REDUCED MINIMAX FILTERING BY MEANS OF DIFFERENTIAL-ALGEBRAIC EQUATIONS

Vivien Mallet
INRIA
Paris-Rocquencourt research center
France
&
CEREA, joint laboratory EDF R&D
and École des Ponts ParisTech,
Université Paris Est
vivien.mallet@inria.fr

Sergey Zhuk
Department of System Analysis an Decision Making Theory
Taras Shevchenko National University of Kiev
Ukraine
beetle@unicyb.kiev.ua
&
CWI, Amsterdam
The Netherlands
sergiy.zhuk@cwi.nl

Abstract

A reduced minimax state estimation approach is proposed for high-dimensional models. It is based on the reduction of the ordinary differential equation with high state space dimension to the low-dimensional Differential-Algebraic Equation (DAE) and on the subsequent application of the minimax state estimation to the resulting DAE. The DAE is composed of a reduced state equation and of a linear algebraic constraint. The latter allows to bound linear combinations of the reduced state's components in order to prevent possible instabilities, originating from the model reduction. The method is robust as it can handle model and observational errors in any shape, provided they are bounded. It allows to compute both the state estimate and the reachability set in the reduced space. We include a quick application to a complex air quality model in order to illustrate the benefit of the minimax reduction compared to the classical reduction.

Key words

minimax, reduction, differential-algebraic equations, estimation, filtering

1 Introduction

Numerical modeling of complex systems such as the Earth's atmosphere involves complex numerical models relying on systems of coupled Partial Differential Equations (PDEs). As an example, consider chemistry-transport models that describe the fate of the pollutants in the atmosphere (e.g., the models described in (Mallet, Quélo, Sportisse, Ahmed de Biasi, Debry, Korsakissok, Wu, Roustan, Sartelet, Tombette and Foudhil, 2007)). For these models, the dimension of the state vector[1] can reach 10^7 or even more, and the time integration has such a large computational cost that only the equivalent of a few dozens of model calls may be affordable. The computational costs of these models and their dimensions raise specific issues when one wants to reduce simulation errors (caused by imperfect model formulation or uncertain inputs) through assimilation of the observed data (sparse observations of the model's state) into the model. Classical assimilation algorithms such as the Kalman filters (Balakrishnan, 1984) can be so demanding in terms of computations that they cannot be applied to these models without a reduction.

Reduced Kalman filters have been developed to address this issue by introducing a reduction of in the filtering algorithm—see (Wu, Mallet, Bocquet and Sportisse, 2008) for an application to the aforementioned chemistry-transport models in air quality simulation. In these filters, the key reduction lies in the propagation of the state error covariance matrix which is intractable[2] in Kalman filter. A popular reduced Kalman filter is the so-called ensemble Kalman filter in which the state error covariance matrix is approximated by the empirical variance of the ensemble (Heemink, Verlaan and Segers, 2001). The particles can be deterministically sampled like in the SEIK versions (Pham, 2001), in the unscented Kalman filter (Julier and Uhlmann, 1997) or in its reduced version (Moireau and Chapelle, 2010). Another example is the reduced-rank square-root Kalman filter based on propagation of the most important modes (Verlaan and Heemink, 1995) of the error covariance matrix.

Another direction is the reduction of the model itself and subsequent application of an appropriate filter-

[1] State vector of the PDE after discretization in space.

[2] Since the propagation involves twice as much calls to the tangent linear model as components in the state

ing technique to the resulting low-dimensional model. The Galerkin projection represents one of the most used techniques for model reduction (Brenner and Scott, 2005). The idea is to find a low dimensional subspace in the model state space and to restrict the model onto that subspace. Of course, there is a loss of information due to restricting the dynamics of the model onto the subspace. One way to minimize the loss is to generate the subspace by means of the Proper Orthogonal Decomposition (POD) (Homescu, Petzold and Serban, 2005).

In this paper, we introduce a reduced minimax filter, designed to be applied to high-dimensional models. Minimax approach allows 1) to filter any model error with bounded energy and observational error either deterministic with bounded energy or stochastic with bounded variance., 2) to estimate the worst-case error and 3) to assess how accurate the link between the model and observed phenomena is. Our approach is to make a reduction of the model itself and to apply the minimax filtering to the reduced model, provided uncertain model error and observation noise are elements of a given bounding set. We introduce a reduced state equation projecting the full state vector onto a subspace which can be generated, e.g., by means of POD. The projection introduces errors that can lead to a reduced state equation with unstable dynamics. In order to address this issue, we introduce an additional energy constraint on the reduced state in the form of a linear algebraic equation. Finally, our reduced model is represented by a Differential-Algebraic Equation (DAE), composed of a reduced state equation and of a linear constraint. We apply an extended version of the minimax filter for DAE (Zhuk, 2010) to the reduced model without further reduction on the filter.

The paper is organized as follows. After the notation is introduced in section 1.1, the extended minimax filter, without reduction, is presented in section 2. This section quickly explains the minimax framework, introduces the filter and comments on the intractability of the computations. The reduction procedure is then derived in section 3. The classical Galerkin projection is first commented. The DAE approach is then introduced, in the linear case and in an extended version for the non-linear case. The benefit of the reduction with our approach is finally illustrated with an air quality model.

1.1 Notation

Let $\mathcal{M}_t : \mathbb{R}^N \to \mathbb{R}^N$ define the model at some time step $t \in \{0, \dots, T-1\}$:

$$x_{t+1} = \mathcal{M}_t(x_t) + e_t\,, \quad x_0 = x_0^g + e\,, \tag{1}$$

where x_0^g is an approximation of the initial condition with error $e \in \mathbb{R}^N$, $x_t \in \mathbb{R}^N$ denotes the state vector, $e_t \in \mathbb{R}^N$ is the model error.

Let $y_t \in \mathbb{R}^m$ denote the observation of the true state x_t at time t. We assume that y_t satisfies

$$y_t = \mathcal{H}_t(x_t) + \eta_t\,, \tag{2}$$

where $\mathcal{H}_t : \mathbb{R}^N \to \mathbb{R}^m$ is the observation operator mapping the state space into observation space, and $\eta_t \in \mathbb{R}^m$ is the observation error.

We assume that the error (e, e_t, η_t) is uncertain but bounded so that

$$\begin{aligned} &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle \\ &+ \sum_{t=0}^{T-1}\langle Q_t^{-1}(e_t-\overline{e}_t), e_t-\overline{e}_t\rangle \\ &+ \sum_{t=0}^{T}\langle R_t^{-1}(\eta_t-\overline{\eta}_t), \eta_t-\overline{\eta}_t\rangle \leqslant 1\,, \end{aligned} \tag{3}$$

where $Q, Q_t \in \mathbb{R}^{N\times N}$ and $R \in \mathbb{R}^{m\times m}$ are symmetric positive-definite matrices, and $\overline{e}, \overline{e}_t \in \mathbb{R}^N$ and $\overline{\eta}_t \in \mathbb{R}^m$ may be viewed as systematic errors.

The tangent linear model is $M_t = \mathrm{D}\mathcal{M}_t(x_t) \in \mathbb{R}^{N\times N}$. Consistently we introduce the associated tangent linear operator $H_t = \mathrm{D}\mathcal{H}_t(x_t) \in \mathbb{R}^{m\times N}$.

The reduction applies to the model state, and the reduced model state is denoted $z_t = F_t^T x_t \in \mathbb{R}^n$, with $n \ll N$. $F_t \in \mathbb{R}^{N\times n}$ is a reduction matrix. The minimax estimator of x_t is denoted $\widehat{x}_t \in \mathbb{R}^N$. The minimax estimator is derived from the reduced minimax estimator with $\widehat{x}_t = F_t\widehat{z}_t$.

The tangent linear operators along the trajectory $\widehat{x}_t$ are denoted $\widehat{M}_t = \mathrm{D}\mathcal{M}_t(\widehat{x}_t)F_t \in \mathbb{R}^{N\times n}$ (for $t \geqslant 0$) and $\widehat{H}_t = \mathrm{D}\mathcal{H}_t(\mathcal{M}_{t-1}(\widehat{x}_{t-1}) + \overline{e}_{t-1})F_t \in \mathbb{R}^{m\times n}$ (for $t > 0$), for the model and the observation operator respectively. We also define $\widehat{H}_0 = \mathrm{D}\mathcal{H}_0(x_0^g + \overline{e})F_0 \in \mathbb{R}^{m\times n}$.

We denote by A^+ the Moore-Penrose pseudoinverse of a matrix A. $I_{k\times k}$ denotes the identity matrix in $\mathbb{R}^{k\times k}$. $\langle\cdot,\cdot\rangle$ denotes the canonical inner product of the Euclidean space.

2 Extended minimax state estimation

2.1 Minimax filter for Ordinary Differential Equations with discrete time

In what follows we present a minimax state estimation algorithm that solves the following filtering problem: given a sequence of observed data $y_0, \dots, y_T$ in the form (2) and given the uncertainty description (3), one should estimate the state x_T of (1).

Our approach is based on the following idea: to describe how the model propagates uncertain parameters verifying (3). The key point is to construct the so-called reachability set $\mathcal{R}_t$ at time t, that is, the set of all states x_t satisfying (1) and compatible with the description of uncertain parameters (3) and the observed data y_t in the form (2). In other words, the state x_t^* belongs to $\mathcal{R}_t$ if and only if there is a sequence $E^* := (e^*, \{e_0^*, \dots, e_{t-1}^*\}, \{\eta_0^*, \dots, \eta_t^*\})$ verifying (3) such that the sequence $x_0^*, \dots, x_t^*$ computed from (1) for $e = e^*$ and $e_s = e_s^*$, $0 \leqslant s < t$, is compatible with observed data $y_0, \dots, y_t$ through (2) with $\eta_s = \eta_s^*$, $0 \leqslant s \leqslant t$. This suggests a way to estimate how the model propagates uncertain parameters (initial-condition error e and model error e_t): it is sufficient to have a description of the dynamics of $\mathcal{R}_t$ in time. The true state can only lie in $\mathcal{R}_t$. Note that the dynamics of $\mathcal{R}_t$ takes into account only those realizations of e, e_t which are compatible with the actual realization of observed data $y_0, \dots, y_t$. Consequently, if $\mathcal{R}_t$ is empty, one can conclude that the errors were wrongly described by (3).

Any point of $\mathcal{R}_t$ can be the true state. In order to obtain a minimax estimate of this true state, we assign to a point $x \in \mathcal{R}_t$ a worst-case error, that is the maximal distance between x and other points of $\mathcal{R}_t$. The point with minimal worst-case error[3] is the minimax estimate $\hat{x}_t$. Roughly speaking, the worst-case error can be thought of as a "the longest axis" of the minimal ellipsoid containing $\mathcal{R}_t$, and the minimax estimate $\hat{x}_t$ is the central point of that ellipsoid.

The basics of the minimax state estimation were developed by (Bertsekas and Rhodes, 1971), (Milanese and Tempo, 1985), (Chernousko, 1994), (Kurzhanski and Vályi, 1997), (Nakonechny, 2004). The main advantages of minimax estimates are as follows: (1) the possibility to filter any model error and observation noise with bounded energy, (2) the estimation of the worst-case error, (3) fast estimation algorithms in the form of filters, (4) the possibility to evaluate the model, that is, to assess how good the model describes observed phenomena.

In this subsection, we assume that $F_t = I_{N \times N}$. In this case, there is no reduction and the state estimation algorithm operates on the full model. Following (Zhuk, 2010), we introduce an extended version of the linear minimax state estimate $\hat{x}_t$, a minimax gain G_t and the reachability set $\mathcal{R}_t$ for (1):

$$
\begin{aligned}
&G_0 = Q_0^{-1} + \hat{H}_0^T R_0^{-1} \hat{H}_0, \\
&\hat{x}_0 = G_0^{-1} \left(Q_0^{-1} \overline{e} + \hat{H}_0^T R_0^{-1} (y_0 - \overline{\eta}_0) \right), \\
&\beta_0 = \langle R_0^{-1} (y_0 - \overline{\eta}_0), y_0 - \overline{\eta}_0 \rangle, \\
&G_{t+1} = Q_t^{-1} - Q_t^{-1} \widehat{M}_t B_t \widehat{M}_t^T Q_t^{-1} \\
&+ \hat{H}_{t+1}^T R_{t+1}^{-1} \hat{H}_{t+1}, \\
&B_t = (G_t + \widehat{M}_t^T Q_t^{-1} \widehat{M}_t)^{-1}, \\
&\hat{x}_{t+1}^f = \mathcal{M}_t(\hat{x}_t), \\
&\hat{x}_{t+1} = \hat{x}_{t+1}^f + G_{t+1}^{-1} \hat{H}_{t+1}^T R_{t+1}^{-1} [y_{t+1} - \overline{\eta}_{t+1} \\
&- \hat{H}_{t+1} \hat{x}_{t+1}^f] + G_{t+1}^{-1} (Q_t + \widehat{M}_t G_t^{-1} \widehat{M}_t^T)^{-1} \overline{e}_t, \\
&\mathcal{X}_{t+1} = \{ w : \langle G_{t+1} w, w \rangle \leqslant 1 \}, \\
&\beta_{t+1} = \beta_t + \langle R_{t+1}^{-1} (y_{t+1} - \overline{\eta}_{t+1}), y_{t+1} - \overline{\eta}_{t+1} \rangle \\
&- \langle B_t^+ G_t^{-1} \hat{x}_t, G_t^{-1} \hat{x}_t \rangle + \langle \widehat{M}_t^T Q_t \widehat{M}_t \overline{e}_t, \overline{e}_t \rangle, \\
&\mathcal{R}_t = \hat{x}_t + \sqrt{1 - \beta_t + \langle G_t \hat{x}_t, \hat{x}_t \rangle} \mathcal{X}_t .
\end{aligned}
\tag{4}
$$

Here $\mathcal{R}_t$ denotes an ellipsoidal approximation of the reachability set for the model (1) and β_t is a scaling factor. The dynamics of $\mathcal{X}_t$ describes how the model $\mathcal{M}_t$ propagates uncertain parameters from the bounding set (3) compatible with observed data y_t. The observation-dependent scaling factor β_t defines whether $\mathcal{X}_t$ shrinks or expands. If $1 - \beta_t + \langle G_t \hat{x}_t, \hat{x}_t \rangle < 0$, then the observed data is incompatible with our assumption on uncertainty description (3). In the form (4), the minimax state estimation algorithm can be applied to non-linear models, hence we refer to it as an extended minimax filter. Nevertheless, the theory only supports the algorithm in the linear case—that is, the reachability set is known to contain all possible true states only in the linear case.

The algorithm is far too expensive for high-dimensional systems: it requires to propagate a minimax gain $G_t \in \mathbb{R}^{N \times N}$, where N is the dimension of the state space of the model (1). For instance, with $N = 10^7$ like in air quality applications, the dimension of G_t is $10^7 \times 10^7$, which cannot be manipulated by modern computers because of huge computational loads and out-of-reach memory requirements. Hence a reduction is necessary to carry out the computations for high dimensional systems.

2.2 Minimax filter for Differential-Algebraic Equations with discrete time

A more general form of the filter was derived in (Zhuk, 2010) for DAE problems. The filter addresses the problem

$$
\begin{aligned}
&F_{t+1} z_{t+1} = \mathcal{M}_t(F_t z_t) + r_t \,, \\
&y_t = \mathcal{H}_t(F_t z_t) + \eta_t, F_0 z_0 = F_0 F_0^T (x_0^g + e) \,,
\end{aligned}
\tag{5}
$$

[3] This point will coincide with the Tchebysheff center of the smallest ellipsoid containing the reachability set.

with

$$
\begin{aligned}
&\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle \\
&\quad + \sum_{t=0}^{T-1} \langle Q_t^{-1}(r_t-\overline{r}_t), r_t-\overline{r}_t\rangle \\
&\quad + \sum_{t=0}^{T} \langle R_t^{-1}(\eta_t-\overline{\eta}_t), \eta_t-\overline{\eta}_t\rangle \leqslant 1\,.
\end{aligned} \tag{6}
$$

Here $F_t \in \mathbb{R}^{N\times n}$ can be any rectangular matrix and $z_t \in \mathbb{R}^n$ denotes the state of the DAE. If $F_t = I_{N\times N}$, the problem statement is the same as in section 2.1.

Following (Zhuk, 2010), we consider the equation for the minimax gain G_t, for any time $t \in \{0,\dots,T-1\}$:

$$
\begin{aligned}
G_{t+1} &= F_{t+1}^T\left[Q_t^{-1} - Q_t^{-1}\widehat{M}_t B_t \widehat{M}_t^T Q_t^{-1}\right] F_{t+1} \\
&\quad + \hat{H}_{t+1}^T R_{t+1}^{-1}\hat{H}_{t+1}, \\
B_t &= \left(G_t + \widehat{M}_t^T Q_t^{-1}\widehat{M}_t\right)^+ ,
\end{aligned} \tag{7}
$$

with the following initialization:

$$
G_0 = F_0^T Q^{-1} F_0 + \hat{H}_0^T R_0^{-1}\hat{H}_0\,. \tag{8}
$$

For any time $t \in \{0,\dots,T\}$, the minimax estimator is defined as

$$
\hat{z}_t = G_t^+ v_t\,, \tag{9}
$$

with

$$
v_0 = F_0^T Q^{-1}\overline{e} + \hat{H}_0^T R_0^{-1}(y_0 - \overline{\eta}_0)\,, \tag{10}
$$

and, for $t \in \{1,\dots,T\}$,

$$
\begin{aligned}
v_t &= F_t^T Q_{t-1}^{-1}\mathcal{M}_{t-1}(F_{t-1}B_{t-1}v_{t-1}) \\
&\quad + F_t^T\left[Q_{t-1}^{-1} - Q_{t-1}^{-1}\widehat{M}_{t-1}B_{t-1}\widehat{M}_{t-1}^T Q_{t-1}^{-1}\right]\overline{r}_{t-1} \\
&\quad + \hat{H}_t^T R_t^{-1}(y_t - \overline{\eta}_t).
\end{aligned} \tag{11}
$$

For any time $t \in \{0,\dots,T\}$, the reachability set $\mathcal{R}_t$ is defined as

$$
\begin{aligned}
\mathcal{R}_t &= \hat{z}_t + \sqrt{1-\beta_t + \langle G_t\hat{z}_t, \hat{z}_t\rangle}\,\mathcal{X}_t, \\
\mathcal{X}_t &= \{x : \langle G_t x, x\rangle \leqslant 1\}
\end{aligned} \tag{12}
$$

with β_t being a scaling factor depending on observations:

$$
\begin{aligned}
\beta_{t+1} &= \beta_t + \langle R_{t+1}^{-1}(y_{t+1} - \overline{\eta}_{t+1}), y_{t+1} - \overline{\eta}_{t+1}\rangle \\
&\quad - \langle B_t^+ G_t^{-1}\hat{z}_t, G_t^{-1}\hat{z}_t\rangle + \langle \widehat{M}_t^T Q_t \widehat{M}_t \overline{r}_t, \overline{r}_t\rangle
\end{aligned}
$$

We have that the reachability set is a translation of the set $\mathcal{X}_t$ induced by the minimax gain G_t. The shape of $\mathcal{X}_t$ depends only on the model, observation operator and bounding set. $\mathcal{X}_t$ describes how the model propagates uncertain initial condition and model error from the bounding set (6). In contrast to the case of ODE, G_t could be singular so that $\mathcal{X}_t$ contains the kernel of G_t. In fact, the part of the system state lying in that kernel is not observable.

2.3 The case of the non-singular gain

Assume for simplicity that $\overline{r}_t = 0$ and $\overline{e} = 0$. Let us further assume that G_t is positive definite for all time instants t. This is the case when, for instance, F_t, for $t \in \{0, T\}$, is of full column rank, or $F_t^T F_t + \hat{H}_t^T \hat{H}_t$ is positive definite. If G_t is positive definite, then $Q_t + \widehat{M}_t G_t^{-1}\widehat{M}_t^T$ is positive definite, and according to Sherman-Morrison-Woodbury formula (see section 4), its inverse can be written in the form

$$
\begin{aligned}
&\left(Q_t + \widehat{M}_t G_t^{-1}\widehat{M}_t^T\right)^{-1} = Q_t^{-1} \\
&- Q_t^{-1}\widehat{M}_t(G_t + \widehat{M}_t^T Q_t^{-1}\widehat{M}_t)^{-1}\widehat{M}_t^T Q_t^{-1}\,.
\end{aligned} \tag{13}
$$

Using this identity and the gain equation (7), it is possible to write G_{t+1} as

$$
\begin{aligned}
G_{t+1} &= F_{t+1}^T\left(Q_t + \widehat{M}_t G_t^{-1}\widehat{M}_t^T\right)^{-1} F_{t+1} \\
&\quad + \hat{H}_{t+1}^T R_{t+1}^{-1}\hat{H}_{t+1}\,,
\end{aligned} \tag{14}
$$

which gives an alternative form to the filter. It also proves that G_t is positive definite for all $t \in \{0,\dots,T\}$.

The state estimator can be rewritten so that the model is applied directly to $F_t\hat{z}_t$ instead of $F_t B_t v_t$. Although it is an equivalent formulation in the linear case, it can make a huge difference when the model is non-linear. In addition, this alternative form makes it easier to interpret the action of the filter:

$$
\begin{aligned}
\hat{z}_{t+1} &= F_{t+1}^T\mathcal{M}_t(F_t\hat{z}_t) \\
&\quad + G_{t+1}^{-1}\hat{H}_{t+1}^T R_{t+1}^{-1}(y_{t+1} - \overline{\eta}_{t+1} - \hat{H}_{t+1}F_{t+1}^T\mathcal{M}_t(F_t\hat{z}_t)) \\
&\quad + G_{t+1}^{-1}F_{t+1}^T\left(Q_t + \widehat{M}_t G_t^{-1}\widehat{M}_t^T\right)^{-1} \\
&\quad \times (I - F_{t+1}F_{t+1}^T)\mathcal{M}_t(F_t\hat{z}_t)
\end{aligned}
$$

3 Model reduction

We introduce a reduction method which generalizes the classical Galerkin approach. In the later approach, the model state is projected onto a lower-dimensional subspace so that the dynamics of the full state is represented with a small number of scalars. However this reduction can loose some properties of the full model. For instance, the reduced state equation can introduce instabilities that are not in the full model.

Assume that for each time step t, we have a matrix $F_t \in \mathbb{R}^{n\times n}$ whose columns are linearly-independent orthonormal vectors—we therefore have $F_t^T F_t = I_{n\times n}$. We denote by $\mathcal{F}_t$ the linear span of the columns of F_t. The reduction consists in projecting the true state x_t onto this subspace $\mathcal{F}_t$. We introduce $z_t = F_t^T x_t$, which is the vector of the coefficients of the projection of x_t. Consequently we approximate x_t with $F_t z_t$.

3.1 Classical reduction

The main idea of the classical reduction based on the Galerkin projection is to derive the equation for z_t multiplying (1) by F_{t+1}^T:

$$z_{t+1} = F_{t+1}^T x_{t+1} = F_{t+1}^T \mathcal{M}_t(x_t) + F_{t+1}^T e_t \,. \quad (15)$$

Recalling the definition of z_t we obtain

$$\begin{aligned} z_{t+1} &= F_{t+1}^T \mathcal{M}_t(F_t z_t) + F_{t+1}^T e_t \\ &+ F_{t+1}^T \mathcal{M}_t(x_t) - F_{t+1}^T \mathcal{M}_t(F_t z_t)\,. \end{aligned} \quad (16)$$

Let us define

$$p_t = e_t + \mathcal{M}_t(x_t) - \mathcal{M}_t(F_t F_t^T x_t)\,, \quad (17)$$

so that

$$\begin{aligned} z_{t+1} &= F_{t+1}^T \mathcal{M}_t(F_t z_t) + F_{t+1}^T p_t\,, \\ z_0 &= F_0^T(x_0^g + e)\,. \end{aligned} \quad (18)$$

p_t is the sum of the model error and a reduction error. If we were to apply the extended minimax filter on the reduced state equation (18), we would need to evaluate the range of values that p_t can take. Since p_t is state dependent and since the true state is unknown, it is hard to determine the range of p_t. The natural approach to suppress the state dependence is to bound the reduction error for all plausible states. Hence we may assume that

$$\|p_t\| \leqslant \|e_t\| + \delta_t\,,$$

where, for instance, δ_t is guaranteed to exist for Lipschitz continuous models provided $F_t F_t^T x_t$ approximates x_t with finite error. With possibly modified Q, Q_t and R_t, we write

$$\begin{aligned} &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle + \sum_{t=0}^{T}\langle R_t^{-1}(\eta_t-\overline{\eta}_t), \eta_t-\overline{\eta}_t\rangle \\ &+ \sum_{t=0}^{T-1}\langle Q_t^{-1}(p_t-\overline{p}_t), (p_t-\overline{p}_t)\rangle \leqslant 1 \end{aligned} \quad (19)$$

for p_t defined by (17) and some $\overline{p}_t$ defined as a systematic error of the new model error. Note that only $F_{t+1}^T p_t$ has an impact onto dynamics of z_t. Noting that

$$F_{t+1}^T p_t = F_{t+1}^T F_{t+1} F_{t+1}^T p_t\,,$$

we see that it is enough to have a bound on $F_{t+1}F_{t+1}^T p_t$ only. Thus we can consider an ellipsoid in the form

$$\begin{aligned} &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle + \sum_{t=0}^{T}\langle R_t^{-1}(\eta_t-\overline{\eta}_t), \eta_t-\overline{\eta}_t\rangle \\ &+ \sum_{t=0}^{T-1}\langle (F_t^T Q_t^{-1} F_t) F_t^T(p_t-\overline{p}_t), F_t^T(p_t-\overline{p}_t)\rangle \leqslant 1 \end{aligned}$$

Now we stress that the above procedure could lead to the overestimation of the reachability set of the reduced model (18). This is a consequence of the model reduction (we replace $\mathcal{M}_t$ with $F_{t+1}^T\mathcal{M}_t(F_t)$) and the suppression of state-dependence in the reduction error.

3.2 Generalized reduction by means of DAE

Above we have seen that the state estimation problem for (18) could be affected by instability of the reduced model so that the reachability set could rapidly expand in time although the reachability set of the full model behaves differently. In what follows we propose a way to further constraint the size of the reachability set for the reduced state, while relying on the same reduction-error estimations as previously.

Consider the reduced model

$$\begin{aligned} z_{t+1} &= F_{t+1}^T \mathcal{M}_t(F_t z_t) + F_{t+1}^T p_t\,, \\ z_0 &= F_0^T(x_0^g + e)\,, \end{aligned} \quad (20)$$

and the associated error description

$$\begin{aligned} &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle \\ &+ \sum_{t=0}^{T-1}\langle (F_t^T Q_t^{-1} F_t) F_t^T(p_t-\overline{p}_t), F_t^T(p_t-\overline{p}_t)\rangle \\ &+ \sum_{t=0}^{T}\langle R_t^{-1}(\eta_t-\overline{\eta}_t), \eta_t-\overline{\eta}_t\rangle \leqslant 1\,. \end{aligned} \quad (21)$$

We introduce an additional constraint onto the reduced state:

$$\sum_{t=0}^{T-1}\langle S_t^{-1} L_t z_t, L_t z_t\rangle \leqslant 1 \quad (22)$$

where S_t^{-1} is a $s\times s$-symmetric positive-definite matrix defining the shape of the bounding set for the reduced state, and $L_t \in \mathbb{R}^{s\times n}$ is a design parameter allowing to constraint a desired part of the reduced state or just a linear combination of the reduced state's components. We do not impose any conditions on L_t. Now we note that the energy constraint can be incorporated into (20)–(21) using the following construction:

$$\begin{aligned} z_{t+1} &= F_{t+1}^T \mathcal{M}_t(F_t z_t) + F_{t+1}^T p_t, \\ z_0 &= F_0^T(x_0^g + e), L_t z_t + w_t = 0\,, \end{aligned} \tag{23}$$

with

$$\begin{aligned} &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle \\ &+ \sum_{t=0}^{T-1} \langle (F_t^T Q_t^{-1} F_t) F_t^T (p_t - \overline{p}_t), F_t^T (p_t - \overline{p}_t)\rangle \\ &+ \sum_{t=0}^{T} \langle R_t^{-1}(\eta_t - \overline{\eta}_t), \eta_t - \overline{\eta}_t\rangle \\ &+ \sum_{t=0}^{T-1} \langle S_t^{-1}(w_t - \overline{w}_t), (w_t - \overline{w}_t)\rangle \leqslant 1+1 \end{aligned} \tag{24}$$

where $\overline{w}_t$ is a parameter. The bounding set is defined with (24), which introduces a link between the reduced state and the full state through $(I - F_{t+1}F_{t+1}^T)x_{t+1}$. This allows to limit the artificial increase of the reachability set due to the reduction.

3.3 Extended minimax state estimation for DAE

After the considerations of the previous section, we introduced the following filtering problem:

$$\begin{aligned} &z_{t+1} = F_{t+1}^T \mathcal{M}_t(F_t z_t) + F_{t+1}^T p_t, \\ &z_0 = F_0^T(x_0^g + e), \quad L_t z_t + w_t = 0, \\ &\langle Q^{-1}(e-\overline{e}), e-\overline{e}\rangle \\ &+ \sum_{t=0}^{T} \langle R_t^{-1}(\eta_t - \overline{\eta}_t), \eta_t - \overline{\eta}_t\rangle \\ &+ \sum_{t=0}^{T-1} \langle (F_t^T Q_t^{-1} F_t) F_t^T p_t, F_t^T p_t\rangle \\ &+ \sum_{t=0}^{T-1} \langle S_t^{-1} w_t, w_t\rangle \leqslant 1+1\,. \end{aligned} \tag{25}$$

We define a descriptor matrix $\widetilde{F} = \left[\begin{smallmatrix} I_{n\times n} \\ 0 \end{smallmatrix}\right]$. We can extend the model and its associated error matrix: the new DAE model is $\widetilde{\mathcal{M}_t} = \left[\begin{smallmatrix} F_t^T \mathcal{M}_t F_t \\ L_t \end{smallmatrix}\right]$ and $\widetilde{Q}_t = \left[\begin{smallmatrix} F_t^T Q_t^{-1} F_t & 0 \\ 0 & S_t^{-1} \end{smallmatrix}\right]$. With these definition, we can apply the (extended) DAE minimax filter from section 2.2, simply by substituting F_t with $\widetilde{F}$, $\mathcal{M}_t$ with $\widetilde{\mathcal{M}_t}$ and Q_t with $\widetilde{Q}_t$. Also R_t should be modified to take into account the additional error due to reduction in the observation equation, since $\mathcal{H}_t(F_t z_t)$ is involved instead of $\mathcal{H}_t(x_t)$.

The choice of L_t and S_t is a key point in this approach, since they allow to constrain the reduced state components. This additional constraint is the main difference with the Galerkin approach where instabilities may occur because no constraint is enforced.

3.4 Illustration with an air quality model

We illustrate the difference between the classical reduction and our approach by applying the filter, without observations, to an air quality model from the modeling system Polyphemus (Mallet et al., 2007). The model essentially solves a set of reactive transport equations in three dimensions and with 72 reacting chemical species. The transport is modeled with advection (due to the wind) and diffusion (which models turbulence). The model is used in an operational configuration. The state vector contains over one million components, but we carry out the reduction on $10,000$ components only. The reduced state has 30 components. We take $L_t = (I_{N\times N} - F_{t+1}F_{t+1}^T)M_t F_t$ and $S_t = 10^8 I_{N\times N}$ which means we essentially prevent the model to contribute to the complement of the reduced space. Figure 1 shows the ground-level ozone concentration field at a given time, for the model without reduction, for the model with classical reduction and for the model with minimax filtering. On average (over several time steps), the root mean square discrepancy between the fields with reduction and the field without reduction is $3.8\ \mu\mathrm{g\,m^{-3}}$ in the classical case and $0.15\ \mu\mathrm{g\,m^{-3}}$ with minimax filtering. This shows that, even for a complex model and even without assimilation of observations, the reduction within the minimax framework provides significant improvement over the classical approach.

4 Sherman-Morrison-Woodbury formula

The Sherman-Morrison-Woodbury matrix identity is

$$\begin{aligned} &(S + N_1 W N_2)^{-1} = S^{-1} \\ &- S^{-1}N_1(W^{-1} + N_2 S^{-1} N_1)^{-1} N_2 S^{-1} \end{aligned} \tag{26}$$

if S and W are nonsingular matrices.

5 Conclusion

A new reduction method base on minimax state estimation is presented. The main idea is 1) to convert a high-dimensional ODE into a low-dimensional DAE and 2) to apply an extended version of the state estimation algorithm developed in (Zhuk, 2010) to the resulting low-dimensional DAE. The proposed approach addresses instability issues arising in classical reduction methods based on the Galerkin projection.

The next step is the application to real-life high-dimensional models of the reduced version of the filter with assimilation. One objective is to compare with classical reduced Kalman filters where the reduction is carried out in the propagation of error variances, not

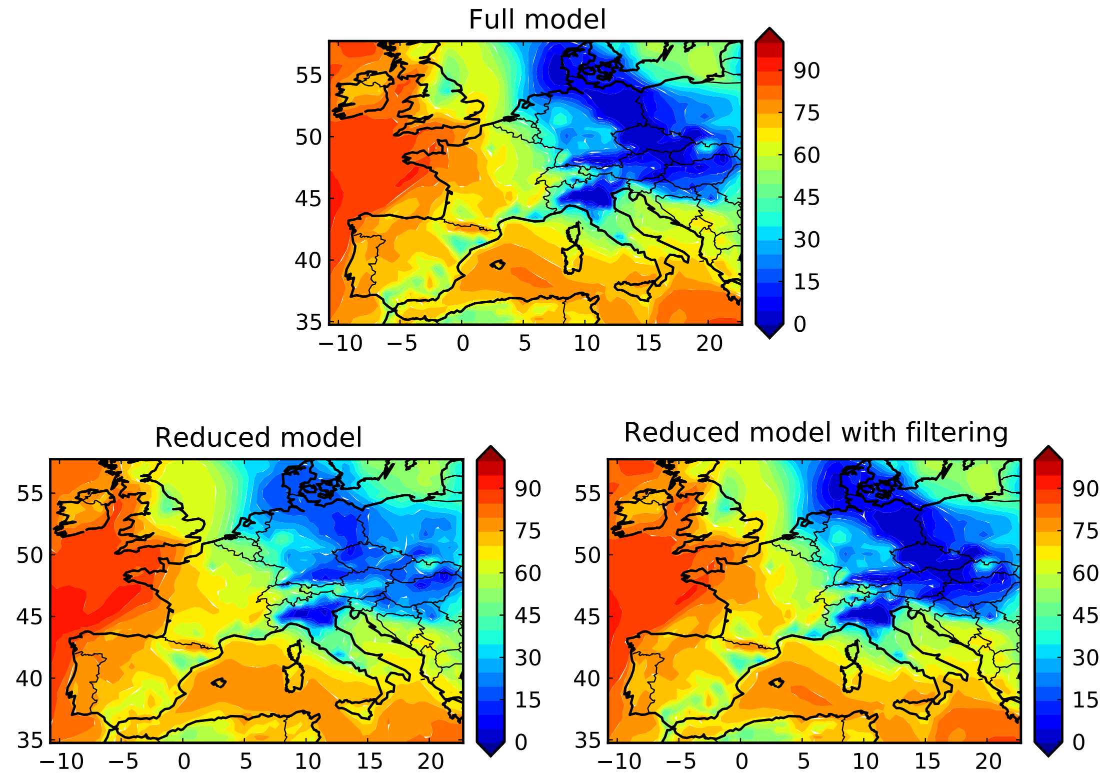

Figure 1. Comparison between ground-level ozone concentrations ($\mu g\ m^{-3}$) as simulated for the 1^{st} January 2001 at 0500 UTC, without reduction (top), with classical reduction (left) and with minimax reduction (right).

on the model state. One target application is the same high-dimensional air quality models as in section 3.4. Another target is to devise an algorithm of reduced basis generation (the matrix F_t) allowing to take into account available observations of the model state. In this direction it is important to generate a robust reduced basis as initial conditions and model errors are supposed to be uncertain.

Acknowledgement

This work was carried out during the second author's tenure of an ERCIM "Alain Bensoussan" Fellowship Programme. This Programme is supported by the Marie Curie Co-funding of Regional, National and International Programmes (COFUND) of the European Commission.

References

Balakrishnan, A.: 1984, *Kalman Filtering Theory*, Opt. Soft., Inc., N.Y.

Bertsekas, D. and Rhodes, I. B.: 1971, Recursive state estimation with a set-membership description of the uncertainty, *IEEE Trans. Automat. Contr.* **AC-16**, 117–128.

Brenner, S. and Scott, R.: 2005, *The Mathematical Theory of Finite Element Methods*, 2nd edn, Springer. ISBN 0-3879-5451-1.

Chernousko, F. L.: 1994, *State Estimation for Dynamic Systems* ., Boca Raton, FL: CRC.

Heemink, A. W., Verlaan, M. and Segers, A. J.: 2001, Variance reduced ensemble Kalman filtering, *Mon. Wea. Rev.* **129**, 1,718–1,728.

Homescu, C., Petzold, L. R. and Serban, R.: 2005, Error estimation for reduced-order models of dynamical systems, *SIAM J. Numer. Anal.* **43**(4), 1,693–1,714.

Julier, S. J. and Uhlmann, J. K.: 1997, A new extension of the Kalman filter to nonlinear systems, *Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*.

Kurzhanski, A. and Vályi, I.: 1997, *Ellipsoidal calculus for estimation and control*, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA.

Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, É., Korsakissok, I., Wu, L., Roustan, Y., Sartelet, K., Tombette, M. and Foudhil, H.: 2007, Technical Note: The air quality modeling system Polyphemus, *Atmos. Chem. Phys.* **7**(20), 5,479–5,487.

Milanese, M. and Tempo, R.: 1985, Optimal algorithms theory for robust estimation and prediction, *IEEE Trans. Automat. Control* **30**(8), 730–738.
URL: *http://dx.doi.org/10.1109/TAC.1985.1104056*

Moireau, P. and Chapelle, D.: 2010, Reduced-order unscented Kalman filtering with application to parameter identification in large-dimensional systems, *ESAIM: Control, Opt. and Calc. Var.* .

Nakonechny, A.: 2004, *Optimal control and estimation for partial differential equations*, National Taras Shevchenko University Publishing. (in Ukrainian).

Pham, D. T.: 2001, Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Mon. Wea. Rev.* **129**, 1,194–1,207.

Verlaan, M. and Heemink, A. W.: 1995, Reduced rank square root filters for large scale data assimilation problems, *Second International Symposium on Assimilation of Observations in Meteorology and Oceanography*, Japan, pp. 247–252.

Wu, L., Mallet, V., Bocquet, M. and Sportisse, B.: 2008, A comparison study of data assimilation algorithms for ozone forecasts, *J. Geophys. Res.* **113**(D20310).

Zhuk, S.: 2010, Minimax state estimation for linear discrete-time differential-algebraic equations, *Automatica J. IFAC* **46**(11), 1785–1789.