School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

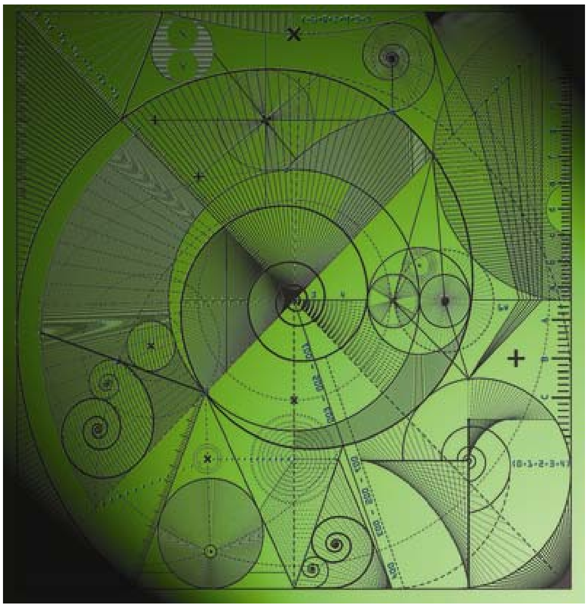
Preprint MPS-2012-05

27 February 2012

On Error Dynamics and Instability in Data Assimilation

by

Roland W.E. Potthast, Alexander J.F. Moodey,
Amos S. Lawless and Peter Jan van Leeuwen



On Error Dynamics and Instability in Data Assimilation

Roland W.E. Potthast[1,2], Alexander J.F. Moodey[2], Amos S. Lawless[2], Peter Jan van Leeuwen[2]

[1] German Meteorological Service - Deutscher Wetterdienst Research and Development, Frankfurter Strasse 135, 63067 Offenbach, Germany

[2] University of Reading, School of Mathematical and Physical Sciences, Whiteknights, PO Box 220, Reading RG6 6AX, United Kingdom

Abstract. Cycled Data Assimilation algorithms are a standard tool to adapt dynamical systems to measurements, usually based on a diversity of different observation systems. Usually, such systems use a *first guess* calculated from previous analysis and then solve some type of *inverse problem* to integrate the measured data into the current state estimate of the dynamical system under consideration. The mapping of the system state φ in its state space X onto the measurements f in a measurement space Y is carried out by a measurement operator $H : X \to Y$. The propagation of the analysis $\varphi^{(a)}$ at some time t_k to the time t_{k+1} is carried out by the model operator M_k.

Errors in the observations, the model and the observation operator will lead to some *analysis error* in each time step. This analysis error can increase over time and lead to severe instabilities in the assimilation process, in particular when the analysis equation $H\varphi = f$ is ill-conditioned or ill-posed. This is the case in particular when the observation operator H includes remote sensing type observations, which often leads to ill-posed analysis equations and a corresponding potential instability in each time-step. Standard regularization provides a tool to stabilise each step, but not necessarily the sequence of cycled inversions.

We will study the set-up in the linear case when the model operator M is amplifying a finite number of modes and damping all sufficiently high modes. In particular, we will derive conditions on the regularization under which the data assimilation system will be stable over time. This includes full proofs for M of Hilbert-Schmidt type and numerical examples to confirm and illustrate the results.

1. Introduction

Data assimilation algorithms are of growing importance for many areas of science, ranging from numerical weather prediction and climate projection to cognitive neuroscience. While the area of inverse problems traditionally is looking more into static inversion[EHN00][KS04], data assimilation is investigating problems with time dynamics and repeated measurements of a system which are changing over time [LLD06].

Usually, some phenomenon is described by a dynamical system with states φ in a state space X. The dynamical system leads to a model operator M, such that M_k takes a state φ_k at time t_k into a new state φ_{k+1} at time t_{k+1}, where we assume that we consider time steps $t_0 < t_1 < t_2 < ....$ Let Y be our measurement space, which we assume to be fixed within this work. For the data assimilation task we are given measurements $f_k \in Y$ at time t_k. Also, we assume that we know some initial guess $\varphi_0 \in X$ at time t_0. The task of data assimilation is to successively calculate some analysis $\varphi_k^{(a)}$ at time t_k which is estimating the true system state $\varphi_k^{(true)}$ at time t_k[LLD06]. Usually, we are also interested in an estimate for the uncertainty of the estimate $\varphi^{(a)}$, or an estimate for the analysis error

$$e_k^{(a)} := \left|\left|\varphi_k^{(a)} - \varphi_k^{(true)}\right|\right|, \quad k \in \mathbb{N}_0. \tag{1.1}$$

Here, we are interested in Hilbert-space type error estimates for the finite-dimensional as well as the *infinite-dimensional* case, which provides a good model for high-dimensional systems. We call a data assimilation system stable, if $e_k^{(a)}$ remains bounded by some constant $C > 0$ for $k \to \infty$. If $e_k^{(a)} \to \infty$ for $k \to \infty$, we call it *unstable*. When we have an estimate

$$\left|\left|f_k - f_k^{(true)}\right|\right| \leq \delta, \tag{1.2}$$

with $f_k^{(true)} = H^{(true)}\varphi^{(true)}$ denoting the values in data space corresponding to the true state $\varphi^{(true)}$ in state space when calculated by the true observation operator $H^{(true)}$, we say that the measurement error is bounded by $\delta \geq 0$. The case $\delta = 0$ corresponds to perfect data. A standard approach to regularize the assimilation is to employ a variant of the Tikhonov inverse $R_\alpha := (\alpha I + H^* H)^{-1} H^*$ (with the adjoint H^* of H) approximating the inverse H^{-1} in equation

$$H\varphi_k = f_k, \quad k \in \mathbb{N}, \tag{1.3}$$

which needs to be solved in every assimilation step.

We will see that in general, when we have model dynamics which is amplifying particular modes and when H is an ill-posed operator, we cannot expect the data assimilation systems to be stable, even if we have arbitrarily small measurement error $\delta > 0$. However, for trace-class model dynamics M_k, $k \in \mathbb{N}_0$ we will show that for sufficiently chosen regularization of the inverse equation (1.3) we will obtain stable data assimilation systems.

Of course, filtering theory has studied the data assimilation task for a long time. In a finite dimensional setting and in the case of the Kalman filter with well-posed observation operators simple versions of our formulas are well-known in control/systems theory of state estimation. Here, we are interested in the cycling of data assimilation schemes, such that we are not obtaining an optimal state estimation, nor do we claim in this work that we are achieving an optimal state estimation. Instead, we are considering a variant of a state observer or a Luenberger observer, see [Lue64],[Lue66]. Since the 1960s/70s much theory has been developed for a Luenberger observer with consideration to the analysis error. However, the theory developed has mostly considered the finite dimensional case, see [Kal60],[Jaz70],[BH75],[AM79],[O'R83],[Rug96],[CF03],[Sim06]. Furthermore, in the finite dimensional setting it is well known that if (M, H) are completely observable then R_α can be chosen such that the analysis error is stable for all time, compare [O'R83][BC85].

The infinite-dimensional state space approach has also been considered since the 1970's. Notably [CP78] developed many results from the finite-dimensional setting into the infinite case. They considered the system dynamics in terms of a strongly continuous semigroup on an appropriate Banach space. Their work, like others, including [But69],[Lio71] focused on the continuous stochastical derivation, with particular attention to optimal state estimation, i.e the Kalman filter. Work has been done in terms of discrete-time systems in the infinite-dimensional setting, which is known as *sampled data systems* in control theory. The work [Dul96][DP00] develop a considerable number of results that relate to this work, however they do not consider the consequence of an ill-posed observation operator. Here, we will focus on the investigation of the practically relevant situation of cycled schemes in a large-scale or infinite-dimensional environment with an ill-posed observation operator. Also, we focus on Hilbert-space type error estimates usually investigated in deterministic applied mathematics. More recently work has considered the state estimation problem as an inverse problem. [Cul12] has looked at the impact of the regularization with cycled data assimilation schemes in the scalar case. Furthermore, more closely related to this work, [BLL$^+$12] have considered the stability of cycled data assimilation schemes, developing theoretical results within the context of strongly nonlinear setting with the 2D incompressible Navier-Stokes equations. [BLL$^+$12] have developed theory for the particular Navier-Stokes system, here we treat any linear model operator. Moreover, [BLL$^+$12] have considered a very particular observation operator, whereby here we study general ill-posed observation operators.

The work will be split into different parts. First, we introduce our set-up and provide a uniform view into three-dimensional variational assimilation (3dVar), four-dimensional variational assimilation (4dVar) and a cycled Kalman filter for linear systems by using a *cycled Tikhonov regularization* approach to data assimilation in Section 2. Then, we work out error estimates and study the error evolution in Section 3. For Hilbert-Schmidt operators we carry out a study of stability depending on the regularization parameter α in Section 4. Finally, numerical results will be provided in Section 5.

2. Cycled Data Assimilation Algorithms

The goal of this section is to introduce our notation and provide a uniform view onto cycled data assimilation by well-known methods like 3dVar, 4dVar and a cycled Kalman filter. We will work out our proofs within this unified view, such that they apply to several well-known schemes.

For simplicity, we assume that the model operator M_k is a linear mapping $M_k : X \to X$ for $k \in \mathbb{N}_0$ and the observation operator H is linear as well and time-invariant.

The nonlinear case is beyond the scope of this work. The state space X is assumed to be a Hilbert space with scalar product $\langle \cdot, \cdot \rangle_X$ and the measurement space Y is a Hilbert space with scalar product $\langle \cdot, \cdot \rangle_Y$. Often, we will drop the indices X or Y, respectively, since it is usually clear which scalar product is used.

At every time step t_k, $k \in \mathbb{N}$ the task of data assimilation is to solve the equation (1.3) given the background $\varphi^{(b)}$, which is a prior estimate calculated from earlier analysis by

$$\varphi_k^{(b)} = M_{k-1} \varphi_{k-1}^{(a)}, \quad k = 0, 1, 2, ... \tag{2.1}$$

This task can be carried out by minimizing

$$J_{Tikhonov}(\varphi) := \alpha \left|\left| \varphi - \varphi_k^{(b)} \right|\right|^2 + ||f_k - H\varphi||^2, \quad \varphi \in X. \tag{2.2}$$

This is known as the *Tikhonov functional*. For linear operators the minimum is given by the *normal equations*, which can be reformulated into the update formula

$$\varphi_k^{(a)} = \varphi_k^{(b)} + R_\alpha \Big(f_k - H \varphi_k^{(b)} \Big) \tag{2.3}$$

with the *regularized inverse*

$$R_\alpha := (\alpha I + H^* H)^{-1} H^*, \quad \alpha > 0, \tag{2.4}$$

of the operator H. The inverse (2.4) is known as *Tikhonov regularization* with *regularization parameter* $\alpha > 0$. The operator R_α is also known as *Kalman gain matrix* in filtering theory literature, compare [Kal60][Jaz70][AM79]. In terms of the analysis fields $\varphi_k^{(a)}$ (2.3) has the form

$$\varphi_k^{(a)} = M_{k-1} \varphi_{k-1}^{(a)} + R_\alpha \Big(f_k - H M_{k-1} \varphi_{k-1}^{(a)} \Big) \tag{2.5}$$

for $k = 0, 1, 2, ...$

In the standard data assimilation literature, usually weighted norms are used in the state space X and the observation space Y. If $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$ these weights can be defined using invertible matrices $B \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$, given as the covariance matrices of the states and the observations with respect to some probability distributions. Here, our analysis will work independently of the particular background on which the weight matrices are defined. We focus on the evolution of the states with respect to the corresponding Hilbert-space norms. We define

$$\langle \varphi, \psi \rangle_{B^{-1}} := \left\langle \varphi, B^{-1} \psi \right\rangle_{l^2}, \quad \langle f, g \rangle_{R^{-1}} := \left\langle f, R^{-1} g \right\rangle_{l^2}, \tag{2.6}$$

where the indices l^2 indicate standard l^2 scalar products on $\mathbb{R}^n$ or $\mathbb{R}^m$, respectively. With the weighted scalar products we can fully apply cycled Tikhonov regularization. For the choice $\alpha = 1$ it is identical to what is known as 3dVar. When H' denotes the adjoint operator with respect to the l^2 scalar product and H^* is the adjoint with respect to the weighted scalar product, with some lines of calculation we obtain

$$H^* = B H' R^{-1}. \tag{2.7}$$

Then, using $H^*(\alpha I + HH^*)^{-1} = (\alpha I + H^*H)^{-1}H^*$ it is straightforward to verify that (2.3) can be written as

$$\begin{aligned}\varphi_k^{(a)} &= \varphi_k^{(b)} + (\alpha B^{-1} + H'R^{-1}H)^{-1}H'R^{-1}\Big(f_k - H\varphi_k^{(b)}\Big) \\ &= \varphi_k^{(b)} + BH'(\alpha R + HBH')^{-1}\Big(f_k - H\varphi_k^{(b)}\Big).\end{aligned} \tag{2.8}$$

For the case $\alpha = 1$, the formula (2.8) is the standard update formula of 3dVar, i.e. the cycled Tikhonov regularization (2.3) is an equivalent way to write 3dVar, when weighted norms (2.6) are used. We need to remark that though formally 3dVar and Tikonov are the same by (2.7), when we explore the spectrum of H*H this contains B and R. So, the covariance structure of B comes into the problem via the spectrum of the operator H^*H.

Four-dimensional data assimilation (4dVar) in its strong-constrained form calculates a state by minimizing the functional

$$J_{4dVar,k}(\varphi) := \alpha \left|\left|\varphi - \varphi_k^{(b)}\right|\right|^2 + \sum_{\ell=1}^{L} \left|\left|f_{k+\ell} - HM_{k,k+\ell}\varphi\right|\right|^2, \tag{2.9}$$

where $M_{k,k+\ell} := M_{k+\ell-1} \circ \ldots \circ M_{k+1} \circ M_k$ mapping states at t_k into states at $t_{k+\ell}$. For linear systems, we collect the states $f_{k+1}, ..., f_{k+L}$ into a vector $\vec{f}_{k+1}$ and the operators $HM_{k,k+1}, ..., HM_{k,k+L}$ into an operator $\vec{H}$. Then, the minimization of (2.9) is equivalent to minimizing

$$\tilde{J}_{4dVar,k}(\varphi) := \alpha \left|\left|\varphi - \varphi_k^{(b)}\right|\right|^2 + \left|\left|\vec{f} - \vec{H}\,\varphi\right|\right|^2. \tag{2.10}$$

When we define

$$\varphi_{k+L}^{(b)} := M_{k,L}\varphi_k^{(a)} \tag{2.11}$$

the 4dVar algorithm is identical to a 3dVar scheme on the time-grid $t_0, t_L, t_{2L}, ...$ and the cycled Tikhonov regularization (2.3). This means that any analysis on the cycled Tikhonov regularization directly applies to cycled 4dVar data assimilation.

Finally, let us consider the Kalman filter. The basic difference between 3dVar and the Kalman filter is an update of the weight matrix B in every analysis step. It is well-known that for linear systems the Kalman filter is equivalent to 4dVar, when the same cycling strategy is used. Usually, the update formula of the Kalman Filter is

$$B_k^{(a)} = (I - R_\alpha H)\, B_k^{(b)}, \quad B_k^{(b)} = M_{k-1}B_{k-1}^{(a)}M_{k-1}^*. \tag{2.12}$$

which is correct for a perfect linear model. Another form of the update formula is given by

$$\left(B_k^{(a)}\right)^{-1} = \left(B_k^{(b)}\right)^{-1} + H'R^{-1}H, \quad k \in \mathbb{N}. \tag{2.13}$$

This means that the norm of the inverse $(B_k^{(a)})^{-1}$ of the matrix $B_k^{(a)}$ becomes larger and larger over time, i.e. for $k \to \infty$. Usually, resetting B to some initial matrix B_0 after a finite number of steps is used or *covariance inflation*, i.e. the multiplication of B by some number to limit the norm of $(B_k^{(a)})^{-1}$. The reset is equivalent to the cycling of

4dVar as described in (2.11). Covariance inflation will have a similar effect, but it is more difficult to treat theoretically and we postpone details to future work.

3. Error Evolution of Data Assimilation Systems

The goal of this part is to study the error evolution of the cycled data assimilation systems which we introduced in Section 2. Since for linear dynamics and linear observation operators all of them can be reduced to the case of cycled Tikhonov regularization, we can restrict our attention to this generic case. Our main goal here is to study the cumulative influence of errors in the case of ill-posed observation operators.

We first need some further preparations. Assume that the true fields at time t_k are given by $\varphi_k^{(true)}$ and that the true measurement values are $f_k^{(true)} = H^{(true)}\varphi_k^{(true)}$ with the true observation operator $H^{(true)}$. Then, we have

$$\varphi_k^{(true)} = M_{k-1}^{(true)}\varphi_{k-1}^{(true)}, \quad k = 1, 2, ... \tag{3.1}$$

with the true dynamical system $M_k^{(true)}$ at time t_k. We denote the error in the data by

$$f_k^{(\delta)} := f_k - f_k^{(true)}, \quad k \in \mathbb{N}. \tag{3.2}$$

We usually assume that some bound on the measurement error is given in the form (1.2), i.e. we have $||f_k^{(\delta)}|| \leq \delta$ for all $k \in \mathbb{N}$.

We are interested in the evolution of the assimilation error over time. The potential errors can be classified into

A. error in the *measurement data* f_k,

B. error in the *observation operator* H,

C. error in the *model dynamics* M_k,

D. error by the *reconstruction operator* R_α which is not identical to H^{-1},

E. cumulated errors from *previous iterations/cycling*.

Clearly, stochastics measures error as the variance of the corresponding distributions. Here, we work out classical Hilbert space error estimates. First, we subtract the true solution and split the terms to identify the role of different types of errors. We employ formula (2.3) to calculate

$$\begin{aligned} e_{k+1}^{(a)} &= \varphi_{k+1}^{(a)} - \varphi_{k+1}^{(true)} \\ &= \varphi_{k+1}^{(b)} - \varphi_{k+1}^{(true)} + R_\alpha\Big(f_{k+1} - f_{k+1}^{(true)}\Big) + R_\alpha\Big(f_{k+1}^{(true)} - H\varphi_{k+1}^{(b)}\Big) \quad (3.3) \\ &= M_k\varphi_k^{(a)} - M_k^{(true)}\varphi_k^{(true)} + R_\alpha\Big(f_{k+1}^{(\delta)}\Big) \\ &\qquad\qquad + R_\alpha\Big(H^{(true)}\varphi_{k+1}^{(true)} - H\varphi_{k+1}^{(b)}\Big) \\ &= M_k\Big(\varphi_k^{(a)} - \varphi_k^{(true)}\Big) + \Big(M_k - M_k^{(true)}\Big)\varphi_k^{(true)} + R_\alpha\Big(f_{k+1}^{(\delta)}\Big) \\ &\quad + R_\alpha\Big((H^{(true)} - H)\varphi_{k+1}^{(true)} + H\Big(\varphi_{k+1}^{(true)} - \varphi_{k+1}^{(b)}\Big)\Big), \end{aligned}$$

where the difference in the last round bracket can be decomposed as the first difference in (3.3), i.e.

$$\varphi_{k+1}^{(true)} - \varphi_{k+1}^{(b)} = M_k\left(\varphi_k^{(true)} - \varphi_k^{(a)}\right) - \left(M_k - M_k^{(true)}\right)\varphi_k^{(true)}. \tag{3.4}$$

This leads to

$$\begin{aligned} e_{k+1}^{(a)} = &\overbrace{(I - R_\alpha H)}^{reconstruction\ error} \overbrace{\left\{M_k e_k^{(a)} + \left(M_k - M_k^{(true)}\right)\varphi_k^{(true)}\right\}}^{propagation\ of\ previous\ error\ +\ model\ error} \\ &+ \overbrace{R_\alpha\left(f_{k+1}^{(\delta)}\right)}^{data\ error\ influence} \\ &+ \overbrace{R_\alpha\left((H^{(true)} - H)\varphi_{k+1}^{(true)}\right)}^{observation\ operator\ error} \end{aligned} \tag{3.5}$$

We first study the simplified situation where we assume that our computational model is correct, i.e. $M^{(true)} = M$, and that we do not have errors in the observation operator, i.e. $H^{(true)} = H$. Further, we assume that M_k does not depend on k. This leads to

$$\begin{aligned} \varphi_{k+1}^{(a)} - \varphi_{k+1}^{(true)} &= M(\varphi_k^{(a)} - \varphi_k^{(true)}) + R_\alpha H M\left(\varphi_k^{(true)} - \varphi_k^{(a)}\right) + R_\alpha f_{k+1}^{(\delta)} \\ &= (I - R_\alpha H)M\left(\varphi_k^{(a)} - \varphi_k^{(true)}\right) + R_\alpha f_{k+1}^{(\delta)}. \end{aligned} \tag{3.6}$$

We define

$$\Lambda := (I - R_\alpha H)M. \tag{3.7}$$

Then, we obtain the iteration formula

$$e_{k+1}^{(a)} = \Lambda e_k^{(a)} + R_\alpha f_{k+1}^{(\delta)}, \quad k = 0, 1, 2, ... \tag{3.8}$$

THEOREM 3.1 *Assume that the error $f_k^{(\delta)}$ does not depend on k, i.e. that we have some time-independent data error for our data assimilation scheme. Then, the error terms $e_k^{(a)}$ with initial error $e_0^{(a)}$ and $e^{(\delta)} := R_\alpha f^{(\delta)}$ described by the update formula (3.8) evolve according to*

$$e_k^{(a)} = \Lambda^k e_0^{(a)} + \left(\sum_{\ell=0}^{k-1} \Lambda^\ell\right) e^{(\delta)}. \tag{3.9}$$

If $(I - \Lambda)^{-1}$ exists, this can be written as

$$e_k^{(a)} = \Lambda^k e_0^{(a)} + (I - \Lambda)^{-1}(I - \Lambda^k)e^{(\delta)} \tag{3.10}$$

Proof. The first formula is obtained by induction. For $k = 1$ we clearly have $e_1^{(a)} = \Lambda e_0^{(a)} + e^{(\delta)}$. Assume the formula is true for $k - 1$. Then, we calculate

$$\begin{aligned} e_{k+1}^{(a)} &= \Lambda e_k^{(a)} + e^{(\delta)} \\ &= \Lambda\left(\Lambda^k e_0^{(a)} + \left(\sum_{\ell=0}^{k-1} \Lambda^\ell\right) e^{(\delta)}\right) + e^{(\delta)} \end{aligned}$$

$$= \Lambda^{k+1} e_0^{(a)} + \left(\sum_{\ell=1}^{k} \Lambda^{\ell} \right) e^{(\delta)} + e^{(\delta)}$$
$$= \Lambda^{k+1} e_0^{(a)} + \left(\sum_{\ell=0}^{k} \Lambda^{\ell} \right) e^{(\delta)}, \tag{3.11}$$

which is the formula for k replaced by $k+1$ and the induction is shown. Then, in the case where $I - \Lambda$ is invertible, the second formula is obtained using the telescopic sum

$$(I - \Lambda) \sum_{\ell=0}^{k-1} \Lambda^{\ell} = I - \Lambda^k \tag{3.12}$$

for any $k \in \mathbb{N}$, and the proof is complete. □

We now obtain an analysis of the evolution of the assimilation error for $k \to \infty$ in the case where $||\Lambda|| < 1$.

Corollary 3.2 *If $||\Lambda|| < 1$, then the error of the data assimilation scheme with constant observation error $f^{(\delta)}$ tends to*

$$e^{(\infty)} := \lim_{k \to \infty} e_k^{(a)} = (I - \Lambda)^{-1} e^{(\delta)} = \Big(I - M + R_\alpha H M \Big)^{-1} R_\alpha f^{(\delta)}. \tag{3.13}$$

The limit $e^{(\infty)}$ depends continuously on the observation error $f^{(\delta)}$. It tends to zero if the constant $f^{(\delta)}$ tends to 0.

Of course, we are not only interested in the special situation of a constant observation error, but for practical applications we need estimates for an observation error which satisfies some bound

$$\left|\left| f_k^{(\delta)} \right|\right| \leq \delta, \quad k \in \mathbb{N}. \tag{3.14}$$

In this case we cannot explicitly derive an evolution equation for the analysis error. But we can still obtain an evolution equation for a bound of the analysis error as follows. We estimate

$$\left|\left| e_{k+1}^{(a)} \right|\right| \leq ||\Lambda|| \cdot \left|\left| e_k^{(a)} \right|\right| + \left|\left| e^{(\delta)} \right|\right|, \quad k \in \mathbb{N}. \tag{3.15}$$

If we define

$$b_{k+1} := \lambda b_k + \tau, \quad k \in \mathbb{N} \tag{3.16}$$

with initial value $b_0 := ||e_0||$ and $\tau := ||e^{(\delta)}||$, we obtain $||e_k|| \leq b_k$ for $k \in \mathbb{N}$, such that the sequence $(b_k)_{k \in \mathbb{N}}$ provides a bound for the analysis error. We can calculate the evolution of b_k analogously to Theorem 3.1.

Theorem 3.3 *We study the situation where we have a perfect model $M = M^{(true)}$ and a perfect observation operator $H = H^{(true)}$. If the observation error $f_k^{(\delta)}$, $k \in \mathbb{N}$, is* *bounded by $\delta > 0$, $\lambda := ||\Lambda||$ and $\tau := ||R_\alpha||\delta$, then the analysis error $e_k^{(a)}$ is estimated by*

$$\left|\left|e_k^{(a)}\right|\right| \leq \lambda^k \left|\left|e_0^{(a)}\right|\right| + \left(\sum_{\ell=0}^{k-1} \lambda^\ell\right) \tau. \tag{3.17}$$

If $\lambda < 1$, we have

$$\left|\left|e_k^{(a)}\right|\right| \leq \lambda^k \left|\left|e_0^{(a)}\right|\right| + \frac{1-\lambda^k}{1-\lambda}\tau \tag{3.18}$$

such that

$$\limsup_{k\to\infty} \left|\left|e_k^{(a)}\right|\right| \leq \frac{||R_\alpha||\delta}{1-\lambda}. \tag{3.19}$$

We now come to the general situation, where M is some approximation to the true model and H is an approximation to the true measurement operator. In this case we will need to take care of additional terms in the error estimate in (3.5). We calculate

$$\begin{aligned}
\sigma_k &:= (I - R_\alpha H)\left(M - M^{(true)}\right)\varphi_k^{(true)} + R_\alpha\left(H^{(true)} - H\right)\varphi_{k+1}^{(true)} && (3.20)\\
&= \left((I - R_\alpha H)\left(M - M^{(true)}\right) + R_\alpha\left(H^{(true)} - H\right)M^{(true)}\right)\varphi_k^{(true)} && (3.21)\\
&= \left(M - M^{(true)} - R_\alpha H M + R_\alpha H^{(true)} M^{(true)}\right)\varphi_k^{(true)} && (3.22)\\
&= \left((I - R_\alpha H)M - \left(I - R_\alpha H^{(true)}\right)M^{(true)}\right)\varphi_k^{(true)} && (3.23)\\
&= \left(\Lambda - \Lambda^{(true)}\right)\varphi_k^{(true)} && (3.24)\\
&= \Lambda^{(error)}\varphi_k^{(true)}, && (3.25)
\end{aligned}$$

where $\Lambda^{(true)} := (I - R_\alpha H^{(true)})M^{(true)}$ and $\Lambda^{(error)} := \Lambda - \Lambda^{(true)}$. This leads to the estimate

$$\begin{aligned}
\left|\left|e_{k+1}^{(a)}\right|\right| &= \left|\left|\Lambda e_k^{(a)} + \Lambda^{(error)}\varphi_k^{(true)} + R_\alpha f_{k+1}^{(\delta)}\right|\right| && (3.26)\\
&\leq ||\Lambda|| \cdot \left|\left|e_k^{(a)}\right|\right| + \left|\left|\Lambda^{(error)}\varphi_k^{(true)}\right|\right| + \left|\left|R_\alpha f_{k+1}^{(\delta)}\right|\right| && (3.27)
\end{aligned}$$

If we define

$$b_{k+1} := \lambda b_k + \tau + \sigma, \quad k \in \mathbb{N} \tag{3.28}$$

with initial value $b_0 := ||e_0||$, $\tau := ||e^{(\delta)}||$ and σ as an upper bound of $||\sigma_k||$, we obtain $||e_k^{(a)}|| \leq b_k$ for $k \in \mathbb{N}$, such that the sequence $(b_k)_{k\in\mathbb{N}}$ provides a bound for the analysis error. We have shown the following result.

Theorem 3.4 *In the case of some general model M approximating $M^{(true)}$ and an observation operator H approximating $H^{(true)}$ we assume that the observation error $f_k^{(\delta)}$, $k \in \mathbb{N}$, is bounded by $\delta > 0$. We use the abbreviations $\lambda := ||\Lambda||$, $\tau := ||R_\alpha||\delta$ and assume that $\sigma > 0$ is some constant bounding the influence of the error of the dynamical model as well as the error of modelling the observation operator via $||\sigma_k|| \leq \sigma$ with σ_k given by (3.20). Then, the analysis error $e_k^{(a)}$ is estimated by*

$$\left|\left|e_k^{(a)}\right|\right| \leq \lambda^k \left|\left|e_0^{(a)}\right|\right| + \frac{1-\lambda^k}{1-\lambda}(\tau + \sigma) \tag{3.29}$$

such that

$$\limsup_{k\to\infty} \left|\left|e_k^{(a)}\right|\right| \leq \frac{||R_\alpha||\delta + \sigma}{1-\lambda}. \tag{3.30}$$

4. Stabilizing Cycled Data Assimilation

We have described error estimates for the analysis error of cycled data assimilation in Theorems 3.3 and 3.4. One key assumption to keep the error bounded is the estimate $\lambda = ||\Lambda|| < 1$. This condition means that $||(I - R_\alpha H)M|| < 1$, i.e. the model dynamics is not increasing the error stronger than the regularized reconstruction can reduce it. Here, we investigate trace-class model operators M to show that we can obtain this condition for appropriately chosen regularization parameter α.

LEMMA 4.1 *For a finite dimensional state space $X = \mathbb{R}^n$ and injective operators H, we can always make $N := I - R_\alpha H$ arbitrarily small in norm.*

Proof. This is due to the fact that H^*H is self-adjoint and thus it has a complete basis $\varphi^{(1)}, ..., \varphi^{(n)}$ of eigenvectors with eigenvalues $\lambda_j > 0$ for $j = 1, ..., n$. We set

$$c := \min_{j=1,\ldots,n} |\lambda_j| > 0. \tag{4.1}$$

We can transform N into

$$N = I - R_\alpha H = I - (\alpha I + H^*H)^{-1}H^*H = \alpha(\alpha I + H^*H)^{-1} \tag{4.2}$$

and estimate

$$||N|| \leq \max_{j=1,\ldots,n} \left|\frac{\alpha}{\alpha + \lambda_j}\right| \leq \frac{\alpha}{\alpha + c}. \tag{4.3}$$

Given c we can always choose $\alpha > 0$ sufficiently small such that the norm $||N||$ of N is arbitrarily small. □

Remark. As a consequence of the previous lemma, given M we can choose $\alpha > 0$ such that $||\Lambda|| \leq ||N||||M|| < 1$.

For the case of infinite dimensions, which is much closer to realistic situations, we need to take more care, since in general the constant c in the above arguments is zero. Here, we will work out the case where M is a trace-class operator. We first collect notations and set-up our scene for further arguments.

Let $\{\psi_\ell : \ell \in \mathbb{N}\}$ be a complete orthonormal system in X. Then, any vector $\varphi \in X$ can be decomposed into its Fourier sum

$$\varphi = \sum_{\ell=1}^{\infty} \alpha_\ell \psi_\ell \tag{4.4}$$

with $\alpha_\ell := \langle \varphi, \psi_\ell \rangle$ for $\ell \in \mathbb{N}$. We apply this to $M\varphi$ to derive

$$M\varphi = \sum_{\ell=1}^{\infty} \langle M\varphi, \psi_\ell \rangle \psi_\ell = \sum_{\ell=1}^{\infty}\sum_{j=1}^{\infty} \alpha_j \langle M\psi_j, \psi_\ell \rangle \psi_\ell = \sum_{\ell=1}^{\infty}\sum_{j=1}^{\infty} \psi_\ell M_{j,\ell} \alpha_j$$

with the infinite matrix $M_{j,\ell} = \langle M\psi_j, \psi_\ell \rangle$. The *Hilbert-Schmidt* or *Frobenius norm* of M with respect to the orthonormal system $\{\psi_\ell : \ell \in \mathbb{N}\}$ is defined as

$$||M||_{Fro} := \sum_{\ell=1}^{\infty} \sum_{j=1}^{\infty} |M_{j,\ell}|^2. \tag{4.5}$$

Our results will be based on estimates with respect to particular orthonormal systems linked to the observation operator. However, to work for general systems we need to rely on the following basic fact, compare [Kna05].

LEMMA 4.2 *For an operator M its Hilbert-Schmidt or Frobenius norm (4.5) is independent of the orthonormal system $\{\psi_\ell : \ell \in \mathbb{N}\}$.*

Let the orthonormal system $\{\psi_\ell : \ell \in \mathbb{N}\}$ in X be given by the singular system of the observation operator $H : X \to Y$. In this case we define an orthogonal decomposition of the space X by

$$X_1^{(n)} := \text{span}\{\psi_1, ..., \psi_n\}, \;\; X_2^{(n)} := \text{span}\{\psi_{n+1}, \psi_{n+2}, ...\}. \tag{4.6}$$

Often, for simplicity we will suppress the index n. Using the orthogonal projection operators P_1 of X onto X_1 and P_2 of X onto X_2, we decompose M into

$$M_1 := P_1 M, \;\; M_2 := P_2 M. \tag{4.7}$$

Using $N = (I - R_\alpha H)$ we obtain

$$\Lambda = N|_{X_1} M_1 + N|_{X_2} M_2. \tag{4.8}$$

The operator N maps X_j, $j = 1, 2$, into itself. This leads to the norm estimate

$$\begin{aligned} ||\Lambda\varphi||^2 &= ||(N|_{X_1} M_1 + N|_{X_2} M_2)\varphi||^2 \\ &= ||N|_{X_1} M_1 \varphi||^2 + ||N|_{X_2} M_2 \varphi||^2, \end{aligned} \tag{4.9}$$

where equality comes from the orthogonality of the spaces X_1 and X_2. This yields

$$||\Lambda||^2 \le ||N|_{X_1}||^2 ||M_1||^2 + ||N|_{X_2}||^2 ||M_2||^2. \tag{4.10}$$

LEMMA 4.3 *On X_1 for $N = I - R_\alpha H$ we have the norm estimate*

$$||N|_{X_1}|| = \sup_{\ell=1,..,n} \left| \frac{\alpha}{\alpha + \mu_n^2} \right| \tag{4.11}$$

where μ_n are the singular values of the operator H ordered according to their size and multiplicity. In particular, given $\epsilon > 0$ and $n \in \mathbb{N}$ we can choose $\alpha > 0$ sufficiently small such that

$$||N|_{X_1}|| < \epsilon. \tag{4.12}$$

Proof. We refer to [MP12] for details about the spectral version of the operator $N = I - R_\alpha H$. In the singular system of H the operator is diagonal with multiplication factors given by

$$N_{(n)} = \frac{\alpha}{\alpha + \mu_n^2}, \;\; n = 1, 2, 3, ..., \tag{4.13}$$

which yields (4.11). The second statement is a result of the order of the singular values $\mu_1 \geq \mu_2 \geq \mu_3 \geq ... > 0$, such that we can choose α sufficiently small to make each term in (4.11) arbitrary small. Since there are only a finite number of such terms, we obtain the estimate (4.12). □

LEMMA 4.4 *The norm of the operator $N|_{X_2}$ is given by*

$$||N|_{X_2}|| = 1 \tag{4.14}$$

for all $n \in \mathbb{N}$ and $\alpha > 0$.

Proof. Since H is compact, we know that $\mu_n \to 0$ for $n \to \infty$. This means that

$$\sup_{\ell=n+1,\ldots,\infty} \left| \frac{\alpha}{\alpha + \mu_n^2} \right| \; = \; 1 \tag{4.15}$$

for all $n \in \mathbb{N}$ and all $\alpha > 0$, which implies the norm estimate (4.14). □

We now need to study the norm of M_1 and M_2. In general, we do not want any limitations on the norm of M_1, since we are interested in a wide range of systems. We will see that this does not generate problems for the stabilization of cycled assimilation schemes.

To estimate the norm of M_2 we derive with the help of the Cauchy-Schwarz inequality

$$\begin{aligned} ||M_2\varphi||^2 &= \sum_{\ell=n+1}^{\infty} |\langle M\varphi, \psi_\ell \rangle|^2 \\ &= \sum_{\ell=n+1}^{\infty} \left| \sum_{j=1}^{\infty} \langle \varphi, \psi_j \rangle \langle M\psi_j, \psi_\ell \rangle \right|^2 \\ &= \sum_{\ell=n+1}^{\infty} \left| \sum_{j=1}^{\infty} \langle \varphi, \psi_j \rangle M_{j,\ell} \right|^2 \\ &\leq \sum_{\ell=n+1}^{\infty} \left(\sum_{j=1}^{\infty} |M_{j,\ell}|^2 \right) \left(\sum_{j=1}^{\infty} |\langle \varphi, \psi_j \rangle|^2 \right) \\ &= \left(\sum_{\ell=n+1}^{\infty} \sum_{j=1}^{\infty} |M_{j,\ell}|^2 \right) ||\varphi||^2. \end{aligned} \tag{4.16}$$

Now, recall that the Hilbert-Schmidt norm of M is finite. This means that

$$|a_\ell|^2 := \sum_{j=1}^{\infty} |M_{j,\ell}|^2, \quad \ell \in \mathbb{N}, \tag{4.17}$$

is a sequence for which

$$\sum_{\ell=1}^{\infty} |a_\ell|^2 < \infty. \tag{4.18}$$

This means that given $\rho > 0$ there is an $n \in \mathbb{N}$ such that

$$\sum_{\ell=n+1}^{\infty} \sum_{j=1}^{\infty} |M_{j,\ell}|^2 = \sum_{\ell=n+1}^{\infty} |a_\ell|^2 < \rho^2. \tag{4.19}$$

This provides an estimate for $||M_2||$ for sufficiently large $n \in \mathbb{N}$. We summarize this result in the following Lemma.

LEMMA 4.5 *If M is a trace-class operator, then given $\rho > 0$ there is $n \in \mathbb{N}$ such that for $M_2 = M_2^{(n)}$ we have $||M_2|| < \rho$.*

Proof. The estimate is given by (4.16) with the help of (4.19). □

We are now able to collect all parts of our analysis to construct cycled data assimilation schemes for ill-posed observation operators which remain stable over time.

THEOREM 4.6 *Assume the system M is trace-class and let α denote the regularization parameter for a cycled data assimilation scheme. Then, for $\alpha > 0$ sufficiently small, i.e. there is α_0 such that for $\alpha < \alpha_0$, we have $||\Lambda|| < 1$. Under the conditions of Theorem 3.4 the analysis error is bounded over time by*

$$\limsup_{k\to\infty} ||e_k|| \leq \frac{||R_\alpha||\delta + \sigma}{1 - ||\Lambda||}. \tag{4.20}$$

Proof. First, we show that we can achieve $||\Lambda|| < 1$. In (4.10) we have $||M_1||$ given as an arbitrary constant c and $||N|_{X_2}|| = 1$ according to Lemma 4.4. First, we use Lemma 4.5 to choose n such that $||M_2||^2 < 1/2$. Now, with fixed n according to Lemma 4.3 the norm $||N|_{X_1}||$ can be made arbitrarily small by choosing α small enough. We choose α such that

$$c^2 \cdot ||N|_{X_1}||^2 \leq \frac{1}{2}. \tag{4.21}$$

Then we obtain $||\Lambda|| < 1$ from (4.10). Now, the bound for the analysis error is given by Theorem 3.4, which also provides the estimate (4.20). □

We have shown that for α sufficiently small we are in a *stable range* for the data assimilation scheme, i.e. the analysis error remains bounded over time. If the observation error tends to zero, this bound will also tend to zero.

We remark that we obtained a *linear* estimate for the dependence of the long-term analysis error depending on the observation error for the stable range of the data assimilation scheme.

The above results have been worked out for the infinite dimensional case. For the finite dimensional case Lemma 4.3 still remains true, but the norm of $N|_{X_2}$ estimated in (4.14) of Lemma 4.4 is not equal to one, but strictly smaller than one. Lemma 4.5 is not really applicable in this case, but we can use Lemma 4.1 such that the result (4.20) applies to this simpler situation with arbitrary linear operator M as well.

Finally, we need to remark that the role of α is to regularize the problem. So it is important to choose α sufficiently large to stay within a stable regime for the inversion of

(1.3). Here, we have shown that we need α sufficiently small to keep the error bounded over time. But the error bound will also depend on α by the norm

$$||R_\alpha|| \le \frac{1}{2\sqrt{\alpha}}$$

of R_α, compare equation (4.20). The bound might in fact be rather large. Given δ there will be some optimal α which leads to the best possible error bounds for a given set-up. However, the maximal α for which $||\Lambda|| < 1$ is achieved might still cause severe numerical instabilities, such that there is the possibility that practically it is impossible to stabilise the scheme, depending on the particular set-up and operators under consideration.

5. Numerical Examples

Here, we present two numerical examples to demonstrate the theory presented in this work. Firstly we present a simple low-dimensional setup which confirms the theoretical results. Then, we investigate a more realistic system, the 2D Eady model [Ead49], which confirms the practical validity of the above results.

The numerics demonstrate that with a small regularization parameter α we can achieve a stable cycled data assimilation scheme. For the simple low-dimensional system we use construct M to be a random $n \times n$-matrix, using Matlab rand function, giving us a singular system by its SVD. To mimic a kind of trace-class operator, we then manipulated the singular values to decay sufficiently strongly.

The interesting case is where some singular values are larger than one, leading to a system with growing modes. Further, we want the rest of the singular values to be smaller than one and tending to zero for larger indices, as a model for a trace-class operator. For example, we pick M randomly and find the singular values

$$\mu(M) = (3.7568, 2.8065, 1.2662, 0.6557, 0.5563).$$

We then construct an observation operator, H in the same way with

$$\mu(H) = (3.1055, 2.5303, 1.7530, 0.0542, 10^{-12}).$$

Here we have growing and damping spectral modes in the model operator and we have simulated the consequence of a compact observation operator. The operator H is severely ill-conditioned with a condition number, $\kappa = 3.1049 \times 10^{12}$ with respect to the l^2 norm. We set up background and observation standard deviations as follows $\sigma_{(b)} = 0.09$ and $\sigma_{(o)} = 0.1$ respectively. Random normally distributed noise is added to the observations with mean, 0 and standard deviation, $\sigma_{(o)}$. Initially we choose $\alpha = \frac{\sigma^2_{(o)}}{\sigma^2_{(b)}} \approx 1.2$. In Figure 1 we observe that over time t_k the analysis error, $||e_k^{(a)}||_{l^2}$ blows up with respect to the l^2 norm. Now in Figure 2 we choose a smaller regularization parameter, $\alpha = \frac{0.1^2}{0.11^2} \approx 0.8$, inflating the background error variance, $\sigma^2_{(b)}$ from 0.09^2 to 0.11^2. Subsequently repeating with the same data, we observe a stable analysis error,

$||e_k^{(a)}||_{l^2}$ over time t_k with respect to the l^2 norm. In Figure 3 we plot $||e_{10000}^{(a)}||_{l^2}$ as α is varied. Here we observe for a fixed time t_{10000} that the analysis error is sensitively dependent on the regularization parameter we choose.

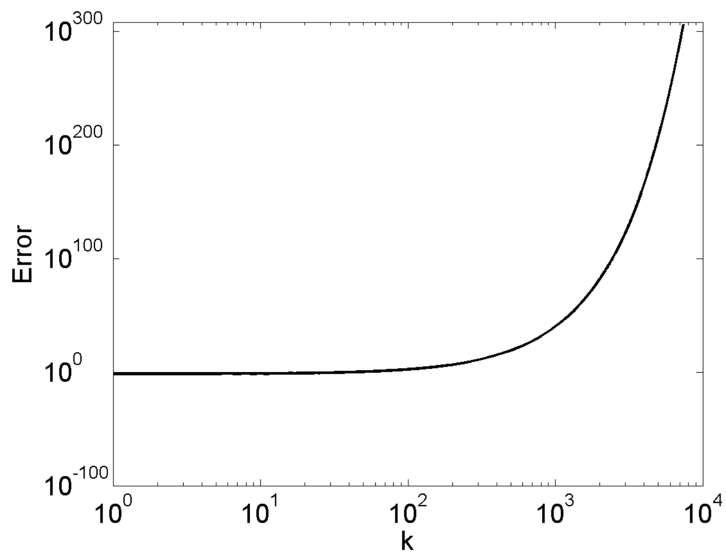


Figure 1. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\alpha = \frac{\sigma_{(o)}^2}{\sigma_{(b)}^2} \approx 1.2$.

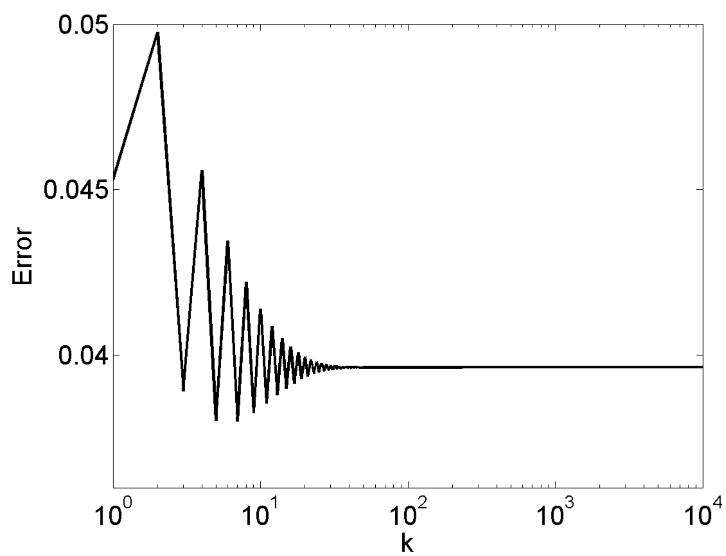


Figure 2. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\alpha \approx 0.8$.

As a second example, we now consider a more realistic system where the model operator M arises from the discretization of a system of partial differential equations.

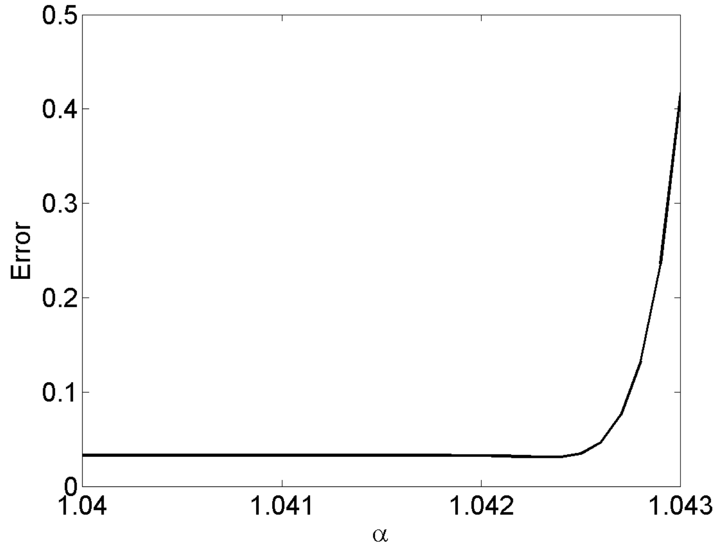


Figure 3. The l^2 norm of the analysis error as the scheme is cycled for a constant time index $k = 10000$, varying the regularization parameter, α.

The system we consider is the two-dimensional Eady model, a simple model of atmospheric instability, see [Ead49] for a detailed introduction. The model is defined in the $x - z$ plane, with periodic boundary conditions in x and $z \in [-1/2, 1/2]$. The state vector consists of the nondimensional buoyancy b on the upper and lower boundaries and the nondimensional potential vorticity in the interior of the domain. For the current study we assume that the interior potential vorticity is zero and thus we only need to consider the dynamics on the boundaries. The buoyancy is advected along the boundaries forced by the nondimensional streamfunction ψ according to the equation

$$\frac{\partial b}{\partial t} + z\frac{\partial b}{\partial x} = \frac{\partial \psi}{\partial x} \qquad \text{on } z = \pm 1/2 \tag{5.1}$$

where the streamfunction satisfies

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} = 0 \qquad \text{in } z \in [-1/2, 1/2], \tag{5.2}$$

with boundary conditions

$$\frac{\partial \psi}{\partial z} = b \qquad \text{on } z = \pm 1/2. \tag{5.3}$$

The equations are discretized as described in [Joh03] and [JHN05] using 40 grid points in the horizontal, giving 80 degrees of freedom. The resulting discrete operator M has a maximum eigenvalue of 1.3066.

We simulate the consequence of compact observation operator H with a random 80×80 matrix with the last 5 singular values $\mu_{76:80} = (10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}, 10^{-14})$ respectively. Therefore, H is severely ill-conditioned with a condition number, $\kappa = 4.1367 \times 10^{15}$ with respect to the l^2 norm. We set up background and observation standard deviations as follows $\sigma_{(b)} = 0.25$ and $\sigma_{(o)} = 1$ respectively. Random normally distributed noise is added to the observations with mean, 0 and standard deviation, $\sigma_{(o)}$. Initially we choose $\alpha = \frac{\sigma^2_{(o)}}{\sigma^2_{(b)}} \approx 16$. In Figure 4 we observe that over time t_k the analysis error, $||e_k^{(a)}||_{l^2}$ blows up with respect to the l^2 norm. Now in Figure 5 we choose a smaller regularization parameter, $\alpha = \frac{1}{0.09} \approx 11.1$, inflating the background error variance, $\sigma^2_{(b)}$ from 0.25^2 to 0.3^2. Subsequently repeating with the same data, we observe a stable analysis error, $||e_k^{(a)}||_{l^2}$ over time t_k with respect to the l^2 norm. In Figure 6 we plot $||e_{10000}^{(a)}||_{l^2}$ as α is varied. Here we observe for a fixed time t_{10000} that the analysis error is largely dependent on the regularization parameter we choose.

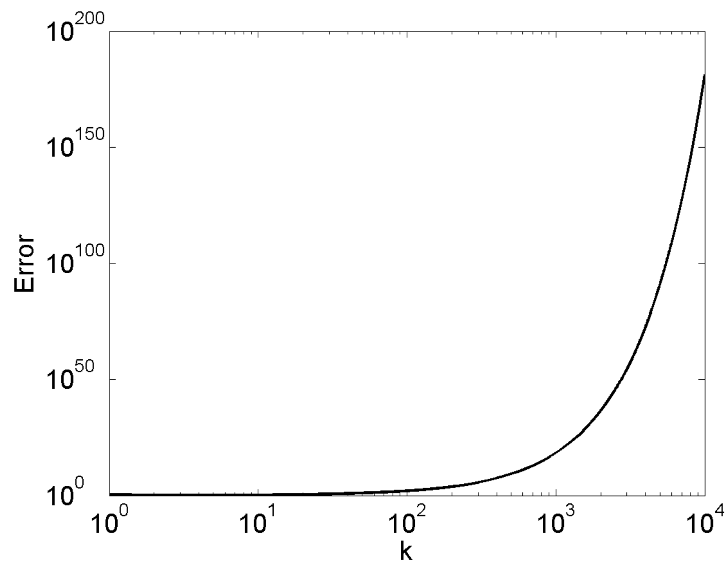


Figure 4. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\alpha = \frac{\sigma^2_{(o)}}{\sigma^2_{(b)}} \approx 16$.

6. Conclusions

The purpose of this work was to study the instability of cycled data assmilation algorithms, in particular when compact measurement operators are employed in large-dimensional (infinite dimensional) systems. We have investigated the long-term behaviour of the analysis error for the general setting of dynamical systems of trace-class, covering a wide range of such systems. The term *stability* is used when the error remains bounded over time. Here, we investigate several standard data assimilation schemes (3dVar, 4dVar and a cycled Kalman Filter), which for linear dynamical systems can be rewritten into a joint framework known as *cycled Tikhonov regularization.*

A key phenomen described and analysed is the dependence of the stability of the cycled data assimilation system on a *scaling parameter* α of the background covariance matrix. We show that when the scaling parameter is sufficiently small, we have boundedness of the analysis error in the long-term limit. If the scaling parameter is

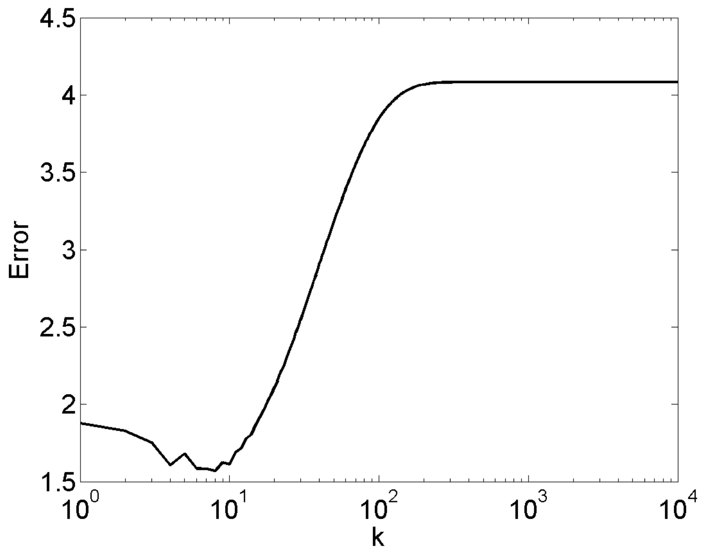


Figure 5. The l^2 norm of the analysis error as the scheme is cycled for the time index k, for a regularization parameter, $\alpha \approx 11.1$.

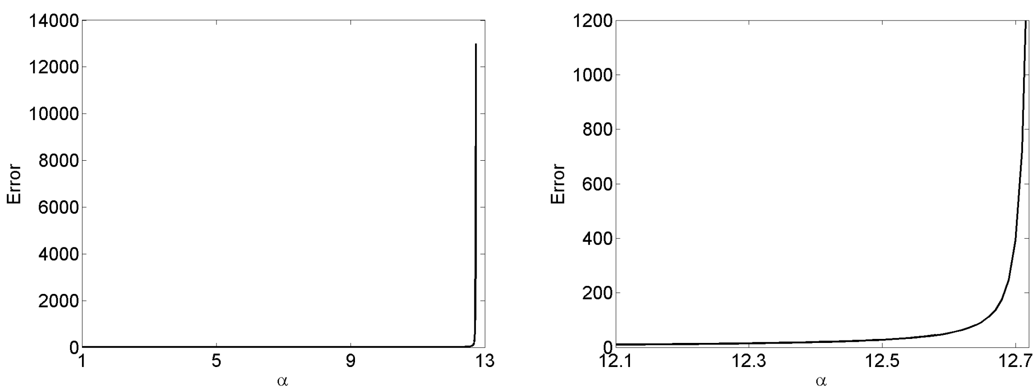


Figure 6. Left: The l^2 norm of the analysis error as the scheme is cycled for a constant time index $k = 10000$, varying the regularization parameter, α. Right: Zoomed version.

too large, error which enters the system via the data in general will be amplified, such that the analysis error growth without limits in its long-term behaviour. This growth happens even for an arbitrarily small data error.

In a numerical part we studied simple examples and also applied the theory to a two-dimensional Eady model, a simple model of atmospheric instability. The numerical results confirm and demonstrate the general theory very well.

Acknowledgements. The authors would like to thank the following institutions for their financial support: Deutscher Wetterdienst (DWD), Engineering and Physical Sciences Research Council (EPSRC) and the National Centre for Earth Observation (NCEO, NERC). The Eady model used in Section 5 was developed by C. Johnson, now at the Met Office, and N.K. Nichols and B. J. Hoskins of the University of Reading. We are also grateful to N.K. Nichols for useful discussions during this work.

References

[AM79] B.D.O. Anderson and J.B. Moore. *Optimal Filtering.* Prentice Hall, 1979.

[BC85] S. Barnett and R.G. Cameron. *Introduction to mathematical control theory.* Oxford applied mathematics and computing science series. Clarendon Press, 1985.

[BH75] A.E. Bryson and Y.C. Ho. *Applied optimal control: optimization, estimation, and control.* Taylor & Francis, 1975.

[BLL+12] C.E.A. Brett, K.F. Lam, K.J.H. Law, D.S. McCormick, M.R. Scott, and A.M. Stuart. Stability of filters for the navier-stokes equation. Submitted, 2012.

[But69] A.G. Butkovskii. *Theory of optimal control of distributed parameter systems.* American Elsevier, 1969.

[CF03] M.J. Corless and A.E. Frazho. *Linear systems and control: an operator perspective.* Monographs and textbooks in pure and applied mathematics. Marcel Dekker, 2003.

[CP78] R.F. Curtain and A. J. Pritchard. *Infinite dimensional linear systems theory.* Lecture notes in control and information sciences. Springer-Verlag, 1978.

[Cul12] M. J. P. Cullen. Analysis of cycled 4d-var with model error. Submitted, 2012.

[DP00] G.E. Dullerud and F.G. Paganini. *A course in robust control theory: a convex approach.* Texts in applied mathematics. Springer, 2000.

[Dul96] G.E. Dullerud. *Control of uncertain sampled-data systems.* Systems & control. Birkhäuser, 1996.

[Ead49] E. T. Eady. Long waves and cyclone waves. *Tellus*, 1:33–52, 1949.

[EHN00] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems.* Mathematics and its applications. Kluwer Academic Publishers, 2000.

[Jaz70] A. H. Jazwinski. *Stochastic processes and filtering theory.* Academic Press, 1970.

[JHN05] C. Johnson, B. J. Hoskins, and N. K. Nichols. A singular vector perspective of 4d-var: Filtering and interpolation. *Q. J. R. Meteorol. Soc.*, 131:1–19, 2005.

[Joh03] C. Johnson. *Information content of observations in variational data assimilation.* PhD thesis, Dept. of Meteorology, University of Reading, 2003.

[Kal60] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, (82 (Series D)):35–45, 1960.

[Kna05] A.W. Knapp. *Advanced real analysis.* Cornerstones Series. Birkhäuser, 2005.

[KS04] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems (Applied Mathematical Sciences).* Springer, 1st edition, December 2004.

[Lio71] J.L. Lions. *Optimal control of systems described by partial differential equations.* Springer, 1971.

[LLD06] J. M. Lewis, S. Lakshmivarahan, and S. Dhall. *Dynamic Data Assimilation: A Least Squares Approach.* Cambridge University Press, August 2006.

[Lue64] D. G. Luenberger. Observing the state of a linear system. *Military Electronics, IEEE Transactions on*, 8(2):74 –80, 1964.

[Lue66] D. G. Luenberger. Observers for multivariable systems. *Automatic Control, IEEE Transactions on*, 11(2):190 – 197, 1966.

[MP12] B. A. Marx and R. W. E. Potthast. On instabilites in data assimilation algorithms. 2012.

[O'R83] J. O'Reilly. *Observers for linear systems.* Mathematics in science and engineering. Academic Press, 1983.

[Rug96] W.J. Rugh. *Linear system theory.* Prentice-Hall information and system sciences series. Prentice Hall, 1996.

[Sim06] D. Simon. *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches.* Wiley-Interscience, 2006.