UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

APPLICATION OF THE FORWARD SENSITIVITY METHOD TO DATA ASSIMILATION OF THE LAGRANGIAN TRACER DYNAMICS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

RAFAL JABRZEMSKI
Norman, Oklahoma
2014

APPLICATION OF THE FORWARD SENSITIVITY METHOD TO DATA ASSIMILATION OF THE LAGRANGIAN TRACER DYNAMICS

A DISSERTATION APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY

Dr. S. Lakshmivarahan, Chair

Dr. Sudarshan Dhall

Dr. Sridhar Radhakrishnan

Dr. Krishnaiyan Thulasiraman

Dr. Ming Xue

DEDICATION

to

My parents

Kazimierz Jabrzemski and Eleonora Jabrzemska,

my wife Marzena Jabrzemski

and daughters Nina and Natalia Jabrzemski

For

Encouragement, belief and patience

Acknowledgements

I wish to express my enormous gratitude to my research advisor, Professor Lakshmivarahan, for teaching me a great deal of computational science, for his advice and support, and for being an excellent mentor. I am very grateful for having had the chance to work with him.

I greatly appreciate guidance and time given by the members of my committee including Dr. Sridhar Radhakrishnan, Dr. Sudarshan Dhall, Dr. Krishnaiyan Thulasiraman and Dr. Ming Xue.

I also wish to thank the Oklahoma Mesonet for supporting the work on my dissertation.

Table of Contents

List of Tables

List of Figures

Abstract

The analysis of the dynamics of a tracer/drifter/buoy floating on the free surface of the water waves in the open ocean whose motion is described by the shallow water model equations is of great interest in Lagrangian data assimilation. A special case of the low/reduced order version of the linearized shallow water model equations gives rise to a class of tracer dynamics given a system of two first order, nonlinear, time varying systems of ordinary differential equations whose flow field is the sum of a time invariant geostrophic mode that depends on a parameter $u_0 \in \mathbb{R}$ and a time varying inertial-gravity mode that depends on a set of three parameters, $\hat{\boldsymbol{\alpha}} = \big(u_1(0), v_1(0), h_1(0)\big)^T \in \mathbb{R}^3$. In this thesis we provide a complete characterization of the properties of the equilibria of the tracer dynamics along with with bifurcation as the four parameters in $\boldsymbol{\alpha} = (u_0, \hat{\boldsymbol{\alpha}})^T \in \mathbb{R}^4$ are varied. It is shown that the impact of the four parameters can be effectively captured by to systems of intersecting hyperbolas in two dimensions. We then apply the Forward Sensitivity Method (FSM) to assimilate data in the twin experiments following the dynamics of the Lagrangian tracers in the shallow water model. In these experiments, we assume that the error results from the incorrect estimation of the control vector $\boldsymbol{\alpha} = \big(u_0, u_1(0), v_1(0), h_1(0)\big)^T \in \mathbb{R}^4$. We also analyze the sensitivity of the model to changes in the elements of the control vector $\boldsymbol{\alpha}$ in order to improve placement of the observations. We have found that sensitivity, together with the condition number of the matrix constructed with sensitivity values, gives a good prognostications about success of the data assimilation.

Chapter 1

Introduction

1.1 Motivation

Dynamics started as a branch of physics in the seventeenth century to deal with description of a change that can be observed for the systems that evolve in time. Ever since, dynamical models are created and used to describe the evolution of real systems. In order to use these dynamical models as a forecasting tool, we must incorporate observations into dynamical system - a process that is known as dynamical data assimilation. With the steady growth in the interest in ocean circulation systems and their impact on climate change, there has been a predictable growth in the number of tracer/drifter/buoy type ocean observing systems. There is a rich and growing literature on the development and testing of data assimilation technology to effectively utilize this new type of data sets. This class of data assimilation has come to be known as Lagrangian data assimilation. In this thesis, we consider problems concerning data assimilation in oceanography while dealing with buoys in Lagrangian models that follow parcels as they move with the flow.

Our goal in this research is twofold. First one is to analyze the shallow-water model behavior following approach presented by Lorenz (1960) [13] that he applied to the minimum hydrodynamic equations, and further expanded by (Lakshmivarahan et al., 2006) [8] in a search of equilibrium points of the minimum hydrodynamic equations model and to explore the bifurcation exhibited by the tracer dynamics induced by the low/reduced order version of the shallow water model obtained using the standard Galerkin type projection method. To this end, instead of relying on numerical methods, we first solve the resulting low order model which is a linear model equations in u, v and h given in Apte et al., (2008) [1] in a closed form. Using this explicit solution, we then express the flow field of the tracer dynamics as a sum of the two parts - a time invariant nonlinear geostrophic mode, $\mathbf{f}(\mathbf{x}, u_0)$ depending on the geostrophic parameter u_0 and a time varying nonlinear part known as the inertial gravity mode, $\mathbf{g}(\mathbf{x}, \hat{\boldsymbol{\alpha}})$ depending on three parameters $\hat{\boldsymbol{\alpha}} = (u_1(0), v_1(0), h_1(0))^T \in \mathbb{R}^3$. It turns out that the tracer dynamics controlled by the four parameters $\boldsymbol{\alpha} = (u_0, \hat{\boldsymbol{\alpha}}) \in \mathbb{R}^4$ exhibits complex behavior.

Second goal is to explore the applicability of the new class of methods called the forward sensitivity method (FSM) (Lakshmivarahan and Lewis 2010 [10]) for assimilating tracer data and test the impact of observations on assimilation procedures as it was noticed by [Lakshmivarahan and Lewis (2011) [11]].

1.2 Previous work

1.2.1 Assimilation of Lagrangian Data into a Numerical Model

In one of the earlier studies, Carter(1989) [2] examined the process of assimilating data from a set of 39 neutrally buoyant floats that followed the Gulf stream, which measured the location and depth collected three times a day, for 45 days. These RAFOS floats were distributed at approximately ten day intervals during 1984 and 1985. Carter combined this data with the nonlinear shallow water model that describes a single active layer as follows:

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - fv = -\frac{\partial h}{\partial x}, \tag{1.1a}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + fu = -\frac{\partial h}{\partial y}, \tag{1.1b}$$

$$\frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} = -h\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \tag{1.1c}$$

where u and v denote the horizontal velocity components, h describes the geopotential height of the active layer and f is the Coriolis parameter. We can notice that this choice of model physics closely follows the components directly observed by RAFOS floats. The assimilation of data was done using the well known extended Kalman filtering model (chapters 27 to 29, Lewis et al., 2006, [12]). This technique allows for incorporation of the estimate of the field $\mathbf{X}$ at time $t-1$ and observations $\mathbf{Z}$ at time t into the estimate $\mathbf{X}$ at time t. This is accomplished with the use of the following equation

$$\mathbf{X}(t|t) = \mathbf{X}(t|t-1) + \mathbf{K}(t)\left[\mathbf{Z}(t) - \mathbf{H}(t)\mathbf{X}(t|t-1)\right]. \tag{1.2}$$

The matrix $\mathbf{H}$ describes transformation between the observations and the model fields. The Kalman gain, $\mathbf{K}(t)$ is the key element of this equation that captures the relative weights used to assimilate the observations into the current model estimate. It is calculated by

$$\mathbf{K}(t) = \mathbf{P}(t|t-1)\mathbf{H}^T(t)\left[\mathbf{H}(t)\mathbf{P}(t|t-1)\mathbf{H}^T(t) + \mathbf{R}\right]^{-1}. \tag{1.3}$$

The measurement noise covariance is given by matrix $\mathbf{R}$. In case of measurements with independent errors, $\mathbf{R}$ is a diagonal matrix. The analysis covariance matrix updated each time the new measurements are available is given by

$$\mathbf{P}(t|t) = \left[\mathbf{I} - \mathbf{K}(t)\mathbf{H}(t)\right]\mathbf{P}(t|t-1), \tag{1.4}$$

where $\mathbf{I}$ is the identity matrix. The description of the evolution of the physical field in time is captured in the system transition matrix $\mathbf{\Phi}$.

$$\mathbf{X}(t|t-1) = \mathbf{\Phi}\mathbf{X}(t-1|t-1), \tag{1.5}$$

Analogously, the forecast covariance matrix evolves in time

$$\mathbf{P}(t|t-1) = \mathbf{\Phi}\mathbf{P}(t-1|t-1)\mathbf{\Phi}^T + \mathbf{Q}, \tag{1.6}$$

where $\mathbf{Q}$ represents the covariance of the model errors. We have to note that $\mathbf{Q}$ is usually not known accurately. Carter in his paper stresses the importance of choosing a numerical method that would lend itself to Kalman filter application (chapters 27 to 29, Lewis et al., 2006, [12]). The preferred numerical methods that are used with the Kalman filter should not increase the size of the state vector

by using more than two fields for two time steps. Carter noticed that observation influences only a very limited region and that they have a different impact on the overall improvement to the forecast. He has also reported a difference whether observations are taken and assimilated earlier or later during the model evolution. It is also noted in the paper that one has to deal with the problem of inertia-gravity waves that can be excited by the data insertion into the model. This has to be taken under consideration during the design of the Kalman filter.

1.2.2 Assimilation of drifter observations for the reconstruction of the Eulerian circulation field

Molcard et al., (2003) [14] using a quasi-geostrophic reduced gravity model equations in a twin experiment set up generated circulation related data and developed an assimilation scheme that is based on the classical optimum interpolation (OI) technique (chapter 19, Lewis et al., 2006 [12]) that follows the general Bayesian theory. The quasi-geostrophic reduced gravity model is given by

$$\frac{\partial q}{\partial t} + J\left(\psi, q\right) = \frac{f_0}{H} w_E + \nu \nabla^4 \psi - r \nabla^2 \psi, \tag{1.7}$$

where the potential vorticity q is given by

$$q = \nabla^2 \psi + \beta y - \frac{1}{R_d^2} \psi. \tag{1.8}$$

The geostrophic stream function is denoted by ψ, f_0 gives the Coriolis parameter at a reference latitude, β is the meridional gradient of the Coriolis parameter, the radius of deformation is $R_d = \sqrt{g'H}/f_0$, where g' is the reduced gravity, H is the layer depth, w_E is the Ekman velocity field proportional to the wind stress curl, ν

denotes the horizontal eddy viscosity, r is the interfacial friction coefficient. The Jacobian operator is given by $J(\psi, q) = \frac{\partial \psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial q}{\partial x}\frac{\partial \psi}{\partial y}$.

In their paper, they treat the model and observations as equal contributors and try to find their linear combination to represent the true field. Introducing the derivative of the model-to-observation functional $\mathbf{G}$ defined as follows

$$\mathbf{G} = \frac{\partial \mathbf{H}(\mathbf{u}^b)}{\partial \mathbf{u}^b}, \tag{1.9}$$

and variables: $\mathbf{u}^a$ as the model velocity vector after assimilation, $\mathbf{u}^b$ as the model velocity vector before assimilation, $\mathbf{y}$ representing the vector of observations, $\mathbf{H}(\mathbf{u}^b)$ as the functional that relates model state variables to the observations, $\mathbf{R}^0$ as the observation error covariance matrix, and $\mathbf{R}^b$ is the covariance matrix of the model uncertainty, they use the following equation to calculate the assimilated vector field.

$$\mathbf{u}^a = \mathbf{u}^b + \mathbf{R}^b\mathbf{G}^T\left(\mathbf{G}\mathbf{R}^b\mathbf{G}^T + \mathbf{R}^0\right)^{-1}\left(\mathbf{y} - \mathbf{H}(\mathbf{u}^b)\right) \tag{1.10}$$

Superscript T indicates transposition. Equation (1.10) is optimal under several conditions:

- The true vector $\mathbf{u}$ has the prior distribution that is Gaussian with mean $\mathbf{u}^b$ and covariance $\mathbf{R}^b$.

- The observation vector $\mathbf{y}$ has also the Gaussian distribution with the mean $\mathbf{H}(\mathbf{u})$ and covariance $\mathbf{R}^0$. It is assumed that the observation vector has error characterized by $\mathbf{R}^0$ and that the is error depends on instrument resolution and accuracy.

- It is assumed that the functional $\mathbf{H}(\mathbf{u})$ is linear, which may hold true only locally. This condition is often not met in case of nonlinear problems. It can be noted that there is an analogy between (1.10) and the extended Kalman filter.

Their assimilation algorithm follows M Lagrangian particles released at the same initial time $t = 0$ from different positions $\mathbf{r}_1^0, \mathbf{r}_2^0, \ldots, \mathbf{r}_M^0$ at the same plane. Motion of these particles can be described as

$$\begin{aligned} \frac{d\mathbf{r}_m}{dt} &= \mathbf{u}(t, \mathbf{r}_m),\ \mathbf{r}_m(0) = \mathbf{r}_m^0,\ m = 1 \ldots M, \\ \mathbf{v}_m(t) &= \frac{d\mathbf{r}_m}{dt}. \end{aligned} \tag{1.11}$$

Here, $\mathbf{u}(t, \mathbf{r})$ represents the Eulerian velocity field, while $\mathbf{v}_m(t)$ stands to represent the horizontal Lagrangian velocity of the m-th particle. Trajectories of Lagrangian particles are measured at discrete times equal to $n\Delta t$, $n = 1, \ldots, N$. These observations are denoted as $\mathbf{r}_m^0(n)$. Their model counterparts are represented as $\mathbf{r}_m^b(n)$. In their paper, Molcard et al., (2003) [14] introduce finite difference Lagrangian velocity, both for the observations and model

$$\begin{aligned} \mathbf{v}_m^0(n) &= \frac{\Delta \mathbf{r}_m^0}{\Delta t} = \frac{\mathbf{r}_m^0(n) - \mathbf{r}_m^0(n-1)}{\Delta t}, \\ \mathbf{v}_m^b(n) &= \frac{\Delta \mathbf{r}_m^b}{\Delta t} = \frac{\mathbf{r}_m^b(n) - \mathbf{r}_m^b(n-1)}{\Delta t}. \end{aligned} \tag{1.12}$$

There is an assumption made that the frequency of measurements is high enough to capture the spatial gradients of the current. The zero-order assimilation for-

mulas used are as follows (from equation (5) Molcard et al. 2003):

$$
\begin{aligned}
u_{ij}^{a}(n) &= u_{ij}^{b}(n) + \alpha^{-1} \sum_{m=1}^{M} \gamma_{ijm} \Big(u_{m}^{0}(n) - u_{m}^{b}(n) \Big), \\
v_{ij}^{a}(n) &= v_{ij}^{b}(n) + \alpha^{-1} \sum_{m=1}^{M} \gamma_{ijm} \Big(v_{m}^{0}(n) - v_{m}^{b}(n) \Big).
\end{aligned}
\tag{1.13}
$$

Here, variables u, v with the single subscript (m) denote the Lagrangian velocity component of the m-th drifter, while the velocity components u, v with subscripts (ij) represent the Eulerian velocities at the corresponding grid point, and coefficient γ approximates the delta function from the derivative of the Gauss function

$$\gamma_{ijm} = E_h \left(x_m^b(n) - ih, y_m^b(n) - jh \right), \tag{1.14}$$

$$E_h\left(x, y\right) \equiv \exp\left(-\frac{x^2}{2h^2} - \frac{y^2}{2h^2} \right), \tag{1.15}$$

$$\alpha = 1 + \frac{\sigma_o^2}{\sigma_b^2}, \;\; \text{where} \;\; \sigma_o^2 = \frac{\sigma_r^2}{\Delta t^2}, \tag{1.16}$$

where σ_b^2 is the modeling velocity mean square error and σ_o^2 denotes the error of the Lagrangian velocity related to the error σ_r^2 (the error of independent position $\sigma_o^2 = \sigma_r^2/\Delta t^2$). It is assumed that model and observed variables have errors that are uncorrelated in space and time. It is worth noting that equation (1.13) takes only two successive data points, that is only one time step Δt; the particle path is not represented, just two consecutive locations; authors indicated that the more advanced algorithms that focus on complete path information are possible. In this formulation, the position of the Lagrangian particle is converted into the Lagrangian velocity v information along the particle trajectory that is averaged over sampling time Δt by using two endpoints of the particle path and converting it into v^b.

Using this methodology, Molcard et al., (2003) examined sensitivity to the sampling period Δt, since the sampling periods for Lagrangian tracers varies from minutes to weeks. Experiments for $\Delta t = 2, 5$ and 10 days were conducted. They all showed trajectories improved by the process of data assimilation, with the error being the smallest for $\Delta t = 2$ days. For $\Delta t > 2$ days, Molcard proposed repeating the assimilation procedure in an iterative way, which gave improved results when compared to only one assimilation run. Sensitivity to model forcing was also investigated, since control runs and assimilation runs were subject to a difference in a wind forcing. It was shown that this assimilation technique is effective even in a presence of large errors in the wind forcing influencing the ocean circulation model. In addition, numerical experiments were conducted to investigate the sensitivity of the assimilation to the number of drifters that was varied (9, 16, 36, 49, 100, 144 and 196); it was noted that to a naked eye, difference is not very noticeable for runs where the number of drifters is larger than 25 over the area under their study. Lastly, sensitivity to the initial distribution of drifters was addressed by Molcard et al., (2003). For this purpose, 25 drifters were distributed over the domain. Their impact was higher when they were placed in the areas with high average kinetic energy of the subdomain where they were released, when averaged by the total kinetic energy of the full domain. But even then, data assimilation experiments showed high sensitivity to their launch location. Overall, the importance of the initial sampling location is high; a homogeneous sampling is less efficient than a sampling that is aimed at the high energy regions of the real ocean. This problem is even more complex in the real ocean, where oftentimes, drifters are advected away from the energetic regions.

1.2.3 Assimilation of drifter observations in primitive equation models of midlatitude ocean circulation

Özgökmen et al., (2003) [15] examined the use of Kalman filter based method to assimilate drifter observations into the Miami Isopycnic Coordinate Ocean Model (MICOM). It is a comprehensive large-scale ocean model, that is an idealized reduced-gravity layered primitive equation model. They describe an effort to assimilate the Lagrangian data (the drifter position $\mathbf{r}$) into ocean general circulation model (OGCM) to correct the Eulerian surface velocity field $\mathbf{u}$. This is done at the time interval Δt. In the so called "Pseudo-Lagrangian" approach, this is done by approximation of the Eulerian field by $\Delta\mathbf{r}/\Delta t$, assuming that the sampling period is much smaller than the Lagrangian correlation timescale (order of magnitude of time scale in which the flow forgets its past behavior). Problem arises when the sampling time is not much smaller. In their work, Özgökmen et al., (2003) [15] build on the work of Molcard et al., (2003) [14], by extending the Lagrangian assimilation procedure to primitive equations and comparing the Lagrangian and the Pseudo-Lagrangian assimilation techniques. The ocean model used for their work is a comprehensive large-scale ocean model MICOM, in its reduced-gravity version. The momentum and the layer-thickness conservation equations are as follows:

$$\frac{\partial u}{\partial t} + \mathbf{u}\cdot\nabla u - fv = -g'\frac{h}{x} + \frac{1}{\rho}\frac{\partial \tau^x}{\partial z} + \frac{\nu_H}{h}\nabla\cdot h\nabla u, \tag{1.17a}$$

$$\frac{\partial v}{\partial t} + \mathbf{u}\cdot\nabla v + fu = -g'\frac{h}{y} + \frac{1}{\rho}\frac{\partial \tau^y}{\partial z} + \frac{\nu_H}{h}\nabla\cdot h\nabla v, \tag{1.17b}$$

$$\frac{\partial h}{\partial t} + \nabla\cdot(\mathbf{u}h) = 0. \tag{1.17c}$$

Here, h is the thickness of a layer of a constant density, $\mathbf{u} = (u, v)$ is the layer-

averaged horizontal velocity vector, $g' = g\frac{\delta\rho}{\rho}$ denotes the reduced gravity, g is gravitational acceleration, f is the Coriolis parameter expressed with the β plane approximation given by $f = f_0 + \beta y$, the lateral viscosity coefficient is represented by ν_H, and the wind stress vector is given by $\tau = (\tau^x, \tau^y)$. After rewriting the momentum equations (1.17a) and (1.17b), the geostrophic momentum balance is represented by

$$\begin{aligned} T^x - f\Delta v &= -g'\frac{\partial(\Delta h)}{\partial x}, \\ T^y + f\Delta v &= -g'\frac{\partial(\Delta h)}{\partial y}. \end{aligned} \tag{1.18}$$

This is done to use the velocity correction $(\Delta u, \Delta v)$ from the Lagrangian drifters to find the Δh correction to the layer thickness that approximately satisfies the momentum equation. Here (T^x, T^y) represent the ageostrophic momentum terms in each direction; they capture the acceleration related to time dependence, nonlinearity, forcing, and dissipation. After simplifications due to scale analysis, the correction to layer thickness is calculated from geostrophic balance

$$\nabla^2\left(\Delta h_g\right) = \frac{f}{g'}\left[\frac{\partial(\Delta v)}{\partial x} - \frac{\partial(\Delta u)}{\partial y}\right], \tag{1.19}$$

with the homogeneous boundary condition $\Delta h_g \to 0$. The mass conservation constrains require that the correction to the layer thickness depends on the correctional velocity field. Here we see that there is a strong dependence on data distribution, since correction to (1.18) will be mainly done at the location of drifters. Therefore the net mass deviation stemming from (1.18) is calculated by

$$\Delta m = \frac{1}{|A|}\int_A \Delta h_g dA, \tag{1.20}$$

where A represents the ocean basin with $|A|$ being its area. The final correction

is given by $\Delta h = \Delta h_g - \Delta m$. This leads to the equation for the assimilation of the layer thickness as

$$h^a(n) = h^b(n) + \Delta h. \tag{1.21}$$

Here, the superscript a denotes assimilated data, and the superscript b denotes the background (predicted) field.

Data assimilation part follows an array of M Lagrangian trajectories at times $n\Delta t = 1, 2, \ldots N$. Observations are indicated by $\mathbf{r}_m^0(n)$, while model values are indicated by $\mathbf{r}_m^b(n)$ with $m = 1, \ldots M$. We can express the Lagrangian velocity calculated from observations and the model by the following finite differences:

$$\begin{aligned} \mathbf{v}_m^0(n) &= \frac{\Delta \mathbf{r}_m^0}{\Delta t} = \frac{\mathbf{r}_m^0(n) - \mathbf{r}_m^0(n-1)}{\Delta t}, \\ \mathbf{v}_m^b(n) &= \frac{\Delta \mathbf{r}_m^b}{\Delta t} = \frac{\mathbf{r}_m^b(n) - \mathbf{r}_m^b(n-1)}{\Delta t}. \end{aligned} \tag{1.22}$$

At the same time, the Eulerian velocity obtained from the grid is

$$\mathbf{u}_{ij} = \mathbf{u}(n\Delta t, i\Delta r, j\Delta r), \tag{1.23}$$

where Δr is the grid scale and n is the time index. The assimilation equations for the velocity are given by:

$$\mathbf{u}_{ij}^a(n) = \mathbf{u}_{ij}^b(n) + \beta \sum_{m=1}^{M} \gamma_{ijm} \Big(\mathbf{v}_m^0(n) - \mathbf{v}_m^b(n) \Big), \tag{1.24}$$

The velocity components u withs subscripts (ij) represent the Eulerian velocities

at the corresponding grid point, and the assimilation coefficients β and γ are

$$\gamma_{ijm} = \exp\left(-\frac{\left(x_m^b(n) - i\Delta r\right)^2 + \left(y_m^b(n) - j\Delta r\right)^2}{2\Delta r^2}\right), \qquad \beta = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2}, \text{ and } \sigma_o^2 = \frac{\sigma_r^2}{\Delta t^2}, \tag{1.25}$$

where σ_b^2 is the modeling velocity mean square error and σ_o^2 denotes the error of the Lagrangian velocity related to the error σ_r^2 (the error of independent position $\sigma_o^2 \sim \sigma_r^2/\Delta t^2$). It is assumed that model and observed variables have errors that are uncorrelated in space and time, and that the velocity spatial gradients are small in relation to $(\Delta t)^{-1}$. And so, the optimal interpolation based Lagrangian data assimilation is focused on calculating the velocity correction $\Delta\mathbf{u} = (\Delta u, \Delta v)$ field used in (1.24) given by

$$\Delta\mathbf{u}_{ij} = \beta \sum_{m=1}^{M} \gamma_{ijm} \left(\mathbf{v}_m^0(n) - \mathbf{v}_m^b(n)\right) \tag{1.26}$$

and then by applying it as in input to the layer thickness field equation (1.18) to (1.25); this is done following the multivariate dynamic balancing approach.

Özgökmen et al., (2003) [15] have found that Pseudo-Lagrangian and Lagrangian assimilation gives similar results only for $\Delta t \ll T_L$. For bigger Δt values that are between $(T_L/5 \leq \Delta t \leq T_L/2)$, Lagrangian approach is better than Pseudo-Lagrangian. When $\Delta t \geq T_L$, neither of these methods can retrieve useful Eulerian information, since Lagrangian predictability limits are surpassed. In order to resolve this issue, Lagrangian assimilation has to be set in the primitive equations. Twin experiments have shown that the simple dynamical balancing technique developed in this paper, that corrects the model velocity field, and then

in term, corrects the layer thickness gives a positive result.

1.2.4 A method for assimilation of Lagrangian data

Kuznetsov et al. (2003) [7] examined the use of extended Kalman filter methodology (EKF) to assimilate Lagrangian tracer data. Majority of models in oceanography and meteorology use a fixed grid in space. To mesh Lagrangian data with the Eulerian variables computed in a set grid, they have proposed to augment the state space by inclusion of the tracer coordinates as an extra data, and then to apply the EKF to this dynamics. This augmented state vector $\mathbf{x} = (\mathbf{x}_F, \mathbf{x}_D)$ consists of the Eulerian part $\mathbf{x}_F$ (the state of the flow), and the Lagrangian part $\mathbf{x}_D$ (coordinates of the drifters). This approach, in which drifter information is a part of the dynamical model, allows for tracking drifters along with their correlation with the flow. Positions of N_D particles are observed at the regular times and assimilated into the model. And so, the state vector $\mathbf{x}(t)$ has a dimension N. Its dimension is a product of the number of these variables and the number of the discretization elements that include Fourier modes, gird points, etc. The evolution of the state vector can by generalized as follows:

$$\frac{d\mathbf{x}^f}{dt} = M(\mathbf{x}^f, t), \tag{1.27}$$

where $\mathbf{x}^f$ corresponds to the forecasted state, whereas the true state is denoted as $\mathbf{x}^t$, and M represents the dynamic operator. Since (1.27) is usually not a closed system, to represent the subgrid-scale processes we can represent the dynamics in a discrete form as a stochastic system

$$d\mathbf{x}^t = M(\mathbf{x}^t, t)dt + \boldsymbol{\eta}^t(t)dt. \tag{1.28}$$

All of the unresolved processes are described by $\boldsymbol{\eta}$ as a zero-mean Gaussian white noise. In a process of the sequential assimilation, the model is updated each time the observation ($\mathbf{x}_j^t \equiv \mathbf{x}^t(t_j)$) is given at a given time t_j. Observations $\mathbf{y}_j^o$ with their uncorrelated zero-mean Gaussian errors $\boldsymbol{\epsilon}_j^t$ (with a covariance $E[\boldsymbol{\epsilon}_j^t(\boldsymbol{\epsilon})_{\boldsymbol{j}}^{\boldsymbol{t}^T}]$), and their vector $\mathbf{H}_j$ of observation functions can be expressed in terms of

$$\mathbf{y}_j^o = \mathbf{H}_j[\mathbf{x}_j^t] + \boldsymbol{\epsilon}_j^t, \tag{1.29}$$

The number of observations $\mathbf{y}_j^o$, say L_j, can be different for each update time. Data assimilation with the Kalman filter focuses on tracking the evolution of the error covariance matrix defined as

$$\mathbf{P}_f \equiv E[(\mathbf{x}^f - \mathbf{x}^t)(\mathbf{x}^f - \mathbf{x}^t)^T]. \tag{1.30}$$

This is done with the tangent linear model (TLM) with the linearized dynamic operator calculated at $\mathbf{x}^f(t)$ by the following formula

$$\mathbf{M} \equiv \frac{\partial M(\mathbf{x}^t, t)}{\partial \mathbf{x}}, \tag{1.31}$$

and expressing the evolution of the error covariance matrix by

$$\frac{d\mathbf{P}^f}{\partial t} = \mathbf{M}\mathbf{P}^f + (\mathbf{M}\mathbf{P}^f)^T + \mathbf{Q}(t), \tag{1.32}$$

where $\mathbf{Q}(t)$ symbolizes our estimation for system noise covariance $\mathbf{Q}^t(t)$. Data assimilation proceeds by minimization of the mean-square error

$$\mathrm{tr}\mathbf{P}_j^a = E[(\mathbf{x}_j^a - \mathbf{x}_j^t)^T(\mathbf{x}_j^a - \mathbf{x}_j^t)] \tag{1.33}$$

at each update time t_j, where $\mathbf{x}^a$ represents the analysis state. We are using the model error covariance matrix forecasted with the equation (1.32) to get the first-order approximation of the optimal analysis state. The update combines the predicted model state and the innovation vector with the size L_j

$$\begin{aligned} \mathbf{x}_j^a &= \mathbf{x}_j^f + \mathbf{K}_j \mathbf{d}_j, \\ \mathbf{d}_j &\equiv \mathbf{y}_j^o - \mathbf{H}_j[\mathbf{x}_j^f], \\ \mathbf{K}_j &= \mathbf{P}_j^f \mathbf{H}_j^T \left(\mathbf{H}_j \mathbf{P}_j^f \mathbf{H}_j^T + \mathbf{R}_j \right)^{-1}, \\ \mathbf{H}_j &= \frac{\partial \mathbf{H}_j}{\partial \mathbf{x}} \end{aligned} \tag{1.34}$$

where $\mathbf{H}_j$ is the linearized observation matrix, and our estimation of the covariance matrix of the observation error is given by $\mathbf{R}_j$. The Kalman gain matrix is $\mathbf{K}_j$, and the updated error covariance matrix is given by

$$\mathbf{P}_j^a = (\mathbf{I} - \mathbf{K_j H_j}) \mathbf{P}_j^f. \tag{1.35}$$

In their paper, Lagrangian tracers information ($\mathbf{x}_D$) is combined with the model flow ($\mathbf{x}_F$) into one model state vector $\mathbf{x} = (\mathbf{x}_F, \mathbf{x}_D)^T$. The advection of tracers is followed as

$$\frac{d\mathbf{x}_F^f}{dt} = M(\mathbf{x}_F^f, t), \tag{1.36}$$

$$\frac{d\mathbf{x}_D^f}{dt} = M(\mathbf{x}_D^f, \mathbf{x}_F^f, t). \tag{1.37}$$

Equation (1.36) states the model in it original form stated in (1.27), and (1.37)

describes dynamics of the Lagrangian tracers. The operator (1.31) takes form

$$\mathbf{M} = \begin{pmatrix} \mathbf{M_{FF}} & 0 \\ \mathbf{M_{DF}} & \mathbf{M_{DD}} \end{pmatrix}, \tag{1.38}$$

and it has dimension $(N + L) \times (N + L)$. Similarly, the error covariance matrix is given by

$$\mathbf{P} = \begin{pmatrix} \mathbf{P_{FF}} & \mathbf{P_{FD}} \\ \mathbf{P_{DF}} & \mathbf{P_{DD}} \end{pmatrix}. \tag{1.39}$$

We can note that $\mathbf{M_{FF}}$ and $\mathbf{P_{FF}}$ are $(N \times N)$ matrices, while $\mathbf{M_{DD}}$ and $\mathbf{P_{DD}}$ are $(L \times L)$ matrices, their sizes differ a lot, since $N \gg L$ in real applications. Therefore, the addition of the Lagrangian model variables does not increase the overall computational cost. The Kalman gain matrix gets simplified since the observation function $\mathbf{H}$ is linear

$$\mathbf{K} = \begin{pmatrix} \mathbf{P_{FD}} \\ \mathbf{P_{DD}} \end{pmatrix} (\mathbf{P_{DD}} + \mathbf{R})^{-1}, \tag{1.40}$$

For their physical model, authors have used the Euler's equation singular solution depicting point vortices. Point vortex flows can be used to model 2D flows that contain strong coherent vortices. The state vector of a flow with N_F point vortices has a dimension $N = 2N_F$, since it follows their positions on the plane $[x_m(t), y_m(t)]$, $m = 1, \ldots, N_F$. In the complex coordinates $x_m(t) = x_m(t) + \mathrm{i}y_m(t)$, the complex valued state vector is $\mathbf{z} \in \mathbb{C}^{N_F}$.

$$\frac{dz_m^t}{dt} = \frac{\mathrm{i}}{2\pi} \sum_{l=1, l\neq m}^{N_F} \frac{\Gamma_l}{z_m^{t\,\star} - z_l^{t\star}} + \eta_{F_m}(t), \; (m = 1, \ldots, N_F) \tag{1.41}$$

Here, Γ_l is circulation of vortex l, and the last term $\boldsymbol{\eta}_F = \boldsymbol{\eta}_F^{(x)} + \mathrm{i}\boldsymbol{\eta}_F^{(y)}$ captures the unresolved processes.

$$\begin{aligned} E[\boldsymbol{\eta}_F^{(x,y)}(t)\boldsymbol{\eta}_F^{(x,y)}(t')] &= \sigma^2\delta(t-t')\mathbf{I}, \\ E[\boldsymbol{\eta}_F^{(x)}\boldsymbol{\eta}_F^{(y)}] &= 0. \end{aligned} \tag{1.42}$$

Observations are provided by coordinates of the N_D Lagrangian tracers. Here, their coordinates are represented on the complex plane $\boldsymbol{\zeta} \in \mathbb{C}^{N_F}$ with their noise as $\boldsymbol{\eta}_D = \boldsymbol{\eta}_D^{(x)} + \mathrm{i}\boldsymbol{\eta}_D^{(y)}$ with zero mean. The advection of N_D tracers is described by the following equation

$$\frac{d\zeta_m^t}{dt} = \frac{\mathrm{i}}{2\pi}\sum_{l=1}^{N_F}\frac{\Gamma_l}{\zeta_m^{t\,\star} - z_l^{t\star}} + \eta_{D_m}(t),\ (m = 1, \ldots, N_D) \tag{1.43}$$

Analogously to the flow variables, the covariance of tracers position has

$$\begin{aligned} E[\boldsymbol{\eta}_D^{(x,y)}(t)\boldsymbol{\eta}_D^{(x,y)}(t')] &= \sigma^2\delta(t-t')\mathbf{I}, \\ E[\boldsymbol{\eta}_D^{(x)}\boldsymbol{\eta}_D^{(y)}] &= 0. \end{aligned} \tag{1.44}$$

The state vector contains all the model equations with all corresponding dynamical variables:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{z} \\ \boldsymbol{\zeta} \end{pmatrix}, \boldsymbol{x} \in \mathbb{C}^{N_F+N_D} \tag{1.45}$$

The flow variables $\mathbf{x}_F$ are corrected by the first N rows of $\mathbf{K}$, which is proportional to the correlation between the flow state and the drifter positions expressed in $\mathbf{P_{FD}}$.

Numerical experiments focused on comparison of the model without data assimilation process, and one with extra observations coming from different number of tracers, and different sampling frequency. Influence of the launch position was

also investigated, in particular, relation of the initial position and separatrices of the streamfunction in the corotating frame.The results point to dependency of the assimilation on the scales of the motion and noise levels. Success in assimilation is inversely proportional to how chaotic the system is, with biggest tracking errors related to chaotic vortex initial condition.

1.2.5 Using flow geometry for drifter deployment in Lagrangian data assimilation

Salman et al., (2008) [18] recently explored the effectiveness of drifter deployment strategies using a nonlinear reduced gravity shallow water model with external wind forcing and combined it with the Ensemble nonlinear Kalman filter technology.

$$\frac{\partial u}{\partial t} = -\mathbf{u}\cdot\nabla u + fv - g'\frac{h}{x} + F^u + \frac{1}{h}\left(\frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y}\right), \tag{1.46a}$$

$$\frac{\partial v}{\partial t} = -\mathbf{u}\cdot\nabla v - fu - g'\frac{h}{y} + \frac{1}{h}\left(\frac{\partial \tau_{yx}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y}\right), \tag{1.46b}$$

$$\frac{\partial h}{\partial t} = -\frac{\partial hu}{\partial x} - \frac{\partial hv}{\partial y}. \tag{1.46c}$$

Here, h is the surface height, (u, v) is the fluid velocity vector, g' is the reduced gravity, F^u is a horizontal wind-forcing acting in a zonal direction

$$F^u = \frac{-\tau_0}{\rho H_0(t)}\cos\left(\frac{2\pi y}{L_y}\right), \tag{1.47}$$

and f is the Coriolis parameter expressed with the β plane approximation given by $f = f_0 + \beta y$, with β and f_0 being constants, τ_0 is the wind stress. To calculate

the average water depth H_0 following equation

$$H_0(t) = \frac{1}{L_x L_y} \int_0^{L_x} \int_0^{L_y} h(x, y, t) dx\, dy \tag{1.48}$$

is used. L_x and L_y are the dimensions of the zonal and the meridional directions. The dissipation terms are given by the following equation

$$\tau_{ij} = \mu h \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \delta_{ij} \frac{\partial u_k}{\partial x_k} \right), \tag{1.49}$$

where indices i and j take all possible permutations, while μ represents a constant eddy viscosity of the flow. One of the ideas described in this paper is search for Lagrangian coherent structures (LCS) in the numerically generated flow, since it has become understood that they are responsible for the evolution of the motion of material particles. This search is quite complicated because there is the non-linear relation between the flow field and the Lagrangian drifter position. Successful estimation of LCS aids placement of launch position for the Lagrangian drifters and this improves data assimilation procedure. The assimilation of Lagrangian data of N_D drifters is conducted with the augmented state vector $\mathbf{x} = (\mathbf{x_F^T}, \mathbf{x_D^T})^T$, where we have $(N_F \times 1)$ equations describing the flow vector $\mathbf{X_F}$, and the $(2N_D \times 1)$ drifter state vector $\mathbf{X_D}$. For the ensemble forecast with N_E members, following equation is used

$$\frac{d\mathbf{x}_j^f}{dt} = \mathbf{m}_j(\mathbf{x}_j, t),\ j = 1, \ldots, N_E. \tag{1.50}$$

Ensemble members are denoted by subscript j, forecast is denoted by superscript f, and $\mathbf{m}_j$ represents the evolution operator. The augmented system of equations

for ensemble members is used to define the covariance matrix as

$$\mathbf{P}^f = \frac{1}{N_E - 1} \sum_{j=1}^{N_E} \left(\mathbf{x}_j^f - \overline{\mathbf{x}^f} \right) \left(\mathbf{x}_j^f - \overline{\mathbf{x}^f} \right)^T \tag{1.51}$$

The mean of the state vector calculated for the ensemble $\overline{x}$ is calculated by

$$\overline{\mathbf{x}^f} = \frac{1}{N_E} \sum_{j=1}^{N_E} \mathbf{x}_j^f. \tag{1.52}$$

The analysis state of the system is defined as

$$\mathbf{x}_j^a = \mathbf{x}_j^f + \mathbf{K}\mathbf{d}_j, \tag{1.53}$$

with Kalman gain matrix given by

$$\mathbf{K} = \begin{pmatrix} \rho_{FD} \circ \mathbf{P}_{FD} \\ \\ \rho_{DD} \circ \mathbf{P}_{DD} \end{pmatrix} \left(\rho_{DD} \circ \mathbf{P}_{DD} + \mathbf{R} \right)^{-1}, \tag{1.54}$$

while the innovation vector $\mathbf{d}_j$ is represented by

$$\mathbf{d}_j = \mathbf{y}^0 - \mathbf{K}\mathbf{x}_{D,j}^f + \tilde{\boldsymbol{\epsilon}}_j^f. \tag{1.55}$$

The observations vector $\mathbf{y}^0$ holds spatial coordinates of the drifters in the zonal and meridional coordinates. The matrix $\boldsymbol{\rho}$ holds distance-dependent correlation function. The noise in the drifter positions is given by $\tilde{\boldsymbol{\epsilon}}_j^f$ and is based on the gaussian distribution with a covariance matrix equal to $\mathbf{R}$ and

$$\frac{1}{N_E} \sum_{j=1}^{N_E} \tilde{\boldsymbol{\epsilon}}_j^f = 0. \tag{1.56}$$

In equation (1.54) operator $\circ$ indicates the Schur product between two matrices.

Four different drifter locations were analyzed: a uniform drifter deployment within the ocean basin, a saddle point launch strategy, a vortex launch strategy, and a mixed combination of saddle and vortex centre launches. Nine drifters were used for all these experiments. The mixed launch produced the highest convergence for the velocity field whereas the uniform and saddle launches achieved the minimal error in the height estimation. They have found that bifurcations of coherent flow structures can lead to rapid dispersion of drifters placed within such coherent vortices; forecasts made over longer time-scales can differ a great deal from forecast over the shorter time.

1.2.6 A Bayesian approach to Lagrangian data assimilation

Apte et al., (2008) [1] described a Bayesian perspective based approach to data assimilation. Their motivation was to follow the issues related to the highly nonlinear characteristics of the Lagrangian data as it influences data assimilation. This becomes a major problem when the time between the observations is large. For their study, they have used an idealized ocean model given by the inviscid linearized shallow water model given by equation (5) in [1]:

$$\begin{aligned} \frac{\partial u}{\partial t} &= v - \frac{\partial h}{\partial x}, \\ \frac{\partial v}{\partial t} &= -u - \frac{\partial h}{\partial y}, \\ \frac{\partial h}{\partial t} &= -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}. \end{aligned} \tag{1.57}$$

They took into the account first two modes of the Fourier modes for their numerical modeling, with the first component describing the geostrophic mode and the second one describing the inertia-gravity mode given by equation (6) in [1].

$$\begin{aligned} u(x,y,t) &= -2\pi l \quad \sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t), \\ v(x,y,t) &= +2\pi k \quad \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t), \\ h(x,y,t) &= \qquad\quad \sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t). \end{aligned} \tag{1.58}$$

After combining (1.57) and (1.58) they have presented the following equations for amplitudes (equation (7) in [1]):

$$\begin{aligned} \dot{u_0} &= 0, \\ \dot{u_1} &= v_1, \\ \dot{v_1} &= -u_1 - 2\pi m h_1, \\ \dot{h_1} &= 2\pi m v_1, \end{aligned} \tag{1.59}$$

along with initial conditions $[u_0(0), u_1(0), v_1(0), h_1(0)]$; $du_0/dt \equiv \dot{u_0}$ is following standard notation, as is $\dot{u_1}$, $\dot{v_1}$ and $\dot{h_1}$.

Apte et al., (2008) described the use of Bayes theorem as applied to a data assimilation problem with a deterministic dynamic model. The initial conditions of the deterministic model with an n-dimensional state vector given by

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}),\ \mathbf{x}(0) = \mathbf{x}_0 \sim \boldsymbol{\xi}, \tag{1.60}$$

are taken from a prior with probability density function $p_\xi(\mathbf{x}_0)$. Noisy observations $\mathbf{y}_k \in \mathbb{R}^m$ represent the state of the system at a specific time t_k. If we use

the solution operator for the dynamics $\mathbf{\Phi}$, state of the system can be expressed as $\mathbf{x}(t) = \mathbf{\Phi}(\mathbf{x}_0, t)$. Therefore, observations $\mathbf{y}_k$ can be written as

$$\mathbf{y}_k = \mathbf{h}[\mathbf{x}(t_k)] + \boldsymbol{\eta}_k = \mathbf{h}[\mathbf{\Phi}(\mathbf{x}_0, t_k)] + \boldsymbol{\eta}_k, \tag{1.61}$$

where operator $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$. This indicates that the observations can be treated as the non-linear numerical functions of the initial conditons. For the set of noisy observations at various times $t_1, \ldots, t_k$, we can express the total observations vector $\mathbf{y}^T = (\mathbf{y}_1^T, \ldots, \mathbf{y}_k^T)$ as

$$\mathbf{y} = \mathbf{H}(\mathbf{x}_0) + \boldsymbol{\eta}. \tag{1.62}$$

We are using $\mathbf{H}(\mathbf{x}_0)^T = \Big(\mathbf{h}[\mathbf{x}(t_1)]^T, \ldots, \mathbf{h}[(\mathbf{x}(t_m)]^T\Big)$ and $\boldsymbol{\eta}^T = (\boldsymbol{\eta}_1^T, \ldots, \boldsymbol{\eta}_m^T)$. The random vector $\boldsymbol{\eta}$ has a probability density function $p_{\boldsymbol{\eta}} : \mathbb{R}^{mK} \to \mathbb{R}$. Therefore, the conditional probability of observations $\mathbf{y}$ related to the initial data $\mathbf{x}_0$ can be described as:

$$p(\mathbf{y}|\mathbf{x}_0) = p_{\boldsymbol{\eta}}[\mathbf{y} - \mathbf{H}(\mathbf{x}_0)]. \tag{1.63}$$

If we have a prior distribution of initial conditions $p_{\boldsymbol{\zeta}}$ along with a realization of the observations $\hat{\mathbf{y}}$, using Bayes' theorem, we can write a posterior probability for the state vector

$$p(\mathbf{x}_0|\hat{y}) = \frac{p_{\boldsymbol{\eta}}[\hat{\mathbf{y}} - \mathbf{H}(\mathbf{x}_0)]p_{\zeta}(\mathbf{x}_0)}{p(\hat{y})}, \tag{1.64}$$

where

$$p(\mathbf{y}) = \int p_{\boldsymbol{\eta}}[\mathbf{y} - \mathbf{H}(\mathbf{x}_0)]p_{\zeta}(\mathbf{x}_0)d\mathbf{x}_0. \tag{1.65}$$

It can be noticed that $p(\mathbf{y})$ is only a function of observations and that $p(\hat{\mathbf{y}})$ has a constant value for a particular realization of $\hat{\mathbf{y}}$ which is independent of $\mathbf{x}_0$.

Apte et al., (2008) have noted several key features of the this approach to data assimilation:

- The conditional distribution at time $t = 0$ of the state of the model in (1.64) observed during time 0 to t_k is the posterior of (1.64).
- To sample the posterior $p(\mathbf{x}_0|\hat{y})$, there is a need for $p_{\boldsymbol{\eta}}$. This functional can contain non-Gaussian errors than can be correlated.
- The problem was investigated under the twin experiments setup. Observations and posterior's sampling are generated with the same model dynamics.
- True state of the model $\mathbf{x}(t)$is never mentioned, since it is never known. Equation (1.62) is different then from $\mathbf{y} = \mathbf{H}(\mathbf{x}(t)) + \boldsymbol{\eta}$, by the fact, that $\mathbf{H}$ in not a function of the 'true' state. Because of the random nature of errors in the data, all that can be estimated is the probabilistic state of the system, and we don't know the best a priori estimate of the system.
- When we present the posterior distribution for the deterministic model, we have a 'strong constrained formulation' in mind. Oftentimes, a 'weak constrained formulation' may be needed in the oceanographical applications.

Numerical experiments described by Apte et al., (2008) focus on a short trajectories that stay in one cell, and on a trajectories that get close to the saddle point and leave the cell. They have found that these instances have different prior distributions. Examination of results showed that ensemble generated by the model gives information about the variability of estimations due to different initial conditions. The posterior distribution is impacted by the model dynamics together with the assimilated observations. When compared to the Ensemble nonlinear Kalman filter, Bayesian data assimilation works better in presence of

bigger interval between observations. It also performs well in presence of the center point. Three different methods were used to sample the posterior: Langevin stochastic differential equation (LSDE), Metropolis adjusted Langevin algorithm (MALA) and Random walk Metropolis-Hastings (FWMH). For the Lagrangian data assimilation, the Metropolis-Hastings methodology gives the best results.

1.3 Summary

Lagrangian data assimilation has a long history in meteorology and oceanography. The above collected papers give a broad overview of different approaches taken over the years by different authors. There are some underlying common threads in all of them. First, data assimilation of the Lagrangian data is a very important part of any modern model that uses data from sensors following the flow to improve the Eulerian forecast. All of the described assimilation schemes are sensitive to time period between measurements. Second, improvement in models that deal with the shallow water models of a different level of complexity and point vortex systems depends on the initial location of tracers which data is used to improve forecast. This launch location has high significance when it comes to being useful for data assimilation. Third, the number of the Lagrangian tracers used for data assimilation has some importance, however, in some cases, adding more sensors above a certain level does not increase the overall quality of improvements to the data assimilation procedures.

1.4 Organization of the Dissertation

In chapter 2, we give an overview of the shallow water model, and scaling assumption that we make in order to linearize it. In chapter 3, we present a low-order model used to derive the explicit expressions for the tracer dynamics which is a system of two first-order, coupled nonlinear, time varying ordinary differential equations. A complete catalog of all of the equilibria and their character is also presented. It is followed by examination of the bifurcation of the tracer dynamics by succinctly summarizing the dependence on the four dimensional parameter set using a simple two dimensional characterization. In chapter 4, we present a data assimilation approach known as the Forward Sensitivity Method. It is followed in chapter 5 by examination of the sensitivity of the shallow water model to the initial conditions and model parameters, known as the control vector. Thorough evaluation of the Data Assimilation experiments is included in chapter 6. Concluding remarks are contained in chapter 7.

Chapter 2

Shallow Water Model

2.1 Introduction

In this chapter, we give a short description of atmospheric and oceanic motion at a scale that is important to work done in this dissertation. We elaborate on scaling assumptions taken in our work, as they are applied to linearized shallow water model.

2.2 Large scale motion

For our study, we are using a model that can be used to describe a motion in the earth's ocean and atmosphere alike. The disparity between horizontal and vertical scales is well known for a large-scale geophysical motions related to the fluid with stable density. We think of the large-scale motion when it is influenced by the earth's rotation. There is a measure one can use to determine the significance of rotation that is known as the Rossby number. To use it, we need to define L to be a characteristic length scale, and U to be a horizontal velocity scale

characteristic of the motion. The angular velocity of the earth' rotation Ω has value of $|\mathbf{\Omega}| = 7.292 \times 10^{-5}$ rad/s. The non-dimensional parameter Ro is usually used to denote Rossby number, and for a large-scale motion at the latitude Φ, it can be expressed as

$$Ro = \frac{U}{(2\Omega \sin \Phi)L} \leq 1.$$

It is worth indicating that only the component of earth's rotation perpendicular to the planetary surface is used in the estimation of the Rossby number. In the atmosphere, vertical scale is of the order of ten kilometers while horizontal scale is of the order of thousand kilometers. Similarly, the depth of the ocean almost never is bigger than six kilometers, and the horizontal extent of major currents systems is usually huge, that is much longer than six kilometers. Therefore, the motion occurs within an relatively thin sheet of fluid, and given the large extent of the horizontal scale of the motion the geometrical constraint produces fluid trajectories which are very flat. Obviously, the motions described in such way apply to cases in which stratification does not play a major role.

In meteorology and oceanography, there is a simple model called *shallow water model* that describes a motion of this type. We can write a set of equations following Daley (page 194, (1991)) [3] that states the shallow water model in cartesian coordinates, where x is the eastward direction, and y is the northward direction.

$$\frac{\partial U}{\partial t} + U\frac{\partial U}{\partial x} + V\frac{\partial U}{\partial y} - fV + g\frac{\partial H}{\partial x} = 0, \tag{2.1a}$$

$$\frac{\partial V}{\partial t} + U\frac{\partial V}{\partial x} + V\frac{\partial V}{\partial y} + fU + g\frac{\partial H}{\partial y} = 0, \tag{2.1b}$$

$$\frac{\partial H}{\partial t} + U\frac{\partial H}{\partial x} + V\frac{\partial H}{\partial y} + H\left(\frac{\partial U}{\partial x} + \frac{\partial V}{\partial y}\right) = 0. \tag{2.1c}$$

Here, U is the eastward wind component, V is the northward wind component, H indicates the height of the free surface of the fluid, g is the gravitational constant while f is the Coriolis parameter. Equations (2.1a) and (2.1b) constitute momentum equations for the eastward and northward components, while equations (2.1c) represents the continuity equation. Our goal is to linearize the shallow water model. Hence, we focus on analysis of the small-amplitude motions. To this end, we introduce the following $U = \tilde{u} + u$, $V = \tilde{v} + v$ and $H = \tilde{h} + h$ into the equations (2.1a)-(2.1c), where $\tilde{u}, \tilde{v}$ and $\tilde{h}$ indicate a base state, and u, v, and h indicate perturbations. We also assume a constant Coriolis parameter f equal f_0; this is known as the mid-latitude *f-plane assumption*.

$$\begin{aligned}
&\frac{\partial(\tilde{u}+u)}{\partial t} + (\tilde{u}+u)\frac{\partial(\tilde{u}+u)}{\partial x} + (\tilde{v}+v)\frac{\partial(\tilde{u}+u)}{\partial y} \\
&\qquad - f(\tilde{v}+v) + g\frac{\partial(\tilde{h}+h)}{\partial x} = 0, && \text{(2.2a)} \\
&\frac{\partial(\tilde{v}+v)}{\partial t} + (\tilde{u}+u)\frac{\partial(\tilde{v}+v)}{\partial x} + (\tilde{v}+v)\frac{\partial(\tilde{v}+v)}{\partial y} \\
&\qquad + f(\tilde{u}+u) + g\frac{\partial(\tilde{h}+h)}{\partial y} = 0, && \text{(2.2b)} \\
&\frac{\partial(\tilde{h}+h)}{\partial t} + (\tilde{u}+u)\frac{\partial(\tilde{h}+h)}{\partial x} + (\tilde{v}+v)\frac{\partial(\tilde{h}+h)}{\partial y} \\
&\qquad + (\tilde{h}+h)\left(\frac{\partial(\tilde{u}+u)}{\partial x} + \frac{\partial(\tilde{v}+v)}{\partial y}\right) = 0. && \text{(2.2c)}
\end{aligned}$$

Now, we can think of a basic state in which fluid is at rest ($\tilde{u} = \tilde{v} = 0$) and has a free surface with height $\tilde{h}$ that is independent of the position and time:

$$\begin{aligned}
&\frac{\partial\tilde{u}}{\partial t} = \frac{\partial\tilde{v}}{\partial t} = \frac{\partial\tilde{h}}{\partial t} = 0, \\
&\frac{\partial\tilde{h}}{\partial x} = \frac{\partial\tilde{h}}{\partial y} = 0.
\end{aligned}$$

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} - fv + g\frac{\partial h}{\partial x} = 0, \tag{2.3a}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + fu + g\frac{\partial h}{\partial y} = 0, \tag{2.3b}$$

$$\frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + v\frac{\partial h}{\partial y} + \tilde{h}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0. \tag{2.3c}$$

Below we include a shallow water model as stated in Pedlosky([16] page 68.) representing a deterministic dynamical framework appropriate for the motion calculations of large space and time scales; motions of atmospheric and oceanic relevance; this is the set of equations (2.3a)-(2.3a) in which we have dropped the nonlinear terms $u\frac{\partial u}{\partial x}$, $v\frac{\partial u}{\partial y}$, $u\frac{\partial v}{\partial x}$, $v\frac{\partial v}{\partial y}$, $u\frac{\partial h}{\partial x}$, $v\frac{\partial h}{\partial y}$, because $\frac{\partial u}{\partial t} \gg u\frac{\partial u}{\partial x}, v\frac{\partial u}{\partial y}$ and similarly $\frac{\partial v}{\partial t} \gg u\frac{\partial v}{\partial x}, v\frac{\partial v}{\partial y}$. We also can observe that $\tilde{h} \gg h$.

$$\frac{\partial u}{\partial t} - fv = -g\frac{\partial h}{\partial x}, \tag{2.4a}$$

$$\frac{\partial v}{\partial t} + fu = -g\frac{\partial h}{\partial y}, \tag{2.4b}$$

$$\frac{\partial h}{\partial t} + \tilde{h}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 0. \tag{2.4c}$$

The velocity has horizontal components u and v, f denotes the Coriolis parameter, h indicates the height of the surface of the fluid above the reference level $\tilde{h}$.

2.3 Scalling assumptions

Using non-dimensional variables $\hat{h}, \hat{u}, \hat{v}, \hat{t}, \hat{x}, \hat{y}$ we can define

$$\begin{aligned}
&u = u_0\hat{u}, \text{ where } u_0 \sim \mathcal{O}(1\,\mathrm{cm\,s^{-1}}) \sim \mathcal{O}(1\times10^{-2}\,\mathrm{m\,s^{-1}}) \sim \mathcal{O}(0.036\,\mathrm{km\,h^{-1}}),\\
&v = u_0\hat{v},\\
&h = h_0\hat{h}, \text{ where } h_0 \sim \mathcal{O}(1\times10^{-1}\,\mathrm{m}),\\
&t = \hat{t}/f, \text{ where } f \sim \mathcal{O}(1\times10^{-4}\,\mathrm{s^{-1}}),\\
&x = L\hat{x}, \text{ where } L \sim \mathcal{O}(1\times10^{6}\,\mathrm{m}),\\
&y = L\hat{y}.
\end{aligned} \tag{2.5}$$

We can now consider (2.4a)

$$\begin{aligned}
\frac{\partial u}{\partial t} &= \frac{\partial u_0\hat{u}}{\partial \hat{t}/f} = (u_0 f)\,\frac{\partial \hat{u}}{\partial \hat{t}},\\
\frac{\partial h}{\partial x} &= \frac{\partial h_0\hat{h}}{\partial L\hat{x}} = \left(\frac{h_0}{L}\right)\frac{\partial \hat{h}}{\partial \hat{x}},
\end{aligned} \tag{2.6}$$

and so, (2.4a) becomes

$$f u_0\frac{\partial \hat{u}}{\partial \hat{t}} - f u_0\hat{v} = -\frac{g h_0}{L}\frac{\partial \hat{h}}{\partial \hat{x}}, \tag{2.7}$$

and finally

$$\frac{\partial \hat{u}}{\partial \hat{t}} - \hat{v} = -\frac{g h_0}{f L u_0}\frac{\partial \hat{h}}{\partial \hat{x}} = N\frac{\partial \hat{h}}{\partial \hat{x}}, \text{ where } N = \frac{g h_0}{f L u_0}. \tag{2.8}$$

Since $g \sim \mathcal{O}(10\,\mathrm{m\,s^{-2}})$ and f is fixed at $g \sim 10^{-4}$,

$$N = \frac{1\times10^{1}\,\mathrm{m\,s^{-2}}\times1\times10^{-1}\,\mathrm{m}}{1\times10^{-4}\,\mathrm{s^{-1}}\times1\times10^{6}\,\mathrm{m}\times1\times10^{-2}\,\mathrm{m\,s^{-1}}} = 1. \tag{2.9}$$

And this gives us scaled equation (2.4a) as

$$\frac{\partial \hat{u}}{\partial \hat{t}} - \hat{v} = -\frac{\partial \hat{h}}{\partial \hat{x}}. \tag{2.10}$$

Similarly, now we can now consider (2.4b) and express its compnents as

$$\begin{aligned} \frac{\partial v}{\partial t} &= \frac{\partial u_0 \hat{v}}{\partial \hat{t}/f} = (u_0 f)\, \frac{\partial \hat{v}}{\partial \hat{t}}, \\ \frac{\partial h}{\partial y} &= \frac{\partial h_0 \hat{h}}{\partial L \hat{y}} = \left(\frac{h_0}{L}\right) \frac{\partial \hat{h}}{\partial \hat{y}}, \end{aligned} \tag{2.11}$$

and so, (2.4b) becomes

$$f u_0 \frac{\partial \hat{v}}{\partial \hat{t}} + f u_0 \hat{u} = -\frac{g h_0}{L} \frac{\partial \hat{h}}{\partial \hat{y}}, \tag{2.12}$$

and finally, using N defined in (2.8)

$$\begin{aligned} \frac{\partial \hat{v}}{\partial \hat{t}} + \hat{u} &= -\frac{g h_0}{f L u_0} \frac{\partial \hat{h}}{\partial \hat{y}}, \\ \frac{\partial \hat{v}}{\partial \hat{t}} + \hat{u} &= N \frac{\partial \hat{h}}{\partial \hat{y}}. \end{aligned} \tag{2.13}$$

And this gives us equation (2.4b) scaled as

$$\frac{\partial \hat{v}}{\partial \hat{t}} + \hat{u} = -\frac{\partial \hat{h}}{\partial \hat{y}}. \tag{2.14}$$

Finally, we can consider parts of equation (2.4c)

$$\begin{aligned} \frac{\partial h}{\partial t} &= \frac{\partial h_0 \hat{h}}{\partial \hat{t}/f} = (h_0 f)\, \frac{\partial \hat{h}}{\partial \hat{t}}, \\ \tilde{h}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) &= \tilde{h}\left(\frac{\partial u_0 \hat{u}}{\partial L \hat{x}} + \frac{\partial u_0 \hat{v}}{\partial L \hat{y}}\right) = \frac{\tilde{h} u_0}{L}\left(\frac{\partial \hat{u}}{\partial \hat{x}} + \frac{\partial \hat{v}}{\partial \hat{y}}\right), \end{aligned} \tag{2.15}$$

We can combine the above to get

$$\left(\frac{fLh_0}{\tilde{h}u_0}\right)\frac{\partial\hat{h}}{\partial\hat{t}}+\left(\frac{\partial\hat{u}}{\partial\hat{x}}+\frac{\partial\hat{v}}{\partial\hat{y}}\right)=M\frac{\partial\hat{h}}{\partial\hat{t}}+\left(\frac{\partial\hat{u}}{\partial\hat{x}}+\frac{\partial\hat{v}}{\partial\hat{y}}\right)=0, \quad \text{where } M=\frac{fLh_0}{\tilde{h}u_0}. \tag{2.16}$$

Let us analyze M from the above equation using $\tilde{h}\sim\mathcal{O}(1\times10^3\,\mathrm{m})$.

$$M=\frac{1\times10^{-4}\,\mathrm{s}^{-1}\times1\times10^6\,\mathrm{m}\times1\times10^{-1}\,\mathrm{m}}{1\times10^3\,\mathrm{m}\times1\times10^{-2}\,\mathrm{m\,s}^{-1}}=1. \tag{2.17}$$

Finally this gives us (2.4c)

$$\frac{\partial\hat{h}}{\partial\hat{t}}+\left(\frac{\partial\hat{u}}{\partial\hat{x}}+\frac{\partial\hat{v}}{\partial\hat{y}}\right)=0. \tag{2.18}$$

And so, we can combine these three equations (2.10, 2.14 and 2.18) in their non-dimensional form. After dropping the ˆ sign, we have the from as presented in Apte et al., (2008) [1].

$$\begin{aligned}\frac{\partial u}{\partial t}&=v-\frac{\partial h}{\partial x},\\ \frac{\partial v}{\partial t}&=-u-\frac{\partial h}{\partial y},\\ \frac{\partial h}{\partial t}&=-\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}.\end{aligned} \tag{2.19}$$

2.4 Summary

We have introduced shallow water equation in their non-linear form. We have demonstrated the assumptions that lead to linearized model. Thereafter, by applying appropriate scale factors, we have derived a non dimensional form presented in equation (2.19); this will constitutes the basis for our analysis in the subsequent chapters.

Chapter 3

Linearized Shallow Water Model and Tracer Dynamics

3.1 Introduction

In this chapter, we introduce a solution to the low order linearized shallow water model. We find a closed form solution for the time dependent amplitudes and incorporate them into an analytical solution describing tracer dynamics. This is followed by the analysis of the equilibria of tracer dynamics. For this purpose, linearized shallow water model solutions are divided into geostrophic, inertial-gravity and combine modes. Bifurcation analysis is given for each of these modes.

3.2 Model variables

The linear coupled system of three partial differential equations presented in (2.19) establishes the basis of our work in this dissertation. Here $(x, y)^T \in \mathbb{R}^2$ denote the two dimensional space coordinates and $t \geq 0$ denote the time variable.

Let $u(x,y,t)(v(x,y,t))$ denote the east-west (north-south) components of the velocity field at the spacial location $(x,y)^T$ and time t. Let $h(x,y,t)$ denote the height of the free surface of water measured above a pre-specified mean level. Equation (2.19) shows the variation of the two components of the velocity field, $u(x,y,t)$ and $v(x,y,t)$ with respect to the variation of the free surface height measured from the mean level $h(x,y,t)$.

3.3 Low-order model (LOM)

Lorenz (1960) [13] has shown that one can approximate the solution of a complex model with the low-order model. This approach has been used with a great success; a short list of a well know applications in the geophysical domain has been listed by Laksmivarahan et al., (2006) [8]. Our analysis depends on the low-order counter part of the infinite dimensional model in (2.19) obtained by the application of the standard Galerkin type projection method. To this end, we first express u, v and h in the standard two dimensional truncated Fourier series consisting of only two terms given by (Apte et al. 2008) [1] in the equation (6):

$$\begin{aligned} u(x,y,t) &= -2\pi l \quad \sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t), \\ v(x,y,t) &= +2\pi k \quad \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t), \\ h(x,y,t) &= \qquad\quad \sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t). \end{aligned} \tag{3.1}$$

In each equation above, the first time independent term with constant amplitude that is proportional to u_0 is called the *geostrophic mode* and the second time dependent terms with amplitudes $u_1(t), v_1(t), h_1(t)$ are called *inertial-gravity modes*. In the following, we refer to u_0 as the *geostrophic parameter* and $u_1(0), v_1(0), h_1(0)$ as the *inertial parameters*. Without loss of generality we only consider the case

where $n = m = l = 1$ in ((3.1)). Following (Apte et al. 2008) [1] the equations (7)), we combine equations (2.19) and (3.1) to find dynamic equations for the amplitudes

$$\begin{aligned} \dot{u_0} &= 0, \\ \dot{u_1} &= v_1, \\ \dot{v_1} &= -u_1 - 2\pi m h_1, \\ \dot{h_1} &= 2\pi m v_1, \end{aligned} \tag{3.2}$$

along with initial conditions $[u_0(0), u_1(0), v_1(0), h_1(0)]$; $du_0/dt \equiv \dot{u_0}$ is following standard notation, as is $\dot{u_1}$, $\dot{v_1}$ and $\dot{h_1}$. The process of reduction is shown in the Appendix A.

3.3.1 Analytical solution for amplitudes

In order to find an analytical solution for amplitudes we rewrite equations (3.2) in a matrix-vector form with $a = 2\pi m$.

$$\begin{bmatrix} \dot{u_1} \\ \dot{v_1} \\ \dot{h_1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & a \\ 0 & a & 0 \end{bmatrix} \times \begin{bmatrix} u_1 \\ v_1 \\ h_1 \end{bmatrix}. \tag{3.3}$$

This can be expressed succinctly as $\dot{\boldsymbol{\xi}} = \mathbf{A}\,\boldsymbol{\xi}$, with $\xi(0)$ as initial conditions (I.C), where

$$\boldsymbol{\xi} = \begin{bmatrix} u_1 \\ v_1 \\ h_1 \end{bmatrix}, \text{and } \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & a \\ 0 & a & 0 \end{bmatrix}. \tag{3.4}$$

Clearly $\mathbf{A}$ is a *skew-symmetric* matrix.

Property 3.3.1. Let $E(t) = 1/2\,\boldsymbol{\xi}^T(t)\,\boldsymbol{\xi}(t)$ denote the energy associated with the state $\boldsymbol{\xi}(t)$. Then (3.3) conserves energy $E(t)$, that is, $\dot{E}(t) = 0$.

Proof. For,

$$\begin{aligned}\frac{dE(t)}{dt} &= \boldsymbol{\xi}^T(t)\,\mathbf{A}\,\boldsymbol{\xi}(t) \\ &= u_1\dot{u_1} + v_1\dot{v_1} + h_1\dot{h_1} \\ &= u_1[v_1] + v_1[-u_1 - ah_1] + h_1[av_1] \\ &= 0.\end{aligned} \tag{3.5}$$

since $\mathbf{A}$ is a *skew-symmetric* matrix. Stated in other words, the solution $\boldsymbol{\xi}(t)$ of the linear system (3.3) always lies on the surface of a sphere centered at the origin with the radius given by

$$\begin{aligned}E(0)^{1/2} &= \Big[\boldsymbol{\xi}^T(0)\,\boldsymbol{\xi}(0)\Big]^{1/2} \\ &= \Big[u_1^2(0) + v_1^2(0) + h_1^2(0)\Big]^{1/2} \\ &= \Big[u_1^2(t) + v_1^2(t) + h_1^2(t)\Big]^{1/2} = E(t)^{1/2}.\end{aligned} \tag{3.6}$$

□

3.3.2 Eigenvalues and eigenvectors of A

Derivation of eigenvalues and eigenvectors of A

Let us look again at matrix $\mathbf{A}$ defined as

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & a \\ 0 & a & 0 \end{bmatrix}. \tag{3.7}$$

We can proceed to set the *characteristic polynomial*

$$\det(\mathbf{A}-\lambda\mathbf{I}) = |\mathbf{A}-\lambda\,\mathbf{I}| = \begin{vmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & a \\ 0 & a & -\lambda \end{vmatrix} = 0. \tag{3.8}$$

Expanding it, we get

$$-\lambda \begin{vmatrix} -\lambda & -a \\ a & -\lambda \end{vmatrix} - \begin{vmatrix} -1 & -a \\ 0 & -\lambda \end{vmatrix} = -\lambda(\lambda^2 - \lambda + a^2) = 0, \tag{3.9}$$

where solutions are

$$\begin{cases} \lambda_1 = 0 + i\sqrt{1+a^2} \\ \lambda_2 = 0 - i\sqrt{1+a^2} \\ \lambda_3 = 0 \end{cases}. \tag{3.10}$$

That is, the eigenvalues of a skew-symmetric matrix are purely imaginary or zero as shown in equation (3.10). It can be verified that the eigenvectors $\mathbf{V_1}, \mathbf{V_2}, \mathbf{V_3}$ corresponding to $\lambda_1, \lambda_2, \lambda_3$ are:

$$\mathbf{V} = [\mathbf{V_1}, \mathbf{V_2}, \mathbf{V_3}] = \begin{bmatrix} 0 & -1/\lambda & a/\lambda \\ 1 & 0 & 0 \\ 0 & -a/\lambda & -1/\lambda \end{bmatrix}. \tag{3.11}$$

Now, for the first non-zero eigenvalue we have

$$\begin{aligned} \mathbf{A}\mathbf{V_1} &= \lambda_1 \mathbf{V_1}, \\ \mathbf{A}(\mathbf{x}+i\mathbf{y}) &= i\lambda(\mathbf{x}+i\mathbf{y}). \end{aligned} \tag{3.12}$$

Equating the real and imaginary parts we get

$$\begin{aligned} \mathbf{Ax} &= -\lambda \mathbf{y} \\ \mathbf{Ay} &= \lambda \mathbf{x} \end{aligned} \tag{3.13}$$

or

$$\begin{aligned} \mathbf{y} &= -\tfrac{1}{\lambda}\mathbf{Ax}, \\ \mathbf{Ax}^2 &= \lambda^2 \mathbf{x}. \end{aligned}$$

Analogously, for the second non-zero eigenvalue we have

$$\begin{aligned} \mathbf{AV_2} &= \lambda_2 \mathbf{V_2}, \\ \mathbf{A}(\mathbf{x} - i\mathbf{y}) &= i\lambda(\mathbf{x} - i\mathbf{y}). \end{aligned} \tag{3.14}$$

Again, equating the real and imaginary parts we get

$$\begin{aligned} \mathbf{Ax} &= \ \lambda \mathbf{y} \\ \mathbf{Ay} &= \lambda \mathbf{x} \end{aligned}$$

or

$$\begin{aligned} \mathbf{x} &= -\tfrac{1}{\lambda}\mathbf{Ay}, \\ \mathbf{Ay}^2 &= -\lambda^2 \mathbf{y}. \end{aligned}$$

We will use $\mathbf{A}^2$ in our calculations, so let us find it now as

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & a \\ 0 & a & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & a \\ 0 & a & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -a \\ 0 & -(1+a^2) & 0 \\ -a & 0 & -a^2 \end{bmatrix}. \tag{3.15}$$

It can be verified that $\mathbf{A}^2 = \mathbf{AA} = \mathbf{A}^T\mathbf{A}^T = (\mathbf{AA})^T$, therefore $\mathbf{A}^2$ is *symmetric* (one for which $\mathbf{A}^T = \mathbf{A}$) if $\mathbf{A}$ is *skew-symmetric* (one for which $\mathbf{A}^T = -\mathbf{A}$). Now,

we continue on finding eigenvectors:

$$\mathbf{A}\mathbf{x}^2 = -\lambda^2\mathbf{x}, \tag{3.16}$$

$$\begin{bmatrix} -1 & 0 & -a \\ 0 & -(1+a^2) & 0 \\ -a & 0 & -a^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = -(1+a^2) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{3.17}$$

After expanding, we arrive at

$$\begin{cases} -x_1 - ax_3 &= -(1+a^2)x_1 \\ -(1+a^2)x_2 &= -(1+a^2)x_2 \\ -ax_1 - a^2x_3 &= -(1+a^2)x_3 \end{cases}.$$

Now, we can solve it after simplifying

$$\begin{cases} x_3 &= ax_1 \\ x_2 &= x_2 \\ x_3 &= ax_1 \end{cases}.$$

Since we can set x_2 to an arbitrary value, let us choose $x_2 = 1$, and then we can set $x_1 = x_3 = 0$. This would give us the normalized first eigenvector $\mathbf{V_1} = (0,1,0)^T$. For the second eigenvector, we can set $x_2 = 0$, and $x_1 = -1$. This would give us $x_3 = -a$, and after normalization we get get the second eigenvector as $\mathbf{V_2} = -1/\sqrt{(1+a^2)}(1,0,a)^T$, and then we would find the third normalized eigenvector as $\mathbf{V_3} = -1/\sqrt{(1+a^2)}(1,0,-a)^T$. It can be verified that the matrix

of eigenvectors of A

$$V = [V_1, V_2, V_3] = \begin{bmatrix} 0 & -1/\lambda & a/\lambda \\ 1 & 0 & 0 \\ 0 & -a/\lambda & -1/\lambda \end{bmatrix} \tag{3.18}$$

is an orthogonal matrix, and so, $\mathbf{V}^T = \mathbf{V}^{-1}$.

Eigendecomposition of A

The Jordan canonical form of matrix $\mathbf{A}$ is as follows

$$\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T, \tag{3.19}$$

where

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & \lambda & 0 \\ -\lambda & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

or

$$\mathbf{V}^T\mathbf{A}\mathbf{V} = \mathbf{\Lambda}. \tag{3.20}$$

We know that $\dot{\boldsymbol{\xi}} = \mathbf{A}\,\boldsymbol{\xi}$ with $\boldsymbol{\xi}(0)$ as initial conditions. Then, following (chapter 31, Lewis et al., (2006) [12])

$$\begin{aligned} \boldsymbol{\xi}(t) &= e^{\mathbf{A}t}\,\boldsymbol{\xi}(0) \\ &= e^{(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T)t}\,\boldsymbol{\xi}(0) \\ &= [\mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^T]\,\boldsymbol{\xi}(0). \end{aligned} \tag{3.21}$$

The exponential of a matrix

For a moment, we have to shift our attention to finding an exponent of a matrix. Following Hirsch et al., (2004) [4], we know the exponential function can be expressed in a form of the infinite series

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{3.22}$$

Similarly for A, we can define its exponential as

$$\exp \mathbf{A} = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!}. \tag{3.23}$$

In our case, matrix $\mathbf{\Lambda}$ is as follows

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & \lambda & 0 \\ -\lambda & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \boldsymbol{\lambda} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

So, we can express its powers in the following manner

$$\mathbf{\Lambda}^0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{\Lambda}^1 = \lambda \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Lambda}^2 = -\lambda^2 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = -\lambda^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Lambda}^3 = -\lambda^3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = -\lambda^3 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Lambda}^4 = -\lambda^4 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \lambda^4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{\Lambda}^5 = \lambda^5 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \lambda^5 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \cdots$$

Therefore, to express $exp(\mathbf{\Lambda})$ we can look at the few first elements

$$\begin{aligned}
r(1,1) &= 1 + \lambda * 0 - \lambda^2 * 1 - \lambda^3 * 0 + \lambda^4 * 1 + \lambda^5 * 0 + \cdots \\
&= 1 - \lambda^2 + \lambda^4 + \cdots \\
r(1,2) &= 0 + \lambda * 1 - \lambda^2 * 0 - \lambda^3 * 1 + \lambda^4 * 0 + \lambda^5 * 1 + \cdots \\
&= \lambda - \lambda^3 + \lambda^5 + \cdots \\
r(2,1) &= -\lambda * 1 + \lambda^2 * 0 + \lambda^3 * 1 - \lambda^4 * 0 - \lambda^5 * 1 + \cdots \\
&= -\lambda + \lambda^3 - \lambda^5 - \cdots \\
r(2,2) &= 1 + \lambda * 0 - \lambda^2 * 1 - \lambda^3 * 0 + \lambda^4 * 1 + \lambda^5 * 0 + \cdots \\
&= 1 - \lambda^2 + \lambda^4 + \cdots
\end{aligned}$$

Now, we can recall that *cos* function can be expressed in a form on an infinite sum

$$\cos\lambda = \sum_{k=0}^{\infty} (-1)^k \frac{\lambda^{2k}}{(2k)!}.$$

Analogously, *sin* function can be expressed as

$$\sin\lambda = \sum_{k=0}^{\infty} (-1)^k \frac{\lambda^{2k+1}}{(2k+1)!}.$$

Therefore, combining these we can express $\exp(\mathbf{\Lambda})$ as

$$\begin{aligned} e^{\mathbf{\Lambda}} &= \begin{bmatrix} \sum_{k=0}^{\infty}(-1)^k \frac{\lambda^{2k}}{(2k)!} & \sum_{k=0}^{\infty}(-1)^k \frac{\lambda^{2k+1}}{(2k+1)!} & 0 \\ \sum_{k=0}^{\infty}(-1)^k \frac{\lambda^{2k+1}}{(2k+1)!} & \sum_{k=0}^{\infty}(-1)^k \frac{\lambda^{2k}}{(2k)!} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos\lambda & \sin\lambda & 0 \\ -\sin\lambda & \cos\lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{3.24}$$

But we also know that

$$e^{\mathbf{\Lambda} t} = \begin{bmatrix} \cos(\lambda t) & \sin(\lambda t) & 0 \\ -\sin(\lambda t) & \cos(\lambda t) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ where } c = \cos(\lambda t), s = \sin(\lambda t). \tag{3.25}$$

Amplitudes

Combining all of the above terms, we get $\xi(t) = e^{(\mathbf{V}\mathbf{\Lambda}\mathbf{V}^T)t}\boldsymbol{\xi}(0)$ expanded as follows:

$$\begin{bmatrix} u_1(t) \\ v_1(t) \\ h_1(t) \end{bmatrix} = \begin{bmatrix} 0 & -1/\lambda & a/\lambda \\ 1 & 0 & 0 \\ 0 & -a/\lambda & -1/\lambda \end{bmatrix} \begin{bmatrix} c & s & 0 \\ -s & c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1/\lambda & a/\lambda \\ 1 & 0 & 0 \\ 0 & -a/\lambda & -1/\lambda \end{bmatrix} \begin{bmatrix} u_1(0) \\ v_1(0) \\ h_1(0) \end{bmatrix} \tag{3.26}$$

$$= \begin{bmatrix} s/\lambda & -c/\lambda & a/\lambda \\ c & s & 0 \\ as/\lambda & -ac/\lambda & 1 \end{bmatrix} \begin{bmatrix} 0 & -1/\lambda & a/\lambda \\ 1 & 0 & 0 \\ 0 & -a/\lambda & -1/\lambda \end{bmatrix} \begin{bmatrix} u_1(0) \\ v_1(0) \\ h_1(0) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{c+a^2}{\lambda^2} & \frac{s}{\lambda} & \frac{ac-a}{\lambda^2} \\ \frac{-s}{\lambda} & c & \frac{-as}{\lambda} \\ \frac{ac-a}{\lambda^2} & \frac{as}{\lambda} & \frac{1+ca^2}{\lambda^2} \end{bmatrix} \begin{bmatrix} u_1(0) \\ v_1(0) \\ h_1(0) \end{bmatrix}.$$

After we expand the above equation, we have an explicit expressions for the time varying amplitudes $u_1(t), v_1(t), \text{and}\, h_1(t)$ as a function of time

$$\begin{aligned} u_1(t) = &\quad \frac{c+a^2}{\lambda^2}\, u_1(0) + \quad \frac{s}{\lambda}\, v_1(0) + \quad \frac{ac-a}{\lambda^2}\, h_1(0), \\ v_1(t) = &\quad \frac{-s}{\lambda}\, u_1(0) + \quad c\, v_1(0) + \quad \frac{-as}{\lambda}\, h_1(0), \\ h_1(t) = &\quad \frac{ac-a}{\lambda^2}\, u_1(0) + \quad \frac{as}{\lambda}\, v_1(0) + \quad \frac{1+ca^2}{\lambda^2}\, h_1(0). \end{aligned} \tag{3.27}$$

We can notice that for $t = 0$, we have $c = \cos(\lambda t) = 1$ and $s = \sin(\lambda t) = 0$. We recall that $a = 2\pi m$, and $\lambda = \sqrt{1+a^2} = \sqrt{1+(2\pi m)^2}$.

$$\begin{aligned} u_1(0) &= \frac{1+a^2}{\lambda^2}\, u_1(0) = \frac{1+a^2}{1+a^2}\, u_1(0) = u_1(0), \\ v_1(0) &= v_1(0), \\ h_1(0) &= \frac{1+a^2}{\lambda^2}\, h_1(0) = \frac{1+a^2}{1+a^2}\, h_1(0) = h_1(0). \end{aligned} \tag{3.28}$$

We can expand equations (3.27) to get full expressions for amplitudes

$$
\begin{aligned}
u_1(t) = & \frac{\left(\cos\left(t\sqrt{4\,\pi^2\,m^2+1}\right)+4\,m^2\,\pi^2\right)}{4\,\pi^2\,m^2+1}u_1(0) \\
& + \frac{\sin\left(t\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}v_1(0) \\
& + \frac{2\,\pi\,m\,\left(\cos\left(t\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}h_1(0), \\
v_1(t) = & -\frac{\sin\left(t\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}u_1(0) \\
& + \cos\left(t\sqrt{4\,\pi^2\,m^2+1}\right)v_1(0) \\
& - \frac{2\,\pi\,m\,\sin\left(t\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}h_1(0), \\
h_1(t) = & \frac{2\,\pi\,m\,\left(\cos\left(t\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}u_1(0) \\
& + \frac{2\,\pi\,m\,\sin\left(t\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}v_1(0) \\
& \frac{\left(4\,m^2\,\pi^2\,\cos\left(t\sqrt{4\,\pi^2\,m^2+1}\right)+1\right)}{4\,\pi^2\,m^2+1}h_1(0).
\end{aligned}
\tag{3.29}
$$

3.4 Tracer dynamics

We proceed with the analysis of the tracer dynamics after rearranging equations (3.27). We can present a number of interesting features of the shallow water

model using the version of these equations in (3.30).

$$\begin{aligned} u_1(t) &= \frac{a}{\lambda^2}\left[au_1(0) - h_1(0)\right] + \frac{1}{\lambda^2}\left[u_1(0) + ah_1(0)\right]c + \frac{v_1(0)}{\lambda}\,s, \\ v_1(t) &= v_1(0)\,c - \frac{1}{\lambda}\left[ah_1(0) + u_1(0)\right]s, \\ h_1(t) &= \frac{1}{\lambda^2}\left[h_1(0) - au_1(0)\right] + \frac{a}{\lambda^2}\left[u_1(0) + ah_1(0)\right]c + \frac{av_1(0)}{\lambda}\,s. \end{aligned} \tag{3.30}$$

Since the solution of equation (3.3) $\boldsymbol{\xi}(t)$ lies on a sphere, it is immediate that $u_1(t), v_1(t)$, and $h_1(t)$ are not linearly independent.

Let $\mathbf{X}(t) = \big(x(t), y(t)\big)^T \in \mathbb{R}^2$ denote the position of the tracer particle floating on the free surface of the water. The dynamics of motion of this tracer is then given by

$$\begin{aligned} \dot{x} &= u(x, y, t), \\ \dot{y} &= v(x, y, t), \end{aligned} \tag{3.31}$$

where the right hand side of (3.31) is obtained by substituting (3.27) in (3.1) with u_0, a constant. For later reference, define the geostrophic velocity components

$$\begin{aligned} f_1(x, y) &= -2\pi u_0 \sin(2\pi x)\cos(2\pi y)\,, \\ f_2(x, y) &= 2\pi u_0 \cos(2\pi x)\sin(2\pi y)\,, \end{aligned} \tag{3.32}$$

and the inertial-gravity velocity components

$$\begin{aligned} g_1(x, y, t) &= u_1(t)\cos(2\pi m y)\,, \\ g_2(x, y, t) &= v_1(t)\cos(2\pi m y)\,. \end{aligned} \tag{3.33}$$

Let $\mathbf{f} = (f_1, f_2)^T \in \mathbb{R}^2$ and $\mathbf{g} = (g_1, g_2)^T \in \mathbb{R}^2$. Then, (3.31) can be rewritten in

the vector form as

$$\dot{\mathbf{X}}(t) = \mathbf{f}(\mathbf{x}, u_0) + \mathbf{g}(\mathbf{x}, \hat{\boldsymbol{\alpha}}, t), \tag{3.34}$$

where $u_1(t)$, $v_1(t)$, $h_1(t)$ are given in (3.27). We can also recall that $\hat{\boldsymbol{\alpha}} = \big(u_1(0), v_1(0), h_1(0)\big) \in \mathbb{R}^3$. Clearly, the first component $\mathbf{f}$ of the vector field depends on the geostrophic parameter u_0 and the second component $\mathbf{g}$ depends on the inertial parameters $\hat{\boldsymbol{\alpha}} \in \mathbb{R}^3$. One of our goals in this thesis is to characterize and catalog the behavior of the solution of (3.34) as the geostrophic and inertial parameters are varied. It turns out this class of nonlinear time varying dynamics exhibits many interesting bifurcations as the four parameters $(u_0, \hat{\boldsymbol{\alpha}})$ are varied in $\mathbb{R}^4$.

3.5 Analysis of Equilibria of Tracer dynamics

It is convenient to divide the analysis into three cases.

3.5.1 Case 1: Equilibria in geostrophic mode

By setting the inertial parameter $\hat{\boldsymbol{\alpha}} = 0$, we obtain the tracer dynamics given by the nonlinear autonomous system

$$\dot{\mathbf{X}} = \mathbf{f}(\mathbf{X}, u_0) \tag{3.35}$$

controlled by the geostrophic parameter u_0. From (3.32), the Jacobian of the flow field $\mathbf{f}(\mathbf{X}, u_0)$ is given by

$$\mathbf{D}_{\mathbf{x}}(\mathbf{f}) = -4\pi^2 u_0 \begin{bmatrix} \cos(2\pi x)\cos(2\pi y) & -\sin(2\pi x)\sin(2\pi y) \\ \sin(2\pi x)\sin(2\pi y) & -\cos(2\pi x)\cos(2\pi y) \end{bmatrix}. \tag{3.36}$$

We consider two types of equilibria for (3.35)

Type 1 Equilibria:

From (3.35) and (3.32), it follows that $\mathbf{f}(\mathbf{X}, u_0) = 0$ when $sin(2\pi x) = 0 = sin(2\pi y)$, that is, when

$$x = y = \pm\frac{k}{2} \text{ for } k = 0, 1, 2, \ldots \tag{3.37}$$

which is independent of u_0. The Jacobian (3.36) at these equilibria becomes

$$\mathbf{D_x}(\mathbf{f}) = -4\pi^2 u_0 \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{3.38}$$

whose eigenvalues are given by $\lambda_1 = 4\pi^2 u_0$ and $\lambda_2 = -4\pi^2 u_0$, with $\mathbf{e_1} = (1, 0)^T$ and $\mathbf{e_2} = (0, 1)^T$ as their corresponding eigenvectors. Clearly, this family of equilibria corresponds to a *saddle* for all non-zero values of the geostrophic parameter u_0. In a small neighborhood around the equilibria in (3.37), the system (3.35) is equivalent to a linear dynamics given by

$$\dot{\boldsymbol{\eta}} = \mathbf{D_x}(\mathbf{f})\boldsymbol{\eta}, \tag{3.39}$$

with $\mathbf{D_x}(\mathbf{f})$ in (3.38) and $\boldsymbol{\eta}(t) = \big(\eta_1(t), \eta_2(t)\big)^T$.

The solution of (3.39) is given by

$$\eta_1(t) = \eta_1(0)e^{(4\pi^2 u_0)t} \text{ and } \eta_2(t) = \eta_2(0)e^{-(4\pi^2 u_0)t}. \tag{3.40}$$

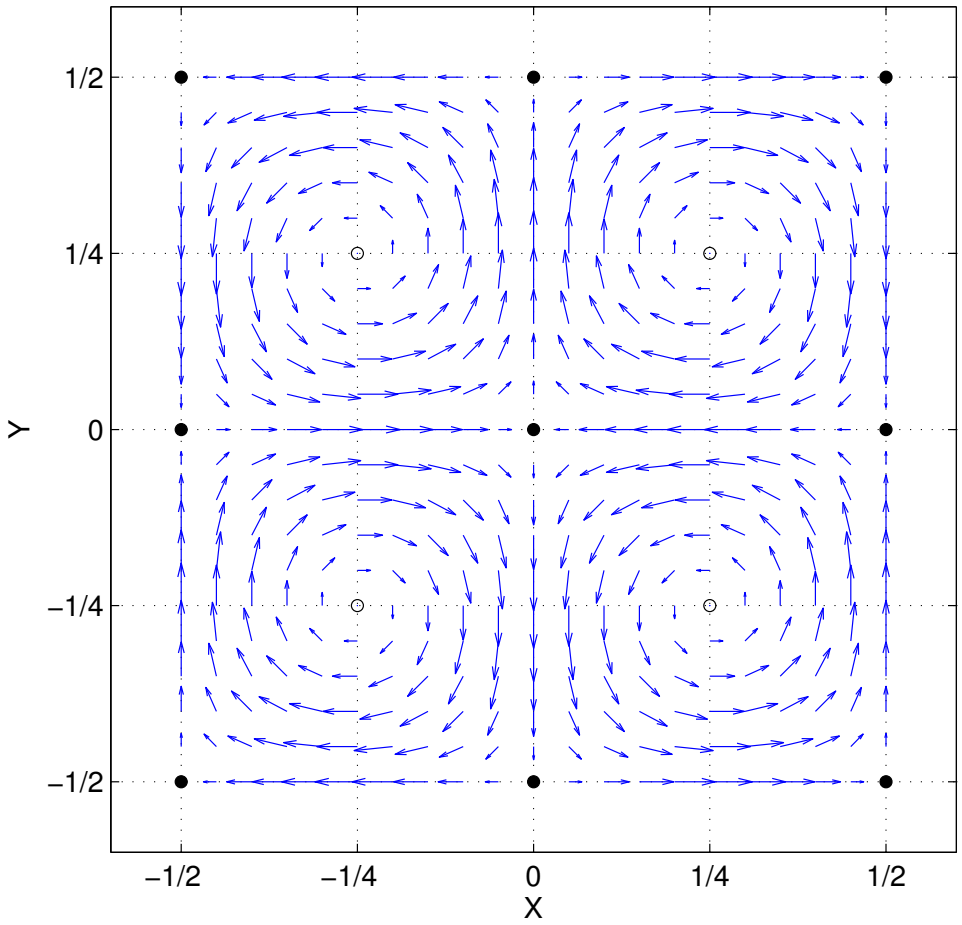


Figure 3.1: A display of equilibria of *Type 1* - filled circles and *Type 2* - unfilled circles along with the velocity field around them. Filled circles are saddle points and unfilled circles are centers. The field plot around these equilibria corresponds to $f(x, y)$ in (3.32) with $u_0 = 1.0$, $u_1(0) = v_1(0) = h_1(0) = 0$ and time $t = 0$.

Type 2 Equilibria:

From (3.35) and (3.32), it again follows that $\mathbf{f}(\mathbf{X}, u_0) = 0$ when $\cos(2\pi x) = 0 = \cos(2\pi y)$, that is, when

$$x = y = \pm\frac{2k+1}{4} \text{ for } k = 0, 1, 2, \ldots \tag{3.41}$$

which is also independent of u_0. The Jacobian (3.36) at these equilibria becomes

$$\mathbf{D_x}(\mathbf{f}) = -4\pi^2 u_0 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{3.42}$$

whose eigenvalues are purely imaginary and are given by $\lambda_1 = i4\pi^2 u_0$ and $\lambda_2 = -i4\pi^2 u_0$, with any pair of unit, orthogonal vectors $\mathbf{e_1} = \frac{1}{\sqrt{a^2+b^2}}(a,b)^T$ and $\mathbf{e_2} = \frac{1}{\sqrt{a^2+b^2}}(b,-a)^T$ as corresponding eigenvectors, where a and b are arbitrary real constants. This family of equilibria correspond to a *center* for all non-zero values of the geostrophic parameter u_0. In a small neighborhood around this equilibria, equation (3.35) is equivalent to a linear dynamics

$$\dot{\boldsymbol{\eta}} = \mathbf{D_x}(\mathbf{f})\boldsymbol{\eta}, \tag{3.43}$$

with $\mathbf{D_x}(\mathbf{f})$ in (3.42). The solution of (3.43) is then given by

$$\begin{bmatrix} \eta_1(t) \\ \eta_2(t) \end{bmatrix} = \begin{bmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{bmatrix} \begin{bmatrix} \eta_1(0) \\ \eta_2(0) \end{bmatrix}, \quad \text{where} \quad b = 4\pi^2 u_0. \tag{3.44}$$

It turns out that equation (3.35) with the field $\mathbf{f}(\mathbf{X}, u_0)$ given in (3.32) can indeed be solved in closed form. It can be verified that (with $a = 2\pi$), (3.35) is equivalent to

$$\frac{dy}{dx} = -\frac{\tan(ay)}{\tan(ax)}. \tag{3.45}$$

Expressing (3.45) as

$$\frac{dy}{\tan(ay)} = -\frac{dx}{\tan(ax)} \tag{3.46}$$

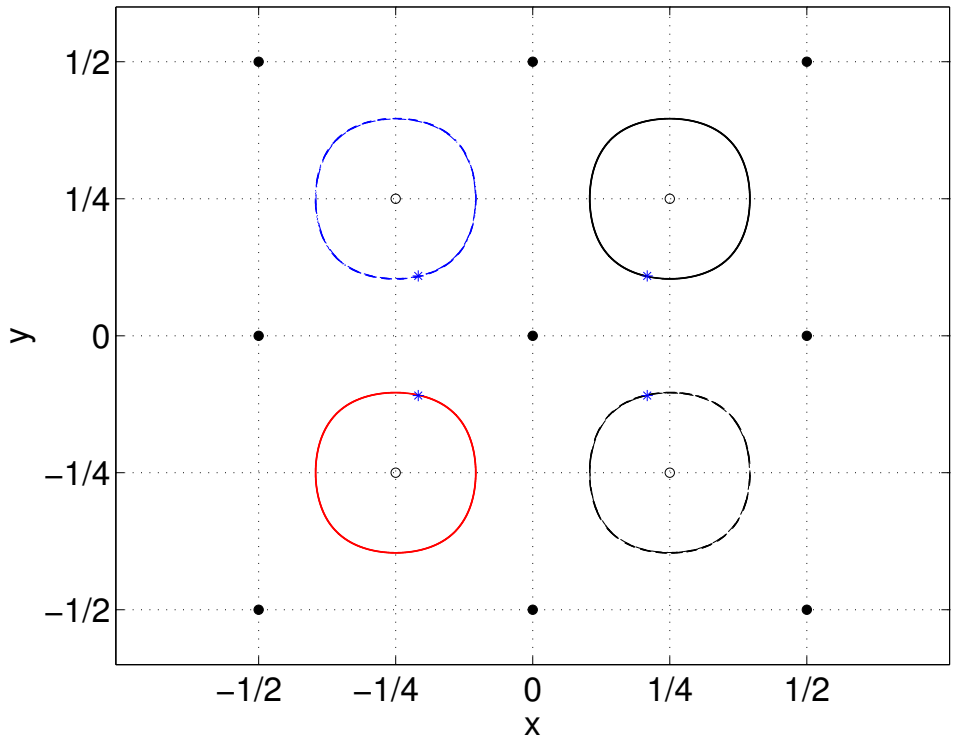


Figure 3.2: A display of trajectories of the pure geostrophic mode in equation (3.35) $\dot{x} = f(x, u_0)$. The values of $(x_0, y_0) = (0.209, 0.109)$, $(x_0, y_0) = (-0.209, 0.109)$, $(x_0, y_0) = (0.209, -0.109)$, $(x_0, y_0) = (-0.209, -0.109)$ are indicated by *.

and integrating, the solution is given by

$$\log\left(\sin(ax)\sin(ay)\right) = c, \tag{3.47}$$

where c is the constant of integration. If (x_0, y_0) is the initial position of the tracer, from (3.47) the solution of (3.35) can be expressed as

$$\sin(ax)\sin(ay) = \sin(ax_0)\sin(ay_0) \tag{3.48}$$

Figure 3.1 is an illustration of the relative disposition *type 1 and* type 2 equilibria along with the vector field around them when $u_0 = 1.0$. Phase plots using

equation (3.48) for various choices of (x_0, y_0) are given in Figure 3.2.

3.5.2 Case 2: Equilibria in inertial-gravity mode

By setting the geostrophic parameter u_0 to zero and ensuring that $\hat{\boldsymbol{\alpha}}$ is not equal to zero, we obtain the second special case for the tracer dynamics given by the nonlinear nonautonomous system

$$\dot{\mathbf{X}} = \mathbf{g}(\mathbf{X}, \hat{\boldsymbol{\alpha}}, t), \tag{3.49}$$

whose behavior is controlled by $\hat{\boldsymbol{\alpha}} \in \mathbb{R}^3$. Referring to (3.33), the Jacobian of this flow field is given by

$$\mathbf{D}_{\mathbf{x}}(\mathbf{g}) = 2\pi \begin{bmatrix} 0 & -u_1(t)\sin(2\pi y) \\ 0 & -v_1(t)\sin(2\pi y) \end{bmatrix}. \tag{3.50}$$

From (3.33) it is immediate that $\mathbf{g}(\mathbf{x}, \hat{\boldsymbol{\alpha}}, t) = 0$, when $\cos(2\pi y) = 0$. Thus, the equilibria for (3.49) are given by

$$x \text{ arbitrary, and } y = \pm\frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots \tag{3.51}$$

Again, it is convenient to consider two types of equilbria.

Type a: Equilibria in the upper half plane

This type of equilibria is given by

$$x \text{ arbitrary, and } y = \frac{2k+1}{4}, k = 0, 1, 2, \ldots \tag{3.52}$$

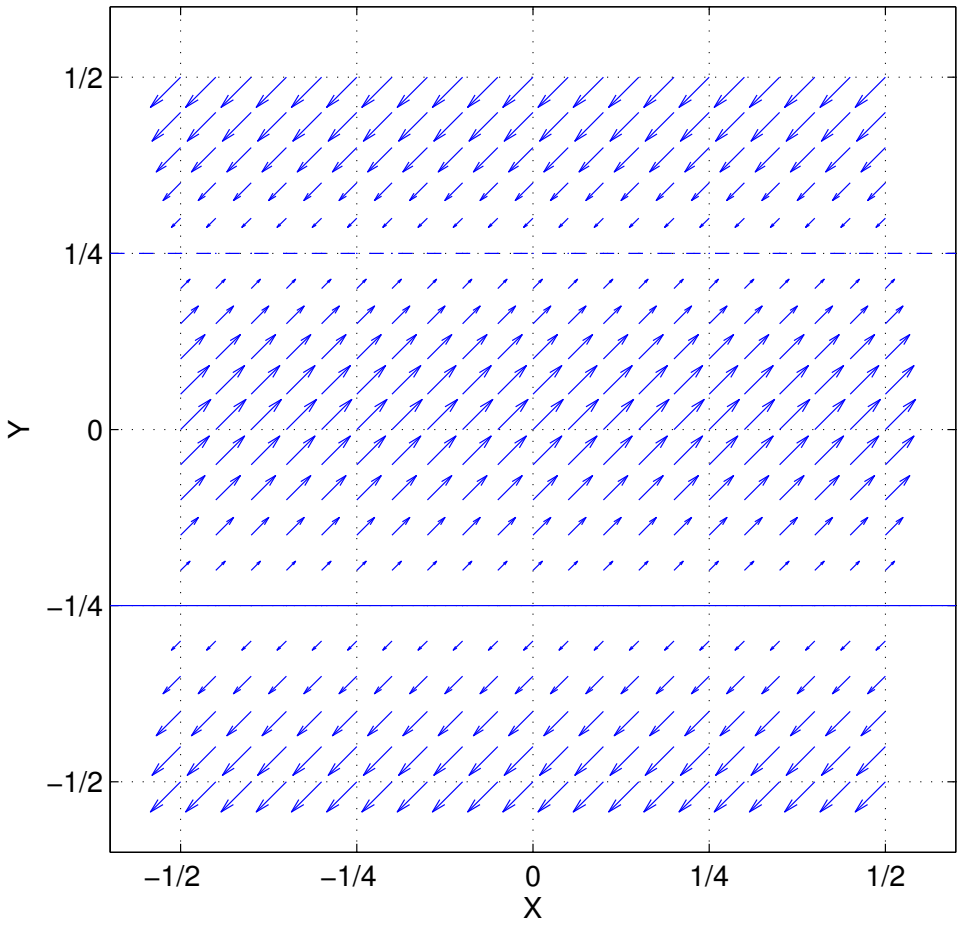


Figure 3.3: A display of equilibria of *Type a* - dashed line and *Type b* - solid line, and the flow field around them. The plot of the time varying vector field around these equilibria corresponding to $g(x, \alpha, t)$ in (3.33) at time $t = 0$ and for values of corresponding parameters $u_0 = 0$, $u_1(0) = v_1(0) = h_1(0) = 1$ and time t = 0.

The Jacobian (3.50) evaluated at these equilibria becomes

$$\mathbf{D_x(g)} = 2\pi \begin{bmatrix} 0 & -u_1(t) \\ 0 & -v_1(t) \end{bmatrix}, \tag{3.53}$$

whose eigenvalues are given by $\lambda_1 = -2\pi v_1(t)$ and $\lambda_2 = 0$, where $v_1(t)$ depends on $\hat{\boldsymbol{\alpha}}$ through (3.27). It can be verified that the eigenvector corresponding to λ_1

is $\mathbf{e_1} = \left(\frac{u_1(t)}{v_1(t)}, 1\right)^T$ and that corresponding to λ_2 is any arbitrary non-zero vector in $\mathbb{R}^2$. Dynamics in (3.49) around the equilibrium (3.52) is equivalent to

$$\dot{\boldsymbol{\eta}} = \mathbf{D_x}(\mathbf{g})\boldsymbol{\eta}, \tag{3.54}$$

with $\mathbf{D_x}(\mathbf{g})$ in (3.53). Clearly,

$$\eta_2(t) = \eta_2(0)e^{-2\pi\Psi(t)}, \tag{3.55a}$$

$$\text{where } \Psi(t) = \int_0^t v_1(s)ds. \tag{3.55b}$$

Hence,

$$\eta_2(t) = -2\pi\eta_2(0)\int_0^t u_1(s)\eta_2(s)ds. \tag{3.56}$$

Type b: Equilibria in the lower half plane

This type of equilibria is given by

$$x \text{ arbitrary, and } y = -\frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots \tag{3.57}$$

The Jacobian (3.50) evaluated at these equilibria becomes

$$\mathbf{D_x}(\mathbf{g}) = 2\pi \begin{bmatrix} 0 & u_1(t) \\ 0 & v_1(t) \end{bmatrix}, \tag{3.58}$$

whose eigenvalues are given by $\lambda_1 = 2\pi v_1(t)$ and $\lambda_2 = 0$, whose eigenvectors are the same as in *Type a*. Again the nonlinear equation (3.49) is equivalent to

$$\dot{\boldsymbol{\eta}} = \mathbf{D_x}(\mathbf{g})\boldsymbol{\eta} \tag{3.59}$$

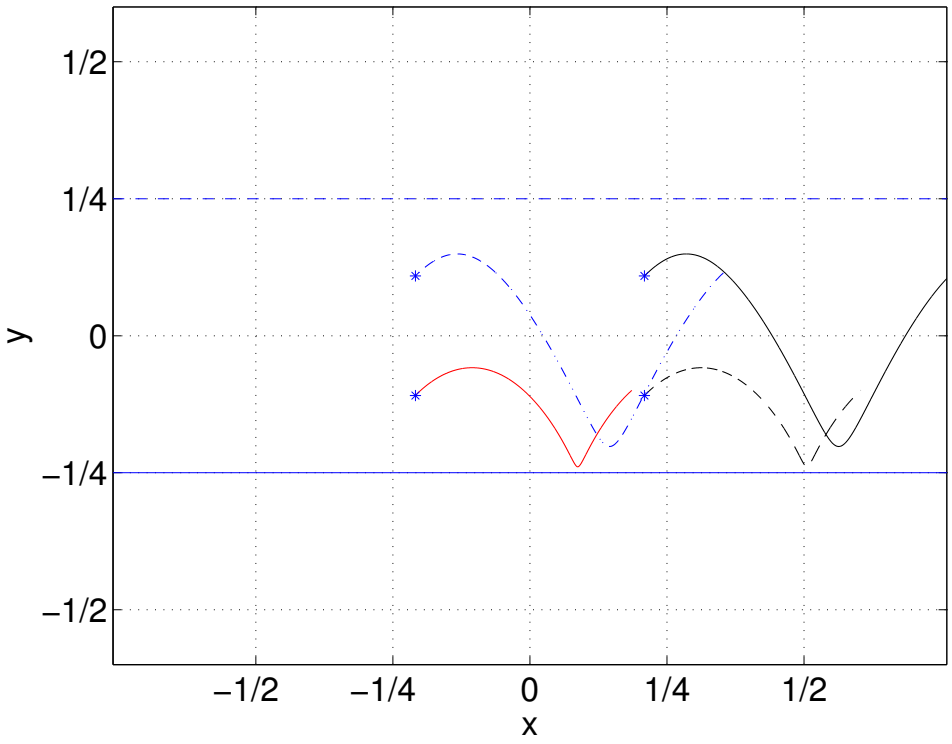


Figure 3.4: A display of trajectories of the pure inertial gravity modes $\dot{x} = g(x, \alpha, t)$ in (3.33). The starting points of various trajectories $(x_0, y_0) = (0.209, 0.109)$, $(x_0, y_0) = (-0.209, 0.109)$,$(x_0, y_0) = (0.209, -0.109)$, $(x_0, y_0) = (-0.209, -0.109)$ are shown by *.

around the equilibria in (3.57) with $D_x(g)$ in (3.58). The solution of (3.59) is given by

$$\begin{aligned} \eta_2(t) &= \eta_2(0) e^{2\pi\Psi(t)} \\ \eta_1(t) &= 2\pi\eta_1(0) \int_0^t u_1(s)\eta_2(s)ds \end{aligned} \tag{3.60}$$

Again it turns out that we can solve (3.49) from (3.33). The second equation in (3.49) becomes

$$\frac{dy}{\cos ay} = v_1(t)dt, \tag{3.61}$$

which on integrating becomes

$$\frac{1}{a}\ln\left(\tan\left(\frac{\pi}{4}+\frac{ay}{2}\right)\right)=c+\Psi(t), \tag{3.62}$$

where $\Psi(t)$ is given in (3.55b). If (x_0, y_0) is the initial position of the tracer, the solution of (3.62) becomes

$$\tan\left(\frac{\pi}{4}+\frac{ay}{2}\right)=e^{2\pi\Psi(t)}\tan\left(\frac{\pi}{4}+\frac{ay_o}{2}\right),\ \text{ or} \tag{3.63a}$$

$$y(t)=\frac{2}{a}\left[\arctan\left(e^{2\pi\Psi(t)}\tan\left(\frac{\pi}{4}+\frac{ay_o}{2}\right)\right)-\frac{\pi}{4}\right]. \tag{3.63b}$$

The first component of the solution of (3.49) is given by

$$x(t)=x(0)+\int_0^t u_1(s)\cos\left(2\pi y(s)\right)ds. \tag{3.64}$$

An illustration of the relative disposition of the equilibria and the field corresponding to (3.49) are given in Figure 3.3. Figure 3.4 contains the phase plot $\big(x(t), y(t)\big)$ obtained from (3.63) and (3.64) for different initial points.

3.5.3 Case 3: Equilibria in general case

When $u_0 \neq 0$ and $\hat{\boldsymbol{\alpha}} \neq 0$, we obtain the interesting general case where the geostrophic and the inertial-gravity modes exert their own influence. Depending on the relative strength of these component fields, we can obtain a variety of behavior of the tracer dynamics

$$\dot{\mathbf{X}}=\mathbf{f}(\mathbf{X}, u_0)+\mathbf{g}(\mathbf{X}, \hat{\boldsymbol{\alpha}}, t). \tag{3.65}$$

The Jacobian of this flow field is given by

$$\begin{aligned}
&\mathbf{D_x}(\mathbf{f}+\mathbf{g}) = \mathbf{D_x}(\mathbf{f}) + \mathbf{D_x}(\mathbf{g}) \\
&= 2\pi \begin{bmatrix} -2\pi u_0 \cos(2\pi x)\cos(2\pi y) & 2\pi u_0 \sin(2\pi x)\sin(2\pi y) - u_1(t)\sin(2\pi y) \\ -2\pi u_0 \sin(2\pi x)\sin(2\pi y) & 2\pi u_0 \cos(2\pi x)\cos(2\pi y) - v_1(t)\sin(2\pi y) \end{bmatrix} \\
&= 2\pi \begin{bmatrix} -P_C(x,y) & P_S(x,y) - u_1(t)\sin(2\pi y) \\ -P_S(x,y) & P_C(x,y) - v_1(t)\sin(2\pi y) \end{bmatrix},
\end{aligned} \tag{3.66}$$

where $P_S(x,y) = 2\pi u_0 \sin(2\pi x)\sin(2\pi y)$, and $P_C(x,y) = 2\pi u_0 \cos(2\pi x)\cos(2\pi y)$. It can be verified that $\mathbf{f}(\mathbf{X}, u_0) + \mathbf{g}(\mathbf{X}, \hat{\alpha}) = 0$ when

$$x = y = \pm\frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots \tag{3.67}$$

To simplify the analysis we consider four types of equilibria.

Type A: Equilibria in the first quadrant

This type of equilibria is given by

$$x = y = \frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots$$

In this case, the Jacobian (3.66) becomes

$$\mathbf{D_x}(\mathbf{f}+\mathbf{g}) = 2\pi \begin{bmatrix} 0 & 2\pi u_0 - u_1(t) \\ -2\pi u_0 & -v_1(t) \end{bmatrix}, \tag{3.68}$$

whose eigenvalues are given by

$$\lambda_{1,2}(t) = \frac{-v_1(t) \pm \sqrt{\Delta_1(t)}}{2}, \tag{3.69}$$

where

$$\Delta_1(t) = \left(v_1(t)\right)^2 - 8\pi u_0 \left(2\pi u_0 - u_1(t)\right). \tag{3.70}$$

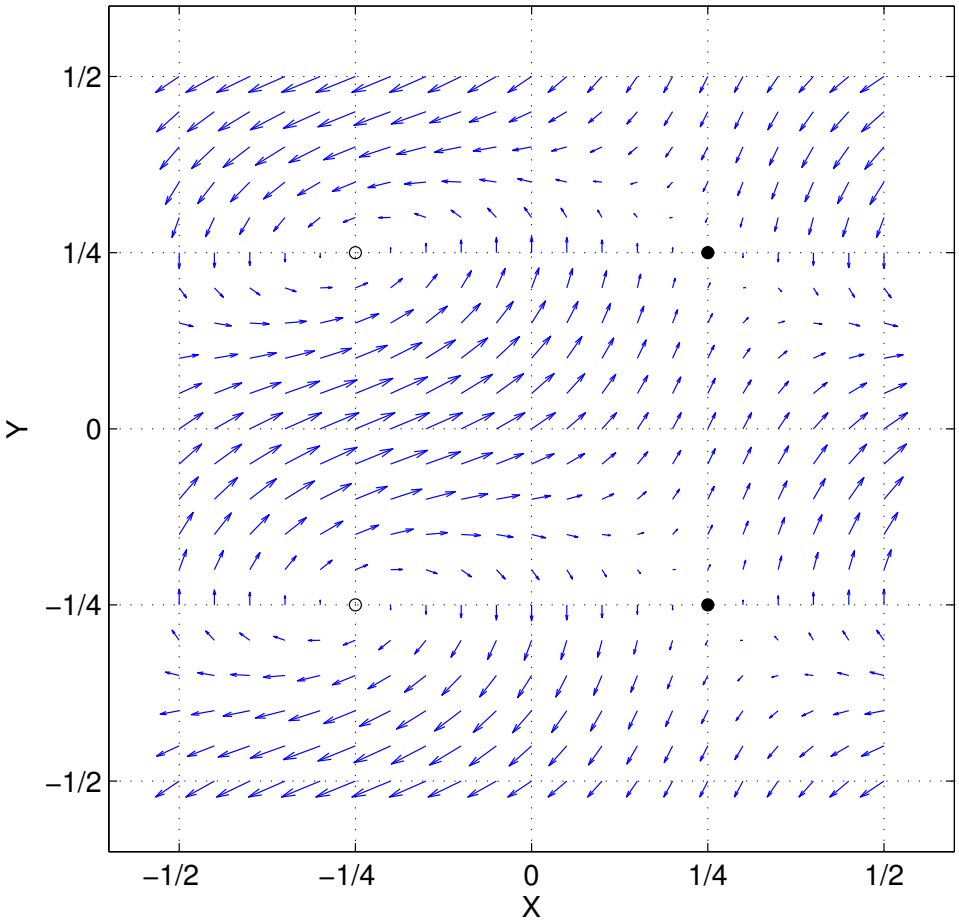


Figure 3.5: A display of equilibria of *Type A* (1/4, 1/4) and *B* (1/4, - 1/4)-solid circles and *Type C* (-1/4, 1/4) and *D* (-1/4, -1/4) - empty circles , and the flow field around them. A snapshot of the time varying vector field given by (3.35) at time $t = 0$ where $u_0 = 1.0$, $u_1(0) = 9.4248, v_1(0) = \lambda$ and $h_1(0) = -1.5$.

Type B: Equilibria in the fourth quadrant

This type of equilibria is given by

$$x = -y = \frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots$$

In this case, the Jacobian (3.66) takes the form

$$\mathbf{D_x}(\mathbf{f}+\mathbf{g}) = 2\pi \begin{bmatrix} 0 & -2\pi u_0 + u_1(t) \\ 2\pi u_0 & v_1(t) \end{bmatrix}, \tag{3.71}$$

whose eigenvalues are given by

$$\lambda_{1,2}(t) = \frac{v_1(t) \pm \sqrt{\Delta_1(t)}}{2}, \tag{3.72}$$

where Δ_1 is given in (3.70).

Type C: Equilibria in the second quadrant

$$x = -y = -\frac{2k+1}{4}, \text{ for } k = 0, 1, 2, \ldots \tag{3.73}$$

The Jacobian (3.66) becomes

$$\mathbf{D_x}(\mathbf{f}+\mathbf{g}) = 2\pi \begin{bmatrix} 0 & -2\pi u_0 - u_1(t) \\ 2\pi u_0 & -v_1(t) \end{bmatrix}, \tag{3.74}$$

whose eigenvalues are given by

$$\lambda_{1,2}(t) = \frac{-v_1(t) \pm \sqrt{\Delta_2(t)}}{2}, \tag{3.75}$$

where

$$\Delta_2(t) = \left(v_1(t)\right)^2 - 8\pi u_0 \left(2\pi u_0 + u_1(t)\right). \tag{3.76}$$

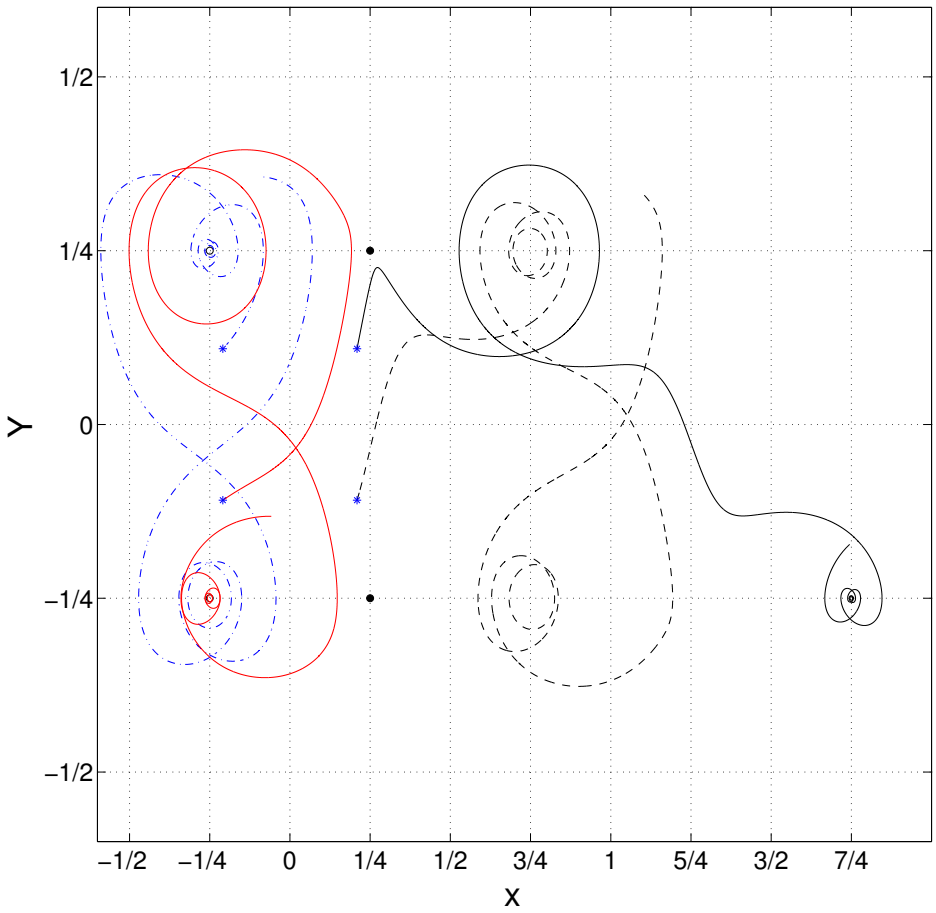


Figure 3.6: A display of trajectories of (3.35) are given where the starting points $(x_0, y_0) = (0.209, 0.109)$, $(x_0, y_0) = (-0.209, 0.109)$, $(x_0, y_0) = (0.209, -0.109)$, $(x_0, y_0) = (-0.209, -0.109)$ are indicated by *.

Type D: Equilibria in the third quadrant

$$x = y = -\frac{2k+1}{4}, \quad \text{for } k = 0, 1, 2, \ldots$$

The Jacobian (3.66) at this equilibrium becomes

$$\mathbf{D_x}(\mathbf{f}+\mathbf{g}) = 2\pi \begin{bmatrix} 0 & 2\pi u_0 + u_1(t) \\ 2\pi u_0 & v_1(t) \end{bmatrix}, \tag{3.77}$$

whose eigenvalues are given by

$$\lambda_{1,2}(t) = \frac{v_1(t) \pm \sqrt{\Delta_2(t)}}{2}, \tag{3.78}$$

where Δ_2 is given in (3.76).

The character of the equilibria of *Types A* and *B* depends on the sign of $\Delta_1(t)$ in (3.70) and that of the equilibria of *Types C* and *D* depends on the sign of $\Delta_2(t)$ in (3.76).

3.5.4 Conditions for the sign definiteness of $\Delta_i(t)$

From (3.70) and (3.76), since $v_1^2(t) \geq 0$, a necessary and sufficient condition for $\Delta_i(t) > 0$ for all $t \geq 0$ is given by

$$\Delta_1(t) \geq 0 \text{ when } \quad u_1(t) \geq 2\pi u_0 \qquad \text{if } \quad u_0 > 0, \tag{3.79a}$$

$$u_1(t) \leq 2\pi |u_0| \qquad \text{if } \quad u_0 < 0, \tag{3.79b}$$

and

$$\Delta_2(t) \geq 0 \text{ when } \quad u_1(t) \leq -2\pi u_0 \qquad \text{if } \quad u_0 > 0, \tag{3.80a}$$

$$u_1(t) \geq 2\pi |u_0| \qquad \text{if } \quad u_0 < 0. \tag{3.80b}$$

Since $u_1(t)$ depends on $\hat{\boldsymbol{\alpha}}$, for a given value of $(u_0, \hat{\boldsymbol{\alpha}})$, inequalities (3.79a) and (3.80a) cannot hold simultaneously. Similarly, (3.79b) and (3.80b) cannot also hold simultaneously. Thus, for a given $(u_0, \hat{\boldsymbol{\alpha}})$, if $\Delta_1(t) \geq 0$ then $\Delta_2(t) \leq 0$, and vice versa. This is in term implies that if the eigenvalues of *Type A* and *B* equilibria are real, then those of *Type C* and *D* are complex conjugates and vice versa. To further understand the nature of the real eigenvalues, consider the case when $u_0 > 0$ and $u_1(t) \geq 2\pi u_0$. For this choice,

$$\begin{aligned} \Delta_1(t) &= (v_1(t))^2 - 8\pi u_0(2\pi u_0 - u_1(t)) \\ &= (v_1(t))^2 + 8\pi u_0(u_1(t) - 2\pi u_0). \end{aligned} \tag{3.81}$$

Hence $\sqrt{\Delta_1(t)} > v_1(t)$ and the two eigenvalues are such that

$$\lambda_1(t) = \frac{-v_1(t) + \sqrt{\Delta_1(t)}}{2} \text{ and } \lambda_2(t) = \frac{-v_1(t) - \sqrt{\Delta_1(t)}}{2}. \tag{3.82}$$

This is, when $u_0 > 0$ and $u_1(t) > 2\pi u_0$, the *Type A* and *B* equilibria are *saddle* points while *Type C* and *D* are *centers*. Stated in other words, *Type A* and *Type B*, and *Types C* and *D* have a mutually complementary character. A similar argument carry over to the case when $u_0 < 0$.

A display of these four types of equilibria along with the flow field around them is given in Figure 3.5. Some sample solutions of (3.65) obtained numerically are also given in Figure 3.6.

3.6 Analysis of bifurcations

In Section 3.5 we have cataloged the properties of the set of all equilibria of the tracer dynamics in different regions of the four dimensional parameter space, $\mathbb{R}^4$

containing $(u_0, \hat{\boldsymbol{\alpha}})$ namely, $u_0 \neq 0$ and $\hat{\boldsymbol{\alpha}} = 0$ in Case 1, $u_0 = 0$ and $\hat{\boldsymbol{\alpha}} \neq 0$ in Case 2, $u_0 \neq 0$ and $\hat{\boldsymbol{\alpha}} \neq 0$ in Case 3. In this section we examine the transition between the equilibria as the parameters are varied continuously.

To this end we start by translating the conditions for the positive definiteness of $\Delta_i(t)$ in ((3.79a), (3.79b)) and ((3.80a), (3.80b)) directly in term of $(u_0, \hat{\boldsymbol{\alpha}})$ using (3.27). Define a new set of parameters U, H and V through a linear transformation suggested by the expressions for $u_1(t)$ in (3.30) as

$$\begin{aligned} U &= \frac{a^2}{\lambda^2} u_1(0) - \frac{a}{\lambda^2} h_1(0), \\ H &= \frac{1}{\lambda^2} u_1(0) + \frac{a}{\lambda^2} h_1(0), \end{aligned} \tag{3.83}$$

$$V = \frac{1}{\lambda} v_1(0). \tag{3.84}$$

Rewriting (3.83) in matrix form, it can be verified that

$$\begin{bmatrix} U \\ H \end{bmatrix} = \frac{1}{\lambda^2} \begin{bmatrix} a^2 & -a \\ 1 & a \end{bmatrix} \begin{bmatrix} u_1(0) \\ h_1(0) \end{bmatrix}. \tag{3.85}$$

The 2×2 matrix in (3.85) is non-singular. Hence (3.83) and (3.84) define an invertible linear transformation between $\big(u_1(0), v_1(0)), h_1(0)\big)^T \in \mathbb{R}^3$ to $(U, V, H) \in \mathbb{R}^3$.

In terms of these new variables, $u_1(t)$ in (3.27) now becomes

$$u_1(t) = U + H \cos(\lambda t) + V \sin(\lambda t). \tag{3.86}$$

Now invoking the results from Appendix B, we obtain a uniform (in time) lower

and upper bound on $u_1(t)$ given by

$$U - \sqrt{H^2 + V^2} \leq u_1(t) \leq U + \sqrt{H^2 + V^2}. \tag{3.87}$$

Using these bounds, we now translate the conditions for sign definiteness of $\Delta_i(t)$ in (3.79) - (3.80) in terms of these new set of parameters. It can be verified that

$$\Delta_1(t) \geq 0 \text{ when } U - \sqrt{H^2 + V^2} \geq 2\pi u_0 \text{ and } U \geq 2\pi u_0 \quad \text{if } u_0 > 0, \tag{3.88a}$$

$$U + \sqrt{H^2 + V^2} \leq -2\pi|u_0| \text{ and } U \leq -2\pi|u_0| \quad \text{if } u_0 < 0, \tag{3.88b}$$

and

$$\Delta_2(t) \geq 0 \text{ when } U + \sqrt{H^2 + V^2} \leq -2\pi u_0 \text{ and } U \leq 2\pi u_0 \quad \text{if } u_0 > 0, \tag{3.89a}$$

$$U - \sqrt{H^2 + V^2} \geq 2\pi|u_0| \text{ and } U \geq 2\pi|u_0| \quad \text{if } u_0 < 0. \tag{3.89b}$$

Without the loss of generality, in the following we examine the bifurcation in the new parameter space $(U, V, H) \in \mathbb{R}^3$ and $u_0 \in \mathbb{R}$. We consider two cases.

3.6.1 Case 1

To visualize these conditions graphically, consider first the condition for $\Delta_1(t) \geq 0$ where $u_0 > 0$, namely,

$$U - \sqrt{H^2 + V^2} \geq 2\pi u_0, \text{ and } U \geq 2\pi u_0. \tag{3.90}$$

Rewriting (3.90) as

$$\frac{(U-2\pi u_0)^2}{V^2} - \frac{H^2}{V^2} \geq 1 \text{ and } U \geq 2\pi u_0, \tag{3.91}$$

and referring to Appendix C, it turns out that (3.91) with the equality sign indeed represents the equation for a standard hyperbola in the $U-H$ (two dimensional) plane with the following key characteristics: the hyperbola is centered at $(2\pi u_0, 0)$, its semi-major and semi-minor axes are equal and are given by V; its eccentricity is $e=\sqrt{2}$ and the slopes of its asymptotes are ± 1. Since $U \geq 2\pi u_0$, we only have to consider the right branch of hyperbola. Notice that the geostrophic parameter u_0 only affects the location of the center of the hyperbola and $v_1(0)$ (through V) only affects the length of the semi-axes. It is interesting to note that while the parameter space is really four dimensional, we can succinctly represent the effects of all the parameters using the standard hyperbola in two dimensions. Refer to Figure 3.7 for an illustration.

Now consider the case for $\Delta_2(t) \geq 0$ when $u_0 > 0$, namely

$$U + \sqrt{H^2+V^2} \leq -2\pi u_0, \quad \text{and } U \leq -2\pi u_0. \tag{3.92}$$

Rewriting (3.92) as

$$\frac{(U+2\pi u_0)^2}{V^2} - \frac{H^2}{V^2} \leq 1, \quad \text{and } U \leq -2\pi u_0, \tag{3.93}$$

it can be verified that (3.93) with equality sign denotes a hyperbola which shares all the characteristics of the one described above but centered $(-2\pi u_0, 0)$. Refer to Figure 3.8 for an illustration. Again, in view of the constraint $U \leq 2\pi u_0$, we only have to consider the left branch of hyperbola.

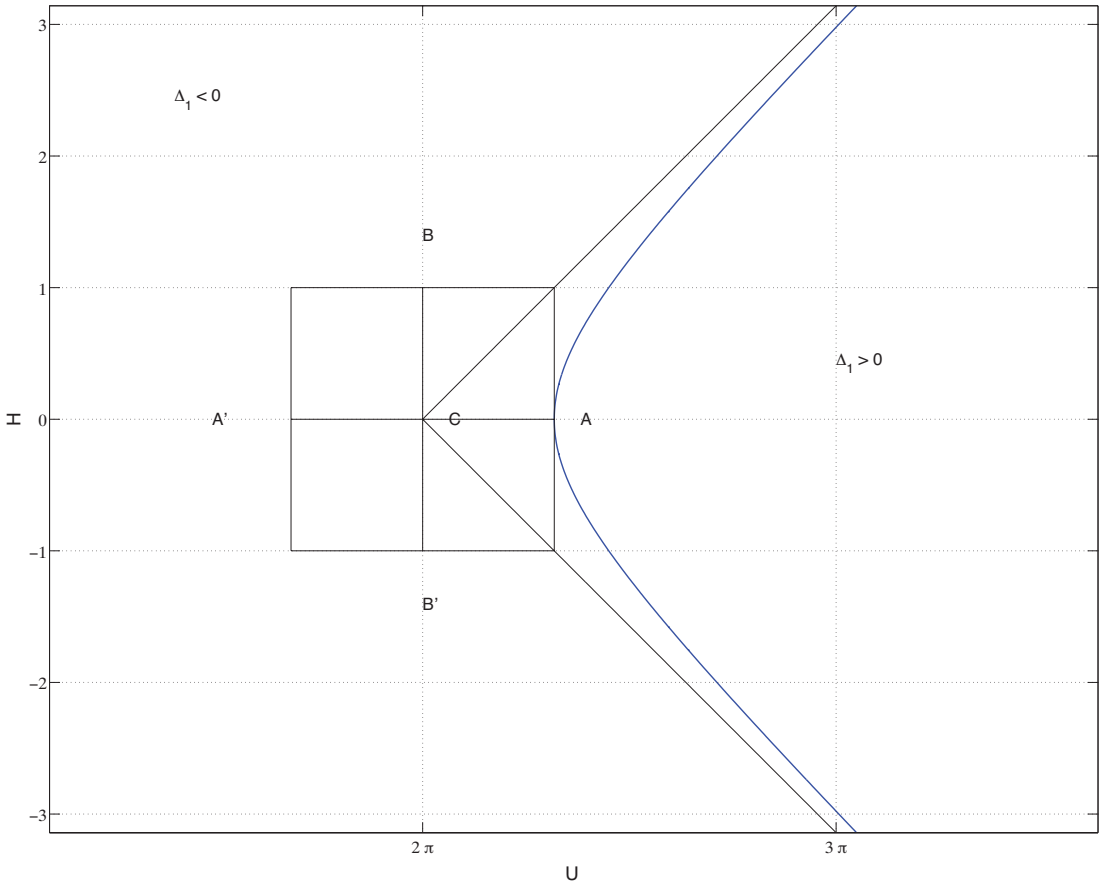


Figure 3.7: Hyperbola corresponding to (14.9) C in the center at $(2\pi u_0, 0)$ for $u_0 = 1$ Let $v_1(0) = 1$ and the semi axes $AC = AC' = BC = 1$. The asymptotes have slope ± 1.

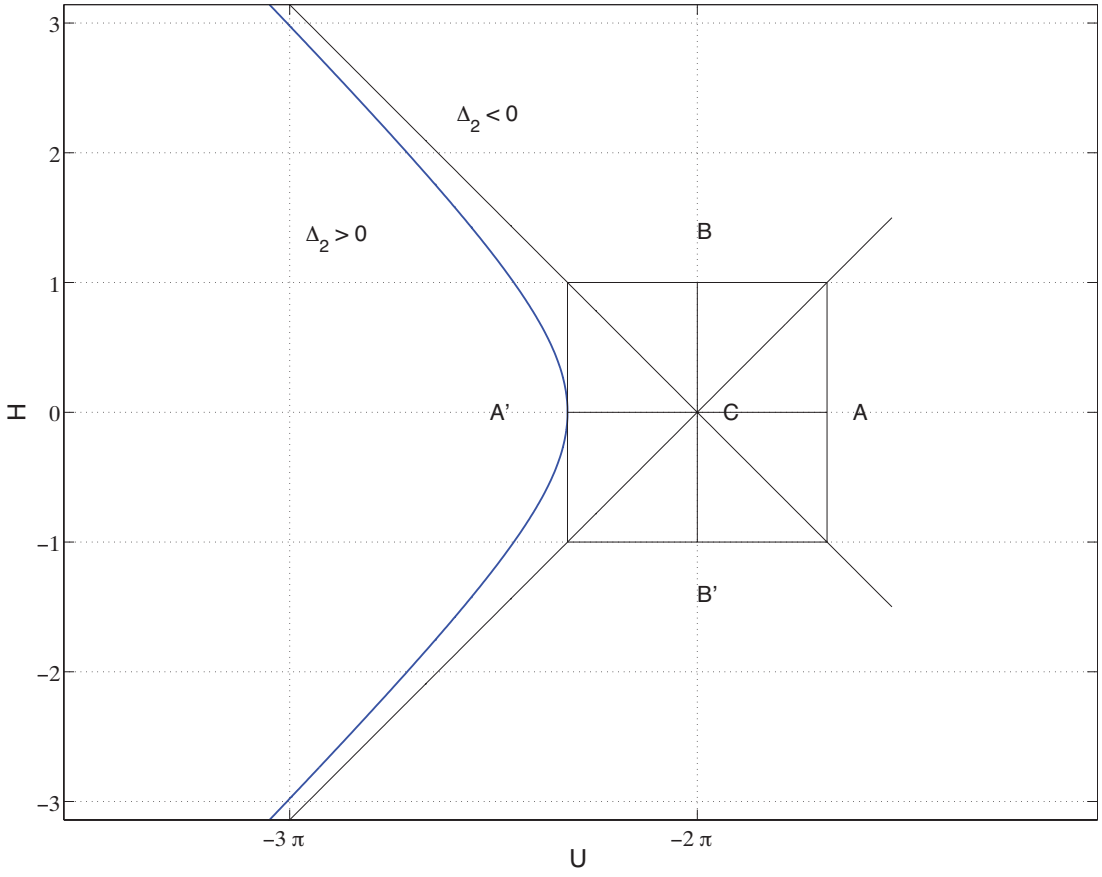


Figure 3.8: Hyperbola corresponding to (14.11) C in the center at $(-2\pi u_0, 0)$ with $u_0 = 1$ Let $v_1(0) = 1$ and the semi axes $CA = CA' = CB = CB' = 1$. The asymptotes have slope ± 1.

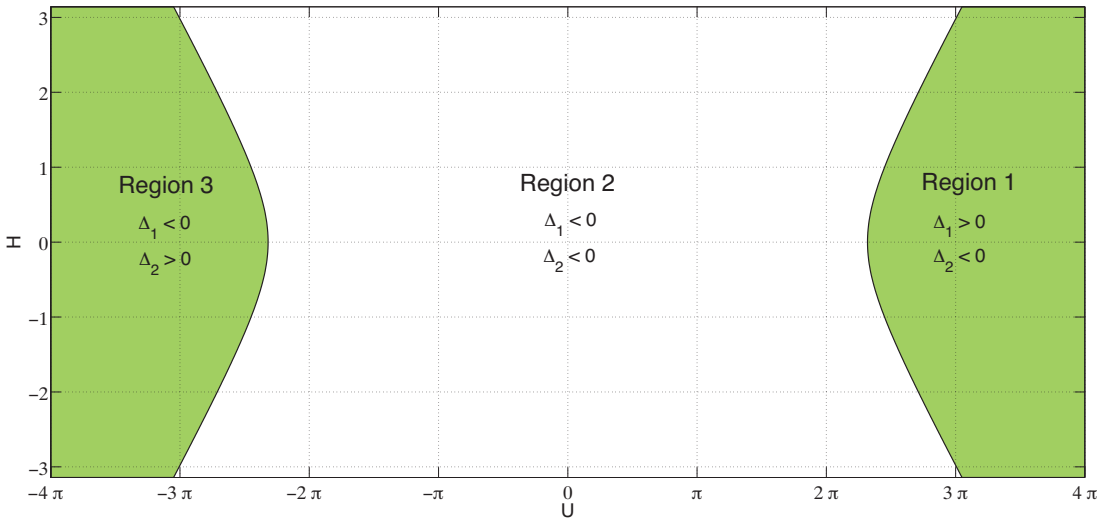


Figure 3.9: The combined system of hyperbolas from Figure 3.7 and 3.8. Regions corresponding to different signs of $\Delta_1(t)$ and $\Delta_2(t)$ are shown. Points on the hyperbola are the bifurcation points.

The complete characterization of the four types of equilibria in Case 3 in Section 4 is obtained by superimposing two systems of hyperbolas in Figures 3.7 and 3.8. This combined system along with the complete characteristic of various regions are given in Figure 3.9.

When $u_0 \to 0$, the center of the hyperbolas move towards the origin of the $U - H$ plane. Similarly, when $v_1(0) \to 0$, while the eccentricity $e = \sqrt{2}$ the slopes of the asymptotes remain constant at ± 1, the lengths of the semi axes shrink to zero.

3.6.2 Case 2

From (3.88), the condition for $\Delta_1(t) \geq 0$, when $u_0 \leq 0$, becomes

$$\frac{(U + 2\pi\,|u_0|)^2}{V^2} - \frac{H^2}{V^2} \leq 1, \quad \text{and } U \leq -2\pi|u_0|. \tag{3.94}$$

and from (3.89) the conditions for $\Delta_2(t) \geq 0$, and $u_0 \leq 0$, becomes

$$\frac{(U - 2\pi\,|u_0|)^2}{V^2} - \frac{H^2}{V^2} \geq 1, \quad \text{and } U \leq -2\pi|u_0|. \tag{3.95}$$

Equations (3.94)-(3.95) with equality sign again represent the system of hyperbolas whose properties are quite similar to the one for the case $u_0 > 0$. For completeness, we provide a snapshot of the field plot at time $t = 0$ and $t = 0.5$ for values of parameters in *Region 1, 2 and* 3, in figures 3.10, 3.11 and 3.12 (in pages 76-80) respectively. Figure 3.13 gives the field plot at time $t = 0$ and $t = 0.5$ corresponding to a bifurcation point on the boundary between *Regions 1* and *2.* Similarly, figure 3.14 provides the field plot at time $t = 0$ and $t = 0.5$ for a bifurcation point on the boundary between *Regions 2* and *3.*

3.7 Tracer Dynamics in the Linearized Shallow Water Model

Now substituting (3.27) into (3.1), we get explicit expression for $u(t), v(t)$, and, $h(t)$, which is the solution of (2.19). Given that we know $u(t), v(t)$, and $h(t)$ we can now compute the solution of (3.31) setting

$$\begin{aligned} \dot{x}_i(t) &= u(t), \\ \dot{y}_i(t) &= v(t), \end{aligned} \tag{3.96}$$

and integrating

$$\begin{aligned} x_i(t) &= x_i(0) + \int_0^t u(\tau)d\tau, \\ y_i(t) &= y_i(0) + \int_0^t v(\tau)d\tau. \end{aligned} \tag{3.97}$$

Therefore, we can compute the position of the i^{th} drifter using equations (3.27) - (3.97).

System before substitution of amplitudes listed above

$$\begin{aligned} u(t) &= -2\pi l \quad \sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t), \\ v(t) &= +2\pi k \quad \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t), \\ h(t) &= \qquad\quad \sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t). \end{aligned} \tag{3.98}$$

Summarizing all our efforts, we get explicit equations for tracer positions as

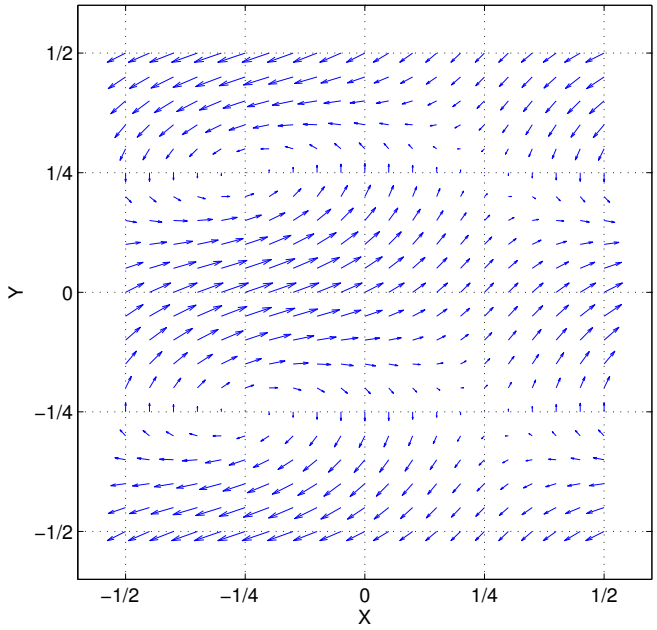
follows:

$$
\begin{aligned}
u(t) = & -2\,\pi\,l\,\sin(2\,\pi\,k\,x)\,\cos(2\,\pi\,l\,y)\,u_0 \\
& + \cos(2\,\pi\,m\,y)\left(\frac{\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+4\,m^2\,\pi^2\right)}{4\,\pi^2\,m^2+1}u_1(0)\right. \\
& + \frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}v_1(0) \\
& \left. + \frac{2\,\pi\,m\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}h_1(0)\right), \\
v(t) = & \ 2\,\pi\,k\ \cos(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,u_0 \\
& - \cos(2\,\pi\,m\,y)\left(\frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}u_1(0)\right. \\
& - \cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)v_1(0) \\
& \left. + \frac{2\,\pi\,m\ \sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}h_1(0)\right), \\
h(t) = & \ \sin(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,u_0 \\
& + \sin(2\,\pi\,m\,y)\left(\frac{2\,\pi\,m\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}u_1(0)\right. \\
& + \frac{2\,\pi\,m\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}v_1(0) \\
& \left. + \frac{\left(4\,m^2\,\pi^2\,\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+1\right)}{4\,\pi^2\,m^2+1}h_1(0)\right).
\end{aligned}
\tag{3.99}
$$

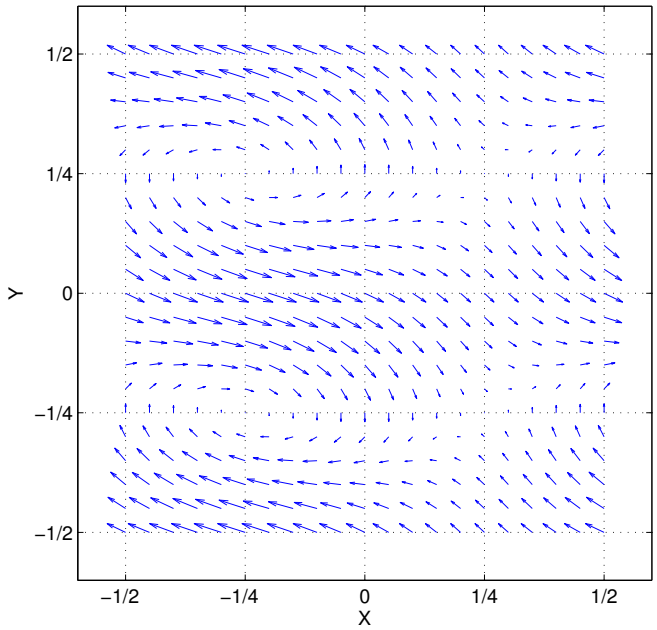
3.8 Summary

In this chapter we have presented a new analytical solution of the linear coupled system of three PDE's that describe amplitudes of the linear shallow water model (LSWM). This, in term, allowed us to give a complete stability analysis of the LSWM, after dividing model into three distinct modes (geostrophic, inertial-gravity and combined). It can be noted that the *geostrophic mode* is related to a cellular flow field with hyperbolic fixed points located at $(x, y) = (i/2k, j/2l), i, j \in \mathbb{Z}$. This flow is prevented from mixing by separatrices. Introduction of the time-dependent *inertial-gravity modes* brings mixing. We were able to find equilibria points of SWM at hand, for each of the three cases. The bifurcation analysis was done; the four dimensional parameter space was converted into two dimensional space with the system of hyperbolas indicating bifurcations.

In addition, we have expressed tracer dynamics in a closed form with the use of the above mentioned analytical expressions for amplitudes. These findings were presented by Jabrzemski and Lakshmivarahan (2013) at [5]. These analytical expressions for tracer dynamics are going to be used for data assimilation and sensitivity analysis in the subsequent chapters.

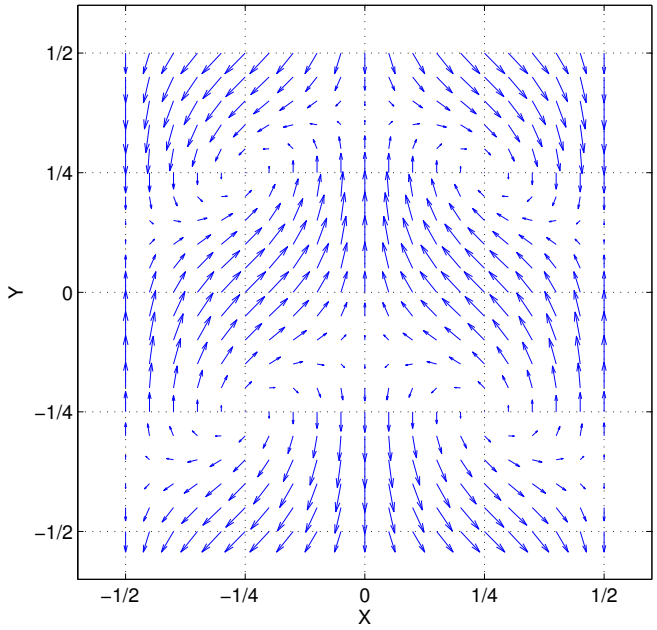


(a) Time t = 0.0

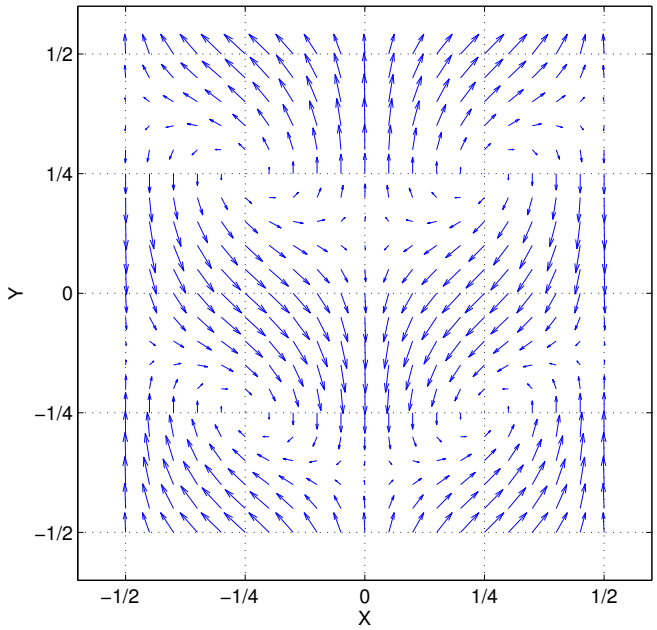


(b) Time t = 0.5

Figure 3.10: A snapshot of the time varying vector field given by (3.31) at time $t = 0$ and $t = 0.5$, where $u_0 = 1.0$, $u_1(0) = 12.5664, v_1(0) = \lambda$ and $h_1(0) = -2.0$. This corresponds to $V(0) = 1$, $U(0) = 4\pi$ and $H(0) = 0$, which is a point in *Region 1* in Figure 9.

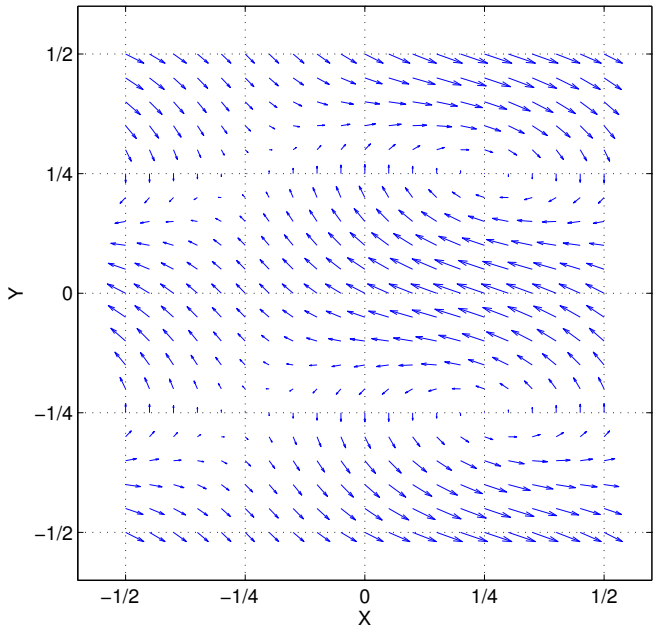


(a) Time t = 0.0

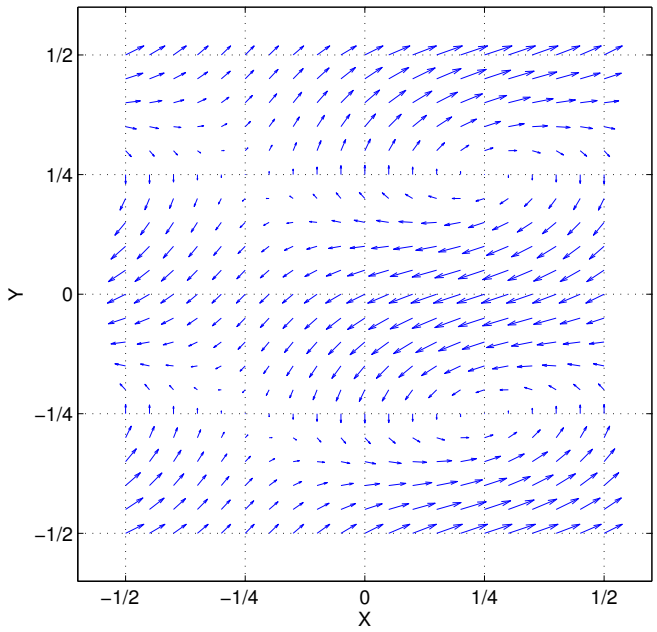


(b) Time t = 0.5

Figure 3.11: A snapshot of the time varying vector field given by (3.31) $t = 0$ and $t = 0.5$, where $u_0 = 1.0$, $u_1(0) = 0.0, v_1(0) = \lambda$ and $h_1(0) = 0.0$ This corresponds to $V(0) = 1$, $U(0) = 0$ and $H(0) = 0$, which is a point in *Region 2* in Figure 9.

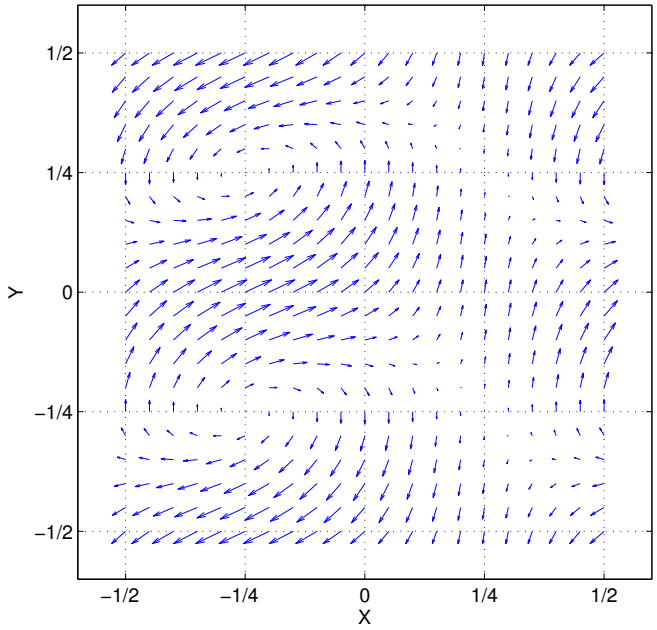


(a) Time t = 0.0

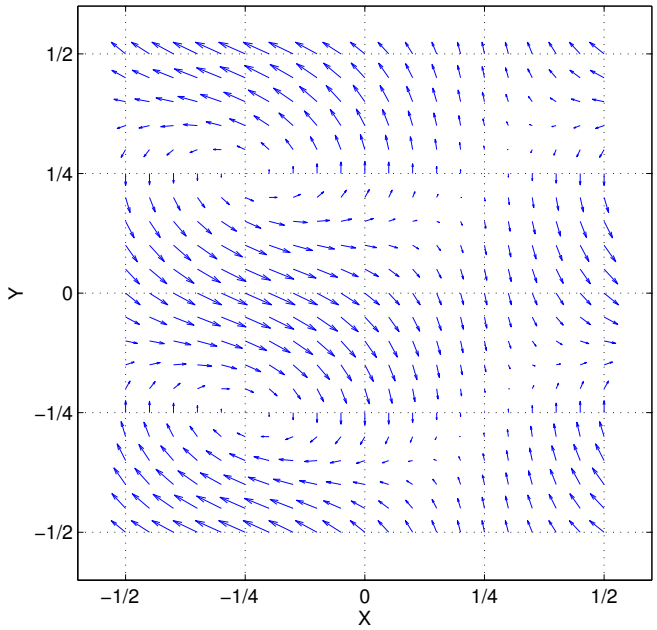


(b) Time t = 0.5

Figure 3.12: A snapshot of the time varying vector field given by (3.31) A snapshot of the time varying vector field given by (3.31) at time $t = 0$ and $t = 0.5$, where $u_0 = 1.0$, $u_1(0) = -12.5664, v_1(0) = \lambda$ and $h_1(0) = 2.0$. This corresponds to $V(0) = 1, U(0) = -4\pi$ and $H(0) = 0$, which is a point in *Region 3* in Figure 9.

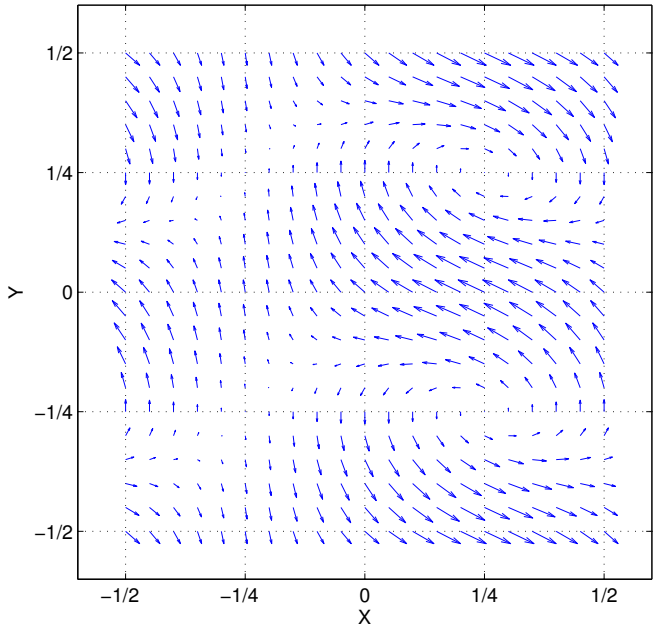


(a) Time t = 0.0

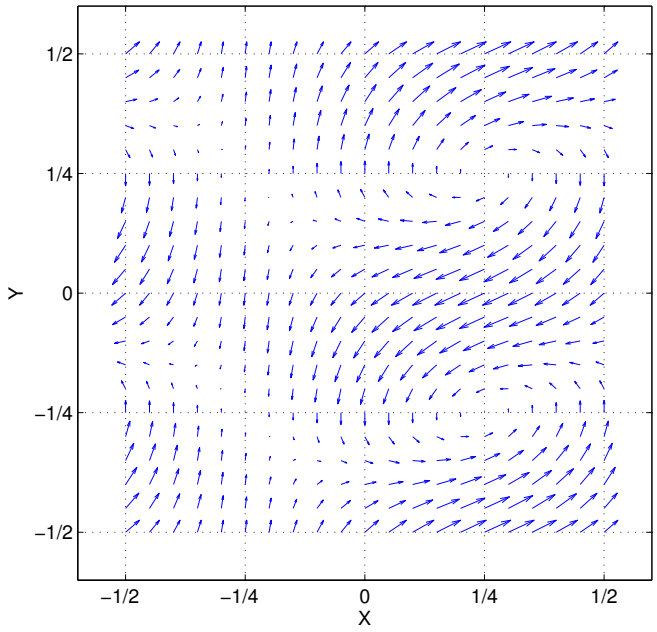


(b) Time t = 0.5

Figure 3.13: A snapshot of the time varying vector field given by (3.31) A snapshot of the time varying vector field given by (3.31) at time $t = 0$ and $t = 0.5$, where $u_0 = 1.0$, $u_1(0) = 7.2832, v_1(0) = \lambda$ and $h_1(0) = -1.1592$. This corresponds to $V(0) = 1, U(0) = 2\pi + 1$ and $H(0) = 0$, which is a bifurcation point on the hyperbola separating *Region 1* and *2* in Figure 9.



(a) Time t = 0.0



(b) Time t = 0.5

Figure 3.14: A snapshot of the time varying vector field given by (3.31) A snapshot of the time varying vector field given by (3.31) $t = 0$ and $t = 0.5$, where $u_0 = 1.0$, $u_1(0) = -7.2832, v_1(0) = \lambda$ and $h_1(0) = 1.1592$. This corresponds to $V(0) = 1$, $U(0) = -2\pi - 1$ and $H(0) = 0$, which is a bifurcation point on the hyperbola separating *Region 2* and *3* in Figure 9.

Chapter 4

A Framework for Data Assimilation

4.1 Introduction

We are going to follow the approach known as a forward sensitivity method (FSM) to data assimilation as its value was demonstrated by Laksmivarahan and Lewis in [10]. The have demonstrated the utility of FSM to improve correspondence of model results and observations, and pointed out that sensitivity of the model to elements of control vector provides insight into model dynamics. They have also demonstrated that there is a dual relationship to equations stemming from 4D-Var/adjoint equations using a simplified air-sea interaction model in Lakshmivarahan (2010) [9] and [10].

4.2 Model

In our dissertation we are going to focus on horizontal components of tracer positions. Therefore, our model can be categorized as a deterministic dynamical system, where $x(t)$ and $y(t)$ describe its state, as shown in equations (4.1). Here, $t \geq 0$ denotes the time, and $\boldsymbol{\alpha}$ belongs to the parameter space $\mathbb{R}^4$. We state this system of ordinary differential equations again for completeness

$$\begin{aligned} u = \dot{x} &= F_1(x, y, \boldsymbol{\alpha}), \\ v = \dot{y} &= F_2(x, y, \boldsymbol{\alpha}), \end{aligned} \tag{4.1}$$

We can define

$$\begin{aligned} \mathbf{X} = \begin{pmatrix} x \\ y \end{pmatrix}, \dot{\mathbf{X}} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix}, \text{ and} \\ \mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}, \end{aligned} \tag{4.2}$$

where $\mathbf{X} \in \mathbb{R}^2$, $\mathbf{F} \in \mathbb{R}^2$, and $\mathbf{X}(0) \in \mathbb{R}^2$ gives initial conditions. We assume that that the solution $\mathbf{X}(t)$ exists and is unique. The time derivatives of $x(t)$ and $y(t)$ are indicated respectively by $\dot{x}$ and $\dot{y}$. The equation (4.1) becomes

$$\dot{\mathbf{X}} = \mathbf{F}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right). \tag{4.3}$$

Discretizing the above equation using an Euler scheme (k indicates the time step), we get:

$$\frac{\mathbf{X}(k+1) - \mathbf{X}(k)}{\Delta t} = \mathbf{F}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right),$$

or

$$\mathbf{X}(k+1) = \mathbf{X}(k) + \mathbf{F}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right) \Delta t. \tag{4.4}$$

More generally, this can be stated as

$$\mathbf{X}(k+1) = \mathbf{M}(\mathbf{X}(k), \boldsymbol{\alpha}), \text{ where } \mathbf{M}(\mathbf{X}(k), \boldsymbol{\alpha}) = \mathbf{X}(k) + \mathbf{F}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right) \Delta t. \tag{4.5}$$

So, our *physical model* is given by the above equation, where

$$\mathbf{M}: \ \mathbb{R}^2 \times \mathbb{R}^4 \to \mathbb{R}^2, \tag{4.6}$$

$$\mathbf{M}(\mathbf{X}(k), \boldsymbol{\alpha}) = \begin{bmatrix} M_1(\mathbf{X}(k), \boldsymbol{\alpha}) \\ M_2(\mathbf{X}(k), \boldsymbol{\alpha}) \end{bmatrix}. \tag{4.7}$$

4.2.1 Observations

If it is assumed that we have a set of N observations of the position of the tracer

$$\mathbf{z}(k) = \mathbf{h}(\hat{\mathbf{X}}(k), \boldsymbol{\alpha}) + \boldsymbol{\nu}(k), \text{where } \ \mathbf{h}: \mathbb{R}^2 \times \mathbb{R}^4 \ \text{ and } \ \boldsymbol{\nu}(k) \in N(0, \sigma^2) \tag{4.8}$$

The mapping $\mathbf{h}(\cdot)$ can be called the *observation operator*, or the *forward operator*. We assume that $\boldsymbol{\nu}(t)$ is a white Gaussian noise.

4.2.2 Objective function

We define our objective function as

$$
\begin{aligned}
&J(\mathbf{X}(0), \boldsymbol{\alpha}) = \frac{1}{2\sigma^2} \sum_{k=1}^{N} [\mathbf{z}(k) - \mathbf{h}(\mathbf{X}(k), \boldsymbol{\alpha})]^2 = \frac{1}{2\sigma^2} \sum_{k=1}^{N} J_k(\mathbf{X}(0), \boldsymbol{\alpha}), \text{ where} \\
&\mathbf{X}(0) = \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} \text{ is the initial position of the tracers,} \\
&\boldsymbol{\alpha} = (u_0, u_1(0), v_1(0), h_1(0))^T, \text{ and } \sigma^2 \text{ is the variance of observations.}
\end{aligned} \tag{4.9}
$$

Our goal is to estimate $\mathbf{X}(0)$ and/or $\boldsymbol{\alpha}$ based on the set of N observations (4.8) to minimize (4.9) w.r.t $\mathbf{X}(0)$ and $\boldsymbol{\alpha}$, where $\mathbf{X}(k)$ are constrained by the model dynamics. Therefore, we have a *constrained optimization problem with equality constraints*.

Fact 1

Given $\mathbf{h}(\mathbf{X}, \boldsymbol{\alpha})$, if $\mathbf{X}$ is perturbed by $\delta\mathbf{X}$ and $\boldsymbol{\alpha}$ by $\delta\boldsymbol{\alpha}$, then

$$
\begin{aligned}
&\mathbf{h}(\mathbf{X} + \delta\mathbf{X}, \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) \approx \mathbf{h}(\mathbf{X}, \boldsymbol{\alpha}) + \mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X} + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}, \text{ where} \\
&\mathbf{D}_{\mathbf{X}}(\mathbf{h}) = \left[\frac{\partial \mathbf{h}}{\partial x}, \frac{\partial \mathbf{h}}{\partial y} \right] \text{ and} \\
&\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h}) = \left[\frac{\partial \mathbf{h}}{\partial u_0}, \frac{\partial \mathbf{h}}{\partial u_1(0)}, \frac{\partial \mathbf{h}}{\partial v_1(0)}, \frac{\partial \mathbf{h}}{\partial h_1(0)} \right].
\end{aligned} \tag{4.10}
$$

Here $\mathbf{D}_{\mathbf{X}}(\mathbf{h}) \in \mathbb{R}^{1\times 2}$ indicates the Jacobian of the *forward operator* $\mathbf{h}(\cdot)$ with respect to $\mathbf{X}$, and $\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h}) \in \mathbb{R}^{1\times 4}$ represents the Jacobian of the *forward operator* $\mathbf{h}(\cdot)$ with respect to $\boldsymbol{\alpha}$. Therefore, from (4.10), the variation in $\mathbf{h}$ denoted by $\delta\mathbf{h}$

is given by

$$\delta\mathbf{h} = h(\mathbf{X} + \delta\mathbf{X}, \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) - \mathbf{h}(\mathbf{X}, \boldsymbol{\alpha}) = \mathbf{D}_{\mathbf{X}}(\mathbf{h})\,\delta\mathbf{X} + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\,\delta\boldsymbol{\alpha}. \tag{4.11}$$

Fact 2

Let as define

$$J_k(\mathbf{X}(0), \boldsymbol{\alpha}) = \frac{1}{2\sigma^2}[\mathbf{z}(k) - \mathbf{h}(\mathbf{X}(k), \boldsymbol{\alpha})]^2. \tag{4.12}$$

Then we can express its first variation as

$$\delta J_k = \frac{-1}{\sigma^2}[\mathbf{z}(k) - \mathbf{h}(\mathbf{X}(k), \boldsymbol{\alpha})]\delta\mathbf{h}, \tag{4.13}$$

and further, using equation (4.11), we get

$$\delta J_k = \frac{-1}{\sigma^2}\left[\mathbf{z}(k) - \mathbf{h}(\mathbf{X}(k), \boldsymbol{\alpha})\right]\left[\mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X}\Big|_{\mathbf{X}=\mathbf{X}(k)} + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}\Big|_{\mathbf{X}=\mathbf{X}(k)}\right]. \tag{4.14}$$

If we define a vector of forecast errors

$$\mathbf{e}(\mathbf{k}) = \mathbf{z}(k) - \mathbf{h}(\mathbf{X}(k), \boldsymbol{\alpha}), \tag{4.15}$$

then we can express variation (4.14) as

$$\delta J_k = \frac{-\mathbf{e}(k)}{\sigma^2}\left[\mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X}(k) + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}\right]. \tag{4.16}$$

Fact 3

We can relate $\delta\mathbf{X}(k)$ to $\delta\mathbf{X}(0)$ and $\delta\boldsymbol{\alpha}$ using a simple model, that we depict below:

$$\begin{array}{ccc} (\mathbf{X}(0), \boldsymbol{\alpha}) & \longrightarrow & (\mathbf{X}(k)) \\ \downarrow & & \downarrow \\ (\mathbf{X}(0) + \delta\mathbf{X}(0), \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}) & \longrightarrow & (Y(k) = \mathbf{X}(k) + \delta\mathbf{X}(k)) \end{array} \tag{4.17}$$

Clearly $\delta\mathbf{X}(k)$ is the first variation in $\mathbf{X}(k)$ resulting from a first variation $\delta\mathbf{X}(0)$ in $\mathbf{X}(0)$ and from $\delta\boldsymbol{\alpha}$ in $\boldsymbol{\alpha}$. From first principles:

$$\delta\mathbf{X}(k) = \left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\delta\mathbf{X}(0) + \left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\delta\boldsymbol{\alpha}, \tag{4.18}$$

with the forward sensitivity of the solution at time k w.r.t $\mathbf{X}(0)$ is given by the Jacobian

$$\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right] = \mathbf{D}_{\mathbf{X}(0)}[\mathbf{X}(k)],$$

and the forward sensitivity of the solution at time k w.r.t $\boldsymbol{\alpha}$ is represented by the Jacobian

$$\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha})}\right] = \mathbf{D}_{\boldsymbol{\alpha}}[\mathbf{X}(k)].$$

Fact 4

Consider the first variation of J in equation (4.9) from *Fact 2*.

$$\begin{aligned}
\delta J &= \frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\delta\mathbf{h} \ \{\text{from Fact 2}\} \\
&= \frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\left[\mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X}(k)+\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}\right] \ \{\text{from Fact 1}\} \\
&= \frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\left[\mathbf{D}_{\mathbf{X}}(\mathbf{h})\left(\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\delta\mathbf{X}(0)+\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\delta\boldsymbol{\alpha}\right)+\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}\right] \\
&= \frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X}(0)+\frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\boldsymbol{\alpha} \\
&\quad+\frac{-1}{\sigma^2}\sum_{k=1}^{N}\mathbf{e}(k)\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}.
\end{aligned} \tag{4.19}$$

Let $<\mathbf{a},\mathbf{b}>$ denote the inner product. We can apply it as follows:

$$\begin{aligned}
\delta J &= \frac{-1}{\sigma^2}\sum_{k=1}^{N}\left\langle\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\mathbf{D}_{\mathbf{X}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\mathbf{X}(0)\right\rangle \\
&\quad-\frac{1}{\sigma^2}\sum_{k=1}^{N}\left\langle\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\mathbf{D}_{\mathbf{X}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\boldsymbol{\alpha}\right\rangle \\
&\quad-\frac{1}{\sigma^2}\sum_{k=1}^{N}\left\langle\mathbf{D}_{\boldsymbol{\alpha}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\boldsymbol{\alpha}\right\rangle \\
&= \left\langle\frac{-1}{\sigma^2}\sum_{k=1}^{N}\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\mathbf{D}_{\mathbf{X}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\mathbf{X}(0)\right\rangle \\
&\quad-\left\langle\frac{1}{\sigma^2}\sum_{k=1}^{N}\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\mathbf{D}_{\mathbf{X}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\boldsymbol{\alpha}\right\rangle \\
&\quad-\left\langle\frac{1}{\sigma^2}\sum_{k=1}^{N}\mathbf{D}_{\boldsymbol{\alpha}}^{T}(\mathbf{h})\mathbf{e}(k),\delta\boldsymbol{\alpha}\right\rangle.
\end{aligned} \tag{4.20}$$

Using first principles, we can express it as

$$\delta J = \left\langle \nabla_{\mathbf{X}(0)} J, \delta \mathbf{X}(0) \right\rangle + \langle \nabla_{\boldsymbol{\alpha}} J, \delta \boldsymbol{\alpha} \rangle \,. \tag{4.21}$$

Comparing the like terms on the r.h.s of (4.19) and (4.20), we can get

$$\nabla_{\mathbf{X}(0)} J = \frac{-1}{\sigma^2} \sum_{k=1}^{N} \left[\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)} \right]^T \mathbf{D}_X^T(\mathbf{h})\, \mathbf{e}(k), \tag{4.22}$$

$$\begin{aligned} \nabla_{\boldsymbol{\alpha}} J = \frac{-1}{\sigma^2} \sum_{k=1}^{N} \left[\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}} \right]^T \mathbf{D}_X^T(\mathbf{h})\, \mathbf{e}(k) \\ - \frac{1}{\sigma^2} \sum_{k=1}^{N} \mathbf{D}_{\boldsymbol{\alpha}}^T(\mathbf{h}) \Big|_{\mathbf{X}=\mathbf{X}(k)} \mathbf{e}(k). \end{aligned} \tag{4.23}$$

Fact 5

Left hand side of (4.22) is called the *adjoint sensitivity*, if it is expressed as the sum of the products of the forward sensitivity $\left[\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)}\right]^T \mathbf{D}_{\mathbf{X}}^T(\mathbf{h})$ and the forecast errors $\mathbf{e}(k)$. Thus, $\nabla_{\mathbf{X}(0)} J$ will be non-zero at time k only if both $\left[\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)}\right]^T \mathbf{D}_{\mathbf{X}}^T(\mathbf{h})$ and $\mathbf{e}(k)$ are non-zero. If one of them is small, their contribution will be small. Similarly, the same can be noted for the other gradient $\nabla_{\boldsymbol{\alpha}} J$. Therefore, it is of interest to examine the evolution of $\left[\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)}\right]$ and $\left[\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}}\right]$ with time k.

Fact 6

Consider equation (4.5)

$$\mathbf{X}(k+1) = \mathbf{M}(\mathbf{X}(k), \boldsymbol{\alpha}).$$

We can differentiate both sides with respect to $\mathbf{X}(0)$ and get

$$\left[\frac{\partial \mathbf{X}(k+1)}{\partial \mathbf{X}(0)}\right] = D_{\mathbf{X}(k)}(\mathbf{M}) \left[\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)}\right] \text{ with I.C. } \frac{\partial \mathbf{X}(0)}{\partial \mathbf{X}(0)} = \mathbf{I}. \tag{4.24}$$

Solution of (4.24) gives the forward sensitivity $\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)}$ for all k. We can note that $\frac{\partial \mathbf{X}(k)}{\partial \mathbf{X}(0)} \in \mathbb{R}^{2\times 2}$. Analogously, we can differentiate it with respect to $\boldsymbol{\alpha}$ and obtain

$$\begin{aligned}\left[\frac{\partial \mathbf{X}(k+1)}{\partial \boldsymbol{\alpha}}\right] &= \mathbf{D}_{\mathbf{X}(k)}(\mathbf{M}) \left[\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}}\right] + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M}) \\ &= \mathbf{D}_{\mathbf{X}(k)}(\mathbf{M}) \left[\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}}\right] + \left[\frac{\partial \mathbf{M}_i(k)}{\partial \boldsymbol{\alpha}_j}\right] \text{ with I.C. } \frac{\partial \mathbf{X}(0)}{\partial \boldsymbol{\alpha}} \equiv 0.\end{aligned} \tag{4.25}$$

Solution of (4.25) gives the forward sensitivity with respect to $\boldsymbol{\alpha}$ for all k. Note that $\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}} \in \mathbb{R}^{2\times 4}$.

4.3 Data Assimilation using Forward Sensitivity Method (FSM)

Let $\mathbf{X}(k)$ be the solution at time k starting with $[\mathbf{X}(0), \boldsymbol{\alpha}]$, such that forecast error can be represented as

$$\mathbf{e}(k) = \mathbf{z}(k) - \mathbf{h}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right) \neq 0, \tag{4.26}$$

where $\mathbf{h}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right)^T$ is the forward operator with $\mathbf{X}(k) = \left(x(k), y(k)\right)^T$ and $\boldsymbol{\alpha} = \left(u_0, u_1(0), v_1(0), h_1(0)\right)$. Our goal is to find a correction $\left(\delta\mathbf{X}(0), \delta\boldsymbol{\alpha}\right)$ such that the new solution $\mathbf{Y}(k) = \mathbf{X}(k) + \delta\mathbf{X}(k)$ at time k will annihilate the forecast

error $\mathbf{e}(k)$.

$$\begin{aligned} \mathbf{z}(k) - \mathbf{h}\left(\mathbf{Y}(k), \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}\right) &= 0, \\ \mathbf{z}(k) - \mathbf{h}\left(\mathbf{X}(k) + \delta\mathbf{X}(k), \boldsymbol{\alpha} + \delta\boldsymbol{\alpha}\right) &= 0, \\ \underbrace{\mathbf{z}(k) - \mathbf{h}\left(\mathbf{X}(k), \boldsymbol{\alpha}\right)}_{\mathbf{e}(k)} - \left[\mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\delta\mathbf{X}(k) + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}\right] &= 0. \end{aligned} \tag{4.27}$$

This gives as expression for the forecast error as

$$\mathbf{e}(k) = \mathbf{D}_{\mathbf{X}}(\mathbf{h})\delta\mathbf{X}(k) + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha}. \tag{4.28}$$

But from *Fact 3* (equation (4.18)) we get

$$\delta\mathbf{X}(k) = \left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\delta\mathbf{X}(0) + \left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\delta\boldsymbol{\alpha}.$$

When we substitute equation (4.18) in equation (4.28) we express the forecast error as

$$\begin{aligned} \mathbf{e}(k) &= \mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\left(\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\delta\mathbf{X}(0) + \left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\delta\boldsymbol{\alpha}\right) + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha} \\ &= \mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\delta\mathbf{X}(0) + \mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right]\delta\boldsymbol{\alpha} + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h})\delta\boldsymbol{\alpha} \\ &= \left[\underbrace{\mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]}_{\mathbf{H}_1}, \quad \underbrace{(\mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\left[\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}\right] + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h}))}_{\mathbf{H}_2}\right]\begin{bmatrix}\delta\mathbf{X}(0)\\ \delta\boldsymbol{\alpha}\end{bmatrix}. \end{aligned} \tag{4.29}$$

Let $\mathbf{H}_1 = \mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\Big|_{1\times 2}\left[\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}\right]\Big|_{2\times 2}$ be a row vector of size two. We can compute $\frac{\partial\mathbf{X}(k)}{\partial\mathbf{X}(0)}$ using equation (4.24), and $\frac{\partial\mathbf{X}(k)}{\partial\boldsymbol{\alpha}}$ using equation (4.25).

Let $\mathbf{H}_2 = \mathbf{D}_{\mathbf{X}(k)}(\mathbf{h})\Big|_{1\times 2}\left[\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}}\right]\Big|_{2\times 4} + \mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{h}))\Big|_{1\times 4}$ be a row vector of size four. We can compute $\frac{\partial \mathbf{X}(k)}{\partial \boldsymbol{\alpha}}$ using equation (4.25).

Let

$$\boldsymbol{\zeta} = \begin{bmatrix} \partial \mathbf{X}(0) \\ \hline . \ . \ . \ . \\ \hline \partial \boldsymbol{\alpha} \end{bmatrix} \begin{matrix} \Big\} 2 \\ . \ . \in \mathbb{R}^6, \\ \Big\} 4 \end{matrix}$$

$$\begin{aligned} \Big[\mathbf{H}_1(k), \mathbf{H}_2(k)\Big] \begin{bmatrix} \partial \mathbf{X}(0) \\ \hline . \ . \ . \ . \\ \hline \partial \boldsymbol{\alpha} \end{bmatrix} &= \mathbf{e}(k), \ \text{ or} \\ \mathbf{H}(k)\boldsymbol{\zeta} &= \mathbf{e}(k) \end{aligned} \tag{4.30}$$

with $\big[\mathbf{H}_1(k), \mathbf{H}_2(k)\big]$ is a row vector of size 1×6. Equation (4.30) is a *linear least squares problem* (chapter 5, Lewis et at. (2006) [12]).

4.3.1 Multiple observations

Let there be N observations of the position of the tracer as it floats in the flow. Then we have one equation like (4.30) for each time, and we can combine them to get

$$\begin{aligned} \begin{bmatrix} \mathbf{H}(1) \\ \mathbf{H}(2) \\ \hline . \ . \ . \ . \\ \hline \mathbf{H}(N) \end{bmatrix} \boldsymbol{\zeta} &= \begin{bmatrix} \mathbf{e}(1) \\ \mathbf{e}(2) \\ \hline . \ . \ . \ . \\ \hline \mathbf{e}(N) \end{bmatrix}, \ \text{ or} \\ \mathbf{H}\boldsymbol{\zeta} &= \mathbf{e}, \end{aligned} \tag{4.31}$$

where

$$\begin{bmatrix} \mathbf{H}(1) \\ \mathbf{H}(2) \\ \cdot \cdot \cdot \cdot \\ \mathbf{H}(N) \end{bmatrix} \in \mathbb{R}^{N \times 6}, \text{ and } \begin{bmatrix} \mathbf{e}(1) \\ \mathbf{e}(2) \\ \cdot \cdot \cdot \cdot \\ \mathbf{e}(N) \end{bmatrix} \in \mathbb{R}^{N}. \tag{4.32}$$

If $N > 6$, the there are more equations than the number of unknowns, and the system is over-determined. If we apply the method of normal equations, then the solution is given by:

$$\boldsymbol{\zeta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{e}. \tag{4.33}$$

The equation giving the optimal solution can be expressed as

$$\left(\mathbf{H}^T\mathbf{H}\right)\zeta = \mathbf{H}^T\mathbf{e}. \tag{4.34}$$

This can also be solved by QR decomposition or using SVD. They are described in [12] in chapters 5 and 9. Thus, FSM based method gives the correction $\boldsymbol{\zeta}$. We can improve the estimate by iteratively repeating the above procedure numerous times.

4.4 Summary

We have provided a general description of the methodology that we are going to use for our data assimilation experiments. As indicated by Lakshmivarahan and Lewis (2010) [10], FSM allows for correction of the control vector and initial conditions given the governing equations of the physical model. We will use it as a diagnostic tool to better understand measurement placement within a time domain, and compare sensitivity following different starting points for our

experiments. We will start off with the investigation of the evolution of the sensitivity functions derived for the linearized shallow water model that is the focus of this dissertation. In addition, the FSM methodology outlined in the chatter will server as a tool for data assimilation that will follow.

Chapter 5

Sensitivity

5.1 Introduction

In this chapter we will provide the most fundamental definition of sensitivity to bring it together with FSM and apply ideas from FSM to find sensitivity of the linearized shallow water model to initial conditions and the control vector. We will derive analytical formulas for sensitivity, since we have a closed form solution for trajectories. Using these closed form expressions, we will conduct a series of numerical experiments to investigate sensitivity of the linearized shallow water model, and illustrate how placement of measurements influences our ability to improve model error.

The dependence of the model prediction on initial conditions and parameters has always been a part of the sensitivity analysis. A comprehensive overview of sensitivity analysis used for chemical models has been given by Rabitz et al. (1983) [17]. Rabitz has indicated that sensitivity analysis gives a measure of model error propagation, and a great insight into fine physical structure of the model. Obviously, there are some simple calculations that can be applied

repeatedly to gain a very general idea about this structure, but sensitivity analysis conducted with well-defined sensitivity gradients answers a wide variety of questions concerning mutual relation of all of the dependent and independent variables in the model.

5.2 Basic ideas

We can start by presenting a fundamental measure of sensitivity following (Appendix E, Lewis et al. (2006) [12]). For a scalar valued function F of a scalar variable x, we indicate $\Delta F(x)$ as a change stemming from a change Δx in x, that is, $\Delta F(x) = F(x + \Delta x) - F(x)$. As a next step, we introduce a relative change in $F(x)$ as $\frac{\Delta F(x)}{F(x)}$, and relative change in x as $\frac{\Delta x}{x}$. And finally we can formulate a measure of sensitivity of the dependent variable F to changes in the independent variable x, and apply the standard finite difference approximation to the derivative of $F(x)$ w.r.t x

$$S_F(x) = \frac{\frac{\Delta F(x)}{F(x)}}{\frac{\Delta x}{x}} = \frac{\Delta F(x)}{F(x)} \frac{x}{\Delta x} \approx \frac{dF(x)}{dx} \left(\frac{x}{F(x)} \right). \tag{5.1}$$

Clearly, $S_F(x)$ as a measure of sensitivity is related to the first derivative of F w.r.t x. Along these lines, we can relate the gradient of the dependent variable F to the measure of the first-order sensitivity.

5.3 Evolution of sensitivity of the shallow water model with respect to initial conditions and parameters

Let us present equations that reflect the sensitivity of the linearized shallow water model to the control vector and initial conditions. These equations listed below are used in our DA experiments using FSM and are derived from equations (3.99) that were presented in the closed form. We are only focusing on (x, y) components.

Let us bring equation (4.5) again; this and (3.99) gives a starting point for sensitivity analysis of the linearized shallow water model.

$$\mathbf{X}(k+1) = \mathbf{M}\left(\mathbf{X}(k), \alpha\right), \quad \text{where } \mathbf{M}\left(\mathbf{X}(k), \alpha\right) = \mathbf{X}(k) + \mathbf{F}\left(\mathbf{X}(k), \alpha\right) \Delta t.$$

Our *physical model* describes $x(t)$ and $y(t)$; it was mentioned in (4.6),

$$\mathbf{M}: \ \mathbb{R}^2 \times \mathbb{R}^4 \to \mathbb{R}^2,$$
$$\mathbf{M}(\mathbf{X}(k), \alpha) = \begin{bmatrix} M_1(\mathbf{X}(k), \boldsymbol{\alpha}) \\ M_2(\mathbf{X}(k), \boldsymbol{\alpha}) \end{bmatrix}.$$

$$
\begin{aligned}
M_1\left(\mathbf{X}(k),\alpha\right) &= x(t+1)\\
&= x(t) + \Big[-2\,\pi\,l\,\sin\big(2\,\pi\,k\,x(t)\big)\,\cos\big(2\,\pi\,l\,y(t)\big)\,u_0\\
&\quad + \cos\big(2\,\pi\,m\,y(t)\big)\,\left(\frac{\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+4\,m^2\,\pi^2\right)}{4\,\pi^2\,m^2+1}u_1(0)\right.\\
&\quad + \frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}v_1(0)\\
&\quad \left.+\ \frac{2\,\pi\,m\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}h_1(0)\right)\Big]\,\Delta t,
\end{aligned}
\tag{5.2}
$$

$$
\begin{aligned}
M_2\left(\mathbf{X}(k),\alpha\right) &= y(t+1)\\
&= y(t) + \Big[2\,\pi\,k\ \cos\big(2\,\pi\,k\,x(t)\big)\,\sin\big(2\,\pi\,l\,y(t)\big)\,u_0\\
&\quad - \cos\big(2\,\pi\,m\,y(t)\big)\,\left(\frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}u_1(0)\right.\\
&\quad - \cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)v_1(0)\\
&\quad \left.+\frac{2\,\pi\,m\ \sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}h_1(0)\right)\Big]\,\Delta t.
\end{aligned}
\tag{5.3}
$$

We have given general equations to find sensitivity in (4.10); here we provide Jacobians of the solution $\mathbf{X}(t)$ that will be used to trace sensitivity of the shallow water model:

$$
\begin{aligned}
\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M}) &= \left[\frac{\partial\mathbf{M}}{\partial u_0},\frac{\partial\mathbf{M}}{\partial u_1(0)},\frac{\partial\mathbf{M}}{\partial v_1(0)},\frac{\partial\mathbf{M}}{\partial h_1(0)}\right],\ \text{and}\\
\mathbf{D}_{\mathbf{X}}(\mathbf{M}) &= \left[\frac{\partial\mathbf{M}}{\partial x},\frac{\partial\mathbf{M}}{\partial y}\right].
\end{aligned}
\tag{5.4}
$$

5.3.1 Sensitivity to elements of the control vector α

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(1,1)} = -2\,\pi\,l\,\cos(2\,\pi\,l\,y)\,\sin(2\,\pi\,k\,x)\,, \tag{5.5}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(1,2)} = \frac{\cos(2\,\pi\,m\,y)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+4\,m^2\,\pi^2\right)}{4\,\pi^2\,m^2+1}, \tag{5.6}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(1,3)} = \frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)\,\cos(2\,\pi\,m\,y)}{\sqrt{4\,\pi^2\,m^2+1}}, \tag{5.7}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(1,4)} = \frac{2\,\pi\,m\,\cos(2\,\pi\,m\,y)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}, \tag{5.8}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(2,1)} = 2\,\pi\,k\,\cos(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,, \tag{5.9}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(2,2)} = -\frac{\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)\,\cos(2\,\pi\,m\,y)}{\sqrt{4\,\pi^2\,m^2+1}}, \tag{5.10}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(2,3)} = \cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)\,\cos(2\,\pi\,m\,y)\,, \tag{5.11}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(2,4)} = -\frac{2\,\pi\,m\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)\,\cos(2\,\pi\,m\,y)}{\sqrt{4\,\pi^2\,m^2+1}}, \tag{5.12}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(3,1)} = \sin(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,, \tag{5.13}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(3,2)} = \frac{2\,\pi\,m\,\sin(2\,\pi\,m\,y)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}, \tag{5.14}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(3,3)} = \frac{2\,\pi\,m\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)\,\sin(2\,\pi\,m\,y)}{\sqrt{4\,\pi^2\,m^2+1}}, \tag{5.15}$$

$$\mathbf{D}_{\boldsymbol{\alpha}}(\mathbf{M})_{(3,4)} = \frac{\sin(2\,\pi\,m\,y)\,\left(4\,m^2\,\pi^2\,\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+1\right)}{4\,\pi^2\,m^2+1}. \tag{5.16}$$

5.3.2 Sensitivity to elements of the initial conditions $X(0)$

$$\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(1,1)} = -4\,k\,l\,\pi^2\,u_0\,\cos(2\,\pi\,k\,x)\,\cos(2\,\pi\,l\,y)\,, \tag{5.17}$$

$$\begin{aligned}\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(1,2)} = {} & 4\,l^2\,\pi^2\,u_0\,\sin(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y) \\ & - 2\,\pi\,m\,\sin(2\,\pi\,m\,y)\,\left(\frac{v_1(0)\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}\right. \\ & + \frac{u_1(0)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+4\,m^2\,\pi^2\right)}{4\,\pi^2\,m^2+1} \\ & \left. + \frac{2\,\pi\,m\,h_1(0)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}\right),\end{aligned} \tag{5.18}$$

$$\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(2,1)} = -4\,k^2\,\pi^2\,u_0\,\sin(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,, \tag{5.19}$$

$$\begin{aligned}\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(2,2)} = {} & +4\,k\,l\,\pi^2\,u_0\,\cos(2\,\pi\,k\,x)\,\cos(2\,\pi\,l\,y) \\ & + 2\,\pi\,m\,\sin(2\,\pi\,m\,y)\,\left(\frac{u_1(0)\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}\right. \\ & - v_1(0)\,\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right) \\ & \left. + \frac{2\,\pi\,m\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}h_1(0)\right),\end{aligned} \tag{5.20}$$

$$\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(3,1)} = 2\,\pi\,k\,u_0\,\cos(2\,\pi\,k\,x)\,\sin(2\,\pi\,l\,y)\,, \tag{5.21}$$

$$\mathbf{D}_{\mathbf{X}(0)}(\mathbf{M})_{(3,2)} = 2\,\pi\,l\,u_0\,\sin(2\,\pi\,k\,x)\cos(2\,\pi\,l\,y) \tag{5.22}$$
$$+\,2\,\pi\,m\,\cos(2\,\pi\,m\,y)\left(\frac{h_1(0)\,\left(4\,m^2\,\pi^2\,\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)+1\right)}{4\,\pi^2\,m^2+1}\right.$$
$$+\,\frac{2\,\pi\,m\,v_1(0)\,\sin\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)}{\sqrt{4\,\pi^2\,m^2+1}}$$
$$\left.+\,\frac{2\,\pi\,m\,u_1(0)\,\left(\cos\left(t\,\sqrt{4\,\pi^2\,m^2+1}\right)-1\right)}{4\,\pi^2\,m^2+1}\right).$$

5.4 Numerical experiments

5.4.1 Experiment 5.1

In this experiment, we are running a combined mode, with control vector and initial conditions given in Table 5.1. The goal of this experiment is to show and analyze the sensitivity of the model to the initial conditions and elements of the control vector.

Table 5.1: Experiment 5.1: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.209	0.109	0.5	2.5e-05

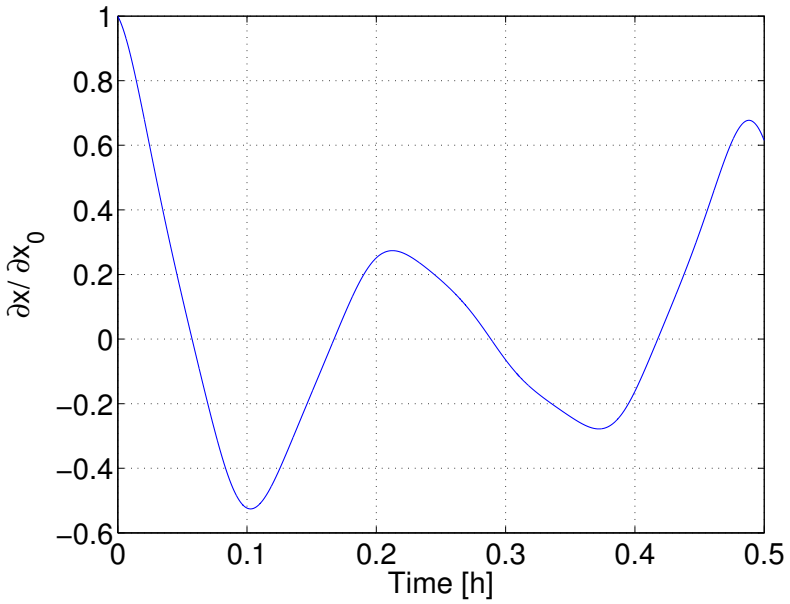


Figure 5.1: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. $x(0)$

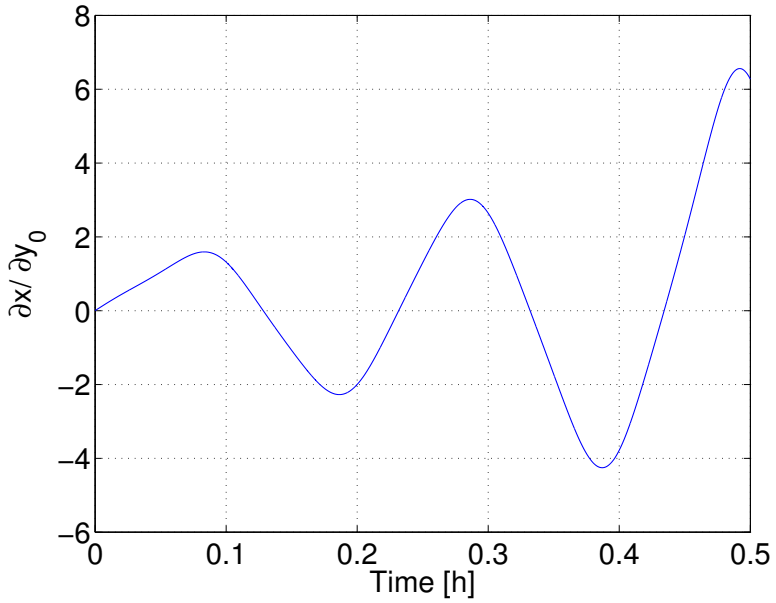


Figure 5.2: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. $y(0)$

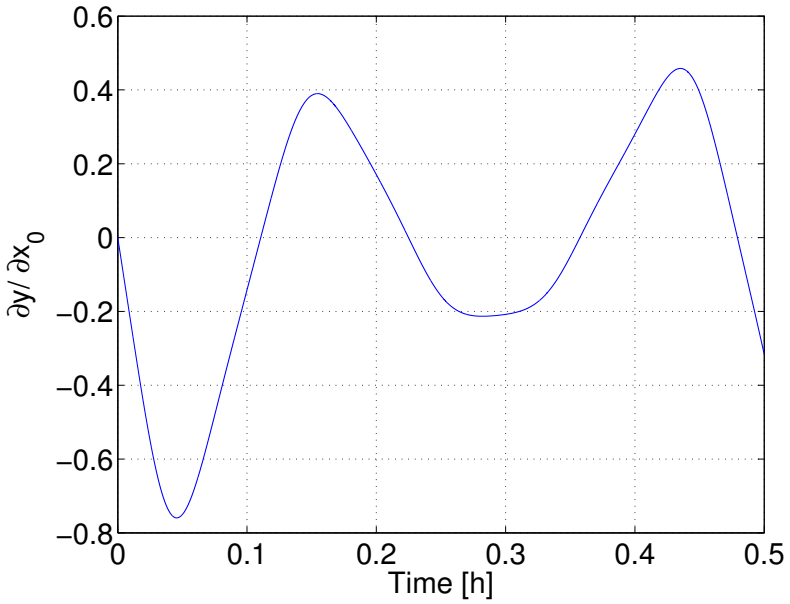


Figure 5.3: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. $x(0)$

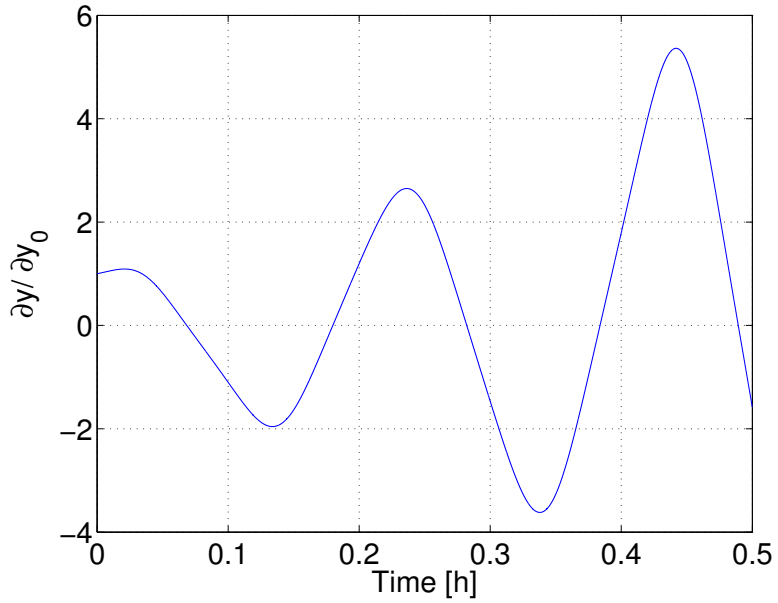


Figure 5.4: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. $y(0)$

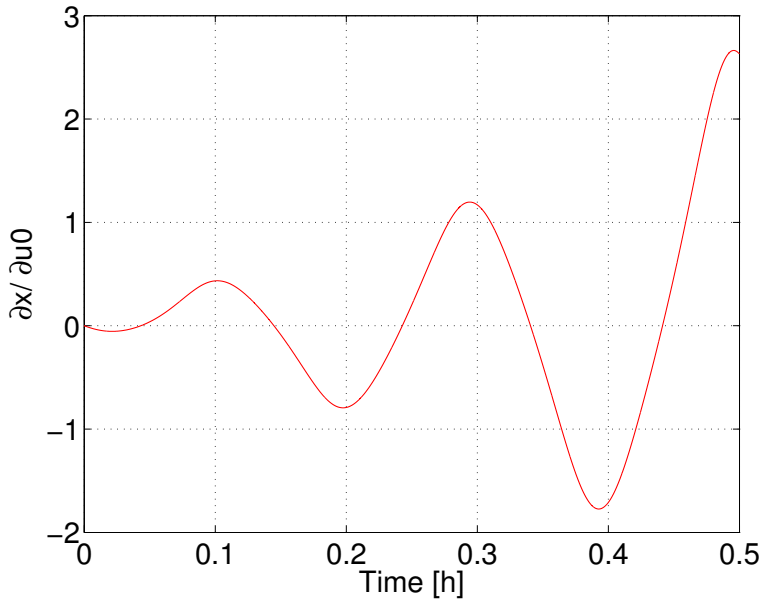


Figure 5.5: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. u_0

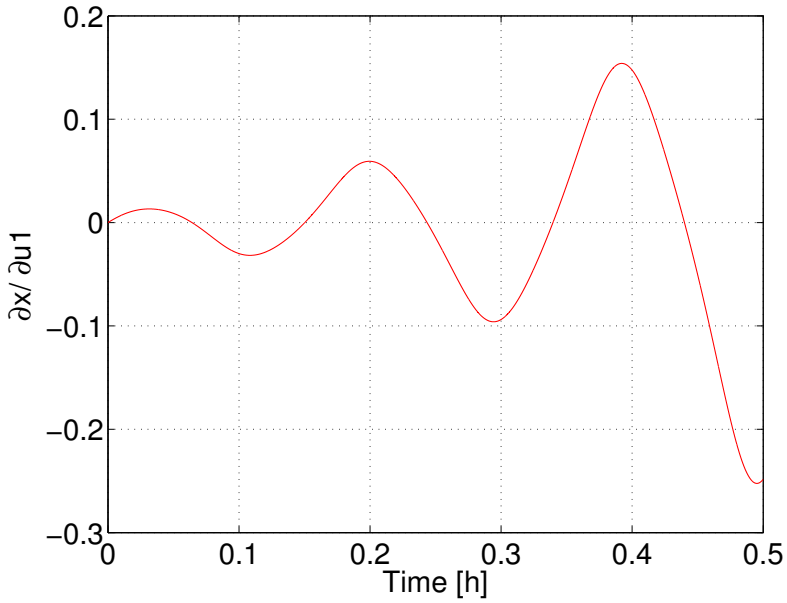


Figure 5.6: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. $u_1(0)$

We can make several general comments about the sensitivity functions for this particular experiment. When it comes to initial conditions, x depends initially on x_0 as seen in Figure 5.1, but not on y_0 as depicted in Figure 5.2. However, by the

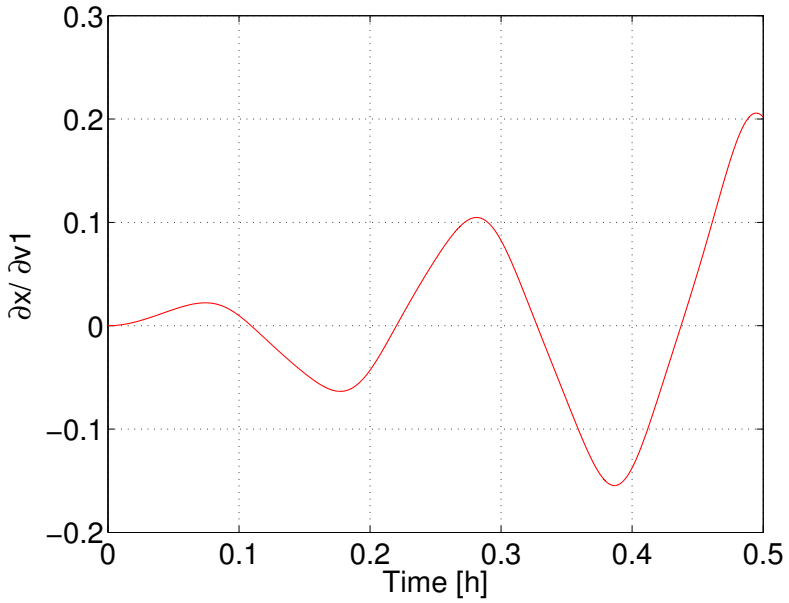


Figure 5.7: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. $v_1(0)$

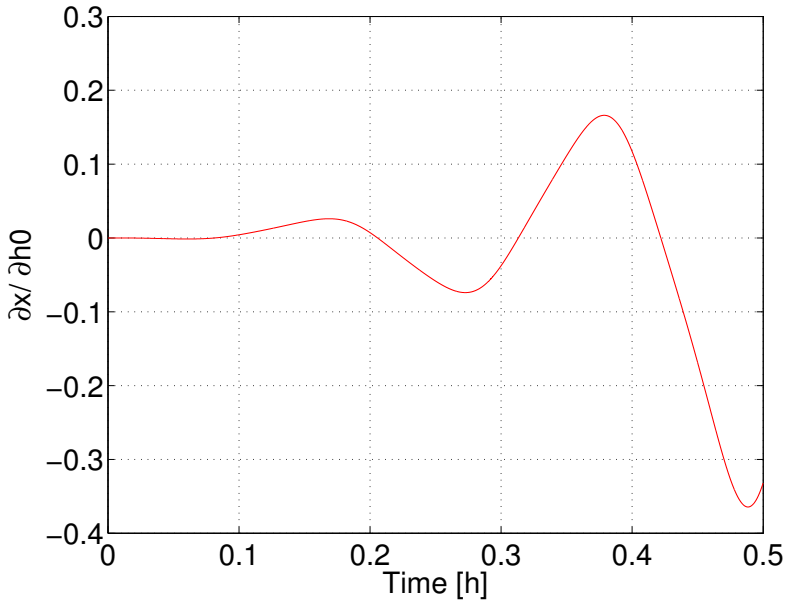


Figure 5.8: Experiment 5.1: Sensitivity of $x(t)$ w.r.t. $h_1(0)$

end of our simulation, x depends much more on y_0 than on x_0. This dependence is similar in the middle part of the run, when it initially decreases, then increase and decreases again before reaching high values at the end of the run. We have a

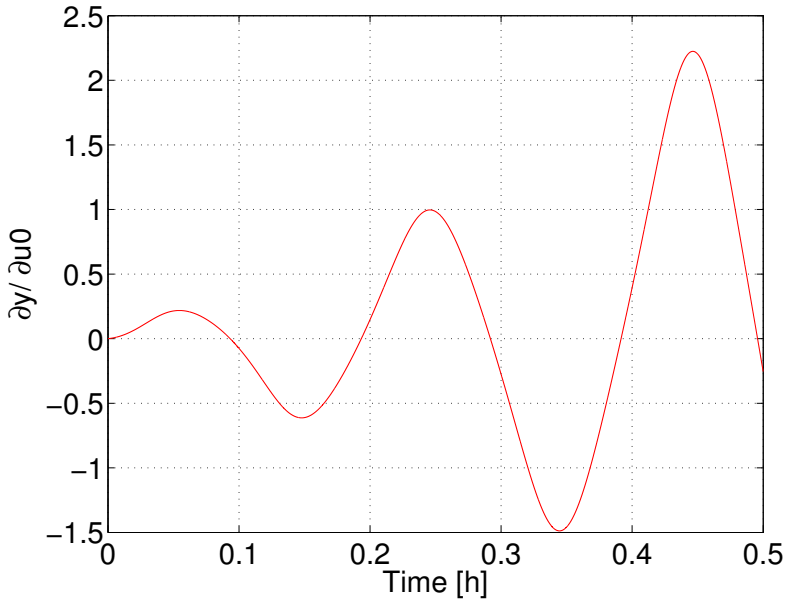


Figure 5.9: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. u_0

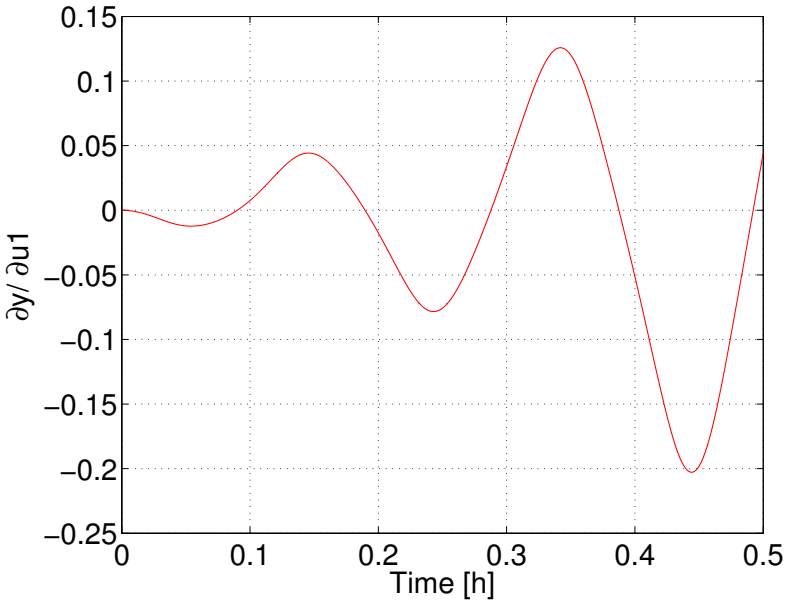


Figure 5.10: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. $u_1(0)$

reverse situation with values of y when it comes to initial conditions as depicted in Figures 5.3 and 5.4.

Sensitivity to the control vector elements shares common treats. It starts

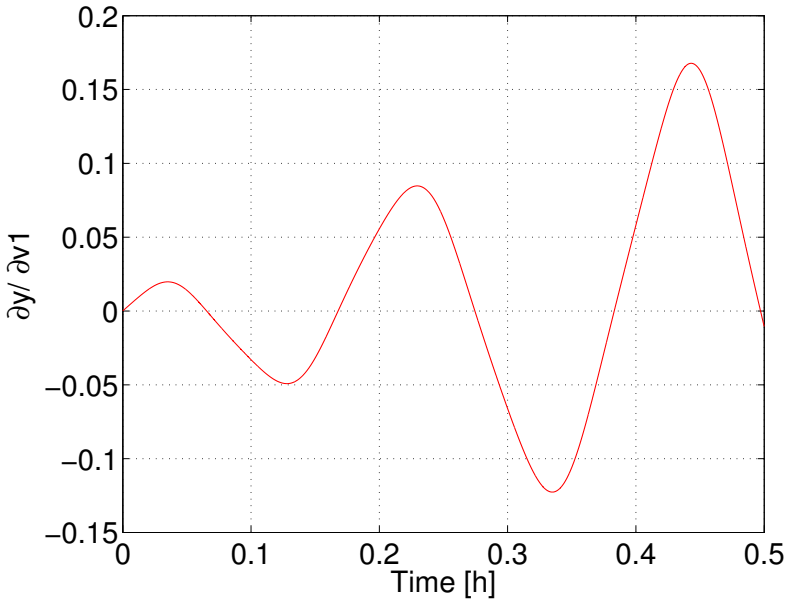


Figure 5.11: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. $v_1(0)$

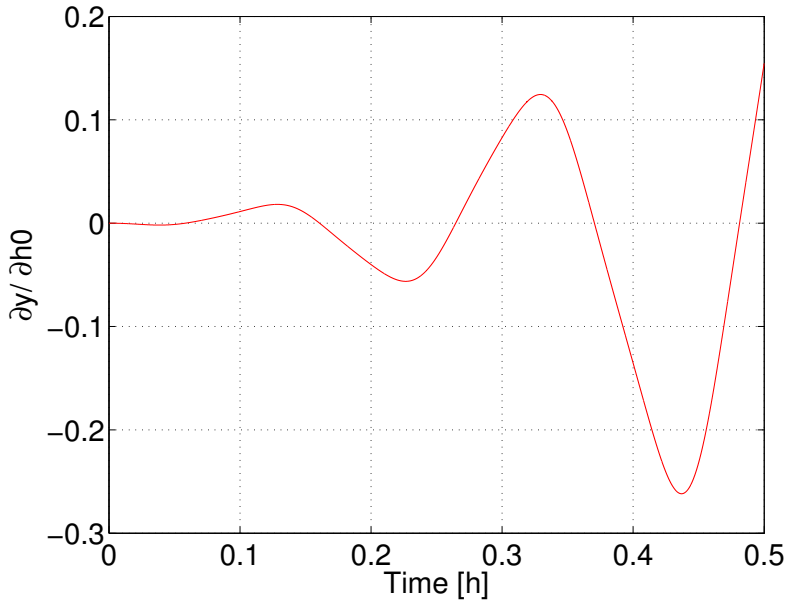


Figure 5.12: Experiment 5.1: Sensitivity of $y(t)$ w.r.t. $h_1(0)$

from nearly zero, and after several zero crossings goes to a maximum value at the end of the simulation at time 0.5. We can see that at the end of the run, sensitivity of x to u_0 seen in Figure 5.5 is much bigger than to $u_1(0)$ as depicted

in Figure 5.6, and they are reverse in sign. There is not much initial sensitivity of x to $h_1(0)$, but it intensifies by the end of the run. Similar pattern of sensitivity characterizes y.

5.4.2 Experiment 5.2

In this experiment, we are running a combined mode, with control vector and initial conditions given in Table 5.2. We have kept u_0, $u_1(0)$, $v_1(0)$ and $h_1(0)$ identical to Experiment 5.1, but we have changed the initial conditions x_0 , y_0.

Table 5.2: Experiment 5.2: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.0	0.0	0.5	2.5e-05

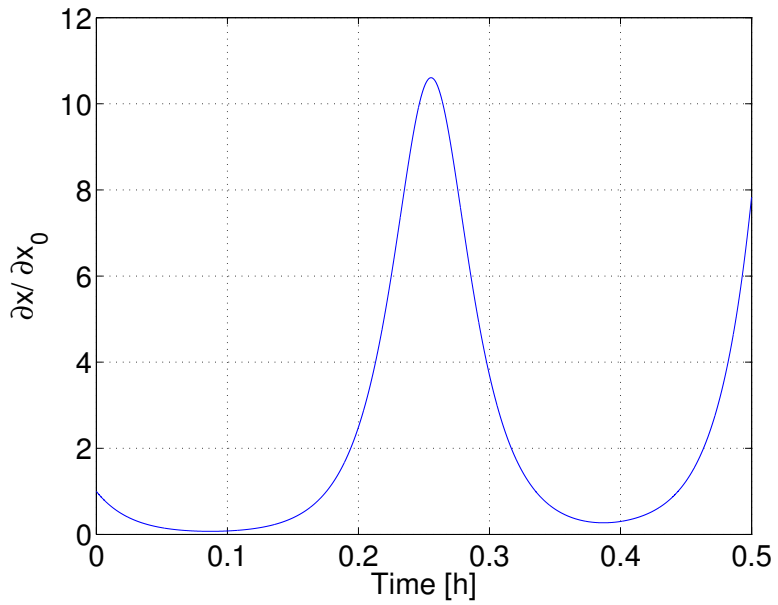


Figure 5.13: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. $x(0)$

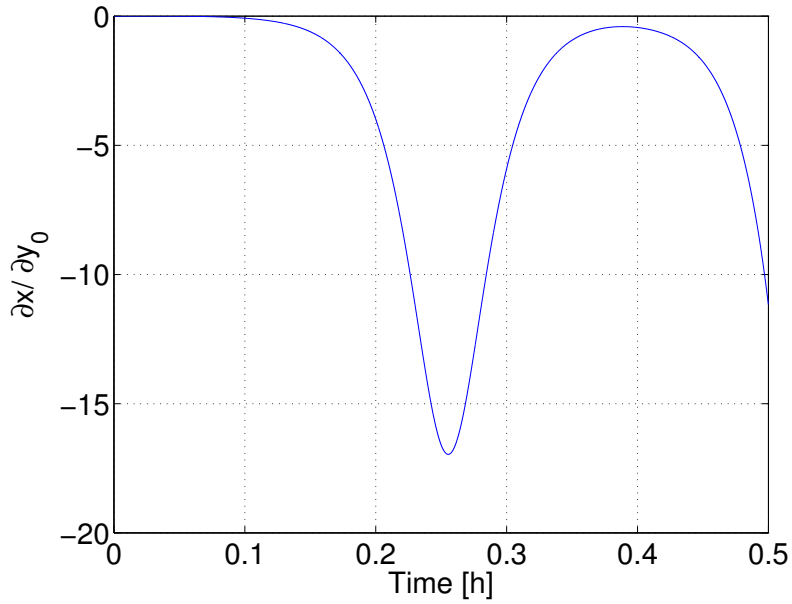


Figure 5.14: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. $y(0)$

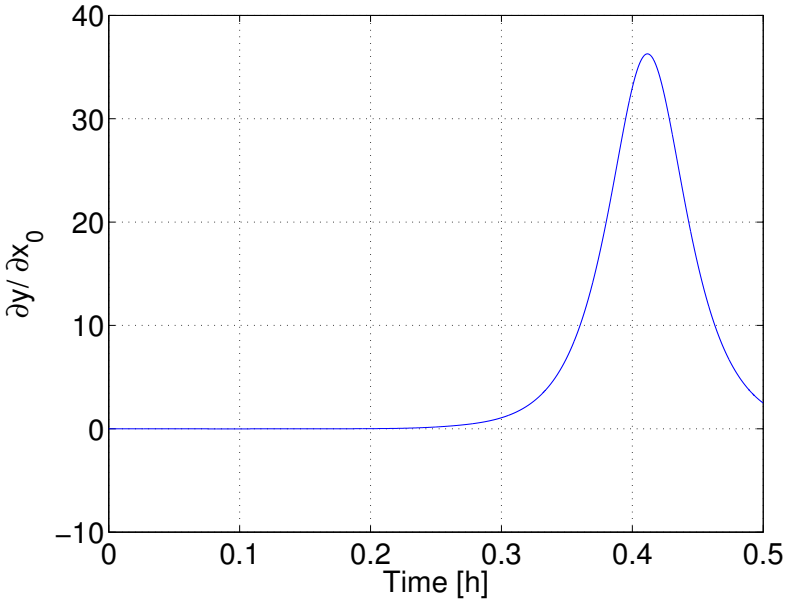


Figure 5.15: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. $x(0)$

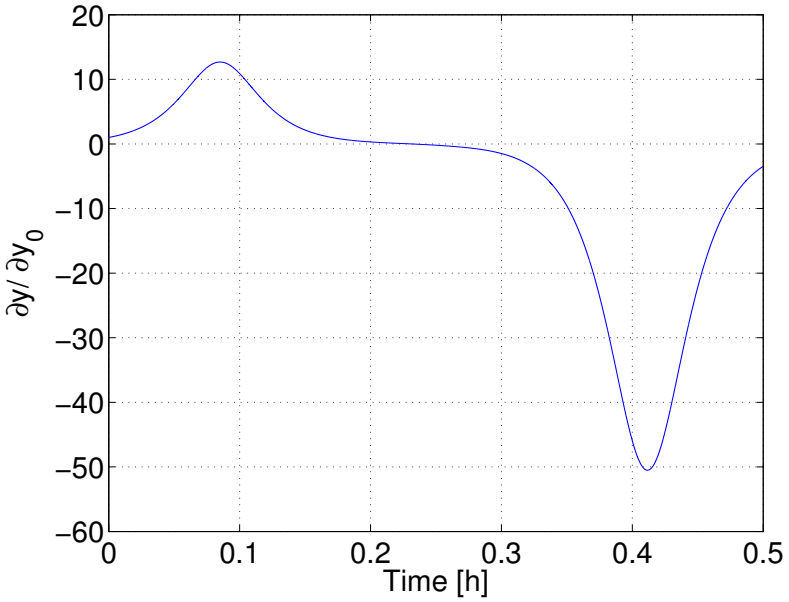


Figure 5.16: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. $y(0)$

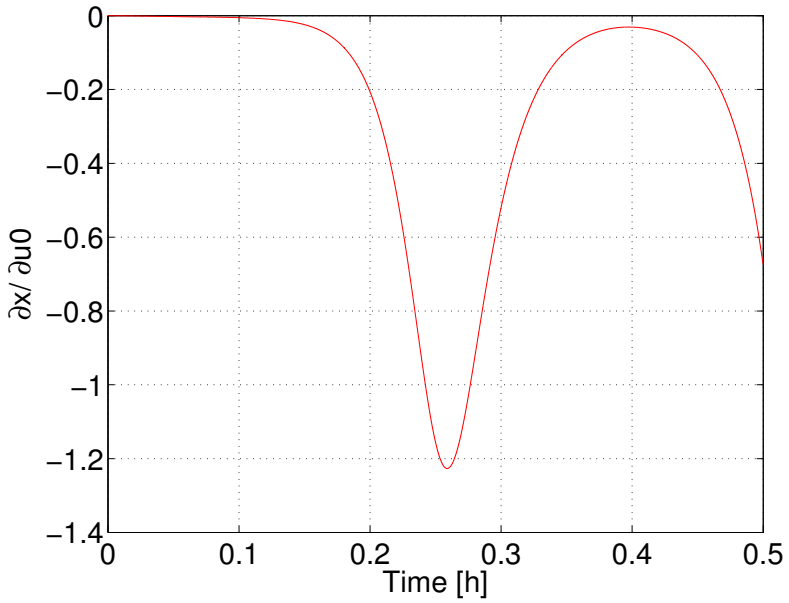


Figure 5.17: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. u_0

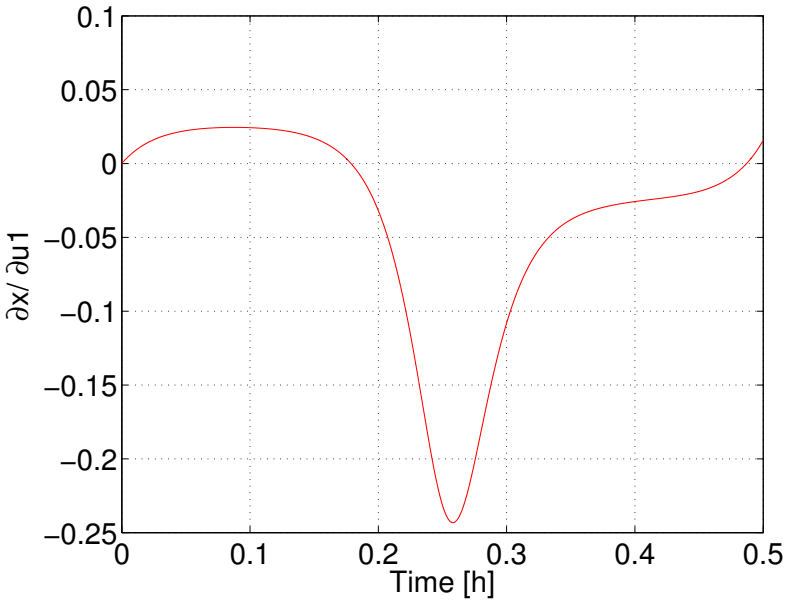


Figure 5.18: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. $u_1(0)$

We can notice that the character of sensitivity functions have changed a great deal when compared to Experiment 5.1. There is a small initial dependence of $x(t)$ on initial condition x_0 that grows and stay generally large for a duration

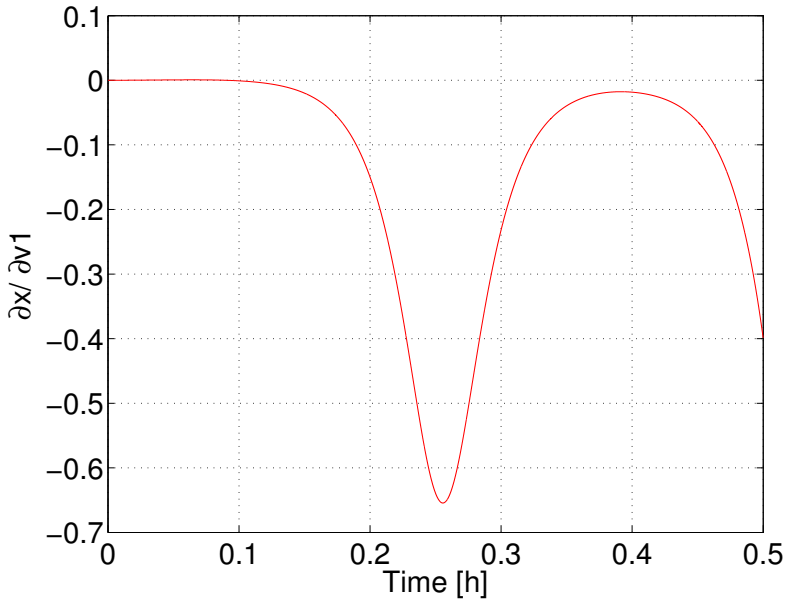


Figure 5.19: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. $v_1(0)$

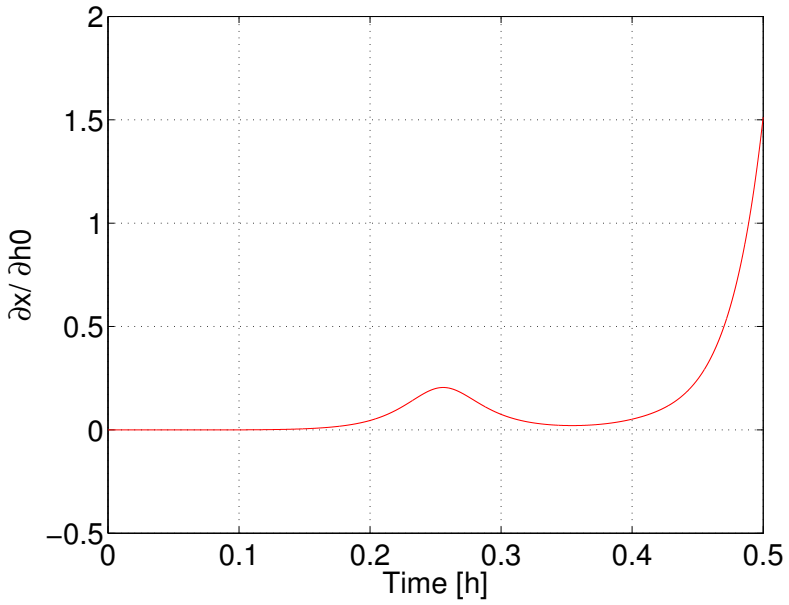


Figure 5.20: Experiment 5.2: Sensitivity of $x(t)$ w.r.t. $h_1(0)$

of the experiment, as seen in Figure 5.13. By looking at Figure 5.14, we can state that this is not the case for dependence of $x(t)$ on y_0 that starts from zero and grows, but is much smaller than dependence on x_0. We can notice that

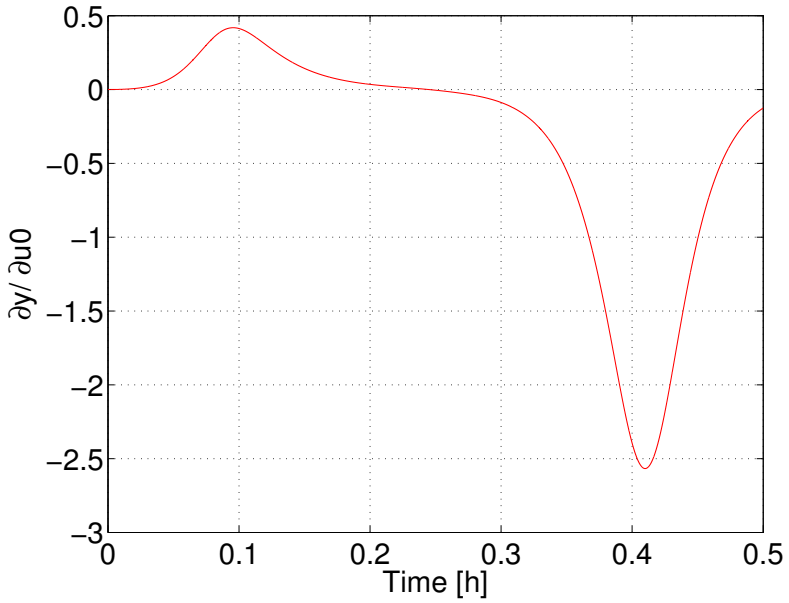


Figure 5.21: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. u_0

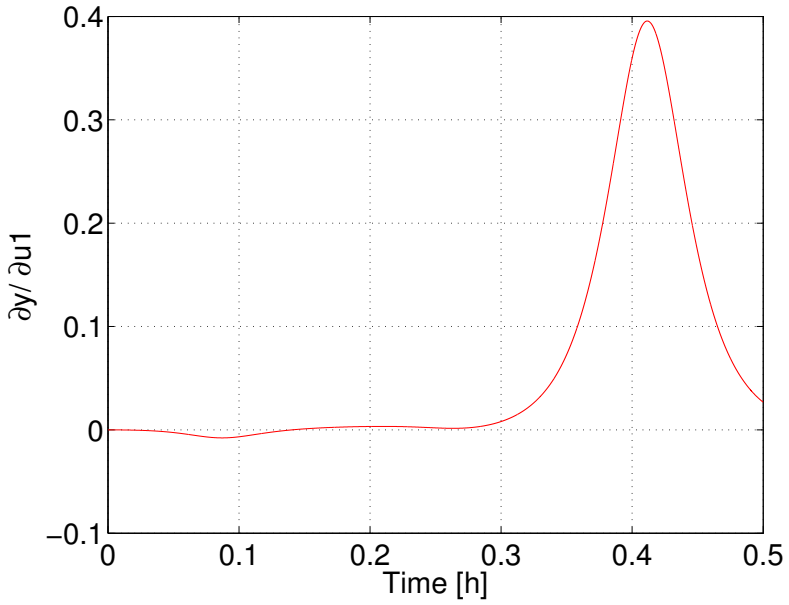


Figure 5.22: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. $u_1(0)$

$y(t)$ depends on initial conditions in a different way. For almost one third of the simulation, $y(t)$ does not depends on x_0, but depends a great deal on y_0, as seen in Figures 5.15 and 5.16. This dependence subsides when it comes to $y(0)$, since

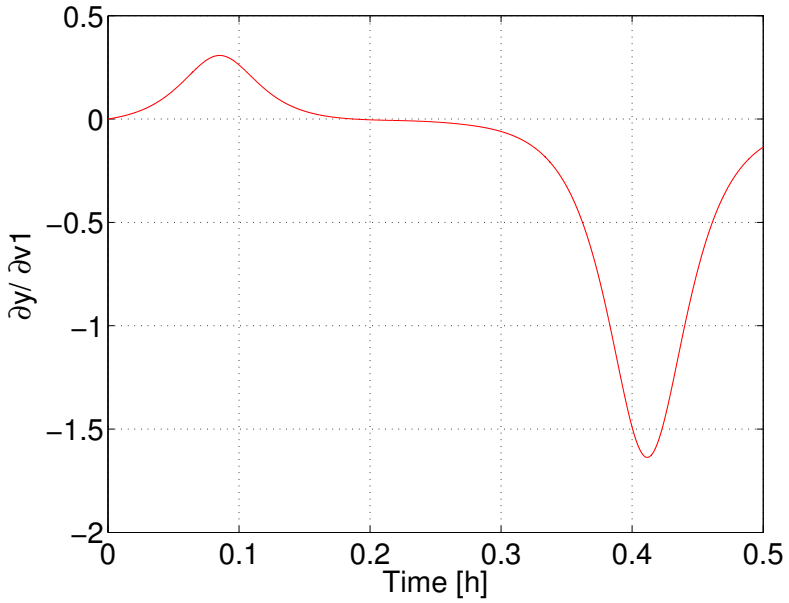


Figure 5.23: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. $v_1(0)$

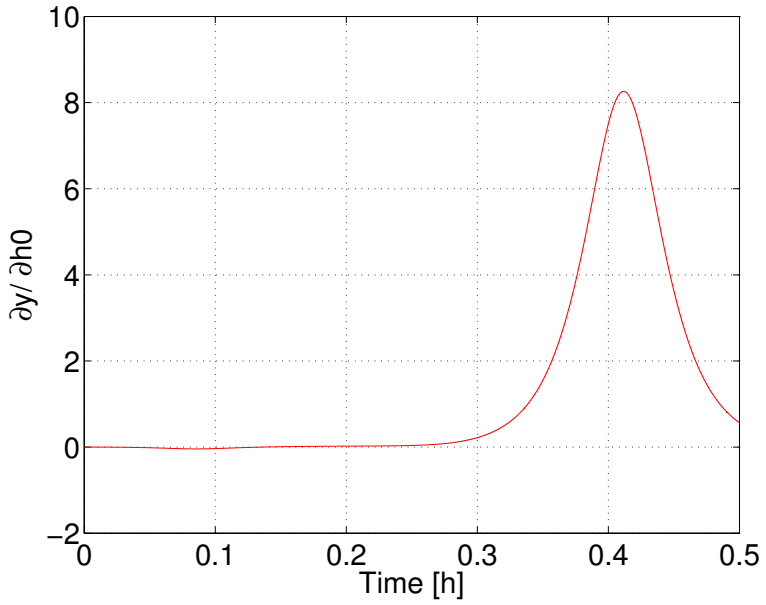


Figure 5.24: Experiment 5.2: Sensitivity of $y(t)$ w.r.t. $h_1(0)$

it falls to nearly zero levels after one fifth of the simulation time.

We can notice in Figures 5.17 and 5.18 that $u(t)$ does not depend initially on u_0 or $u_1(0)$. This dependence grows with the simulation time, and is inversely

dependndent on u_0, and directly dependent on $u_1(t)$. We can also notice that for this simulation, $x(t)$ is sensitive more to $u_1(0)$ than to $v_1(0)$.

Judging from Figure 5.22, $y(t)$ is sensitive to $u_1(0)$ more than to $v_1(0)$ that is depicted in Figure 5.23.

5.4.3 Experiment 5.3

In this experiment, we are running a combined mode, with control vector and initial conditions given in Table 5.3. While the control vector α is the same as the two previous experiments, initial conditions are moved to the point $(0.1, 0.1)$.

Table 5.3: Experiment 5.3: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.1	0.1	0.5	2.5e-05

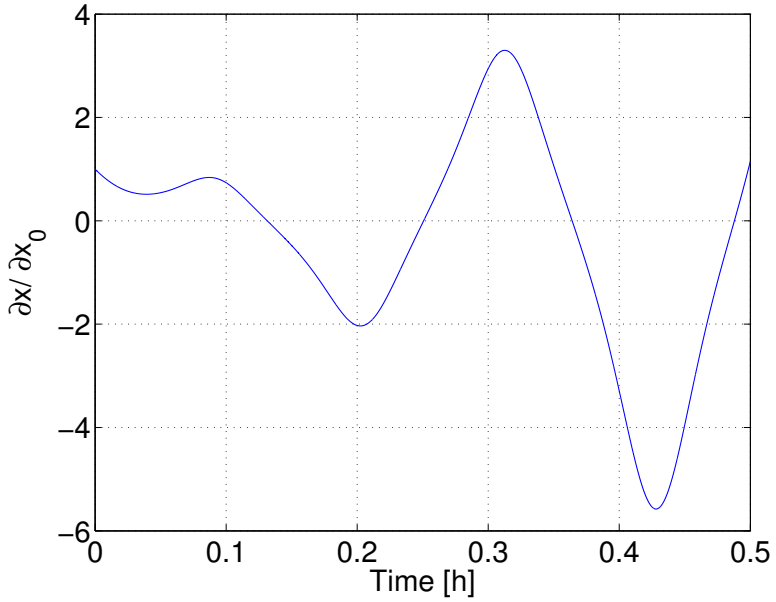


Figure 5.25: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. $x(0)$

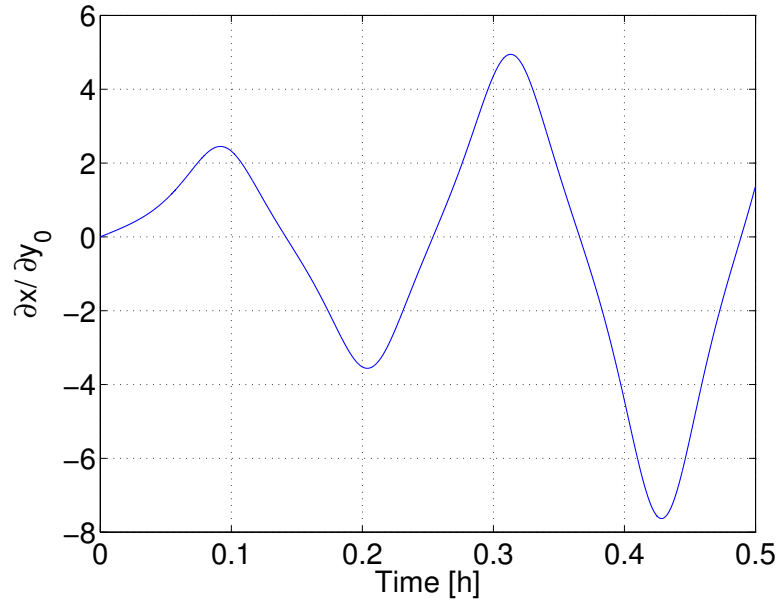


Figure 5.26: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. $y(0)$

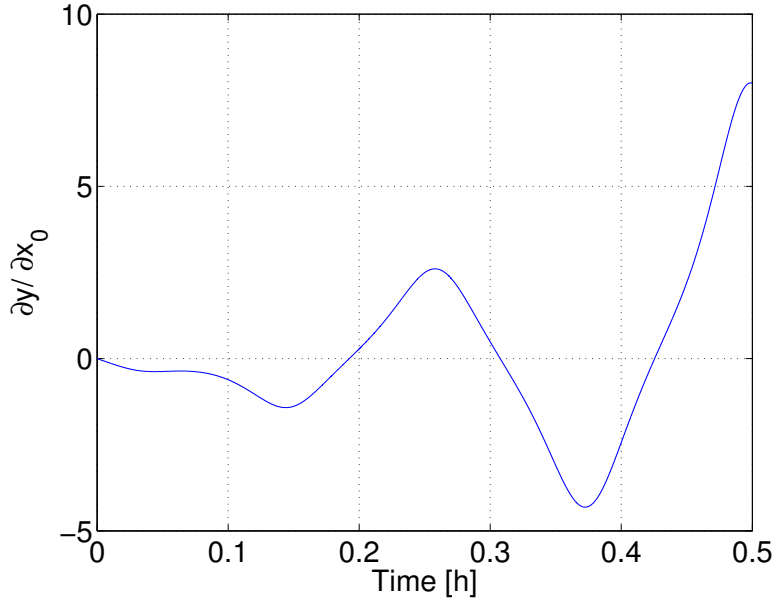


Figure 5.27: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. $x(0)$

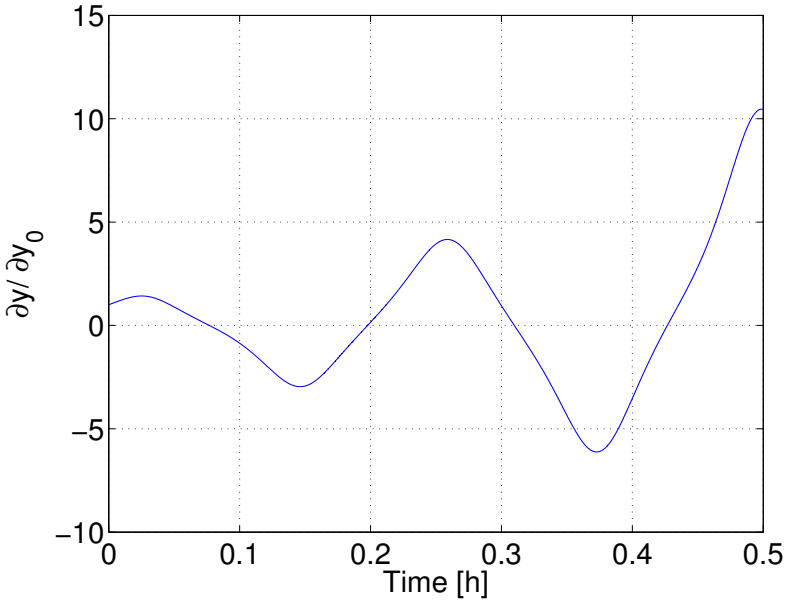


Figure 5.28: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. $y(0)$

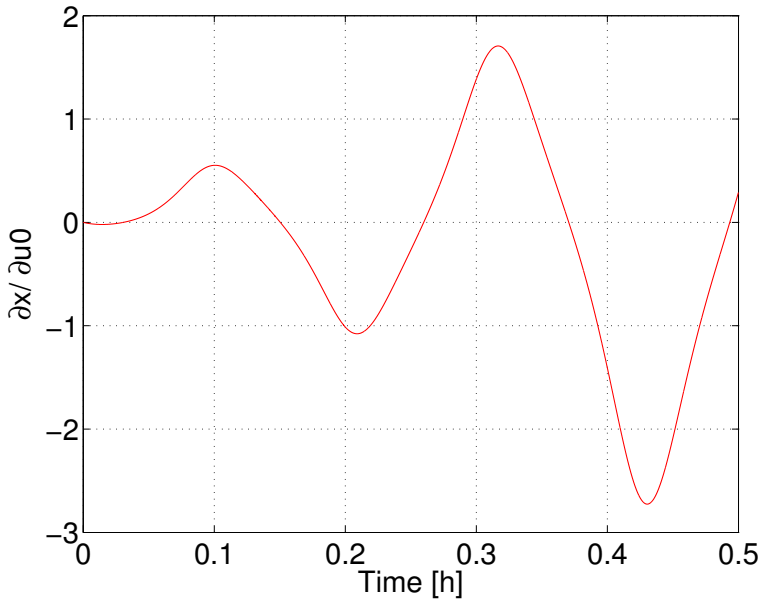


Figure 5.29: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. u_0

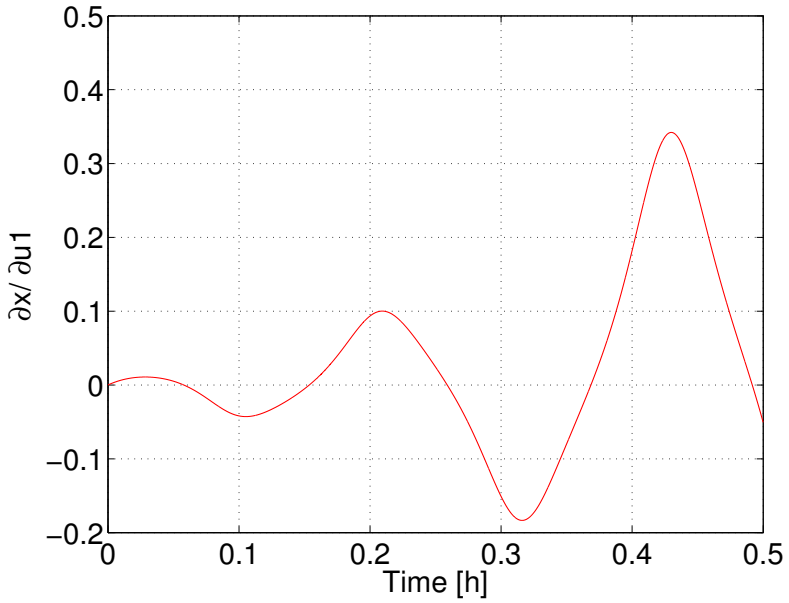


Figure 5.30: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. $u_1(0)$

We can notice in Figure 5.25 that $x(t)$ sensitivity starts around one, then goes through a zero crossing point at time 0.13 to reach value -2, then it grows to about 3.5 to have a zero crossing again at time 0.36 do drop to -5.5. While

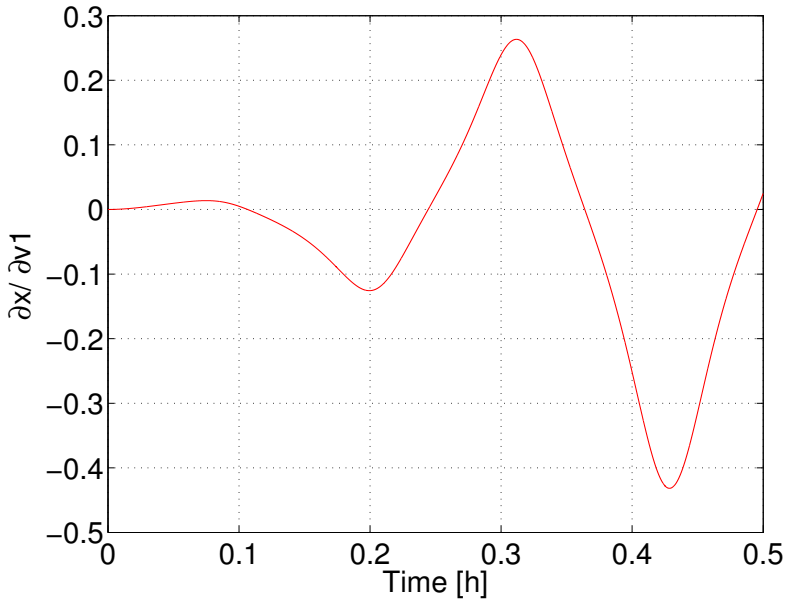


Figure 5.31: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. $v_1(0)$

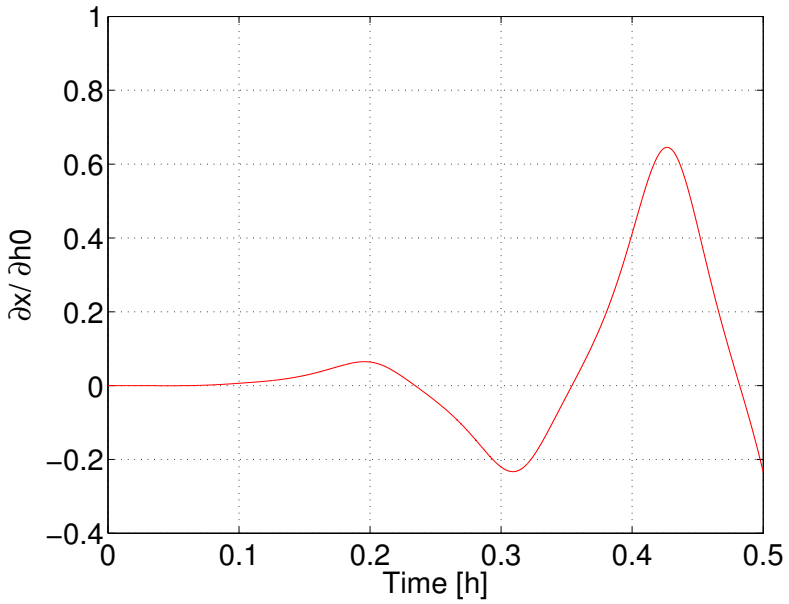


Figure 5.32: Experiment 5.3: Sensitivity of $x(t)$ w.r.t. $h_1(0)$

behavior is similar to this in Figure 5.1, values of sensitivity are bigger than in the Experiment 5.1. Sensitivity of $x(t)$ w.r.t. y_0 start from zero, but it reaches values bigger than in the Experiment 5.1 as well. We can see in Figure 5.27 that $y(t)$ is

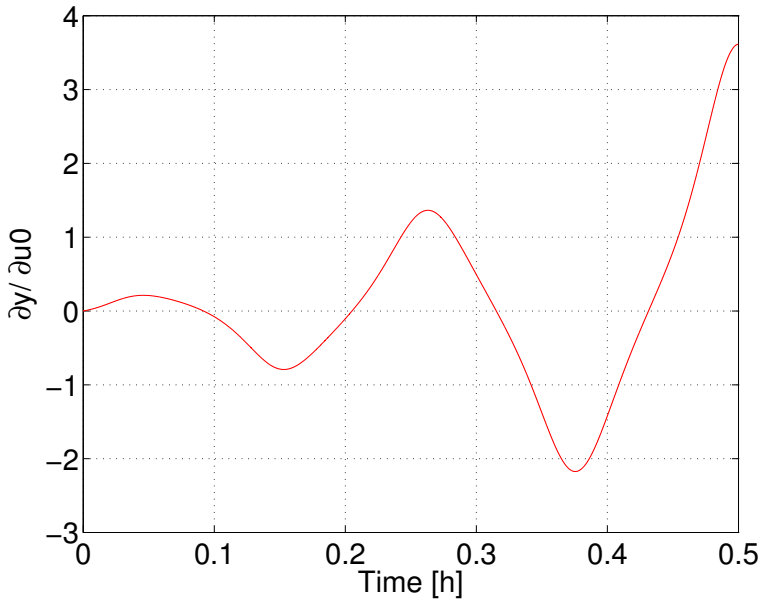


Figure 5.33: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. u_0

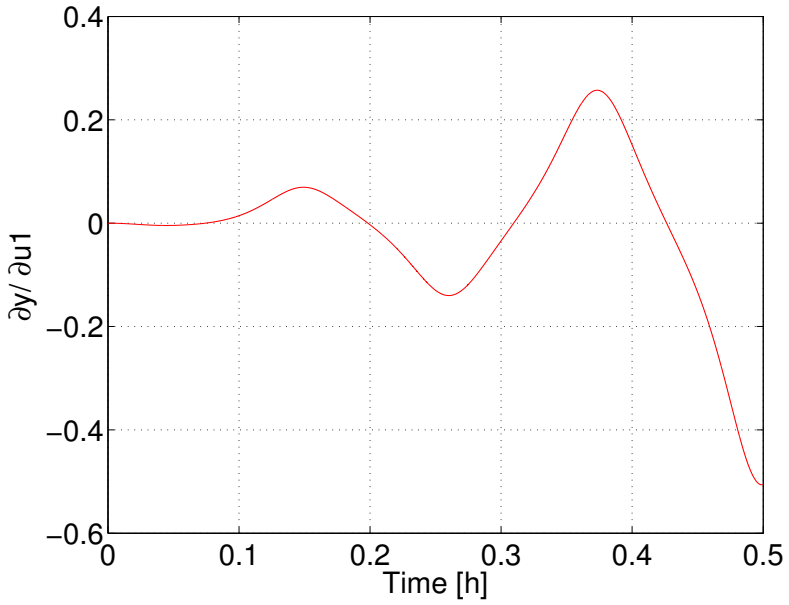


Figure 5.34: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. $u_1(0)$

initially not sensitive to value of $x(0)$, but it reaches value in the order of 7.5 by the end of the simulation. Its graph has three zeros crossings, and behaves quite differently than the one from Figure 5.3 in Experiment 5.1. When we compare

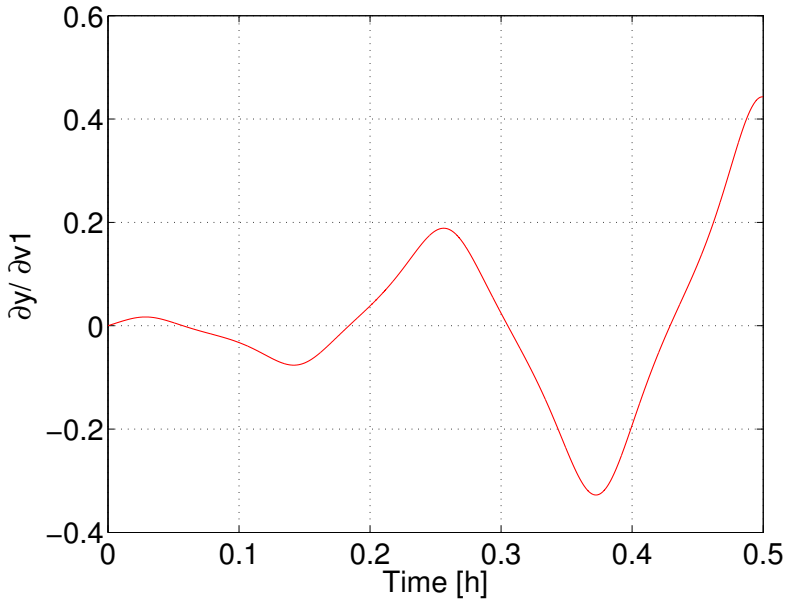


Figure 5.35: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. $v_1(0)$

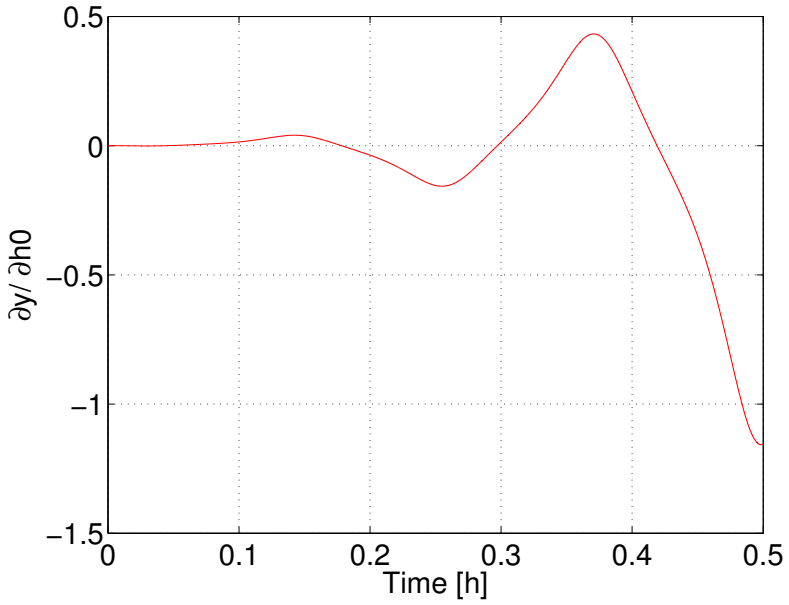


Figure 5.36: Experiment 5.3: Sensitivity of $y(t)$ w.r.t. $h_1(0)$

sensitivity $y(t)$ from Figure 5.28 with the one from the Experiment 5.1 Figure 5.4, we can notice than they have similar shape but $y(t)$ sensitivity values in the Experiment 5.3 are almost twice as big.

5.5 Summary

In this chapter, we have analyzed model sensitivity, that is the dependence of the linearized shallow water model on initial conditions and parameters. We have shown that the shallow water model can exhibit different behavior that depends on the initial conditions and the control vector. This gives us a great insight into the fine structure of the tracer dynamics. By analyzing sensitivity plots, we can answer in the quantitative way, which parameters control tracer position for a given prediction to a bigger degree. We can also see in included plots that for the linearized shallow water model, there is no identical dependence on parameters and initial conditions; they have to be analyzed case by case. This determination of parametric sensitivity is very important for the successful data assimilation. This is studied in the next chapter.

Chapter 6

Numerical experiments in Data Assimilation

6.1 Introduction

In the previous chapter, we have conducted numerical experiments in which we have analyzed sensitivity of tracer dynamics to the initial conditions $\mathbf{X}(0)$ and elements of the control vector $\boldsymbol{\alpha}$. Our findings would suggest that there is a reason for different influence of measurements taken with in a temporal domain.

In this chapter, we demonstrate the effectiveness of the forward sensitivity method (FSM) by conducting a series of numerical experiments conducting DA. We differ the number of observations, their temporal distribution and the observational error variance. We need to bring some definitions we can use to compare different data assimilation experiments.

6.1.1 Root-mean-square error

We use the *root-mean-square* (RMSE) error to estimate the goodness of our data assimilation scheme, that is, to evaluate a difference between the values that are observed and values that are estimated by the model. It represents a very valuable measure of forecast errors and it allows for comparison between forecast and different number of observations and their placement. Predicted values $\hat{y}_t$ at n different times are compared with the observed values

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \tag{6.1}$$

6.1.2 Condition number of a matrix

Let us bring up a definition of the *condition number* of a matrix, usually indicated as κ, from (page 667, Lewis et al. (2006)) [12]. For matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$, where $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$,

$$\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|. \tag{6.2}$$

The condition number of a matrix describes how the sensitivity of the solution of a system of linear equations depends on error in the data. It can be said that matrices with small condition number are well conditioned; singular matrices have infinite condition numbers. We will use a *spectral condition number* defined on (page 668, Lewis et al. (2006))

$$\kappa_2(A) = \|\mathbf{A}\|_2\|\mathbf{A}^{-1}\|_2 = \frac{\max_i |\lambda_i|}{\min_i |\lambda_i|}. \tag{6.3}$$

6.1.3 Methodology

The observations that are used for data assimilation experiments are created using "Base" trajectory by adding the noise with the observational error variance σ of different magnitudes given in Table 6.1. In every experiment, control vectors used for "Base" trajectory are shown alongside the perturbed control vectors that we try to improve. We start each numerical experiment with the incorrect control vector. The forward sensitivity method uses sensitivity functions that use that erroneous solution. We use FSM to assimilate a different number of observations, and since we have four elements of the control vector, for number of observations bigger than four, we have an *overdetermined system*. Our numerical experiments use observations from four different temporal ranges described as: START, MIDDLE, FINISH, UNIFORM. They are described in the next section. We keep track of the condition number, the ratio of the largest to the smallest of $\mathbf{H}^T\mathbf{H}$, since the inverse of $\mathbf{H}^T\mathbf{H}$ influences the optimal adjustments to the control vector. We have to look at the observation spacing in some temporal ranges. We can infer some characteristic of the measurement distribution and compare condition numbers.

We use the iterative process correcting elements of the control vector. After first iteration, we apply the correction and repeat the process, that is, we make another forecast using the corrected control vector. We have shown results of the firs three steps of the iteration. Clearly, FSM improves forecast at each step; we track improvements by following RMS error between observations and the forecast.

Table 6.1: Values of the observational error variance σ and σ^2 used in data assimilation experiments to create perturbed observations

	σ	σ^2
1	0.0707	0.0050
2	0.0866	0.0075
3	0.1000	0.0100
4	0.1118	0.0125
5	0.1225	0.0150
6	0.1323	0.0175
7	0.1414	0.0200

6.2 Data assimilation

Each of our data assimilation experiments consist of four separate data assimilations that deal with different distribution of measurements called: START, MIDDLE, FINISH and UNIFORM. They differ by placing measurements as follows: START puts all measurements in the first third of the temporal domain, MIDDLE puts all measurements in the second third or the temporal domain, FINISH puts all measurements in the last third of the temporal domain, and finally UNIFORM distributes measurements over the entire period of modeling.

6.2.1 Experiment 6.1

Experiment base configuration is shown in Table 6.2. These values are used to create observations in a twin-experiment by adding a random noise with different σ^2 values shown in Table 6.1. Then, control vector $\boldsymbol{\alpha}$ values are modified as shown in 6.3. Then, four separate experiments are conducted with different distribution of measurements. We can see a trajectory of the Lagrangian tracer, and data assimilation.

Table 6.2: Experiment 6.1: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.209	0.109	0.5	2.5e-05

Table 6.3: Experiment 6.1: Perturbed control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
0.986523	-0.001833	0.546079	0.136232	0.209	0.109	0.5	2.5e-05

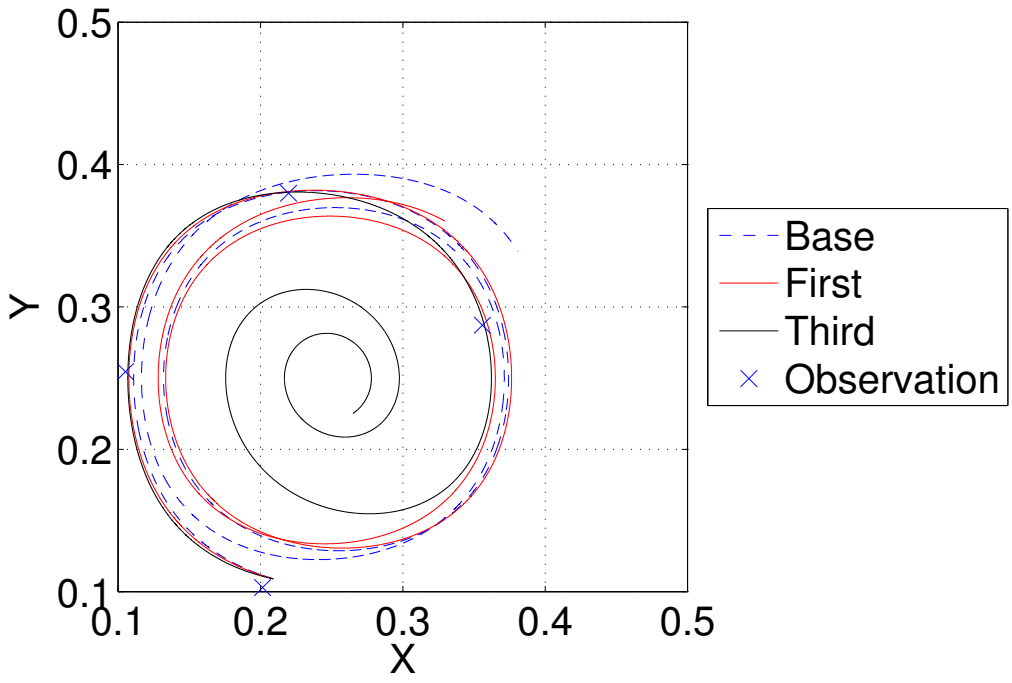


Figure 6.1: Experiment 6.1: Trajectory with 4 measurements, START distribution

In table 6.4, we can see a summary of similar experiments described so far with 4 measurements. They differ by increasing value of σ^2, by which we perturb observations in the twin-experiment. We can see in all experiments except one,

Table 6.4: Experiment 6.1: Comparison of data assimilations for a set of distributions and errors with 4 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0190	0.0056	1.7e+04	1.7e+04
	0.007	0.0579	0.0138	2.5e+04	1.8e+04
	0.010	0.0565	0.0162	2.4e+04	2.0e+04
	0.013	0.0361	0.0122	1.7e+04	1.8e+04
	0.015	0.0418	0.0129	1.7e+04	1.8e+04
	0.018	0.0211	0.0095	1.9e+04	1.7e+04
	0.020	0.0479	0.0288	2.0e+04	3.0e+04
MIDDLE	0.005	0.0395	0.0073	3.5e+04	3.1e+04
	0.007	0.1411	0.0511	5.7e+04	3.0e+04
	0.010	0.1428	0.0600	5.8e+04	3.2e+04
	0.013	0.0675	0.0188	3.8e+04	3.1e+04
	0.015	0.1059	0.0468	3.5e+04	3.2e+04
	0.018	0.0609	0.0177	3.6e+04	3.0e+04
	0.020	0.0429	0.0171	3.3e+04	3.4e+04
FINISH	0.005	0.0574	0.0145	1.1e+05	1.4e+05
	0.007	0.2377	0.1652	1.5e+05	1.1e+05
	0.010	0.2335	0.1798	1.5e+05	1.3e+05
	0.013	0.1174	0.0609	9.2e+04	1.2e+05
	0.015	0.1774	0.1651	8.1e+04	1.0e+05
	0.018	0.1185	0.0568	9.5e+04	1.3e+05
	0.020	0.1067	0.1812	1.1e+05	1.2e+05
UNIFORM	0.005	0.0319	0.0071	2.2e+05	1.3e+05
	0.007	0.1400	0.0592	4.8e+04	8.4e+04
	0.010	0.1260	0.0613	5.2e+04	1.4e+05
	0.013	0.0773	0.1497	2.5e+05	5.1e+04
	0.015	0.1153	0.5134	1.0e+06	6.0e+04
	0.018	0.0628	0.0269	3.9e+04	9.7e+04
	0.020	0.0513	0.0209	4.6e+04	1.7e+05

consecutive iterations improve match between the model and observations. In a run with a uniform distribution, and σ^2 set to 0.0125, RMS error increase and

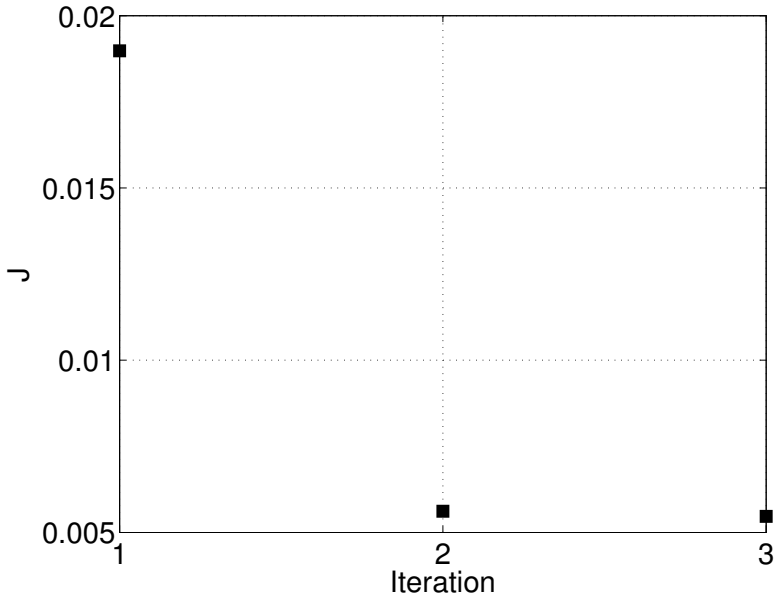


Figure 6.2: Experiment 6.1: Cost function for three steps of data assimilation, START measurement distribution

condition number κ is about three times bigger than in most other experiments.

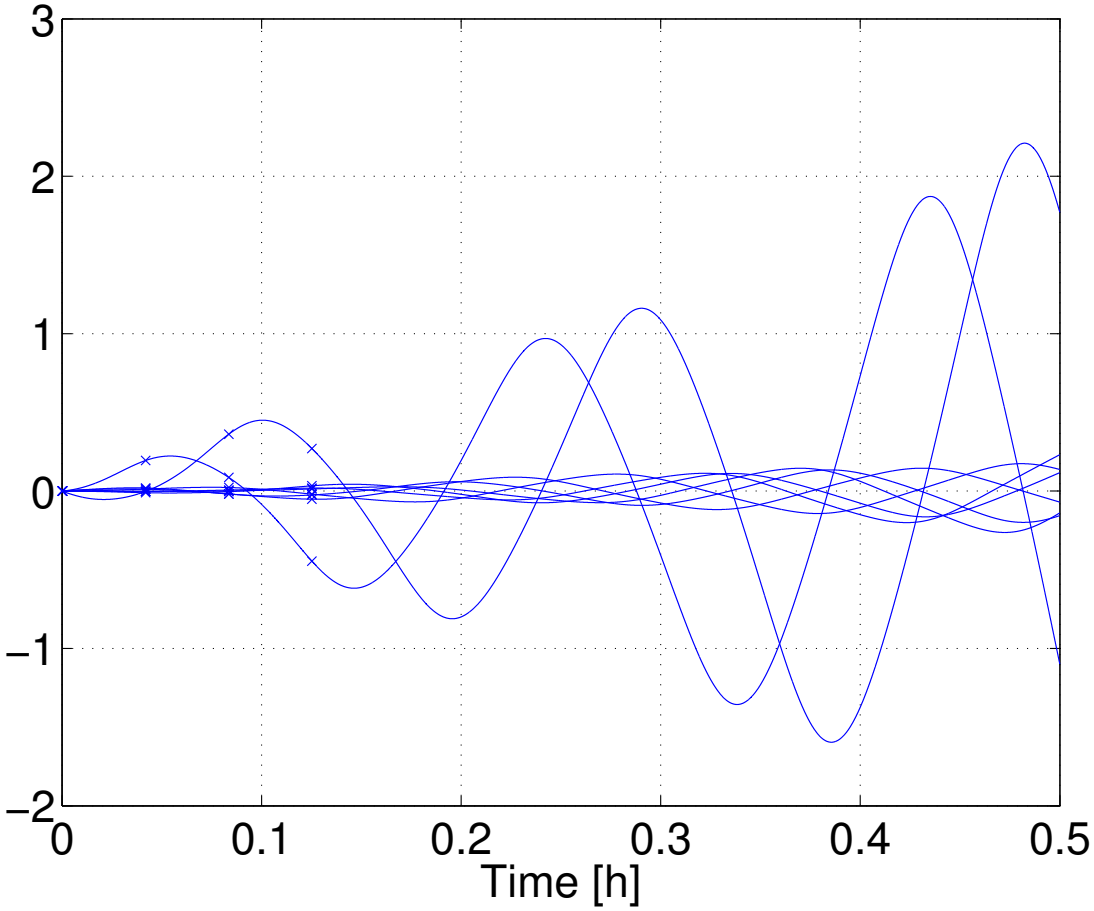


Figure 6.3: Experiment 6.1: Sensitivity functions, START measurement distribution, measurements times indicated along the sensitivity plot by *

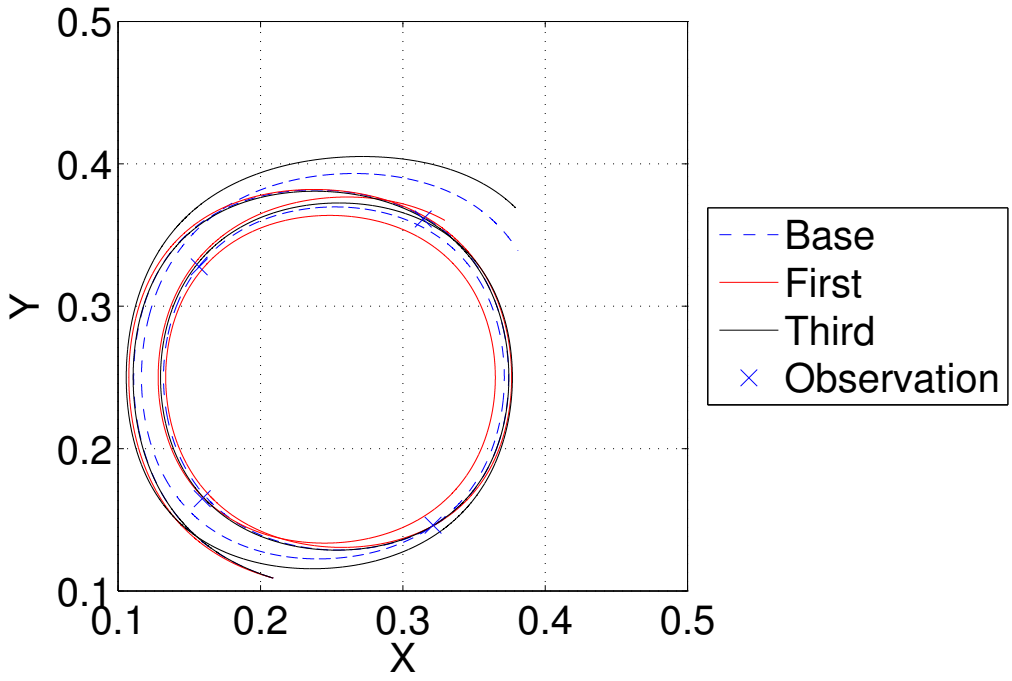


Figure 6.4: Experiment 6.1: Trajectory with 4 measurements, MIDDLE distribution

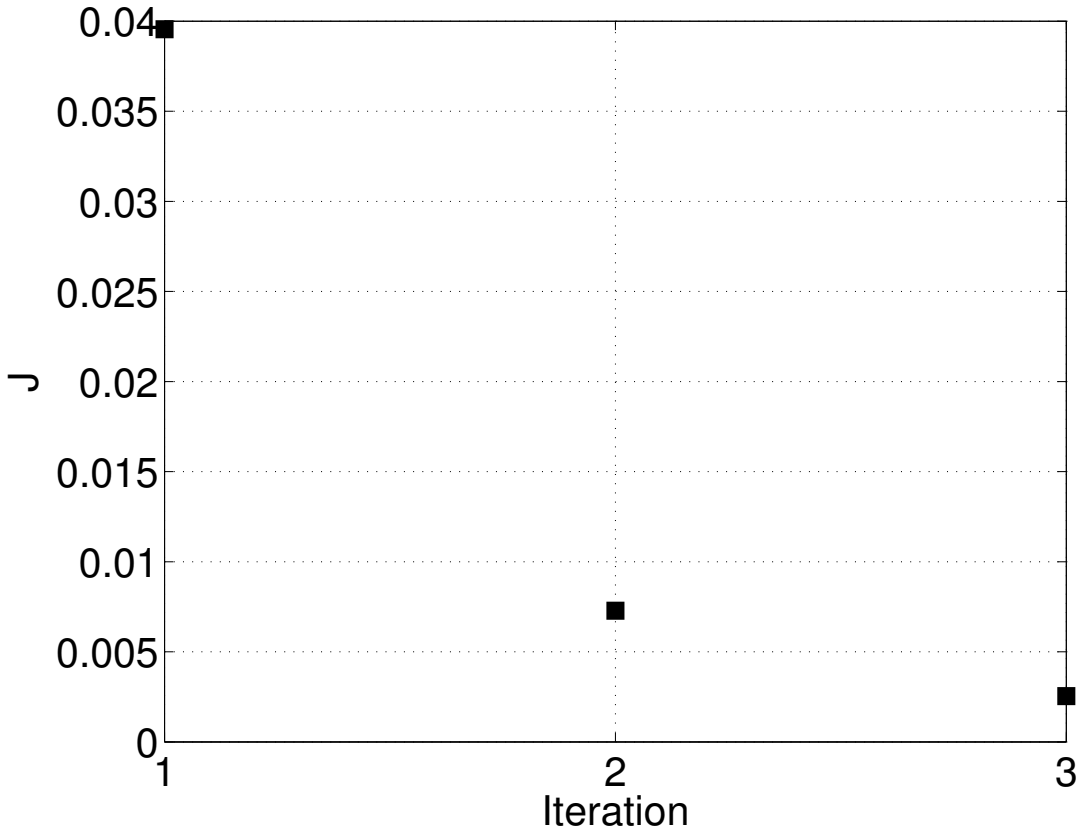


Figure 6.5: Experiment 6.1: Cost function for three steps of data assimilation, MIDDLE measurement distribution

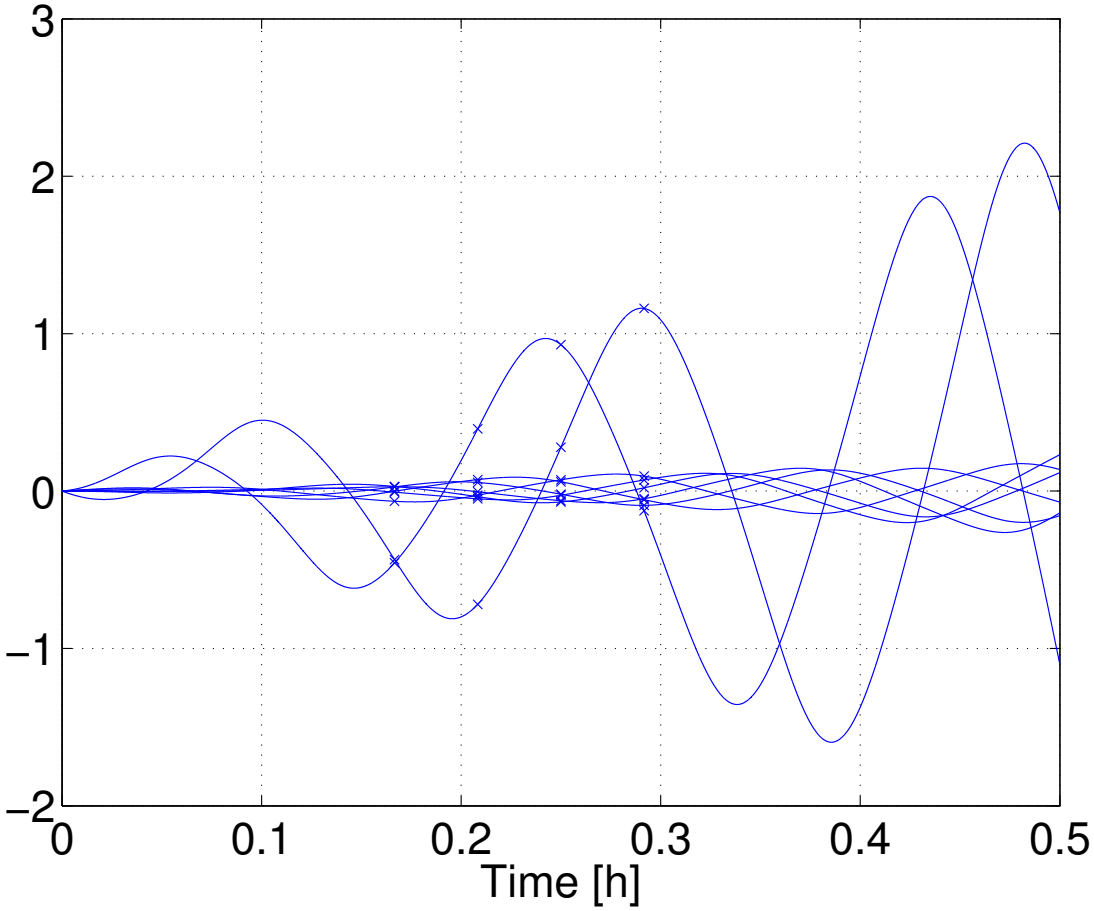


Figure 6.6: Experiment 6.1: Cost function for three steps of data assimilation, MIDDLE measurement distribution, measurements times indicated along the sensitivity plot by *

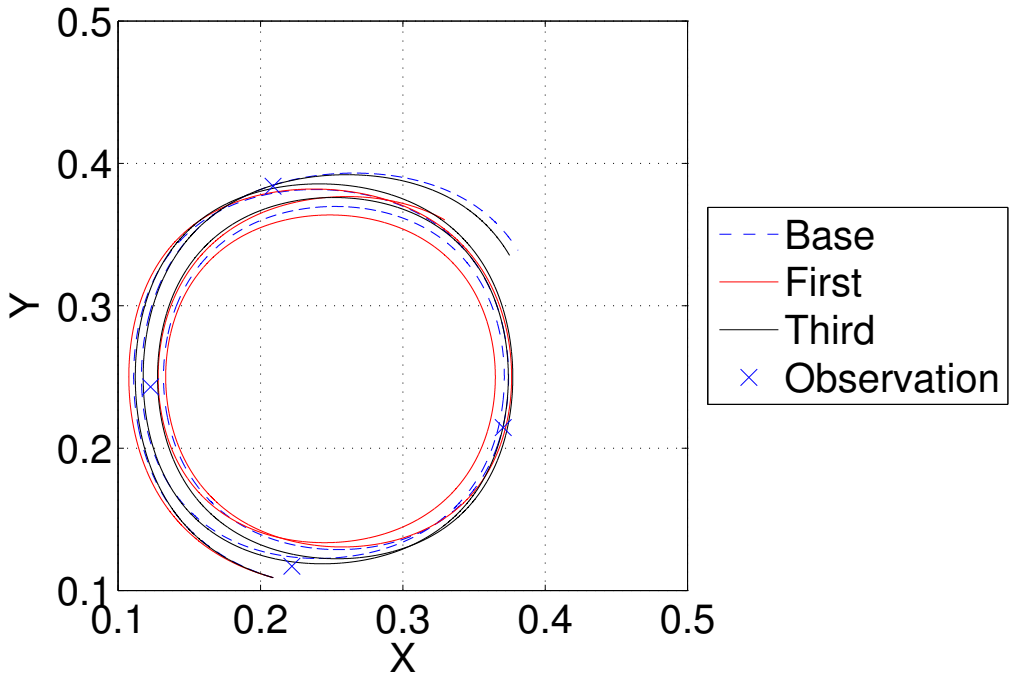


Figure 6.7: Experiment 6.1: Trajectory with 4 measurements, FINISH distribution

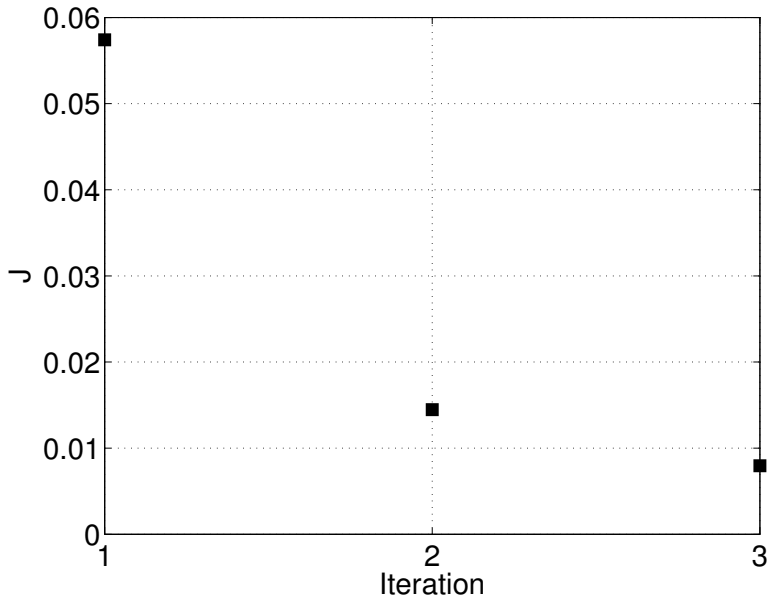


Figure 6.8: Experiment 6.1: Cost function for three steps of data assimilation, FINISH measurement distribution

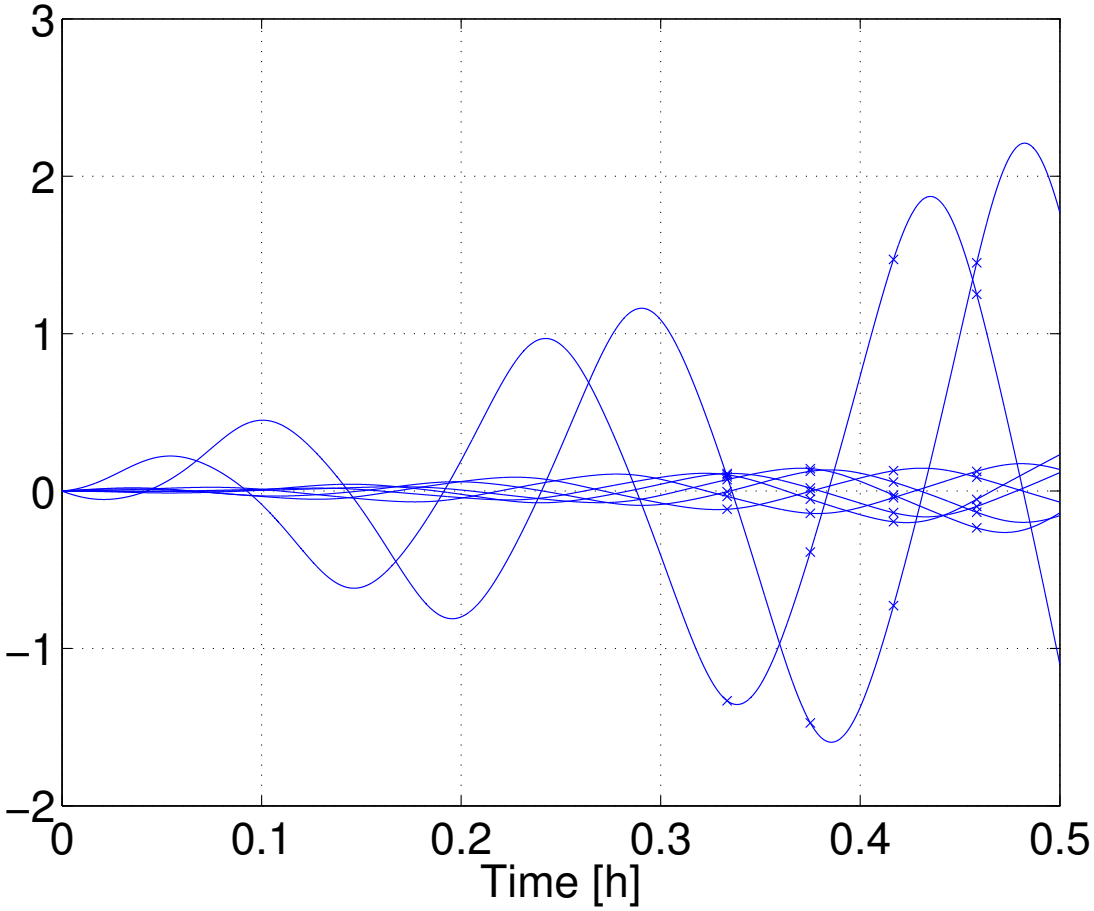


Figure 6.9: Experiment 6.1: Sensitivity functions, FINISH measurement distribution, measurements times indicated along the sensitivity plot by *

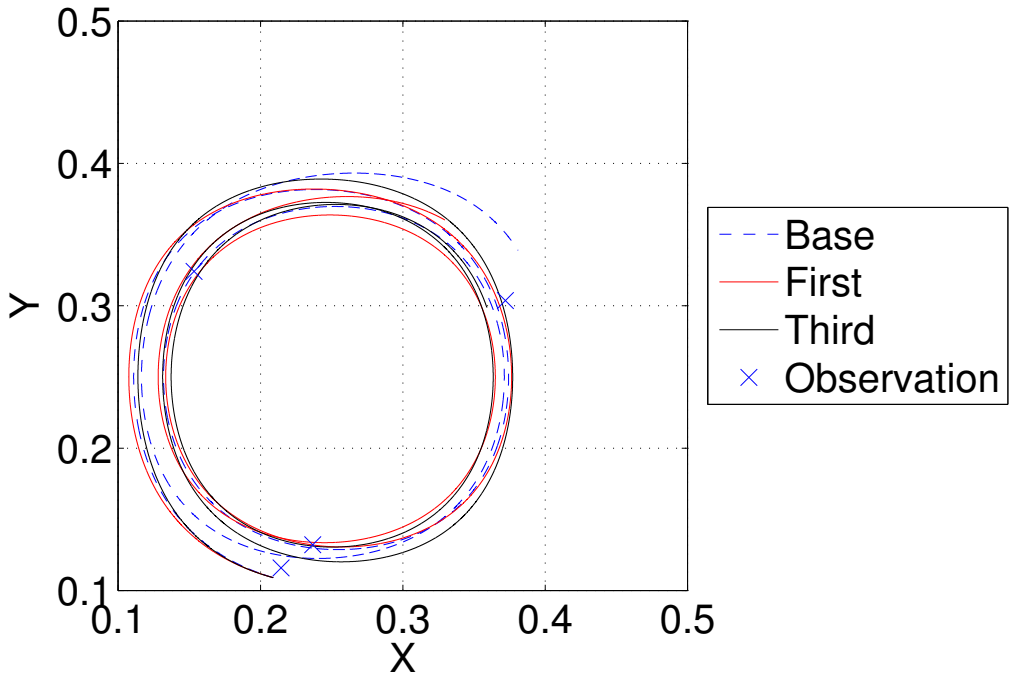


Figure 6.10: Experiment 6.1: Trajectory with 4 measurements, UNIFORM distribution

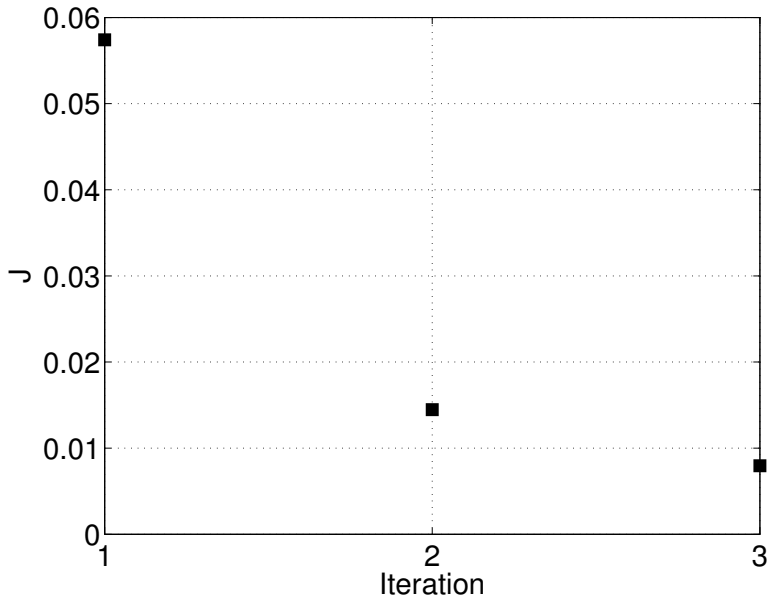


Figure 6.11: Experiment 6.1: Cost function for three steps of data assimilation, UNIFORM measurement distribution

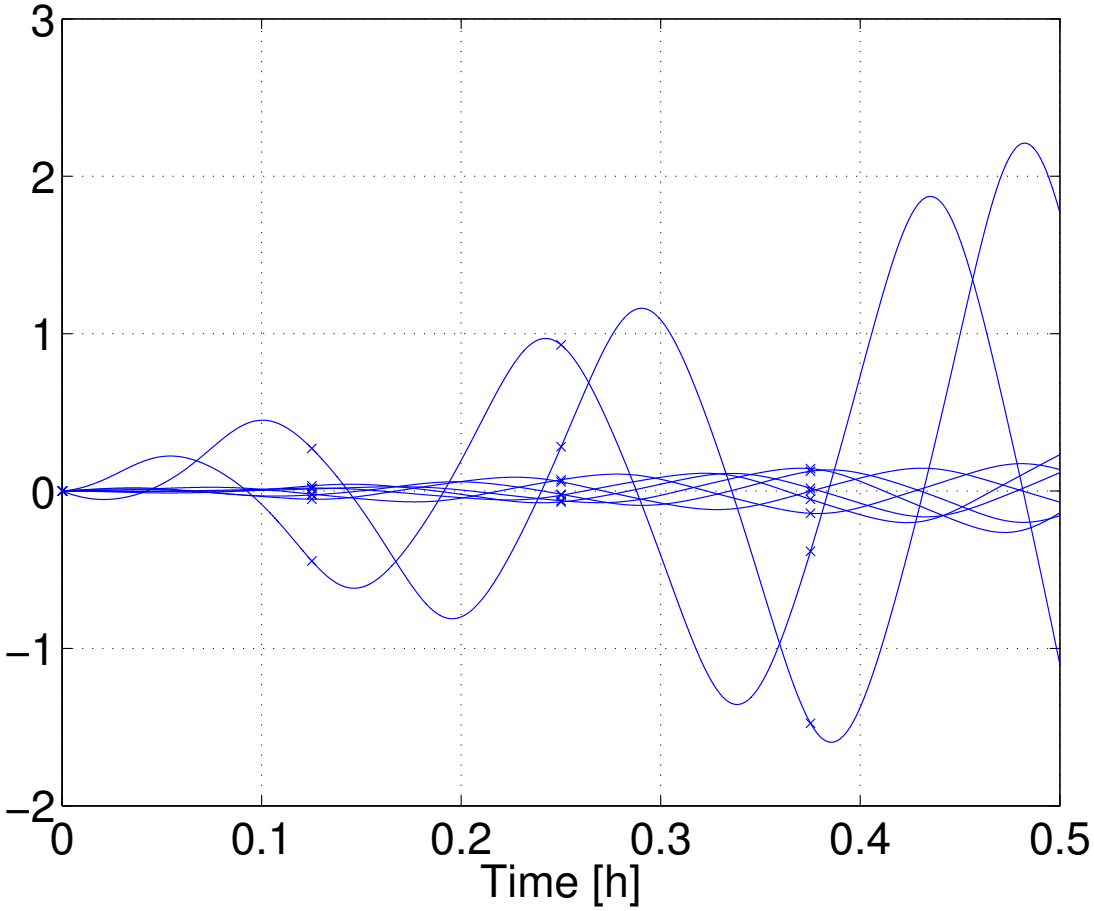


Figure 6.12: Experiment 6.1: Sensitivity functions, UNIFORM measurement distribution, measurements times indicated along the sensitivity plot by *

6.2.2 Experiment 6.2

In this experiment, we are going to continue with the way it was organized in Experiment 6.1, and we are going to investigate whether the increased number of measurements within the set period always lead to better correction of the control vector $\boldsymbol{\alpha}$. We are going to use 4, 8, 16, 32, 64 and 128 measurements within periods described in the Experiment 6.1. Base control vector and perturbed control vector are like the one described in the Experiment 6.1.

Table 6.5: Experiment 6.2: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.209	0.109	0.5	2.5e-05

Table 6.6: Experiment 6.2: Perturbed control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
0.986523	-0.001833	0.546079	0.136232	0.209	0.109	0.5	2.5e-05

Table 6.7: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 4 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0190	0.0056	1.7e+04	1.7e+04
	0.007	0.0579	0.0138	2.5e+04	1.8e+04
	0.010	0.0565	0.0162	2.4e+04	2.0e+04
	0.013	0.0361	0.0122	1.7e+04	1.8e+04
	0.015	0.0418	0.0129	1.7e+04	1.8e+04
	0.018	0.0211	0.0095	1.9e+04	1.7e+04
	0.020	0.0479	0.0288	2.0e+04	3.0e+04
MIDDLE	0.005	0.0395	0.0073	3.5e+04	3.1e+04
	0.007	0.1411	0.0511	5.7e+04	3.0e+04
	0.010	0.1428	0.0600	5.8e+04	3.2e+04
	0.013	0.0675	0.0188	3.8e+04	3.1e+04
	0.015	0.1059	0.0468	3.5e+04	3.2e+04
	0.018	0.0609	0.0177	3.6e+04	3.0e+04
	0.020	0.0429	0.0171	3.3e+04	3.4e+04
FINISH	0.005	0.0574	0.0145	1.1e+05	1.4e+05
	0.007	0.2377	0.1652	1.5e+05	1.1e+05
	0.010	0.2335	0.1798	1.5e+05	1.3e+05
	0.013	0.1174	0.0609	9.2e+04	1.2e+05
	0.015	0.1774	0.1651	8.1e+04	1.0e+05
	0.018	0.1185	0.0568	9.5e+04	1.3e+05
	0.020	0.1067	0.1812	1.1e+05	1.2e+05
UNIFORM	0.005	0.0319	0.0071	2.2e+05	1.3e+05
	0.007	0.1400	0.0592	4.8e+04	8.4e+04
	0.010	0.1260	0.0613	5.2e+04	1.4e+05
	0.013	0.0773	0.1497	2.5e+05	5.1e+04
	0.015	0.1153	0.5134	1.0e+06	6.0e+04
	0.018	0.0628	0.0269	3.9e+04	9.7e+04
	0.020	0.0513	0.0209	4.6e+04	1.7e+05

Table 6.8: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 8 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0621	0.0184	1.2e+04	1.4e+04
	0.007	0.0182	0.0085	1.6e+04	1.4e+04
	0.010	0.0190	0.0131	1.4e+04	1.6e+04
	0.013	0.0309	0.0141	1.8e+04	1.5e+04
	0.015	0.0874	0.0375	2.7e+04	1.4e+04
	0.018	0.0339	0.0220	1.9e+04	1.5e+04
	0.020	0.0456	0.0266	1.8e+04	1.4e+04
MIDDLE	0.005	0.1538	0.0984	2.8e+04	3.2e+04
	0.007	0.0502	0.0167	4.1e+04	3.5e+04
	0.010	0.0156	0.0139	3.5e+04	3.6e+04
	0.013	0.0872	0.0358	4.5e+04	3.6e+04
	0.015	0.1810	0.0865	6.6e+04	2.8e+04
	0.018	0.0927	0.0475	4.9e+04	3.3e+04
	0.020	0.0830	0.0402	4.3e+04	3.7e+04
FINISH	0.005	0.2532	0.1322	7.9e+04	6.9e+04
	0.007	0.0900	0.0312	1.1e+05	1.2e+05
	0.010	0.0261	0.0213	1.2e+05	1.2e+05
	0.013	0.1243	0.0636	1.1e+05	1.4e+05
	0.015	0.2422	0.1288	1.7e+05	1.2e+05
	0.018	0.1412	0.0913	1.1e+05	1.5e+05
	0.020	0.1350	0.0658	1.1e+05	1.3e+05
UNIFORM	0.005	0.1642	0.1307	2.9e+04	1.6e+04
	0.007	0.0546	0.0173	4.7e+04	3.3e+04
	0.010	0.0080	0.0069	3.8e+04	3.9e+04
	0.013	0.0807	0.0274	5.1e+04	3.4e+04
	0.015	0.1776	0.0880	5.2e+04	8.6e+03
	0.018	0.0946	0.0429	5.6e+04	3.4e+04
	0.020	0.0977	0.0407	4.8e+04	2.7e+04

Table 6.9: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 16 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0496	0.0108	2.3e+04	1.5e+04
	0.007	0.0494	0.0132	1.1e+04	1.4e+04
	0.010	0.0178	0.0139	1.7e+04	1.5e+04
	0.013	0.0355	0.0157	2.0e+04	1.5e+04
	0.015	0.0481	0.0254	1.2e+04	1.5e+04
	0.018	0.0489	0.0209	1.1e+04	1.5e+04
	0.020	0.0237	0.0187	1.4e+04	1.5e+04
MIDDLE	0.005	0.1127	0.0351	5.1e+04	3.6e+04
	0.007	0.1253	0.0669	3.4e+04	3.3e+04
	0.010	0.0293	0.0112	4.0e+04	3.7e+04
	0.013	0.0858	0.0303	5.0e+04	3.8e+04
	0.015	0.1033	0.0461	3.5e+04	3.5e+04
	0.018	0.1131	0.0573	3.4e+04	3.4e+04
	0.020	0.0389	0.0304	3.6e+04	4.0e+04
FINISH	0.005	0.1752	0.1834	1.2e+05	1.1e+05
	0.007	0.2175	0.1758	8.0e+04	1.3e+05
	0.010	0.0394	0.0133	1.1e+05	1.2e+05
	0.013	0.1481	0.0998	1.2e+05	1.4e+05
	0.015	0.1801	0.1638	8.9e+04	1.1e+05
	0.018	0.1905	0.1706	8.8e+04	1.2e+05
	0.020	0.0384	0.0252	1.1e+05	1.2e+05
UNIFORM	0.005	0.1182	0.0536	4.4e+04	3.8e+04
	0.007	0.1429	0.1263	2.8e+04	2.7e+04
	0.010	0.0249	0.0105	4.4e+04	4.1e+04
	0.013	0.0967	0.0384	4.3e+04	3.5e+04
	0.015	0.1163	0.0786	2.9e+04	3.2e+04
	0.018	0.1264	0.0964	2.8e+04	3.3e+04
	0.020	0.0326	0.0293	3.7e+04	3.9e+04

Table 6.10: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 32 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0087	0.0066	1.6e+04	1.5e+04
	0.007	0.0397	0.0134	1.2e+04	1.5e+04
	0.010	0.0359	0.0130	2.0e+04	1.5e+04
	0.013	0.0595	0.0203	2.3e+04	1.5e+04
	0.015	0.0256	0.0238	1.6e+04	1.6e+04
	0.018	0.0417	0.0273	2.0e+04	1.5e+04
	0.020	0.0303	0.0271	1.7e+04	1.6e+04
MIDDLE	0.005	0.0143	0.0053	4.1e+04	3.9e+04
	0.007	0.0995	0.0425	3.5e+04	3.8e+04
	0.010	0.0792	0.0250	5.0e+04	4.0e+04
	0.013	0.1336	0.0557	5.8e+04	3.6e+04
	0.015	0.0199	0.0186	3.9e+04	3.9e+04
	0.018	0.0892	0.0428	5.2e+04	3.9e+04
	0.020	0.0410	0.0339	4.0e+04	3.9e+04
FINISH	0.005	0.0213	0.0058	1.1e+05	1.1e+05
	0.007	0.1716	0.1376	9.3e+04	1.2e+05
	0.010	0.1507	0.0915	1.3e+05	1.3e+05
	0.013	0.2217	0.1664	1.6e+05	9.9e+04
	0.015	0.0180	0.0172	1.1e+05	1.1e+05
	0.018	0.1564	0.0867	1.3e+05	1.3e+05
	0.020	0.0324	0.0298	1.1e+05	1.1e+05
UNIFORM	0.005	0.0160	0.0066	4.4e+04	4.2e+04
	0.007	0.1151	0.0704	3.0e+04	3.0e+04
	0.010	0.1020	0.0408	4.1e+04	3.1e+04
	0.013	0.1499	0.0844	4.6e+04	2.4e+04
	0.015	0.0197	0.0192	4.3e+04	4.2e+04
	0.018	0.1019	0.0497	4.2e+04	3.2e+04
	0.020	0.0313	0.0262	4.4e+04	4.2e+04

Table 6.11: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 64 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0542	0.0161	2.2e+04	1.5e+04
	0.007	0.0145	0.0110	1.5e+04	1.6e+04
	0.010	0.0800	0.0300	1.0e+04	1.4e+04
	0.013	0.0709	0.0265	1.1e+04	1.5e+04
	0.015	0.0239	0.0198	1.4e+04	1.5e+04
	0.018	0.0303	0.0251	1.8e+04	1.5e+04
	0.020	0.0282	0.0258	1.7e+04	1.6e+04
MIDDLE	0.005	0.1354	0.0650	6.4e+04	3.6e+04
	0.007	0.0183	0.0114	3.9e+04	3.9e+04
	0.010	0.1866	0.1402	2.7e+04	3.1e+04
	0.013	0.1592	0.1037	3.0e+04	3.0e+04
	0.015	0.0380	0.0209	3.6e+04	4.0e+04
	0.018	0.0495	0.0265	4.6e+04	3.9e+04
	0.020	0.0442	0.0249	4.5e+04	3.9e+04
FINISH	0.005	0.2312	0.1393	1.7e+05	1.2e+05
	0.007	0.0220	0.0115	1.2e+05	1.1e+05
	0.010	0.2840	0.1718	8.2e+04	5.6e+05
	0.013	0.2506	0.1277	7.9e+04	9.5e+04
	0.015	0.0520	0.0190	1.2e+05	1.1e+05
	0.018	0.0794	0.0313	1.1e+05	1.2e+05
	0.020	0.0830	0.0398	1.1e+05	1.2e+05
UNIFORM	0.005	0.1555	0.0801	5.3e+04	1.4e+04
	0.007	0.0190	0.0105	4.4e+04	4.2e+04
	0.010	0.2010	0.1273	3.3e+04	5.4e+04
	0.013	0.1757	0.1409	2.7e+04	2.2e+04
	0.015	0.0416	0.0197	3.9e+04	3.9e+04
	0.018	0.0573	0.0264	4.3e+04	3.8e+04
	0.020	0.0509	0.0274	3.9e+04	4.0e+04

Table 6.12: Experiment 6.2: Comparison of data assimilations for a set of distributions and errors with 128 measurements

Distribution of measurements	σ^2	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0530	0.0141	1.1e+04	1.5e+04
	0.007	0.0368	0.0119	1.2e+04	1.5e+04
	0.010	0.0441	0.0169	1.2e+04	1.5e+04
	0.013	0.0227	0.0182	1.5e+04	1.6e+04
	0.015	0.0349	0.0185	1.2e+04	1.6e+04
	0.018	0.0469	0.0268	1.2e+04	1.6e+04
	0.020	0.0807	0.0379	2.3e+04	1.5e+04
MIDDLE	0.005	0.1235	0.0631	3.4e+04	3.5e+04
	0.007	0.0971	0.0397	3.6e+04	3.7e+04
	0.010	0.0981	0.0429	3.4e+04	3.9e+04
	0.013	0.0306	0.0210	4.1e+04	3.9e+04
	0.015	0.0748	0.0311	3.5e+04	4.1e+04
	0.018	0.0952	0.0441	3.5e+04	3.9e+04
	0.020	0.1770	0.0844	6.7e+04	3.4e+04
FINISH	0.005	0.2040	0.1753	8.7e+04	9.9e+04
	0.007	0.1871	0.1869	8.6e+04	9.8e+04
	0.010	0.1527	0.1052	1.0e+05	1.3e+05
	0.013	0.0327	0.0355	1.1e+05	1.1e+05
	0.015	0.1251	0.0664	1.1e+05	1.3e+05
	0.018	0.1643	0.1274	9.9e+04	1.3e+05
	0.020	0.2742	0.1449	1.7e+05	7.9e+04
UNIFORM	0.005	0.1406	0.1152	2.8e+04	2.5e+04
	0.007	0.1230	0.0923	3.2e+04	3.7e+04
	0.010	0.1088	0.0577	3.0e+04	2.9e+04
	0.013	0.0280	0.0226	3.9e+04	4.4e+04
	0.015	0.0862	0.0363	3.5e+04	3.6e+04
	0.018	0.1153	0.0670	3.2e+04	3.1e+04
	0.020	0.1919	0.1158	6.1e+04	2.1e+04

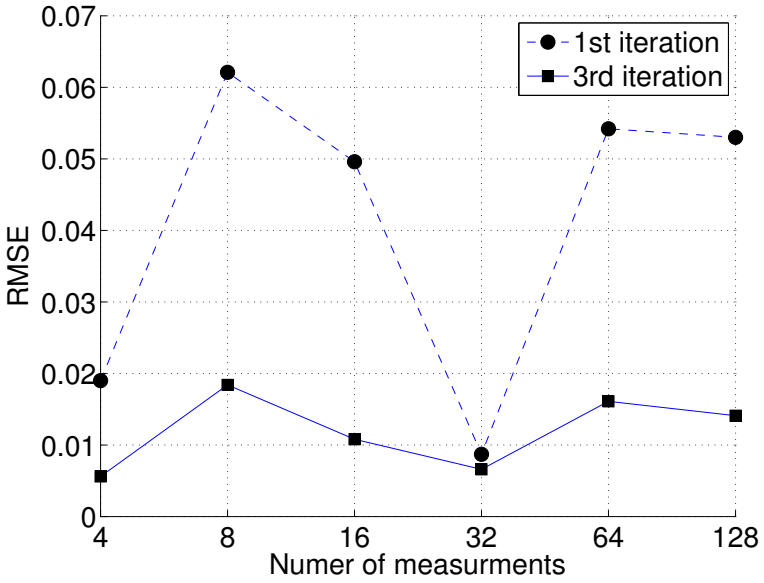


Figure 6.13: Experiment 6.2: Root-mean-square (RMSE) error of observations as a function of number of experiments $\sigma^2 = 0.005$, START measurement distribution.

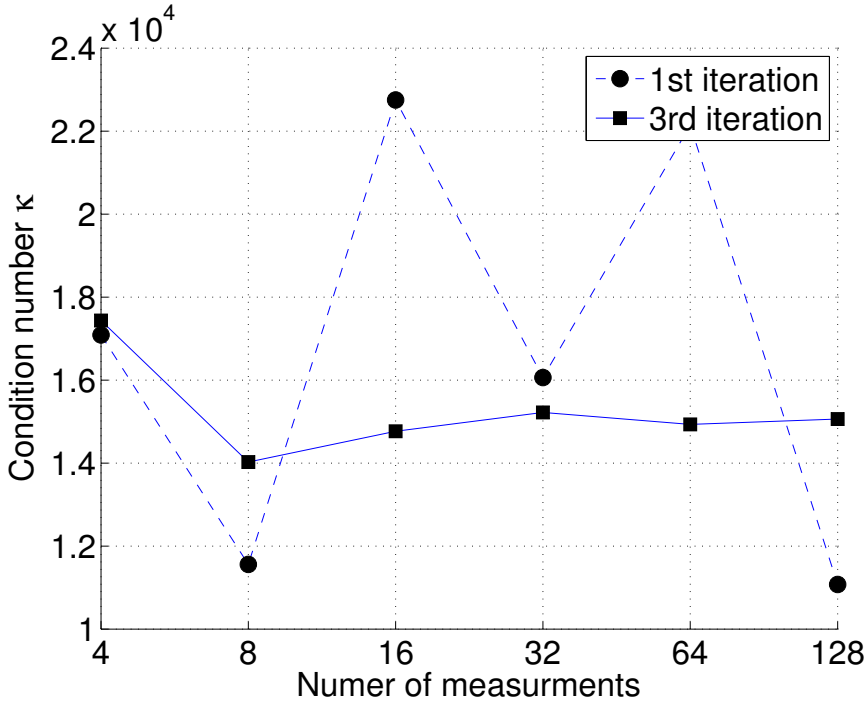


Figure 6.14: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.005$, START measurement distribution.

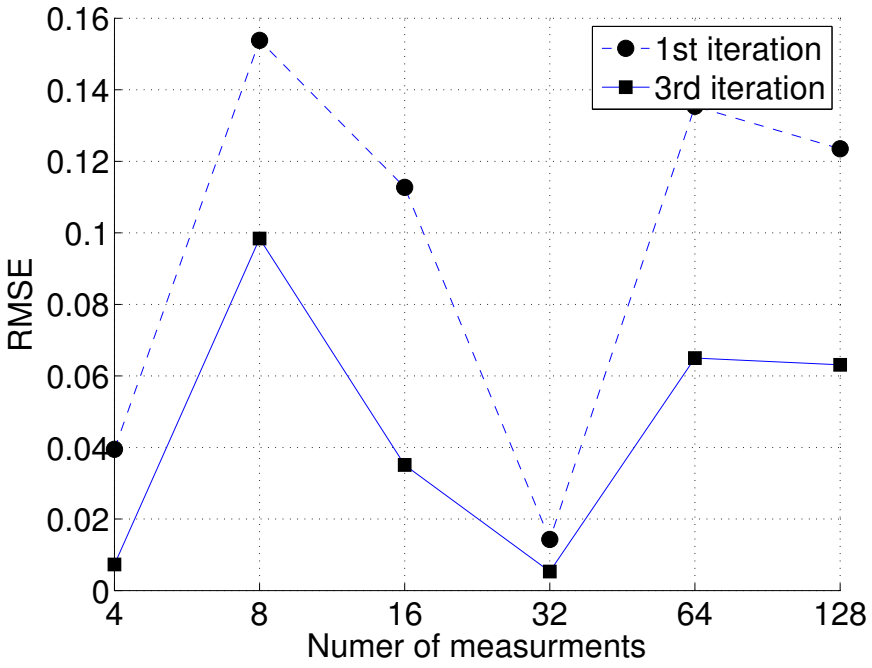


Figure 6.15: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.005$, MIDDLE measurement distribution.

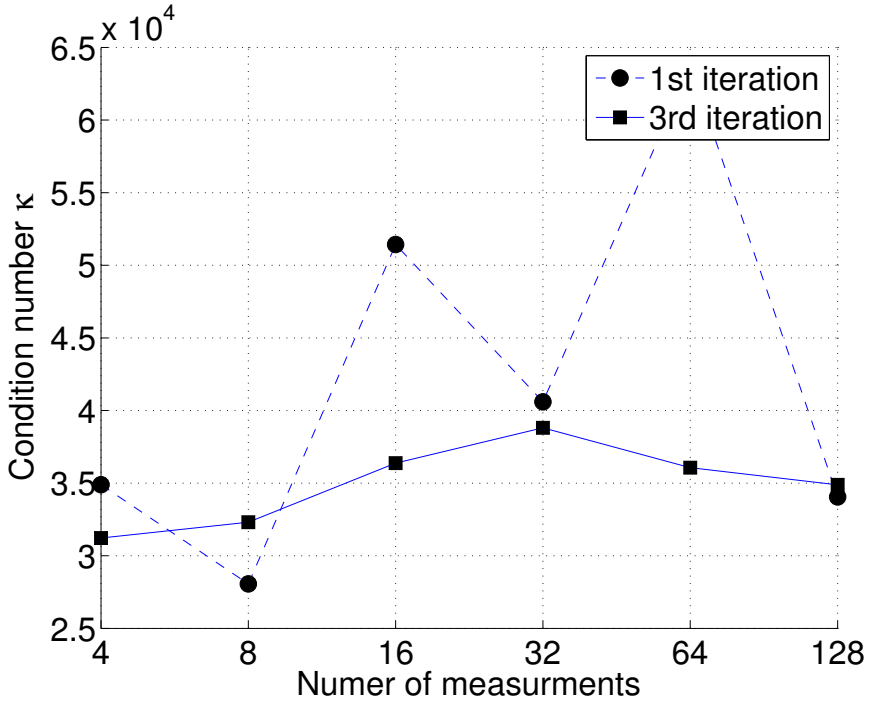


Figure 6.16: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.005$, MIDDLE measurement distribution.

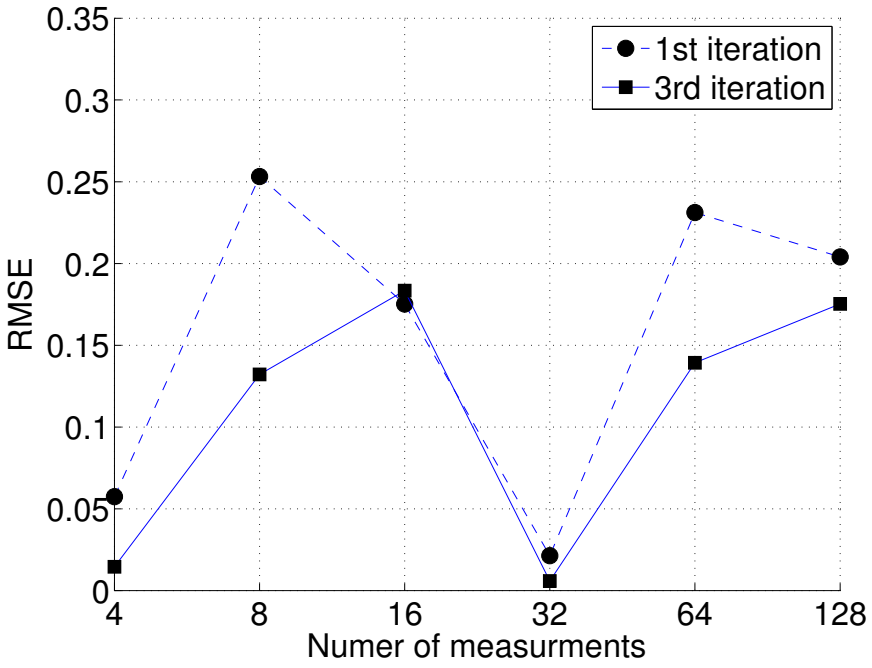


Figure 6.17: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.005$, FINISH measurement distribution.

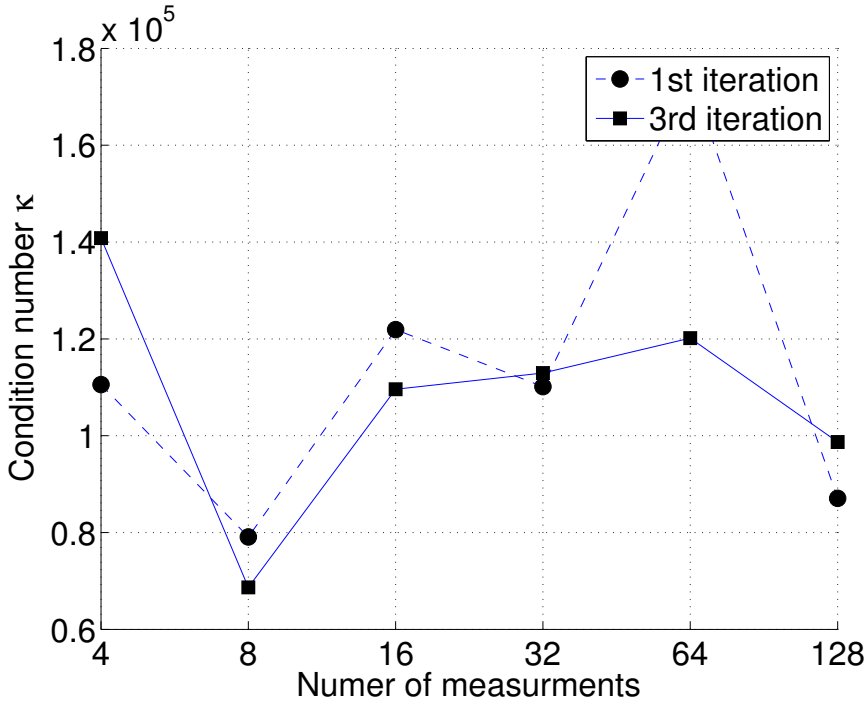


Figure 6.18: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.005$, FINISH measurement distribution.

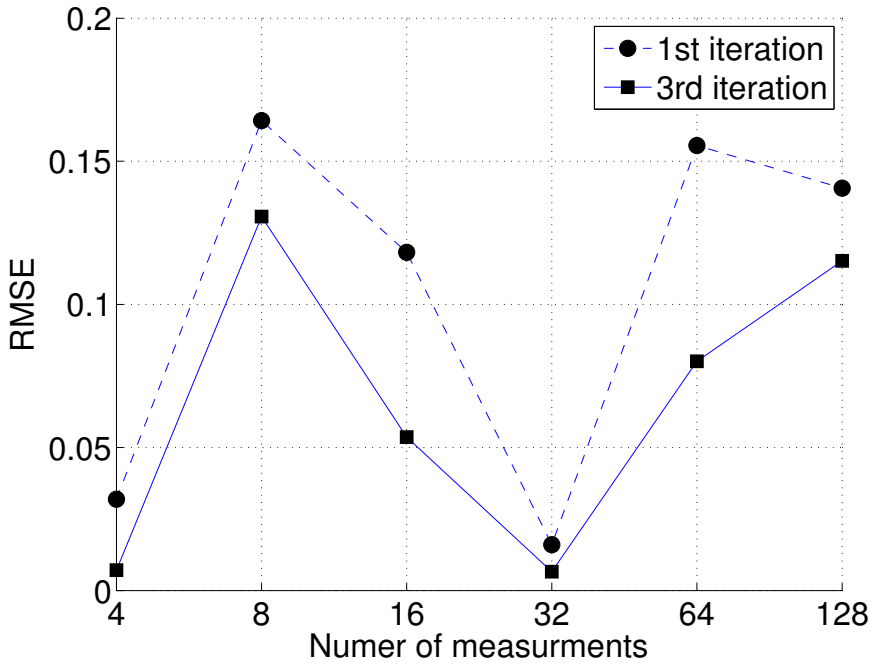


Figure 6.19: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.005$, UNIFORM measurement distribution.

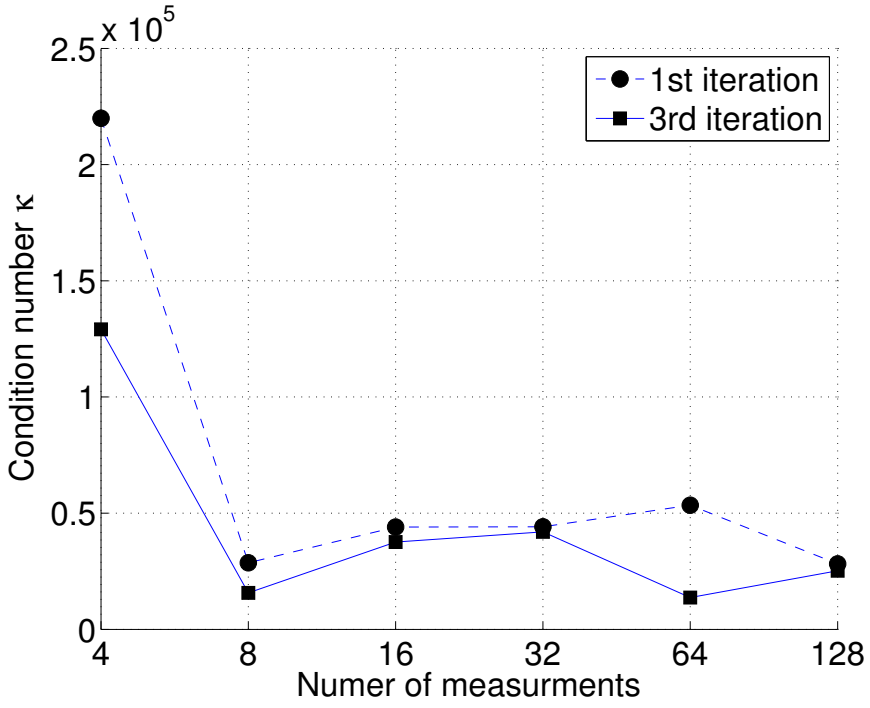


Figure 6.20: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.005$, UNIFORM measurement distribution.

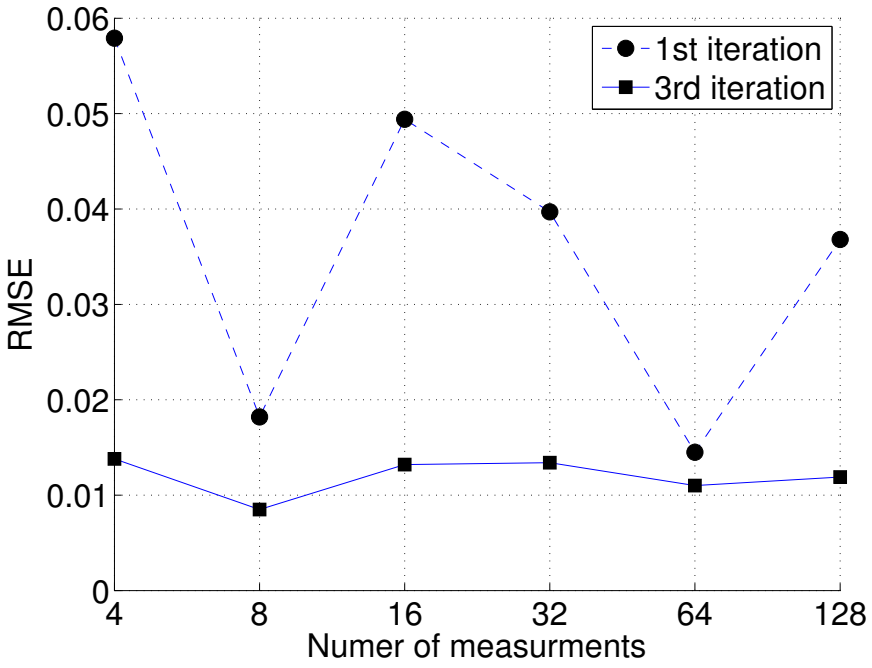


Figure 6.21: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.0075$, START measurement distribution.

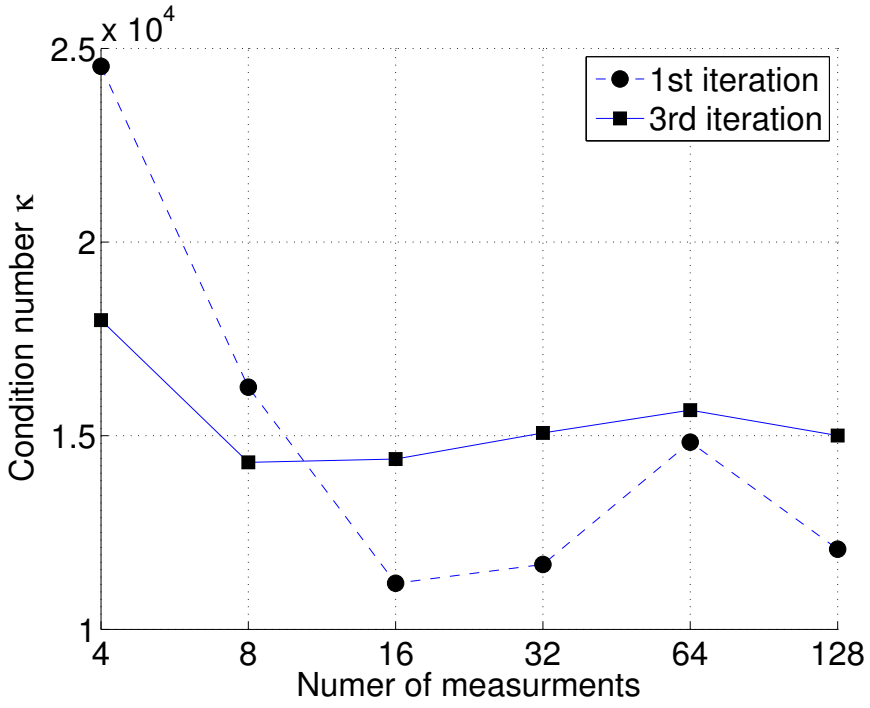


Figure 6.22: Experiment 6.2: Condition number κ a function of number of experiments $\sigma^2 = 0.0075$, START measurement distribution.

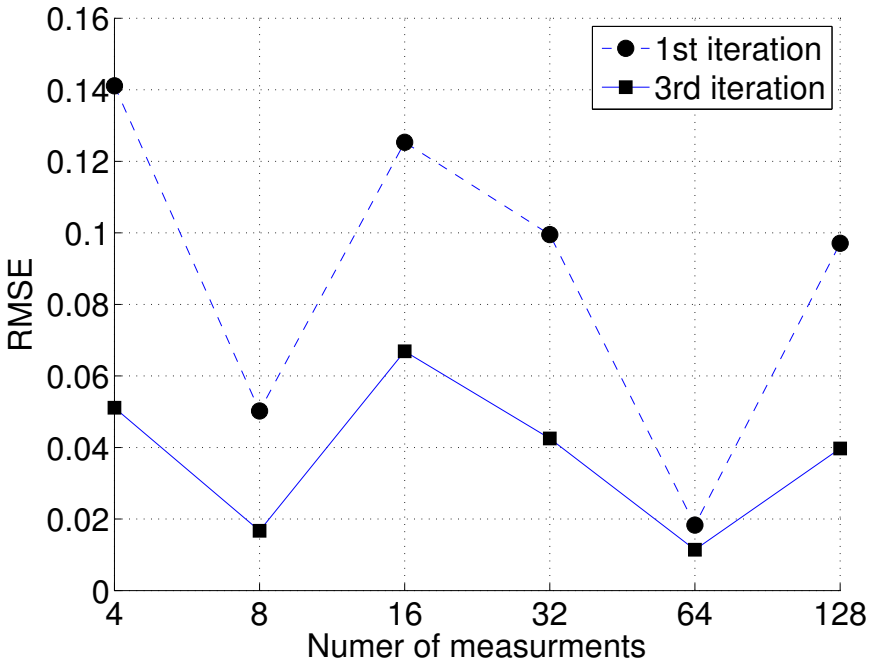


Figure 6.23: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.0075$, MIDDLE measurement distribution.

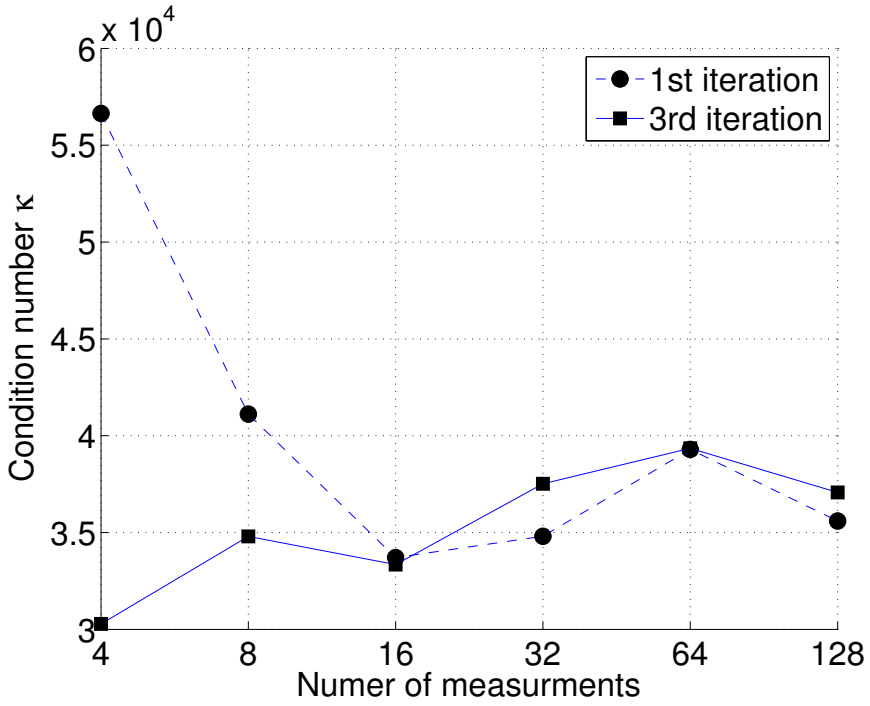


Figure 6.24: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.0075$, MIDDLE measurement distribution.

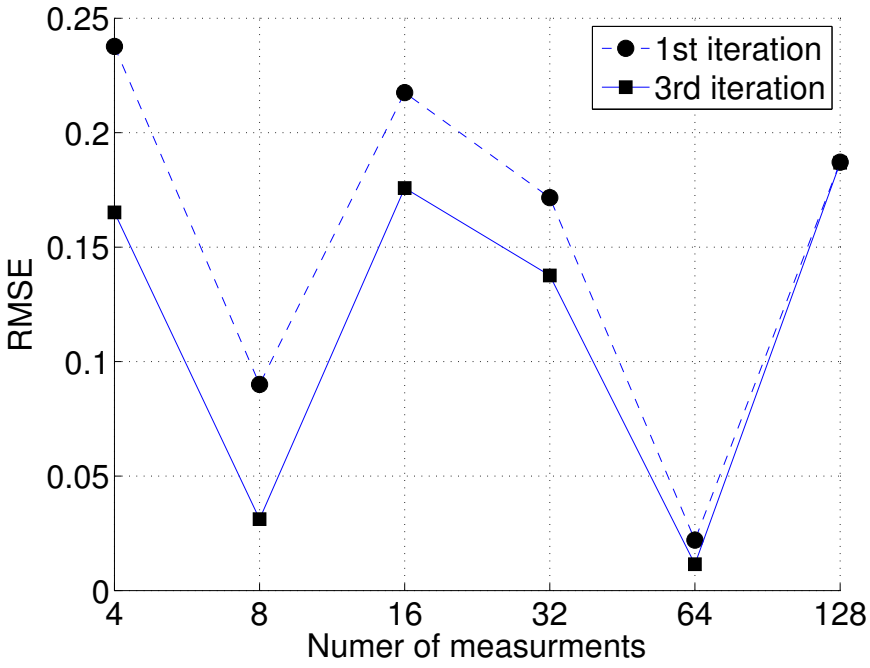


Figure 6.25: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.0075$, FINISH measurement distribution.

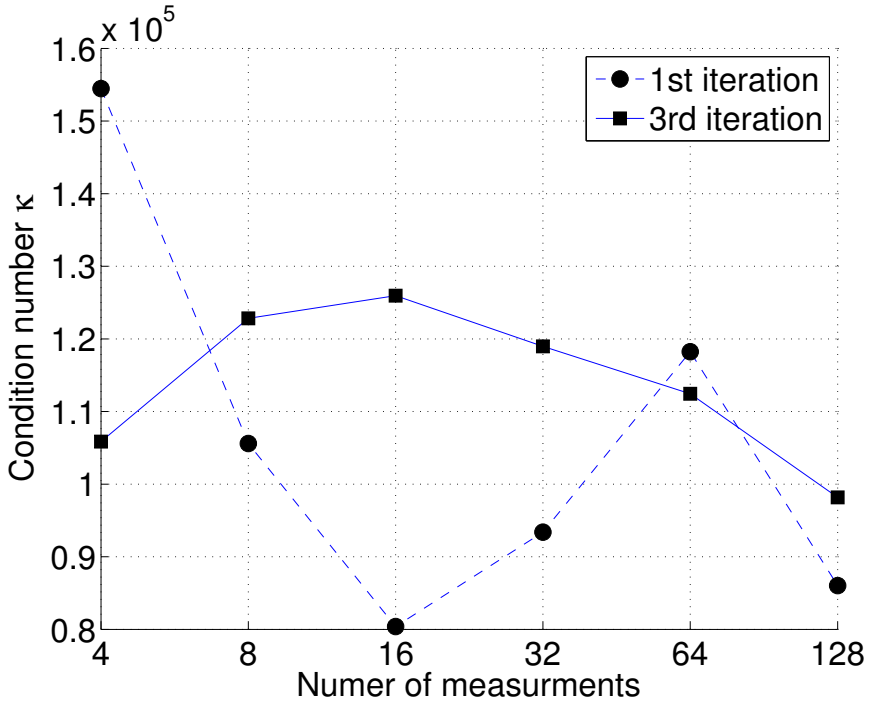


Figure 6.26: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.0075$, FINISH measurement distribution.

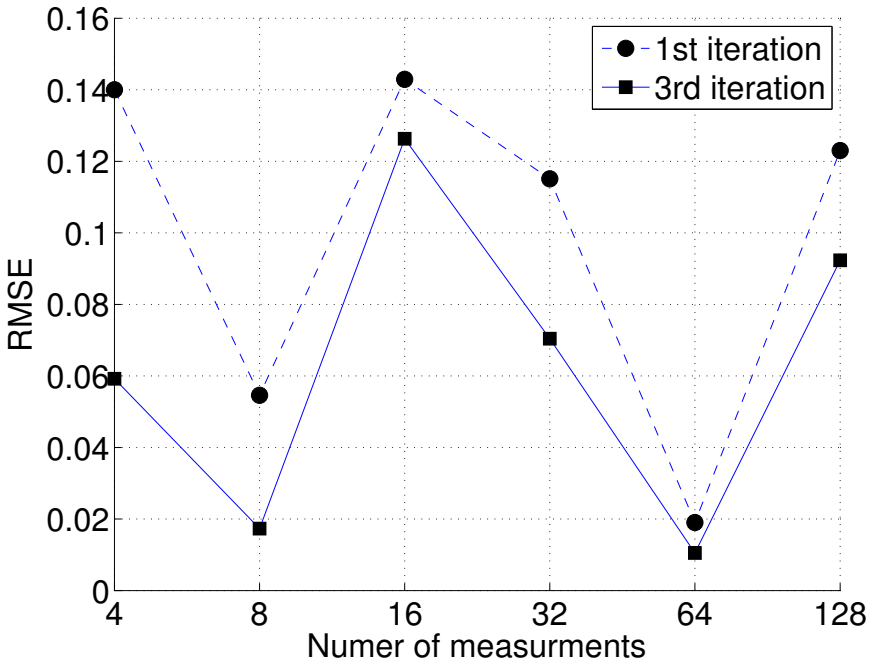


Figure 6.27: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.0075$, UNIFORM measurement distribution.

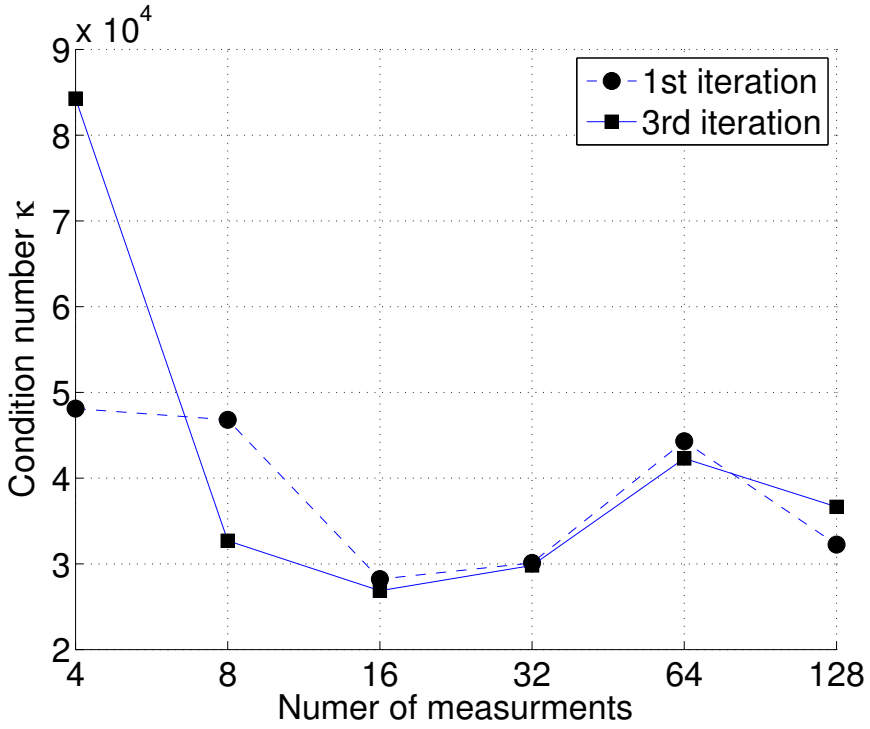


Figure 6.28: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.0075$, UNIFORM measurement distribution.

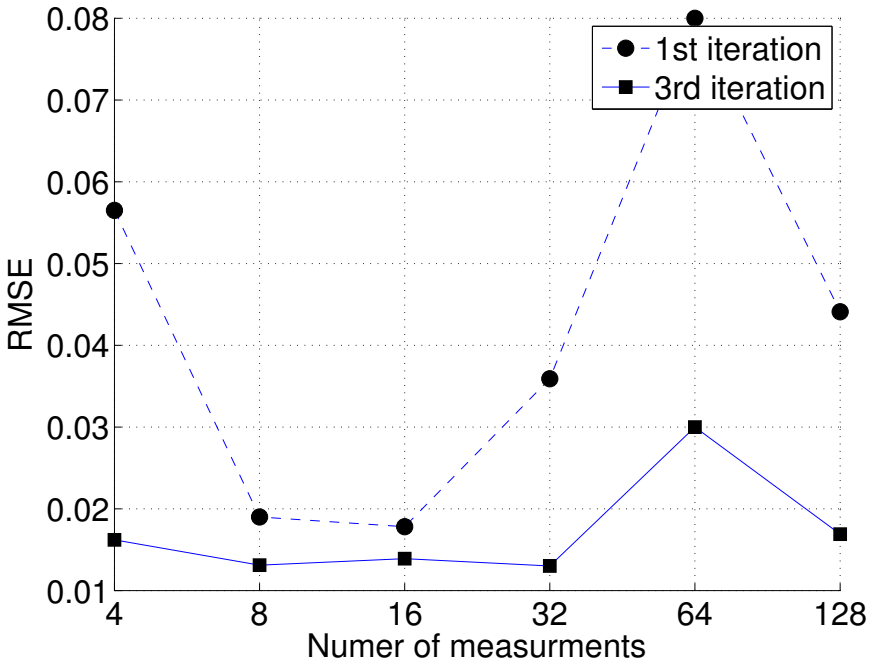


Figure 6.29: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.01$, START measurement distribution.

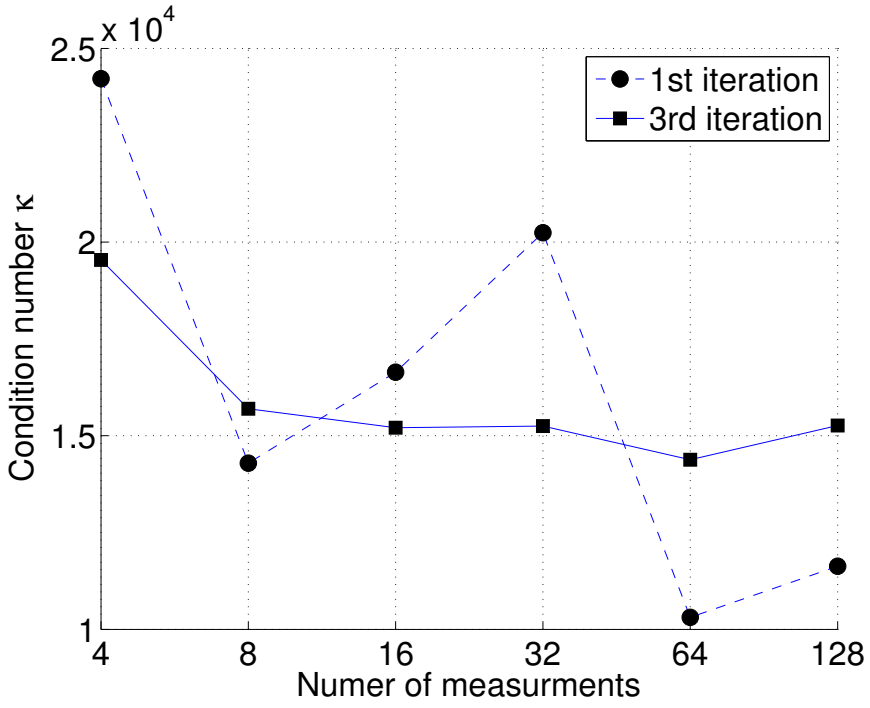


Figure 6.30: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.01$, START measurement distribution.

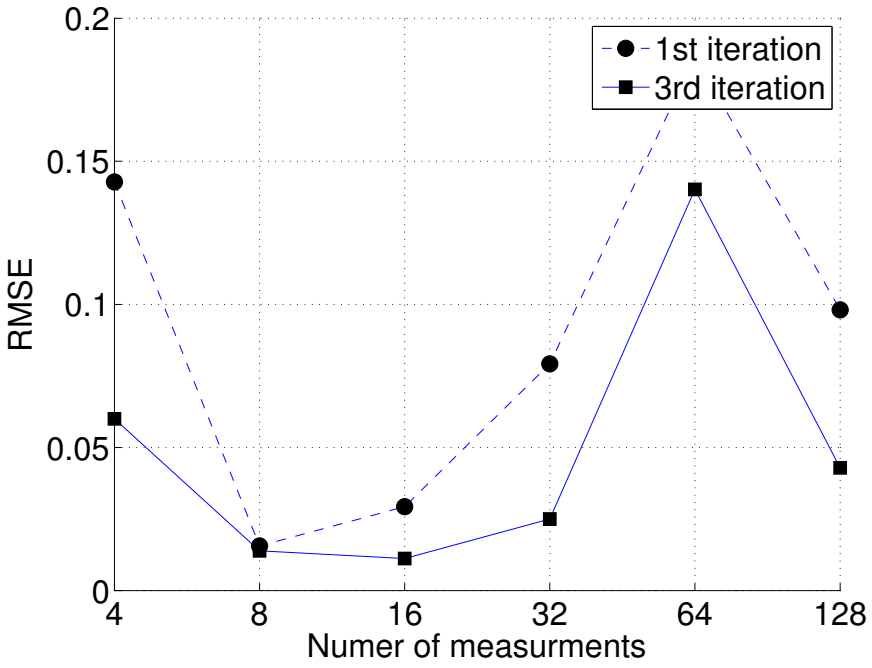


Figure 6.31: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.01$, MIDDLE measurement distribution.

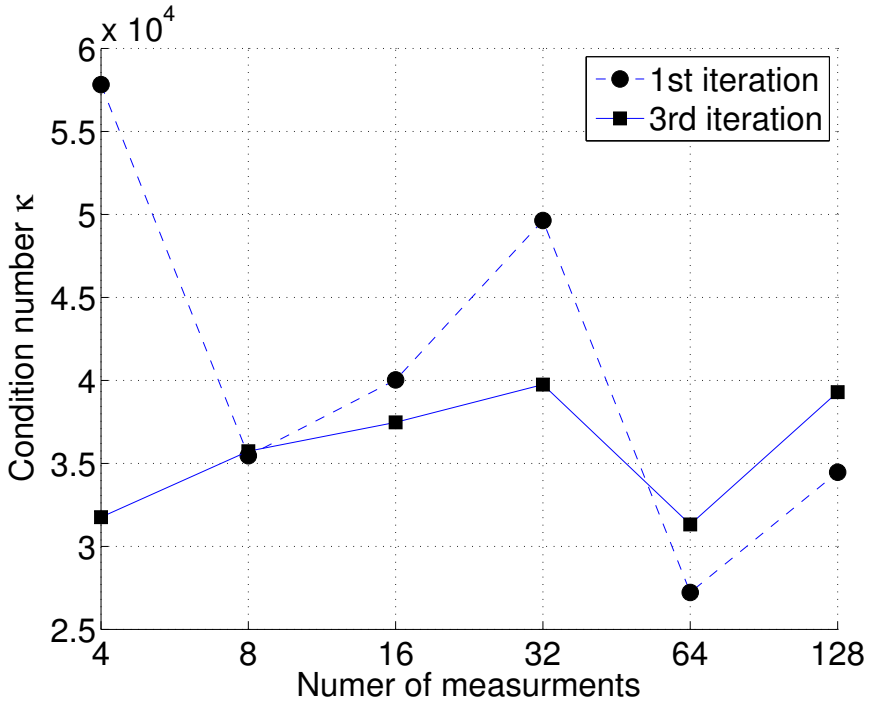


Figure 6.32: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.01$, MIDDLE measurement distribution.

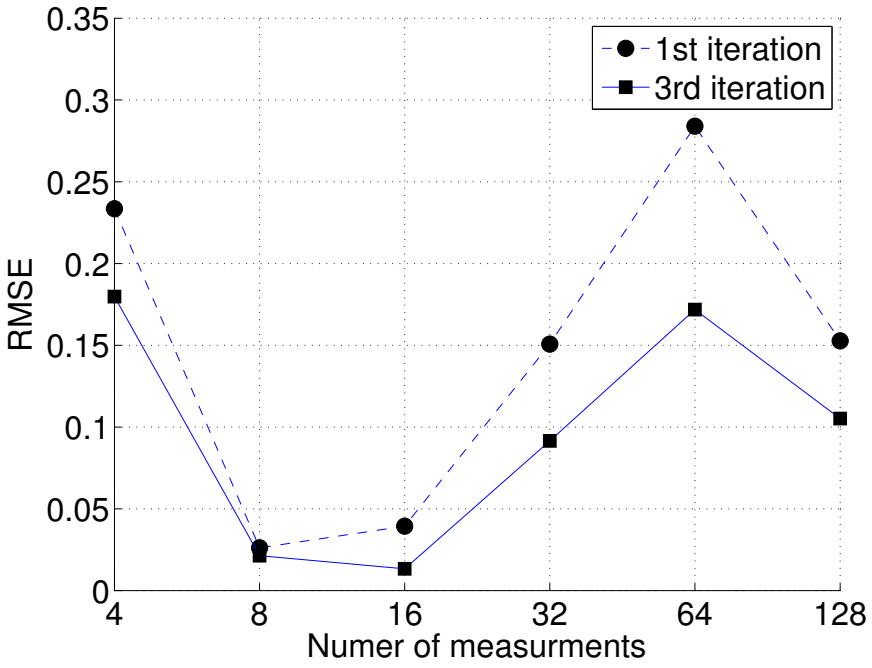


Figure 6.33: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.01$, FINISH measurement distribution.

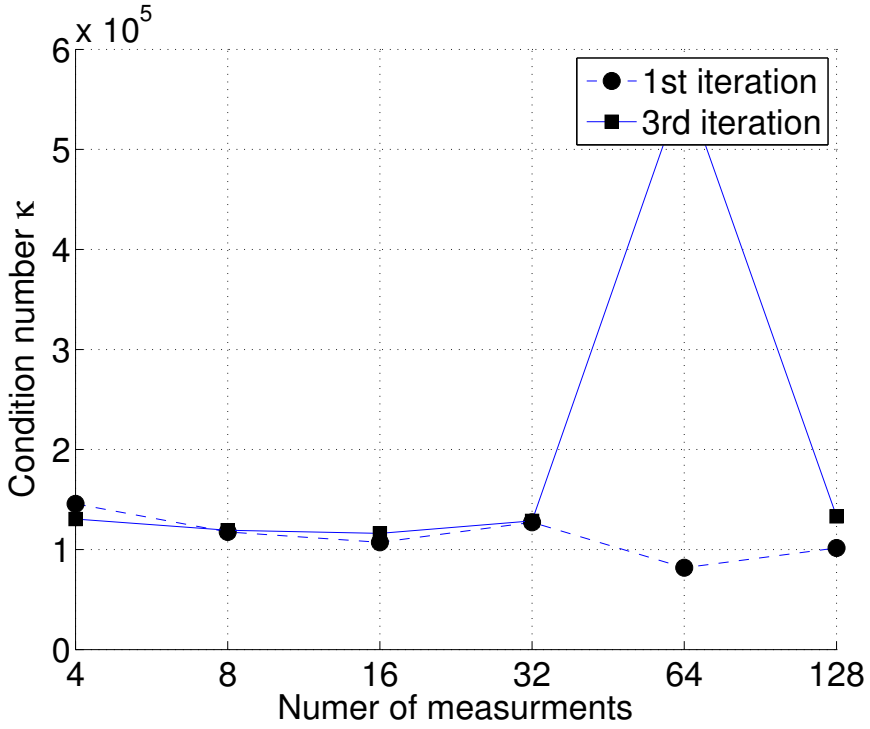


Figure 6.34: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.01$, FINISH measurement distribution.

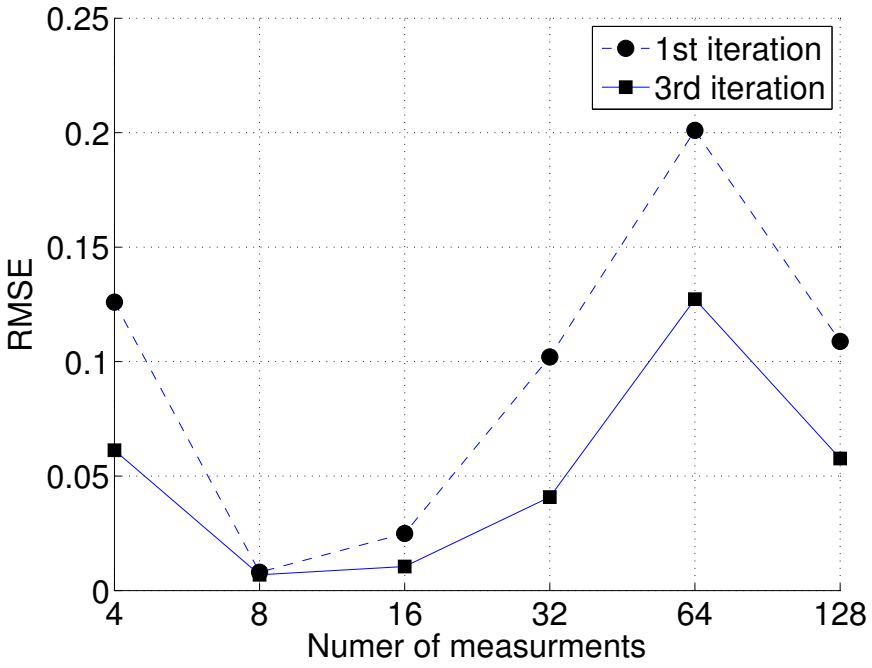


Figure 6.35: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.01$, UNIFORM measurement distribution.

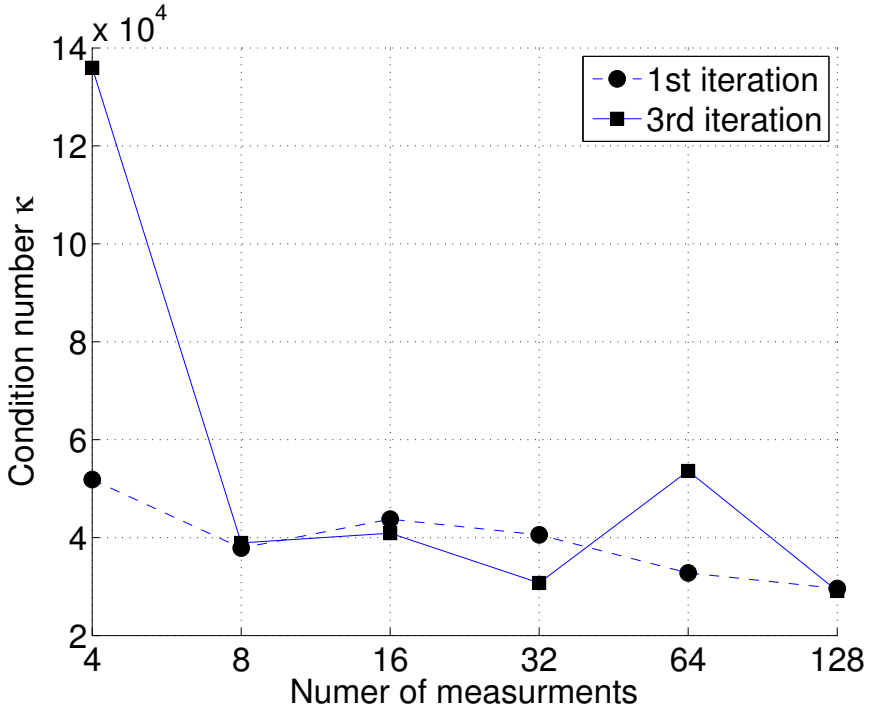


Figure 6.36: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.01$, UNIFORM measurement distribution.

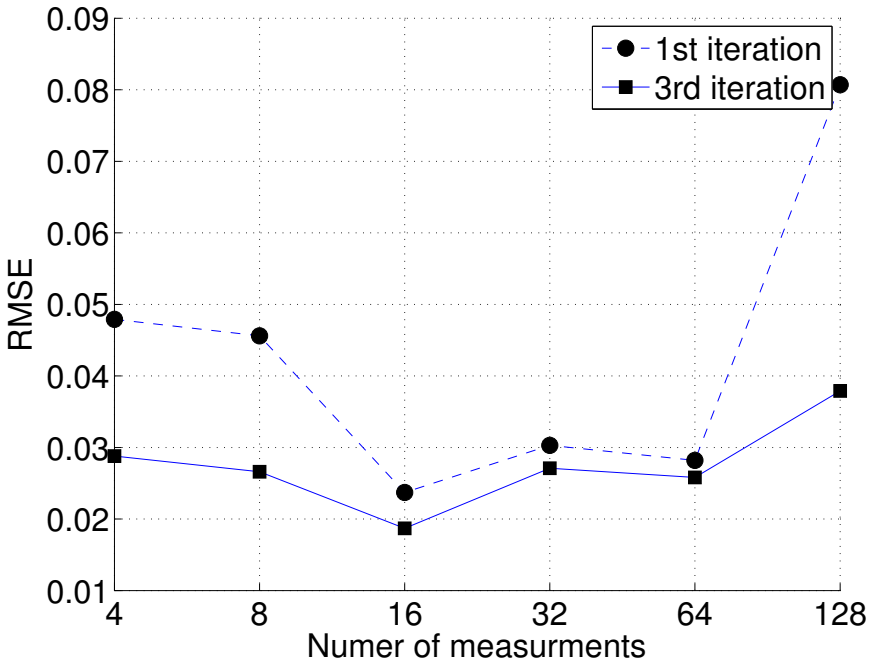


Figure 6.37: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.02$, START measurement distribution.

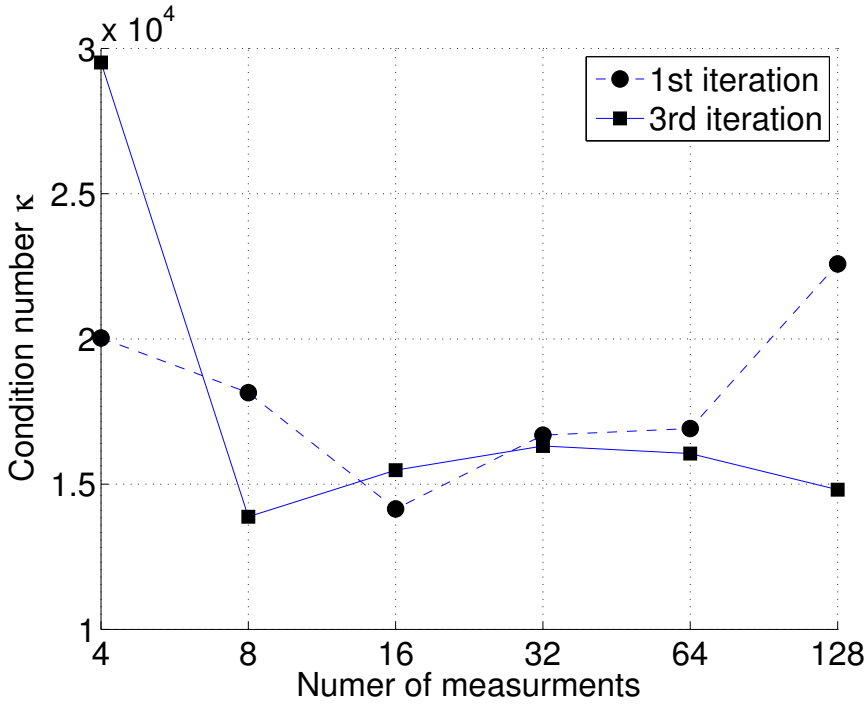


Figure 6.38: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.02$, START measurement distribution.

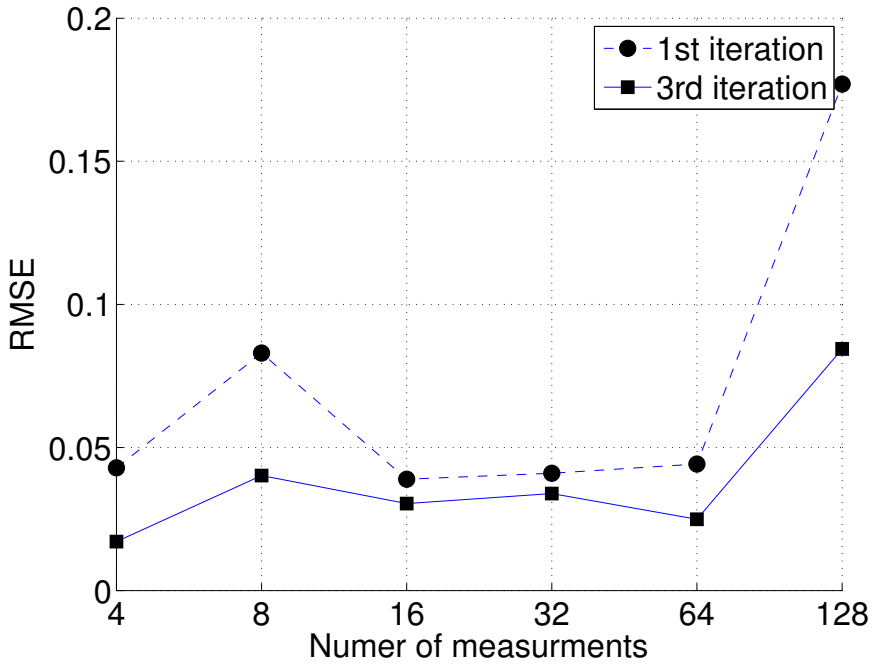


Figure 6.39: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.02$, MIDDLE measurement distribution.

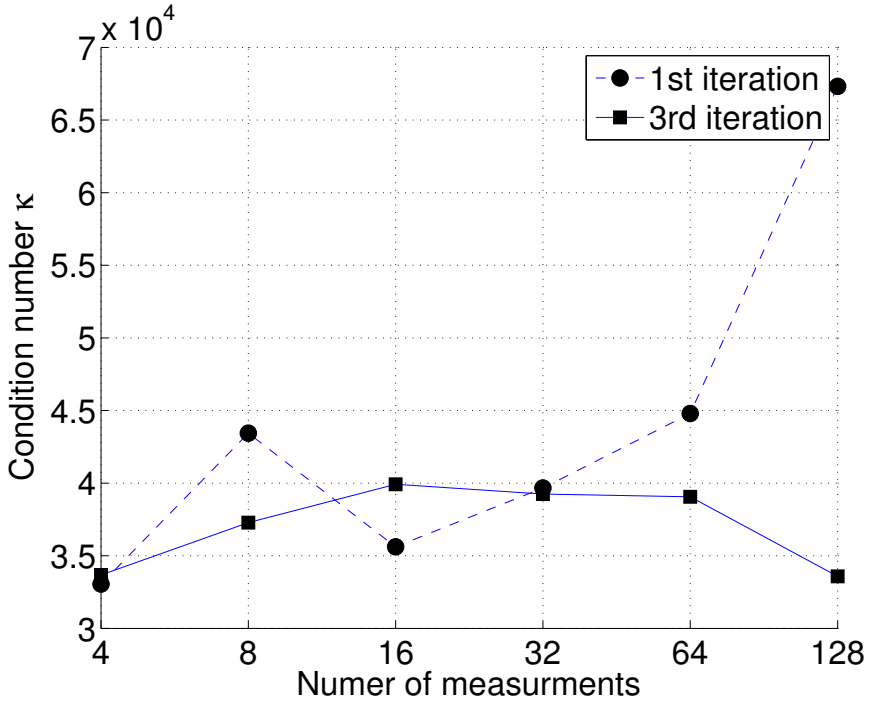


Figure 6.40: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.02$, MIDDLE measurement distribution.

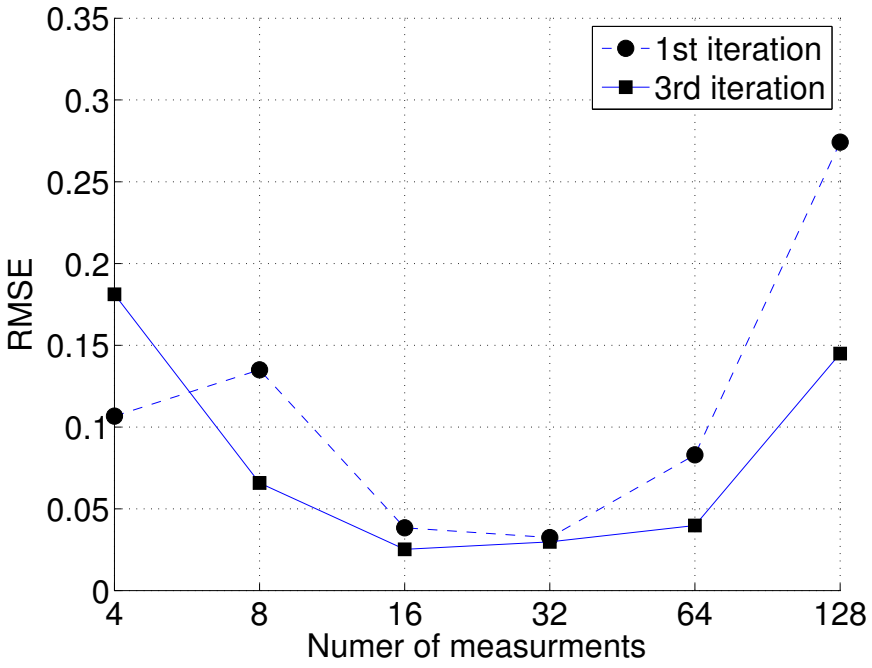


Figure 6.41: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.02$, FINISH measurement distribution.

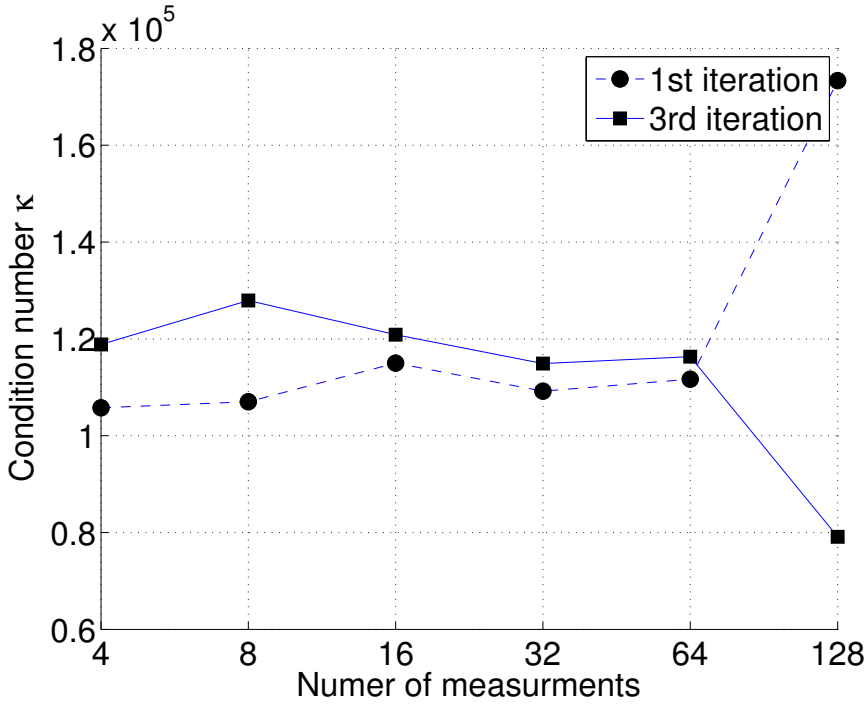


Figure 6.42: Experiment 6.2: Condition number κ as a function of number of experiments $\sigma^2 = 0.02$, FINISH measurement distribution.

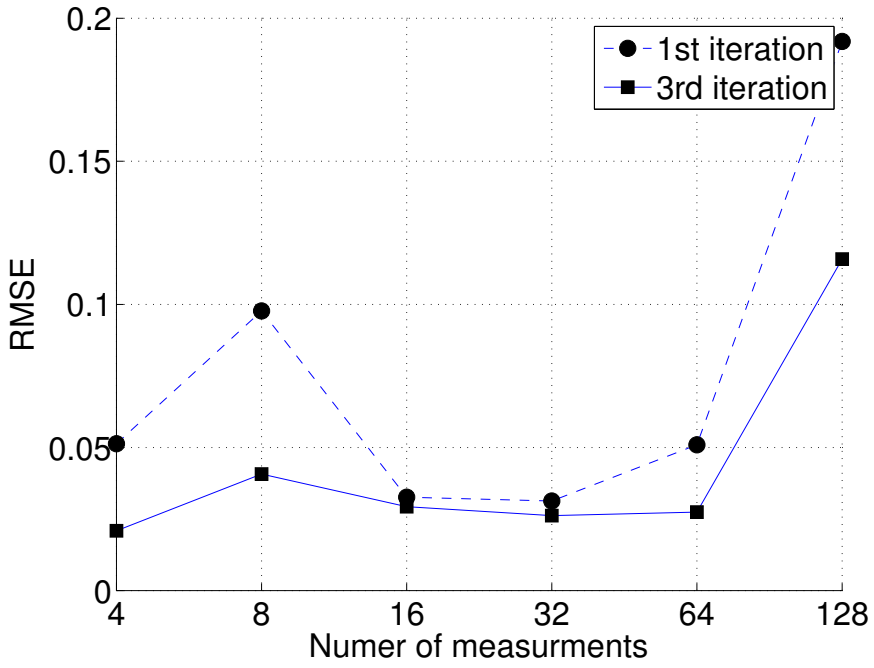


Figure 6.43: Experiment 6.2: Root-mean-square error of observations as a function of number of experiments $\sigma^2 = 0.02$, UNIFORM measurement distribution.

Table 6.13: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.005$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0190	0.0056	1.7e+04	1.7e+04
	8	0.0621	0.0184	1.2e+04	1.4e+04
	16	0.0496	0.0108	2.3e+04	1.5e+04
	32	0.0087	0.0066	1.6e+04	1.5e+04
	64	0.0542	0.0161	2.2e+04	1.5e+04
	128	0.0530	0.0141	1.1e+04	1.5e+04
MIDDLE	4	0.0395	0.0073	3.5e+04	3.1e+04
	8	0.1538	0.0984	2.8e+04	3.2e+04
	16	0.1127	0.0351	5.1e+04	3.6e+04
	32	0.0143	0.0053	4.1e+04	3.9e+04
	64	0.1354	0.0650	6.4e+04	3.6e+04
	128	0.1235	0.0631	3.4e+04	3.5e+04
FINISH	4	0.0574	0.0145	1.1e+05	1.4e+05
	8	0.2532	0.1322	7.9e+04	6.9e+04
	16	0.1752	0.1834	1.2e+05	1.1e+05
	32	0.0213	0.0058	1.1e+05	1.1e+05
	64	0.2312	0.1393	1.7e+05	1.2e+05
	128	0.2040	0.1753	8.7e+04	9.9e+04
UNIFORM	4	0.0319	0.0071	2.2e+05	1.3e+05
	8	0.1642	0.1307	2.9e+04	1.6e+04
	16	0.1182	0.0536	4.4e+04	3.8e+04
	32	0.0160	0.0066	4.4e+04	4.2e+04
	64	0.1555	0.0801	5.3e+04	1.4e+04
	128	0.1406	0.1152	2.8e+04	2.5e+04

Table 6.14: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.075$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0579	0.0138	2.5e+04	1.8e+04
	8	0.0182	0.0085	1.6e+04	1.4e+04
	16	0.0494	0.0132	1.1e+04	1.4e+04
	32	0.0397	0.0134	1.2e+04	1.5e+04
	64	0.0145	0.0110	1.5e+04	1.6e+04
	128	0.0368	0.0119	1.2e+04	1.5e+04
MIDDLE	4	0.1411	0.0511	5.7e+04	3.0e+04
	8	0.0502	0.0167	4.1e+04	3.5e+04
	16	0.1253	0.0669	3.4e+04	3.3e+04
	32	0.0995	0.0425	3.5e+04	3.8e+04
	64	0.0183	0.0114	3.9e+04	3.9e+04
	128	0.0971	0.0397	3.6e+04	3.7e+04
FINISH	4	0.2377	0.1652	1.5e+05	1.1e+05
	8	0.0900	0.0312	1.1e+05	1.2e+05
	16	0.2175	0.1758	8.0e+04	1.3e+05
	32	0.1716	0.1376	9.3e+04	1.2e+05
	64	0.0220	0.0115	1.2e+05	1.1e+05
	128	0.1871	0.1869	8.6e+04	9.8e+04
UNIFORM	4	0.1400	0.0592	4.8e+04	8.4e+04
	8	0.0546	0.0173	4.7e+04	3.3e+04
	16	0.1429	0.1263	2.8e+04	2.7e+04
	32	0.1151	0.0704	3.0e+04	3.0e+04
	64	0.0190	0.0105	4.4e+04	4.2e+04
	128	0.1230	0.0923	3.2e+04	3.7e+04

Table 6.15: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.01$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0565	0.0162	2.4e+04	2.0e+04
	8	0.0190	0.0131	1.4e+04	1.6e+04
	16	0.0178	0.0139	1.7e+04	1.5e+04
	32	0.0359	0.0130	2.0e+04	1.5e+04
	64	0.0800	0.0300	1.0e+04	1.4e+04
	128	0.0441	0.0169	1.2e+04	1.5e+04
MIDDLE	4	0.1428	0.0600	5.8e+04	3.2e+04
	8	0.0156	0.0139	3.5e+04	3.6e+04
	16	0.0293	0.0112	4.0e+04	3.7e+04
	32	0.0792	0.0250	5.0e+04	4.0e+04
	64	0.1866	0.1402	2.7e+04	3.1e+04
	128	0.0981	0.0429	3.4e+04	3.9e+04
FINISH	4	0.2335	0.1798	1.5e+05	1.3e+05
	8	0.0261	0.0213	1.2e+05	1.2e+05
	16	0.0394	0.0133	1.1e+05	1.2e+05
	32	0.1507	0.0915	1.3e+05	1.3e+05
	64	0.2840	0.1718	8.2e+04	5.6e+05
	128	0.1527	0.1052	1.0e+05	1.3e+05
UNIFORM	4	0.1260	0.0613	5.2e+04	1.4e+05
	8	0.0080	0.0069	3.8e+04	3.9e+04
	16	0.0249	0.0105	4.4e+04	4.1e+04
	32	0.1020	0.0408	4.1e+04	3.1e+04
	64	0.2010	0.1273	3.3e+04	5.4e+04
	128	0.1088	0.0577	3.0e+04	2.9e+04

Table 6.16: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.0125$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0361	0.0122	1.7e+04	1.8e+04
	8	0.0309	0.0141	1.8e+04	1.5e+04
	16	0.0355	0.0157	2.0e+04	1.5e+04
	32	0.0595	0.0203	2.3e+04	1.5e+04
	64	0.0709	0.0265	1.1e+04	1.5e+04
	128	0.0227	0.0182	1.5e+04	1.6e+04
MIDDLE	4	0.0675	0.0188	3.8e+04	3.1e+04
	8	0.0872	0.0358	4.5e+04	3.6e+04
	16	0.0858	0.0303	5.0e+04	3.8e+04
	32	0.1336	0.0557	5.8e+04	3.6e+04
	64	0.1592	0.1037	3.0e+04	3.0e+04
	128	0.0306	0.0210	4.1e+04	3.9e+04
FINISH	4	0.1174	0.0609	9.2e+04	1.2e+05
	8	0.1243	0.0636	1.1e+05	1.4e+05
	16	0.1481	0.0998	1.2e+05	1.4e+05
	32	0.2217	0.1664	1.6e+05	9.9e+04
	64	0.2506	0.1277	7.9e+04	9.5e+04
	128	0.0327	0.0355	1.1e+05	1.1e+05
UNIFORM	4	0.0773	0.1497	2.5e+05	5.1e+04
	8	0.0807	0.0274	5.1e+04	3.4e+04
	16	0.0967	0.0384	4.3e+04	3.5e+04
	32	0.1499	0.0844	4.6e+04	2.4e+04
	64	0.1757	0.1409	2.7e+04	2.2e+04
	128	0.0280	0.0226	3.9e+04	4.4e+04

Table 6.17: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.015$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0418	0.0129	1.7e+04	1.8e+04
	8	0.0874	0.0375	2.7e+04	1.4e+04
	16	0.0481	0.0254	1.2e+04	1.5e+04
	32	0.0256	0.0238	1.6e+04	1.6e+04
	64	0.0239	0.0198	1.4e+04	1.5e+04
	128	0.0349	0.0185	1.2e+04	1.6e+04
MIDDLE	4	0.1059	0.0468	3.5e+04	3.2e+04
	8	0.1810	0.0865	6.6e+04	2.8e+04
	16	0.1033	0.0461	3.5e+04	3.5e+04
	32	0.0199	0.0186	3.9e+04	3.9e+04
	64	0.0380	0.0209	3.6e+04	4.0e+04
	128	0.0748	0.0311	3.5e+04	4.1e+04
FINISH	4	0.1774	0.1651	8.1e+04	1.0e+05
	8	0.2422	0.1288	1.7e+05	1.2e+05
	16	0.1801	0.1638	8.9e+04	1.1e+05
	32	0.0180	0.0172	1.1e+05	1.1e+05
	64	0.0520	0.0190	1.2e+05	1.1e+05
	128	0.1251	0.0664	1.1e+05	1.3e+05
UNIFORM	4	0.1153	0.5134	1.0e+06	6.0e+04
	8	0.1776	0.0880	5.2e+04	8.6e+03
	16	0.1163	0.0786	2.9e+04	3.2e+04
	32	0.0197	0.0192	4.3e+04	4.2e+04
	64	0.0416	0.0197	3.9e+04	3.9e+04
	128	0.0862	0.0363	3.5e+04	3.6e+04

Table 6.18: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.0175$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0211	0.0095	1.9e+04	1.7e+04
	8	0.0339	0.0220	1.9e+04	1.5e+04
	16	0.0489	0.0209	1.1e+04	1.5e+04
	32	0.0417	0.0273	2.0e+04	1.5e+04
	64	0.0303	0.0251	1.8e+04	1.5e+04
	128	0.0469	0.0268	1.2e+04	1.6e+04
MIDDLE	4	0.0609	0.0177	3.6e+04	3.0e+04
	8	0.0927	0.0475	4.9e+04	3.3e+04
	16	0.1131	0.0573	3.4e+04	3.4e+04
	32	0.0892	0.0428	5.2e+04	3.9e+04
	64	0.0495	0.0265	4.6e+04	3.9e+04
	128	0.0952	0.0441	3.5e+04	3.9e+04
FINISH	4	0.1185	0.0568	9.5e+04	1.3e+05
	8	0.1412	0.0913	1.1e+05	1.5e+05
	16	0.1905	0.1706	8.8e+04	1.2e+05
	32	0.1564	0.0867	1.3e+05	1.3e+05
	64	0.0794	0.0313	1.1e+05	1.2e+05
	128	0.1643	0.1274	9.9e+04	1.3e+05
UNIFORM	4	0.0628	0.0269	3.9e+04	9.7e+04
	8	0.0946	0.0429	5.6e+04	3.4e+04
	16	0.1264	0.0964	2.8e+04	3.3e+04
	32	0.1019	0.0497	4.2e+04	3.2e+04
	64	0.0573	0.0264	4.3e+04	3.8e+04
	128	0.1153	0.0670	3.2e+04	3.1e+04

Table 6.19: Experiment 6.2: Comparison of data assimilations for a set of distributions with different number of measurements and errors with $\sigma^2 = 0.02$

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	4	0.0479	0.0288	2.0e+04	3.0e+04
	8	0.0456	0.0266	1.8e+04	1.4e+04
	16	0.0237	0.0187	1.4e+04	1.5e+04
	32	0.0303	0.0271	1.7e+04	1.6e+04
	64	0.0282	0.0258	1.7e+04	1.6e+04
	128	0.0807	0.0379	2.3e+04	1.5e+04
MIDDLE	4	0.0429	0.0171	3.3e+04	3.4e+04
	8	0.0830	0.0402	4.3e+04	3.7e+04
	16	0.0389	0.0304	3.6e+04	4.0e+04
	32	0.0410	0.0339	4.0e+04	3.9e+04
	64	0.0442	0.0249	4.5e+04	3.9e+04
	128	0.1770	0.0844	6.7e+04	3.4e+04
FINISH	4	0.1067	0.1812	1.1e+05	1.2e+05
	8	0.1350	0.0658	1.1e+05	1.3e+05
	16	0.0384	0.0252	1.1e+05	1.2e+05
	32	0.0324	0.0298	1.1e+05	1.1e+05
	64	0.0830	0.0398	1.1e+05	1.2e+05
	128	0.2742	0.1449	1.7e+05	7.9e+04
UNIFORM	4	0.0513	0.0209	4.6e+04	1.7e+05
	8	0.0977	0.0407	4.8e+04	2.7e+04
	16	0.0326	0.0293	3.7e+04	3.9e+04
	32	0.0313	0.0262	4.4e+04	4.2e+04
	64	0.0509	0.0274	3.9e+04	4.0e+04
	128	0.1919	0.1158	6.1e+04	2.1e+04

6.2.3 Experiment 6.3

In a setup similar to Experiment 5.2, we are going to conduct data assimilation experiments.

Table 6.20: Experiment 6.3: Base control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.000	0.000	0.500	0.000	0.0	0.0	0.5	2.5e-05

Table 6.21: Experiment 6.3: Perturbed control vector $\boldsymbol{\alpha}$ and initial conditions $\mathbf{X}(0)$

u_0	$u_1(0)$	$v_1(0)$	$h_1(0)$	x_0	y_0	t	time step
1.129419	0.170145	0.549086	-0.040774	0.000	0.000	0.5	2.5e-05

Table 6.22: Experiment 6.3: Comparison of data assimilations for a set of distributions and errors with 4 measurements, start (0,0)

Distribution of measurements	Number of measurements	RMS Error of observations		κ	
		Iteration		Iteration	
		1^{st}	3^{rd}	1^{st}	3^{rd}
START	0.005	0.0359	0.0275	5.3e+19	5.0e+19
	0.007	0.0691	0.0426	1.7e+20	1.4e+20
	0.010	0.0357	0.0281	2.5e+19	1.0e+20
	0.013	0.0169	0.0163	2.6e+20	1.1e+20
	0.015	0.0143	0.0143	1.4e+20	5.7e+19
	0.018	0.0469	0.0360	6.6e+20	3.7e+19
	0.020	0.0144	0.0143	8.3e+20	7.1e+20
MIDDLE	0.005	0.1249	0.0568	2.6e+21	2.7e+19
	0.007	0.1776	0.0482	1.3e+19	1.6e+19
	0.010	0.0890	0.0292	1.7e+18	3.5e+18
	0.013	0.0448	0.0164	9.4e+18	1.1e+19
	0.015	0.0367	0.0115	1.5e+19	3.9e+18
	0.018	0.0706	0.0347	5.8e+18	8.5e+19
	0.020	0.0593	0.0406	5.2e+19	2.2e+19
FINISH	0.005	0.7448	0.5471	2.9e+18	6.3e+17
	0.007	0.3430	0.0838	8.9e+18	1.1e+18
	0.010	0.7248	0.5323	9.6e+18	1.6e+18
	0.013	0.6104	0.4216	1.8e+20	9.1e+18
	0.015	0.4433	0.0225	1.8e+21	1.4e+19
	0.018	0.5804	0.3634	1.5e+18	4.0e+18
	0.020	0.5687	0.0187	1.0e+19	2.8e+19
UNIFORM	0.005	0.2856	0.1408	4.8e+20	1.4e+21
	0.007	0.1190	0.1053	2.6e+20	5.0e+19
	0.010	0.2664	0.0503	1.1e+20	7.9e+20
	0.013	0.1738	0.0402	7.9e+20	4.7e+22
	0.015	0.0559	0.0277	3.4e+21	4.5e+22
	0.018	0.1651	0.0862	1.3e+21	1.2e+20
	0.020	0.1081	0.0178	1.8e+21	2.3e+21

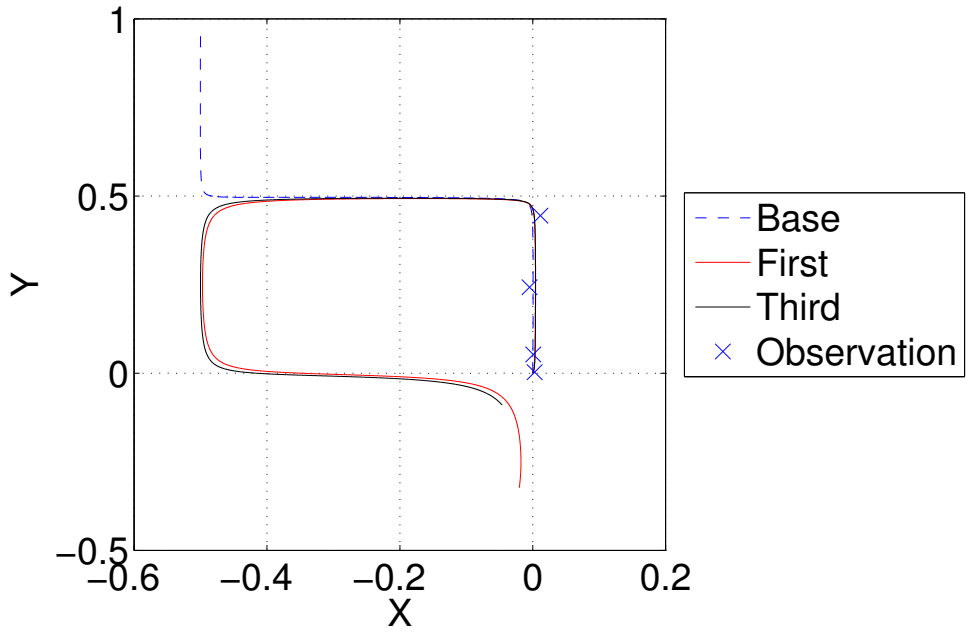


Figure 6.44: Experiment 6.3: Track START $\sigma^2 = 0.005$

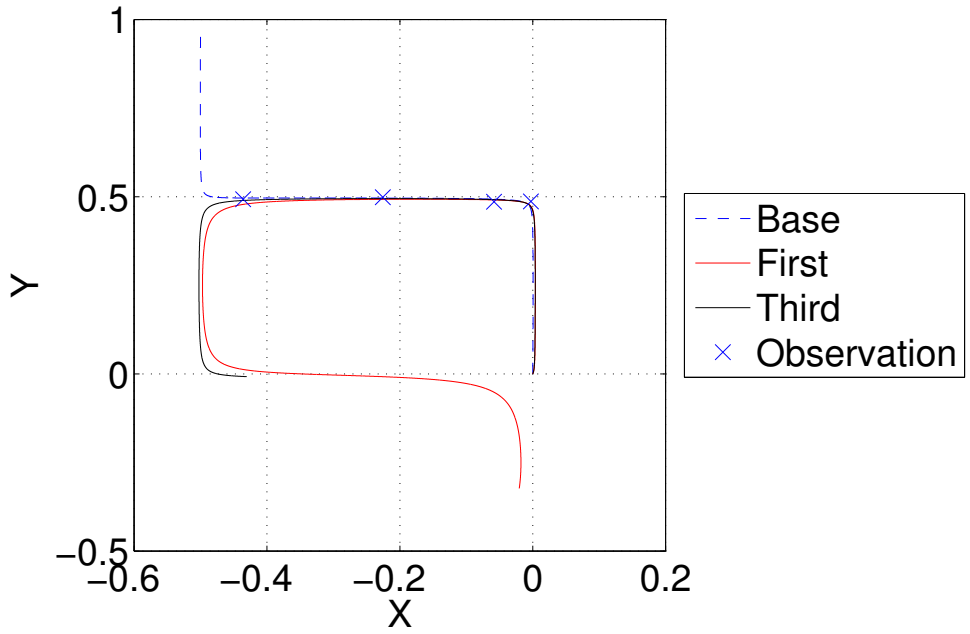


Figure 6.45: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.005$

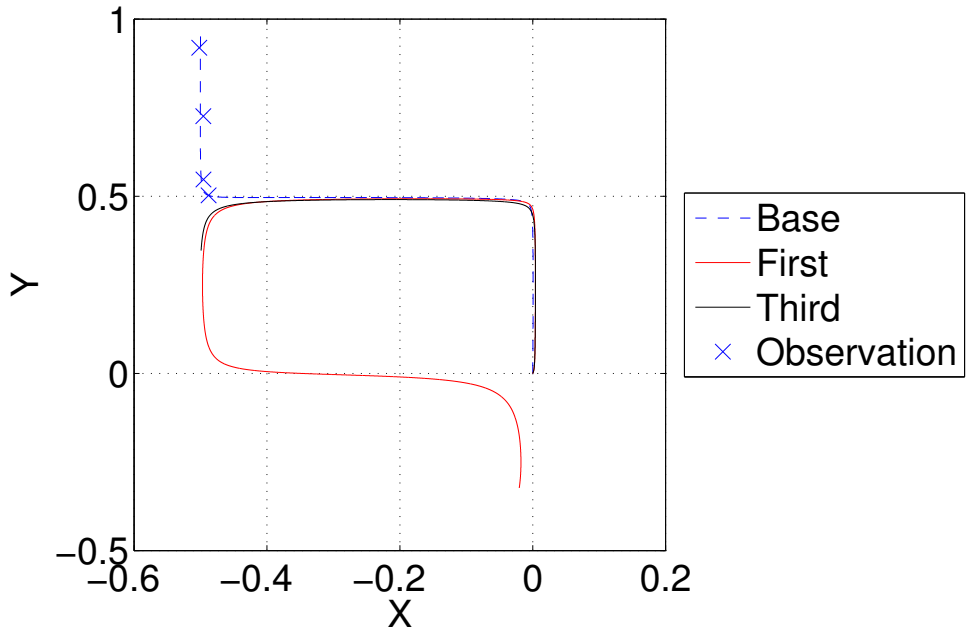


Figure 6.46: Experiment 6.3: Track FINISH $\sigma^2 = 0.005$

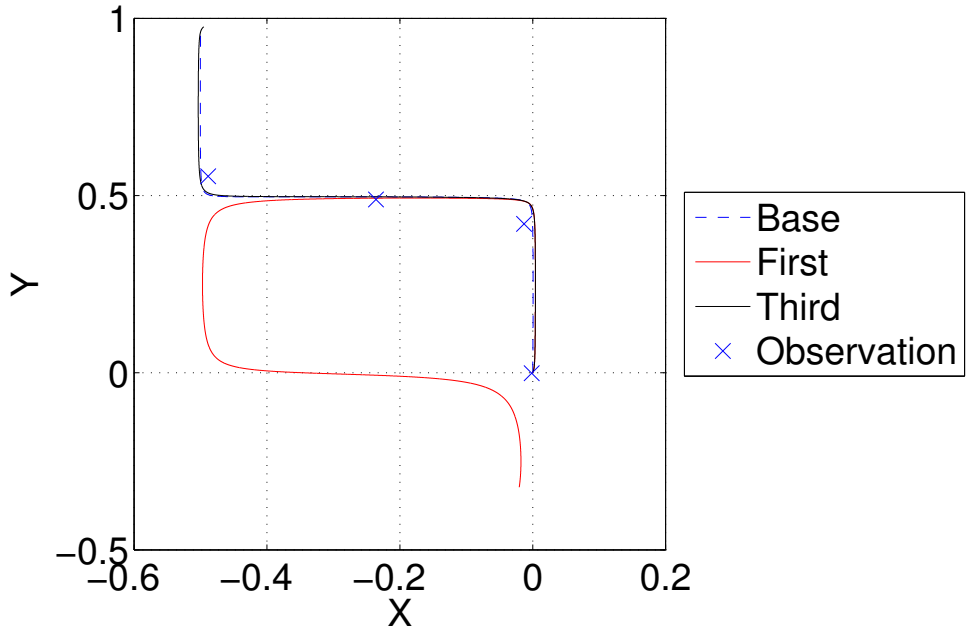


Figure 6.47: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.005$

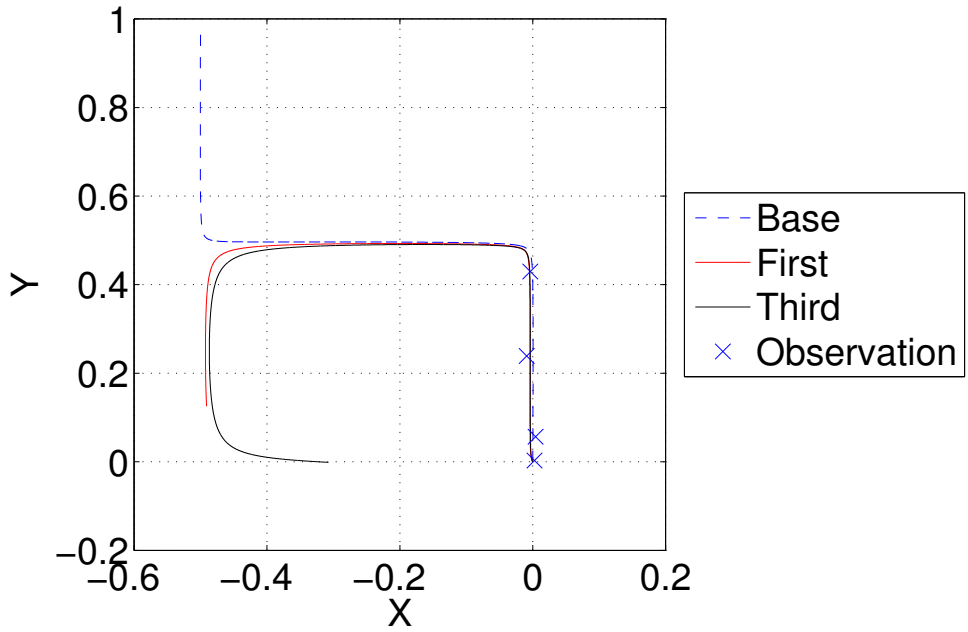


Figure 6.48: Experiment 6.3: Track START $\sigma^2 = 0.0075$

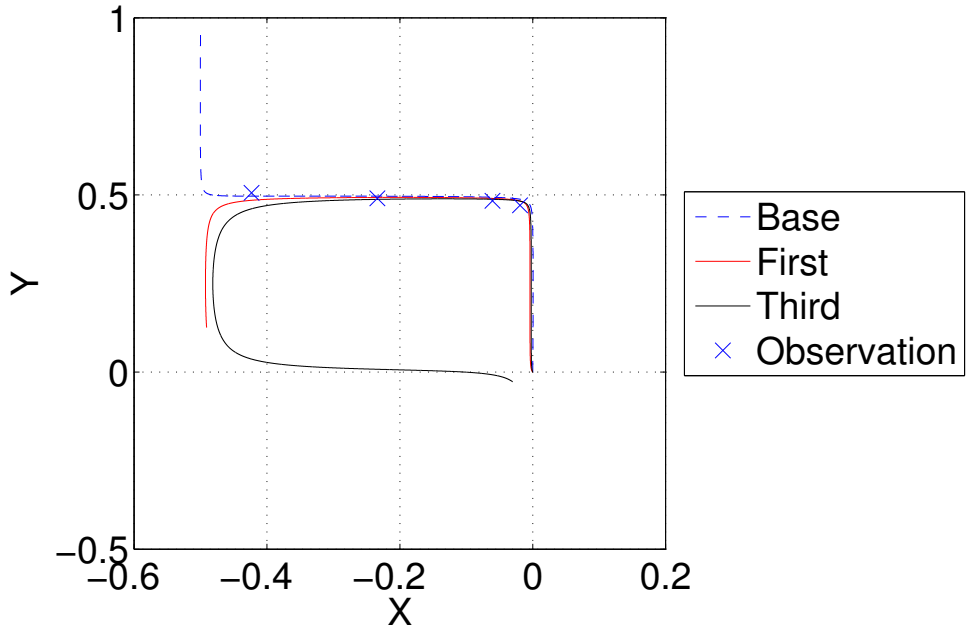


Figure 6.49: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.0075$

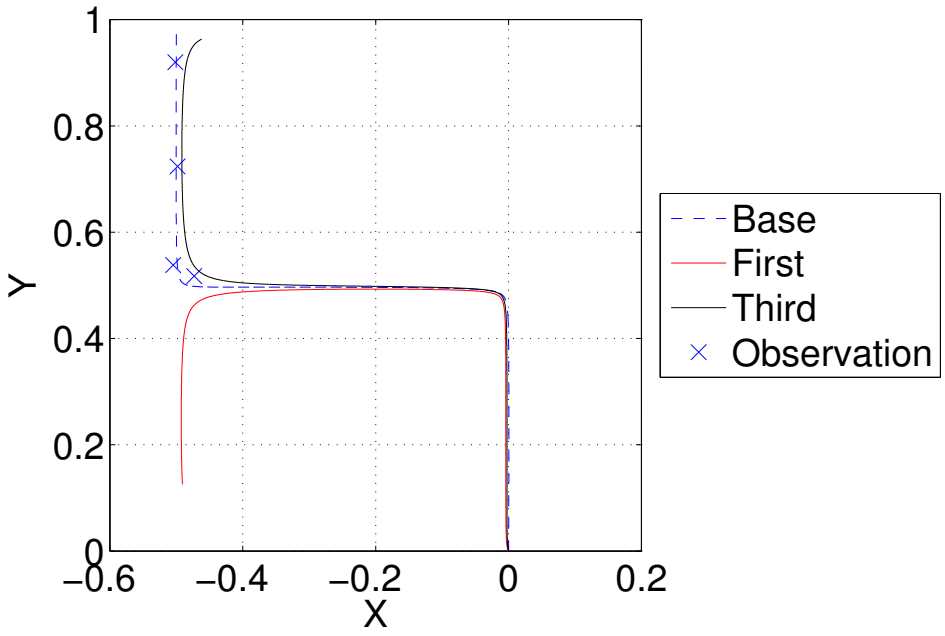


Figure 6.50: Experiment 6.3: Track FINISH $\sigma^2 = 0.0075$

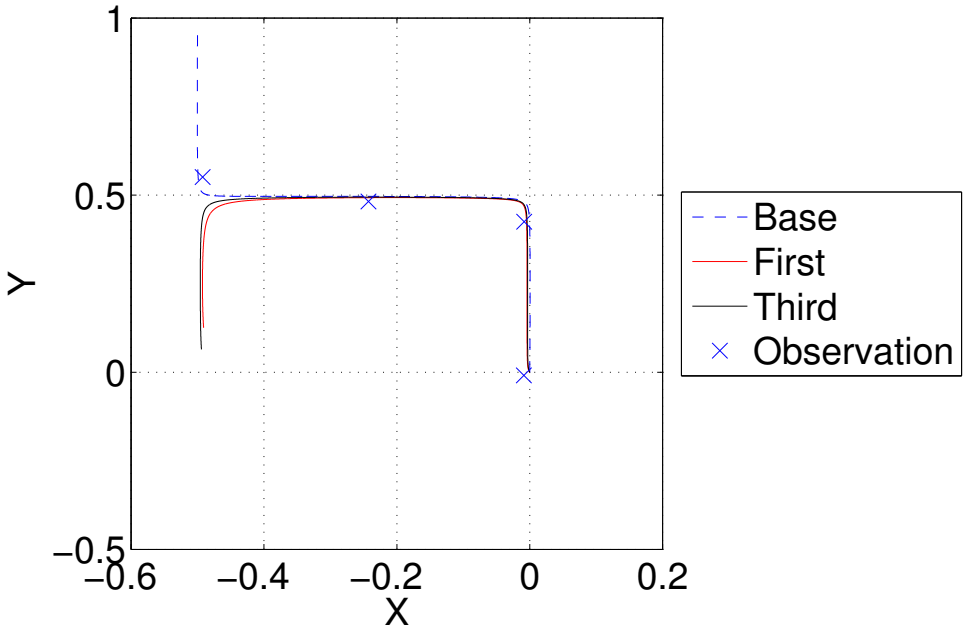


Figure 6.51: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.0075$

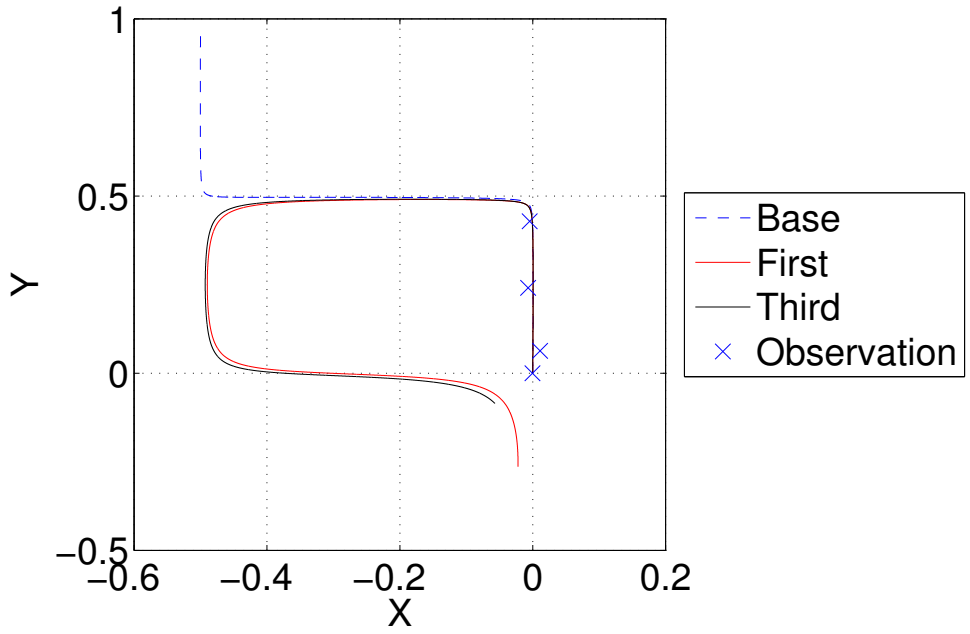


Figure 6.52: Experiment 6.3: Track START $\sigma^2 = 0.01$

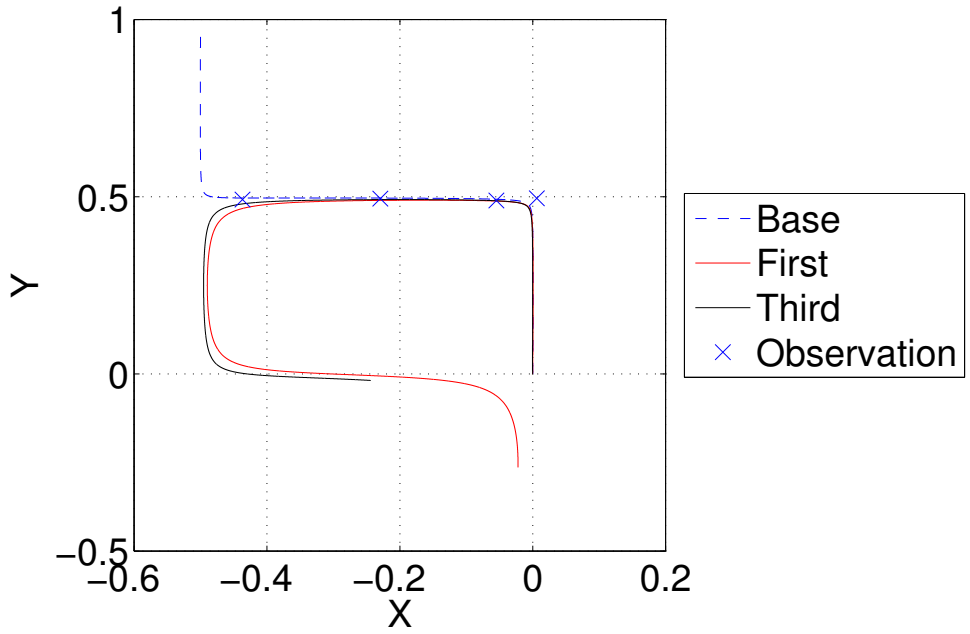


Figure 6.53: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.01$

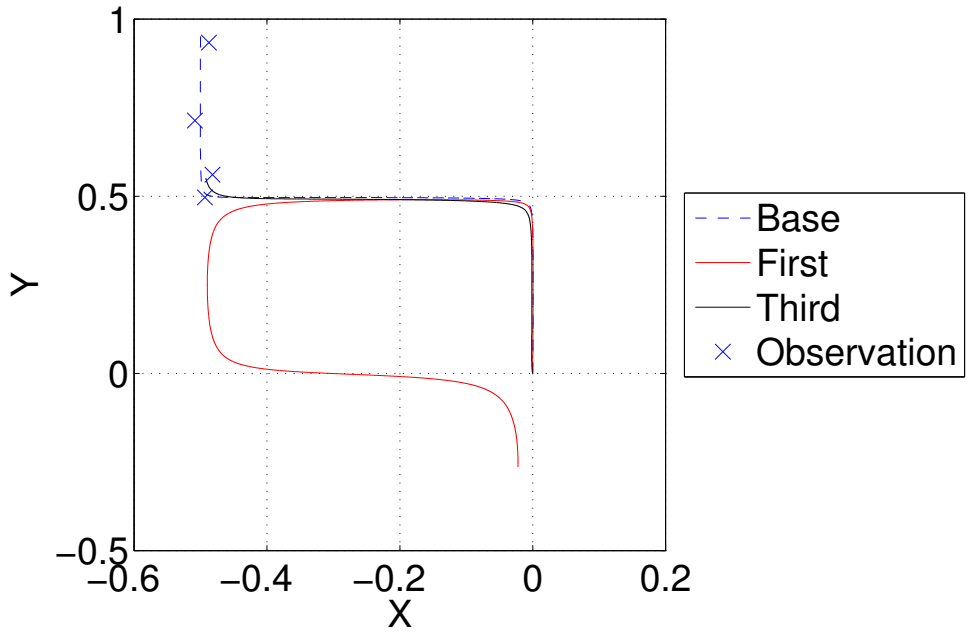


Figure 6.54: Experiment 6.3: Track FINISH $\sigma^2 = 0.01$

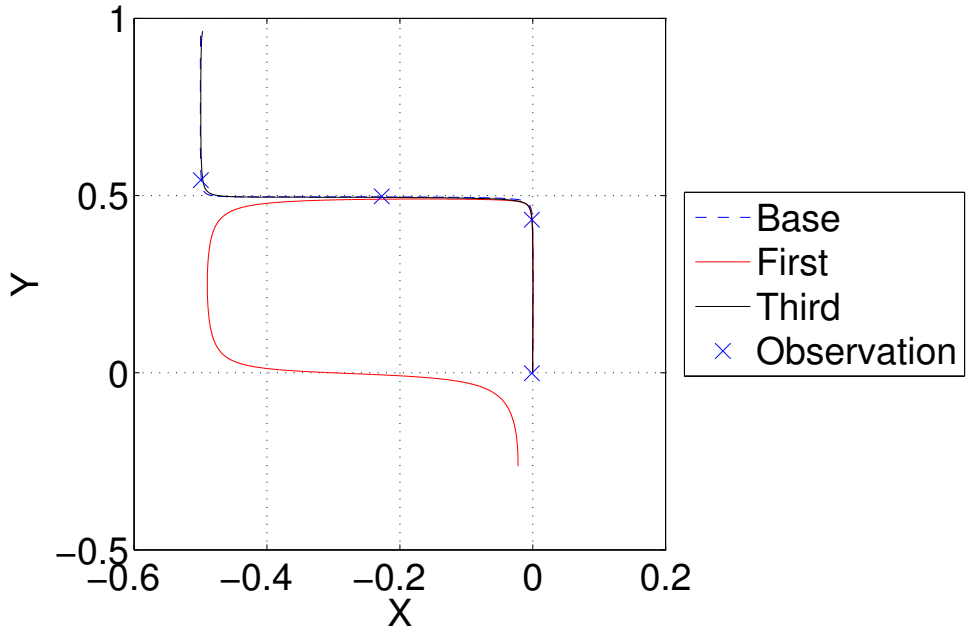


Figure 6.55: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.01$

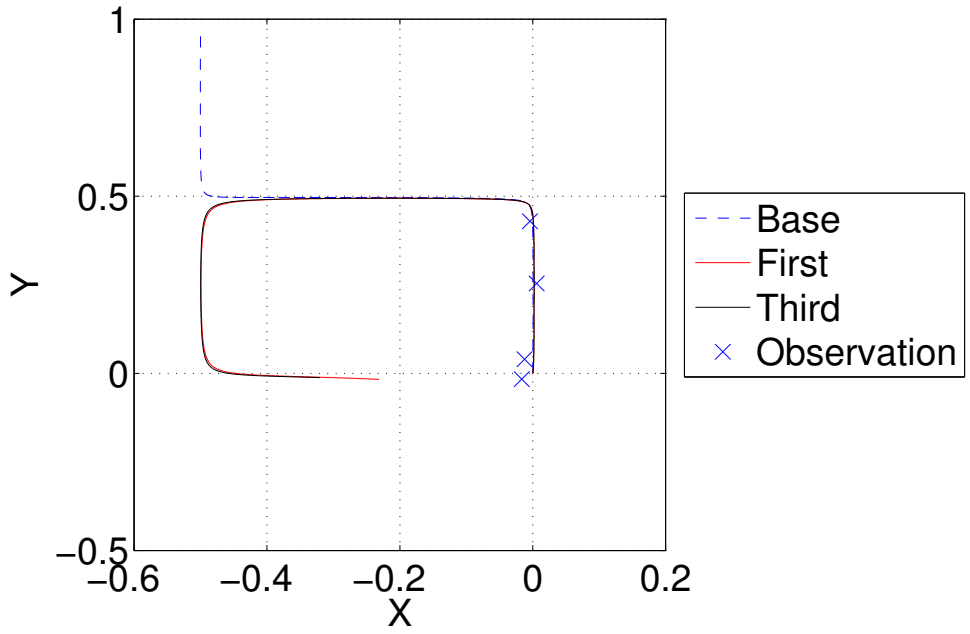


Figure 6.56: Experiment 6.3: Track START $\sigma^2 = 0.0125$

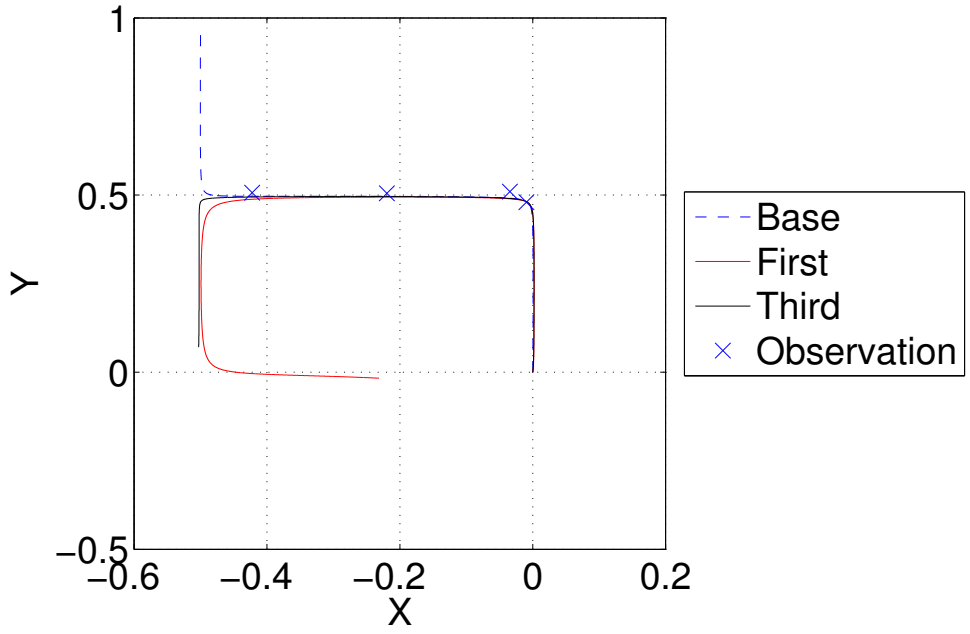


Figure 6.57: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.0125$

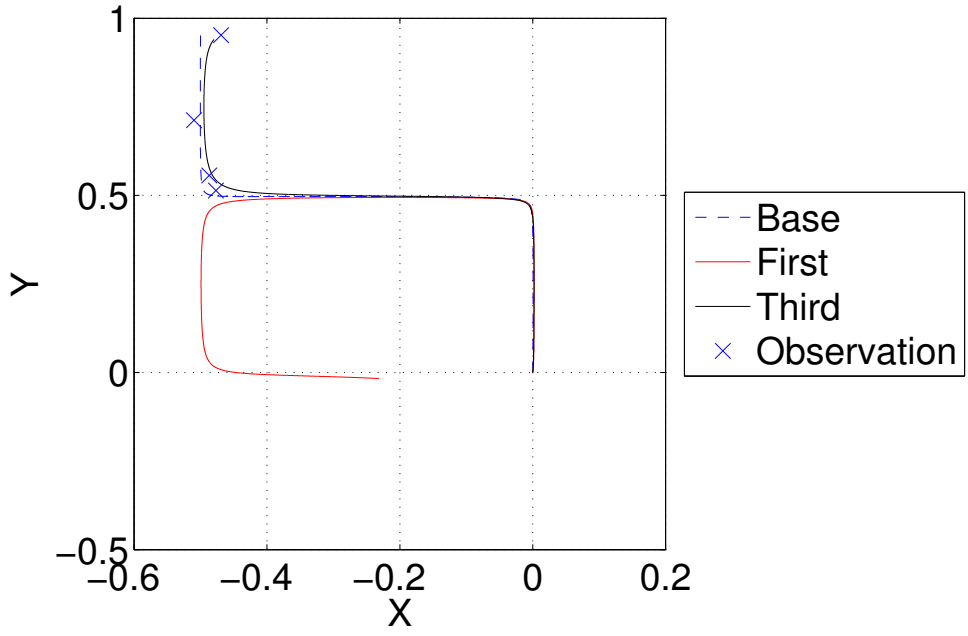


Figure 6.58: Experiment 6.3: Track FINISH $\sigma^2 = 0.0125$

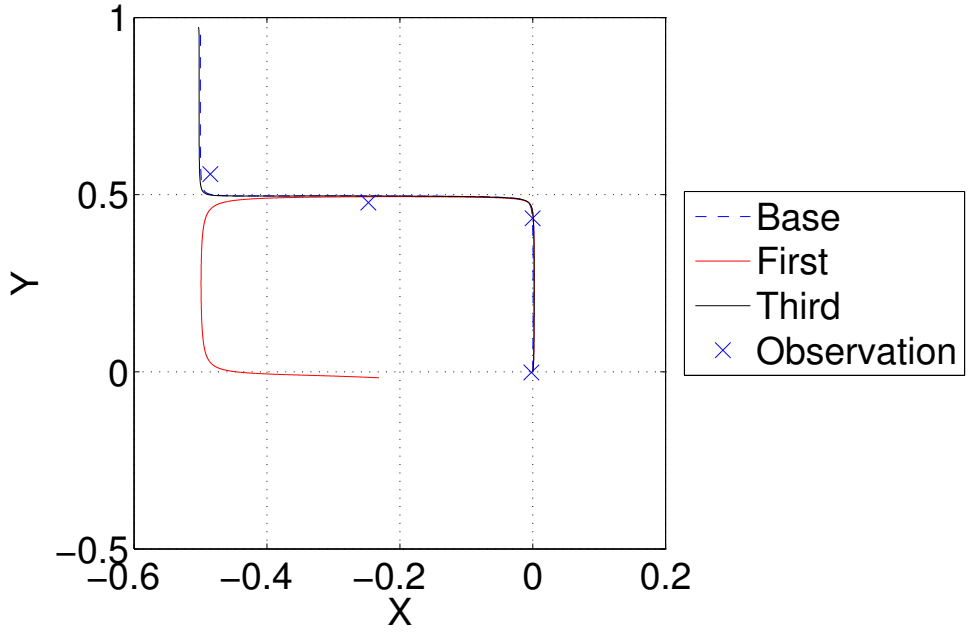


Figure 6.59: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.0125$

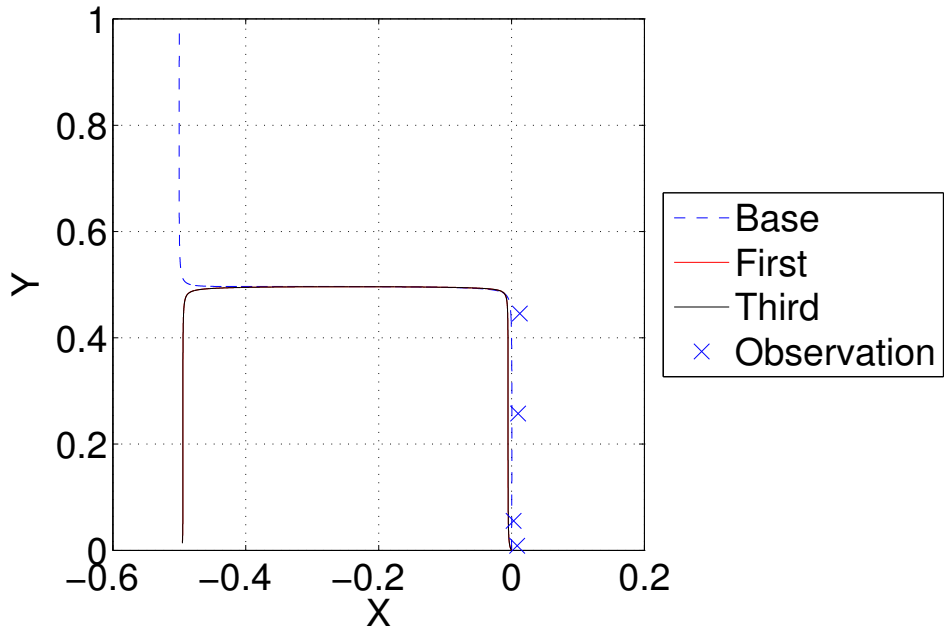


Figure 6.60: Experiment 6.3: Track START $\sigma^2 = 0.015$

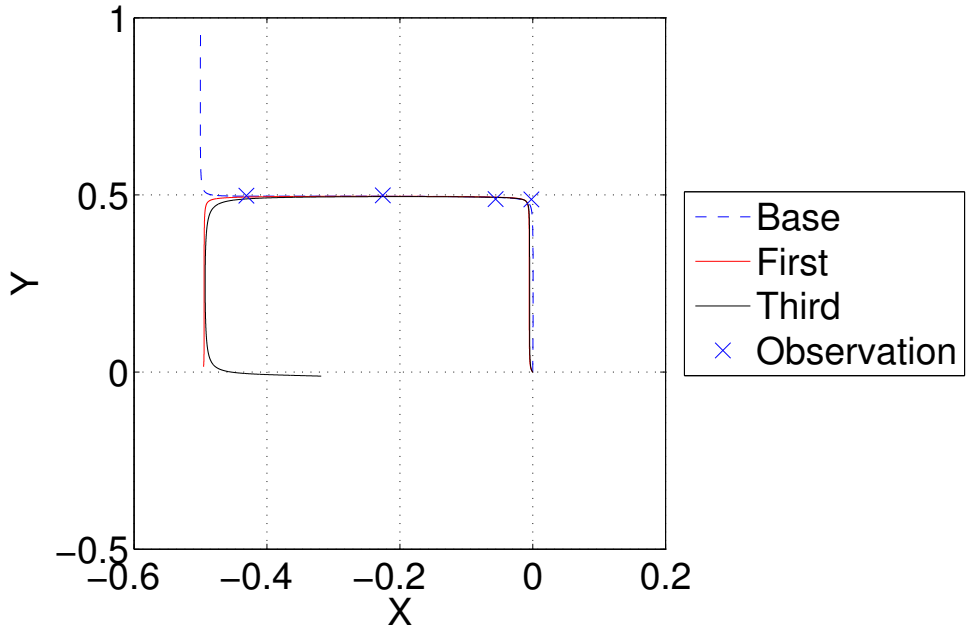


Figure 6.61: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.015$

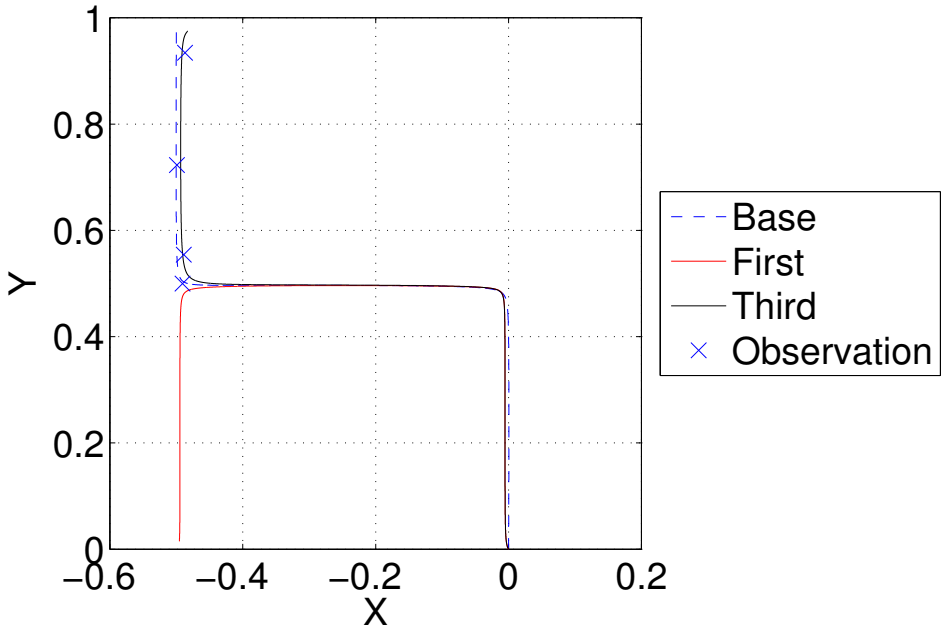


Figure 6.62: Experiment 6.3: Track FINISH $\sigma^2 = 0.015$

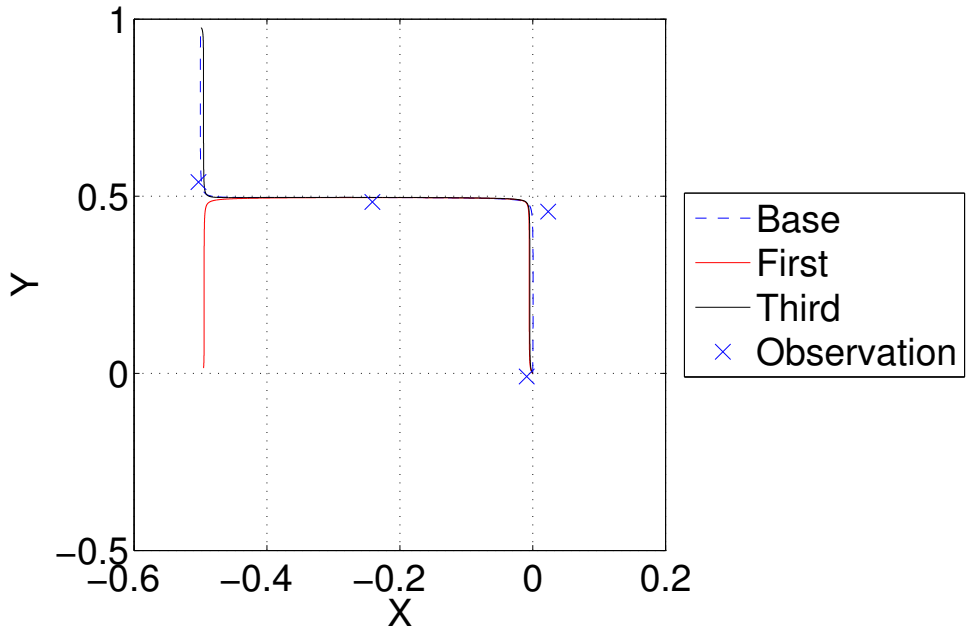


Figure 6.63: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.015$

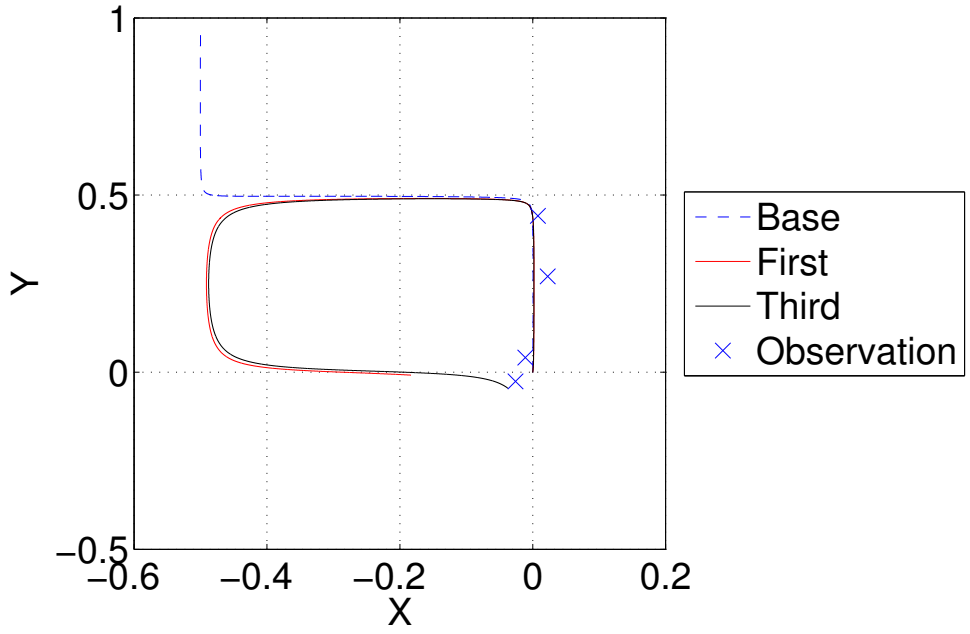


Figure 6.64: Experiment 6.3: Track START $\sigma^2 = 0.0175$

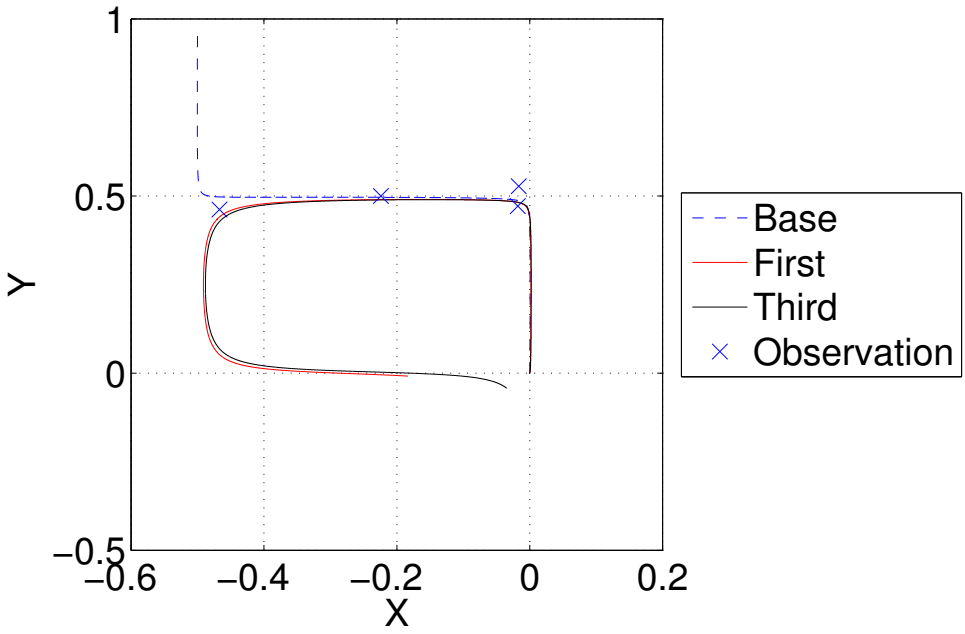


Figure 6.65: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.0175$

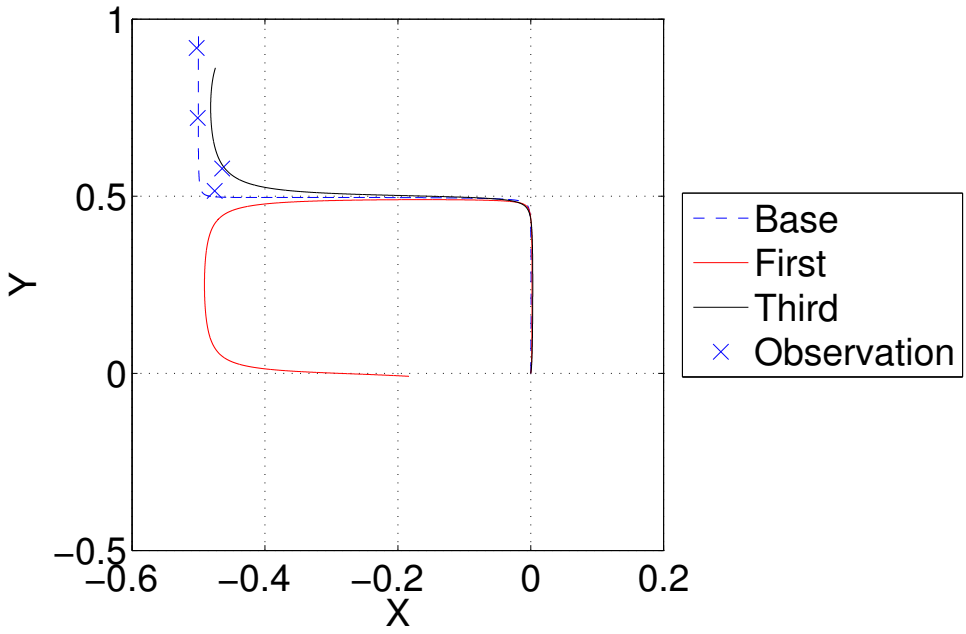


Figure 6.66: Experiment 6.3: Track FINISH $\sigma^2 = 0.0175$

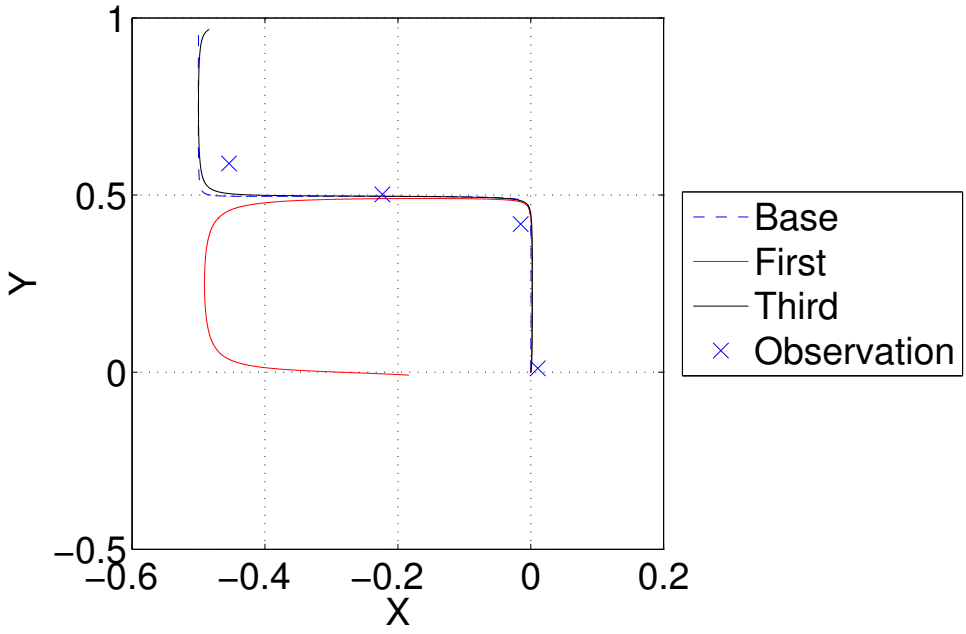


Figure 6.67: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.0175$

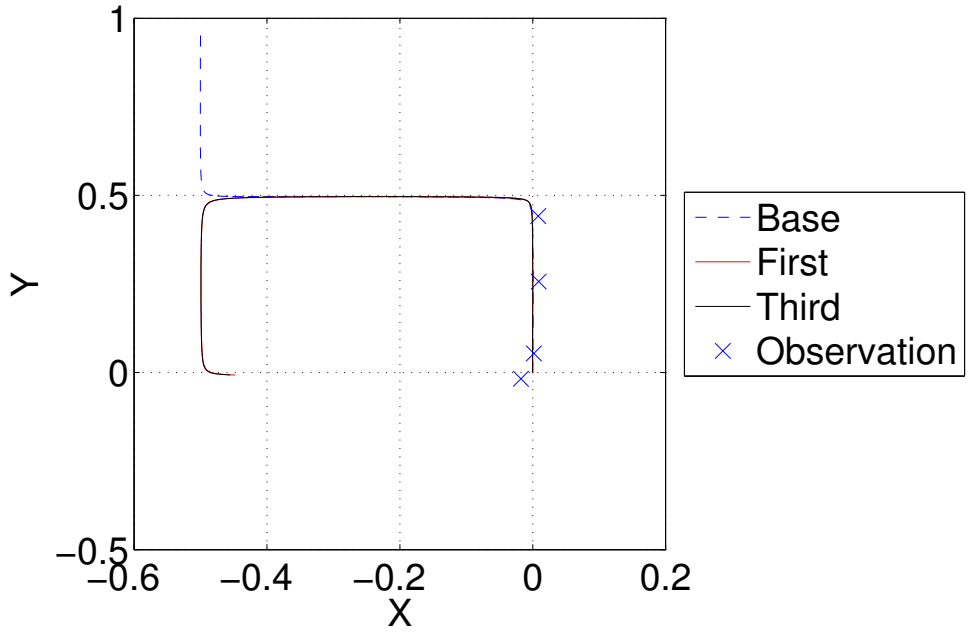


Figure 6.68: Experiment 6.3: Track START $\sigma^2 = 0.02$

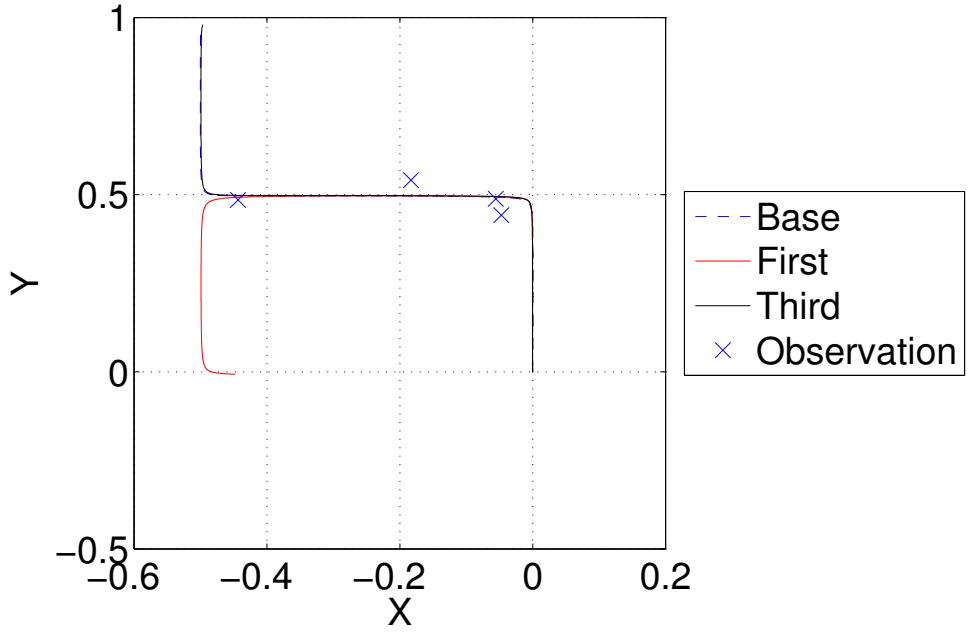


Figure 6.69: Experiment 6.3: Track MIDDLE $\sigma^2 = 0.02$

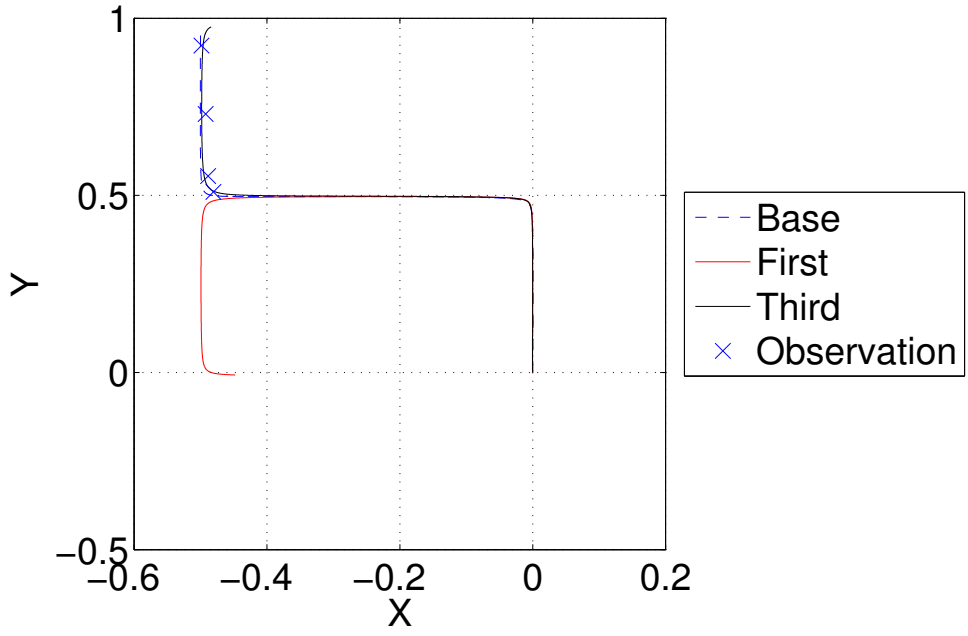


Figure 6.70: Experiment 6.3: Track FINISH $\sigma^2 = 0.02$

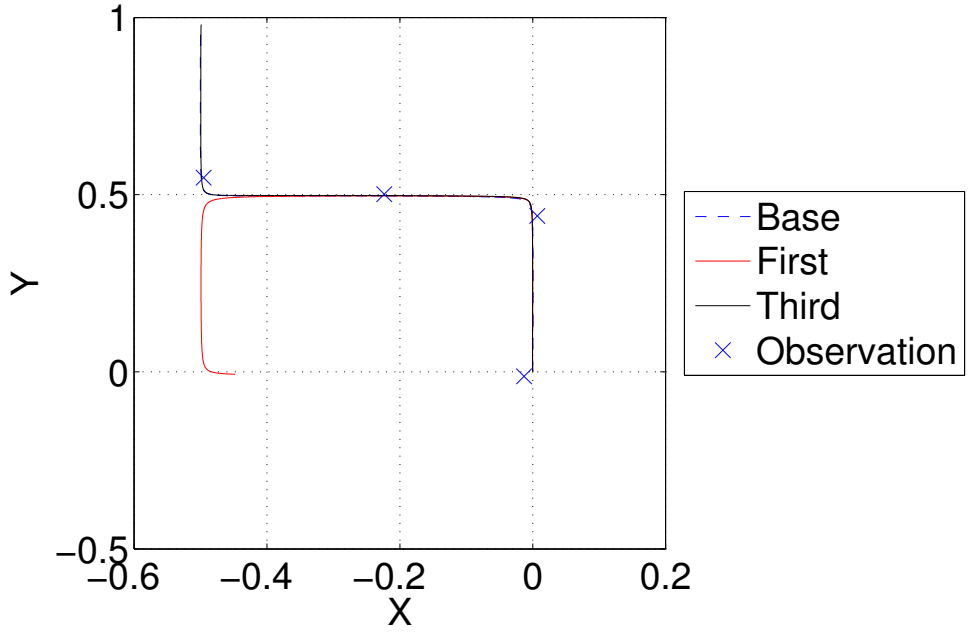


Figure 6.71: Experiment 6.3: Track UNIFORM $\sigma^2 = 0.02$

6.3 Summary

Experiments 6.1 and 6.2 using (4, 8, 16, 32, 64 and 128) measurements distributed in four distinct patterns: START, MIDDLE, FINISH and UNIFORM prove effectiveness of the forward sensitivity method applied to data assimilation of the linear shallow water model tracers. In almost all cases, after third iteration improving match between observations and model prediction, root-square mean error is reduced. Method proves effective in a presence of wide selection of the observational error variance used to generate observations. Visual inspection of the forecast error shows its close relation to model sensitivity and distribution of observations (START, MIDDLE, FINISH and UNIFORM). In cases where observations are taken in a period when sensitivity is higher, overall forecast error is smaller when compared to cases when sensitivity is lower. For cases when trajectories are well within the cell, condition number κ is in order of 10^4 to 10^5. It is noted that increasing the number of observations does not necessarily reduce the root-square mean error by much; this fact depends however on the observational error variance σ.

Experiments are indicating that even when the condition number $\boldsymbol{\kappa}$ is of order of 10^{20}, there can be an improvement made to the control vector, if the location of measurements is in the time period when the model has high sensitivity to the control vector, as shown in Experiment 6.3.

Results illustrating the model sensitivity to initial conditions are the elements of the control vector together with the data assimilation experiments described by Jabrzemski and Lakshmivarahan (2014) [6] were presented at the American Meteorological Society Annual Meeting.

Chapter 7

Conclusions

Our work in this dissertation is centered on the linearized shallow water model used by other researchers to evaluate Lagrangian data assimilation. It is divided into two major parts.

First one, focuses on finding the closed form solution to the linearized shallow water model. This in turn, allows us to provide a complete characterization of the properties of equilibria of the tracer dynamics given by a system of two first order nonlinear time varying ordinary differential equations. This dynamics is derived from a low order or reduced order spectral model obtained using Galerkin type projection of the linearized shallow water equations, while the tracer dynamics is controlled by four parameters, we provide a succinct characterization of the bifurcation using a system of hyperbolas in two dimensions. The shallow water model is thoroughly analyzed after looking separately at the geostrophic mode, the inertia-gravity mode, and both of them combined. We find that most applications of the shallow water model in meteorology and oceanography use the model within region 2, and we don't have to be concerned with the bifurcations in the control space.

Second part focuses on application of the forward sensitivity method to a problem that is historically done in meteorology and oceanography with the use of 4D-Var method. Several experiments investigated sensitivity of the LSWM to initial conditions and control vector elements. They gave a good illustration of the fact that the sensitivity has to be considered on the case by case basis. Therefore, model sensitivity gives a great insight into the model dynamics and its relation to the forecasting error. In this respect, FSM provides information about how each element of the control vector influences the solution; it shows the relation of the magnitudes of sensitivity together with their temporal distribution. This in turn helps to guide the placement of observations, or determines the impact of the observations on the forecast improvement.

Number of numerical experiments conducted with the use of FSM to assimilate data explains possible difficulties with the judicious placement of the observations by highlighting the fine physical structure of the model; there is a big value in using the closed form solution derived in the first part of this dissertation in accomplishing this second goal. Influence of the temporal distribution of the observations can be explained by evaluation of the sensitivity functions that are also derived in the closed form. There is a definite relation between a distribution of the observations and forecast error. Data assimilation of observations taken in the temporal domain related to the high sensitivity has smaller error forecast and better match of forecasted and measured tracer positions. It is worth noting that increasing the number of observations does not always lead to the improvement of the forecast error as a difference between the observations and predicted values.

Future work with the use of FSM and LSWM should focus on analyzing possible improvement to the initial condition estimation subject to the initial incorrect tracer position. Next, the forward sensitivity method can be used to

improve the initial conditions and the elements of the control vector, when both of them are incorrect. We have investigated only the measurements related to the horizontal position of a tracer, data assimilation of the height measurements can also be of interest. Data assimilation experiments can investigate FSM usefulness in analyzing data assimilation in the model while the small changes in the control vector can lead to bifurcations of the model.

Bibliography

[1] Apte, A., Jones, C. K., and Stuart, A. A bayesian approach to lagrangian data assimilation. *Tellus A 60*, 2 (2008), 336–347.

[2] Carter, E. F. Assimilation of lagrangian data into a numerical model. *Dynamics of atmospheres and oceans 13*, 3 (1989), 335–348.

[3] Daley, R. *Atmospheric data analysis*. Cambridge university press, 1991.

[4] Hirsch, M. W., Smale, S., and Devaney, R. L. *Differential equations, dynamical systems and an introduction to chaos*, vol. 60. Access Online via Elsevier, 2004.

[5] Jabrzemski, R., and Lakshmivarahan, S. Lagrangian data assimilation: Part I analysis of bifurcations in tracer dynamics. In *Proc. 17th Conference on Integrated Observing and Assimilation Systems for the Atmosphere, Oceans, and Land Surface (IOAS-AOLS)* (2013), American Meteorological Society.

[6] Jabrzemski, R., and Lakshmivarahan, S. Lagrangian data assimilation: Part II data assimilation with the forward sensitivity method. In *Proc. 18th Conference on Integrated Observing and Assimilation Systems for*

the Atmosphere, Oceans, and Land Surface (IOAS-AOLS) (2014), American Meteorological Society.

[7] Kuznetsov, L., Ide, K., and Jones, C. A method for assimilation of lagrangian data. *Monthly Weather Review 131*, 10 (2003), 2247–2260.

[8] Lakshmivarahan, S., Baldwin, M. E., and Zheng, T. Further analysis of lorenz's maximum simplification equations. *Journal of the atmospheric sciences 63*, 11 (2006), 2673–2699.

[9] Lakshmivarahan, S., and Lewis, J. Forecast bias correction using forward sensitivity analysis: A framework. Tech. rep., School of Computer Science, University of Oklahoma, 2009.

[10] Lakshmivarahan, S., and Lewis, J. Forward sensitivity approach to dynamic data assimilation. *Advances in Meteorology 2010* (2010).

[11] Lakshmivarahan, S., and Lewis, J. M. Impact of observations on adjoint sensitivity. Septermber 2011.

[12] Lewis, J. M., Lakshmivarahan, S., and Dhall, S. *Dynamic data assimilation: a least squares approach*, vol. 13. Cambridge University Press, 2006.

[13] Lorenz, E. N. Maximum simplification of the dynamic equations. *Tellus 12*, 3 (1960), 243–254.

[14] Molcard, A., Piterbarg, L. I., Griffa, A., Özgökmen, T. M., and Mariano, A. J. Assimilation of drifter observations for the reconstruction of the eulerian circulation field. *Journal of geophysical research 108*, C3 (2003), 3056.

[15] Özgökmen, T. M., Molcard, A., Chin, T. M., Piterbarg, L. I., and Griffa, A. Assimilation of drifter observations in primitive equation models of midlatitude ocean circulation. *Journal of geophysical research 108*, C7 (2003), 3238.

[16] Pedlosky, J. Geophysical fluid dynamics. *New York and Berlin, Springer-Verlag, 1982. 636 p. 1* (1982).

[17] Rabitz, H., Kramer, M., and Dacol, D. Sensitivity analysis in chemical kinetics. *Annual review of physical chemistry 34*, 1 (1983), 419–461.

[18] Salman, H., Ide, K., and Jones, C. K. Using flow geometry for drifter deployment in lagrangian data assimilation. *Tellus A 60*, 2 (2008), 321–335.

Appendix A

Reduction process

We start with the linearized shallow water model equations (2.19):

$$\begin{aligned} \frac{\partial u}{\partial t} &= v - \frac{\partial h}{\partial x}, && \text{(A.1)} \\ \frac{\partial v}{\partial t} &= -u - \frac{\partial h}{\partial y}, && \text{(A.2)} \\ \frac{\partial h}{\partial t} &= -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}. && \text{(A.3)} \end{aligned}$$

We express u, v and h in the standard two dimensional truncated Fourier series consisting of only two terms given (3.1)

$$\begin{aligned} u(x,y,t) &= -2\pi l \quad \sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t), \\ v(x,y,t) &= +2\pi k \quad \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t), \\ h(x,y,t) &= \qquad\quad \sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t). \end{aligned} \tag{A.4}$$

Let us take (A.1) and substitute corresponding parts of (A.4):

$$\begin{aligned}
L.H.S &= \frac{\partial u(x,y,t)}{\partial t} \\
&= \frac{\partial(-2\pi l \sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t))}{\partial t} \\
&= \cos(2\pi my)\frac{\partial u_1(t)}{\partial t}, \\
R.H.S &= v - \frac{\partial h}{\partial x} \\
&= 2\pi k \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t) \\
&- \frac{\partial\big(\sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t)\big)}{\partial x} \\
&= 2\pi k \cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t) \\
&- 2\pi k \cos(2\pi kx)\sin(2\pi ly)u_0 \\
&= \cos(2\pi my)v_1(t).
\end{aligned}$$

Now we can combine it

$$L.H.S = \cos(2\pi my)\frac{\partial u_1(t)}{\partial t} = \cos(2\pi my)v_1(t) = R.H.S,$$

and simplify to get

$$\frac{\partial u_1(t)}{\partial t} = v_1(t).$$

Similarly, we can look at (A.2) and substitute corresponding parts of (A.4):

$$
\begin{aligned}
L.H.S &= \frac{\partial v(x,y,t)}{\partial t} \\
&= \frac{\partial\big(2\pi k\cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t)\big)}{\partial t} \\
&= \cos(2\pi my)\frac{\partial v_1(t)}{\partial t}, \\
R.H.S &= -u - \frac{\partial h}{\partial y} \\
&= 2\pi l\sin(2\pi kx)\cos(2\pi ly)u_0 - \cos(2\pi my)u_1(t) \\
&- \frac{\partial\big(\sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t)\big)}{\partial y} \\
&= 2\pi l\sin(2\pi kx)\cos(2\pi ly)u_0 - \cos(2\pi my)u_1(t) \\
&- 2\pi l\sin(2\pi kx)\cos(2\pi ly)u_0 \\
&- 2\pi m\cos(2\pi my)h_1(t) \\
&= -\cos(2\pi my)u_1(t) - 2\pi m\cos(2\pi my)h_1(t).
\end{aligned}
$$

Now we can combine it

$$
L.H.S = \cos(2\pi my)\frac{\partial v_1(t)}{\partial t} = -\cos(2\pi my)u_1(t) - 2\pi m\cos(2\pi my)h_1(t) = R.H.S,
$$

and simplify to get

$$
\frac{\partial v_1(t)}{\partial t} = -u_1(t) - 2\pi m h_1(t).
$$

Finally, we can look at (A.3) and substitute corresponding parts of (A.4):

$$\begin{aligned}
L.H.S &= \frac{\partial h(x,y,t)}{\partial t} = \frac{\partial\big(\sin(2\pi kx)\sin(2\pi ly)u_0 + \sin(2\pi my)h_1(t)\big)}{\partial t} \\
&= \sin(2\pi my)\frac{\partial h_1(t)}{\partial t}, \\
R.H.S &= -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = -\frac{\partial\big(-2\pi l\sin(2\pi kx)\cos(2\pi ly)u_0 + \cos(2\pi my)u_1(t)\big)}{\partial x} \\
&\quad - \frac{\partial\big(2\pi k\cos(2\pi kx)\sin(2\pi ly)u_0 + \cos(2\pi my)v_1(t)\big)}{\partial y} \\
&= 2\pi l 2\pi k\cos(2\pi kx)\cos(2\pi ly)u_0 + \\
&\quad - 2\pi k 2\pi l\cos(2\pi kx)\cos(2\pi ly)u_0 + 2\pi m\sin(2\pi my)v_1(t) \\
&= 2\pi m\sin(2\pi my)v_1(t).
\end{aligned}$$

Now we can combine it

$$L.H.S = \sin(2\pi my)\frac{\partial h_1(t)}{\partial t} = 2\pi m\sin(2\pi my)v_1(t) = R.H.S,$$

and simplify to get

$$\frac{\partial h_1(t)}{\partial t} = 2\pi m v_1(t).$$

Using the fact that u_0 is constant, we can get the equations describing amplitudes as follows

$$\begin{aligned}
\dot{u_0} &= 0, \\
\dot{u_1} &= v_1, \\
\dot{v_1} &= -u_1 - 2\pi m h_1, \\
\dot{h_1} &= 2\pi m v_1,
\end{aligned}$$

which is exactly the set of equations presented in (3.2).

Appendix B

Bounds on $u_1(t)$ in (2.14)

Let A and B be two real numbers with at least one of them non-zero. For $0 \leq x \leq 2\pi$, consider the function

$$f(x) = A\cos x + B\sin x. \tag{B.1}$$

The behavior of f(x) in the interval $[0, 2\pi]$ is given in the following

Property B.0.1.

(a) $f(x)$ attains its extremum values at x^* where $\tan x^* = B/A$

(b) The distribution of $\big(x^*, f(x^*)\big)$ pair for various values and signs of A and B are given in Table A.

(c) If $f(x^*)$ is a maximum / minimum, then $f(x^*+\pi)$ is a minimum / maximum.

(d) For all $x \in [0, 2\pi]$,

$$-\sqrt{A^2+B^2} = f_{min} \leq f(x) \leq f_{max} = \sqrt{A^2+B^2}. \tag{B.2}$$

Proof. Setting the derivative of $f(x)$ to zero, the claim (a) follows. Hence

$$\cos x^* = \frac{A}{\sqrt{A^2+B^2}} \text{ and } \sin x^* = \frac{B}{\sqrt{A^2+B^2}}. \tag{B.3}$$

□

Consider the case when $A > 0$ and $B > 0$ and $A > B$. For this choice of A and B, it can be verified that $0 \leq x^* \leq \pi/4$. Further, it is easy to verify that the second derivative of $f(x)$ is negative at x^* and positive at $(x^* + \pi)$. Hence $f(x^*)$ is a maximum and $f(x^* + \pi)$ is a minimum from which one of the entries in Table A and claim (c) follows. Considering these and (A.3) with (A.1), the claim (d) immediately follows

Table B.1: Distribution of the pairs $\big(x^*, f(x^*)\big)$

Ratio	A	B	x^* belongs to	$f(x) =$
$\|A\| > \|B\|$ or $0 \leq \frac{\|B\|}{\|A\|} \leq 1$	+	+	$[0,\ \pi/4]$	f_{max}
	-	-	$[0,\ \pi/4]$	f_{min}
	-	+	$[-\pi/4, 0]$	f_{min}
	+	-	$[-\pi/4, 0]$	f_{max}
$\|A\| < \|B\|$ or $\frac{\|B\|}{\|A\|} \geq 1$	+	+	$[\pi/4, \pi/2]$	f_{max}
	-	-	$[\pi/4, \pi/2]$	f_{min}
	-	+	$[-\pi/2, -\pi/4]$	f_{max}
	+	-	$[-\pi/2, -\pi/4]$	f_{min}

Appendix C

Definition and properties of standard hyperbola

Equation for a standard hyperbola with center at $x_0, y_0)$ is given by (Spigel (1968))

$$\frac{(x-x_0)^2}{a^2} - \frac{(y-y_0)^2}{b^2} = 1 \tag{C.1}$$

where a and b are called the length of the semi major and semi minor axes respectively. The graph of the hyperbola in (C.1) is given in Figure B1.

C is the center of the hyperbola whose coordinates are (x_0, y_0) with respect to the origin 0. $AA' = 2a$ is the major axis, and $BB' = 2b$ is the minor axis. F and F' are called the foci and $CF = CF' = \sqrt{a^2 + b^2}$. The eccentricity e of this hyperbola is given by $e = \frac{\sqrt{a^2+b^2}}{a} > 1$. The slope of the asymptote HG' is (b/a) and the of the asymptote $H'G$ is $-(b/a)$. The hatched region indicate where

$$\frac{(x-x_0)^2}{a^2} - \frac{(y-y_0)^2}{b^2} > 1.$$

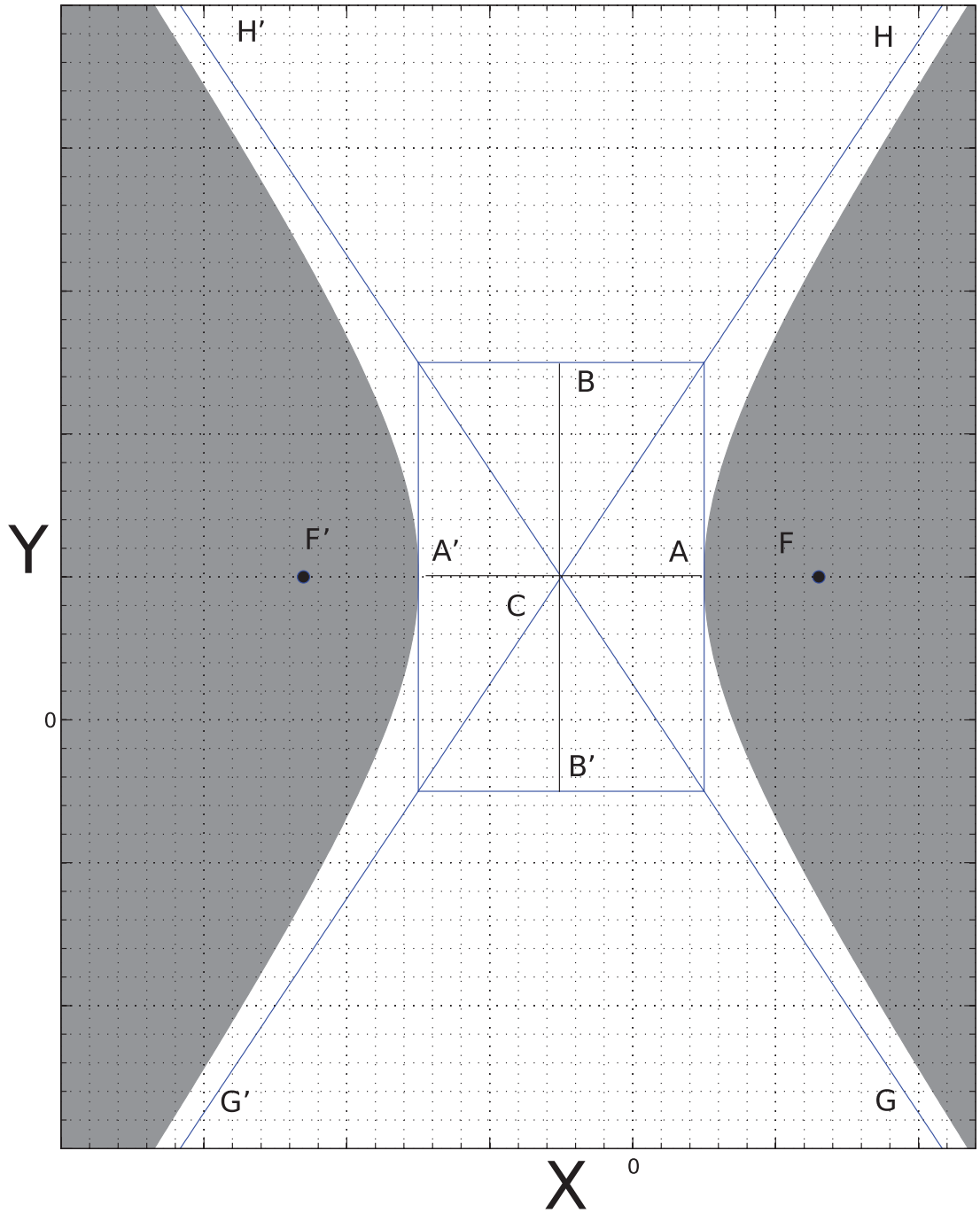


Figure B.1: An illustration of the standard hyperbola given by (C.1). $AA' = 2a$, $BB' = 2b, c$ is the center, F and F' are foci. $CF = CF' = \sqrt{a^2 + b^2}$. The eccentricity $c = \frac{\sqrt{a^2+b^2}}{a} > 1$.