A Reduced Adjoint Approach to Variational Data Assimilation

Mohamad El Gharamti

M. U. Altaf, I. Hoteit, A. W. Heemink

King Abdullah University of Science and Technology

"The Ninth International Workshop on Adjoint Model Applications in Dynamic Meteorology"

October 13, 2011

Outline

Overview & Motivation

- 4D-Var or the adjoint method is widely known and have been applied in several studies in oceanography and shallow water flows.
- Several studies used a modified version of it called model-reduced 4D-Var based on Proper Orthogonal Decomposition (POD) also known as PCA and EOF.
- Given their importance, both approaches have some issues to be discussed regarding estimation accuracy, computational effort and implementation.
- In this work, an effective variational data assimilation scheme is proposed using reduced adjoint method based on POD.

Overview & Motivation

- 4D-Var or the adjoint method is widely known and have been applied in several studies in oceanography and shallow water flows.
- Several studies used a modified version of it called model-reduced 4D-Var based on Proper Orthogonal Decomposition (POD) also known as PCA and EOF.
- Given their importance, both approaches have some issues to be discussed regarding estimation accuracy, computational effort and implementation.
- In this work, an effective variational data assimilation scheme is proposed using reduced adjoint method based on POD.

Overview & Motivation

- 4D-Var or the adjoint method is widely known and have been applied in several studies in oceanography and shallow water flows.
- Several studies used a modified version of it called model-reduced 4D-Var based on Proper Orthogonal Decomposition (POD) also known as PCA and EOF.
- Given their importance, both approaches have some issues to be discussed regarding estimation accuracy, computational effort and implementation.
- In this work, an effective variational data assimilation scheme is proposed using reduced adjoint method based on POD.

Overview & Motivation

- 4D-Var or the adjoint method is widely known and have been applied in several studies in oceanography and shallow water flows.
- Several studies used a modified version of it called model-reduced 4D-Var based on Proper Orthogonal Decomposition (POD) also known as PCA and EOF.
- Given their importance, both approaches have some issues to be discussed regarding estimation accuracy, computational effort and implementation.
- In this work, an effective variational data assimilation scheme is proposed using reduced adjoint method based on POD.

Overview & Motivation

- 4D-Var or the adjoint method is widely known and have been applied in several studies in oceanography and shallow water flows.
- Several studies used a modified version of it called model-reduced 4D-Var based on Proper Orthogonal Decomposition (POD) also known as PCA and EOF.
- Given their importance, both approaches have some issues to be discussed regarding estimation accuracy, computational effort and implementation.
- In this work, an effective variational data assimilation scheme is proposed using reduced adjoint method based on POD.

State-Space Models

Consider the following strong constrained discrete system

$$X(t_{i+1}) = M_i X(t_i), i = 1, ..., m-1 \quad (1)$$

where $X(t_{i+1})$ is the model state vector $(N \times 1)$, M_i is a non-linear operator of size $(N \times N)$ propagating the states in time.

Imperfect observations $Y(t_i)$ are related to the model states

$$Y(t_i) = H(X(t_i)) + \eta(t_i), \quad (2)$$

where H is a linear observation operator of size $(q \times N)$ and $\eta(t_i)$ is a Gaussian noise $\sim \mathcal{N}(0, R)$.

Full Adjoint Method

Objective function and its gradient:

$$
\begin{aligned}
J(X_0) &= \frac{1}{2}\sum_{i=1}^{m}\left[HX(t_i)-Y(t_i)\right]^T R_i^{-1}\left[HX(t_i)-Y(t_i)\right] + \\
&\quad \frac{1}{2}(X_0-X^b)^T B_0^{-1}(X_0-X^b), \qquad (3a)\\
\nabla J(X_0) &= B_0^{-1}(X_0-X^b)+\nu(t_0), \qquad (3b)
\end{aligned}
$$

where X^b is the background estimate of the initial condition X_0, $(X_0 - X^b)$ are usually assumed to be Gaussian with covariance B_0 and $\nu(t_0)$ is the adjoint state at the initial time.

Adjoint backward in time:

$$
\begin{aligned}
\nu(t_i) &= \mathbf{M}_i^T\nu(t_{i+1}) - H^T R_i^{-1}\left[HX(t_i)-Y(t_i)\right], \qquad (4a)\\
\nu(t_m) &= 0, \qquad (4b)
\end{aligned}
$$

where $\mathbf{M}_i$ are the Jacobians of M_i with respect to X_i.

Model-Reduced Adjoint Method

Objective function and its gradient:

$$\begin{aligned}\hat{J}(r_0) &= \frac{1}{2}\sum_{i=1}^{m}\left[H\hat{X}(t_i)-Y(t_i)\right]^T R_i^{-1}\left[H\hat{X}(t_i)-Y(t_i)\right]+ \\ &\quad \frac{1}{2}\left[P^T(X_0-X^b)\right]^T P^T B_0^{-1} P\left[P^T(X_0-X^b)\right], \qquad (5a)\\ \nabla\hat{J}(r_0) &= P^T B_0^{-1} P\left[P^T(X_0-X^b)\right]+\hat{\nu}(t_0), \qquad (5b)\end{aligned}$$

where P is space reduction operator from POD and $\hat{X}(t_i)$ is the approximate linearized state.

Adjoint backward in time:

$$\begin{aligned}\hat{\nu}(t_i) &= \widetilde{M}_i^T\hat{\nu}(t_{i+1})-\hat{H}^T R_i^{-1}\left[H\hat{X}(t_i)-Y(t_i)\right], \qquad (6a)\\ \nu(t_m) &= 0, \qquad (6b)\end{aligned}$$

where $\widetilde{M}_i$ is the simplified model dynamics.

Some Issues with these 2 Methods

1. *The Full Adjoint method:*
   - A single run of the adjoint model backward in time is expensive and is usually equivalent to several forward model simulations.
   - Programming effort and implementation.

2. *The Model-Reduced Adjoint method:*
   - Loss of accuracy because of using a simplified model.
   - Solution from the reduced model might not converge to the same global minimum when it's projected back as the optimization is performed in reduced space.

Model Reduction Tool: POD (PCA or EOF)

Given an ensemble $E_{N\times s}$ of centered model states, we construct the (inner product, if N is very large) matrix Q as follows

$$Q = E^T E. \tag{7}$$

In order to reduce the space of the system, we solve the eigenvalue problem to obtain the leading POD modes according to (9)

$$E^T E Z_i = \lambda_i Z_i, \tag{8}$$

$$\Rightarrow P = EZ\Lambda^{-1/2}, \tag{9}$$

where $Z = \{Z_1, Z_2, \ldots, Z_s\}$, $P = \{p_1, p_2, \ldots, p_s\}$ and Λ is a diagonal matrix with the eigenvalues. The relative energy of the chosen eigenvalues can be calculated as

$$\psi_i = \frac{\lambda_i}{\sum_{j=1}^{s} \lambda_j} \times 100, i = 1, 2, ..., r. \tag{10}$$

$\rightarrow$ Note that $\psi_1 > \psi_2 > ... > \psi_r$.

Reduced Adjoint Method

The idea is to use the full space in the objective function as in full adjoint method (7a) and run the adjoint model in reduced space

$$\hat{\nu}(t_i) = \widetilde{M}_i^T \hat{\nu}(t_{i+1}) - \hat{H}^T R_i^{-1}[H \overbrace{X(t_i)}^{\text{Full state}} - Y(t_i)], \tag{11}$$

where $\widetilde{M}_i$ is the reduced operator approximating the full jacobians $\mathbf{M}_i$;

$$\widetilde{M}_i = P^T(\overbrace{\frac{\partial M_i}{\partial X(t_i)} p_1}^{\alpha_1}, \ldots, \frac{\partial M_i}{\partial X(t_i)} p_r) \; ; \; \alpha_1 = \frac{M_i[X(t_i) + \varepsilon p_1] - M_i[X(t_i)]}{\varepsilon}$$

Once the adjoint state at t_0 is obtained in the reduced space, it's projected back in full space as follows

$$\nabla J(X_0) = B_0^{-1}(X_0 - X^b) + P\hat{\nu}(t_0), \tag{12}$$

the minimization process is then carried out in the full space.

Procedural Flow

Flow of the problem

1. Initial guess X_0, to generate an ensemble of forward model simulations.
2. Solve the eigenvalue problem to get the dominant eigenmodes.
3. The reduced dynamic operator $\widetilde{M}_i$ is obtained on the dominant POD modes.
4. Optimization is done using Quasi-Newton method. Each iteration includes running the original model (1) and the reduced adjoint model to get the optimal solution of the objective function J.
5. After minimization, the initial condition X_0 is updated and the process could be started again from step 4 and occasionally from step 1 (to get optimized initial conditions).

Procedural Flow

Flow of the problem

1. Initial guess X_0, to generate an ensemble of forward model simulations.
2. Solve the eigenvalue problem to get the dominant eigenmodes.
3. The reduced dynamic operator $\widetilde{M}_i$ is obtained on the dominant POD modes.
4. Optimization is done using Quasi-Newton method. Each iteration includes running the original model (1) and the reduced adjoint model to get the optimal solution of the objective function J.
5. After minimization, the initial condition X_0 is updated and the process could be started again from step 4 and occasionally from step 1 (to get optimized initial conditions).

Procedural Flow

Flow of the problem

1. Initial guess X_0, to generate an ensemble of forward model simulations.
2. Solve the eigenvalue problem to get the dominant eigenmodes.
3. The reduced dynamic operator $\widetilde{M}_i$ is obtained on the dominant POD modes.
4. Optimization is done using Quasi-Newton method. Each iteration includes running the original model (1) and the reduced adjoint model to get the optimal solution of the objective function J.
5. After minimization, the initial condition X_0 is updated and the process could be started again from step 4 and occasionally from step 1 (to get optimized initial conditions).

Procedural Flow

Flow of the problem

1. Initial guess X_0, to generate an ensemble of forward model simulations.
2. Solve the eigenvalue problem to get the dominant eigenmodes.
3. The reduced dynamic operator $\widetilde{M}_i$ is obtained on the dominant POD modes.
4. Optimization is done using Quasi-Newton method. Each iteration includes running the original model (1) and the reduced adjoint model to get the optimal solution of the objective function J.
5. After minimization, the initial condition X_0 is updated and the process could be started again from step 4 and occasionally from step 1 (to get optimized initial conditions).

Procedural Flow

Flow of the problem

1. Initial guess X_0, to generate an ensemble of forward model simulations.
2. Solve the eigenvalue problem to get the dominant eigenmodes.
3. The reduced dynamic operator $\widetilde{M}_i$ is obtained on the dominant POD modes.
4. Optimization is done using Quasi-Newton method. Each iteration includes running the original model (1) and the reduced adjoint model to get the optimal solution of the objective function J.
5. After minimization, the initial condition X_0 is updated and the process could be started again from step 4 and occasionally from step 1 (to get optimized initial conditions).

The Coupled Flow and Transport Model

A 2D flow model that solves the Darcy (13) and the continuity (14) equations is considered

$$U = -\frac{\mathbf{K}}{\mu}\nabla P \tag{13}$$

$$\frac{\partial(\phi\rho)}{\partial t} = -\nabla\cdot(\rho U) \tag{14}$$

On top of this, we add the contaminant transport equation

$$\frac{\partial(\phi\mathbf{C})}{\partial t} + \nabla\cdot(U\mathbf{C} - \phi D(U)\nabla\mathbf{C}) = \phi r(\mathbf{C}) + q\mathbf{C}^* \tag{15}$$

The upwind scheme of CCFD is utilized to descritize the PDEs in (13), (14) and (15). In modeling, we will consider only the advection part of (15) to get a linearized model.

Flow Model

(50×50) grid rectangular domain $[0, 1000] \times [0, 500]$ m. Two main rocks, where one is embedded in the other, having different permeabilities; $\mathbf{K}_1 = 100$md and $\mathbf{K}_2 = 10$md. Dirichlet B.C. with a head drop of 90m.

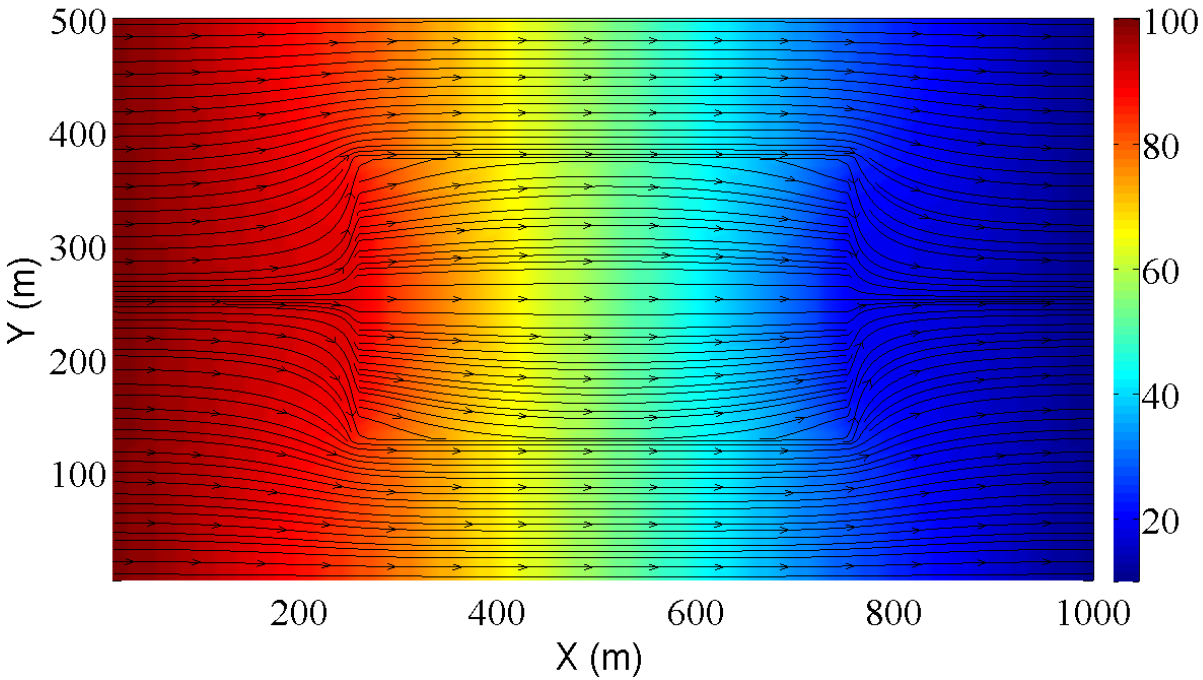


Fig. 1: Pressure head and velocity streamlines in the domain.

Contaminant Model

Initially, a contaminant plume of 100ppm is inserted close to the western boundary and 0ppm (pure water) elsewhere. The region is modeled by running the linear transport model for 50 years with a time step of 2 months.

Loading "Contaminant Transport" ...

Pseudo-Observations & Initial Guess

Pseudo-observations were extracted from 160 grid cells out of 2500 as in figure (2, right panel). The initial guess, in figure (2, left panel), for the optimization problem is taken as the average of the the simulated state vectors from a long model run.

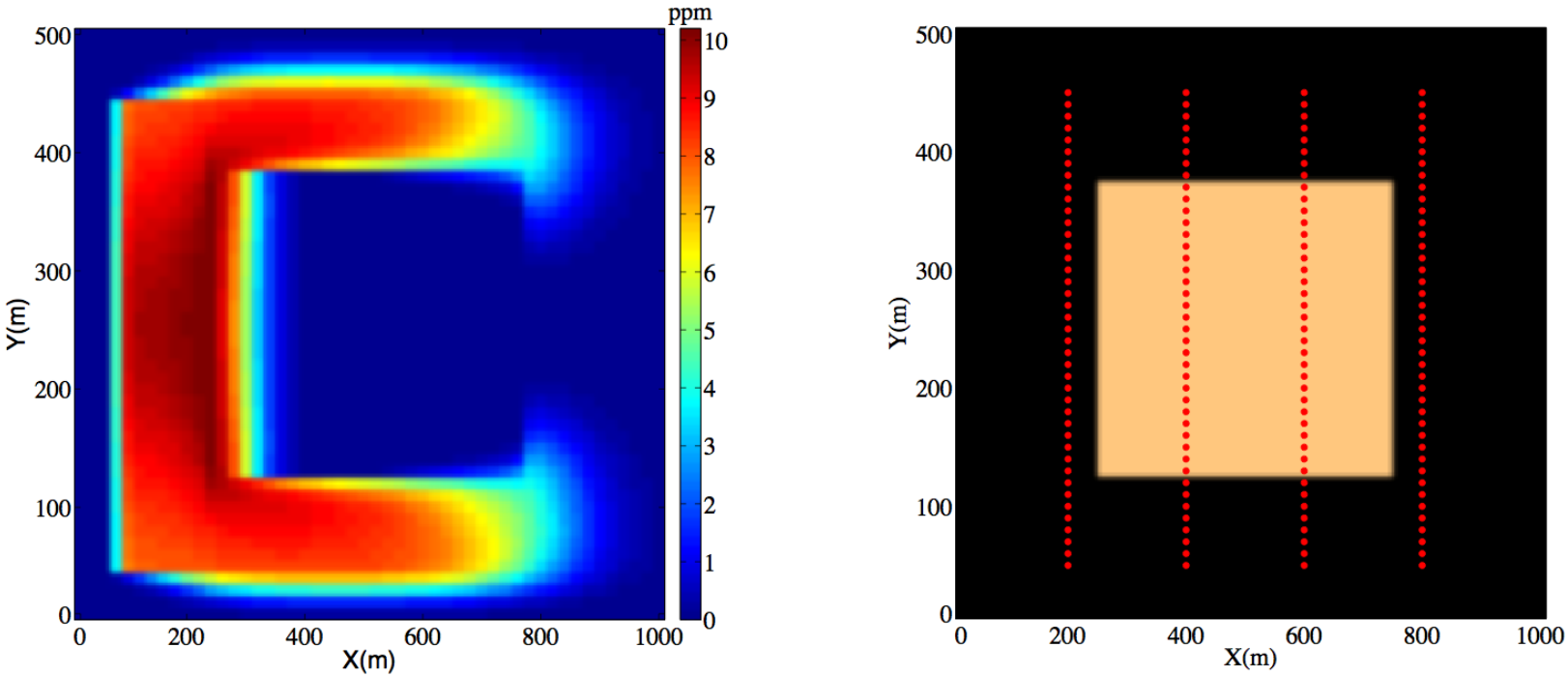


Fig. 2: System Configuration.

Experiments

1. *Experiment I*
   - Ensemble of 60 snapshots of the simulated states.
   - Perfect observations (no perturbations).
   - Observations available every 2 months, i.e. every time step.
2. *Experiment II*
   - Ensemble of 60 snapshots of the simulated states.
   - Perturbed observations; additive Gaussian noise of mean 0 and covariance of 10% of the total variance.
   - Observations available every 1 year, i.e. every 6 time steps.

Convergence Criterion:

$$\mu = \sum_{k=1} \|\nabla J_i^k\| / \|\nabla J_0^k\| \leq \underbrace{\kappa}_{\approx 10^{-4}} \tag{16}$$

Results: Dominant POD Modes

Using the proposed ensemble, a basis consisting of only 10 POD modes was collected. They captured $\sim 99.999\%$ of the relative energy.

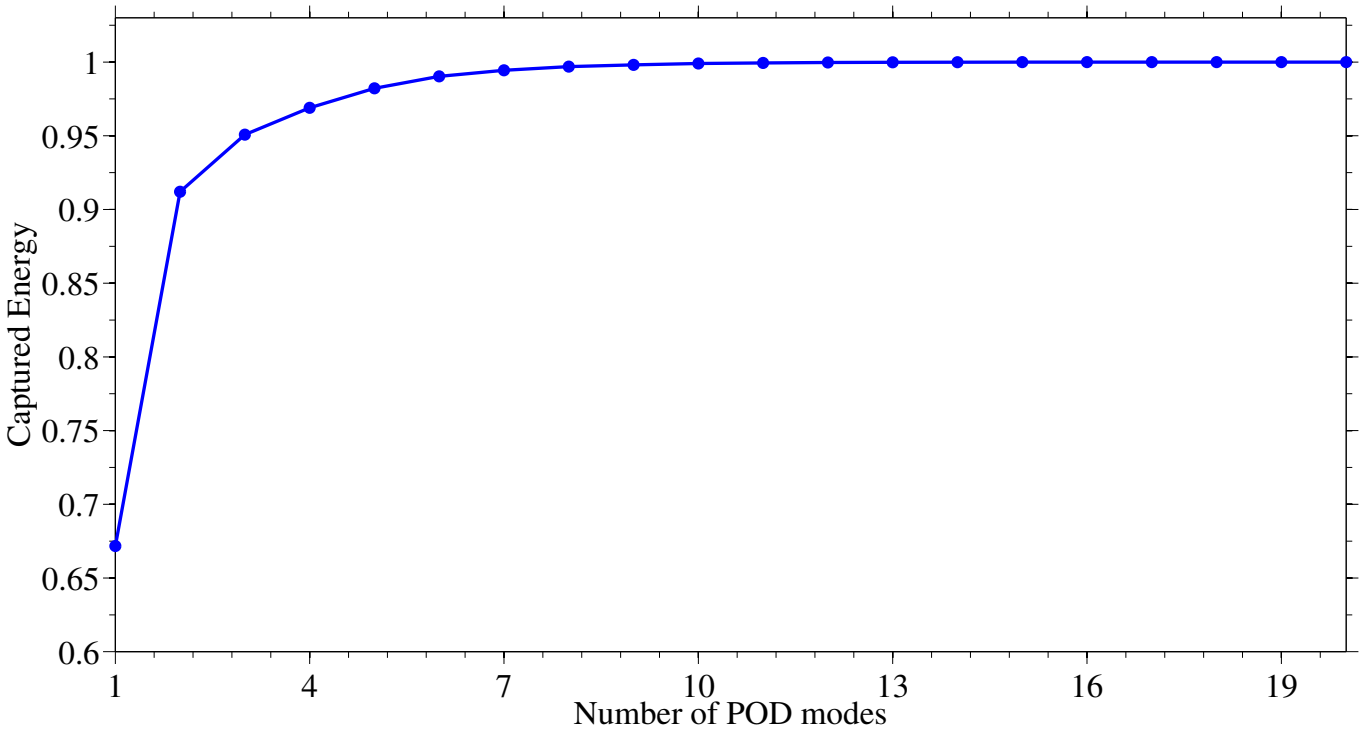


Fig. 3: Leading eigenmodes.

Experiment I - Initial State Solution

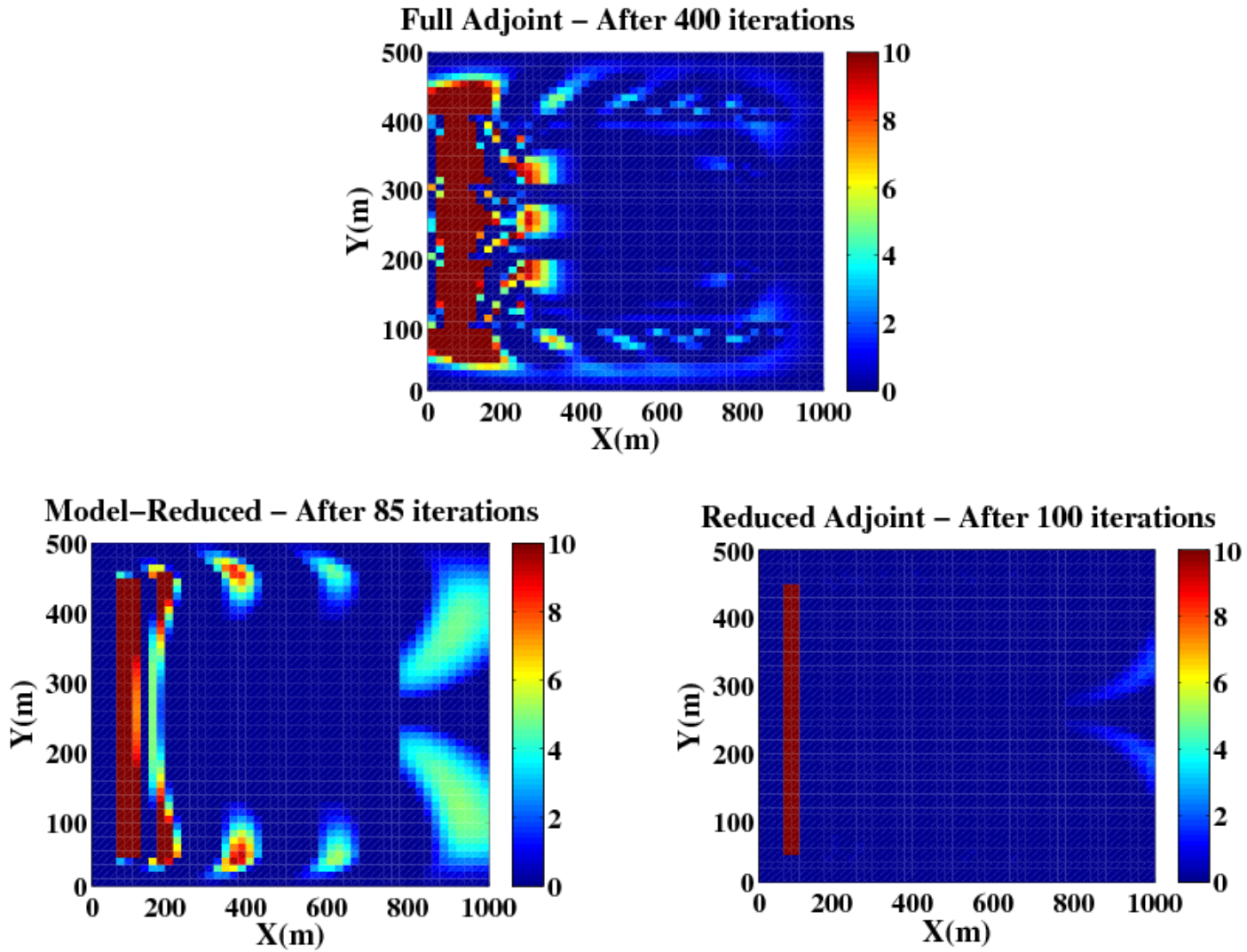


Fig. 4: Solution from the 3 methods after satisfying convergence.

Experiment I - Convergence

Full adjoint is slowly converging, reduced-model adjoint is slightly inaccurate whereas the reduced adjoint is fast and highly accurate.

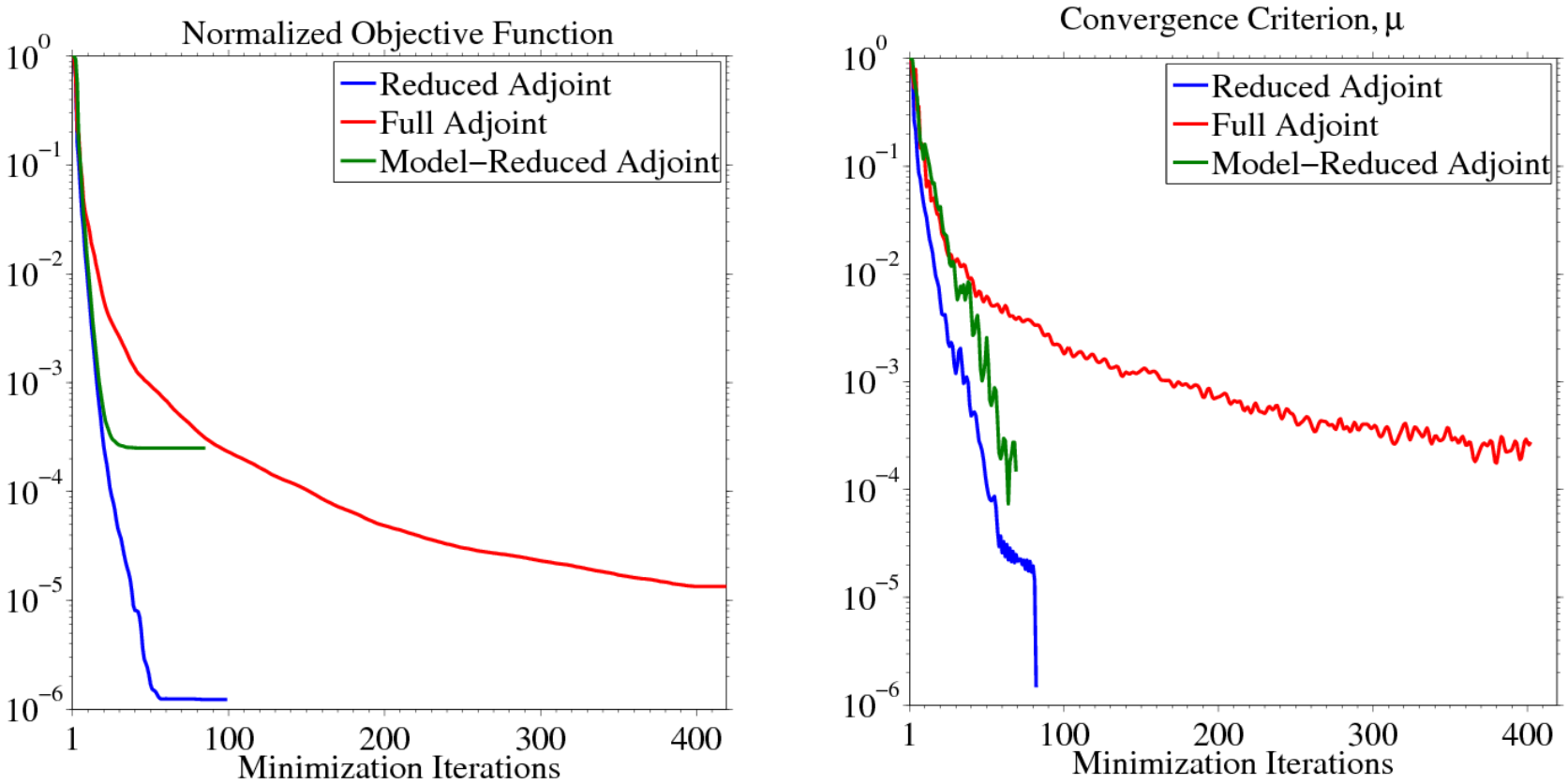
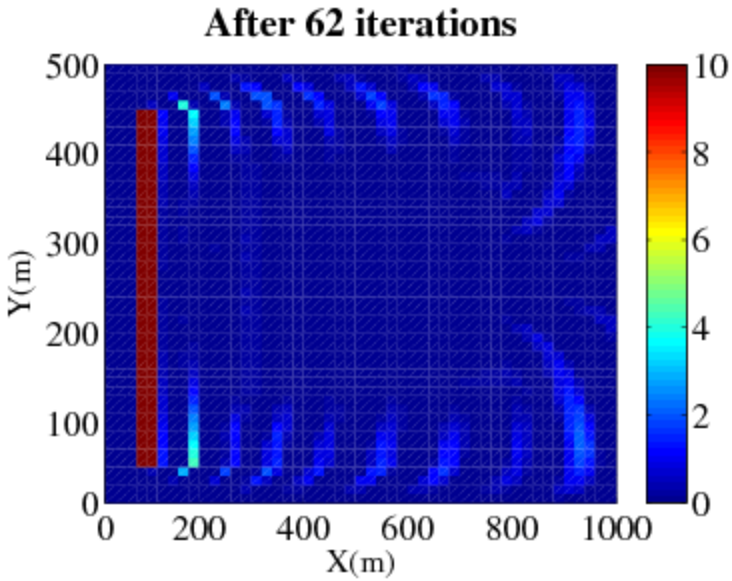


Fig. 5: Cost function and gradient minimization history.

Experiment II

When the observations are perturbed and not available every time step, the reduced adjoint method was still effective.



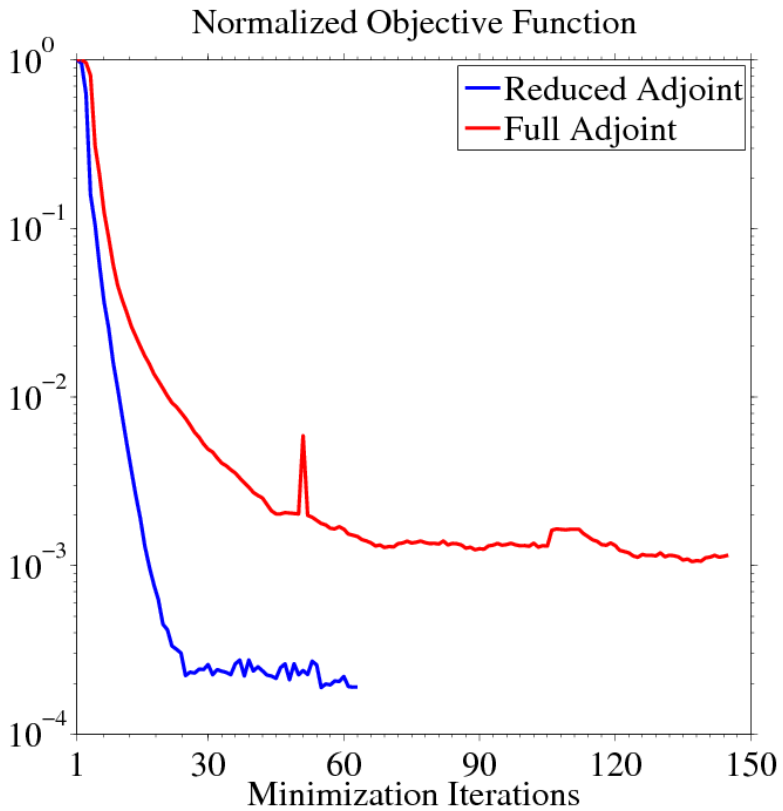


Fig. 6: Cost function history and initial state estimate.

Conclusion

- Adjoint method is an efficient tool for sensitivity analysis.
- Adjoint method is computationally expensive as one iteration backward in time is usually equivalent to several forward model simulations. Also, implementing the adjoint requires significant programming effort.
- Model-reduced adjoint method mainly suffers from accuracy problems.
- The proposed adjoint method is based on running the forward simulations in full space while running the adjoint model in reduced space which can be utilized using POD.
- Preliminary results suggest that the reduced adjoint method is efficient, computationally less expensive and highly reliable.

Conclusion

- Adjoint method is an efficient tool for sensitivity analysis.
- Adjoint method is computationally expensive as one iteration backward in time is usually equivalent to several forward model simulations. Also, implementing the adjoint requires significant programming effort.
- Model-reduced adjoint method mainly suffers from accuracy problems.
- The proposed adjoint method is based on running the forward simulations in full space while running the adjoint model in reduced space which can be utilized using POD.
- Preliminary results suggest that the reduced adjoint method is efficient, computationally less expensive and highly reliable.

Conclusion

- Adjoint method is an efficient tool for sensitivity analysis.
- Adjoint method is computationally expensive as one iteration backward in time is usually equivalent to several forward model simulations. Also, implementing the adjoint requires significant programming effort.
- Model-reduced adjoint method mainly suffers from accuracy problems.
- The proposed adjoint method is based on running the forward simulations in full space while running the adjoint model in reduced space which can be utilized using POD.
- Preliminary results suggest that the reduced adjoint method is efficient, computationally less expensive and highly reliable.

Conclusion

- Adjoint method is an efficient tool for sensitivity analysis.
- Adjoint method is computationally expensive as one iteration backward in time is usually equivalent to several forward model simulations. Also, implementing the adjoint requires significant programming effort.
- Model-reduced adjoint method mainly suffers from accuracy problems.
- The proposed adjoint method is based on running the forward simulations in full space while running the adjoint model in reduced space which can be utilized using POD.
- Preliminary results suggest that the reduced adjoint method is efficient, computationally less expensive and highly reliable.

Conclusion

- Adjoint method is an efficient tool for sensitivity analysis.
- Adjoint method is computationally expensive as one iteration backward in time is usually equivalent to several forward model simulations. Also, implementing the adjoint requires significant programming effort.
- Model-reduced adjoint method mainly suffers from accuracy problems.
- The proposed adjoint method is based on running the forward simulations in full space while running the adjoint model in reduced space which can be utilized using POD.
- Preliminary results suggest that the reduced adjoint method is efficient, computationally less expensive and highly reliable.

THE END ...
THANK YOU

mohamad.elgharamti@kaust.edu.sa