Different Regularisation Techniques for Four Dimensional Data Assimilation (4DVar)

Melina A. Freitag

Department of Mathematical Sciences, University of Bath, UK[†]

email: `m.freitag@maths.bath.ac.uk`

Introduction

Four dimensional data assimilation aims to minimise the cost function

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^B) + \sum_{i=0}^{n} (\mathbf{y}_i - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{y}_i - H_i(\mathbf{x}_i))$$

subject to model dynamics $\mathbf{x}_i = M_{0\to i}\mathbf{x}_0$, where $\mathbf{x}_0$ is the sought-after initial state. $\mathbf{x}_0^B$ is the initial background state and $\mathbf{B}$, $\mathbf{R}_i$ are covariance matrices. The observations are given by $\mathbf{y}_i$ and H_i is the observation operator.

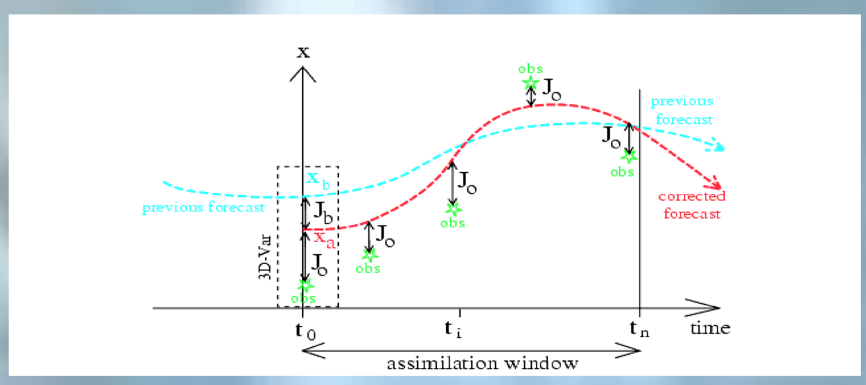

Fig. 1: 4DVar (Copyright:ECMWF)

Tikhonov regularisation

The operator $\mathbf{A}$ in the operator equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

it is well-posed if

- a solution to the operator equation exits
- the solution is unique
- the solution is stable ($\mathbf{A}^{-1}$ continuous)

Equation is ill-posed if it is not well-posed. In finite dimensions existence and uniqueness can be imposed, but

- discrete problem becomes ill-conditioned
- singular values of $\mathbf{A}$ decay to zero $\Rightarrow \mathbf{A}^{-1}$ is unstable!

Use regularisation parameter α to stabilise the problem

$$\mathbf{x}_\alpha = \arg\min \left\{ \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 + \alpha\|\mathbf{x}\|^2 \right\} = (\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\mathbf{b}.$$

Using the SVD of $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ the regularised solution in Tikhonov regularisation is given by

$$\mathbf{x}_\alpha = (\mathbf{A}^T\mathbf{A} + \alpha\mathbf{I})^{-1}\mathbf{A}^T\mathbf{b} = \mathbf{V}\mathrm{diag}\left(\frac{s_i^2}{s_i^2 + \alpha}\frac{1}{s_i}\right)\mathbf{U}^T\mathbf{b} = \sum_{i=1}^{n} \frac{s_i^2}{s_i^2+\alpha}\frac{u_i^T\mathbf{b}}{s_i}\mathbf{v}_i.$$

Linking 4DVar and Tikhonov regularisation

Rewrite the cost function as

$$J(\mathbf{x}_0) = (\mathbf{x}_0 - \mathbf{x}_0^B)^T \mathbf{B}^{-1} (\mathbf{x}_0 - \mathbf{x}_0^B) + (\hat{\mathbf{y}} - \hat{\mathbf{H}}(\mathbf{x}_0))^T \hat{\mathbf{R}}^{-1} (\hat{\mathbf{y}} - \hat{\mathbf{H}}(\mathbf{x}_0))$$

where $\hat{\mathbf{y}} = [\mathbf{y}_0^T, \ldots, \mathbf{y}_n^T]^T$ and

$$\hat{\mathbf{H}} = [H_0^T, (H_1 M(t_1, t_0))^T, \ldots (H_n M(t_n, t_0))^T]^T$$

and $\hat{\mathbf{R}}$ is block diagonal with $\mathbf{R}_i$ on diagonal. Linearise about $\mathbf{x}_0$ then the solution to the optimisation problem is given by

$$\mathbf{x}_0 = \mathbf{x}_0^B + (\mathbf{B}^{-1} + \hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{H}})^{-1}\hat{\mathbf{H}}^T\hat{\mathbf{R}}^{-1}\hat{\mathbf{d}},$$
$$\hat{\mathbf{d}} = \hat{\mathbf{H}}(\mathbf{x}_0^B - \hat{\mathbf{y}})$$

Assume $\mathbf{B} = \sigma_B^2\mathbf{I}$ and $\hat{\mathbf{R}} = \sigma_O^2\mathbf{I}$ and $\hat{\mathbf{H}} = \mathbf{U}\Sigma\mathbf{V}^T$. Then the optimal analysis can be written as

$$\mathbf{x}_0 = \mathbf{x}_0^B + \sum_j \frac{s_j^2}{\mu^2 + s_j^2}\frac{\mathbf{u}_j^T\hat{\mathbf{d}}}{s_j}\mathbf{v}_j,$$

where $\mu^2 = \dfrac{\sigma_O^2}{\sigma_B^2}$. For nondiagonal covariance matrices variable transformation with $\mathbf{B} = \sigma_B^2\mathbf{F}_B$, $\hat{\mathbf{R}} = \sigma_O^2\mathbf{F}_R$ and $\mathbf{z} := \mathbf{F}_B^{-1/2}(\mathbf{x}_0 - \mathbf{x}_0^B)$ gives

$$\hat{J}(\mathbf{z}) = \mu^2\|\mathbf{z}\|_2^2 + \|\mathbf{F}_R^{-1/2}\hat{\mathbf{d}} - \mathbf{F}_R^{-1/2}\hat{\mathbf{H}}\mathbf{F}_B^{-1/2}\mathbf{z}\|_2^2$$

where μ^2 can be interpreted as a regularisation parameter.

Results from image deblurring

In image processing, L_1-norm regularisation and Total Variation regularisation provides edge preserving image deblurring!

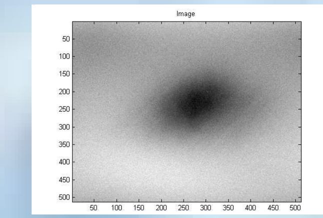

Fig. 2: Blurred picture

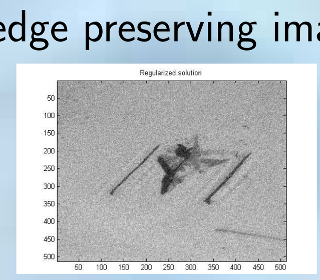

Fig. 2: Tikhonov regularisation

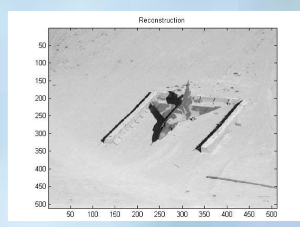

Fig. 2: L_1-norm regularisation

- 4DVar/Tikhonov regularisation smears out sharp fronts
- L_1-norm/TV regularisation beneficial in data assimilation

L_1 regularisation within 4DVar

Burger's equation

$$u_t + u\frac{\partial u}{\partial x} = u + f(u)_x = 0, \quad f(u) = \frac{1}{2}u^2$$

with initial conditions

$$u(x,0) = \begin{cases} 2 & 0 \le x < 2.5 \\ 0.5 & 2.5 \le x \le 10. \end{cases}$$

Discretising

$$x(j) = 10(j - 1/2)\Delta x; \quad U^0(x(j)) = \begin{cases} 2 & 0 \le x(j) < 2.5 \\ 0.5 & 2.5 \le x(j) \le 10. \end{cases}$$

with $\Delta x = \frac{1}{100}$ and $j = 1, \ldots, N$. The exact solution is obtained using the method of characteristics (Riemann problem)

$$u(x,t) = \begin{cases} 2 & 0 \le x < 2.5 + st \\ 0.5 & 2.5 + st \le x \le 10, \end{cases}$$

where $s = 1.25$ The numerical solution (which introduces model error) is obtained using

- the Lax-Friedrich method (smearing out the shock)

$$U_j^{n+1} = \frac{1}{2}(U_{j-1}^n + U_{j+1}^n) - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)).$$

- the Lax-Wendroff method (oscillations near the shock).

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{2\Delta x}(f(U_{j+1}^n) - f(U_{j-1}^n)) + \frac{\Delta t^2}{2\Delta x^2}\left(A_{j+\frac{1}{2}}(f(U_{j+1}^n) - f(U_j^n)) - A_{j-\frac{1}{2}}(f(U_j^n) - f(U_{j-1}^n))\right).$$

We use 3 regularisation methods:

4DVar

$$J(U^0) = \frac{1}{2}\|U_B^0 - U^0\|_{\alpha B}^2 + \frac{1}{2}\sum_{i=1}^{N}\|Y_i - H_i(U_i)\|_{R_i}^2$$

L_1-regularisation

$$J(U^0) = \frac{1}{2}\|Z_B^0 - Z^0\|_p^p + \frac{1}{2}\sum_{i=1}^{N}\|Y_i - H_i(U_i)\|_{R_i}^2$$

where $p = 1$ (or $p = 1.0001$) and $Z = (\alpha B)^{-1/2}U$.

TV regularisation

$$J(U^0) = \frac{1}{2}\|D(Z_B^0 - Z^0)\|_p^p + \frac{1}{2}\sum_{i=1}^{N}\|Y_i - H_i(U_i)\|_{R_i}^2$$

where D is a matrix approximating the first derivative.

- $\Delta t = 0.001$
- length of the assimilation window: 100 time steps
- perfect observations
- use use quadratic programming tools for $p = 1$.

Acknowledgment

This research project is supported by

Lax-Friedrich method

4DVar - observations everywhere

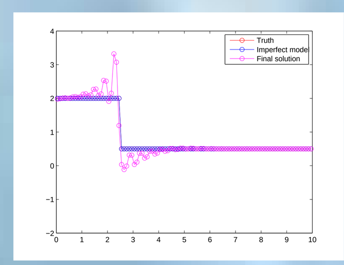

Fig. 3: $t = 0$

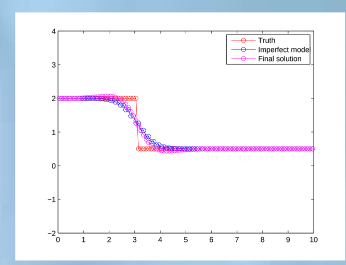

Fig. 3: $t = 50$

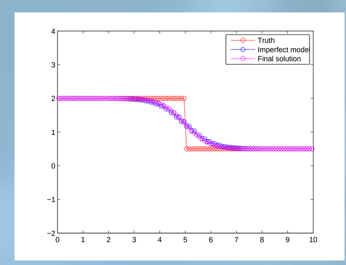

Fig. 3: $t = 200$

L_1/TV regularisation - observations everywhere

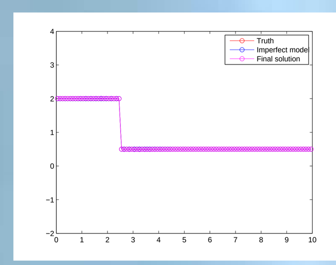

Fig. 4: $t = 0$

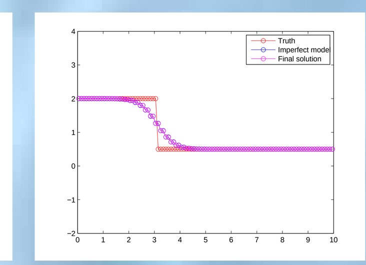

Fig. 4: $t = 50$

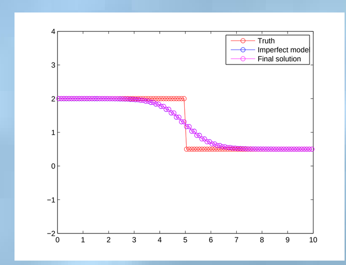

Fig. 4: $t = 200$

Root mean square error

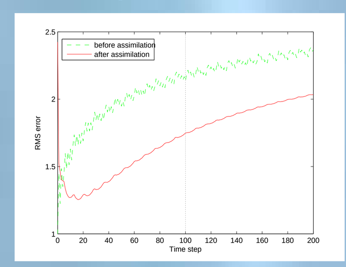

Fig. 5: 4DVar/Tikhonov.

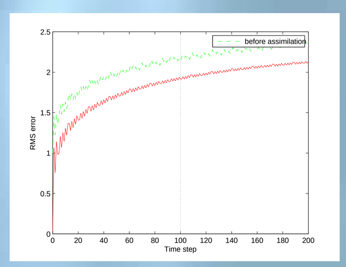

Fig. 5: L_1/TV regularisation.

4DVar - observations every 2 time steps and 20 points in space

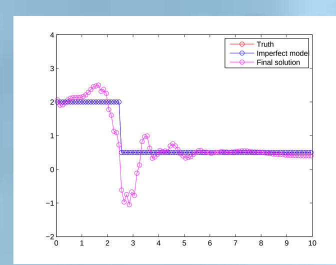

Fig. 6: $t = 0$

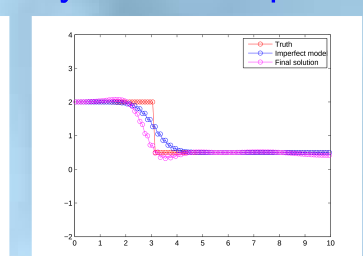

Fig. 6: $t = 50$

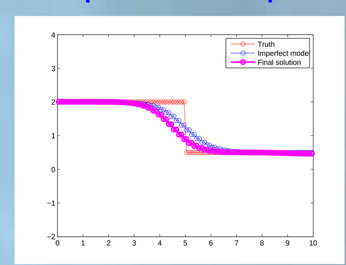

Fig. 6: $t = 200$

L_1/TV regularisation - observations every 2 time steps and 20 points in space

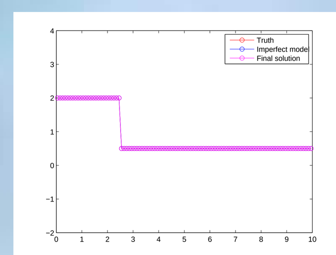

Fig. 7: $t = 0$

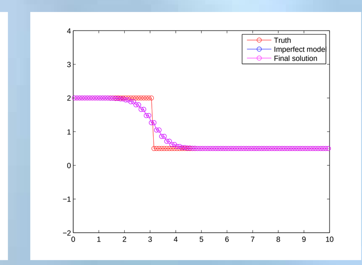

Fig. 7: $t = 50$

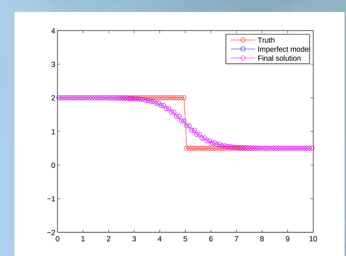

Fig. 7: $t = 200$

Root mean square error

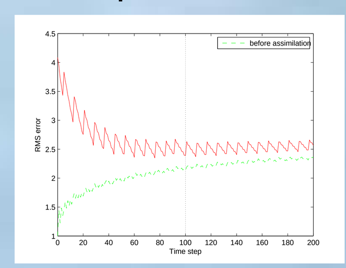

Fig. 8: 4DVar/Tikhonov.

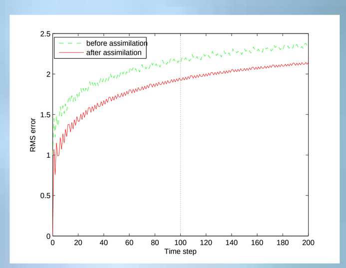

Fig. 8: L_1/TV regularisation.

Lax-Wendroff method

Root mean square error (observations everywhere)

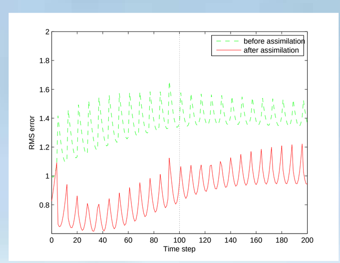

Fig. 9: 4DVar/Tikhonov.

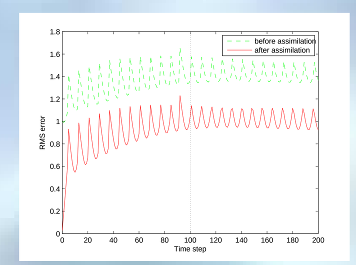

Fig. 9: L_1/TV regularisation.

Root mean square error (observations every 2 time steps and 20 points in space)

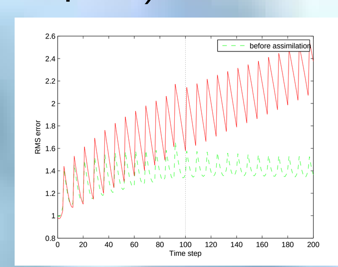

Fig. 10: 4DVar/Tikhonov.

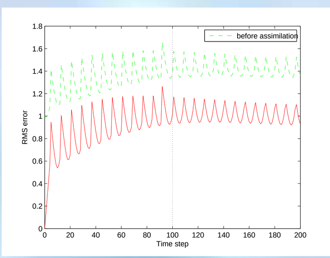

Fig. 10: L_1/TV regularisation.

Observations

- 4DVar is very sensitive to the regularisation parameter α, whereas L_1-norm and TV regularisation are very robust with respect to different values of α.
- experiments with noisy observations and/or different B matrices give similar results.
- TV regularisation converges (generally) faster than L_1-norm regularisation.

Conclusions

Both L_1-norm and TV regularisation recover discontinuity better than 4DVar.

[†]Address for correspondence: Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, BA2 7AY