School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

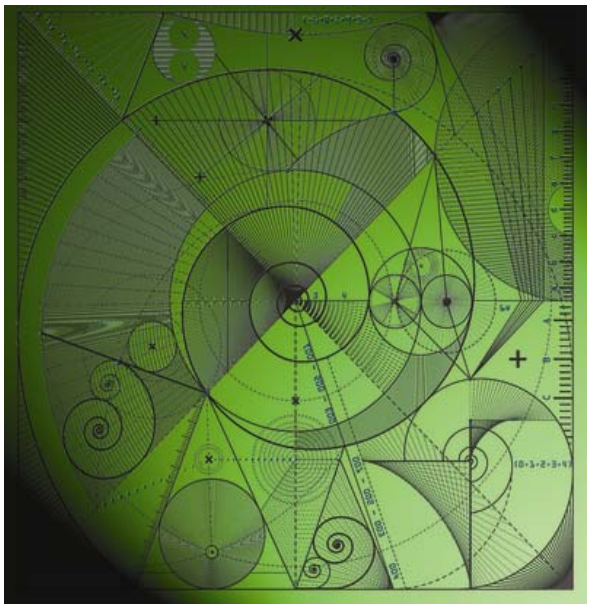
Preprint MPS-2012-06

28 February 2012

On Instabilities in Data Assimilation Algorithms

by

Boris A. Marx and Roland W.E. Potthast



On Instabilities in Data Assimilation Algorithms

Boris A. Marx[1] **and Roland W.E. Potthast**[2,3]

[1] University of Göttingen, Germany, Graduiertenkolleg 1023 "'Identification in Mathematical Models: Synergy of Stochastic and Numerical Methods"'
[2] German Meteorological Service - Deutscher Wetterdienst Research and Development, Frankfurter Strasse 135, 63067 Offenbach, Germany
[3] University of Reading, Department of Mathematics, Whiteknights, PO Box 220, Reading RG6 6AX, United Kingdom

Abstract. Data assimilation algorithms are a crucial part of operational systems in numerical weather prediction, hydrology and climate science, but are also important for dynamical reconstruction in medical applications and quality control for manufactoring processes. Usually, a variety of diverse measurement data are employed to determine the state of the atmosphere or to a wider system including land and oceans. Modern data assimilation systems use more and more remote sensing data, in particular radiances measured by satellites, radar data and integrated water vapor measurements via GPS/GNSS signals. The inversion of some of these measurements are ill-posed in the classical sense, i.e. the inverse of the operator H which maps the state onto the data is unbounded. In this case, the use of such data can lead to significant instabilities of data assimilation algorithms.

The goal of this work is to provide a rigorous mathematical analysis of the instability of well-known data assimilation methods. Here, we will restrict our attention to particular linear systems, in which the instability can be explicitly analyzed. We investigate the three-dimensional variational assimilation and four-dimensional variational assimilation. A theory for the instability is developed using the classical theory of ill-posed problems in a Banach space framework. Further, we demonstrate by numerical examples that instabilities can and will occur, including an example from dynamic magnetic tomography.

1. Introduction

Data assimilation algorithms in combination with the reconstruction of quantities from remote sensing data are important in many areas like *numerical weather prediction* [War11], *oceanic and hydrologic applications* [PX09] as well as *process tomography* (for example using *magnetic tomography* [KKP02], [HKP05], [HPWS05], [HP07], [HPW08], [PW09], [Wan09]) and in cognitive neuroscience [PbG09].

Today, there is a variety of algorithms for data assimilation. There are classical variational approaches (compare [LLD06], [LSK10]) known as three-dimensional variational assimilation (3dVar) and four-dimensional variational assimilation (4dVar). Many data assimilation schemes can be considered as special cases of statistical inversion methods [Jaz70], [KS04], i.e. Bayesian estimation for the case of linear systems [RL00].

In the case of Gaussian errors and Gaussian background distributions for linear systems Bayes' formula leads to the well known *Kalman Filter* [Kal60].

Here, we study data assimilation algorithms as *iterative* or *cycled* schemes. Our viewpoint is driven by the field of inverse problems, which rather than focusing on stochastical properties of the fields under consideration, has placed strong emphasis on the analysis of the ill-posedness of the reconstruction problem.

In fact, using remote sensing data we will find severe instabilities of these algorithms. The origin of this instable behavior may either be due to the nonlinear or chaotic systems dynamics as for examle presented in [CGTU08] or because of the ill-posedness of the observation operator H. The second cause of instability is our key concern in this work.

To provide a thorough analysis we investigate a *dynamic (or cycled) Tikhonov regularization* scheme (compare also [Mar11]), which is a Tikhonov regularization with a dynamic background which is updated by propagating the analysis of the previous step to the next point in time where data are provided. We use the Hilbert spaces X, Y and the discrete time-slices

$$t_0 < t_1 < t_2 < ... < t_k < \dots \tag{1.1}$$

where we consider the system states $x_k \in X$ at time t_k and the corresponding data $y_k \in Y$. They are given via the measurement operator $H : X \to Y$ by

$$y_k = y_k^{(true)} + y^{(\delta)} = Hx_k^{(true)} + y^{(\delta)}, \quad k \in \mathbb{N} \tag{1.2}$$

Let our measurement operator H be *compact* and linear and X be of *infinite dimension*. Then, H cannot have a bounded inverse H^{-1} (c.f. [Kre99]). We need to employ regularization which is implicitly carried out by classical variational data assimilation schemes minimizing for example

$$J(x) := \alpha||x - x_k^{(b)}||^2 + ||y_k - Hx_k^{(b)}||^2 \tag{1.3}$$

by the background term $\alpha||x - x_k^{(b)}||^2$ given some state $x_k^{(b)}$ which is usually denoted as *the background* at time t_k. The minimizer $x_k^{(a)}$ of (1.3) is called *the analysis* at time t_k. The propagation of the analysis $x_k^{(a)}$ at time t_k to the time t_{k+1} is carried out using the model operator $M_{k+1|k} : X \to X$ by

$$x_{k+1}^{(b)} = M_{k+1|k}x_k^{(a)}. \tag{1.4}$$

The equations (1.3) and (1.4) define the *cycled Tikhonov regularization*.

Our plan is as follows. In Section 2 we describe the setup of our systems. We will base our arguments on the fact that the classical *three-dimensional variational assimilation* (3dVar) and the *fourdimensional variational assimilation* (4dVar) for linear systems can be rewritten as a cycled Tikhonov regularization when weighted norms are introduced. So we can restrict our analysis to the case (1.3) - (1.4) and obtain results which hold for key classical data assimilation methods.

In Section 3 we study the asymptotic behavior of $x_k^{(a)}$ for $k \to \infty$ when particular model systems M are chosen for which we can explicitly carry out the asymptotic

analysis. These systems serve as key examples to prepare the investigation of more complex nonlinear systems. In particular, we will study a constant system and spectrally expanding or collapsing systems. We will show by functional analytic tools that for compact observation operators H and data $y^{(\delta)} \notin H(X)$, the analysis will diverge for $k \to \infty$.

Further, we provide a spectral view on the evolution in Section 3.2, for which we provide explicit evolution formulas for the spectral coefficients of the analysis with respect to the singular system of the observation operator. These formulas are also provided for spectrally expanding or collapsing systems with a system dynamics which is diagonal in the above singular system.

For the particular spectrally collapsing systems we show divergence of the relative error when constant error terms are used in Section 3.3. For spectrally expanding systems the relative error will tend to zero and convergence of the analysis towards the dynamic true solution is shown.

The final Section 4 shows numerical examples for the asymptotic behavior. In particular, we demonstrate the behavior for a spectral system where we have relatively fast convergence of the spectral coefficients $a_{n,k}^{(a)}$ of the analysis $x_k^{(a)}$ for small modes and relatively slow convergence of the spectral analysis coefficients for large modes. This leads to a particular behavior where we have a reduction of the analysis error $||x_k^{(a)} - x_k^{(true)}||$ over some time, until the contribution of larger modes and the condition $y^{(\delta)} \notin H(X)$ takes over and leads to a divergence of the analysis. We finally provide a practical example from dynamic magnetic tomography which confirms this behavior.

2. System and Assimilation Setup

In this section we first introduce the model system setup which we will use as basis for the subsequent instability study and then summarize the data assimilation algorithms under consideration. We will also summarize the basic arguments why for linear model dynamics both 3dVar and 4dVar can be considered as a cycled Tikhonov regularization in a space with weighted norms, such that our analysis is valid for these assimilation approaches.

2.1. Constant, spectrally expanding or collapsing systems.

Here, we will put a particular emphasis on situations where the phenomena can be explicitly calculated. These are in particular

a) *Constant systems.* These are systems where $M_{k+1|k}$ is the identity operator. Clearly, constant dynamics is only a limiting case for very short time scales or slowly changing situations. But the phenomena which appear here arising from ill-posed observation operators will also be visible when the dynamics is no longer constant. Also, the development of errors and iterates can be explicitly calculated.

b) *Spectrally diagonal systems*, i.e. systems for which the dynamics is given by

$$M_{k+1|k} = U^{-1}DU \tag{2.1}$$

with some orthonormal transformation $U : X \to X$ and a diagonal mapping D. We call a mapping diagonal in an infinite dimensional space, when there is some orthonormal system $\{\varphi_n : n \in \mathbb{N}\}$ such that

$$D\varphi_n = d_n\varphi_n, \quad n \in \mathbb{N}. \tag{2.2}$$

c) *Spectrally expanding systems.* In the case where all d_n satisfy

$$|d_n| \geq q > 1, \quad n \in \mathbb{N}, \tag{2.3}$$

with some constant q we call the system *expanding*.

d) *Spectrally collapsing systems.* These are systems with a representation of type (2.1), (2.2) for which

$$|d_n| \leq q < 1, \quad n \in \mathbb{N} \tag{2.4}$$

with some constant q is satisfied.

We will also be interested in systems for which the columns of U are the singular vectors of the measurement operator H, i.e. $H^*H\varphi_n = \mu_n^2\varphi_n$ for $n \in \mathbb{N}$. In this case, the complete dynamics of the data assimilation system can be diagonalized by the singular vectors of H.

2.2. Cycled Regularization

Tikhonov regularization is a well-known scheme to solve an ill-posed operator equation of the type $Hx = y$ by minimization of the cost functional

$$J(x) = ||Hx - y||^2 + \alpha||x||^2. \tag{2.5}$$

To solve such a data equation on successive points in time t_k, we employ a modified regularization term using the *background*

$$x_{k+1}^{(b)} = M_{k+1|k}x_k^{(a)}, \quad k \in \mathbb{N} \tag{2.6}$$

which is calculated by propagating the *analysis* of the kth step by the model to the time t_{k+1}. An iterated Tikhonov scheme was analyzed in [Eng87] to find an optimal regularization parameter for this iterated scheme. This is a special case of our cycled scheme for a constant background. The cost functional for *cycled Tikhonov regularization* is then

$$J_{Tik}(x) = ||y_{k+1} - Hx||^2 + \alpha||x - x_{k+1}^{(b)}||^2 \tag{2.7}$$

given measurements y_k, $k = 1, 2, 3, ...$ at time t_k. The minimum of (2.7) is denoted as $x_{k+1}^{(a)}$ (called *the analysis*). It is well-known that the normal equations for the quadratic functional (2.7) lead to the update formula

$$x_{k+1}^{(a)} = x_{k+1}^{(b)} + R_\alpha\left(y_{k+1} - Hx_{k+1}^{(b)}\right), \quad k = 0, 1, 2, ... \tag{2.8}$$

with

$$R_\alpha = (\alpha I + H^* H)^{-1} H^*. \tag{2.9}$$

starting from some initial state $x_0^{(a)}$. The parameter $\alpha > 0$ is denoted as regularization parameter (compare [CK97] Theorem 4.14). We call (2.8) together with (2.6) the *cycled Tikhonov regularization.*

2.3. Three-dimensional Variational Assimilation (3dVar).

Three-dimensional variational data assimilation (3dVar) is basically a cycled Tikhonov scheme with weighted norms. For this section let us restrict our arguments to the n-dimensional case where $X = R^n$ with some $n \in \mathbb{N}$ and $Y = \mathbb{R}^m$, $m \in \mathbb{N}$. For some symmetric positive definite matrix Γ we define the *weighted scalar product*

$$\langle \cdot, \cdot \rangle_\Gamma := \langle \cdot \, , \Gamma \cdot \rangle \, , \tag{2.10}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard L^2 scalar product in $\mathbb{R}^n$. Given a linear operator $H : X \to Y$, we denote the adjoint operator with respect to the standard scalar products by H' and the adjoint with respect to the weighted scalar products by H^*.

Let B be the covariance matrix for the background system and R the covariance matrix for the measurements. Then the cost function for the *cycled 3dVar* scheme (compare [LLD06] Chapter 20) is given by

$$J_{3dVar}(x) = ||x - x_{k+1}^{(b)}||_{B^{-1}}^2 + ||y_{k+1} - Hx||_{R^{-1}}^2, \tag{2.11}$$

which is minimized in each step by the analysis

$$x_{k+1}^{(a)} = x_{k+1}^{(b)} + K_{k+1} \left(y_k - H x_{k+1}^{(b)} \right) \tag{2.12}$$

with the Kalman gain

$$K_{k+1} = BH' \left(HBH' + R \right)^{-1}. \tag{2.13}$$

THEOREM 2.1 *For the common L^2 scalar product $\langle \cdot \, , \cdot \rangle$ and the weighted norms (2.10)*

$$\langle \cdot \, , \cdot \rangle_B := \left\langle \cdot \, , B^{-1} \cdot \right\rangle \text{ on } X, \qquad \langle \cdot \, , \cdot \rangle_R := \left\langle \cdot \, , R^{-1} \cdot \right\rangle \text{ on } Y \tag{2.14}$$

with self-adjoint, positive definite matrices B and R the Kalman gain (2.13) corresponds to the Tikhonov projection (2.9) with the adjoint H^* with respect to the weighted scalar product $\langle \cdot \, , \cdot \rangle_B$, where the adjoint is given by

$$H^* = BH'R^{-1} \tag{2.15}$$

Proof. The particular form (2.15) is obtained from

$$\begin{aligned} \langle \varphi \, , H\psi \rangle_R &= \left\langle \varphi \, , R^{-1} H \psi \right\rangle \\ &= \left\langle H'R^{-1}\varphi \, , \psi \right\rangle \\ &= \left\langle H'R^{-1}\varphi \, , BB^{-1}\psi \right\rangle \\ &= \left\langle BH'R^{-1}\varphi \, , B^{-1}\psi \right\rangle \\ &= \left\langle BH'R^{-1}\varphi \, , \psi \right\rangle_B . \end{aligned} \tag{2.16}$$

We now study the relation between the Kalman gain (2.13) and the Tikhonov projection (2.9). Using

$$H^* \left(\alpha I + HH^*\right) = \left(\alpha I + H^* H\right) H^* \tag{2.17}$$

which leads to

$$\left(\alpha I + H^* H\right)^{-1} H^* = H^* \left(\alpha I + HH^*\right)^{-1} \tag{2.18}$$

we transform the Kalman gain (2.13) by

$$\begin{aligned} BH' \left(HBH' + R\right)^{-1} &= BH' R^{-1} \left(HBH' R^{-1} + I\right)^{-1} \\ &= H^* \left(HH^* + I\right)^{-1} \\ &= \left(I + H^* H\right)^{-1} H^* \end{aligned} \tag{2.19}$$

which is the Tikhonov operator (2.9) for $\alpha = 1$. □

2.4. Four-dimensional Variational Assimilation (4dVar)

The 4dVar scheme minimizes a cost functional of the form

$$J_{4dVar}(x) := ||x - x_0||^2_{B^{-1}} + \sum_{k=1}^{K} ||y_k - GM_{k|0} x||^2_{R_k^{-1}}, \tag{2.20}$$

where we now call the observation operator G for reasons which will be clear immediately. For linear systems M it can be rewritten as the 3dVar scheme by redefining the variables as follows. We define the vectorial observation space $\underline{Y}$ by

$$\underline{Y} := \underbrace{Y \times \ldots \times Y}_{K \text{ times}} \tag{2.21}$$

and a vectorial observation operator

$$H := \begin{pmatrix} GM_{1|0} \\ \vdots \\ GM_{K|0} \end{pmatrix} \tag{2.22}$$

and

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_K \end{pmatrix} \tag{2.23}$$

as well as the block diagonal covariance matrix for the measurements

$$R = \begin{pmatrix} R_1 & & \\ & \ddots & \\ & & R_K \end{pmatrix} \tag{2.24}$$

with block diagonal entries. A norm in $\underline{Y}$ is given by

$$||y||^2_{R^{-1}} = ||y_1||^2_{R_1^{-1}} + \ldots + ||y_K||^2_{R_K^{-1}} \tag{2.25}$$

with $y_j \in Y$. Then we can write (2.20) as

$$J_{4dVar}(x) := ||x - x_0||^2_{B^{-1}} + ||y - Hx||^2_{R^{-1}}, \tag{2.26}$$

which corresponds to (2.11) for one iteration step. Thus, 4dVar for linear systems can be written as a 3dVar algorithm and in a second step as a cycled Tikhonov regularization, such that results for cycled Tikhonov regularization apply both 3dVar and 4dVar.

3. Instability of Variational Data Assimilation

The goal of this section is to carry out the stability analysis for cycled variational assimilation algorithms with ill-posed observation operators, in particular for a cycled Tikhonov regularization (which according to the above arguments then also hold for 3dVar and 4dVar). We will show that all cycled assimilation schemes can exhibit strong instabilities when remote sensing operators are involved.

3.1. Cycled Tikhonov Regularization

We have shown that for linear models choosing appropriate norms the cycled 3dVar or 4dVar can be written as a cycled Tikhonov regularization. Thus, without loss of generality, we can restrict our attention to the update formula (2.8). We denote the true system by $x_k^{(true)}$ with true data $y_k^{(true)} = Hx_k^{(true)}$. The real measured data is assumed to be of the form

$$y_k = Hx_k^{(true)} + y_k^{(\delta)}, \quad k \in \mathbb{N}. \tag{3.1}$$

Further, for simplicity we study a uniform time grid and use M for the system dynamics from one time step to the next. In this case, we transform the update formula (2.8) into

$$\begin{aligned} x_{k+1}^{(a)} &= x_{k+1}^{(b)} + R_\alpha\Big(y_{k+1} - Hx_{k+1}^{(b)}\Big) \\ &= Mx_k^{(a)} + R_\alpha H\Big(Mx_k^{(true)} - Mx_k^{(a)}\Big) + R_\alpha y_{k+1}^{(\delta)} \\ &= (I - R_\alpha H)Mx_k^{(a)} + R_\alpha HMx_k^{(true)} + R_\alpha y_{k+1}^{(\delta)}. \end{aligned} \tag{3.2}$$

Subtracting the true solution $x_{k+1}^{(true)}$ on both sides leads to

$$x_{k+1}^{(a)} - x_{k+1}^{(true)} = (I - R_\alpha H)M\Big(x_k^{(a)} - x_k^{(true)}\Big) + R_\alpha y_{k+1}^{(\delta)} \tag{3.3}$$

for $k = 0, 1, 2, ...$ The formula (3.3) is a general formula which holds for a wide range of systems. The case where we have the same type of error $y^{(\delta)} = y_k^{(\delta)}$ for all steps t_k, $k = 1, 2, 3, ...$ constitutes a kind of *worst case* scenario, as we will see in more detail below.

We denote the *error* between the analysis and true solution at time t_k as e_k. Then, from (3.3) we obtain the error evolution

$$e_{k+1} = \Lambda e_k + S, \quad k = 0, 1, 2, ..., \tag{3.4}$$

where we abbreviate

$$\Lambda := (I - R_\alpha H)M, \quad S := R_\alpha y^{(\delta)}. \tag{3.5}$$

The evolution of the error governed by the formula (3.4) can be carried out by an induction proof as follows.

LEMMA 3.1 *Consider the evolution of some term $e_k \in X$, $k = 0, 1, 2, ...$, in some Banach space X given by (3.4) with initial value e_0. Then, we obtain*

$$e_k = \Lambda^k e_0 + \Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S, \quad k \in \mathbb{N}. \tag{3.6}$$

If $(I - \Lambda)$ is invertible in X, then the formula can be transformed into

$$e_k = \Lambda^k + (1 - \Lambda)^{-1}(1 - \Lambda^k)S, \quad k \in \mathbb{N}. \tag{3.7}$$

Proof. For $k = 1$ the formula (3.6) is clearly identical to the general update formula (3.4). Assume that it is true for some $k \in \mathbb{N}$, then we calculate

$$\begin{aligned} e_{k+1} &= \Lambda e_k + S \\ &= \Lambda\Big(\Lambda^k e_0 + \Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S\Big) + S \\ &= \Lambda^{k+1} e_0 + \Big(\sum_{\xi=0}^{k} \Lambda^\xi\Big) S, \end{aligned} \tag{3.8}$$

which is the formula for $k+1$ and proves (3.6). If $I-\Lambda$ is invertible, we use the telescopic sum

$$(I - \Lambda) \sum_{\xi=0}^{k-1} \Lambda^\xi = I - \Lambda^k \tag{3.9}$$

and multiply by $(I - \Lambda)^{-1}$ from the left to arrive at (3.7), which ends the proof. □

As a consequence of the previous lemma we obtain the asymptotic behavior of the error for cycled assimilation schemes. Here, we first study the case where $M = I$, i.e. the model dynamics is the identity, as a basic reference situation. If things are instable for this case, we cannot expect them to be stable in a more general setup. Here, we have

$$\Lambda = I - R_\alpha H, \quad I - \Lambda = R_\alpha H, \quad S = R_\alpha y^{(\delta)}, \tag{3.10}$$

leading to

$$\begin{aligned} \Lambda &= I - R_\alpha H = I - (\alpha I + H^* H)^{-1} H^* H \\ &= \alpha(\alpha I + H^* H)^{-1} = (I + \alpha^{-1} H^* H)^{-1}. \end{aligned} \tag{3.11}$$

Clearly, the asymptotic behavior of e_k for $k \to \infty$ is governed by the behavior of Λ^k for $k \to \infty$. We carry out the convergence analysis in the following theorem.

THEOREM 3.2 *Assume that $H : X \to Y$ is a compact injective operator from a Hilbert space X into a Hilbert space Y. Then we have the* pointwise convergence

$$\left(I + \alpha^{-1} H^* H\right)^{-k} \varphi \to 0, \quad k \to \infty \tag{3.12}$$

for every fixed element $\varphi \in X$.

Proof. To show the convergence (3.12) we use $\alpha = 1$ to keep things readable, the general case is carried out analogously. We have

$$\begin{aligned} \left\langle (I + H^*H)\psi, (I + H^*H)\psi \right\rangle &= ||\psi||^2 + 2||H\psi||^2 + ||H^*H\psi||^2 \\ &> ||\psi||^2 \end{aligned} \tag{3.13}$$

if $H\psi \neq 0$. If we use $\psi = (I + H^*H)^{-1}\varphi$, we estimate

$$||(I + H^*H)^{-1}\varphi|| < ||\varphi||. \tag{3.14}$$

This observation can be generalized. We calculate

$$\left\langle (I + H^*H)^k \psi, (I + H^*H)^k \psi \right\rangle = ||\psi||^2 + 2k||H\psi||^2 + U_k \tag{3.15}$$

with some terms U_k for which we show later that they are positive or zero. As a consequence we obtain

$$||(I + H^*H)^{-k}\varphi||^2 + 2k||H(I + H^*H)^{-k}\varphi||^2 < ||\varphi||^2. \tag{3.16}$$

Now, we know that $\lambda_k := (I + H^*H)^{-k}\varphi$ is bounded in X. Assume that it is not convergent towards zero. Then there is a weakly convergent subsequence $(\lambda_{k_j})_{j\in\mathbb{N}}$, which tends weakly towards $\psi \in X$, $\psi \neq 0$. But now $H\lambda_{k_j} \to H\psi$ in X. From (3.16) we then obtain $H\psi = 0$ and thus $\psi = 0$, which is a contradiction to our assumption and shows that (3.12) must be satisfied.

Finally, we need to show that the terms U_k are positive. We use the binomial formula to calculate

$$\begin{aligned} &\left\langle (I + H^*H)^k \psi, (I + H^*H)^k \psi \right\rangle \\ &= \left\langle \sum_{\xi=0}^{k} \binom{k}{\xi} (H^*H)^\xi \psi, \sum_{\eta=0}^{k} \binom{k}{\eta} (H^*H)^\eta \psi \right\rangle \\ &= \sum_{\xi=0}^{k} \sum_{\eta=0}^{k} \binom{k}{\xi} \binom{k}{\eta} \left\langle (H^*H)^\xi \psi, (H^*H)^\eta \psi \right\rangle. \end{aligned} \tag{3.17}$$

Each scalar product in (3.17) can be seen to be positive by transforming them into terms of the form

$$\left\langle (H^*H)^l \psi, (H^*H)^l \psi \right\rangle = ||(H^*H)^l \psi||^2 \tag{3.18}$$

with $l = (\xi + \eta)/2$ for $\xi + \eta$ even or

$$\left\langle (H^*H)^{l+1} \psi, (H^*H)^l \psi \right\rangle = ||H(H^*H)^l \psi||^2 \tag{3.19}$$

with $l = (\xi + \eta - 1)/2$ for $\xi + \eta$ odd, and the proof is complete. □

For H injective we have shown the pointwise convergence $\Lambda^k e_0 \to 0$ for $k \to \infty$ for any $e_0 \in X$. To fully study the behaviour of e_k we further need to calculate the second term in (3.6) or (3.7), respectively. We need to investigate

$$(I - \Lambda)^{-1} S = (R_\alpha H)^{-1} R_\alpha y^{(\delta)}. \tag{3.20}$$

Under the condition that H^*H is invertible, we calculate

$$\begin{aligned}(R_\alpha H)^{-1} R_\alpha &= \Big((\alpha I + H^*H)^{-1} H^*H\Big)^{-1} (\alpha I + H^*H)^{-1} H^* \\ &= (H^*H)^{-1} H^*, \end{aligned} \tag{3.21}$$

which is the Moore-Penrose pseudo inverse. Since H and H^* are compact, in general invertibility is not given, but we obtain invertibility of $R_\alpha H$ on the subspace

$$Z := (\alpha I + H^*H)^{-1} H^*H(X) \subset Y.$$

We capture the phenomena in the following lemma.

LEMMA 3.3 *Assume that H and H^* are injective. If $y^{(\delta)} \in H(X)$, i.e. there is $x^{(\delta)} \in X$ such that $y^{(\delta)} = Hx^{(\delta)}$, then with $S = R_\alpha y^{(\delta)}$ we obtain*

$$\Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S \to x^{(\delta)}, \quad k \to \infty. \tag{3.22}$$

In the case where $y^{(\delta)} \notin H(X)$, we have

$$\Big|\Big|\Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S\Big|\Big| \to \infty, \quad k \to \infty. \tag{3.23}$$

Proof. In the case $y^{(\delta)} \in H(X)$ we have

$$(I - \Lambda)^{-1} S = (R_\alpha H)^{-1} R_\alpha H x^{(\delta)} = x^{(\delta)}.$$

We use Theorem 3.2 for injective H^*H to derive

$$(I - \Lambda^k)(I - \Lambda)^{-1} S = (I - \Lambda^k) x^{(\delta)} \to x^{(\delta)}. \tag{3.24}$$

Using (3.9) and noting that the terms under consideration commute this yields

$$\Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S \to x^{(\delta)}, \tag{3.25}$$

which shows (3.22). To prove the second statement assume that

$$\lambda_k := \Big(\sum_{\xi=0}^{k-1} \Lambda^\xi\Big) S, \quad k \in \mathbb{N}, \tag{3.26}$$

is bounded in X. Then, there is a weakly convergent subsequence and an element $\psi \in X$ such that $\lambda_{k_j} \rightharpoonup \psi$ for $j \to \infty$. Since H is compact, it maps the weakly

convergent sequence into a strongly convergent sequence, i.e. $H\lambda_{k_j} \to H\psi$. We further have the convergence

$$\begin{aligned}(R_\alpha H)\Big(\sum_{\xi=0}^{k_j-1}\Lambda^\xi\Big)S &= (I-\Lambda)\Big(\sum_{\xi=0}^{k_j-1}\Lambda^\xi\Big)S \\ &= (I-\Lambda^{k_j})S \\ &\to R_\alpha y^{(\delta)}, \quad j\to\infty. \end{aligned} \tag{3.27}$$

and

$$\begin{aligned}(R_\alpha H)\Big(\sum_{\xi=0}^{k_j-1}\Lambda^\xi\Big)S &= R_\alpha H\lambda_{k_j} \\ &\to R_\alpha H\psi, \quad j\to\infty. \end{aligned} \tag{3.28}$$

Since R_α is boundedly invertible, from (3.27) and (3.28) we obtain the identity $H\psi = y^{(\delta)}$, i.e. $y^{(\delta)}$ is in the images space $H(X)$. This shows that under the condition $y^{(\delta)} \notin H(X)$ the sequence $(\lambda_k)_{k\in\mathbb{N}}$ cannot be uniformly bounded and the proof is complete. □

Finally, we summarize our results and apply them to variational data assimilation schemes as a corollary.

Corollary 3.4 (Instability of Assimilation Schemes) *The cycled Tikhonov regularization and thus also the data assimilation schemes 3dVar and 4dVar for a constant model M and an injective compact observation operators are unstable when we feed in data $y_k = Hx_k^{(true)} + y^{(\delta)}$, where $y^{(\delta)} \notin R(H)$. If we have $y^{(\delta)} = 0$, we obtain convergence towards the true solution $x^{(true)}$. If $y^{(\delta)} = Hx^{(\delta)}$, we have convergence towards $x^{(true)} + x^{(\delta)}$.*

3.2. Spectral Cycled Tikhonov Scheme

The spectral approach (compare e.g. [CK97], Chapter 4.3) applied to the cycled data assimilation provides further insight. We employ a singular system (φ_n, g_n, μ_n) with an orthonormal basis φ_n in X and an orthonormal basis g_n in Y such that

$$H\varphi_n = \mu_n g_n, \quad H^* g_n = \mu_n \varphi_n. \tag{3.29}$$

For a state vector $x \in X$ we denote the coefficients for the singular value decomposition by $a_n = \langle x, \varphi_n\rangle_X$. The coefficients of a true solution x_{true} are defined by

$$x_{true} = \sum_{n=1}^{\infty} a_n^{(true)}\varphi_n. \tag{3.30}$$

To include the temporal dimension we add the index k as in (2.6), i.e. we write

$$x_k = \sum_{n=0}^{\infty} a_{n,k}\varphi_{n,k} \qquad k = 0, 1, \ldots \tag{3.31}$$

with coefficients $a_{n,k}$. The coefficients of the data vectors y_k at time t_k are defined by

$$y_k = \sum_{n=0}^{\infty} b_{n,k} g_n, \qquad k = 1, 2, \ldots \tag{3.32}$$

According to (3.29) an application of the operators H or H^* corresponds to a spectral multiplication by μ_n, $n \in \mathbb{N}$ and an application of H^{-1} corresponds to $1/\mu_n$, $n \in \mathbb{N}$. We use the notation

$$q_n := \frac{\alpha}{\mu_n^2 + \alpha}, \qquad 1 - q_n = \frac{\mu_n^2}{\mu_n^2 + \alpha}, \quad n \in \mathbb{N}. \tag{3.33}$$

Then, in spectral terms we carry out the application of the following operators by spectral multiplication in the sense of (3.29). For $R_\alpha = (I + H^*H)^{-1}H^*$ we spectrally multiply by

$$\frac{\mu_n}{\alpha + \mu_n^2} = \frac{(1 - q_n)}{\mu_n}, \quad n \in \mathbb{N}.$$

In general, we have the spectral operations

$$\begin{array}{rrl} & H: & a_n \mapsto \mu_n a_n \\ & R_\alpha: & b_n \mapsto \frac{1-q_n}{\mu_n} b_n \\ I - \Lambda = & R_\alpha H: & a_n \mapsto (1 - q_n) a_n \\ \Lambda = & I - R_\alpha H: & a_n \mapsto q_n a_n \\ \Lambda^k = & (I - R_\alpha H)^k: & a_n \mapsto (q_n)^k a_n. \end{array} \tag{3.34}$$

With this, the spectral version of the analysis step of the cycled Tikhonov scheme (2.8) takes the form

$$\begin{aligned} a_{n,k}^{(a)} &= a_{n,k}^{(b)} + \frac{1 - q_{n,k}}{\mu_n} \left(b_n - \mu_n a_{n,k}^{(b)} \right) \\ &= q_{n,k} a_{n,k}^{(b)} + (1 - q_{n,k}) \frac{b_n}{\mu_n}. \end{aligned} \tag{3.35}$$

For a constant model operator we have $a_{n,k}^{(b)} = a_{n,k-1}^{(a)}$ for $k = 1, 2, 3, \ldots$. We can then formulate the spectral version of Lemma 3.1.

LEMMA 3.5 *For the cycled Tikhonov regularization scheme (3.35) with constant model operator we have the spectral formula*

$$a_{n,k} - a_n^{(true)} = q_n^k \left(a_{n,0} - a_n^{(true)} \right) + \left(\sum_{\xi=0}^{k-1} q_n^\xi \right) \frac{1 - q_n}{\mu_n} b_n^{(\delta)} \tag{3.36}$$

and

$$a_{n,k} - a_n^{(true)} = q_n^k \left(a_{n,0} - a_n^{(true)} \right) + (1 - q_n^k) \frac{b_n^{(\delta)}}{\mu_n} \tag{3.37}$$

Proof. We rewrite Lemma 3.1 using the spectral operations introduced above. □

In spectral terms, the operator $\alpha^k(\alpha I + H^*H)^{-k}$ is given by the multiplication by $\alpha^k(\alpha + \mu_n^2)^{-k}$. This now leads to an extended version and alternative proof for Theorem 3.2.

THEOREM 3.6 *Assume that $H : X \to Y$ is a compact operator from a Hilbert space X into a Hilbert space Y. Then*

$$\left(I + \alpha^{-1} H^* H\right)^{-k} \varphi \to 0, \quad k \to \infty \tag{3.38}$$

for every fixed element $\varphi \in X$, $\varphi \in N(H^*H)^\perp$. For $\varphi \in N(H^*H)$ we obtain

$$\left(I + \alpha^{-1} H^* H\right)^{-k} \varphi = \varphi \tag{3.39}$$

for all $k \in \mathbb{N}$.

Proof. For $\varphi \in N(H^*H)^\perp$ we represent the vector φ in the orthonormal system φ_n by

$$\varphi = \sum_{n=1}^{\infty} a_n \varphi_n. \tag{3.40}$$

The application of $(I + \alpha^{-1} H^* H)^{-1}$ to φ_n is given by multiplication with

$$\left(1 + \frac{\mu_n^2}{\alpha}\right)^{-1}, \quad n \in \mathbb{N}. \tag{3.41}$$

Since $\mu_n^2 > 0$, the term is bounded by 1 uniformly for all $n \in \mathbb{N}$, which yields

$$||(I + \alpha^{-1} H^* H)^{-1}|| \leq 1.$$

Since $\mu_n^2 \to 0$ for $n \to \infty$, we also know that for all $\epsilon > 0$ we can find some n such that

$$\left(1 + \frac{\mu_n^2}{\alpha}\right)^{-1} > (1 - \epsilon),$$

from which we obtain

$$||(I + \alpha^{-1} H^* H)^{-1}|| = 1. \tag{3.42}$$

The same argument applied to the k-th power of the term shows that we cannot obtain norm convergence towards zero in (3.38). Clearly $\mu_n^2/\alpha > 0$ for every fixed n, thus we obtain

$$\left|\left(1 + \frac{1}{\alpha}\mu_n^2\right)^{-k}\right| \to 0, \quad k \to \infty \tag{3.43}$$

for fixed $n \in \mathbb{N}$. To show (3.38) we now decompose φ into

$$\varphi = \sum_{n=1}^{n_0} \beta_n \varphi_n \quad + \quad \sum_{n=n_0+1}^{\infty} \beta_n \varphi_n. \tag{3.44}$$

Since $\varphi \in X$, we have

$$\sum_{n=n_0+1}^{\infty} |\beta_n|^2 \to 0, \quad n_0 \to \infty. \tag{3.45}$$

This yields

$$\left|\left| \left(I + \alpha^{-1} H^* H\right)^{-k} \sum_{n=n_0+1}^{\infty} \beta_n \varphi_n \right|\right|^2 \leq \sum_{n=n_0+1}^{\infty} |\beta_n|^2 \to 0 \tag{3.46}$$

for $n_0 \to \infty$. Given $\epsilon > 0$ we first choose n_0 such that $\sum_{n=n_0+1}^{\infty} \beta_n \varphi_n$ is smaller than $\epsilon/2$ in norm. Then, we choose $k_0 > 0$ such that

$$\sum_{n=1}^{n_0} \left| \left(1 + \frac{1}{\alpha}\mu_n^2\right)^{-k} \beta_n \right|^2 < \frac{\epsilon}{2} \tag{3.47}$$

for $k \geq k_0$. This yields

$$\left|\left| \left(I + \alpha^{-1} H^* H\right)^{-k} \varphi \right|\right|^2 \leq \epsilon, \tag{3.48}$$

thus (3.38) is satisfied for $k \to \infty$. To show (3.39) we employ the trivial identity $I = (I + \alpha^{-1}H^*H) - \alpha^{-1}H^*H$ to calculate

$$\begin{aligned} (I + \alpha^{-1}H^*H)^{-1}\varphi &= (I + \alpha^{-1}H^*H)^{-1} \left\{ (I + \alpha^{-1}H^*H) - \alpha^{-1}H^*H \right\} \varphi \\ &= \varphi - (I + \alpha^{-1}H^*H)^{-1}\alpha^{-1}H^*H\varphi \\ &= \varphi \end{aligned} \tag{3.49}$$

for $\varphi \in N(H^*H)$, which ends the proof. □

In particular, for the cycled Tikhonov scheme (2.8) with true data

$$b_n = \mu_n a_n^{(true)}$$

we obtain

$$a_{n,k} = a_n^{(true)} - q_n^k \left(a_n^{(true)} - a_{n,0}\right) \tag{3.50}$$

Since $q_n^k \to 0$, $k \to \infty$ we obtain the spectral version of the convergence (3.22).

COROLLARY 3.7 *For true data $b_n = \mu_n a_{n,true}$ the cycled regularization scheme (3.35) is given by (3.50) and the iterated coefficients converge toward the true solution*

$$a_{n,k} \to a_{n,true}, \qquad k \to \infty. \tag{3.51}$$

We now consider noisy data, i.e. $y = y^{(true)} + y^{(\delta)}$ or the corresponding spectral formula $b_n = \mu_n a_n^{(true)} + b_n^{(\delta)}$. Applying Lemma 3.5 to measured data y with $y^{(\delta)} \in H(X)$ we obtain the following result.

COROLLARY 3.8 *For data $y = y^{(true)} + y^{(\delta)}$ with $y^{(\delta)} \in H(X)$, i.e. its spectral coefficient $b_n^{(\delta)}$ is given by $b_n^{(\delta)} = \mu_n a_n^{(\delta)}$, the dynamic Tikhonov scheme (3.35) generates iterated coefficients which converge toward some disturbed solution*

$$a_{n,k} \to a_n^{(true)} + a_n^{(\delta)}, \qquad k \to \infty. \tag{3.52}$$

Thus for disturbed data the Tikhonov scheme converges but not to the true solution, which is of course what we expect.

We now consider the case for $y^{(\delta)} \notin H(X)$ for which we will show the divergence of the regularization scheme, now by spectral arguments. The following basic estimate serves as a tool.

LEMMA 3.9 *Let $a, b \in \mathbb{R}$, then we have*

$$|a-b|^2 \geq \frac{|a|^2}{2} - |b|^2 \tag{3.53}$$

Proof. We note

$$a^2 = (a-b+b)^2 = (a-b)^2 + 2(a-b)b + b^2 \tag{3.54}$$

and from

$$0 \leq (a-b-b)^2 = (a-b)^2 - 2(a-b)b + b^2$$

derive

$$2(a-b)b \leq (a-b)^2 + b^2. \tag{3.55}$$

Combining the two equations (3.54) and (3.55) we obtain (3.53). □

We are now prepared to show the basic divergence result.

LEMMA 3.10 *If $y_\delta \notin Range(H)$ and $y_k = y_{true} + y_\delta$, then the dynamic Tikhonov scheme (2.8) or (3.35), respectively, diverges, i.e.*

$$||x_k^{(a)}|| \to \infty, \qquad k \to \infty. \tag{3.56}$$

Proof. For $y^\delta \notin H(X)$, due to Picard's theorem (compare [CK97] Theorem 4.8), we have

$$\sum_{n=0}^{L} \left| \frac{b_n}{\mu_n} \right|^2 \to \infty, \qquad L \to \infty. \tag{3.57}$$

We further have the following estimates for the coefficients: For every $n \in \mathbb{N}$ and since $0 < q_n < 1$ we have

$$\left|q_n^k\right| \leq 1, \ \left|1 - q_n^k\right| \leq 1, \qquad \forall n, k \in \mathbb{N}_0. \tag{3.58}$$

Using the representation (3.37), the bounds (3.58) and the general estimate from Lemma 3.9 we obtain

$$\begin{aligned}
||x_k^{(a)}||^2 &\geq \sum_{n=0}^{L} |a_{n,k}|^2 \\
&= \sum_{n=0}^{L} \left| q_n^k a_{n,0} + (1-q_n^k) a_n^{(true)} + \frac{b_n}{\mu_n}(1-q_n^k) \right|^2 \\
&\geq \frac{1}{2} \sum_{n=0}^{L} \left| \frac{b_n}{\mu_n}(1-q_n^k) \right|^2 - \sum_{n=0}^{L} \left(\left|q_n^k a_{n,0}\right| + \left|(1-q_n^k)\right| \left|a_n^{(true)}\right| \right)^2 \\
&\geq \frac{1}{2} \sum_{n=0}^{L} \left| \frac{b_n}{\mu_n} \right|^2 (1-q_n^k)^2 - c
\end{aligned} \tag{3.59}$$

for some constant c. For each fixed n we have $(1-q_n^k) \to 0$ for $k \to \infty$. For some constant C we first choose L such that the sum in (3.57) is larger than $C + 2c$. Then for all sufficiently large k the sum in (3.59) is larger than C and therefore we have the divergence (3.56). □

3.3. Instability for spectrally expanding or collapsing systems

Finally, we will study the behaviour of the error of data assimilation algorithms with a compact observation operator in the case of spectrally expanding or collapsing systems as introduced in (2.1). We start with the spectral update formula (3.35) and use the dynamics introduced in (2.2)

$$a_{n,k+1}^{(b)} = d_n a_{n,k}^{(a)}, \quad k = 0, 1, 2, ... \tag{3.60}$$

which is spectrally diagonal with respect to the singular system of the observation operator H. In this case the coefficients show an exponential behaviour

$$a_{n,k}^{(true)} = d_n^k a_{n,0}^{(true)}, \quad k = 1, 2, 3, ... \tag{3.61}$$

Then, the assimilation (3.35) with data $y_k = Hx_k^{(true)} + y_k^{(\delta)}$, $k = 1, 2, 3, ...$, leads to the update formula

$$a_{n,k+1}^{(a)} = q_n d_n a_{n,k}^{(a)} + (1 - q_n) a_{n,k+1}^{(true)} + (1 - q_n) \frac{b_{n,k+1}^{(\delta)}}{\mu_n}. \tag{3.62}$$

Note that the true solution is dynamical as well and we need to build this dynamics into our induction formula to study the dynamical evolution of the analysis and the errors.

One key question is how the data error evolves. Here, we assume that the measurements are remote sensing data which are independent of the system development. This means that the data error is basically constant and does not change over time, but the system expands or collapses. Then, the system development leads to changes in the relationship between the size of the error and the size of the system, such that the relative error changes significantly with time.

LEMMA 3.11 *Assume that we have data given by*

$$y_k = y_k^{(true)} + y^{(\delta)} = Hx_k^{(true)} + y^{(\delta)}. \tag{3.63}$$

Then the updates (3.62) lead to an evolution of the analysis

$$a_{n,k}^{(a)} = q_n^k d_n^k a_{n,0} + (1 - q_n^k) d_n^k a_{n,0}^{(true)} + \Big(\sum_{\xi=0}^{k-1} q_n^\xi d_n^\xi \Big) (1 - q_n) \frac{b_n^{(\delta)}}{\mu_n} \tag{3.64}$$

$$= q_n^k d_n^k a_{n,0} + (1 - q_n^k) a_{n,k}^{(true)} + (1 - d_n^k q_n^k) \frac{(1 - q_n)}{(1 - d_n q_n)} \frac{b_n^{(\delta)}}{\mu_n}. \tag{3.65}$$

Proof. For $k = 1$ the formula (3.64) corresponds to (3.62). Assume it to be true for some $k \in \mathbb{N}$. Then we have

$$\begin{aligned} a_{n,k+1}^{(a)} &= q_n d_n a_{n,k}^{(a)} + (1 - q_n) a_{n,k+1}^{(true)} + (1 - q_n) \frac{b_n^{(\delta)}}{\mu_n} \\ &= q_n d_n \Big\{ q_n^k d_n^k a_{n,0} + (1 - q_n^k) d_n^k a_{n,0}^{(true)} + \Big(\sum_{\xi=0}^{k-1} q_n^\xi d_n^\xi \Big) (1 - q_n) \frac{b_n^{(\delta)}}{\mu_n} \Big\} \\ &\quad + (1 - q_n) d_n^{k+1} a_{n,0}^{(true)} + (1 - q_n) \frac{b_n^{(\delta)}}{\mu_n} \end{aligned}$$

$$= q_n^{k+1} d_n^{k+1} a_{n,0} + (1 - q_n^{k+1}) d_n^{k+1} a_{n,0}^{(true)} + \Big(\sum_{\xi=0}^{k} q_n^{\xi} d_n^{\xi} \Big)(1 - q_n) \frac{b_n^{(\delta)}}{\mu_n},$$

which is the desired formula with k replaced by $k+1$. Finally, (3.65) is obtained by standard geometric series, and the proof is complete. □

1) The Convergence Case. We can now investigate the asymptotic behavior, where two main situations arise. We have convergent terms in (3.65) if the condition

$$|d_n q_n| = \left| \frac{\alpha d_n}{\alpha + \mu_n^2} \right| < 1 \tag{3.66}$$

is satisfied. The condition (3.66) can be rewritten as

$$|d_n| < 1 + \frac{\mu_n^2}{\alpha}, \quad n \in \mathbb{N}. \tag{3.67}$$

The spectral expansion coefficient d_n needs to be bounded by (3.67) for the initial guess term to decay. This reflects the general understanding that if the dynamical increase of a mode is stronger than the decrease of its influence over time, the influence of the past errors will increase in our current analysis.

The condition (3.66) is also needed for the third spectral term to converge. Otherwise, the error which is coming in from $y^{(\delta)}$ is increased in every step of the time evolution.

(i) If the data error $y^{(\delta)}$ is zero, then we obtain exponential convergence of the analysis towards the dynamics of the true solution $x_k^{(true)}$, since $q_n^k \to 0$ for $k \to \infty$.

(ii) If the measurement error in each step is given by $y^{(\delta)} \neq 0$, its spectral limit is

$$\frac{(1 - q_n)}{(1 - d_n q_n)} \frac{b_n^{(\delta)}}{\mu_n} = \Big(\frac{\mu_n^2}{(1 - d_n)\alpha + \mu_n^2} \Big) \frac{b_n^{(\delta)}}{\mu_n}, \quad n \in \mathbb{N}, \tag{3.68}$$

which leads to a constant regularized solution $x_\alpha^{(\delta)} \in X$ in the case of *strictly spectrally collapsing dynamics*, i.e. if $|d_n| \leq \rho < 1$ for all $n \in \mathbb{N}$ with some constant ρ. In this case, where the true solution $x_k^{(true)}$ is spectrally collapsing, i.e.

$$\left| x_k^{(true)} \right| \leq \rho^k \left| x_0^{(true)} \right|, \quad k \in \mathbb{N}, \tag{3.69}$$

the error limit $x_\alpha^{(\delta)}$ will gain more and more relative importance. The relative error diverges.

(iii) If we have a spectrally expanding system where

$$|d_n q_n| \leq \rho < 1, \quad n \in \mathbb{N}, \tag{3.70}$$

we obtain

$$|d_n| = \frac{\rho}{|q_n|} \to \rho < 1, \quad n \to \infty, \tag{3.71}$$

which contradicts $d_n > 1$, i.e. there cannot be spectrally expanding systems for which (3.70) is satisfied.

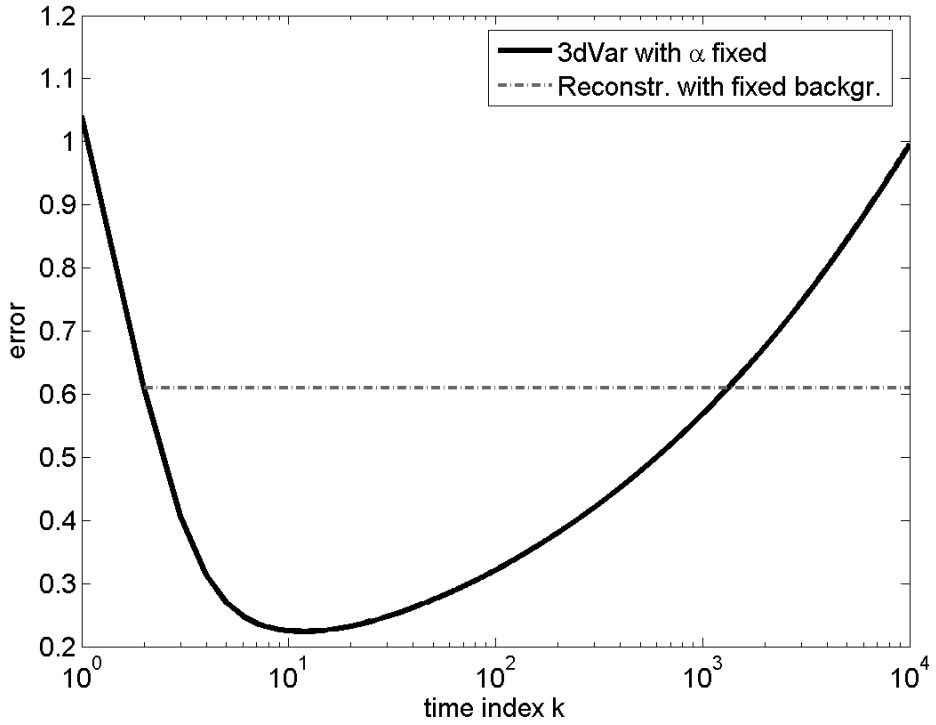


Figure 1. The image displays the evolution of the analysis error for the cycled Tikhonov regularization simulating the development of the analysis $x_k^{(a)}$ for $k \to \infty$ when $y^{(\delta)} \notin H(X)$. We observe that the analysis error first decays to some minimum and then increases and becomes even larger than the error of a non-dynamic plain Tikhonov regularization where we use a fixed background (dashed line).

2) Strong exponential system growth. If the expansion given by d_n is too strong, then it is not sufficient to look at the unscaled system, but again the relative error needs to be studied. Here, we scale the growth by $1/d_n^k$, i.e. we rescale everything to the size of the true mode. Then

$$\begin{aligned} \frac{a_{n,k}^{(a)}}{d_n^k} &= q_n^k a_{n,0} + (1-q_n^k) a_{n,0}^{(true)} + \left(\frac{1}{d_n^k} - q_n^k\right) \frac{(1-q_n)}{(1-d_n q_n)} \frac{b_n^{(\delta)}}{\mu_n} \\ &\to a_{n,0}^{(true)}, \quad k \to \infty, \end{aligned} \tag{3.72}$$

i.e. in this case we have convergence of the scaled assimilation towards the true solution.

4. Numerical Examples

The task of this final part is to provide practical examples to demonstrate the behavior described by the above theory. We first show a realization of the above spectral system in Example 1, a second example is taken from dynamic magnetic tomography.

Example 1: Spectral System. Consider the application of 3dVar for data assimilation where we assume a constant model system $M_{k+1|k} = I$ and measured data

$y = Hx^{(true)} + y^{(\delta)} \notin H(X)$. We have shown in Lemma 3.10 that in this case we have

$$||x_k^{(a)}|| \to \infty, \quad k \to \infty. \tag{4.1}$$

Further, from (3.37) we deduce that the spectral coefficient of the analysis error is exponentially convergent towards

$$\frac{b_n^{(\delta)}}{\mu_n}, \quad n \in \mathbb{N}. \tag{4.2}$$

The speed of this convergence is proportional to

$$q_n = \frac{\alpha}{\alpha + \mu_n^2}, \quad n \in \mathbb{N}, \tag{4.3}$$

i.e. it is depending on the particular mode. Since $|\mu_n| \to 0$ for $n \to \infty$ the convergence is quick for small modes and decreases when n gets larger.

- When we have large initial error on small modes, this error will be quickly reduced by the 3dVar cycling steps.
- Slowly we have the approximation of the higher error modes, such that there is increasing error on the higher modes which tends to infinity for $k \to \infty$.

This can lead to a reduction of the error over some period of time $[0, T_{min}]$. Then, when the error of higher modes takes over, the error grows again in $[T_{min}, \infty)$ with limit ∞ at the right-hand side of the unbounded interval.

The results of the simulation are shown in Figure 1, where we see the behavior of the total error of our system. The thick black line first shows a decay of the total error until it reaches a minimum at T_{min}. This corresponds to the part where the error in the higher modes of of the error is small. In the second part we see that the total error increases. The results confirm our theoretical considerations.

In the figure we also display a dotted line, where we employ Tikhonov regularization for reconstruction without an update of the background, i.e. with a fixed background chosen by the initial guess. So from the second step of the iteration we obtain a constant reconstruction (due to our constant system and constant data we feed in). The cycled 3dVar starts with the same error, then improves and leads to a better reconstruction for some time before it becomes distorted by the error $y^{(\delta)} \notin H(X)$. This means that after some time T_2 the error of the cycled scheme becomes larger than the error of the plain non-dynamic Tikhonov regularization.

Example 2: Dynamic Magnetic Tomography. We provide a realistic numerical example using the dynamic magnetic tomography problem as introduced and analyzed in [Mar11]. Figure 2 shows the setup of fuel cell or fuel cell stack, and corresponds to the domain $\Omega \in \mathbb{R}^3$. In Ω we have the anisotropic conductivity distribution $x \in (L^\infty(\Omega))^{3\times 3}$ and currents $j \in (L^2(\Omega))^3$. Further we use the measurement operator

$$\begin{array}{ccc} W : (L^\infty(\Omega))^{3\times 3} & \longrightarrow & \left(L^2(\partial G)\right)^3, \\ x & \longmapsto & y. \end{array} \tag{4.4}$$

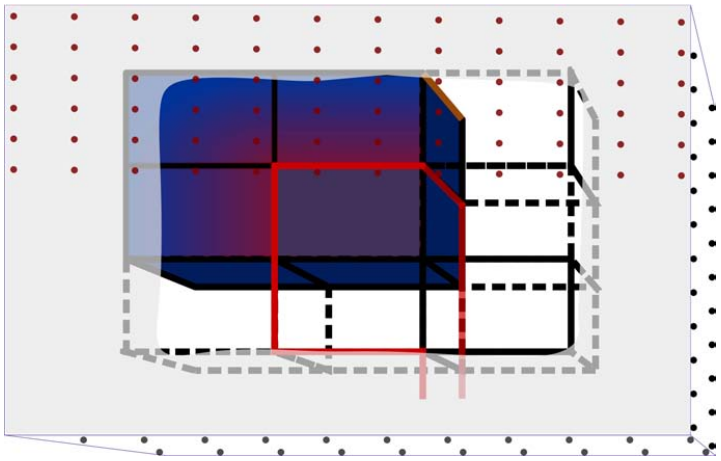


Figure 2. Scheme representing the fuel cell domain with in- and outflowing Nodes, conductivity (solid and dashed lines) and current density flow (red path) current density distribution (colored boxes) and surrounding measurement surface (transparent box surface) and measurement points.

with $\bar{\Omega} \subset G$, which maps the conductivity distribution x on the domain Ω onto the magnetic field y on ∂G. We measure magnetic field data y on the boundary ∂G. Our task is the dynamic reconstruction of the conductivity x.

From the literature we know that the classical problems consists of three major parts, i.e. the model for the current density distribution give by a boundary value problem as presented in [PK03]. The direct model given by the Biot-Savart operator as e.g. also analyzed in [PK03] with several properties stated in [HKP05] and the reconstruction step presented in [KKP02]. In [Mar11] the model was extended to use it with data assimilation algorithms, i.e. we combine the boundary value formulation with the Biot-Savart operator to state the direct magnetic tomography problem (4.4) and solve it by using data assimilation algorithms from Section 2. For all details on the dynamic magnetic tomography problem we refer to [Mar11].

Nonlinear Observation Operators. To treat nonlinear observation operators, we use the Fréchet derivative as follows. From [Mar11] we know that W is Fréchet differentiable, i.e. we have

$$W(x+\delta x) = W(x) + W'(x)\delta x + \mathcal{O}(\delta x), \quad \delta x \in X, \tag{4.5}$$

with the *Fréchet derivative* $W'(x)$ of W at the point $x \in X$. Then, at time t_{k+1} we need to solve $W(x+\delta x) = y$, which is approximated by

$$W(x) + W'(x)\delta x = y \quad \Leftrightarrow W'(x)\delta x = y - W(x). \tag{4.6}$$

This leads to the linearized version of the update, given by (2.12) with

$$K_{k+1} = B_{k+1}^{(b)}(W'(x_{k+1}^{(b)}))^* \left(W'(x_{k+1}^{(b)})B_{k+1}^{(b)}(W'(x_{k+1}^{(b)}))^* + R\right)^{-1}. \tag{4.7}$$

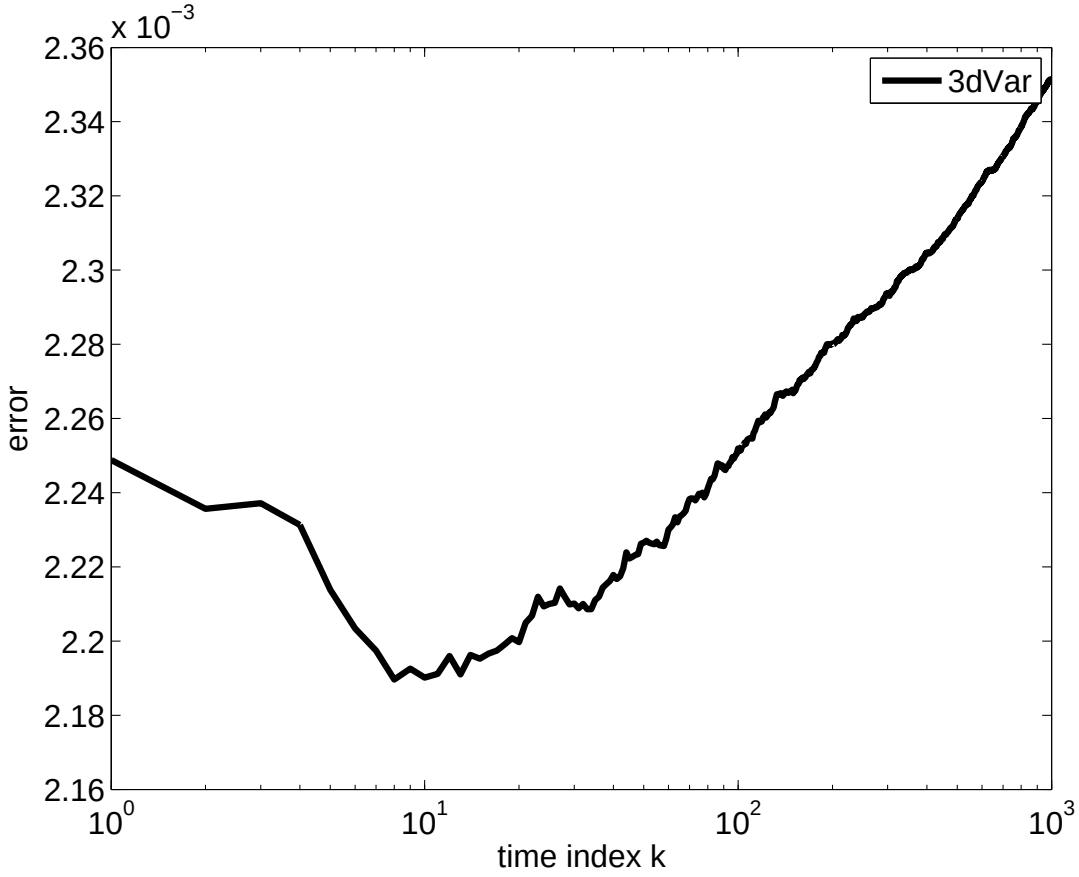


Figure 3. Dynamic Magnetic Tomography problem relative error for 3dVar reconstructions with divergent behavior due to errors on singular values. The parameters are $\alpha^{(x)} := 1$ as diagonal value for the covariance matrix B and $\alpha^{(r)} = 1$ as diagonal value for the covarince matrix R which corresponds to the Tikhonov case for $\alpha = \alpha^{(x)}$.

Simulation and Noise. We use a fixed grid, set a initial conductivity x_0 configuration on Ω and choose a fixed in- and outflowing current $j_{in/out}$. Then, by solving the non-linear part of the direct problem, we obtain the current density distribution inside Ω. This is mapped onto the magnetic field data $y \in G$. To test our theory we employ a constant system, i.e. the model operator is given by the identity $M_{k+1|k} := I$ for all $k \in \mathbb{N}$.

We use noisy data to construct some data error $y^{(\delta)}$ using its singular values μ_n. In particular, setting the additive noise to $a_n^{(\delta)} := c \cdot (\mu_n)^\lambda$ for $n = 1, ..., N$ with the size N of our numerical discretization and with some $0 < \lambda < 1$ simulates the condition $y^{(\delta)} \notin H(X)$.

Reconstructions using 3dVar. We study the cycled 3dVar algorithm for inversion. Choosing the measurement covariance matrix $R := I$ and the background covariance matrix $B_k := \alpha I$. We choose the initial conductivity x_0 to be a diagonal conductivity matrix with constant entries for each spatial direction. We then use (4.7) with linearization around x_0 and start the 3dVar scheme.

Figure 3 shows the relative reconstruction error. We see that the error decreases until it reaches a minmum after 11 steps. After that point the sum of the error according to (3.57) takes over and the full system behaves as seen in (3.59). This confirms the evolution given by Lemma 3.10 for a practically relevant example.

5. Summary

We have investigated the instability which can occur when measurements are assimilated into a dynamical system evolution which are linked to the system state φ in an infinite dimensional state space X by a compact measurement operator H. For simple systems we have shown that cycling of standard data assimilation schemes can lead to severe instabilities of the analysis, i.e. small measurement errors accumulate over time and can lead to large analysis errors. We have worked out explicit spectral formulas which show the instable behaviour. Further, numerical results in simple cases and from dynamical magnetic tomography confirm the results. These results are interesting also in seismology and earth sciences.

The systems investigated in this work can be seen as a simple model if the speed of change in the dynamical system is small compared to the frequency of measurements. But if even these quite stable systems can show severe instable behaviour, more general systems are even more likely to show similar behaviour. Subsequent work has already been carried out by Potthast, Moodey, Lawless and van Leeuwen [PMLvL], where dynamical systems M of trace class are investigated. Note that the systems here, in particular the constant system $M = I$, are not of trace class, but trace class systems damp higher modes as it is usually carried out in global atmospheric models. For trace class systems the authors show similar results, but also provide a stable assimilation setup to control the analysis error over time. Further research in this direction is highly interesting to learn more about possible instabilities of operational data assimilation systems, which by causing large forecast error can have a significant impact on many parts of society.

References

[CGTU08] Alberto Carrassi, Michael Ghil, Anna Trevisan, and Francesco Uboldi. Data assimilation as a nonlinear dynamical systems problem: Stability and convergence of the prediction-assimilation system. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 18(2):023112, 2008.

[CK97] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, 2nd ed. edition, December 1997.

[Eng87] Heinz W. Engl. On the choice of the regularization parameter for iterated tikhonov regularization of III-posed problems. *Journal of Approximation Theory*, 49(1):55–63, January 1987.

[HKP05] Karl-Heinz Hauer, Lars Kühn, and Roland Potthast. On uniqueness and non-uniqueness for current reconstruction from magnetic fields. *Inverse Problems*, 21(3):955–967, 2005.

[HP07] Karl-Heinz Hauer and Roland Potthast. Magnetic tomography for fuel cells - current status and problems. *Journal of Physics: Conference Series*, 73:012008 (17pp), 2007.

[HPW08] Karl-Heinz Hauer, Roland Potthast, and Martin Wannert. Algorithms for magnetic tomography?on the role of a priori knowledge and constraints. *Inverse Problems*, 24, 2008.

[HPWS05] Karl-Heinz Hauer, Roland Potthast, Thorsten Wüster, and Detlef Stolten. Magnetotomography – a new method for analysing fuel cell performance and quality. *Journal of Power Sources*, 143(1-2):67–74, April 2005.

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

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

[KKP02] Rainer Kress, Lars Kühn, and Roland Potthast. Reconstruction of a current distribution from its magnetic field. *Inverse Problems*, 18(4):1127–1146, 2002.

[Kre99] Rainer Kress. *Linear Integral Equations: v.82: Vol 82*. Springer, 2nd ed. edition, April 1999.

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

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

[LSK10] William Lahoz, Richard Swinbank, and Boris Khattatov. *Data Assimilation: Making Sense of Observations*. Springer, June 2010.

[Mar11] Boris A. Marx. *Dynamic Magnetic Tomography*. Der Andere Verlag, May 2011.

[PbG09] Roland Potthast and Peter beim Graben. Inverse problems in neural field theory. *SIAM J. Appl. Dyn. Syst.*, 8(4):1405–1433, 2009.

[PK03] Roland Potthast and Lars Kühn. On the convergence of the finite integration technique for the anisotropic boundary value problem of magnetic tomography. *Mathematical Methods in the Applied Sciences*, 26(9):739–757, 2003.

[PMLvL] Roland Potthast, Alexander J.F. Moodey, Amos S. Lawless, and Peter Jan van Leeuwen. On error dynamics and instability in data assimilation. *Preprint http://www.reading.ac.uk/web/FILES/maths/preprint*$_1 2_0 5_L awless.pdf$.

[PW09] Roland Potthast and Martin Wannert. Uniqueness of current reconstructions for magnetic tomography in multilayer devices. *SIAM Journal on Applied Mathematics*, 70(2):563–578, 2009.

[PX09] Seon K. Park and Liang Xu. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications*. Springer, 2009.

[RL00] Allan R. Robinson and Pierre F.J. Lermusiaux. *Overview of data assimilation*. Harvard Reports in Physical/Interdisciplinary (Ocean Sciences); The Division of Engineering and Applied Sciences Harvard University: Cambridge, Massachusetts, USA, 2000.

[Wan09] Martin Wannert. *Stabile Algorithmen für die Magnetotomographie an Brennstoffzellen*. PhD thesis, Universität Göttingen, 2009.

[War11] Thomas Tomkins Warner. *Numerical Weather and Climate Prediction*. Cambridge University Press, January 2011.