdot,\cdot) \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n_e}$ is the approximation of $\mathbf{B}$ on $\mathbb{R}^n$.

Different types of emission sources, for example, the emissions from traffic and photosynthesis, have different diurnal profiles. We categorize the emission rates according

the distinction of profiles of emissions and assume that there are k_e different kinds of emissions. here $\delta e(t) = (\delta e_1^\top(t), \cdots, \delta e_k^\top(t))^\top \in \mathbb{R}^{n_e}$, where $n_e = k_e \times n$ is the dimension of the partial phase space of k_e different kinds of emission rates from different sources for one species.

In addition, let $y(t)$ be the observation vector of $c(t)$ and define

$$\delta y(t) = y(t) - \mathcal{H}(t)c_b(t),$$

where $\mathcal{H}(t)$ is a nonlinear forward observation operator mapping the model space to the observation space. Then by linearizing the nonlinear operator $\mathcal{H}$ as H and define $\delta y_c(t) = H(t)\delta c(t)$, the linearized model equivalents of observation configurations can thus be presented as

$$\delta y(t) = \delta y_c(t) + \nu(t),$$

where $\delta y(t) \in \mathbb{R}^{m(t)}$, $m(t)$ the dimension of the phase space of observation configurations at time t. $\nu(t)$ is the observation error at time t following the Gaussian distribution with zero mean and covariance $R(t) \in \mathbb{R}^{m(t)\times m(t)}$.

Under the constraint of diurnal profiles of emission rates and based on the background knowledge of emission rates and only optimizing the amplitude of emission rates, with the same idea in (3.8), we establish the dynamic model of each type of emission rates as

$$\delta e_i(t) = M_{ei}(t,s)\delta e_i(s), \quad i = 1, \cdots, k, \tag{4.36}$$

where $M_{ei}(t,s)$ is the diagonal matrix defined as

$$M_{ei}(t,s) = \begin{pmatrix} \frac{e_{bi}^1(t)}{e_{bi}^1(s)} & & & \\ & \frac{e_{bi}^2(t)}{e_{bi}^2(s)} & & \\ & & \ddots & \\ & & & \frac{e_{bi}^n(t)}{e_{bi}^n(s)} \end{pmatrix}$$

and $e_{bi}(\cdot) \in \mathbb{R}^n$ is the background vector of i^{th} type of emissions, of which the j^{th} element is denoted by $e_{bi}^j(\cdot)$.

If we rewrite B as the block form $B = (B_1, \cdots, B_{k_e})$, where $B_i \in \mathbb{R}^{n\times n}$, $i = 1, \cdots, k_e$, (4.35) can be written as

$$\delta c(t) = M_c(t,t_0)\delta c(t_0) + \sum_{i=1}^{k_e} \int_{t_0}^{t} M_c(t,s)B_i M_{ei}(s,t_0)\delta e_i(t_0)ds. \tag{4.37}$$

Hence, we obtain the model extended by emission rates

$$\begin{pmatrix} \delta c(t) \\ \delta e_1(t) \\ \vdots \\ \delta e_{k_e}(t) \end{pmatrix} = M(t,t_0) \begin{pmatrix} \delta c(t_0) \\ \delta e_1(t_0) \\ \vdots \\ \delta e_{k_e}(t_0) \end{pmatrix}, \tag{4.38}$$

where

$$M(t,t_0) := \begin{pmatrix} M_c(t,t_0) & \int_{t_0}^{t} M_c(t,s)B_1M_{e1}(s,t_0)ds & \cdots & \int_{t_0}^{t} M_c(t,s)B_kM_{ek}(s,t_0)ds \\ & M_{e1}(t,t_0) & & \mathbf{0} \\ & & \ddots & \\ & \mathbf{0} & & M_{ek_e}(t,t_0) \end{pmatrix}.$$

Typically, there is no direct observation for emissions. Therefore, we reconstruct the observation mapping as

$$\delta y_c(t) = (H(t), 0_{n\times n_e}) \begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix}.$$

4.4.2 Singular vector analysis for emission source apportionments

If we denote

$$\delta x(t) = (\delta c^\top(t), \delta e_1^\top(t), \cdots, \delta e_{k_e}^\top(t))^\top,$$

by observing (4.38), we only need to consider the generalized discrete-time linear system:

$$\delta x(t_{k+1}) = M(t_{k+1}, t_k)\delta x(t_k) \tag{4.39}$$

and the covariance of $\delta x(t_0)$ is denoted by $P(t_0)$.

Correspondingly, we have the observation mapping of (4.39) with the normalized observation error,

$$R^{-\frac{1}{2}}(t_k)\delta y(t_k) = R^{-\frac{1}{2}}(t_k)\delta y_c(t_k) + R^{-\frac{1}{2}}(t_k)\nu(t_k), \quad \nu(t_k) \sim \mathcal{N}(0, R(t_k)). \tag{4.40}$$

If we attempt to determine the emission source apportionments based on the model (4.38), it is adequate to consider the influence between the initial perturbation including the emission rates and the normalized observation perturbation $\{R^{-\frac{1}{2}}(t_k)\delta y_c(t_k)\}_{k=1}^N$ instead of $\delta y(t_k)$, since the characteristic of the observation error, or equivalently, observations are statistically captured by $\{R(t_k)\}_{k=1}^N$. It indicates that the problem is already independent with the real observation data now. Therefore, we define

$$\delta y_c = \begin{pmatrix} \delta y_c(t_0) \\ \delta y_c(t_1) \\ \vdots \\ \delta y_c(t_k) \end{pmatrix}, \quad \mathcal{R} = \begin{pmatrix} R(t_0) & & & \\ & R(t_1) & & \\ & & \ddots & \\ & & & R(t_N) \end{pmatrix}.$$

To find a sequence of orthogonal directions to investigate how the initial perturbation influence on the observations, we consider the ratio to measure the perturbation growth

$$g^2 = \frac{\|\mathcal{R}^{-\frac{1}{2}}\delta y_c\|^2}{\|P^{-\frac{1}{2}}(t_0)\delta x(t_0)\|^2}. \tag{4.41}$$

According to Section 4.3, we have

$$g^2 = \frac{\langle \delta x(t_0), \sum_{k=0}^{N} M^\top(t_k, t_0) H^\top(t_k) R^{-1}(t_k) H(t_k) M(t_k, t_0) \delta x(t_0) \rangle}{\langle \delta x(t_0), P^{-1}(t_0) \delta x(t_0) \rangle} \tag{4.42}$$

and the singular vectors referring to the following singular value decomposition shows that the directions of the error growth in a descend sequence with respect to the descent singular values. Hence, in order to search the maximal directions of

$$\begin{aligned} & P^{\frac{1}{2}}(t_0) \mathcal{G}^\top \mathcal{R}^{-1} \mathcal{G} P^{\frac{1}{2}}(t_0) v_k = s_k^2 v_k; \\ & \mathcal{R}^{-\frac{1}{2}} \mathcal{G} P(t_0) \mathcal{G}^\top \mathcal{R}^{-\frac{1}{2}} u_k = s_k^2 u_k, \quad k = 1, \cdots, r, \end{aligned} \tag{4.43}$$

where v_k and u_k are singular vectors in the state space and the observation space related to the singular value s_k, r is number of the positive singular values and $\mathcal{G}$ is defined as (4.7).

The singular values s_i associated with the evolution of the corresponding singular vectors in state space v_i to the observation space u_i, are the amplification of the impact of the initial state to the observation configurations during the entire time interval. However, in the real world, the initial perturbation is unlikely to be one of the singular vectors exactly but usually a linear combination of the basis consisted by the singular vectors. We know that it is not theoretically sufficient to determine the portion of each emission source in the potential observations if we only consider few dominant singular values and vectors, although the final result might be numerically approximated by few dominant singular values and vectors. Besides, independent of the real observation data, it becomes essential to interpret this problem with the clear statistical significance. Thus, we now start with an arbitrary initial perturbation $\delta x(t_0)$. On one hand, we have shown in (4.42) that the corresponding observation perturbation in the entire time interval $[t_0, \cdots, t_N]$ given by

$$\mathcal{R}^{-\frac{1}{2}} \delta y = \mathcal{R}^{-\frac{1}{2}} \mathcal{G} \delta x(t_0),$$

where $\delta y = (\delta y^\top(t_0), \cdots, \delta y^\top(t_N))^\top$. On the other hand,

$$\mathcal{R}^{-\frac{1}{2}} \delta y_c = UC,$$

where $C = U^\top \mathcal{R}^{-\frac{1}{2}} \delta y_c$ is the coordinate vector of $\mathcal{R}^{-\frac{1}{2}} \delta y_c$ with the basis in U. Thus, by the singular value decomposition in (4.43), we have

$$UC = USV^\top P^{-\frac{1}{2}}(t_0) \delta x(t_0),$$

and

$$C = SV^\top P^{-\frac{1}{2}}(t_0) \delta x(t_0). \tag{4.44}$$

It is obvious from (4.44) that C is able to completely reflect the initial perturbation $\delta x(t_0)$. Therefore, we consider the covariance between C and $\delta x(t_0)$

$$\begin{aligned} \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0) \delta x(t_0)) &= \mathrm{cov}(SV^\top P^{-\frac{1}{2}}(t_0) \delta x(t_0), P^{-\frac{1}{2}}(t_0) \delta x(t_0)) \\ &= SV^\top P^{-\frac{1}{2}}(t_0) P(t_0) P^{-\frac{1}{2}}(t_0) \\ &= SV^\top. \end{aligned} \tag{4.45}$$

If we denote

$$P = \text{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0))^\top \text{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) = VS^\top SV^\top,$$

in order to investigate to what extend, each entry in $\delta x(t_0)$ contributes to, or namely is sensitive to the observation perturbations in the whole time interval, we consider the square root matrix of P, which is given by

$$P^{\frac{1}{2}} = VSV^\top.$$

Then we denote the sensitivity of each entry by $\text{SST}(\delta x_i(t_0))$ and define as the i^{th} element of the diagonal of P, given by

$$\text{SST}(\delta x_i(t_0)) = \sum_{j=1}^{r} s_j v_{ji}^2.$$

where v_{ji} is the i^{th} element of the singular vector v_j related to s_j.

Furthermore, in order to determine the total sensitivity or contribution of the concentration and each category of emission rates in $\delta x(t_0)$, we divide V and $P^{\frac{1}{2}}$ into the block form, according to the dimensions of $\delta c^\top(t_0)$ and $\delta e_i^\top(t_0), i = 1, \cdots, k_e$, as

$$V = \begin{pmatrix} V^c \\ V^{e_1} \\ \vdots \\ V^{e_{k_e}} \end{pmatrix}, \quad P^{\frac{1}{2}} = \begin{pmatrix} P_c^{\frac{1}{2}} & P_{ce_1}^{\frac{1}{2}} & \cdots & P_{ce_{k_e}}^{\frac{1}{2}} \\ P_{e_1c}^{\frac{1}{2}} & P_{e_1}^{\frac{1}{2}} & \cdots & P_{e_1e_{k_e}}^{\frac{1}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ P_{e_{k_e}c}^{\frac{1}{2}} & \cdots & P_{e_{k_e-1}e_{k_e}}^{\frac{1}{2}} & P_{e_{k_e}}^{\frac{1}{2}} \end{pmatrix}.$$

Denoting the normalized block initial vector by

$$\delta\tilde{x}(t) := P^{-\frac{1}{2}}(t_0)\delta x(t) = (\delta\tilde{c}^\top(t), \delta\tilde{e}_1^\top(t), \cdots, \delta\tilde{e}_{k_e}^\top(t))^\top,$$

we define the *total sensitivity* (TSST) or *contribution* by

$$\text{TSST}(\delta c(t_0)) = \|P_c^{\frac{1}{2}}\|_1 = \|V^c S (V^c)^\top\|_1 = \sum_{j=1}^{r} s_j \|v_j^c\|^2,$$

$$\text{TSST}(\delta e_i(t_0)) = \|P_{e_i}^{\frac{1}{2}}\|_1 = \|V^{e_i} S (V^{e_i})^\top\|_1 = \sum_{j=1}^{r} s_j \|v_j^{e_i}\|^2, \quad i = 1, \cdots, k_e.$$

where the j^{th} singular vector v_j with the block form $v_j = ((v_j^c)^\top, (v_j^{e_1})^\top, \cdots, (v_j^{e_{k_e}})^\top)^\top$ and here $\|\cdot\|$ is the Euclidean norm of vector.

Further, since V is an unitary matrix, we can normalize the total sensitivity as *the total sensitivity degree* (TSST%) by

$$\text{TSST}\%(\delta c(t_0)) = \frac{\|P_c^{\frac{1}{2}}\|_1}{\|P^{\frac{1}{2}}\|_1} = \frac{\sum_{j=1}^{r} s_j \|v_j^c\|^2}{\sum_{j=1}^{r} s_j},$$

$$\text{TSST}\%(\delta e_i(t_0)) = \frac{\|P_{e_i}^{\frac{1}{2}}\|_1}{\|P^{\frac{1}{2}}\|_1} = \frac{\sum_{j=1}^{r} s_j \|v_j^{e_i}\|^2}{\sum_{j=1}^{r} s_j}, \quad i = 1, \cdots, k_e.$$

Obviously, it is more intuitive and straightforward to consider

$$\mathrm{cov}(\mathcal{R}^{-\frac{1}{2}}\delta y_c, P^{-\frac{1}{2}}(t_0)\delta x(t_0))$$

rather than $\mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0))$ in (4.45). In fact,

$$\begin{aligned} P &= \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0))^\top \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) \\ &= VS^2V^\top \\ &= VSU^\top USV^\top \\ &= \mathrm{cov}(\mathcal{R}^{-\frac{1}{2}}\delta y_c, P^{-\frac{1}{2}}(t_0)\delta x(t_0))^\top \mathrm{cov}(\mathcal{R}^{-\frac{1}{2}}\delta y_c, P^{-\frac{1}{2}}(t_0)\delta x(t_0)). \end{aligned}$$

Due to the uniqueness of singular value decomposition and C is uniquely determined by U, it is equivalent to define SST, TSST and TSST% based on

$$\mathrm{cov}(\mathcal{R}^{-\frac{1}{2}}\delta y_c, P^{-\frac{1}{2}}(t_0)\delta x(t_0))$$

The various indices can meet our objective from different aspects. On one hand, compared with SSTs or TSSTs, which are absolute values of contributions, the relative index TSST%s can be more directly applied to compare the concentration and emission source apportionments. On the other hand, the relativity of TSST%s causes the loss of information about the amplitude of the impact of concentration and emission rates to the potential observations, which is indicated by singular values and implicitly included in SSTs and TSSTs.

Let us take the real data into account now. Because of the noises of observations $\{\nu(t_k)\}_{k=1}^N$, it could occur that some components of observation configurations δy_c may trap in the observation noises and be not effective to observations. Thus, our present obstacle is how to judge the amplitudes of singular values and further decide which independent directions included in SSTs are reliable and beyond observation noises.

As mentioned before, the observation error in (4.40) has been normalized to take the identity matrix as the covariance. If we define

$$\delta\tilde{y} = ((R^{-\frac{1}{2}}(t_0)\delta y(t_0))^\top, \cdots, (R^{-\frac{1}{2}}(t_N)\delta y(t_N))^\top)^\top,$$

then we have

$$\mathrm{cov}(\delta\tilde{y}, \delta\tilde{y}) = \mathcal{R}^{-\frac{1}{2}}\mathcal{G}P(t_0)\mathcal{G}\mathcal{R}^{-\frac{1}{2}} + I.$$

According to the similar way to determine the degree of freedom of observation in [85, Section 2.4.1], the effective independent components of the initial perturbation to the observations are the singular vectors of which the relative singular values are larger than the unity. It not only provides a criteria to judge the amplitude of singular values and reliability of each direction of the initial perturbation, but also gives an access to reasonably approximate SSTs by the singular values larger than one and their corresponding singular vectors.

In order to reduce the computation cost of solving the singular vector problem (4.31), we project the original state space to a lower-rank sampling space. Thus, we assume that there are q samplings of the initial perturbation, denoted by

$$X(t_0) = (\delta x_1(t_0), \delta x_2(t_0), \cdots, \delta x_q(t_0)),$$

then its sampling covariance is given by

$$\bar{P}(t_0) = \frac{1}{q-1}\tilde{X}(t_0)\tilde{X}^\top(t_0),$$

where

$$\tilde{X}(t_0) = X(t_0) - \frac{1}{q}X(t_0)\mathbb{1}_{q\times q}.$$

As shown in (4.46)

$$\bar{P}^{\frac{1}{2}}(t_0)\mathcal{G}^\top\mathcal{R}^{-\frac{1}{2}} = \bar{P}^{\dagger\frac{1}{2}}(t_0)\bar{P}^f_{xy}\mathcal{R}^{-\frac{1}{2}}, \tag{4.46}$$

where $\bar{P}^{\dagger\frac{1}{2}}(t_0)$ is the pseudo inverse of $\bar{P}^{\frac{1}{2}}(t_0)$, $\bar{P}^f_{xy} = \bar{P}(t_0)\mathcal{G}^\top$ is the sampling covariance between initial perturbation $\delta x(t_0)$ and the corresponding forecasting observation perturbations δy in the entire time interval. Hence, we only need to find out the singular values and singular vectors of $\bar{P}^{\dagger\frac{1}{2}}(t_0)\bar{P}^f_{xy}\mathcal{R}^{-\frac{1}{2}}$, which can be easily computed since the explicit pattern of $\mathcal{G}^\top$ can be avoided and its rank must be less than the ensemble number.

4.4.3 Example

Consider a linear advection-diffusion model with periodic horizontal boundary condition and Neumann boundary condition in the vertical direction on the domain $[0, 14]\times[0, 14]\times[0, 4]$,

$$\frac{\partial\delta c}{\partial t} = -v_x\frac{\partial\delta c}{\partial x} - v_y\frac{\partial\delta c}{\partial y} + \frac{\partial}{\partial z}(K(z)\frac{\partial\delta c}{\partial z}) + \sum_{i=1}^{3}\delta e_i - \delta d, \tag{4.47}$$

where δc, δe_i and δd are the perturbations of the concentration, emission rate and deposition rate of a species respectively. v_x and v_y are constants and $K(z)$ is a differentiable function of height z.

Assume $\triangle t = 0.5$, the numerical solution is based on the symmetric operator splitting technique [102] with the following operator sequence

$$\delta c(t + \triangle t) = T_xT_yD_zAD_zT_yT_x\delta c(t),$$

where T_x and T_y are transport operators in horizontal directions (x, y), D_z is the diffusion operator in vertical direction z. The parameters of emission and deposition rates are included in A. The Lax-Wendroff algorithm is chosen as the discretization method for horizontal advection with $\triangle x = \triangle y = 1$. The vertical diffusion is discretized by Crank-Nicolson discretization with the Thomas algorithm as the solver. The horizontal domain is $[0, 14] \times [0, 14]$ with the horizontal space discretization interval, while the vertical domain is $[0, 4]$ with $\triangle z = 1$. The number of the grid points is $N_g = 1125$.

In addition, we assume that e_i, $i = 1, 2, 3$ in (4.47) are three different sorts of emissions for one species with distinct profiles. Their background estimates are denoted by $e_{bi}(t)$ respectively. They generate their corresponding evolution operator $M_{e_i}(t, t_0)$ such

that with the same assumptions of $\triangle t$ and grid points in the 3D domain for $\delta c(t)$, the discrete dynamic model of emission rates is given by

$$\delta e_i(t+\triangle t) = M_{ei}(t+\triangle t, t)\delta e_i(t), \quad i = 1, 2, 3,$$

where the j^{th} element on the diagonal of the diagonal matrix $M_{ei}(t+\triangle t, t)$ is given by $e_{bi}(t+\triangle t)/e_{bi}(t)$.

Here we consider the deposition rate as the input of model and then assume $\delta d(t) = 0$.

According to the discretization of the phase space, we assume the only one observation configuration is time-invariant in this example. It indicates that the observation operator mapping the state space to the observation space is a $1 \times 4N_g$ time-invariant matrix. Meanwhile, it shows that our objective in this example is to investigate the portions of emissions e_1, e_2 and e_3 in the location with observations.

The background phenomenon of the concentration and emission rates at the initial time by $c_b(t_0), e_{b1}(t_0), e_{b2}(t_0), e_{b3}(t_0)$. Then we set 800 (the ensemble number q) samplings respectively for the initial concentration and emission rates.

We produce the samplings by pseudo independent random numbers following the multivariate Gaussian distribution with zero mean and diagonal initial covariance of which the diagonal is given by

$$0.1(c_b(t_0), e_{b1}(t_0), e_{b2}(t_0), e_{b3}(t_0)),$$

which is 10% of the background phenomenon of the concentration and emission rates at the initial time, and further make the states correlated with their adjacent states by the moving average technique.

TSST	Fig. 4.13	Fig. 4.14	Fig. 4.15	Fig. 4.16	Fig. 4.17	Fig. 4.18	Fig. 4.19 Nonperiodic	Fig. 4.19 Periodic
δc	0.6687	0.6668	0.6750	0.6658	0.7240	1.0708	0.5300	1.0802
δe_1	0.5147	0.5157	0.5157	0.4895	0.0284	0.1386	0.0020	0.3414
δe_2	0.4758	0.4745	0.4794	0.4548	0.6094	0.1637	0.0019	0.2974
δe_3	0.3520	0.3575	0.3504	0.4496	0.6042	0.1369	0.0016	0.3308

Table 4.3: TSST of Fig. 4.13 to Fig. 4.19.

TSST%	Fig. 4.13	Fig. 4.14	Fig. 4.15	Fig. 4.16	Fig. 4.17	Fig. 4.18	Fig. 4.19 Nonperiodic	Fig. 4.19 Periodic
δc	0.3324	0.3310	0.3342	0.3232	0.3682	0.7091	0.9897	0.5270
δe_1	0.2560	0.2560	0.2550	0.2377	0.0145	0.0918	0.0038	0.1665
δe_2	0.2366	0.2355	0.2373	0.2208	0.3100	0.1084	0.0036	0.1451
δe_3	0.1750	0.1775	0.1735	0.2183	0.3073	0.0907	0.0029	0.1614

Table 4.4: TSST% of Fig. 4.13 to Fig. 4.19.

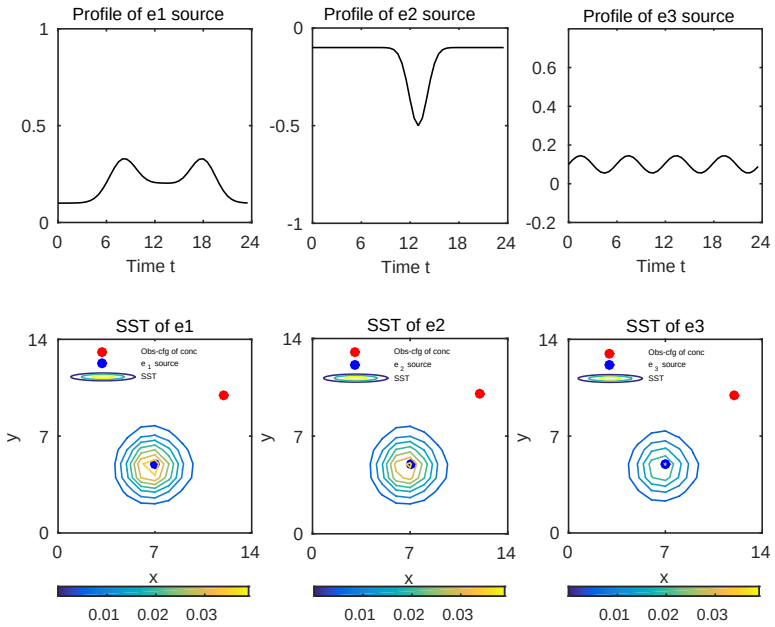

Figure 4.13: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

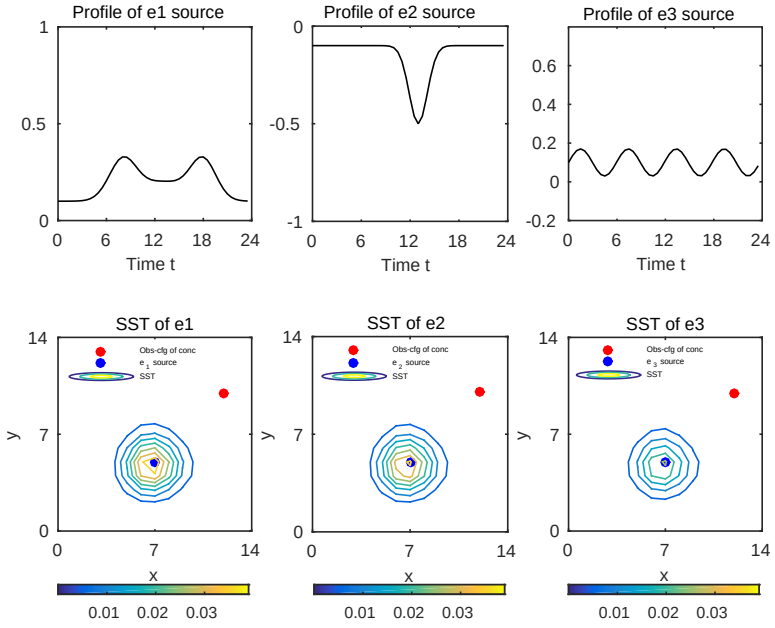

Figure 4.14: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

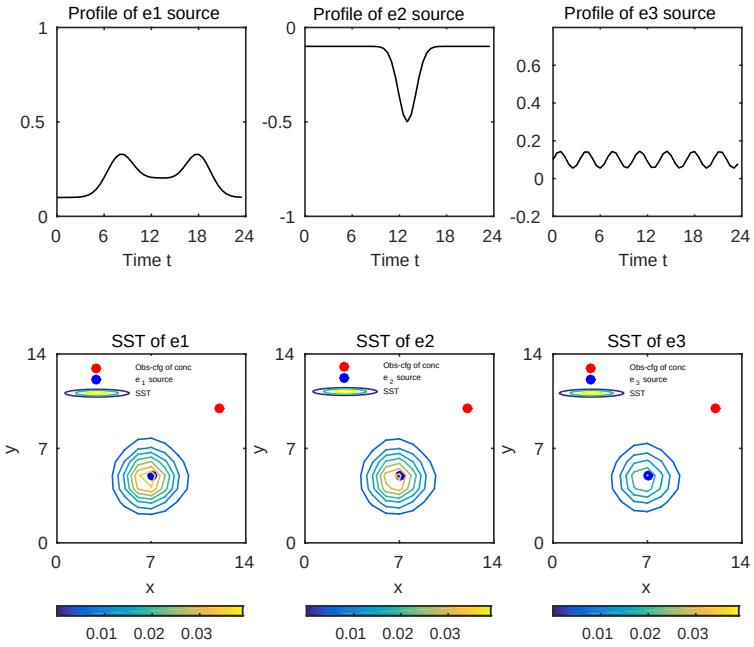

Figure 4.15: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

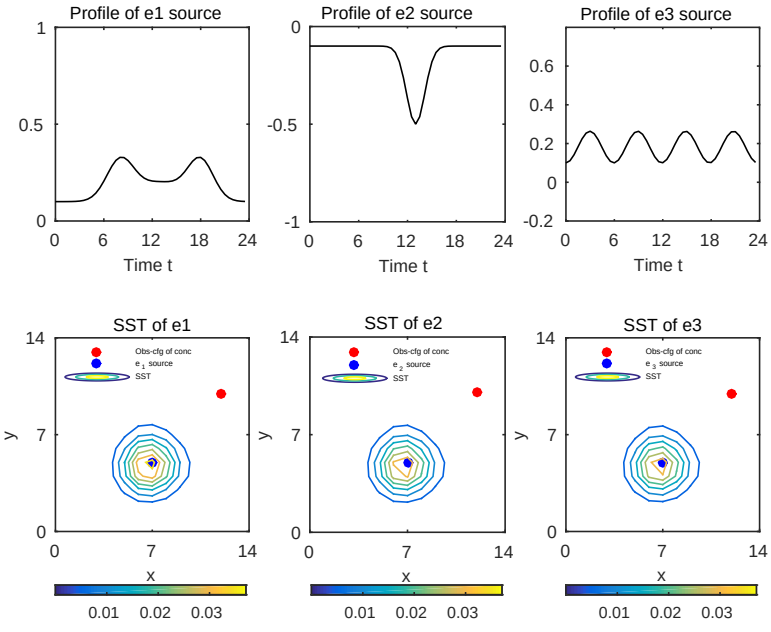

Figure 4.16: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

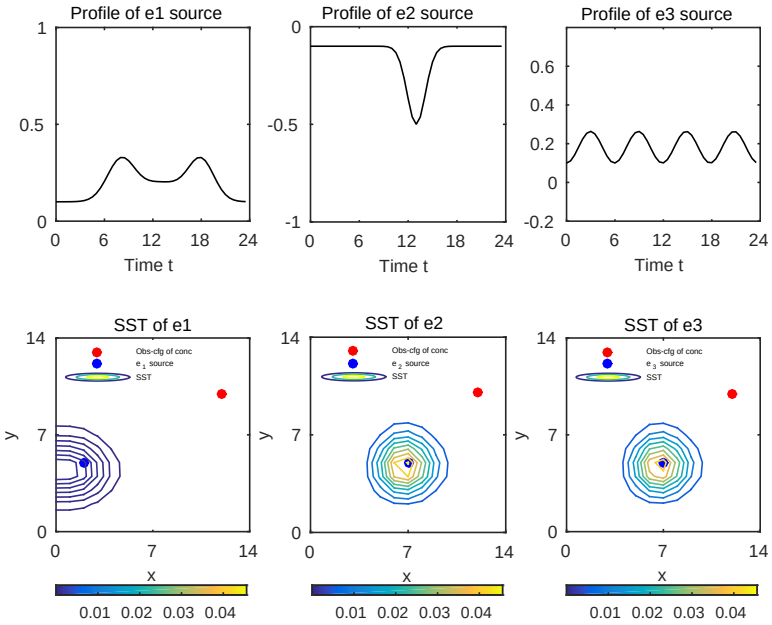

Figure 4.17: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(2, 5, 0)$ and $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

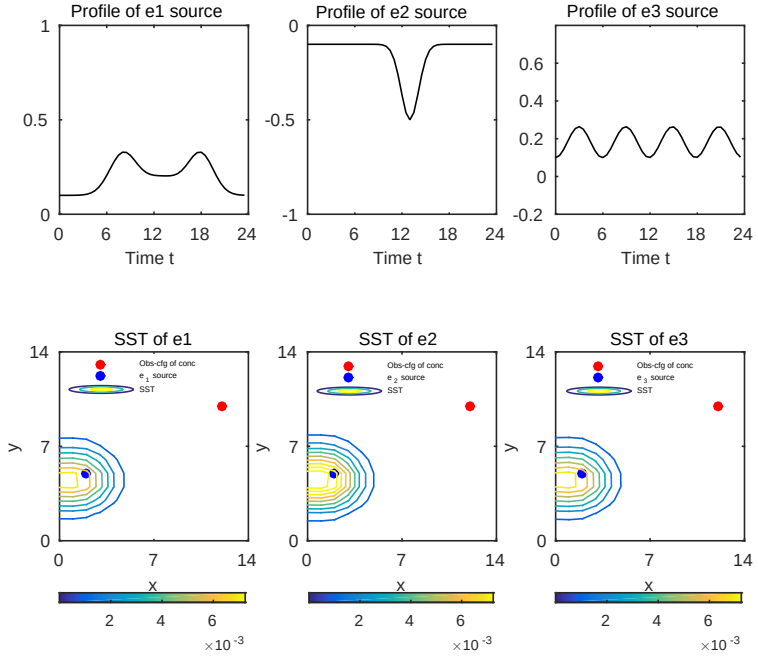

Figure 4.18: Contour lines: SST at $z = 0$. Wind direction: southwest. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(2, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

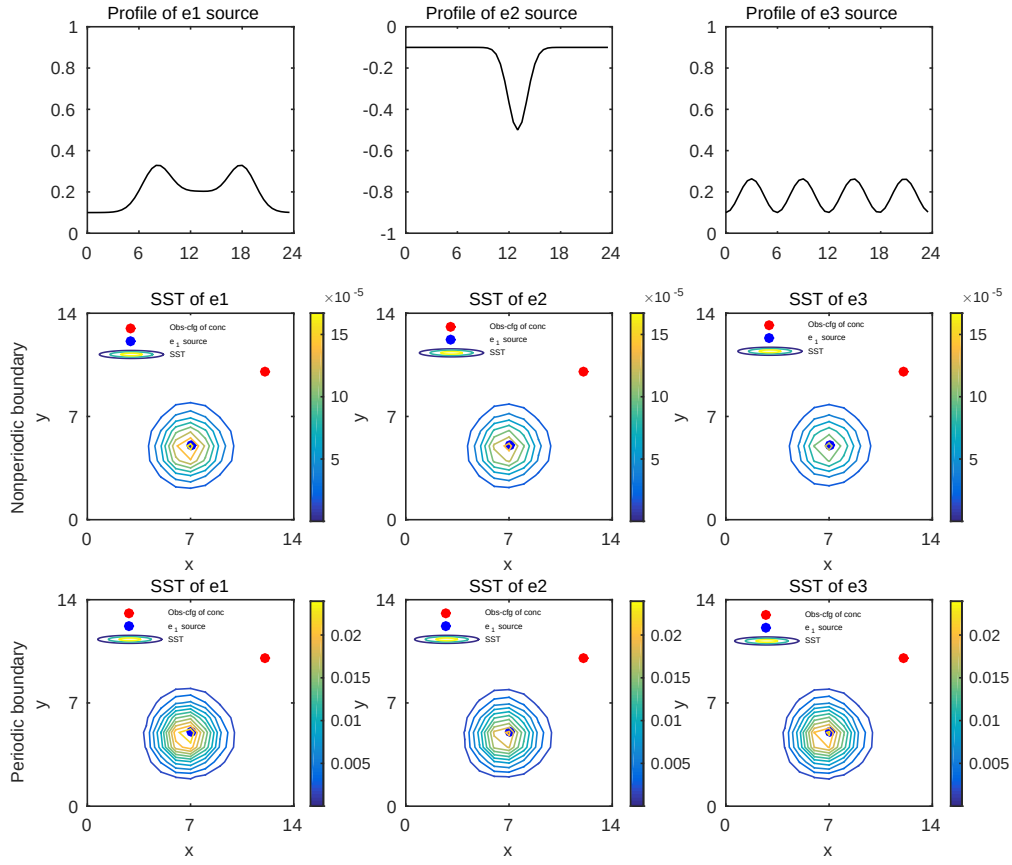

Figure 4.19: SST at $z = 0$ of different boundary conditions Wind direction: northeast. Dot at $(12, 10, 0)$: The single observation configuration. Dot at $(7, 5, 0)$: Emission source obtained from $\{e_{bi}(t_0)\}_{i=1}^{3}$.

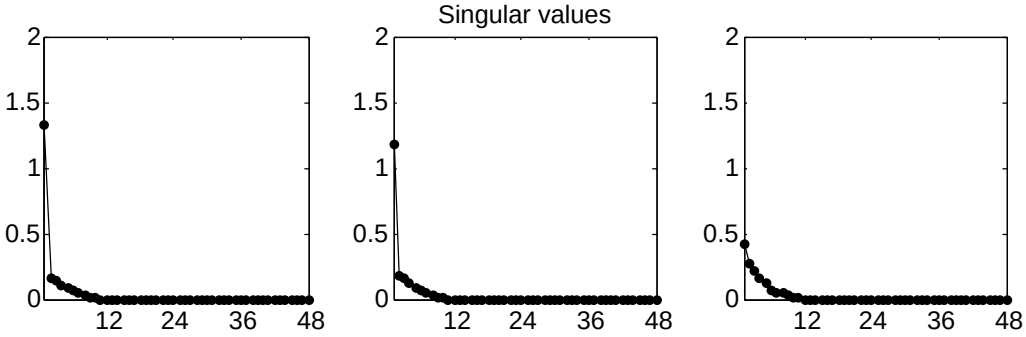

Figure 4.20: Singular values of Fig. 4.16, Fig. 4.17 and Fig. 4.18.

	Fig. 4.16		Fig. 4.17		Fig. 4.18	
	$\widetilde{\text{TSST}}$	$\widetilde{\text{TSST}}\%$	$\widetilde{\text{TSST}}$	$\widetilde{\text{TSST}}\%$	$\widetilde{\text{TSST}}$	$\widetilde{\text{TSST}}\%$
δc	0.0707	0. 0533	0.0554	0.0466	0	0
δe_1	0.4511	0.3397	0.0142	0.0119	0	0
δe_2	0.3878	0.2921	0.5562	0.4680	0	0
δe_3	0.4183	0.3150	0.5626	0.4734	0	0

Table 4.5: $\widetilde{\text{TSST}}$ and $\widetilde{\text{TSST}}\%$ of Fig. 4.16 to Fig. 4.18.

From Fig. 4.13 to Fig. 4.16, we clarify how the profiles and amount of emissions are related and contribute to the potential observations, or namely observation configurations. The only difference between the assumptions in Fig. 4.13 to Fig. 4.16 is the profile of e_3, which is described by a sine function. It can be seen from Table 4.3 and 4.4 that either the amplitude (Fig. 4.13) or frequency (Fig. 4.15) of the profile of e_3 is enlarged, TSST and TSST% of e_3 almost keep the same, which further make TSST and TSST% among e_1, e_2 and e_3 invariant. In fact, if we denote the profiles of e_3 in Fig. 4.13 to Fig. 4.15 by $M^1_{e_3}(t,t_0)$, $M^2_{e_3}(t,t_0)$, $M^3_{e_3}(t,t_0)$, respectively. With the same initial perturbation $\delta e_3(t_0)$, we have

$$\int_{t_0}^{t} M^1_{e_3}(s,t_0)\delta e_3(t_0)ds \approx \int_{t_0}^{t} M^2_{e_3}(s,t_0)\delta e_3(t_0)ds \approx \int_{t_0}^{t} M^3_{e_3}(s,t_0)\delta e_3(t_0)ds. \tag{4.48}$$

We can see from (4.48) that the variations of the profile of e_3 do not lead to the change of the total amount of e_3 during the data assimilation window, and further the change of TSSTs and TSST%s of $\delta e_1(t_0)$, $\delta e_2(t_0)$, $\delta e_3(t_0)$. In addition, compared to Fig. 4.13, we can find that the emission e_3 contributes more in Fig. 4.16. Denoting the profile of e_3 in Fig. 4.16 by $M^4_{e_3}(t,t_0)$, it is obvious that

$$\int_{t_0}^{t} M^4_{e_3}(s,t_0)\delta e_3(t_0)ds > \int_{t_0}^{t} M^i_{e_3}(s,t_0)\delta e_3(t_0)ds, \quad i=1,2,3. \tag{4.49}$$

(4.49) shows that the total amount of e_3 during the data assimilation window in Fig. 4.15 increases though the amplitude and frequency of the profile of e_3 are same to the profile in Fig. 4.13. Thus, it is clear that the profiles of emissions are closely related to the amount of emissions under the same initial perturbation of emissions. Therefore, we can conclude that the variations of profiles of emissions which bring changes of the emission amounts can effect SST, TSST and TSST% significantly, equivalently, change the emission apportionments.

It can be found in Fig. 4.13 to Fig. 4.16, the apportionment of e_1 is always dominant. Following the same assumptions in Fig. 4.16 except for giving a distinct information about the location of emission source of $e_1(t_0)$ from $e_{b1}(t_0)$, it is shown in Fig. 4.17 that the apportionment of e_1 decreases significantly. Actually, due to the southwesterly wind, we can foresee that the apportionment of e_1 will decrease, which is verified in Fig. 4.17.

In Fig. 4.18, we consider a more extreme case that all emissions are not well-detectable by setting the emission sources of e_2 and e_3 at the same locations as e_1 in Fig. 4.17. We

can clearly see from TSST% in Table 4.4 that all of emissions totally contribute only around 29% to the observations in Fig. 4.18. In other word, the concentration take around 70% of the observations in the location shown in Fig. 4.18. This property completely benefits from our extended model which both emissions and concentrations are included in the state vector of the model.

Distinct with Fig. 4.13 to Fig. 4.18, we assume there is a northeasterly wind in the case of Fig. 4.19 and show the influence of the boundary condition on the emission apportionments. It is easy to analyze that with the northeasterly wind, it is hard for the single observation at each time step in Fig. 4.19 to detect the three-type of emissions if the advection boundary condition is nonperiodic Dirichlet boundary condition. However, if the advection boundary condition is periodic and the wind speed is high enough, the emissions have the opportunity to be detected. This conclusion is perfectly shown in Fig. 4.19. The quantitative results are shown in Table 4.3 and Table 4.4.

In Fig. 4.20 we show the singular values in a decreasing sequence of the cases in Fig. 4.16, 4.17 and 4.18. It is obvious that there is only one leading singular value larger than 1 in the cases of Fig. 4.16 and 4.17, and no singular value larger than 1 in the situation of Fig. 4.18. If we only consider the contribution to TSST and TSST% of those directions dominant the normalized observation error, which are denoted by $\widetilde{\text{TSST}}$ and $\widetilde{\text{TSST}}\%$, respectively, we obtain the comparison among TSST, TSST% , $\widetilde{\text{TSST}}$ and $\widetilde{\text{TSST}}\%$ in Table 4.3 to Table 4.5 for Fig. 4.16 to Fig. 4.18.

It is worth noticing that $\widetilde{\text{TSST}}\%$s of the emissions are usually larger than the corresponding TSST%s since in those directions in which the variability of states are weaker than the variability of the observation error and within the first several time steps the emissions are hardly detected while the concentration is dominant. This leads us to "underestimate" $\widetilde{\text{TSST}}$ of the concentration since those directions, which are weak for emission but contain more information for the concentration, are not considered. This can be verified by observing the results in Table 4.5. Compared with TSST in Table 4.3 and TSST%s in Table 4.4, $\widetilde{\text{TSST}}$s of the concentration decrease more than emissions. Besides, we show in Table 4.5 that $\widetilde{\text{TSST}}$s and $\widetilde{\text{TSST}}\%$s are zeros for the reason that there is no singular value larger than 1 in the case of Fig. 4.18. This extreme case indicates, though we can gain the potential contributions of the concentration and emissions, it is probably unreliable if we apply the result into the real observation data, since the influence of observation errors during the data assimilation window on the data is stronger than the influence of the effective observations of the state.

4.4.4 Joint influence of observation configurations

We now assume that there is a sequence of observation operators $H_n(t_k)$, $n = 1, \cdots, m$, at each time step t_k, $k = 1, \cdots, N$. Each row of $H_n(t_k)$, $n = 1, \cdots, m$ represents a observation at different locations or of different species. Rearranging all the rows of $\{H_n(t_k)\}_{n=1}^m$ according to the state vector, we gain a new observation operator $H(t_k)$ including all the rows of $\{H_n(t_k)\}_{n=1}^m$. Thus, the sensitivities or contributions of different emission rates under the observation operator $\{H(t_k)\}_{k=1}^N$ can be viewed as ones the

jointly influenced by $H_n(t_k)$, $n = 1, \cdots, m$. In this section we study how the jointly sensitivities are related with the sensitivities under $H_n(t_k)$, $n = 1, \cdots, h$, respectively.

As an analogy to the notations in the last section, we denote the observation mapping related with $H_n(\cdot)$ as

$$R_n^{-\frac{1}{2}}(t_k)\delta y_n(t_k) = R_n^{-\frac{1}{2}}(t_k)\delta y_{c,n}(t_k) + R_n^{-\frac{1}{2}}(t_k)\nu(t_k), \quad \nu_n(t_k) \sim \mathcal{N}(0, R_n(t_k)).$$

Correspondingly, with above $H_n(\cdot)$ and $R_n(\cdot)$, $\mathcal{G}_n$ and $\mathcal{R}_n$ are similarly defined as $\mathcal{G}$ and $\mathcal{R}$. By singular value decomposition, we have

$$\mathcal{R}_n^{-\frac{1}{2}}\mathcal{G}_n P_n^{\frac{1}{2}}(t_0) = U_n S_n V_n^\top,$$

where U_n and V_n are orthogonal matrices consisted by singular vectors and S_n is rectangular matrix of which the diagonal consists of singular values. If we assume

$$\bar{U} = \begin{pmatrix} U_1 & & \\ & \ddots & \\ & & U_m \end{pmatrix},$$

then there exists an invertible matrix L such that $U = \bar{U}L$, since $\bar{U}$ is also an unitary matrix. Denoting the coefficients of observation perturbations under the basis U_n and U by C_n and C respectively, we obtain

$$\mathcal{R}_n^{-\frac{1}{2}}\delta y_n = U_n C_n, \quad \mathcal{R}^{-\frac{1}{2}}\delta y = UC.$$

Then by denoting

$$\bar{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix},$$

we have $\bar{C} = LC$. Thus, on one hand, similar with (4.45)

$$\mathrm{cov}(\bar{C}, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) = \begin{pmatrix} S_1 V_1^\top \\ \vdots \\ S_m V_m^\top \end{pmatrix}.$$

On the other hand,

$$\begin{aligned} \mathrm{cov}(\bar{C}, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) &= \mathrm{cov}(LC, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) \\ &= L \cdot \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) = LSV^\top. \end{aligned}$$

Since $\bar{U}$ and U are unitary matrices, by $L^\top L = U^\top \bar{U}\bar{U}^\top U = I$, then L is an unitary matrix as well. Hence,

$$\begin{aligned} & \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0))^\top \mathrm{cov}(C, P^{-\frac{1}{2}}(t_0)\delta x(t_0)) \\ = \; & VS^\top SV^\top \\ = \; & VS^\top L^\top LSV^\top \\ = \; & \mathrm{cov}(\bar{C}, P^{-\frac{1}{2}}(t_0)\delta x(t_0))^\top \mathrm{cov}(\bar{C}, P^{-\frac{1}{2}}(t_0)\delta x(t_0)). \end{aligned}$$

We can find that the joint SST, TSST and further TSST% from multiple observation configuration H can be calculated by the singular values and singular vectors based on $\{H_n\}$. It means that through the singular value decomposition based on a sequence of low-rank observation configurations, we can obtain their SSTs, TSST%s but also SST of the high-rank observation without additional high computational costs. It makes the selection of observation configurations more flexible.

Chapter 5

Optimal Control Locations for Time-Varying Systems in Hilbert Spaces on a Finite-Time Horizon

The choice of the locations of control hardware, such as actuators, plays an important role in the designs of control systems for many physical and engineering problems. Proper locations of actuators are essential to improve the performance of the controlled system.

In this chapter, we will start from the infinite-dimensional state space to consider the optimal location problem of controllers for time-varying systems on a finite-time horizon. In Section 5.2 we study the linear-quadratic optimal location control problem for both deterministic and stochastic systems and develop conditions guaranteeing the existence of optimal locations of linear quadratic control problems. Associated with practical applications, since optimal control problems cannot be solved directly in infinite-dimensional spaces, a sequence of approximations of the original time-varying system have to be considered. Thus, in Section 5.3, as an analogy to the approximation theory of time-invariant systems, we introduce the similar approximation conditions of evolution operators in order to ensure that the approximated control problems converge to the optimal control problem of the original infinite-dimensional time-varying system. Further, we show the convergence of the sequence of minimal costs and a subsequence of optimal locations of approximations.

5.1 Linear-quadratic optimal control problem

Throughout this chapter, we will always assume that the state space of the time-varying system is a real separable Hilbert space X, and the input and output space are Hilbert spaces denoted by U and Y, respectively. We firstly introduce the notion of mild evolution operators for the time-varying system.

Definition 5.1.1. *Denote* $\Gamma_a^b : \{(t,s)| -\infty < a \leqslant s \leqslant t \leqslant b < \infty\}$. *We call* $T(\cdot,\cdot) : \Gamma_a^b \to \mathcal{L}(X)$ *a* mild evolution operator *if*

1. $T(t,t) = I$,
2. $T(t,r)T(r,s) = T(t,s), \quad a \leqslant s \leqslant r \leqslant t \leqslant b$,
3. $T(\cdot,s) : [s,b] \to \mathcal{L}(X)$ *and* $T(t,\cdot) : [a,t] \to \mathcal{L}(X)$ *are strongly continuous,*
4. $\lambda := \sup\limits_{(t,s)\in\Gamma_a^b} \|T(t,s)\| < \infty$.

In the following we assume that $T(\cdot,\cdot) : \Gamma_a^b \to \mathcal{L}(X)$ is a mild evolution operator and $B \in L_s^\infty(a,b;U,X)$ with $B^* \in L_s^\infty(a,b;X,U)$. Here

$$L_s^\infty(a,b;X,Y) := \{F : [a,b] \to \mathcal{L}(X,Y) \mid F \text{ is strongly measurable and } \|F\|_\infty := \operatorname{esssup}_{t\in[a,b]} \|F(t)\| < \infty\}.$$

For the uniformly bounded mild evolution operator, the following perturbation theorem will be very helpful.

Theorem 5.1.2. *[43, Theorem 2.1] Let $T(\cdot,\cdot)$ be a mild evolution operator with uniformly bounded λ, $B \in L_{s,\infty}(\tau,t;X,X)$ and $(t,\tau) \in \Gamma_a^b$. Then there exists an uniquely determined mild evolution operator $T_B(\cdot,\cdot) : \Gamma_a^b \to \mathcal{L}(X)$ satisfying the integral equation*

$$T_B(t,\tau)x = T(t,\tau)x + \int_\tau^t T(t,s)B(s)T_B(s,\tau)x ds, \quad x \in X$$

$T_B(\cdot,\cdot)$ is called the perturbed evolution operator corresponding to the perturbation of $T(\cdot,\cdot)$ by B and

$$\|T_B(\cdot,\cdot)\| \leqslant \lambda e^{\lambda\|B\|_\infty (t-\tau)}.$$

Moreover, $T_B(t,\tau)$ is also the unique solution of

$$T_B(t,\tau)x = T(t,\tau)x + \int_\tau^t T_B(t,s)B(s)T(s,\tau)x ds, \quad x \in X.$$

For an initial time $t_0 \in [a,b]$, we consider the time-varying system described by

$$x(t) = T(t,t_0)x_0 + \int_{t_0}^t T(t,s)B(s)u(s)ds, \quad t \in [t_0,b], \tag{5.1}$$

where $x_0 \in X$ and $u \in L^2(t_0,b;U)$. We are interested in the following linear-quadratic optimal control problem.

Linear-Quadratic Optimal Control Problem: Find for $x_0 \in X$ a control $u_0 \in L^2(t_0,b;U)$ which minimizes the cost functional

$$\begin{aligned} & J(t_0,x_0,u) \\ = \; & \langle x(b), Gx(b)\rangle + \int_{t_0}^b \langle C(s)x(s), C(s)x(s)\rangle + \langle u(s), F(s)u(s)\rangle ds, \end{aligned} \tag{5.2}$$

where the function x is given by (5.1). Here $C \in L_s^\infty(a,b;X,Y)$, $G \in \mathcal{L}(X)$ and $F \in L_s^\infty(a,b;U,U)$ are self-adjoint and nonnegative for fixed t, and $F^{-1} \in L_s^\infty(a,b;U,U)$.

It is well known [43], that the linear-quadratic optimal control problem possesses for $x_0 \in X$ a unique solution u_0, which is given by

$$u_0(t) = -L(t)x(t), \quad t \in [t_0, b],$$

where $L(t) = F^{-1}(t)B^*(t)\Pi(t)$, such that the minimum of the cost functional is given by

$$\min_{u \in L^2(t_0,b;U)} J(t_0, x_0, u) = J(t_0, x_0, u_0) = \langle x_0, \Pi(t_0)x_0 \rangle,$$

where the self-adjoint nonnegative operator $\Pi(t)$ is the unique solution of the first integral Riccati equation (IRE)

$$\begin{aligned}\Pi(t)x = {} & T^*(b,t)GT(b,t)x \\ & + \int_t^b T^*(s,t)[C^*(s)C(s) - \Pi(s)B(s)F^{-1}(s)B^*(s)\Pi(s)]T(s,t)x ds \end{aligned} \tag{5.3}$$

and the second IRE

$$\begin{aligned}\Pi(t)x = {} & T_L^*(b,t)GT_L(b,t)x \\ & + \int_t^b T_L^*(s,t)[C^*(s)C(s) + \Pi(s)B(s)F^{-1}(s)B^*(s)\Pi(s)]T_L(s,t)x ds, \end{aligned} \tag{5.4}$$

where according to Theorem 5.1.2, we simply denote

$$\begin{aligned} T_L(t,\tau)x &:= T_{-BL}(t,\tau) \\ &= T(t,\tau)x - \int_\tau^t T(t,s)B(s)F^{-1}(s)B^*(s)\Pi(s)T_L(s,\tau)x ds, \ (t,\tau) \in \Gamma_a^b. \end{aligned}$$

5.2 Existence of optimal control locations

In this section we consider the situation having the opportunity to choose m locations to control and each location varies over a compact set $\Omega \subset \mathbb{R}^l$. We indicate these m locations by the parameter $r \in \Omega^m$, and denote the location-dependent input operator $B(\cdot)$ by $B_r(\cdot)$. Throughout this chapter, for the location-dependent control problem, we consider the time-varying system

$$x(t) = T(t,t_0)x(t_0) + \int_{t_0}^t T(t,s)B_r(s)u(s)ds, \quad t \in [t_0, b], r \in \Omega^m \tag{5.5}$$

with the corresponding cost functional

$$\begin{aligned} & J(t_0, x_0, u) \\ = {} & \langle x(b), Gx(b) \rangle + \int_{t_0}^b \langle C(s)x(s), C(s)x(s) \rangle + \langle u(s), F(s)u(s) \rangle ds. \end{aligned} \tag{5.6}$$

The unique feedback control such that

$$\min_{u\in L^2(t_0,b;U)} J(t_0,x_0,u) = J(t_0,x_0,u_0) = \langle x_0, \Pi_r(t_0)x_0\rangle,$$

is given by

$$u(t) = -L_r(t)x(t),$$

where $L_r(t) = F^{-1}(t)B_r^*(t)\Pi_r(t)$ and $\Pi_r(t)$ is the unique solution of

$$\begin{aligned}\Pi_r(t)x = {} & T^*(b,t)GT(b,t)x \\ & + \int_t^b T^*(s,t)[C^*(s)C(s) - L_r^*(s)F(s)L_r^*(s)]T(s,t)xds \end{aligned} \tag{5.7}$$

and

$$\begin{aligned}\Pi_r(t)x = {} & T_{L,r}^*(b,t)GT_{L,r}(b,t)x \\ & + \int_t^b T_{L,r}^*(s,t)[C^*(s)C(s) + L_r^*(s)F(s)L_r^*(s)]T_{L,r}(s,t)xds, \end{aligned} \tag{5.8}$$

where

$$\begin{aligned} T_{L,r}(t,\tau)x &:= T_{-B_rL_r}(t,\tau) \\ &= T(t,\tau)x - \int_\tau^t T(t,s)B_r(s)L_r(s)T_{L,r}(s,\tau)xds, \quad (t,\tau)\in\Gamma_a^b. \end{aligned}$$

In most cases, the initial state x_0 is not fixed. This indicates several different ways to define the optimal actuator location problem. We take two possible ways into account here. The first one is to minimize the cost with the worst choice of initial value, which is

$$\max_{\|x_0\|=1}\min_{u\in L^2(t_0,b;U)} J_r(x_0,u) = \max_{\|x_0\|=1}\langle x_0,\Pi_r(t_0)x_0\rangle = \|\Pi_r(t_0)\|.$$

Let $\ell^r(t_0) := \|\Pi_r(t_0)\|$, the optimal performance of r, if it exists, is defined as

$$\hat{\ell}(t_0) = \inf_{r\in\Omega^m}\|\Pi_r(t_0)\|.$$

The second one is to assume that the system is stochastic. Thus, we need to consider the trace of $\Pi_r(t_0)$ instead, since the trace indicates the sum of the deviation of the state vector in each coordinate. Thus the evaluation of the particular performance of r is given by the nuclear norm of $\Pi_r(t_0)$, which is denoted by $\ell_1^r(t_0) := \|\Pi_r(t_0)\|_1$. Further, the optimal performance, if it exists, is defined as

$$\hat{\ell}_1(t_0) = \inf_{r\in\Omega^m}\|\Pi_r(t_0)\|_1.$$

It is worth noting that the optimal location r for $\Pi(t)$ in both norms relies on the time t. We simplify the notation $r(t)$ as r in the rest of this thesis.

For time-invariant systems on an infinite time horizon this optimal location problem is studied in [75]. In this section we prove the existence of optimal control locations for deterministic as well as stochastic time-varying systems on a finite-time horizon. For these proofs, the following theorem in [43, Theorem 5.1] is needed.

Theorem 5.2.1. *[43, Theorem 5.1] For Hilbert spaces X, U, Y with the operators B_i, C, G, F satisfying the same assumptions in Section 5.1, we consider a sequence of optimal control problem with the initial states $x_i(t_0)$ and let $x_i(t_0) \to x(t_0)$, the optimal controls $u_i(\cdot)$, the optimal trajectories $x_i(\cdot)$ have the corresponding cost function*

$$J(t_0, x_i(t_0), u) = \langle x(b), Gx(b)\rangle + \int_{t_0}^{b} \langle C(t)x(t), C(t)x(t)\rangle + \langle u(t), F(t)u(t)\rangle dt, \quad (5.9)$$

where

$$x(t) = T_i(t, t_0)x(t_0) + \int_{t_0}^{t} T_i(t, s)B_i(s)u(s)ds.$$

For the sequence of optimal problems, assume there exists a optimal control problem with the cost functional (5.9) *governed by*

$$x(t) = T_i(t, t_0)x(t_0) + \int_{t_0}^{t} T_i(t, s)B_i(s)u(s)ds,$$

where $T(\cdot, \cdot)$ the uniformly bounded mild evolution operator and input operator $B \in L_{s,\infty}(t_0, b; X, U)$, such that
(1) for each $x \in X$,

$$(i)\ \ T_i(t, s)x \to T(t, s)x; \quad (ii)\ \ T_i^*(t, s)x \to T^*(t, s)x, \quad t_0 \leqslant s \leqslant t \leqslant b,$$

(2) for each $u \in U$, $x \in X$

$$(i)\ \ B_i(t)u \to B(t)u; \quad (ii)\ \ B_i^*(t)x \to B^*(t)x,\ a.e.\ t \in [t_0, b],$$

(3) for each $x \in X$,

$$(i)\ \ C_i(t)x \to C(t)x; \quad (ii)\ \ C_i^*(t)y \to C^*(t)y,\ a.e. \quad t \in [t_0, b],$$

(4) for each $x \in X$,

$$G_i x \to Gx,$$

(5) for each $u \in U$,

$$F_i(t)u \to F(t)u;\ a.e.$$

and $\|T_i(\cdot, \cdot)\|$, $\|B_i\|_\infty$, $\|C_i\|_\infty$, $\|G_i\|$, $\|F_i\|_\infty$ are uniformly bounded, then we have

$$u_i(t) \to u(t),$$
$$x_i(t) \to x(t),$$
$$\Pi_i(t)x \to \Pi(t)x,$$

where $\Pi_i(\cdot)$ are the solutions of the sequence of the integral Riccati equations and $\Pi(\cdot)$ is the solution of the Riccati equation of the original optimal problem.

Theorem 5.2.2. *Let $\{B_r\}_{r\in\Omega^m}$ be a family of compact operators valued functions with the property that $\lim_{r\to r_0}\|B_r - B_{r_0}\|_\infty = 0$ for any $r_0 \in \Omega^m$. Then the solution of the corresponding integral Riccati equation Π_r satisfies*

$$\lim_{r\to r_0}\|\Pi_r(t) - \Pi_{r_0}(t)\| = 0, \quad t \in [a,b]$$

and for any initial time $t_0 \in [a,b]$, there exists an optimal location $\hat{r}$ such that

$$\hat{\ell}(t_0) = \|\Pi_{\hat{r}}(t_0)\| = \inf_{r\in\Omega^m}\|\Pi_r(t_0)\|.$$

Proof. Due to the assumptions on B_r, there exists $\delta > 0$ such that

$$\lambda_B := \sup\{\|B_r(t)\| \mid t \in [a,b], \|r - r_0\| \leqslant \delta\} < \infty.$$

We denote

$$\mathbf{B}(r_0,\delta) := \{r \in \Omega^m : \|r - r_0\| \leqslant \delta\}.$$

Theorem 5.2.1 implies for every $x \in X$

$$\Pi_r(t)x \to \Pi_{r_0}(t)x, \quad r \to r_0.$$

For any feedback control $\tilde{u}(t) = \tilde{L}(t)x(t)$, $\tilde{L} \in L_s^\infty(a,b;X,U)$,

$$\begin{aligned}
&\langle x(t), \Pi_r(t)x(t)\rangle = \min_{u\in L_2([t_0,b];U)} J(t,x(t),u) \leqslant J(t,x(t),\tilde{u}) \\
= \;& \langle x(b), Gx(b)\rangle + \int_t^b \|C(s)x(s)\|^2 + \langle \tilde{L}(s)x(s), F(s)\tilde{L}(s)x(s)\rangle ds \\
= \;& \|G^{\frac{1}{2}}T_{\tilde{L},r}(b,t)x(t)\|^2 \\
&+ \int_t^b \|C(s)T_{\tilde{L},r}(s,t)x(t)\|^2 + \|F^{\frac{1}{2}}(s)\tilde{L}(s)T_{\tilde{L},r}(s,t)x(t)\|^2 ds, \qquad (5.10)
\end{aligned}$$

where

$$T_{\tilde{L},r}(t,\tau)x := T(t,\tau)x + \int_\tau^t T(t,s)B_r(s)\tilde{L}(s)T_{\tilde{L},r}(s,\tau)x ds, \quad (t,\tau) \in \Gamma_a^b.$$

Since the family $\{B_r\}$ is uniformly bounded by λ_B on $\mathbf{B}(r_0,\delta)$, Theorem 5.1.2 implies for all $r \in \mathbf{B}(r_0,\delta)$, $(t,\tau) \in \Gamma_a^b$,

$$\|T_{\tilde{L},r}(t,\tau)\| \leq \lambda e^{(\lambda\lambda_B\|\tilde{L}\|_\infty(t-\tau))}.$$

Further, because $C \in L_s^\infty(a,b;X,Y)$, $F \in L_s^\infty(a,b;U,U)$, there exists a constant λ_Π, independent of t and $r \in \mathbf{B}(r_0,\delta)$, such that $\|\Pi_r\|_\infty \leqslant \lambda_\Pi$.

For $S_r = C^*C - L_r^*FL_r$, where $L_r = F^{-1}B_r^*\Pi_r$, we obtain

$$\Pi_r(t)x - \Pi_{r_0}(t)x = \int_t^b T^*(s,t)\left(S_r(s) - S_{r_0}(s)\right)T(s,t)x ds, \quad x \in X.$$

Since $F^{-1} \in L_s^\infty(a,b;U,U)$ and the operator $B_{r_0}(t)$ is compact for any $t \in [a,b]$, we have

$$\begin{aligned}\|L_r^*(t) - L_{r_0}^*(t)\| &= \|L_{r_0}^*(t)F(t)L_{r_0}(t) - L_r^*(t)F(t)L_r(t)\| \\ &\leqslant \|F^{-1}\|_\infty(\|\Pi_r(t)\|\|B_r(t) - B_{r_0}(t)\| \\ &\quad +\|(\Pi_r(t) - \Pi_{r_0}(t))B_{r_0}(t)\|) \longrightarrow 0, \quad r \to r_0,\end{aligned}$$

which shows

$$\begin{aligned}\|S_r(t) - S_{r_0}(t)\| &\leqslant \|L_{r_0}^*(t) - L_r^*(t)\|\|F(t)L_{r_0}(t)\| \\ &\quad +\|L_r^*(t)F(t)\|\|L_{r_0}(t) - L_r(t)\| \to 0, \quad r \to r_0.\end{aligned}$$

From the uniform boundedness of F, B_r and Π_r on $\mathbf{B}(r_0,\delta)$, L_r and further S_r are uniformly bounded for all $t \in [a,b]$ and $\mathbf{B}(r_0,\delta)$. According to the dominated convergence theorem, we obtain

$$\|\Pi_r(t) - \Pi_{r_0}(t)\| \to 0, \quad r \to r_0.$$

Additionally, since $r \in \Omega^m$, Ω^m is a compact set, there exists an optimal location $\hat{r}$ depending on t_0 such that

$$\|\Pi_{\hat{r}}(t_0)\| = \inf_{r\in\Omega^m} \|\Pi_r(t_0)\|.$$

□

Theorem 5.2.2 shows the continuity of optimal actuator locations and existence of the optimal location in the operator norm. For stochastic systems, the above problem leads to the nuclear norm. Thus, first we develop conditions which guarantee that the Riccati operator is a nuclear operator. Similar to [23, Theorem 3.1], we have

Theorem 5.2.3. *Let $T(\cdot,\cdot)$ be a mild evolution operator on X, $B \in L_s^\infty(a,b;\mathbb{C}^p,X)$, and $C \in L_s^\infty(a,b;X,\mathbb{C}^q)$. Then for any $t_0 \in [a,b]$ we have:*
(1) The observability operator $\mathcal{C}_{t_0} : X \to L^2(t_0,b;\mathbb{C}^q)$ defined by

$$(\mathcal{C}_{t_0}x_0)(\cdot) = C(\cdot)T(\cdot,t_0)x_0, \quad x_0 \in X$$

is a Hilbert-Schmidt operator.
(2) The controllability operator $\mathcal{B}_{t_0} : L^2(t_0,b;\mathbb{C}^p) \to X$ defined by

$$\mathcal{B}_{t_0}u = \int_{t_0}^b T(b,s)B(s)u(s)ds$$

is a Hilbert-Schmidt operator.
(3) $\mathcal{C}_{t_0}^*\mathcal{C}_{t_0}$ and $\mathcal{B}_{t_0}\mathcal{B}_{t_0}^*$ are nuclear operators.

Proof. (1) Defining $\mathcal{C}_{t_0,i} : X \to L^2(t_0,b)$, $i \in \{1,\dots,q\}$

$$(\mathcal{C}_{t_0,i}x_0)(s) = \langle C(s)T(s,t_0)x_0, e_i\rangle, \quad s \geqslant t_0,$$

where $\{e_i\}$ is the standard orthogonal basis of $\mathbb{C}^q$. We have

$$\begin{aligned} |(\mathcal{C}_{t_0,i}x_0)(s)| &= |\langle C(s)T(s,t_0)x_0, e_i\rangle| \leqslant \|C(s)T(s,t_0)x_0\|\|e_i\| \\ &\leqslant \|C(s)\|\|T(s,t_0)\|\|x_0\| < \infty. \end{aligned}$$

The uniform boundedness of $C, T(\cdot,\cdot)$ and [98, Theorem 6.12] imply that $\mathcal{C}_{t_0,i}$ is Hilbert-Schmidt, that is, for any orthogonal basis $\{\bar{e}_i\}$ of X, we have

$$\sum_{i=1}^{q}\sum_{j=1}^{\infty} \|\mathcal{C}_{t_0,i}\bar{e}_j\|^2_{L^2(t_0,b)} < \infty.$$

Since $\|\mathcal{C}_{t_0}\bar{e}_j\|^2_{L^2(t_0,b)} = \sum_{i=1}^{q} \|\mathcal{C}_{t_0,i}\bar{e}_j\|^2_{L^2(t_0,b)}$, we have

$$\sum_{j=1}^{\infty} \|\mathcal{C}_{t_0}\bar{e}_j\|^2_{L^2(t_0,b;\mathbb{C}^q)} = \sum_{i=1}^{q}\sum_{j=1}^{\infty} \|\mathcal{C}_{t_0,i}\bar{e}_j\|^2_{L^2(t_0,b)} < \infty,$$

which shows that $\mathcal{C}_{t_0}$ is a Hilbert-Schmidt operator.

(2) According to [98, Theorem 6.9], $\mathcal{B}_{t_0}$ is a Hilbert-Schmidt operator if and only if $\mathcal{B}^*_{t_0}$ is a Hilbert-Schmidt operator. An easy calculation shows $\mathcal{B}^*_{t_0} : X \to L^2(t_0, b; U)$,

$$(\mathcal{B}^*_{t_0}x)(\cdot) = B^*_{t_0}(\cdot)T^*(b,\cdot)x.$$

From *(1)*, $\mathcal{B}^*_{t_0}$ is Hilbert-Schmidt, and so is $\mathcal{B}_{t_0}$.

(3) Since

$$\begin{aligned} \|\mathcal{C}^*_{t_0}\mathcal{C}_{t_0}\|_1 &\leqslant \|\mathcal{C}^*_{t_0}\|_{HS}\|\mathcal{C}_{t_0}\|_{HS} < \infty, \\ \|\mathcal{B}_{t_0}\mathcal{B}^*_{t_0}\|_1 &\leqslant \|\mathcal{B}^*_{t_0}\|_{HS}\|\mathcal{B}_{t_0}\|_{HS} < \infty, \end{aligned}$$

$\mathcal{C}^*_{t_0}\mathcal{C}_{t_0}$ and $\mathcal{B}_{t_0}\mathcal{B}^*_{t_0}$ are nuclear operators. □

Corollary 5.2.4. *Assume that the input space U and the output space Y are finite-dimensional spaces and G is a nuclear operator, then the unique nonnegative self-adjoint solution $\Pi(t_0)$ of the integral Riccati equation is a nuclear operator.*

Proof. Defining the bounded operator $\mathcal{C}_{t_0} : X \to L^2(t_0, b; U \times Y)$ by

$$(\mathcal{C}_{t_0}x_0)(\cdot) = \begin{pmatrix} C(\cdot) \\ F^{\frac{1}{2}}(\cdot)L(\cdot) \end{pmatrix} T_L(\cdot, t_0)x_0, \quad L = F^{-1}B^*\Pi.$$

$\mathcal{C}_{t_0}$ is Hilbert-Schmidt by Theorem 5.2.3 *(1)* The second IRE (5.4) can be rewritten as

$$\Pi(t_0)x = T^*_L(b,t_0)GT_L(b,t_0)x + \mathcal{C}^*_{t_0}\mathcal{C}_{t_0}x, \quad x \in X.$$

From Theorem 5.2.3 *(3)* and the nuclearity of G, $\Pi(t)$ is a nuclear operator. □

Lemma 5.2.5. *Assume $T(\cdot,\cdot)$ and $T_i(\cdot,\cdot)$, $i \in \mathbb{N}$, are mild evolution operators which are uniformly bounded by λ_T, $D_i, D \in L_s^\infty(a,b;X,X)$ satisfy*

$$\|D_i(t)x - D(t)x\| \to 0, \quad i \to \infty.$$

for every $x \in X$ and $\sup_i\{\|D_i\|_\infty, \|D\|_\infty\} \leqslant \lambda_D$. $T_{D_i}(\cdot,\cdot)$, $T_D(\cdot,\cdot)$ denote the perturbed evolution operators corresponding to the perturbation of $T_i(\cdot,\cdot)$ by D_i and $T(\cdot,\cdot)$ by D. If $\|T_i(t,\tau)x - T(t,\tau)x\| \to 0$ as $i \to \infty$ for $x \in X$, then for any $(t,\tau) \in \Gamma_a^b$ and $x \in X$,

$$\|T_{D_i}(t,\tau)x - T_D(t,\tau)x\| \to 0, \quad i \to \infty.$$

Proof. As in [20], we construct $T_{D_i}(t,\tau)$ as $T_{D_i}(t,\tau) = \sum_{n=0}^\infty T_{D_i,n}(t,\tau)$, where

$$T_{D_i,0}(t,\tau) = T_i(t,\tau), \quad T_{D_i,n}(t,\tau)x = \int_\tau^t T_i(t,s)D_i(s)T_{D_i,n-1}(s,\tau)x ds,\ x \in X.$$

By induction we obtain

$$\|T_{D_i,n}(t,\tau)\| \leqslant \lambda_T(\lambda_T\lambda_D)^n \frac{(t-\tau)^n}{n!}.$$

$T_D(t,\tau)$ can be constructed in a similar manner with the same upper bound.

Defining $d_{i,n}(t,\tau) = T_{D_i,n}(t,\tau) - T_{D,n}(t,\tau)$, we have

$$\begin{aligned}
d_{i,0}(t,\tau) =& T_i(t,\tau) - T(t,\tau),\\
d_{i,n}(t,\tau) =& \int_\tau^t T_i(t,s)D_i(s)d_{i,n-1}(s,\tau)ds\\
&+ \int_\tau^t T_i(t,s)[D_i(s) - D(s)]T_{D,n-1}(s,\tau)ds\\
&+ \int_\tau^t [T_i(t,s) - T(t,s)]D(s)T_{D,n-1}(s,\tau)ds.
\end{aligned}$$

The uniform boundedness of $\{T_{D_i}(t,\tau)\}_{i\in\mathbb{N}}$ and $T_D(t,\tau)$ implies

$$\begin{aligned}
\|\sum_{n=0}^\infty d_{i,n}(t,\tau)\| &= \|\sum_{n=0}^\infty (T_{D_i,n}(t,\tau) - T_{D,n}(t,\tau))\|\\
&\leqslant \|T_{D_i}(t,\tau)\| + \|T_D(t,\tau)\| < \infty.
\end{aligned}$$

Due to $\sup_i\{\|D_i\|_\infty, \|D\|_\infty\} \leqslant \lambda_D$ and $T(\cdot,\cdot)$, $T_i(\cdot,\cdot)$ are uniformly bounded, the mild evolution operators $T_D(\cdot,\cdot)$, $T_{D_i}(\cdot,\cdot)$ are uniformly bounded. Further for any $n \in \mathbb{N}$,

$$\sup_i \sup_{(t,\tau)\in\Gamma_{t_0}^b} \|d_{i,n}(t,\tau)\| < \infty.$$

Meanwhile, since $\|D_i(t)x - D(t)x\| \to 0$, we gain

$$\|d_{i,0}(t,\tau)x\| = \|T_i(t,\tau)x - T(t,\tau)x\| \to 0,\ i \to \infty.$$

Hence,

$$\begin{aligned}\|d_{i,n}(t,\tau)x\| &\leqslant \int_\tau^t \|T_i(t,s)\|\|D_i(s)\|\|d_{i,n-1}(t,\tau)x\|ds \\ &\quad + \int_\tau^t \|T_i(t,s)\|\|[D_i(s)-D(s)]T_{D,n-1}(s,\tau)x\|ds \qquad (5.11)\\ &\quad + \int_\tau^t \|(T_i(t,s)-T(t,s))(s)T_{D,n-1}(s,\tau)x\|ds \longrightarrow 0, \quad i\to\infty.\end{aligned}$$

By the dominated convergence theorem, we have

$$\begin{aligned}\|T_{D_i}(t,\tau)x - T_D(t,\tau)x\| &= \sum_{n=0}^{\infty}\|T_{D_i,n}(t,\tau)x - T_{D,n}(t,\tau)x\| \\ &\leqslant \sum_{n=0}^{\infty}\|d_{i,n}(t,\tau)x\| \to 0, \quad i\to\infty.\end{aligned}$$

□

Corollary 5.2.6. *For any mild evolution operator $T(\cdot,\cdot)$ with uniform bound λ_T and $D_i, D \in L_s^\infty(t_0,b;X,X)$ with $\sup_i\{\|D_i\|_\infty,\|D\|_\infty\} \leqslant \lambda_D$, if $\|D_i(t)-D(t)\| \to 0$. $T_D(\cdot,\cdot)$ and $T_{D_i}(\cdot,\cdot)$ are the perturbations of $T(\cdot,\cdot)$ by D and D_i, then we have*

$$\|T_{D_i}(t,\tau) - T_D(t,\tau)\| \to 0, \quad i\to\infty.$$

Proof. We let $T_i = T$ in Lemma 5.2.5 and replace (5.11) by

$$\begin{aligned}\|d_{i,n}(t,\tau)\| &\leqslant \int_\tau^t \|T(t,s)\|\|D_i(s)\|\|d_{i,n-1}(t,\tau)\|ds \\ &\quad + \int_\tau^t \|T(t,s)\|\|D_i(s)-D(s)\|\|T_{D,n}(s,\tau)\|ds \to 0, \quad i\to\infty.\end{aligned}$$

Then, we can prove the uniform convergence of $T_{D_i}(t,\tau)$ by the dominated convergence theorem in the similar way with Lemma 5.2.5. □

Theorem 5.2.7. *We consider the time-varying system* (5.5) *with the location-dependent input operators and the cost functional* (5.6)*. Assume B_r, $r \in \Omega^m$ satisfies*

$$\lim_{r\to r_0}\|B_r - B_{r_0}\|_\infty = 0, \quad r_0 \in \Omega^m,$$

U and Y are finite-dimensional spaces and G is a nuclear operator, then

$$\lim_{r\to r_0}\|\Pi_r(t) - \Pi_{r_0}(t)\|_1 = 0, \quad t\in[t_0,b]$$

and there exists an optimal location $\hat{r}$ depending on t such that

$$\hat{\ell}_1(t) = \|\Pi_{\hat{r}}(t_0)\|_1 = \inf_{r\in\Omega^m}\|\Pi_r(t_0)\|_1.$$

Proof. Similar to Theorem 5.2.2, there exists $\delta > 0$ such that $\sup_{r\in\mathbf{B}(r_0,\delta)} \|B_r\| < \infty$, $r_0 \in \Omega^m$ and for every $x \in X$ and $t \in [t_0, b]$,

$$\Pi_r(t)x \to \Pi_{r_0}(t)x, \quad r \to r_0.$$

From (5.10), we know that Π_r are uniformly bounded with λ_Π for any $t \in [t_0, b]$ and $r \in \mathbf{B}(r_0, \delta)$.

Defining the operator $\mathcal{C}_{t,r} : X \to L^2(t, b; U \times Y)$, $t \in [t_0, b]$,

$$(\mathcal{C}_{t,r}x(t))(\cdot) = \begin{pmatrix} C(\cdot) \\ F^{-\frac{1}{2}}(\cdot)B_r^*(\cdot)\Pi_r(\cdot) \end{pmatrix} T_{L,r}(\cdot, t)x(t). \tag{5.12}$$

Corollary 5.2.4 has shown that $\mathcal{C}_{t,r}$ is a Hilbert-Schmidt operator and

$$\Pi_r(t) = T_{L,r}^*(b, t)GT_{L,r}(b, t) + \mathcal{C}_{t,r}^*\mathcal{C}_{t,r}.$$

is nuclear if G is nuclear.

Next we show that $\mathcal{C}_{t,r}$ uniformly converges to $\mathcal{C}_{t,r_0}$ in the Hilbert-Schmidt norm. Let $\{e_i\}_{i=1}^{p+q}$ and $\{\bar{e}_i\}_{i=1}^{\infty}$ be the orthogonal basis of $U \times Y$ and X respectively, then

$$\begin{aligned}
&\|\mathcal{C}_{t,r} - \mathcal{C}_{t,r_0}\|_{HS} \\
=&\sum_{i=1}^{\infty}\int_t^b \sum_{j=1}^{p+q} \langle(\mathcal{C}_{t,r}\bar{e}_i)(s) - (\mathcal{C}_{t,r_0}\bar{e}_i)(s), e_j\rangle_{U\times Y} ds \\
=&\sum_{i=1}^{\infty}\int_t^b \sum_{j=1}^{p+q} |\langle \bar{e}_i, T_{L,r}^*(s,t)[C^*(s), L_r^*(s)F^{\frac{1}{2}}(s)]e_j \\
&- T_{L,r_0}^*(s,t)[C^*(s), L_{r_0}^*(s)F^{\frac{1}{2}}(s)]e_j\rangle_X|^2 ds \\
=&\int_t^b \sum_{i=1}^{\infty}\sum_{j=1}^{p+q} |\langle \bar{e}_i, T_{L,r}^*(s,t)[C^*(s), L_r^*(s)F^{\frac{1}{2}}(s)]e_j \\
&- T_{L,r_0}^*(s,t)[C^*(s), L_{r_0}^*(s)F^{\frac{1}{2}}(s)]e_j\rangle_X|^2 ds \\
=&\sum_{j=1}^{p+q}\int_t^b \|T_{L,r}^*(s,t)[C^*(s), L_r^*(s)F^{\frac{1}{2}}(s)]e_j \\
&- T_{L,r_0}^*(s,t)[C^*(s), L_{r_0}^*(s)F^{\frac{1}{2}}(s)]e_j\|_X^2 ds,
\end{aligned}$$

where $L_r = F^{-1}B_r^*\Pi_r$.

From Theorem 5.2.2, we have $\lim_{r\to r_0} \|L_r(t) - L_{r_0}(t)\| = 0$ and $\|L_r\|_\infty < \infty$. Then,

$$\begin{aligned}
&\|B_r(t)L_r(t) - B_{r_0}(t)L_{r_0}(t)\| \\
\leqslant&\|B_r(t)\|\|L_r(t) - L_{r_0}(t)\| + \|B_r(t) - B_{r_0}(t)\|\|L_{r_0}(t)\| \to 0, \quad r \to r_0
\end{aligned} \tag{5.13}$$

and $\|B_r L_r\|_\infty < \infty$. Hence, from Corollary 5.2.6, for any $(s,t) \in \Gamma_{t_0}^b$, $T_{L,r}(s,t)$ uniformly converges to $T_{L,r_0}(s,t)$. Therefore,

$$\begin{aligned} & \|T_{L,r}^*(s,t)[C^*(s), L_r^*(s)F^{\frac{1}{2}}(s)]e_j - T_{L,r_0}^*(s,t)[C^*(s), L_{r_0}^*(s)F^{\frac{1}{2}}(s)]e_j\|_X \\ \leqslant\ & \|(T_{L,r}^*(s,t) - T_{L,r_0}(s,t)^*)[C^*(s), L_r^*(s)F^{\frac{1}{2}}(s)]e_j]\| \\ & + \|T_{L,r_0}^*(s,t)[0, (L_r^*(s) - L_{r_0}^*(s))F^{\frac{1}{2}}(s)]e_j\| \to 0, \quad r \to r_0. \end{aligned} \tag{5.14}$$

By the dominated convergence theorem, we obtain

$$\|\mathcal{C}_{t,r} - \mathcal{C}_{t,r_0}\|_{HS} \to 0, \quad r \to r_0.$$

Further, if G is a nuclear operator,

$$\begin{aligned} & \|\Pi_r(t) - \Pi_{r_0}(t)\|_1 \\ \leqslant & \|T_{L,r}^*(b,t) - T_{L,r_0}^*(b,t)\|\|GT_{L,r}(b,t)\|_1 \\ & + \|T_{L,r_0}^*(b,t)G\|_1\|T_{L,r}(b,t) - T_{L,r_0}(b,t)\| \\ & + \|\mathcal{C}_{t,r}^* - C_{t,r_0}^*\|_{HS}\|\mathcal{C}_{t,r}\|_{HS} + \|C_{t,r_0}^*\|_{HS}\|\mathcal{C}_{t,r} - C_{t,r_0}\|_{HS} \to 0, \quad r \to r_0. \end{aligned}$$

By the compactness of Ω^m, the optimal location $\hat{r}$ exists in the nuclear norm. □

5.3 Convergence of optimal control locations

In practice, the integral Riccati equation in an infinite-dimensional space cannot be solved directly. Usually, one approximates and solves it in finite-dimensional spaces by a sequence of approximations from various numerical methods with the convergence in different norms, such as [40], [41], [43] and [81]. Without loss of generality, we let X_n be a family of finite-dimensional subspaces of X and $\mathbf{P}_n$ be the corresponding orthogonal projection of X onto X_n. The finite-dimensional spaces $\{X_n\}$ inherit the norm from X. For every $n \in \mathbb{N}$, let $T_n(\cdot,\cdot)$ be a mild evolution operator on X_n, $B_n(t) \in L_s^\infty(t_0, b; U, X_n)$ and $C_n(t) = C(t)\mathbf{P}_n$, $G_n \in \mathcal{L}(X_n)$. We thus define a sequence of approximations

$$x(t) = T_n(t,t_0)x(t_0) + \int_{t_0}^t T_n(t,s)B_n(s)u(s)ds, \quad t \in [t_0, b]$$

with the cost functional

$$J_n(t,x,u) = \langle x(b), G_n x(b)\rangle + \int_t^b \langle C_n(s)x(s), C_n(s)x(s)\rangle + \langle u(s), F(s)u(s)\rangle ds,$$

where $G_n \in \mathcal{L}(X_n, X_n)$.

From Section 5.1, there exists the unique optimal control trajectory $u(t) = -L_n(t)x(t)$, where $L_n(t) = F(t)^{-1}B_n^*(t)\Pi_n(t)$, $t \in [t_0, b]$, such that the minimum of the cost functional is

$$\begin{aligned} \min_{u \in L_2([t,b];U)} J_n(t, x(t), u) &= J(t, x(t), -L_n(\cdot)x(\cdot)) \\ &= \langle x(t), \mathbf{P}_n\Pi_n(t)\mathbf{P}_n x(t)\rangle, \quad t \in [t_0, b] \end{aligned}$$

and the self-adjoint nonnegative operator $\Pi(t)$ is the unique solution of the integral Riccati equation (IRE)

$$\begin{aligned}\Pi_n(t)x &= T^*_{L_n}(b,t)GT_{L_n}(b,t)x \\ &+ \int_t^b T^*_{L_n}(s,t)[C^*_n(s)C_n(s)+\Pi_n(s)B_n(s)F^{-1}(s)B^*_n(s)\Pi_n(s)]T_{L_n}(s,t)xds,\end{aligned}$$

where

$$\begin{aligned}T_{L_n}(t,\tau)\mathbf{P}_n x &:= T_{-B_nL_n}(t,\tau)\mathbf{P}_n x \\ &= T_n(t,\tau)\mathbf{P}_n x - \int_\tau^t T_{L_n}(t,s)B_n(s)L_n(s)T_n(s,\tau)\mathbf{P}_n x ds,\end{aligned}$$

such that

$$x(t) = T_{L_n}(t,t_0)\mathbf{P}_n x(t_0).$$

In order to guarantee that $\Pi_n(t)$ converges to $\Pi(t)$, the following assumptions are needed in the approximations of the control problem for partial differential equations [43].
(a1) For each $x \in X$,

$$(i)\ T_n(t,s)\mathbf{P}_n x \to T(t,s)x; \quad (ii)\ T^*_n(t,s)\mathbf{P}_n x \to T^*(t,s)x, \quad t_0 \leqslant s \leqslant t \leqslant b$$

and $\sup_n \|T_n(t,s)\| < \infty$, for any $(t,s) \in \Gamma^b_{t_0}$.
(a2) For each $u \in U$, $x \in X$

$$(i)\ B_n(t)u \to B(t)u; \quad (ii)\ B^*_n(t)\mathbf{P}_n x \to B^*(t)x,\ a.e.\ t \in [t_0,b].$$

(a3) For each $x \in X$, $y \in Y$

$$(i)\ C_n(t)\mathbf{P}_n x \to C(t)x; \quad (ii)\ C^*_n(t)y \to C^*(t)y,\ a.e.\ t \in [t_0,b].$$

(a4) For each $x \in X$,

$$G_n\mathbf{P}_n x \to Gx.$$

These assumptions are rather standard and typical for the approximations of partial differential equations. For instance, (a1)(i) and the uniform boundedness in (a1) are analogue to the condition of Trotter-Kato theorem for the approximations of time-invariant systems. The strong convergence of the projection operator $\mathbf{P}_n$ indicates that (a2)(i) and (a3)(i) can be easily satisfied. (a2)(ii) and (a3)(ii) are necessary for the dual problem to consider the optimal observation locations. By the uniform boundedness principle, (a4) implies that $\sup_n \|G_n\| < \infty$.

Before we study the uniform convergence from $\Pi_n(t)$ to $\Pi(t)$, we study under which condition the compactness of $\Pi(t)$ can be guaranteed.

Lemma 5.3.1. *We consider the time-varying system* (5.1) *with the cost functional* (5.2). *If $B(t)$, $C(t)$, $t \in [t_0,b]$ and G are compact operators, then the unique solution $\Pi(t)$ of the integral Riccati equation* (5.4) *is compact.*

Proof. We denote $S = C^*C + \Pi BF^{-1}B^*\Pi$,

$$\Pi(t) = T_L^*(b,t)GT_L(b,t) + \int_t^b T_L^*(s,t)S(s)T_L(s,t)ds.$$

Since $B(t)$, $C(t)$ and G are compact, $T_L^*(b,t)GT_L(b,t)$ and $T_L^*(s,t)S(s)T_L(s,t)$, $(s,t) \in \Gamma_{t_0}^b$ are compact. First, let us only consider the integral part of $\Pi(t)$. It is clear that there exists a set of orthogonal projections $\{\mathbf{P}_n\}$ to some finite-dimensional spaces X_n, $n \in \mathbb{N}$ such that

$$\lim_{n\to\infty} \|\mathbf{P}_n T_L^*(s,t)S(s)T_L(s,t) - T_L^*(s,t)S(s)T_L(s,t)\| = 0.$$

Then, since $T_L(\cdot,\cdot)$ and $S(\cdot)$ are uniformly bounded, it is easy to obtain $\mathbf{P}_n T_L^* S T_L$ is also uniformly bounded in any time and n. By the dominated convergence theorem,

$$\lim_{n\to\infty} \| \int_t^b \mathbf{P}_n T_L^*(s,t)S(s)T_L(s,t)ds - \int_t^b T_L^*(s,t)S(s)T_L(s,t)ds\| = 0.$$

Obviously,

$$\int_t^b \mathbf{P}_n T_L^*(s,t)S(s)T_L(s,t)ds$$

is still finite-rank operator and bounded, so it is compact.

Therefore,

$$\int_t^b T_L^*(s,t)S(s)T_L(s,t)ds$$

is compact. Further, $\Pi(t)$ is compact. □

The following theorem shows the uniform convergence of $\Pi_n(t)$.

Theorem 5.3.2. *For the sequence of approximations under the assumptions* $(a1)-(a4)$*, if* $B(t)$*,* $C(t)$*,* $t \in [t_0, b]$ *and* G *are compact operators and* $\lim_{n\to\infty} \|B_n - \mathbf{P}_n B\|_\infty = 0$*, then*

$$\lim_{n\to\infty} \|\Pi_n(t)\mathbf{P}_n - \Pi(t)\| = 0, \quad t \in [t_0, b].$$

Proof. From $\lim_{n\to\infty} \|B_n - \mathbf{P}_n B\|_\infty = 0$ and $\sup_{t\in[t_0,b]} \|B(t)\| < \infty$, we have

$$\sup_{n\in\mathbb{N}, t\in[t_0,b]} \|B_n(t)\| < \infty.$$

Moreover, because $B(t)$ is compact and $\mathbf{P}_n$ is strongly convergent to the identity operator I,

$$\lim_{n\to\infty} \|\mathbf{P}_n B(t) - B(t)\| = 0, \quad t \in [t_0, b].$$

Further,

$$\begin{aligned} &\|B_n(t) - B(t)\| \\ \leqslant &\|B_n(t) - \mathbf{P}_n B(t)\| + \|\mathbf{P}_n B(t) - B(t)\| \to 0, \quad t \in [t_0, b], \quad n \to \infty. \end{aligned}$$

By the uniform boundedness of $\|T_n(\cdot,\cdot)\|$, $\|C_n\|_\infty$ and $\|G_n\|$ and Theorem 5.2.1, for any $x \in X$, we have

$$\lim_{n\to\infty} \|\Pi_n(t)x - \Pi(t)x\| = 0, \quad t \in [t_0, b].$$

Similar to the proof of the uniform boundedness of Π_r in Theorem 5.2.2, for the approximations with arbitrary feedback control

$$\tilde{u}_n(t) = \tilde{L}_n(t)x(t) = \tilde{L}(t)\mathbf{P}_n x(t), \quad \tilde{L} \in L_s^\infty(t_0, b; X, U),$$

there exists $\lambda_\Pi > 0$ such that $\sup_n \|\Pi_n\|_\infty < \lambda_\Pi$.

To prove the uniform convergence of $\Pi_n(t)$, we define $S_n = C_n^* C_n + \Pi_n B_n F^{-1} B_n^* \Pi_n$ and S with the similar way, then

$$\begin{aligned}
& \|\Pi_n(t)\mathbf{P}_n - \Pi(t)\| \\
\leqslant\ & \|T_{L_n}^*(b,t)G_n T_{L_n}(b,t)\mathbf{P}_n - T_L^*(b,t)GT_L(b,t)\| \\
& + \| \int_t^b T_{L_n}^*(s,t)S_n(s)T_{L_n}(s,t)\mathbf{P}_n - T_L^*(s,t)S(s)T_L(s,t)ds\| \\
\leqslant\ & \|(T_{L_n}^*(b,t) - T_L^*(b,t))G_n\mathbf{P}_n\|\|T_{L_n}(b,t)\mathbf{P}_n\| \\
& + \|T_L^*(b,t)\|\|(G_n\mathbf{P}_n - G)T_{L_n}(b,t)\mathbf{P}_n\| \\
& + \|T_L^*(b,t)\|\|G(T_{L_n}(b,t)\mathbf{P}_n - T_L(b,t))\| \\
& + \int_t^b \|(T_{L_n}^*(s,t) - T_L^*(s,t))S_n(s)\mathbf{P}_n\|\|T_{L_n}(s,t)\mathbf{P}_n\|ds \\
& + \int_t^b \|T_L^*(s,t)\|\|S_n(s)\mathbf{P}_n - S(s)\|\|T_{L_n}(s,t)\mathbf{P}_n\|ds \\
& + \int_t^b \|T_L(s,t)\|\|S(s)(T_{L_n}(s,t)\mathbf{P}_n - T_L(s,t))\|ds.
\end{aligned}$$

As a result of the uniform boundedness of $\|T_n(\cdot,\cdot)\|$, $\|\Pi_n\|_\infty$ and $\|B_n\|_\infty$ in n, $\|L_n\|_\infty$ is uniform bounded and

$$\begin{aligned}
\|L_n^*(t) - \mathbf{P}_n L^*(t)\| \leqslant\ & \|F^{-1}\|_\infty(\|\Pi_n(t)\|\|B_n(t) - \mathbf{P}_n B(t)\| \\
& + \|\Pi_n(t) - \mathbf{P}_n\Pi(t)\|\|B(t)\| \to 0, \quad r \to r_0,
\end{aligned}$$

so $\lim_{n\to\infty} \|L_n(t)\mathbf{P}_n - L(t)\| = 0$ and further

$$\lim_{n\to\infty} \|B_n(t)L_n(t)\mathbf{P}_n - B(t)L(t)\| = 0.$$

According to Lemma 5.2.5 and assumption (a1),

$$\begin{aligned}
&\lim_{n\to\infty} \|T_{L_n}(t,s)\mathbf{P}_n x - T_L(t,s)x\| = 0, \\
&\lim_{n\to\infty} \|T_{L_n}^*(t,s)\mathbf{P}_n x - T_L^*(t,s)x\| = 0, \quad x \in X.
\end{aligned}$$

Because of the compactness of the self-adjoint operator G_n and G, we have

$$\lim_{n\to\infty} \|(T_{L_n}(t,s) - T_L(t,s))G_n\mathbf{P}_n\| = 0,$$
$$\lim_{n\to\infty} \|G(T_{L_n}(t,s)\mathbf{P}_n - T_L(t,s))\| = 0.$$

Meanwhile,

$$\|S_n\|_\infty \leqslant \|C_n^*C_n\|_\infty + \|\Pi_n B_n F^{-1} B_n^*\Pi_n\|_\infty < \infty, \quad n \in \mathbb{N}.$$

Since $C_n = C\mathbf{P}_n$ is compact,

$$\begin{aligned}
& \|S_n(t)\mathbf{P}_n - S(t)\| \\
\leqslant\ & \|C_n^*(t)C_n(t) - C^*(t)C(t)\| + \|L_n^*(t)F(t)L_n(t)\mathbf{P}_n - L(t)F(t)L(t)\| \\
\leqslant\ & \|C^*(t)\mathbf{P}_n - C^*(t)\|\|C_n\|_\infty + \|C^*\|_\infty\|C(t)\mathbf{P}_n - C(t)\| \\
+\ & \|L_n^*(t) - L^*(t)\|\|F\|_\infty\|L_n\|_\infty \\
+\ & \|L^*\|_\infty\|F\|_\infty\|L_n(t)\mathbf{P}_n - L(t)\| \to 0, \quad n \to \infty.
\end{aligned}$$

By the dominated convergence theorem, we obtain $\|\Pi_n(t)\mathbf{P}_n - \Pi(t)\| \to 0, n \to \infty$. □

Next we show that the optimal control locations of approximations converge to the optimal control location of the original system.

Theorem 5.3.3. *Under the assumptions* $(a1) - (a4)$ *and further assume* $B_{r,n} = \mathbf{P}_n B_r$, $r \in \Omega^m$, *if* $B_r(t)$, $C(t)$ *and* G, $t \in [t_0, b]$ *are compact operators and* $\lim_{r\to r_0} \|B_r - B_{r_0}\| = 0$, *then*

$$\hat{\ell}_n(t) \to \hat{\ell}(t), \quad n \to \infty$$

and there exists a subsequence of the optimal locations $\hat{r}_{n_k}$ *depending on* $\hat{\ell}_{n_k}(t)$ *such that*

$$\hat{r}_{n_k} \to \hat{r}, \quad k \to \infty.$$

Proof. From Theorem 5.3.2, we have

$$\lim_{n\to\infty} \|\Pi_{r,n}(t)\mathbf{P}_n - \Pi_r(t)\| = 0, \quad r \in \Omega^m.$$

Since $\lim_{r\to r_0} \|B_r - B_{r_0}\|_\infty = 0$,

$$\|B_{r,n} - B_{r_0,n}\|_\infty \leqslant \|\mathbf{P}_n\|\|B_r - B_{r_0}\|_\infty \to 0, \quad r \to r_0.$$

From Theorem 5.2.2, for any $n \in \mathbb{N}$, there exists $\hat{\ell}_n(t) = \inf_{r\in\Omega^m} \|\Pi_{r,n}(t)\|$.
On one hand,

$$\begin{aligned}
\hat{\ell}_n(t) &= \inf_{r\in\Omega^m} \|\Pi_{r,n}(t)\| \leqslant \|\Pi_{\hat{r},n}(t)\| \leqslant \|\Pi_{\hat{r},n}(t) - \Pi_{\hat{r}}(t)\| + \|\Pi_{\hat{r}}(t)\| \\
&\to \|\Pi_{\hat{r}}(t)\| = \hat{\ell}(t), \quad n \to \infty,
\end{aligned}$$

so $\lim_{n\to\infty} \sup_n \hat{l}_n(t) \leqslant \hat{l}$.

On the other hand, there exists a subsequence $\{\hat{\ell}_{n_k}(t)\}$ such that

$$\lim_{k\to\infty} \hat{\ell}_{n_k}(t) = \lim_{n\to\infty} \inf_n \hat{\ell}_n(t),$$

where

$$\hat{\ell}_{n_k}(t) = \inf_{r\in\Omega^m} \|\Pi_{r,n_k}(t)\| = \|\Pi_{r_{n_k},n_k}(t)\|.$$

Due to the compactness of Ω^m, without loss of the generality, we assume

$$\lim_{k\to\infty} \hat{r}_{n_k} = \bar{r},$$

then we have

$$\|B_{\hat{r}_{n_k},n_k} - B_{\bar{r}}\|_\infty \leqslant \|\mathbf{P}_{n_k}\|\|B_{\hat{r}_{n_k}} - B_{\bar{r}}\|_\infty + \|\mathbf{P}_{n_k}B_{\bar{r}} - B_{\bar{r}}\|_\infty \to 0, \quad k\to\infty$$

and

$$\begin{aligned}&\|\Pi_{\hat{r}_{n_k},n_k}(t) - \Pi_{\bar{r}}(t)\|\\ \leqslant&\|\Pi_{\hat{r}_{n_k},n_k}(t) - \Pi_{r_{n_k}}(t)\| + \|\Pi_{r_{n_k}}(t) - \Pi_{\bar{r}}(t)\| \longrightarrow 0, \quad k\to\infty. \end{aligned} \tag{5.15}$$

Hence,

$$\lim_{n\to\infty} \inf_n \hat{\ell}_n(t) = \lim_{k\to\infty} \hat{\ell}_{n_k}(t) = \lim_{k\to\infty} \|\Pi_{\hat{r}_{n_k},n_k}(t)\| = \|\Pi_{\bar{r}}(t)\| \geqslant \|\Pi_{\hat{r}}(t)\| = \hat{\ell}_r(t),$$

which implies that we obtain

$$\lim_{n\to\infty} \hat{\ell}_n(t) = \hat{\ell}(t). \tag{5.16}$$

Further, $\lim_{n\to\infty} \hat{\ell}_n(t) = \lim_{n\to\infty} \inf_n \hat{\ell}_n(t) = \hat{\ell}(t)$ implies

$$\lim_{k\to\infty} \|\Pi_{\hat{r}_{n_k},n_k}(t)\| = \|\Pi_{\bar{r}}(t)\| = \|\Pi_{\hat{r}}(t)\|.$$

Without loss of the generality, we simply denote $\bar{r}$ by $\hat{r}$. It is clear now that there exists a subsequence $\hat{r}_{n_k}$ of $\hat{r}_n$ converges to the optimal location $\hat{r}$ of the original control problem when $k\to\infty$. □

Associated with Corollary 5.2.4, the following theorem guarantees the uniform convergence of the Riccati operators of approximations to the Riccati operator of the original system in the nuclear norm. We firstly see a lemma about the uniform convergence in the nuclear norm.

Lemma 5.3.4. *G is a nuclear operator in Hilbert space X, T_n strongly converges to T, T_n, $T\in\mathcal{L}(X)$ and are uniformly bounded by λ_T, then*

$$\lim_{n\to\infty} \|(T_n - T)G\|_1 = 0.$$

Proof. Assume $\{e_i\}$ is the orthogonal basis in X and there exist a partial isometry V such that

$$G = V|G|,$$

where $|G| = (G^*G)^{\frac{1}{2}}$, then,

$$\|(T_n - T)V|G|^{\frac{1}{2}}e_i\| \leqslant \|T_n - T\|\|V|G|^{\frac{1}{2}}e_i\| \leqslant 2\lambda_T\|V|G|^{\frac{1}{2}}e_i\|$$

and because of the strong convergence of T_n to T,

$$\lim_{n\to\infty} \|(T_n - T)V|G|^{\frac{1}{2}}e_i\| = 0.$$

Since G is a nuclear operator, then $|G|^{\frac{1}{2}}$ is a Hilbert-Schmidt operator, so

$$\sum_{i=1}^{\infty} \|(T_n - T)V|G|^{\frac{1}{2}}e_i\| = 2M_T \sum_{i=1}^{\infty} \|V|G|^{\frac{1}{2}}e_i\| < \infty.$$

By the dominated convergence theorem,

$$\begin{aligned} \lim_{n\to\infty} \|(T_n - T)V|G|^{\frac{1}{2}}\|_{HS} &= \lim_{n\to\infty} \sum_{i=1}^{\infty} \|(T_n - T)V|G|^{\frac{1}{2}}e_i\| \\ &= \sum_{i=1}^{\infty} \lim_{n\to\infty} \|(T_n - T)V|G|^{\frac{1}{2}}e_i\| \\ &= 0. \end{aligned}$$

Then we obtain

$$\|(T_n - T)G\|_1 \leqslant \|(T_n - T)V|G|^{\frac{1}{2}}\|_{HS}\||G|^{\frac{1}{2}}\|_{HS} \to 0, \ \ n \to \infty.$$

□

With the different proof, from [46, Chapter III, Theorem 6.3], we can obtain the same conclusion as Lemma 5.3.4.

Theorem 5.3.5. *For the sequence of approximations under the assumptions* $(a1) - (a4)$*, if* U *and* Y *are finite dimensional,* $\lim_{n\to\infty} \|B_n - \mathbf{P}_n B\|_\infty = 0$*,* G *is nuclear operator and* $\lim_{n\to\infty} \|G_n\mathbf{P}_n - G\|_1 = 0$*, then*

$$\lim_{n\to\infty} \|\Pi_n(t)\mathbf{P}_n - \Pi(t)\|_1 = 0.$$

Proof. We define $\mathcal{C}_t$ in the same way with Corollary 5.2.4 and similarly define $\mathcal{C}_{t,n} : X_n \to L_2([t, b]; U \times Y)$ satisfying

$$(\mathcal{C}_{t,n}x(t))(\cdot) = \begin{pmatrix} C_n(\cdot) \\ F^{\frac{1}{2}}(\cdot)L_n(\cdot) \end{pmatrix} T_{L_n}(\cdot, t)x(t),$$

where $L_n = F^{-1}B_n^*\Pi_n$. then, from Theorem 5.2.3 (1), $\mathcal{C}_{t,n}$ is Hilbert-Schmidt and

$$\Pi_n(t) = T_{L_n}^*(b,t)G_nT_{L_n}(b,t) + \mathcal{C}_{t,n}^*\mathcal{C}_{t,n}.$$

From Theorem 5.2.3 *(1)*, $\mathcal{C}_{t,n}$ is Hilbert-Schmidt and further

$$\Pi_n(t) = T_{L_n}^*(b,t)G_nT_{L_n}(b,t) + \mathcal{C}_{t,n}^*\mathcal{C}_{t,n}$$

is nuclear. The same with Theorem 5.3.2, we also have the uniform boundedness of $\|T_n(\cdot,\cdot)\|$, $\|\Pi_n\|_\infty$ $\|B_n\|_\infty$, $\|L_n\|_\infty$ in n and

$$\begin{aligned}\lim_{n\to\infty}\|L_n(t)\mathbf{P}_n - L(t)\| = 0,\\ \lim_{n\to\infty}\|T_{L_n}(t,s)\mathbf{P}_nx - T_L(t,s)x\| = 0.\end{aligned}$$

Hence, similar to Theorem 5.2.7, for any $(s,t)\in\Gamma_{t_0}^b$, we have

$$\begin{aligned}&\|\mathcal{C}_{t,n} - \mathcal{C}_t\|_{HS}\\ \leqslant\ &\sum_{j=1}^{p+q}\int_t^b \|(T_{L_n}^*(s,t) - T_L(s,t)^*)[C_n^*(s), L_n^*(s)F^{\frac{1}{2}}(s)]e_j\|\\ +\ &\|T_L^*(s,t)[C_n(s) - C(s), (L_n^*(s) - L^*(s))F^{\frac{1}{2}}(s)]e_j\|ds \to 0, \quad r\to r_0.\end{aligned}$$

Then, since G is nuclear operator with $\lim_{n\to\infty}\|G_n\mathbf{P}_n - G\|_1 = 0$, by Lemma 5.3.4, we obtain

$$\begin{aligned}&\|\Pi_n(t)\mathbf{P}_n - \Pi(t)\|_1\\ \leqslant\ &\|T_{L_n}^*(b,t)\|\|G_n(T_{L_n}(b,t)\mathbf{P}_n - \mathbf{P}_nT_L(b,t))\|_1\\ +\ &\|T_{L_n}^*(b,t)\|\|G_n\mathbf{P}_n - G\|_1\|T_L(b,t)\|\\ +\ &\|(T_{L_n}^*(b,t) - T_L^*(b,t))G\|_1\|T_L(b,t)\| + \|\mathcal{C}_{t,n}^* - C_t^*\|_{HS}\|\mathcal{C}_{t,n}\|_{HS}\\ +\ &\|C_t^*\|_{HS}\|\mathcal{C}_{t,n} - C_t\|_{HS} \to 0, \quad n\to\infty.\end{aligned}$$

□

Theorem 5.3.6. *Under the assumptions* $(a1)-(a4)$ *and further assume* $B_{r,n} = \mathbf{P}_nB_r$, $r\in\Omega^m$, *if the input space* U *and the output space* Y *are finite dimensional,* $\lim_{r\to r_0}\|B_r - B_{r_0}\| = 0$, G *is nuclear operator and* $\lim_{n\to\infty}\|G_nP_n - G\|_1 = 0$, *then*

$$\hat{\ell}_{1,n}(t) \to \hat{\ell}_1(t), \quad n\to\infty$$

and there exists a subsequence of the optimal locations $\hat{r}_{n_k}$ *depending on* $\hat{\ell}_{1,n_k}(t)$ *such that*

$$\hat{r}_{n_k} \to \hat{r}, \quad k\to\infty.$$

Proof. From Theorem 5.2.7 and Theorem 5.3.5, we have

$$\lim_{r\to r_0} \|\Pi_r(t) - \Pi_{r_0}(t)\|_1 = 0,$$
$$\lim_{n\to\infty} \|\Pi_{r,n}(t)\mathbf{P}_n - \Pi_r(t)\|_1 = 0.$$

The same with Theorem 5.3.3, we have $\hat{\ell}_{1,n}(t) \leqslant \|\Pi_{\hat{r}}(t)\|_1 = \hat{\ell}_1(t)$, $n \to \infty$. Besides, there exists a subsequence $\{\hat{\ell}_{1,n_k}(t)\}$ such that

$$\begin{aligned}\lim_{n\to\infty} \inf_n \hat{\ell}_{1,n}(t) &= \lim_{k\to\infty} \hat{\ell}_{1,n_k}(t) = \lim_{k\to\infty} \|\Pi_{\hat{r}_{n_k},n_k}(t)\|_1 = \|\Pi_{\bar{r}}(t)\|_1 \\ &\geqslant \|\Pi_{\hat{r}}(t)\|_1 = \hat{\ell}_1(t).\end{aligned}$$

Therefore,

$$\lim_{n\to\infty} \hat{\ell}_{1,n}(t) = \hat{\ell}_1(t).$$

Since $\lim_{n\to\infty} \hat{\ell}_{1,n}(t) = \lim_{n\to\infty} \inf_n \hat{\ell}_{1,n}(t) = \hat{\ell}(t)$ and we denote $\bar{r}$ by $\hat{r}$, there exists $\lim_{k\to\infty} \hat{r}_{n_k} = \hat{r}$ such that $\lim_{k\to\infty} \|\Pi_{\hat{r}_{n_k},n_k}(t)\|_1 = \|\Pi_{\hat{r}}(t)\|_1$. □

Chapter 6

Optimal Observation Locations for Time-Varying Systems in Hilbert Spaces on a Finite-Time Horizon

In this chapter we develop the optimal problems of observation locations based on the Kalman filter and smoother in Hilbert spaces. The issue of observations is also of great importance of many estimation problems for stochastic systems, such as weather forecasting and data assimilation problems in meteorology. For this kind of problems, observations always have low temporal and spatial density. The lack of observations is a major barrier of preventing the improvement of estimations and leading to the inaccuracy of predictions. On one hand, based on the insufficient observations, many works make efforts to improve approaches of estimations in recent years. On the other hand, one possibility to improve the predictive or estimation skill for specific problems is to target the locations of observations which can potentially result in the largest forecast improvement in order to make observations more efficient. The better choice of locations of observations can help making more progress of the predictive or estimation skills. In contrast, improper observations probably make no sense to the accuracy of predictions and lead to the waste of resources by optimizing the improper parameters. Motivated by problems of data assimilation in meteorology, we estimate unknown random variables by the Kalman filter and smoother, which has been theoretical foundation of one of the most popular data assimilation approaches in last decades. It provides us an opportunity to define and search for optimal locations of observations by minimizing the proper norm of the covariance.

We firstly state the main results of the Kalman filter and smoother of time-varying systems in the integral form in Hilbert spaces. By the duality between the Kalman filter and linear-quadratic optimal control problem, under certain conditions, the nuclearity of the covariances from the Kalman filter and further from the Kalman smoother can be guaranteed. At the same time, the existence of the minimal cost and optimal observation locations of the estimation of the model state for stochastic systems will be shown. By a sequence of approximations of the original system, we obtain the convergence of the sequence of minimal costs and a subsequence of optimal observation locations based on the Kalman filter and smoother. Finally, we apply the obtained results to a three-

dimensional advection-diffusion model extended by emission rates in Section 6.4. In this example, the operator splitting technique with spatial and temporal discretization is applied to simulate the practical application in meteorology.

6.1 Kalman filter in Hilbert spaces

There are several works, for instance [16], [19], [20], [38], [55], focusing on filtering and smoothing problems in Hilbert spaces from different perspectives. In this section without the differentiability of evolution operators, we study the Kalman filter of time-varying systems in real separable Hilbert spaces driven by white noises.

We always let $\mathcal{X}$, $\mathcal{E}$ and $\mathcal{Y}$ be real separable Hilbert spaces. Assuming μ is a Borel measure in $\mathcal{X}$, we firstly give some basic concepts of probability theory in Hilbert spaces.

Definition 6.1.1. *μ is a* totally finite measure *on $\mathcal{X}$ if for any x, $\int_\Omega \|x\| d\mu < \infty$. Further, if there exists $\bar{x} \in \mathcal{X}$ such that*

$$\langle \bar{x}, h\rangle = \mathbb{E}\langle x, h\rangle = \int_{\mathcal{X}} \langle x, h\rangle \mu(dx), \quad \forall h \in \mathcal{X},$$

then $\bar{x}$ is called the mean *or* expectation *of x and denoted by $\mathbb{E}x$.*

Definition 6.1.2. *The* covariance operator *P of x in $\mathcal{X}$, also denoted by cov(x), if it exists, is given by*

$$\langle Ph_1, h_2\rangle = \langle h_1, Ph_2\rangle = \int_{\mathcal{X}} \langle x - \mathbb{E}x, h_1\rangle\langle x - \mathbb{E}x, h_2\rangle \mu(dx), \quad \forall h_1, h_2 \in \mathcal{X}.$$

Definition 6.1.3. *The random variables x, y whose expectations exist are* independent *if* $\mathbb{E}(\langle x, y\rangle) = \langle \mathbb{E}(x), \mathbb{E}(y)\rangle$.

Definition 6.1.4. *If for any $x \in \mathcal{X}$, the random variable $\langle x, \cdot\rangle$ has a Gaussian distribution, then μ is called a* Gaussian measure*. Further, we denote x of the Gaussian measure with mean $\bar{x}$ and covariance P by $x \sim \mathcal{N}(\bar{x}, P)$.*

It can be concluded that the covariances of Gaussian measures must be nuclear operator. In order to deal with the case that the covariances of Gaussian random variables are not of nuclearity, we need the following definitions, see [2] and [3].

Definition 6.1.5. *Let B be a Borel set of the subspace $\mathcal{X}_n$ of $\mathcal{X}$ generated by $x_i \in \mathcal{X}, i = 1, \cdots, n$ and B_n be a Borel set of $\mathbb{R}^n$ isomorphic to B.* A cylinder set *of $\mathcal{X}$ with the base B means*

$$\{x|([x, x_1], \cdots, [x, x_n]) \in B_n\}.$$

Definition 6.1.6. *Let C be a cylinder set with base B in $\mathcal{X}_n$.* The cylinder measure *μ is defined by*

$$\mu(C) = \nu_n(B), \tag{6.1}$$

where ν_n is a countably additive probability measure on the $\sigma-$ algebra of $\mathcal{X}_n$.

Definition 6.1.7. *The cylinder measure μ on $\mathcal{X}$ is called* Gauss measure *if its characteristic function is given by*

$$\int e^{i\langle x,f\rangle}d\mu = e^{-\frac{1}{2}\|f\|}, \quad x,\ f \in \mathcal{X}.$$

Definition 6.1.8. *Let μ denote the Gauss measure on $L^2([0,T],\mathcal{E})$. The process $\omega(t)$, $\omega \in L^2([0,T],\mathcal{E})$ is called* white noise.

It is worth noting that, for any $f \in L^2([0,T],\mathcal{E})$, ω has the obvious properties [2]:

1. $\mathbb{E}\langle\omega, f\rangle = 0$,
2. $\int_0^T \langle f(t), w(t)\rangle dt$ defines a Gaussian random variable with the variance $\|f\|^2$,
3. $\mathrm{cov}(\omega) = I$.

We consider the integral form of time-varying systems in Hilbert spaces given by

$$x(t) = M(t,t_0)x(t_0) + \int_{t_0}^{t} M(t,s)[B(s)u(s) + D(s)\omega(s)]ds, \quad (t,t_0) \in \Gamma_{t_0}^b, \quad (6.2)$$

where $M(\cdot,\cdot)$ is a mild evolution operator on $\mathcal{X}$. $x(t)$ and $\omega(t)$ are random variables with values in $\mathcal{X}$ and $\mathcal{E}$, respectively and $\omega(t)$ is the white noise with Definition 6.1.8. Further, we assume $u \in L^2(t_0,b;U)$, $B \in L_s^\infty(t_0,b;U,\mathcal{X})$, $B^* \in L_s^\infty(t_0,b;\mathcal{X},U)$, $D \in L_s^\infty(t_0,b;\mathcal{E},\mathcal{X})$.

The observation system of the time-varying system above is given by

$$y(t) = H(t)x(t) + E(t)\nu(t), \quad t \in [t_0,b], \quad (6.3)$$

where $H \in L_s^\infty(t_0,b;\mathcal{X},\mathcal{Y})$, $E \in L_s^\infty(t_0,b;\mathcal{E},\mathcal{Y})$, $y(t)$ and $\nu(t)$ are random variables with values in $\mathcal{Y}$ and $\mathcal{E}$, respectively. Besides, $\nu(t)$ is the white noise with Definition 6.1.8 and $R(t) := E(t)E^*(t)$ is coercive.

We still assume $\hat{x}(t_0|t_{-1}) = \mathbb{E}(x(t_0))$, $P(t_0|t_{-1}) = \mathrm{cov}(x(t_0) - \hat{x}(t_0|t_{-1}))$, $\tilde{x}(t|t) = x(t) - \hat{x}(t|t)$, $P(t|t) = \mathrm{cov}(\tilde{x}(t))$ and $Y_t = \{y(s), t_0 \leqslant s \leqslant t \leqslant b\}$.

The filter problem is to find the best linear unbiased estimate of the state $x(t)$ based on the observations Y_t, given by $\hat{x}(t|t) = \mathbb{E}(x(t)|Y_t)$, which has the form

$$\begin{aligned}\hat{x}(t|t) &= M(t,t_0)\hat{x}(t_0|t_{-1}) \\ &\quad + \int_{t_0}^{t} M(t,s)B(s)u(s)ds + \int_{t_0}^{t} K_f(t,s)[y(s) - H(s)\hat{x}(s|s)]ds\end{aligned}$$

and minimizes the nuclear norm of $P(t|t)$ if it exists. Here $K_f(\cdot,\cdot)$ is unknown.

Theorem 6.1.9. *For the time-varying system* (6.2) *with the observation system* (6.3)*, the linear unbiased estimation of the filter problem $\hat{x}(t|t)$ of $x(t)$ is optimal if the linear gain operator is given by*

$$K_f(t,\tau) = M(t,\tau)P(\tau|\tau)H^*(\tau)R^{-1}(\tau), \quad \tau \leqslant t. \quad (6.4)$$

Proof. By the Wiener-Hopf equation [38], [56], $\hat{x}(t|t)$ is the optimal linear unbiased estimation if and only if $\mathbb{E}\langle\tilde{x}(t|t), h_1\rangle\langle y(\tau), h_2\rangle = 0, \quad \tau \leqslant t,\ h_1,\ h_2 \in \mathcal{X}$. Further, according to [20, Corollary 6.3], $\mathbb{E}\langle\hat{x}(t|t), h_1\rangle\langle\tilde{x}(t|t), h_2\rangle = 0$. Hence,

$$
\begin{aligned}
&\mathbb{E}\langle\tilde{x}(t|t), h_1\rangle\langle y(\tau), h_2\rangle\\
=&\mathbb{E}\langle M(t,\tau)\tilde{x}(\tau|\tau) - \int_\tau^t K_f(t,s)[H(s)\tilde{x}(s|s) + E(s)\nu(s)]ds, h_1\rangle\langle y(\tau), h_2\rangle\\
=&\mathbb{E}\langle M(t,\tau)\tilde{x}(\tau|\tau), h_1\rangle\langle H(\tau)x(\tau), h_2\rangle\\
&- \mathbb{E}\langle\int_\tau^t K_f(t,s)[H(s)\tilde{x}(s|s) + E(s)\nu(s)]ds, h_1\rangle\langle y(\tau), h_2\rangle\\
=&\mathbb{E}\langle M(t,\tau)\tilde{x}(\tau|\tau), h_1\rangle\langle H(\tau)\tilde{x}(\tau|\tau), h_2\rangle - \mathbb{E}\langle\int_\tau^t K_f(t,s)E(s)\nu(s)ds, h_1\rangle\langle y(\tau), h_2\rangle\\
=&\langle h_1, M(t,\tau)P(\tau|\tau)H^*(\tau)h_2\rangle - \mathbb{E}\langle\int_\tau^t K_f(t,s)E(s)\nu(s)ds, h_1\rangle\langle E(\tau)\nu(\tau), h_2\rangle\\
=&\langle h_1, M(t,\tau)P(\tau|\tau)H^*(\tau)h_2\rangle\\
&- \mathbb{E}\langle\int_\tau^t K_f(t,s)E(s)\nu(s)ds, h_1\rangle\langle\int_\tau^t E(s)\nu(s)\delta(s-\tau)ds, h_2\rangle\\
=&\langle h_1, M(t,\tau)P(\tau|\tau)H^*(\tau)h_2\rangle - \langle h_1, K_f(t,\tau)R(\tau)h_2\rangle.
\end{aligned}
$$

Therefore, $K_f(t,\tau)R(\tau) = M(t,\tau)P(\tau|\tau)H^*(\tau)$. Since $R(t)$ is coercive, we obtain

$$K_f(t,\tau) = M(t,\tau)P(\tau|\tau)H^*(\tau)R^{-1}(\tau), \quad \tau < t.$$

If $t = \tau$, by the strong continuity of $K_f(t,\cdot)$, $K_f(t,t) = P(t|t)H^*(t)R^{-1}(t)$. □

We define

$$K(t) := K_f(t,t) = P(t|t)H^*(t)R^{-1}(t),$$

Theorem 6.1.9 implies that

$$
\begin{aligned}
\tilde{x}(t|t) \quad = \quad & M(t,t_0)\tilde{x}(t_0|t_{-1}) - \int_{t_0}^t M(t,s)K(s)H(s)\tilde{x}(s|s)ds\\
&+ \int_{t_0}^t M(t,s)[D(s)\omega(s) - K(s)E(s)\nu(s)]ds. \qquad (6.5)
\end{aligned}
$$

Theorem 6.1.10. *Equation* (6.5) *is equivalent to*

$$\tilde{x}(t|t) = M_K(t,t_0)\tilde{x}(t_0|t_{-1}) + \int_{t_0}^t M_K(t,s)\left(D(s)\omega(s) - K(s)E(s)\nu(s)\right)ds, \quad (6.6)$$

where $M_K(t,\tau)x = M(t,\tau)x - \int_\tau^t M_K(t,s)K(s)H(s)M(s,\tau)xds$, $(t,\tau) \in \Gamma_{t_0}^b$

Proof. From (6.5), we have

$$\begin{aligned}
&\tilde{x}(t|t)\\
=&M_K(t,t_0)\tilde{x}(t_0|t_{-1}) + \int_{t_0}^{t} M_K(t,s)K(s)H(s)M(s,t_0)\tilde{x}(t_0|t_{-1})ds\\
&- \int_{t_0}^{t} M_K(t,s)K(s)H(s)\tilde{x}(s|s)ds + \int_{t_0}^{t} M_K(t,s)[D(s)\omega(s) - K(s)E(s)\nu(s)]ds\\
&- \int_{t_0}^{t}\int_{s}^{t} M_K(t,\eta)K(\eta)H(\eta)M(\eta,s)K(s)H(s)\tilde{x}(s|s)d\eta ds\\
&+ \int_{t_0}^{t}\int_{s}^{t} M_K(t,\eta)K(\eta)H(\eta)M(\eta,s)\left(D(s)\omega(s) - K(s)E(s)\nu(s)\right)d\eta ds\\
=&M_K(t,t_0)\tilde{x}(t_0|t_{-1}) + \int_{t_0}^{t} M_K(t,s)\left(D(s)\omega(s) - K(s)E(s)\nu(s)\right)ds\\
&- \int_{t_0}^{t} M_K(t,s)K(s)H(s)\tilde{x}(s|s)ds + \int_{t_0}^{t} M_K(t,s)K(s)H(s)M(s,t_0)\tilde{x}(t_0|t_{-1})ds\\
&- \int_{t_0}^{t} M_K(t,s)K(s)H(s)\int_{t_0}^{s} M(s,\eta)K(\eta)H(\eta)\tilde{x}(\eta|\eta)d\eta ds\\
&+ \int_{t_0}^{t} M_K(t,s)K(s)H(s)\int_{t_0}^{s} M(s,\eta)\left(D(\eta)\omega(\eta) - K(\eta)E(\eta)\nu(\eta)\right)d\eta ds\\
=&M_K(t,t_0)\tilde{x}(t_0|t_{-1}) + \int_{t_0}^{t} M_K(t,s)\left(D(s)\omega(s) - K(s)E(s)\nu(s)\right)ds
\end{aligned}$$

□

For finite-dimensional systems, the trace of the covariance of $\tilde{x}(t|t)$ is considered as an evaluation of the estimation errors. For time-varying systems in Hilbert spaces, we also consider the nuclear norm of the covariance of $\tilde{x}(t|t)$. Defining $Q(t) := D(t)D^*(t)$, we obtain the following theorem.

Theorem 6.1.11. *The covariance (if exists) of $\tilde{x}(t|t)$ satisfies the IRE*

$$\begin{aligned}
P(t|t) =&M_K(t,t_0)P(t_0|t_{-1})M_K^*(t,t_0)\\
&+ \int_{t_0}^{t} M_K(t,s)\left[Q(s) + P(s|s)H^*(s)R^{-1}(s)H(s)P(s|s)\right]M_K^*(t,s)ds. \quad (6.7)
\end{aligned}$$

Proof. For $\tilde{x}(t|t)$ in (6.6), we assume its covariance $P(t|t)$ exists and define $\mathcal{Q}_t\colon L^2(t_0,t;\mathcal{E}\times\mathcal{E})\to\mathcal{X}$ by

$$\mathcal{Q}_t\begin{pmatrix}\omega\\ \nu\end{pmatrix} = \int_{t_0}^{t}[M_K(t,s)D(s), -M_K(t,s)K(s)E(s)]\begin{pmatrix}\omega(s)\\ \nu(s)\end{pmatrix}ds.$$

Its adjoint operator $\mathcal{Q}_t^*\colon \mathcal{X}\to L^2(t_0,t;\mathcal{E}\times\mathcal{E})$ is given by

$$\mathcal{Q}_t^* x = \begin{pmatrix} D^*(\cdot)M_K^*(t,\cdot)\\ -E^*(\cdot)K^*(\cdot)M_K^*(t,\cdot)\end{pmatrix}x, \quad x\in\mathcal{X}.$$

We obtain

$$\begin{aligned}
& \mathbb{E}\langle\tilde{x}(t|t), h_1\rangle\langle\tilde{x}(t|t), h_2\rangle \\
= \; & \mathbb{E}\langle M_K(t,t_0)\tilde{x}(t_0|t_{-1}), h_1\rangle\langle M_K(t,t_0)\tilde{x}(t_0|t_{-1}), h_2\rangle \\
& + \mathbb{E}\langle \mathcal{Q}_t \begin{pmatrix} \omega \\ \nu \end{pmatrix}, h_1\rangle\langle \mathcal{Q}_t \begin{pmatrix} \omega \\ \nu \end{pmatrix}, h_2\rangle \\
= \; & \langle M_K(t,t_0)P(t_0|t_{-1})M_K^*(t,t_0)h_1, h_2\rangle + \langle \mathcal{Q}_t \text{cov}\left(\begin{pmatrix} \omega \\ \nu \end{pmatrix}\right)\mathcal{Q}_t^* h_1, h_2\rangle \\
= \; & \langle M_K(t,t_0)P(t_0|t_{-1})M_K^*(t,t_0)h_1, h_2\rangle \\
& + \langle \int_{t_0}^{t} M_K(t,s)\left[Q(s) + K(s)R(s)K^*(s)\right] M_K^*(t,s)h_1 ds, h_2\rangle.
\end{aligned}$$

Hence, for any $x \in \mathcal{X}$,

$$\begin{aligned}
P(t|t)x \; = \; & M_K(t,t_0)P(t_0|t_{-1})M_K^*(t,t_0)x \\
& + \int_{t_0}^{t} M_K(t,s)\left[Q(s) + K(s)R(s)K^*(s)\right] M_K^*(t,s)x ds \\
= \; & M_K(t,t_0)P(t_0|t_{-1})M_K^*(t,t_0)x \\
& + \int_{t_0}^{t} M_K(t,s)\left[Q(s) + P(s|s)H^*(s)R^{-1}(s)H(s)P(s|s)\right] M_K^*(t,s)x ds.
\end{aligned}$$

□

A comparison to the main results of the linear-quadratic optimal control problem in Section 5.2 yields: The covariance of $\tilde{x}(t|t)$ of the time-varying system (6.2) with the observations (6.3) equals to the Riccati operator $\Pi(b-t)$ in (5.4) corresponding to the time-varying system

$$x(t) = T(t,t_0)x(t_0) + \int_{t_0}^{t} T(t,s)B(s)u(s)ds$$

with the cost functional

$$J(t,x,u) = \langle x(b), Gx(b)\rangle + \int_t^b \langle C(s)x(s), C(s)x(s)\rangle + \langle u(s), F(s)u(s)\rangle ds,$$

where $T(t,s) = M^*(b-s, b-t)$, $B(s) = H^*(b-s)$, $G = P(t_0|t_{-1})$, $C(s) = D^*(b-s)$, $F(s) = R(b-s)$, $(t,s) \in \Gamma_{t_0}^b$.

By the duality between the linear quadratic control problem and the Kalman filter, Corollary 5.2.4 implies the following condition to guarantee the existence and nuclearity of $P(t|t)$.

Theorem 6.1.12. *For the time-varying system* (6.2) *with the observation system* (6.3)*, if $\mathcal{E}$ and $\mathcal{Y}$ are finite dimensional and $P(t_0|t_{-1})$ is a nuclear operator, then the covariance of $\tilde{x}(t|t)$ based on Y_t satisfying* (6.7) *exists and is a nuclear operator.*

6.2 Kalman smoother in Hilbert spaces

We study the optimal linear unbiased estimation of $x(\tau)$ based on Y_t given by $\hat{x}(\tau|t) = \mathbb{E}(x(\tau)|Y_t)$, $\tau \leqslant t$. We still constrain the linear estimation of $x(\tau|t)$ has the form

$$\hat{x}(\tau|t) = \int_{t_0}^{t} K_s(s,\tau)[y(s) - H(s)\hat{x}(s|s)]ds, \quad \tau \leqslant t, \tag{6.8}$$

where $K_s(\cdot,\cdot)$ is an unknown linear operator.

Since in the case of $\tau = t$, (6.8) with the minimal covariance in the nuclear norm is equivalent to the optimal linear unbiased estimation based on the Kalman filter, in order to determine the optimal estimation of $\hat{x}(\tau|t)$, $\tau \leqslant t$, we can rewrite (6.8) as

$$\hat{x}(\tau|t) = \hat{x}(\tau|\tau) + \int_{\tau}^{t} K_s(s,\tau)[y(s) - H(s)\hat{x}(s|s)]ds. \tag{6.9}$$

Theorem 6.2.1. *For the time-varying system* (6.2) *with the observation system* (6.3)*, the linear unbiased estimation of the filter problem $\hat{x}(\tau|t)$ of $x(\tau)$ is optimal if $K_s(\cdot,\cdot)$ in* (6.9) *is given by*

$$K_s(\eta,\tau) = P(\tau|\tau)M_K^*(\eta,\tau)H^*(\eta)R^{-1}(\eta), \quad \tau \leqslant \eta \leqslant t.$$

Proof. By the Wiener-Hopf equation [56], [38], $\mathbb{E}\langle\tilde{x}(\tau|t), h_1\rangle\langle y(\eta), h_2\rangle = 0$, $h_1 \in \mathcal{X}, h_2 \in \mathcal{Y}$, for any $\eta < t$. In order to derive of Kalman smoother, we only need to consider $\mathbb{E}\langle\tilde{x}(\tau|t), h_1\rangle\langle y(\eta), h_2\rangle = 0$, $\tau \leqslant \eta < t$. Then we have

$$\begin{aligned}
& \mathbb{E}\langle\tilde{x}(\tau|t), h_1\rangle\langle y(\eta), h_2\rangle \\
= \;& \mathbb{E}\langle\tilde{x}(\tau|\eta) - \int_{\eta}^{t} K_s(s,\eta)[y(s) - H(s)\hat{x}(s|s)]ds, h_1\rangle\langle y(\eta), h_2\rangle \\
= \;& \mathbb{E}\langle\tilde{x}(\tau|\eta), h_1\rangle\langle H(\eta)\tilde{x}(\eta|\eta), h_2\rangle \\
& -\mathbb{E}\langle\int_{\eta}^{t} K_s(s,\eta)[H(s)\hat{x}(s|s) + E(s)\nu(s)]ds, h_1\rangle\langle y(\eta), h_2\rangle \\
= \;& \mathbb{E}\langle\tilde{x}(\tau|\tau) - \int_{\tau}^{\eta} K_s(\tau,s)[y(s) - H(s)\hat{x}(s|s)]ds, h_1\rangle\langle H(\eta)\tilde{x}(\eta|\eta), h_2\rangle \\
& -\mathbb{E}\langle\int_{\eta}^{t} K_s(s,\eta)E(s)\nu(s)ds, h_1\rangle\langle E(\eta)\nu(\eta), h_2\rangle \\
= \;& \mathbb{E}\langle\tilde{x}(\tau|\tau), h_1\rangle\langle H(\eta)M_K(\eta,\tau)H^*(\eta)\tilde{x}(\tau|\tau), h_2\rangle \\
& -\mathbb{E}\langle\int_{\eta}^{t} K_s(s,\eta)E(s)\nu(s)ds, h_1\rangle\langle\int_{\eta}^{t} E(s)\nu(s)\delta(s-\eta)ds, h_2\rangle \\
= \;& \langle h_1, P(\tau|\tau)M_K^*(\eta,\tau)H^*(\eta)h_2\rangle - \langle h_1, K_s(\tau,\eta)R(\eta)h_2\rangle.
\end{aligned}$$

By the coercivity of $R(t)$, we obtain $K_s(\tau,\eta) = P(\tau|\tau)M_K^*(\eta,\tau)H^*(\eta)R^{-1}(\eta)$. □

We define $\tilde{x}(\tau|t) = x(\tau) - \hat{x}(\tau|t)$, Theorem 6.2.1 implies

$$\tilde{x}(\tau|t) = \tilde{x}(\tau|\tau) - P(\tau|\tau)\int_{\tau}^{t} M_K^*(s,\tau)H^*(s)R^{-1}(s)[y(s) - H(s)\hat{x}(s|s)]ds.$$

Thus, its covariance can be derived.

Theorem 6.2.2. *The covariance (if exists) of* $\tilde{x}(\tau|t)$, $(t,\tau) \in \Gamma_{t_0}^b$ *is*

$$\begin{aligned} &P(\tau|t)x = P(\tau|\tau)x \\ &- P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s)M_K(s,\tau)P(\tau|\tau)xds, \quad x \in \mathcal{X}. \end{aligned} \tag{6.10}$$

Proof. Denoting the covariance of $\tilde{x}(\tau|t)$ by $P(\tau|t)$, we obtain

$$\begin{aligned} &\langle h_1, P(\tau|t)h_2\rangle \\ =&\mathbb{E}\langle \tilde{x}(\tau|t), h_1\rangle\langle \tilde{x}(\tau|t), h_2\rangle \\ =&\mathbb{E}\langle \tilde{x}(\tau|\tau) - P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)[y(s) - H(s)\hat{x}(s)]ds, h_1\rangle\langle \tilde{x}(\tau|t), h_2\rangle \\ =&\mathbb{E}\langle \tilde{x}(\tau|\tau), h_1\rangle\langle \tilde{x}(\tau|\tau) \\ &- P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)[y(s) - H(s)\hat{x}(s|s)]ds, h_2\rangle \\ =&\mathbb{E}\langle \tilde{x}(\tau|\tau), h_1\rangle\langle \tilde{x}(\tau|\tau), h_2\rangle \\ &- \mathbb{E}\langle \tilde{x}(\tau|\tau), h_1\rangle\langle P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s)\tilde{x}(s|s)ds, h_2\rangle \\ =&\langle h_1, P(\tau|\tau)h_2\rangle - E\langle \tilde{x}(\tau|\tau), h_1\rangle\langle P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s) \\ &\cdot\left(M_K(s,\tau)\tilde{x}(\tau|\tau) + \int_\tau^s M_K(s,\eta)\left(D(\eta)\omega(\eta) - K(\eta)E(\eta)\nu(\eta)\right)d\eta\right)ds, h_2\rangle \\ =&\langle h_1, P(\tau|\tau)h_2\rangle \\ &- \mathbb{E}\langle \tilde{x}(\tau|\tau), h_1\rangle\langle P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s)M_K(s,\tau)\tilde{x}(\tau|\tau)ds, h_2\rangle \\ =&\langle h_1, P(\tau|\tau)h_2\rangle \\ &- \langle h_1, P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s)M_K(s,\tau)P(\tau|\tau)h_2ds\rangle. \end{aligned}$$

Hence, for any $x \in \mathcal{X}$, $(t,\tau) \in \Gamma_{t_0}^b$, we have

$$P(\tau|t)x = P(\tau|\tau)x - P(\tau|\tau)\int_\tau^t M_K^*(s,\tau)H^*(s)R^{-1}(s)H(s)M_K(s,\tau)P(\tau|\tau)xds.$$

□

Theorem 6.2.3. *For the time-varying system* (6.2) *with the observation system* (6.3)*, if* $\mathcal{E}$ *and* $\mathcal{Y}$ *are finite dimensional and* $P(t_0|t_{-1})$ *is a nuclear operator, then* $P(\tau|t)$, $(t,\tau) \in \Gamma_{t_0}^b$ *satisfying* (6.10) *exists and is a nuclear operator.*

Proof. By Theorem 6.1.12 and the uniform boundedness of M_K, H and R^{-1} in $[t_0, b]$,

$$\begin{aligned}&\|P(\tau|t)\|_1\\ \leqslant&\|P(\tau|\tau)\|_1 + \|P(\tau|\tau)\|_1^2 \int_\tau^t \|M_K(s,\tau)\|^2 \|R^{-1}(s)\| \|H(s)\|^2 ds < \infty,\end{aligned}$$

so $P(\tau|t)$ is a nuclear operator for any $(t,\tau) \in \Gamma_{t_0}^b$. □

It is worth noting that the nuclearity of $P(t_0|t_{-1})$ can be guaranteed when the series of the singular values of $P(t_0|t_{-1})$ is summable.

The derivations above of the Kalman filter and smoother in Hilbert spaces are based on the integral linear model (6.2) driven by white noises. In fact, for the linear model driven by the independent finite-dimensional Wiener processes w_1 and w_2,

$$x(t) = M(t,t_0)x(t_0) + \int_{t_0}^t M(t,s)B(s)u(s)ds + \int_{t_0}^t M(t,s)D(s)dw_1(s) \tag{6.11}$$

with the corresponding observation system

$$z(t) = \int_{t_0}^t H(s)x(s) + E(s)dw_2(s), \tag{6.12}$$

it can be easily found in [19] and [20] that the covariances from the Kalman filter and smoother based on (6.11) and (6.12) still satisfy (6.7) and (6.10).

6.3 Optimal locations of observations based on KF and KS

We now take the observation location problem into account. The location parameter r is defined as in Section 5.2. The following theorems show the continuity of $P_r(t|t)$ and $P_r(\tau|t), (t,\tau) \in \Gamma_{t_0}^b$ in the nuclear norm. For the filter problem, due to the duality and Theorem 5.2.7, we obtain the following theorem.

Theorem 6.3.1. *Consider the filter problem of the time-varying system* (6.2) *with location-dependent output operators and the observation system* (6.3). *If H_r has the property that $\lim_{r\to r_0} \|H_r - H_{r_0}\|_\infty = 0$, $\mathcal{E}$ and $\mathcal{Y}$ are finite-dimensional, and $P(t_0|t_{-1})$ is nuclear, then*

$$\lim_{r\to r_0} \|P_r(t|t) - P_{r_0}(t|t)\|_1 = 0, \quad t \in [t_0, b]$$

and there exists an optimal location $\hat{r}^f$ such that

$$\hat{\ell}_1^f(t) = \|P_{\hat{r}^f}(t|t)\|_1 = \inf_{r\in\Omega^m} \|P_r(t|t)\|_1.$$

Theorem 6.3.2. *Consider the smoother problem of the time-varying system* (6.2) *with the location-dependent output operators and the observation system* (6.3). *H_r has the property that $\lim_{r\to r_0} \|H_r - H_{r_0}\|_\infty = 0$. If $\mathcal{E}$ and $\mathcal{Y}$ are finite-dimensional, and $P(t_0|t_{-1})$ is nuclear, then,*

$$\lim_{r\to r_0} \|P_r(\tau|t) - P_{r_0}(\tau|t)\|_1 = 0, \quad (t,\tau) \in \Gamma_{t_0}^b,$$

and there exists an optimal location $\hat{r}^s$ depending on the initial time τ such that

$$\hat{\ell}_1^s(\tau|t) = \|P_{\hat{r}^s}(\tau|t)\|_1 = \inf_{r\in\Omega^m} \|P_r(\tau|t)\|_1.$$

Proof. From Lemma 6.2.3, $P_r(\tau|t)$, $r \in \Omega^m$ are nuclear operators. Hence,

$$\begin{aligned}
&\|P_r(\tau|t) - P_{r_0}(\tau|t)\|_1 \leqslant \|P_r(\tau|\tau) - P_{r_0}(\tau|\tau)\|_1 \\
&+ \int_{t_0}^t \|P_{r_0}(\tau|\tau)M^*_{K,r_0}(s,\tau)H^*_{r_0}(s) - P_r(\tau|\tau)M^*_{K,r}(s,\tau)H^*_r(s)\| \\
&\cdot \|R^{-1}(s)H_{r_0}(s)M_{K,r_0}(s,\tau)P_{r_0}(\tau|\tau)\|_1 ds + \int_{t_0}^t \|P_r(\tau|\tau)M^*_{K,r}(s,\tau)H^*_r(s)R^{-1}(s)\|_1 \\
&\cdot \|H_{r_0}(s)M_{K,r_0}(s,\tau)P_{r_0}(\tau|\tau) - H_r(s)M_{K,r}(s,\tau)P_r(\tau|\tau)\| ds,
\end{aligned}$$

Since $P_r(t|t), r \in \Omega^m$ are nuclear operators and $R^{-1}(t), H_r(t), M_{K,r_0}(t,\tau)$ are uniformly bounded for $(t,\tau) \in \Gamma^b_{t_0}$, then

$$\|R^{-1}(s)H_{r_0}(s)M_{K,r_0}(s,\tau)P_{r_0}(\tau|\tau)\|_1 < \infty \tag{6.13}$$

and so is its adjoint.

From Theorem 6.3.1 and dominated convergence theorem, we obtain

$$\|P_r(\tau|t) - P_{r_0}(\tau|t)\|_1 \to 0, \quad r \to r_0.$$

Because of the compactness of Ω^m, there exists the optimal location of observations such that

$$\hat{\ell}_1^s(\tau|t) = \|P_{\hat{r}^s}(\tau|t)\|_1 = \inf_{r\in\Omega^m} \|P_r(\tau|t)\|_1.$$

□

Next we consider a sequence of approximations of time-varying systems in order to study the convergence of optimal observation locations based on the Kalman filter and smoother. Let $\mathcal{X}_n$ be a family of finite-dimensional subspaces of $\mathcal{X}$ and $\mathbf{P}_n$ be the corresponding orthogonal projection of $\mathcal{X}$ onto $\mathcal{X}_n$. The finite spaces $\{\mathcal{X}_n\}$ inherit the norm from $\mathcal{X}$. For $n \in \mathbb{N}$, let $M_n(\cdot,\cdot)$ be a mild evolution operator on $\mathcal{X}_n$, $D_n(t) = \mathbf{P}_n D(t)$ and $H_n(t) = H(t)\mathbf{P}_n$, $t \in [t_0, b]$. In order to guarantee that $P_n(t|t)$ converges to $P(t|t)$, the following assumptions are needed in the approximation of optimal observation location problems for partial differential equations.

(A1) For each $x \in \mathcal{X}$,

$$(i) \quad M_n(t,s)\mathbf{P}_n x \to M(t,s)x; \quad (ii) \quad M^*_n(t,s)\mathbf{P}_n x \to M^*(t,s)x.$$

and $\sup_n \|M_n(t,s)\| < \infty$, for any $(t,s) \in \Gamma^b_{t_0}$.

(A2) For each $\omega \in \mathcal{E}$,

$$(i) \quad D_n(t)\omega \to D(t)\omega; \quad (ii) \quad D^*_n(t)\mathbf{P}_n x \to D^*(t)x, \ a.e. \quad t \in [t_0, b].$$

(A3) For each $x \in \mathcal{X}$, $y \in \mathcal{Y}$,

$$(i) \quad H_n(t)\mathbf{P}_n x \to H(t)x; \quad (ii) \quad H_n^*(t)y \to H^*(t)y, \ a.e. \quad t \in [t_0, b].$$

(A4) For each $x \in \mathcal{X}$,

$$P_n(t_0|t_{-1})\mathbf{P}_n x \to P(t_0|t_{-1})x.$$

The next theorem shows the uniform convergence of the approximations of covariances of the Kalman filter and smoother in nuclear norm.

Theorem 6.3.3. *Assume that the assumptions* $(A1) - (A4)$ *are satisfied. If* $\mathcal{E}$ *and* $\mathcal{Y}$ *are finite-dimensional spaces,* $\lim_{n\to\infty} \|P_n(t_0|t_{-1})\mathbf{P}_n - P(t_0|t_{-1})\|_1 = 0$ *and* $P(t_0|t_{-1})$ *is nuclear, then*

$$\lim_{n\to\infty} \|P_n(t|t)\mathbf{P}_n - P(t|t)\|_1 = 0,$$
$$\lim_{n\to\infty} \|P_n(\tau|t)\mathbf{P}_n - P(\tau|t)\|_1 = 0, \quad (t,\tau) \in \Gamma_{t_0}^b.$$

Proof. Due to the duality between the Kalman filter and LQ optimal control problem, according to Theorem 5.3.5, we have

$$\lim_{n\to\infty} \|P_n(t|t)\mathbf{P}_n - P(t|t)\|_1 = 0, \quad (t,\tau) \in \Gamma_{t_0}^b. \tag{6.14}$$

Then, we obtain

$$\begin{aligned}
&\|P_n(\tau|t)\mathbf{P}_n - P(\tau|t)\|_1 \leqslant \|P_n(\tau|\tau)\mathbf{P}_n - P(\tau|\tau)\|_1 \\
&+ \int_\tau^t \|P(\tau|\tau)M_K^*(s,\tau)H^*(s) - \mathbf{P}_n P_n(\tau|\tau)M_{K,n}^*(s,\tau)H_n^*(s)\| \\
&\cdot \|R^{-1}(s)H(s)M_K(s,\tau)P(\tau|\tau)\|_1 ds + \int_\tau^t \|\mathbf{P}_n P_n(\tau|\tau)M_{K,n}^*(s,\tau)H_n^*(s)R^{-1}(s)\|_1 \\
&\cdot \|H(s)M_K(s,\tau)P(\tau|\tau) - H_n(s)M_{K,n}(s,\tau)P_n(\tau|\tau)\mathbf{P}_n\| ds,
\end{aligned}$$

where, according to Lemma 5.2.5 and (6.14),

$$\begin{aligned}
&\|H(s)M_K(s,\tau)P(\tau|\tau) - H_n(s)M_{K,n}(s,\tau)P_n(\tau|\tau)\mathbf{P}_n\| \\
\leqslant &\|H(s)M_K(s,\tau)\|\|P(\tau|\tau) - P_n(\tau|\tau)\mathbf{P}_n\| \\
&+ \|H(s) - H_n(s)\|\|M_K(s,\tau)P_n(\tau|\tau)\mathbf{P}_n\| \\
&+ \|H_n(s)\|\|(M_K(s,\tau) - M_{K,n}(s,\tau))P_n(\tau|\tau)\mathbf{P}_n\| \to 0, \quad n \to \infty.
\end{aligned} \tag{6.15}$$

So is its adjoint operator.

By the uniform boundedness of $P(t|t)$, $M_K(t,s)$, $H_n(t)$ for $t \in [t_0, b]$, we have

$$\|P_n(\tau|t)\mathbf{P}_n - P(\tau|t)\|_1 \to 0, \quad n \to \infty.$$

□

Now let us take the location of observations into account and show the convergence of optimal observation locations of approximated covariance of Kalman filter and smoother.

Theorem 6.3.4. *Assume that the assumptions* $(A1) - (A4)$ *holds and* $H_{r,n} = H_r\mathbf{P}_n$ *with* $\lim_{r\to r_0} \|H_r - H_{r_0}\|_\infty = 0$. *If* $\mathcal{E}$ *and* $\mathcal{Y}$ *are finite-dimensional spaces,* $P(t_0|t_{-1})$ *is nuclear and* $\lim_{n\to\infty} \|P_n(t_0|t_{-1})\mathbf{P}_n - P(t_0|t_{-1})\|_1 = 0$, *then*

$$\hat{\ell}^f_{1,n}(t) \to \hat{\ell}^f_1(t), \quad \hat{\ell}^s_{1,n}(\tau|t) \to \hat{\ell}^s_1(\tau|t), \quad (t,\tau) \in \Gamma^b_{t_0}, \quad n \to \infty.$$

and there exists a subsequence of the optimal locations $\hat{r}^f_{n_k}$ *depending on* $\hat{\ell}^f_{1,n_k}(t)$ *and a subsequence* $\hat{r}^s_{n_k}$ *depending on* $\hat{\ell}^s_{1,n_k}(t)$ *such that*

$$\hat{r}^f_{n_k} \to \hat{r}^f, \quad \hat{r}^s_{n_k} \to \hat{r}^s, \quad k \to \infty.$$

Proof. Follows by duality and Theorem 5.3.6. □

6.4 Application

As a popular data assimilation method, the ensemble Kalman filter and smoother have been widely applied in meteorology. In this section, we consider a linear advection-diffusion model with $\Omega := (0,5) \times (0,5) \times (0,1)$ on a fixed time interval $[0,3]$ based on the Kalman filter and smoother, the theoretical foundation of the ensemble Kalman filter and smoother, as an example:

$$\begin{aligned} &\frac{\partial \delta c}{\partial t} = -v_x\frac{\partial \delta c}{\partial x} - v_y\frac{\partial \delta c}{\partial y} + \frac{\partial}{\partial z}(K(z)\frac{\partial \delta c}{\partial z}) + \delta e - \delta d, \\ &\delta c(t_0) = \delta c_0, \quad \delta e(t_0) = \delta e_0, \quad \delta d(t_0) = \delta d_0, \end{aligned} \tag{6.16}$$

where δc, δe and δd are the perturbations of the concentration, the emission rate and deposition rate of a species, respectively. v_x and v_y are constants and $K(z)$ is a continuous differentiable function of z.

Defining $A_x := -v_x\frac{\partial}{\partial x}$, $A_y := -v_y\frac{\partial}{\partial y}$ and $D_z := \dfrac{\partial}{\partial z}(K(z)\dfrac{\partial}{\partial z})$ with domains

$$\begin{aligned} \mathcal{D}(A_x) &= \{f \in L^2(\Omega) \mid A_x f \in L^2(\Omega), f(0,y,z) = f(5,y,z)\}, \\ \mathcal{D}(A_y) &= \{f \in L^2(\Omega) \mid A_y f \in L^2(\Omega), f(x,0,z) = f(x,5,z)\}, \\ \mathcal{D}(D_z) &= \{f \in L^2(\Omega) \mid D_z f \in L^2(\Omega), f_z(x,y,0) = f_z(x,y,1) = 0\} \end{aligned}$$

and denote by S_x, S_y and S_z the semigroups generated by A_x, A_y and D_z. S is the semigroup generated by $A_x + A_y + D_z$ with the domain $\mathcal{D} = \mathcal{D}(A_x) \cap \mathcal{D}(A_y) \cap \mathcal{D}(D_z)$.

In particular, in order to include the emission rate into the state vector as optimized parameter, the dynamic model for emission rates is established as in Section 3.2.1

$$\delta e(t) = M_e(t,s)\delta e(s), \tag{6.17}$$

where $M_e(t,s) = \frac{e_b(t)}{e_b(s)} \in L(L^2(\Omega))$, $e_b(\cdot) \in L^2(\Omega)$ is termed as the background knowledge of the emission rate, which is continuous in time and

$$\sup_{(t,s)\in\Gamma_0^3} \|\frac{e_b(t)}{e_b(s)}\| < \infty.$$

According to Definition 5.1.1, $M_e(\cdot,\cdot)$ is a self-adjoint mild evolution operator.

Ignoring the model error, the model extended with emission rate is given by

$$\begin{aligned} &\begin{pmatrix} \delta c(t+\triangle t) \\ \delta e(t+\triangle t) \end{pmatrix} \\ = \quad & M(t+\triangle t,t)\begin{pmatrix} \delta c(t) \\ \delta e(t) \end{pmatrix} - \begin{pmatrix} \int_t^{t+\triangle t} S(t+\triangle t-s)\delta d(s)ds \\ 0 \end{pmatrix}, \end{aligned} \tag{6.18}$$

where

$$M(t+\triangle t,t) = \begin{pmatrix} S(\triangle t) & \int_t^{t+\triangle t} S(t+\triangle t-s)M_e(s,t)ds \\ 0 & M_e(t+\triangle t,t) \end{pmatrix} \tag{6.19}$$

also satisfies Definition 5.1.1.

The numerical solution is based on the symmetric operator splitting technique [5], [102] with space discretization via finite difference method with discretized intervals $\triangle x$, $\triangle y$ and $\triangle z$ in three dimensions. We assume that the grid points $\{r_i\}_{i=1}^n$ have the coordinates $\{(x_{r_i}, y_{r_i}, z_{r_i})\}$ and define the projection $\mathbf{P}_n : L^2(\Omega) \to \mathbb{R}^n$

$$(\mathbf{P}_n f)_i := \frac{1}{V_i}\int_{\Omega_i} f(\omega)d\omega, \quad i = 1, \cdots, n, \tag{6.20}$$

where

$$\Omega_i = [x_{r_i} - \frac{\triangle x}{2}, x_{r_i} + \frac{\triangle x}{2}] \times [y_{r_i} - \frac{\triangle y}{2}, y_{r_i} + \frac{\triangle y}{2}] \times [z_{r_i} - \frac{\triangle z}{2}, z_{r_i} + \frac{\triangle z}{2}]$$

and V_i is the volume of Ω_i. Defining

$$S_n(\triangle t) := S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(\triangle t)S_{y,n}(\frac{\triangle t}{2})S_{x,n}(\frac{\triangle t}{2}),$$

according to [4, Theorem 3.17], we obtain

$$\lim_{n\to\infty,\triangle t\to 0} \|(S_n(\triangle t))^{\frac{t}{\triangle t}}\mathbf{P}_n f - \mathbf{P}_n S(t)f\| = 0, \quad f \in L^2(\Omega). \tag{6.21}$$

With the same space discretization for δc, the approximation of the emission rate is given by

$$\mathbf{P}_n \delta e(t) = M_{e,n}(t,s)\mathbf{P}_n \delta e(s),$$

where $M_{e,n}(t,s)$ is a diagonal matrix with the diagonal given by

$$\text{diag}(M_{e,n}(t,s)) = (\frac{\int_{\Omega_1} e_b(t,\omega)d\omega}{\int_{\Omega_1} e_b(s,\omega)d\omega}, \cdots, \frac{\int_{\Omega_n} e_b(t,\omega)d\omega}{\int_{\Omega_n} e_b(s,\omega)d\omega}).$$

Then, we can easily get

$$\begin{aligned}
&\left\| \frac{\int_{\Omega_i} e_b(t,\omega)d\omega}{\int_{\Omega_i} e_b(s,\omega)d\omega}(\mathbf{P}_n f)_i - (\mathbf{P}_n M_e(t,s)f)_i \right\| \\
\leqslant &\left\| \frac{\frac{1}{V_i}\int_{\Omega_i} e_b(t,\omega)d\omega}{\frac{1}{V_i}\int_{\Omega_i} e_b(s,\omega)d\omega} - \frac{e_b(t,r_i)}{e_b(s,r_i)} \right\| \left\| \frac{1}{V_i}\int_{\Omega_i} f(\omega)d\omega \right\| \\
+ &\left\| \frac{e_b(t,r_i)}{e_b(s,r_i)}\frac{1}{V_i}\int_{\Omega_i} f(\omega)d\omega - \frac{e_b(t,r_i)}{e_b(s,r_i)}f(r_i) \right\| \\
+ &\left\| \frac{e_b(t,r_i)}{e_b(s,r_i)}f(r_i) - \frac{1}{V_i}\int_{\Omega_i} \frac{e_b(t,\omega)}{e_b(s,\omega)}f(\omega)d\omega \right\| \to 0, \quad n\to\infty, \quad f\in L^2(\Omega),
\end{aligned}$$

so is the adjoint of $M_e(t,s)$. The extended model with operator splitting discretized in space can be written as

$$\begin{aligned}
\begin{pmatrix} \delta c_n(t+\triangle t) \\ \delta e_n(t+\triangle t) \end{pmatrix} &= M_n(t+\triangle t,t)\begin{pmatrix} \delta c_n(t) \\ \delta e_n(t) \end{pmatrix} \\
&\quad - \begin{pmatrix} S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})\int_t^{t+\triangle t} S_{z,n}(t+\triangle t-s)\delta d_n(s)ds \\ 0 \end{pmatrix},
\end{aligned}$$

where $\delta c_n(t) = \mathbf{P}_n\delta c(t)$, $\delta e_n(t) = \mathbf{P}_n\delta e(t)$, $\delta d_n(t) = \mathbf{P}_n\delta d(t)$ and

$$\begin{aligned}
&M_n(t+\triangle t,t) \\
= &\begin{pmatrix} S_n(\triangle t) & S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})\int_t^{t+\triangle t} S_{z,n}(t+\triangle t-s)M_{e,n}(s,t)ds \\ 0 & M_{e,n}(t+\triangle t,t) \end{pmatrix}.
\end{aligned}$$

For any pair of time $(t,s)\in\Gamma_0^3$, assuming $m = \frac{t-s}{\triangle t}\in\mathbb{N}$, we have

$$\begin{aligned}
&\prod_{i=1}^{m} M_n(s+i\triangle t, s+(i-1)\triangle t) \\
&= \begin{pmatrix} (S_n(\triangle t))^m & \sum_{i=1}^m \int_{s+(i-1)\triangle t}^{s+i\triangle t} S_{ce,n}^i(t-h)M_{e,n}(h,s)dh \\ 0 & M_{e,n}(t,s) \end{pmatrix},
\end{aligned}$$

where for $h\in[s+(i-1)\triangle t, s+i\triangle t]$,

$$S_{ce,n}^i(t-h) = (S_n(\triangle t))^{m-i}S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(s+i\triangle t-h).$$

In order to show that $\prod_{i=1}^m M_n(s+i\triangle t, s+(i-1)\triangle t)\mathbf{P}_n$ is strongly convergent to $\mathbf{P}_n M(t,s)$, we only need to show

$$\|S_{ce,n}^i(t-h)M_{e,n}(h,s)\mathbf{P}_n f - \mathbf{P}_n S(t-h)M_e(h,s)f\| \to 0, \quad m,n\to\infty.$$

In fact,

$$\begin{aligned}
&\|S_{ce,n}^i(t-h)M_{e,n}(h,s)\mathbf{P}_n f - \mathbf{P}_n S(t-h)M_e(h,s)f\| \\
&\leqslant \|S_{ce,n}^i(t-h)M_{e,n}(h,s)\mathbf{P}_n f - S_{ce,n}^i(t-h)\mathbf{P}_n M_e(h,s)f\| \\
&+ \|S_{ce,n}^i(t-h)\mathbf{P}_n M_e(h,s)f - \mathbf{P}_n S(t-h)M_e(h,s)f\|,
\end{aligned}$$

where, clearly,

$$\|S^i_{ce,n}(t-h)M_{e,n}(h,s)\mathbf{P}_n f - S^i_{ce,n}(t-h)\mathbf{P}_n M_e(h,s)f\| \to 0,\ m,n\to\infty.$$

Moreover, we have

$$\begin{aligned}
&\|S^i_{ce,n}(t-h)\mathbf{P}_n M_e(h,s)f - \mathbf{P}_n S(t-h)M_e(h,s)f\| \\
\leqslant&\|((S_n(\triangle t))^{m-i} - S(t-s-i\triangle t))S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2}) \\
&\cdot S_{z,n}(s+i\triangle t-h)\mathbf{P}_n M_e(h,s)f\| \\
+&\|S(t-s-i\triangle t)(S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(s+i\triangle t-h)\mathbf{P}_n \\
&-\mathbf{P}_n S(s+i\triangle t-h))M_e(h,s)f\|.
\end{aligned}$$

where, according to (6.21),

$$\begin{aligned}
&\|((S_n(\triangle t))^{m-i} - \mathbf{P}_n S(t-s-i\triangle t)) \\
&\cdot S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(s+i\triangle t-h)\mathbf{P}_n M_e(h,s)f\| \to 0, \quad t\to 0, \quad n\to\infty
\end{aligned}$$

and

$$\begin{aligned}
&\|(S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(s+i\triangle t-h)\mathbf{P}_n - \mathbf{P}_n S(s+i\triangle t-h))M_e(h,s)f\| \\
\leqslant\ &\|S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(s+i\triangle t-h)\| \\
&\cdot\|(I - S_{z,n}(h-s-(i-1)\triangle t))\mathbf{P}_n M_e(h,s)f\| \\
+\ &\|(S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})S_{z,n}(\triangle t) - S_n(\triangle t))\mathbf{P}_n M_e(h,s)f\| \\
+\ &\|(S(h-s-(i-1)\triangle t) - I)S(s+i\triangle t-h)M_e(h,s)f\| \\
+\ &\|(S_n(\triangle t)\mathbf{P}_n - \mathbf{P}_n S(\triangle t))M_e(h,s)f\| \to 0, \quad \triangle t\to 0, \quad n\to\infty. \qquad (6.22)
\end{aligned}$$

Further, we discretize the model in time by the Lax-Wendroff scheme for advection equations in horizontal directions and Crank-Nicolson scheme for the diffusion equation in the vertical direction, such that $S_{x/y/z,n}$ are approximated by

$$\begin{aligned}
\tilde{S}_{x/y,n}(\frac{\triangle t}{2}) &= I + \frac{\triangle t}{2}A_{x/y,n} + \frac{\triangle t^2}{8}A^2_{x/y,n}, \\
\tilde{S}_{z,n}(\triangle t) &= (I - \frac{\triangle t}{2}D_{z,n})^{-1}(I + \frac{\triangle t}{2}D_{z,n}), \\
\tilde{B}^e_{z,n}(t,s)f &= (I - \frac{\triangle t}{2}D_{z,n})^{-1}(\frac{\triangle t}{2}(M_{e,n}(t,s)+I)f),
\end{aligned}$$

where $A_{x/y,n}$ and $D_{z,n}$ is the approximated generators to n-dimensional state space based on finite difference methods.

It is well known [30] that the Lax-Wendroff scheme is consistent and conditional stable for A_x and A_y and the Crank-Nicolson scheme is consistent and stable for D_z,

$(I - \frac{\triangle t}{2} D_{z,n})^{-1}$ is the consistent and conditional stable implicit Euler scheme, by Lax equivalence theorem, that is

$$\lim_{\triangle t \to 0} \|(\tilde{S}_{x/y/z,n}(\triangle t))^{\frac{t}{\triangle t}} f - S_{x/y/z,n}(t) f\| = 0, \quad f \in L^2(\Omega),$$
$$\lim_{\triangle t \to 0} \|((I - \frac{\triangle t}{2} D_{z,n})^{-1})^{\frac{2t}{\triangle t}} f - S_{z,n}(t) f\| = 0, \quad f \in L^2(\Omega).$$

Similarly defining $\tilde{S}_n := \tilde{S}_{x,n}\tilde{S}_{y,n}\tilde{S}_{z,n}\tilde{S}_{y,n}\tilde{S}_{x,n}$,

$$\lim_{n \to \infty, \triangle t \to 0} \|(\tilde{S}_n(\triangle t))^{\frac{t}{\triangle t}} \mathbf{P}_n f - \mathbf{P}_n S(t) f\| = 0, \quad f \in L^2(\Omega). \tag{6.23}$$

Since A_x and A_y are skew-adjoint, which generate unitary groups, and D_z is self-adjoint, $(\tilde{S}_n^*(\triangle t))^{\frac{t}{\triangle t}}$ is also strongly convergent to $S^*(t)$.

Thus, (6.18) is approximated by

$$\begin{aligned}
&\begin{pmatrix} \delta\tilde{c}_n(t+\triangle t) \\ \delta\tilde{e}_n(t+\triangle t) \end{pmatrix} \\
= &\begin{pmatrix} \tilde{S}_n(\triangle t) & \tilde{S}_{x,n}(\frac{\triangle t}{2})\tilde{S}_{y,n}(\frac{\triangle t}{2})\tilde{B}^e_{z,n}(t+\triangle t, t) \\ 0 & M_{e,n}(t+\triangle t, t) \end{pmatrix} \begin{pmatrix} \delta\tilde{c}_n(t) \\ \delta\tilde{e}_n(t) \end{pmatrix} \\
- &\begin{pmatrix} \tilde{S}_{x,n}(\frac{\triangle t}{2})\tilde{S}_{y,n}(\frac{\triangle t}{2})(I - \frac{\triangle t}{2} D_{z,n})^{-1}[\frac{\triangle t}{2}(\delta d_n(t+\triangle t) + \delta d_n(t))] \\ 0 \end{pmatrix}.
\end{aligned}$$

Defining the block evolution operator above as $\tilde{M}_n(t,s)$, $(t,s) \in \Gamma_0^3$, we have

$$\begin{aligned}
&\prod_{i=1}^{m} \tilde{M}_n(s + i\triangle t, s + (i-1)\triangle t) \\
= &\begin{pmatrix} (\tilde{S}_n(\triangle t))^m & \sum_{i=1}^m (\tilde{S}_n(\triangle t))^{m-i} \tilde{S}_{x,n}(\frac{\triangle t}{2})\tilde{S}_{y,n}(\frac{\triangle t}{2})\tilde{B}^e_{z,n}(s+i\triangle t, s) \\ 0 & M_{e,n}(t,s) \end{pmatrix}.
\end{aligned}$$

We define

$$\begin{aligned}
&B^e_{z,n}(s + i\triangle t, s + (i-1)\triangle t, s) f \\
:= &\int_{s+(i-1)\triangle t}^{s+i\triangle t} S_{z,n}(s + i\triangle t - h) M_{e,n}(h, s) f dh, \quad f \in L^2(\Omega).
\end{aligned}$$

By the trapezoidal rule and convergence of the implicit Euler scheme, we have

$$\begin{aligned}
& \|\tilde{B}^e_{z,n}(s+i\triangle t,s)f - B^e_{z,n}(s+i\triangle t,s)f\| \\
\leqslant\ & \|((I-\frac{\triangle t}{2}D_{z,n})^{-1} - S_{z,n}(\triangle t))(\frac{\triangle t}{2}(M_{e,n}(s+i\triangle t,s)+I)f)\| \\
+\ & \|\frac{\triangle t}{2}S_{z,n}(\triangle t)M_{e,n}(s+(i-1)\triangle t,s)f \\
& \qquad +\frac{\triangle t}{2}M_{e,n}(s+i\triangle t,s)f - B^e_{z,n}(s+i\triangle t,s)f\| \\
+\ & \|\frac{\triangle t}{2}S_{z,n}(\triangle t)(M_{e,n}(s+i\triangle t,s)f - M_{e,n}(s+(i-1)\triangle t,s)f)\| \\
+\ & \|\frac{\triangle t}{2}(M_{e,n}(s+i\triangle t,s)+S_{z,n}(\triangle t))f\| \to 0, \quad \triangle t\to 0.
\end{aligned}$$

According to (6.23) and dominated convergence theorem, we obtain that

$$\sum_{i=1}^{m}(\tilde{S}_n(\triangle t))^{m-i}\tilde{S}_{x,n}(\frac{\triangle t}{2})\tilde{S}_{y,n}(\frac{\triangle t}{2})\tilde{B}^e_{z,n}(s+i\triangle t,s)$$

is strongly convergent to

$$\sum_{i=1}^{m}(S_n(\triangle t))^{m-i}S_{x,n}(\frac{\triangle t}{2})S_{y,n}(\frac{\triangle t}{2})B^e_{z,n}(s+i\triangle t,s).$$

Further,

$$\prod_{i=1}^{m}\tilde{M}_n(s+i\triangle t,s+(i-1)\triangle t)$$

is strongly convergent to

$$\prod_{i=1}^{m}M_n(s+i\triangle t,s+(i-1)\triangle t).$$

For the observation system, we assume there is only a single observation during the entire time interval and define the observation mapping $H_r : L^2(\Omega)\to\mathbb{R}$ by

$$H_r f := \frac{1}{V_r}\int_{\Omega_r} f(\omega)d\omega, \quad r=(x_r,y_r,z_r), \quad f\in L^2(\Omega),$$

where Ω_r and V_r are similarly defined as (6.20). Then, the observation system extended by the emission rate is given by

$$\delta y(t) = (H_r,0)\begin{pmatrix}\delta c(t)\\ \delta e(t)\end{pmatrix}+\nu(t),$$

where $\delta y(t)\in\mathbb{R}$ and $\nu(t)$ is the white noise with the distribution $\mathcal{N}(0,1)$.

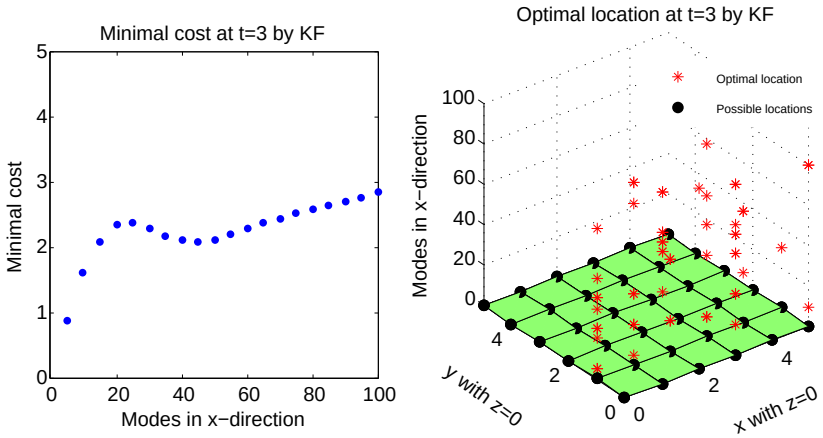

Figure 6.1: Minimal cost $\hat{\ell}^f_{1,n}(3)$ (left) and the corresponding optimal location $\hat{r}^f_n$ at $t = 3$ (right) with $P(t_0|t_{-1}) = e^{-8} I_n$. Points at $z = 0$ in the right figure: Possible observation locations. Stars in the right figure: Optimal locations.

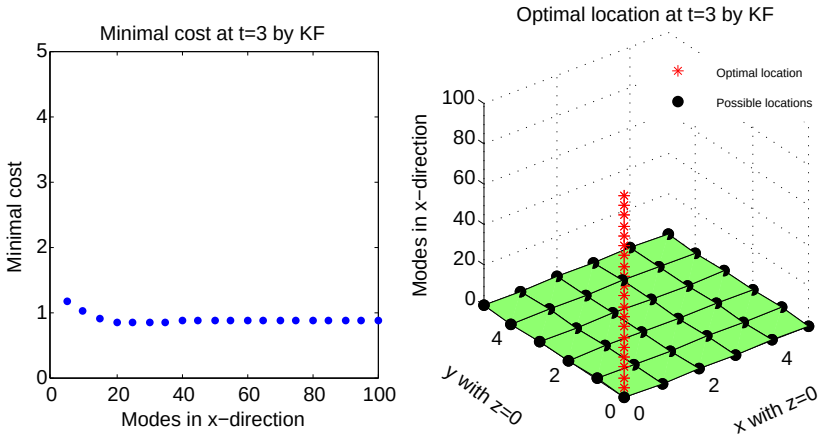

Figure 6.2: Minimal cost $\hat{\ell}^f_{1,n}(3)$ (left) and the corresponding optimal location $\hat{r}^f_n$ at $t = 3$ (right) with $P(t_0|t_{-1}) = \sum_{i=1}^{\infty} e^{-i^2} \langle \cdot, e_i \rangle e_i$. Points at $z = 0$ in the right figure: Possible observation locations. Stars in the right figure: Optimal locations.

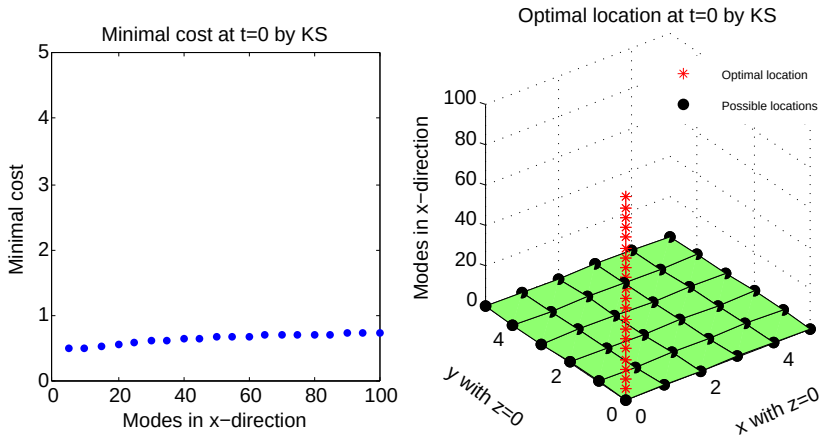

Figure 6.3: Minimal cost $\hat{\ell}^s_{1,n}(0)$ (left) and the corresponding optimal location $\hat{r}^s_n$ at $t = 0$ (right) with $P(t_0|t_{-1}) = \sum_{i=1}^{\infty} e^{-i^2}\langle \cdot, e_i\rangle e_i$. Points at $z = 0$ in the right figure: Possible observation locations. Stars in the right figure: Optimal locations.

According to the spatial discretization of the model, in the vertical direction, $[0, 1]$ is discretized into three layers $\{0, 0.5, 1\}$. Since the diffusion coefficient $K(z)$ is small, we assume that possible locations of the single observation are around the grid points in the first layer $z = 0$, which are shown as the points at $z = 0$ in the right plots in Figure 6.1, 6.2 and 6.3. Besides, we only choose one observation location and display the optimal locations with the increasing number of modes of approximations as stars in the right plots in Figure 6.1, 6.2 and 6.3. Fixing the dimensions of the approximated systems in y and z directions, the $z-$axis in Figure 6.1, 6.2 and 6.3 shows the increasing dimensions of the approximates systems in x direction.

We have already shown that the assumptions $(A1) - (A3)$ in Section 6.3 and the compactness of the possible area of observation locations are satisfied.

In addition, according to the spatial discretization, we assume that the initial covariance is given by $P_n(t_0|t_{-1}) = e^{-8} I_n$, where I_n is the $n \times n$ identity matrix. It implies that $P_n(t_0|t_{-1})$ does not converge to a nuclear operator. It is shown in Figure 6.1 that the optimal locations and minimal costs of a sequence of approximations based on Kalman filter do not converge in this situation.

Next we define the initial covariance as

$$P(t_0|t_{-1})f = \sum_{i=1}^{\infty} e^{-i^2}\langle f, e_i\rangle e_i, \quad f \in L^2(\Omega),$$

where $\{e_i\}$ is an orthogonal basis of $L^2(\Omega)$. The n-dimensional approximation of $P(t_0|t_{-1})$ is given by

$$P_n(t_0|t_{-1})\mathbf{P}_n f = \sum_{i=1}^{n} e^{-i^2}\langle \mathbf{P}_n f, e_i\rangle e_i, \quad f \in L^2(\Omega).$$

With this choice, $P(t_0|t_{-1})$ is nuclear and the assumption $(A4)$ in Section 6.3 is satisfied. From Theorem 6.3.4, the optimal locations of observations and the corresponding

minimal costs of a sequence of approximations are convergent to the optimal location of the observations and the minimal cost of the original model (6.16) extended with (6.17) by the Kalman filter and smoother. They are shown in Figure 6.2 for the filter and Figure 6.3 for the smoother, respectively.

Chapter 7

Conclusion

In this thesis, the observational analysis, including the efficiency analysis, sensitivity analysis and the optimal problem of observation locations, was discussed.

The main contributions of this thesis are as follows. Firstly, we established the tangent linear form of the atmospheric transport model extended by emission rates under the assumption that emissions preserve the invariant diurnal profiles. The initial value and emission rates play the equivalent roles in the extended model since the homogeneity of the extended model is such that emission rates are available to be optimized by the Kalman smoother as model variables.

Secondly, in the context of Kalman smoother, the relative improvement covariance is derived as the criterion to evaluate the potential improvement of each grid point in the state vector and calculated by singular value decomposition. With a statistical interpretation, we can apply it to determine in advance, which parameters can be optimized by the data assimilation procedure. A number of metrics associated with the relative improvement covariance provides us with the quantitative solutions to measure to what extend the parameters can be optimized. Due to its relativity of the normalization, it is uniformly available for any prior initial values of invertible background covariances. Further, the proposal of the ensemble relative improvement covariance, based on EnKS, gives us a computationally feasible access to assess the efficiency of observation networks. An elementary advection-diffusion example illustrated the significance of relative improvements covariances and their various metrics in different situations

Thirdly, the sensitivity of observational networks was formulated by seeking the fastest directions of the perturbation ratio between initial states and observation configurations during the entire time interval. The consistency of efficient and sensitive directions of observation networks complements the two approaches mutually and guarantees the feasibility to target the sensitive states by the weighted leading singular vectors. Further, we applied the sensitivity analysis of observation networks into the emission source apportionment problem.

Fourthly, we studied the optimal problems of control locations for time-varying systems on a finite-time horizon in Hilbert spaces. In the context of linear quadratic control, the minimal costs caused by the worst initial condition in the operator norm and the random initial condition were evaluated in the nuclear norm, respectively. By the compactness of input and output operators, the well-posedness of the optimal control location

problem with the worst initial condition is proved. By restricting input spaces and output spaces into finite-dimensional spaces, the well-posedness of optimal problems of control locations are also guaranteed for random initial conditions. The theorems concerning the convergence of minimal costs and optimal locations of a sequence of approximating systems to the original system allow us to apply the results in practice.

Finally, by the duality between Kalman filters and LQ control problems, the optimal problem of observation locations based on Kalman filters (and further smoothers) has been similarly studied in the nuclear norm. The application to a particular advection-diffusion equation extended with emissions shows the necessity to study the mild integral form of the time-varying system.

In the future, on one hand, we plan to apply the efficiency analysis into the real atmospheric transport model to solve practical problems. On the other hand, we will consider the optimal problems of control and observation locations with unbounded input operators and output operators in order to deal with the boundary control problems and point observation systems.

Bibliography

[1] R. ABIDA, M. BOCQUET, *Targeting of observations for accidental atmospheric atmospheric release monitoring*, Atmos. Env., Vol. 43, pp. 6312–6327, 2009.

[2] A. V. BALAKRISHNAN, *Stochastic optimization theory in Hilbert spaces -1*, Applied Mathematics and Optimization, Vol. 1, No. 2, pp. 97–120, 1974.

[3] A. V. BALAKRISHNAN, *Applied functional analysis*, New York: Springer-Verlag, 1976.

[4] A. BÁTKAI, P. CSOMÓS AND G. NICKEL, *Operator splitting and spatial approximations for evolution equations*, J. Evolution equations, Vol. 9, pp. 613–636, 2009.

[5] A. BÁTKAI, P. CSOMÓS, B. FARKAS, AND G. NICKEL, *Operator splitting with spatial temporal discretization*, Oper. Theory Adv. Appl., Vol. 221. pp. 161–171, 2012.

[6] J. BARKMEIJER, M. V. GIJZEN AND F. BOUTTIER, *Singular vectors and estimates of the analysis-error covariance metric*, Q. J. Roy. Meteorol. Soc., Vol. 124, pp. 1695–1713, 1998.

[7] A. BENSOUSSAN, *Optimization of sensors' location in a distributed filtering problem (ed. R.F. Curtain), Chapter Stability of Stochastic Dynamical Systems*, Springer, 1972.

[8] L. M. BERLINER, Z. LU AND C. SNYDER, *Statistical design for adaptive weather observations*, J. Atmosph. Sci., Vol. 56, pp. 2536–2552, 1998.

[9] C. H. BISHOP AND Z. TOTH, *Ensemble transformation and adaptive ob- servations*, J. Atmosph. Sci., Vol. 56, pp. 1748–1765, 1998.

[10] M. BOCQUET AND P. SAKOV, *Joint state and parameter estimation with an iterative ensemble Kalman smoother*, Nonlinear proc geoph., Vol. 20, pp. 803–818, 2013.

[11] R. W. BROCKETT, *Finite dimensional linear systems*, John Wiley and Sons, Inc, 1970.

[12] R. BUIZZA, *Localization of optimal perturbations using a projection operator*, Quart. J. Roy. Meteor. Soc., Vol. 120, pp. 1647–1681, 1994.

[13] R. BUIZZA AND A. MONTANI, *Targeting observations using singular vectors*, J. Atmos. Sci., Vol. 56, pp. 2965–2985, 1999.

[14] R. BUIZZA, J. BARKMEIJER, T. N. PALMER AND D. S. RICHARDSON, *Current status and future developments of the ECMWF Ensemble Prediction System*, Meteorol. Appl., Vol. 7, No. 2, pp. 163–175, 2000.

[15] R. BUIZZA, T. N. PALMER, *The singular-vector structure of the atmospheric global circulation*, J. Atom. Sci., Vol. 52, No. 9, pp. 1434–1456, 1995.

[16] D. E. CATLIN, *Estimation, control, and the discrete Kalman filter*, Springer-Verlag, 1989.

[17] A. CIOACA AND A. SANDU, *Low-rank approximations for computing observation impact in 4D-Var data assimilation*, Compu. & Math. with Appli., Vol. 67(12), pp. 2112–2126, 2014.

[18] A. CIOACA AND A. SANDU, *An optimization framework to improve 4D-Var data assimilation system performance*, J. Compu. Phy., Vol. 275, pp. 377–389, 2014.

[19] R. F. CURTAIN, *Estimation theory for abstract evolution equations excited by general white noise processes*, SIAM J. Control Optim., Vol. 14, No. 6, pp. 1124–1150, 1976.

[20] R. F. CURTAIN AND A. J. PRITCHARD, *Infinite dimensional linear systems theory*, Lecture Notes in Control and Information Sciences, Vol. 8, 1978.

[21] R. F. CURTAIN AND G. WEISS, *Well posedness of triples of operators (in the sense of linear systems theory)*, International Series of Numerical Mathematics, Vol. 91, pp. 41–58, Birkhäuser-Verlag, 1989.

[22] R. F. CURTAIN AND H. ZWART, *An introduction to infinite-dimensional linear systems theory*, Berlin, Germany: Springer Verlag, 1995.

[23] R. CURTAIN AND A. SASANE, *Compactness and nuclearity of the Hankel operator and internal stability of infinite-dimensional state linear systems*, Int. J. Control, Vol. 74, No. 12, pp. 1260–1270, 2001.

[24] R. CURTAIN, K. MIKKOLA AND A. SASANE, *The Hilbert-Schmidt property of feedback operators*, J. Math. Anal. Appl., Vol. 329, pp. 1145–1160, 2007.

[25] M. CUSACK, N. PÉREZ, J. PEY, A. ALASTUEY AND X. QUEROL, *Source apportionment of fine PM and sub-micron particle number concentrations at a regional background site in the western Mediterranean: a 2.5 year study*, Atmos. Chem. Phys., Vol. 13, pp. 5173–5187, 2013.

[26] D. DAESCU, *On the sensitivity equations of four-dimensional variational (4D-Var) data assimilation*, Mon. Wea. Rev., Vol. 136, pp. 3050–3065, 2008.

[27] D. DAESCU AND I. M. NAVON, *Adaptive observations in the context of 4D-Var data assimilation*, Meteorol. Atmos. Phys., Vol. 55, pp. 205–236, 2004.

[28] R. DALEY, *Atmospheric data analysis*, Cambridge University Press, 1991.

[29] N. DARIVANDI, K. MORRIS AND A. KHAJEPOUR, *An algorithm for LQ optimal actuator location*, Smart Mater. Struct., Vol. 22, Jan. 28, 2013.

[30] R. DAUTRAY AND J. L. LIONS, *Mathematical analysis and numerical methods for science and technoligy: Volume 6 Evolution Problem 2*, Springer, 1999.

[31] M. C. DELFOUR, *The linear quadratic optimal control problem for hereditary differential systems: theory and numerical solution*, Appl. Math. and Opt., Vol. 3, pp. 101–162, 1977.

[32] KJ. ENGEL AND R. NAGEL, *One-parameter semigroup for linear evolution equations*, Springer, 2000.

[33] H. ELBERN, H. SCHMIDT AND A. EBEL, *Variational data assimilation for tropospheric chemistry modeling*, J. Geophys. Res., Vol. 102, (D13), pp. 15967–15985, 1997.

[34] H. ELBERN AND H. SCHMIDT, *A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling*, J. Geophys. Res., Vol. 104, pp. 18583–18598, 1999.

[35] H. ELBERN, H. SCHMIDT, O. TALAGRAND AND A. EBEL, *4D-variational data assimilation with an adjoint air quality model for emission analysis*, Environ modell softw, Vol. 15, pp. 539–548, 2000.

[36] H. ELBERN, A. STRUNK, H. SCHMIDT AND O. TALAGRAND, *Emission rate and chemical state estimation by 4-dimension variational inversion*, Atmos. Chem. Phys., Vol. 7, pp. 3749–3769, 2007.

[37] G. EVENSEN, *Data assimilation: The ensemble Kalman filter, 2th Edition*, Springer, 2009.

[38] P. L. FALB, *Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space*, Info. and Control, Vol. 11, pp. 102–137, 1967.

[39] A. GELB (ED.), *Applied optimal estimation*, Cambridge, Mass.: M.I.T. Press, 1974.

[40] A. GERMANI, L. JETTO AND M. PICCIONI, *Galerkin approximation for optimal linear filtering of infinite-dimensional linear systems*, SIAM J. Control Optim., Vol. 26, No. 6, pp. 1287–1305, 1988.

[41] A. GERMANI, C. MANES AND P. PEPE, *A twofold spline approximation for finite horizon LQG control of hereditary systems*, SIAM J. Control Optim., Vol. 39, No. 4, pp. 1233–1295, 2000.

[42] J. C. GEROMEL, *Convex analysis and global optimization of joint actuator location and control problems*, IEEE Trans. Autom. Control, Vol. 34, No. 7, pp. 711–720, 1989.

[43] J. GIBSON, *The Riccati integral equations for optimal control problems on Hilbert spaces*, SIAM J. Control Optim., Vol. 17, No. 4, pp. 637–665, 1979.

[44] J. GIBSON, *An analysis of optimal model regulation: convergence and stability*, SIAM J. Control Optim., Vol. 19, No. 5, pp. 686–707, 1981.

[45] J. GIBSON, *Linear-quadratic optimal control of hereditary differential systems: Infinite dimensional Riccati equations and numerical approximations*, SIAM J. Control Optim., Vol. 21, No. 1, pp. 95–139, 1983.

[46] I. C. GOHBERG AND M. G. KREĬN, *Introduction to the theory of linear nonselfadjoint operators*, American Mathematical Society, 1969.

[47] N. GORIS AND H. ELBERN, *Singular vector decomposition for sensitivity analyses of tropospheric chemical scenarios*, Atmos. Chem. Phys., Vol. 13, pp. 5063–5087, 2013.

[48] N. GORIS AND H. ELBERN, *Singular vector based targeted observations of chemical constituents: description and first application of EURAD-IM-SVA*, Geosci. Model Dev. Discuss., Vol. 8, pp. 6267–6307, 2015.

[49] F. M. HAM AND R. G. BROWN, *Observability, eigenvalues and Kalman filtering*, IEEE Tran. Aero. and Electronic sys. Vol. AES-19, No. 2, 1983.

[50] N. J. HIGHAM, *Accuracy and stability of numerical algorithms: Second edition*, SIAM, 2002.

[51] K. ITO, *Strong convergence and convergence rates of approximating solutions for algebraic Riccati equations in Hilbert spaces*, Distributed Parameter Systems, Lecture Notes in Control and Information Sciences, Vol. 102, pp. 153–166, Springer-Verlag, 1987.

[52] A. H. JAZWINSKI, *Stochastic processes and filtering theory*, Academic Press, 1970.

[53] C. D. JOHNSON, *Optimization of a certain quality of complete controllability and observability for linear dynamic systems*, Trans. ASME, Vol. 91, series D, pp. 228–238, 1969.

[54] B. B. KHATTATOV, J. GILLE, L. LYJAK, G. BRASSEUR, V. DVORTSOV, A. ROCHE AND J. WATERS, *Assimilation of photochemically active species and a case analysis of UARS data*, J. Geophys. Res., Vol. 22, pp. 18715–18738, 1999.

[55] R. E. KALMAN, *A new approach to linear filtering and prediction problems*, J. Basic Engineering, pp. 35–45, Mar. 1960.

[56] R. E .KALMAN AND R. S. BUCY, *New results in linear filtering and prediction theory*, J. Basic Engineering, pp. 95–108, Mar. 1961.

[57] W. KANG AND L. XU, *Optimal placement of mobile sensors for data assimilations*, Tellus A, Vol. 64, 2012.

[58] D. KASINATHAN AND K. A. MORRIS, *H-infinity optimal actuator location*, IEEE Trans. Autom. Control., Vol. 58, No. 10, pp. 2522–2535, 2013.

[59] T. KATO, *Perturbation theory for linear operators*, Springer, 1991.

[60] T. G. KURTZ, *Extension of Trotter's operator Semigroup approximation theorems*, J. Func. Anal., Vol. 3, pp. 111–132, 1969.

[61] W. LAHOZ, B. KHATTATOV AND R. MÉNARD(EDS), *Data assimilation: Making sense of observations*, Springer, 2010.

[62] I. LASIECKA AND A. MANITIUS, *Differentiability and convergence rates of approximating semi-groups for retarded functional differential equations*, SIAM J. Numer. Anal, Vol. 25, pp. 883–907, 1988.

[63] I. LASIECKA AND R. TRIGGIANI, *Control theory for partial differential equations: continuous and approximation theories*, Cambridge, U.K.: Cambridge Univ. Press, Vol. I, II, 2000.

[64] S. LAWRENCE, R. SOKHI, K. RAVINDRA AND H. MAO, *Source apportionment of traffic emissions of particulate matter using bunnel measurements*, Atmos. Environ., Vol. 77, pp. 548–557, 2013.

[65] W. LIAO, A. SANDU, G. R. CARMICHAEL AND T. CHAI, *Singular vector analysis for atmospheric chemical transport models*, Month. Weather Rev., Vol. 134, pp. 2443–2465, 2006.

[66] J. L. LIONS, *Optimal control of systems governed by partial differential equations*, Springer, New York, 1971.

[67] E. N. LORENZ, *A study of the predictability of a 28 variable atmospheric model*, Tellus, Vol. 17, pp. 321–333, 1965.

[68] P. S. MAYBECK, *Stochastic models, estimation, and control, Volume 1*, Academic Press, 1979.

[69] P. S. MAYBECK, *Stochastic models, estimation, and control, Volume 2*, Academic Press, 1979.

[70] D. Q. MAYNE, *A solution of the smoothing problem for linear dynamic systems*, Automatica, Vol. 4, pp. 73–92, 1966.

[71] R. K. MEHRA, *Optimization of measurement schedules and sensor designs for linear dynamic systems*, IEEE Trans. Autom. Control, Vol. AC-21, No. 1, pp. 55–64, 1976.

[72] B. C. MOORE, *Principal component analysis in linear systems: controllability, observability, and model reduction*, IEEE Trans. Autom. Control, Vol. 26(1), pp. 17–32, 1981.

[73] K. A. MORRIS, *Design of finite-dimensional controllers for infinite-dimensional systems by approximation*, J. Math. Systems Estim. Control, Vol. 4, pp. 1–30, 1994.

[74] K. A. MORRIS, *Convergence of controllers designed using state-space techniques*, IEEE Trans. Autom. Control, Vol. 10, pp. 2100–2104, 1994.

[75] K. MORRIS, *Linear-quadratic optimal actuator location*, IEEE Trans. Autom. Control, Vol. 56, pp. 113–124, 2011.

[76] D. F. MORRISON, *Multivariate statistical methods*, New York: McGraw-Hill, 1967.

[77] C. NAM, Y. KIM AND T. A. WEISHAAR, *Optimal sizing and placement of piezo-actuators for active flutter suppression*, Smart Mater. Struct, Vol. 5, pp. 216–224, 1996.

[78] J. Y. OURARD, *Linear filtering in Hilbert spaces II: An application to the smoothing theory for hereditary systems with observation delays*, SIAM J. Control Optim., Vol. 16, No. 6, pp. 938–952,1978.

[79] S. L. PADULA AND R. K. KINCAID, *Optimization strategies for sensor and actuator placement*, Technical Report NASA/TM-1999-209126, Apr., 1999.

[80] K. R. PARTHASARATHY, *Probability measures on metric spaces*, Academic press, 1967.

[81] A. PAZY, *Semigroup of linear operators and application to partial differential equations*, New York: Springer-Verlag, 1983.

[82] W. PETERS, J. B. MILLER, J. WHITAKER, A. S. DENNING, A. HIRSCH, M. C. KROL, D. ZUPANSKI, L. BRUHWILER AND P. P. TANS, *An ensemble data assimilation system to estimate CO_2 surface fluxes from atmospheric trace gas observations*, J. Geophys. Res., Vol. 110(D24), 2005.

[83] H. E. RAUCH, F. TUNP AND C. T. STRIEBEL, *On the maximum likelihood estimates for linear dynamic systems*, AIAA Journal, Vol. 3, pp. 1445–1450, 1965.

[84] C. D. RODGERS, *Information content and optimization of high spectral resolution measurements*, Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research, Proc. SPIE 2830, pp. 136–147, 1996.

[85] C. D. RODGERS, *Inverse methods for atmospheric sounding: theory and practice (Series on atmospheric, oceanic and planetary physics—Vol.2)*, World Scientific, 2000.

[86] A. SANDU, A. CIOACA AND V. RAO, *Dynamic sensor network configuration in InfoSymbiotic systems using model singular vectors*, Procedia Computer Science 18, pp. 1909–1918, 2013.

[87] A. SANDU, D. N. DAESCU, G. R. CARMICHAEL AND T. CHAI, *Adjoint sensitivity analysis of regional air quality models*, J. Comput. Phys., Vol. 204, pp. 222–252, 2005.

[88] A. DE SANTIS, A. GERMANI, AND L. JETTO, *Approximation of the algebraic Riccati equation in the Hilbert space of Hilbert-Schmidt operators*, SIAM J. Control Optim., Vol. 31, No. 4, pp. 847–874, 1993.

[89] Y. SASAKI, *Some basic formalisams in numerical variational analysis*, Mon. Weather Rev., Vol. 98, pp. 875–883, 1970.

[90] K. SINGH, A. SANDU, M. JARDAK, K. W. BOWMAN AND M. LEE, *A practical method to estimate information content in the context of 4D-Var data assimilation*, SIAM/ASA J. Uncertainty quantification, Vol. 1(1), pp. 106–138, 2013.

[91] E. D. SONTAG, *Mathematical control theory*, Springer, Berlin, 1990.

[92] O. TALAGRAND, *Assimilation of observations, an introduction*, J. Me- teor. Soc. Japan, Vol. 75, (1B), pp. 191–209, 1997.

[93] X. TANG, J. ZHU, Z. F. WANGAND AND A. GBAGUIDI, *Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions*, Atmos. Chem. Phys., Vol. 11, pp. 12901–12916, 2011.

[94] R. D. TORN AND G. J. HAKIM, *Ensemble-based sensitivity analysis*, Mon. Weather Rev., Vol. 136, pp. 663–677, 2008.

[95] M. TUCSNAK AND G. WEISS, *Observation and control for operator semigroup*, Birkhäuser, 2009.

[96] D. UCINSKI, *Optimal sensor location for parameter estimation*, Internat. J. Control, Vol. 73, pp. 1325–1248, 2000.

[97] J. E. WALL, A. S. WILLSKY AND N. R. SANDELL, *On the fixed-interval smoothing problem*, Stochastics, Vol. 5, pp. 1–41, 1981.

[98] J. WEIDMANN, *Linear operators in Hilbert spaces*, New York: Springer-Verlag, 1980.

[99] M. A. WOODBURY, *Inverting modified matrices*, Memorandum Rept. 42, Statistical Research Group, Princeton University, Princeton, NJ, 1950.

[100] X. WU, H. ELBERN AND B. JACOB, *Efficiency and sensitivity analysis of observation networks for atmospheric inverse modelling with emissions*, arXiv:1503.06585, Preprint, March, 2015.

[101] X. WU, B. JACOB AND H. ELBERN, *Optimal control and observation locations for time-varying systems on a finite-time horizon*, SIAM J. Control Optim., Accepted, Dec., 2015.

[102] N. N. YANENKO, *The method of fractional steps: solution of problems of mathematical physics in several variables*, Springer, 1971.

[103] J. ZABCZYK, *Mathematical control theory: an introduction*, Birkhäuser, 2008.