The Back and Forth Nudging algorithm for data assimilation problems: theoretical results on transport equations

Didier Auroux[*†] Maëlle Nodet [‡§]

28th September 2008

Abstract

English:
In this paper, we consider the back and forth nudging algorithm that has been introduced for data assimilation purposes. It consists of iteratively and alternately solving forward and backward in time the model equation, with a feedback term to the observations. We consider the case of 1-dimensional transport equations, either viscous or inviscid, linear or not (Bürgers' equation). Our aim is to prove some theoretical results on the convergence, and convergence properties, of this algorithm. We show that for non viscous equations (both linear transport and Bürgers), the convergence of the algorithm holds under observability conditions. Convergence can also be proven for viscous linear transport equations under some strong hypothesis, but not for viscous Bürgers' equation. Moreover, the convergence rate is always exponential in time. We also notice that the forward and backward system of equations is well posed when no nudging term is considered.

French:
Ce travail étudie l'algorithme du nudging direct et rétrograde, qui a été introduit en assimilation de données. Cet algorithme consiste à résoudre itérativement l'équation du modèle, agrémentée d'un terme de rappel aux observations, dans le sens direct puis dans le sens rétrograde. Dans ce travail nous nous intéressons aux équations de transport en dimension 1, avec ou sans viscosité, linéaires ou non (Bürgers). Notre objectif est d'étudier la convergence éventuelle, et la vitesse de convergence le cas échéant, de cet algorithme. Nous prouvons que, pour les équations

[*]Institut de Mathématiques de Toulouse, Université Paul Sabatier Toulouse 3, 31062 Toulouse cedex 9, France; auroux@math.univ-toulouse.fr

[†]INRIA, Grenoble, France

[‡]Université de Grenoble, Laboratoire Jean Kuntzmann, UMR 5224, Grenoble, France; maelle.nodet@inria.fr

[§]INRIA, Grenoble, France

non visqueuses (linéaire ou Bürgers), la convergence a lieu sous des hypothèses d'observabilité. La convergence peut aussi être démontrée pour des équations de transport linéaires visqueuses sous des hypothèses fortes, mais pas pour l'équation de Bürgers visqueuse. En outre, lorsque la convergence a lieu, la vitesse de convergence est toujours exponentielle en temps. Nous remarquons aussi que le système d'équations directe et rétrograde est toujours bien posé lorsqu'aucun terme de rappel n'est présent.

1 Introduction and main results

Data assimilation is the set of techniques aiming to combine in an optimal way the mathematical information provided by the model equations and the physical information given by observations, in order to retrieve the state of a system. Several types of methods have been widely studied in the past decades. We can cite here interpolation, variational and stochastic methods. The first ones interpolate the measurements from the points of observation towards the grid points, the interpolation being weighted by the statistics of the observations [Kal03]. Variational methods are based on the optimal control theory, and data assimilation is set as being a problem of constrained optimization. The goal is to minimize a cost function measuring the difference between the observations and the corresponding quantities provided by a model integration. The initial condition of the system can then be seen as a control vector [LDT86]. Finally, the basic idea of stochastic methods is to consider the fields as the realization of a stochastic process and carry out Kalman filtering methods [Kal60, EvL99]. We can also mention one of the very first data assimilation schemes: the nudging method. Also known as Newtonian relaxation or dynamic initialization, it consists of adding a feedback term to the observations directly in the model equations [HA76].

All these methods require extensive work, either from the implementation or the computation point of view. For instance, variational methods require the linearization of all operators and also the implementation of the adjoint model. They also need efficient optimization schemes, as the minimization is performed on spaces of huge dimension. On the other side, stochastic methods are somewhat easier to implement, but they require knowledge, storage and manipulations of huge matrices.

The Back and Forth Nudging (BFN) algorithm has recently been introduced as a simple and efficient method for solving data assimilation problems [AB05]. In most geophysical applications, data assimilation consists of estimating a trajectory, solution of a partial differential equation (PDE), from the knowledge of observations. These observations are usually sparse in time and space, and incorrect in the sense that they are not the restriction of a solution of the PDE model. One step of the BFN algorithm consists of solving first the model equation, in which a feedback to the observation solution is added, and then the same equation but backwards in time, with also a feedback term to the observations. Such forward and backward integrations provide a new value of the solution at the initial time $t = 0$ and the aim of the BFN is to improve the quality of the initial condition.

The idea of the back and forth nudging is to use the difference between the observations and the model trajectory as a feedback control of the equations, both in the forward and backward integrations. This makes the numerical scheme extremely easy to implement, in comparison with both variational and stochastic methods, as we usually only consider diagonal (or even scalar) gain matrices. The back and forth nudging scheme can also be seen as an intermediate scheme between variational and stochastic methods, as the standard nudging technique has both variational (minimization of a compromise between the observations and the energy of the system) and stochastic (sub-optimal Kalman filter) interpretations [AB08].

As a first approximation, we consider in this paper that the observations are correct (i.e. no observation error), and hence the observations satisfy the model equation. We consider various observation domains: first we assume that the observations $u_{obs}(t, x)$ are available for any point x and time t, second we assume that they are available for $t \in [t_1, t_2]$ and for all x, and third we consider that they are available for all t over a given space domain. This is done through the time and space dependency of the feedback (or nudging) gain matrix $K(t, x)$ that is equal to 0 when the observations are not available.

Many numerical experiments in almost realistic situations suggest that this algorithm works well, and that the identified solution gets closer to the observations [AB08]. The goal of this paper is to prove some theoretical results and convergence properties in the particular case of transport equations, either viscous or inviscid, either linear or non-linear (Bürgers' equation).

In section 2, we consider one step of the BFN algorithm applied to a linear viscous transport equation:

$$
\begin{array}{ll}
(F) & \left\{ \begin{array}{rcl} \partial_t u - \nu \partial_{xx} u + a(x)\partial_x u & = & -K(u - u_{obs}) \\ u|_{x=0} = u|_{x=1} & = & 0 \\ u|_{t=0} & = & u_0 \end{array} \right. \\
(B) & \left\{ \begin{array}{rcl} \partial_t \tilde{u} - \nu \partial_{xx} \tilde{u} + a(x)\partial_x \tilde{u} & = & K'(\tilde{u} - u_{obs}) \\ \tilde{u}|_{x=0} = \tilde{u}|_{x=1} & = & 0 \\ \tilde{u}|_{t=T} & = & u(T) \end{array} \right.
\end{array}
\tag{1}
$$

where the following notations hold for all further cases:

- the time period considered here is $t \in [0, T]$;
- the first equation (F) is called the forward equation, the second one (B) is called the backward one;
- K and K' are positive and may depend on t and x, but for simplicity reasons, we will always assume that there exists a constant $\kappa \in \mathbb{R}_+^*$ such that $K'(t, x) = \kappa K(t, x)$;
- $a(x) \in W^{1,\infty}(\Omega)$, Ω being the considered space domain, either the interval $[0, 1]$ or the torus $[0, 1]$;

- ν is a constant;
- u_{obs} is a solution of the forward equation with initial condition u_{obs}^0:

$$\left\{ \begin{array}{rcl} \partial_t u_{obs} - \nu \partial_{xx} u_{obs} + a(x) \partial_x u_{obs} & = & 0 \\ u|_{x=0} = u|_{x=1} & = & 0 \\ u|_{t=0} & = & u_{obs}^0 \end{array} \right. \tag{2}$$

The following result holds true:

Theorem 1 *We consider the one-step BFN (1) with observations u_{obs} satisfying (3). We denote*

$$\begin{array}{rcl} w(t) & = & u(t) - u_{obs}(t) \\ \widetilde{w}(t) & = & \widetilde{u}(t) - u_{obs}(t) \end{array} \tag{3}$$

Then we have:

1. *If $K(t,x) = K$, then we have, for all $t \in [0,T]$:*

$$\widetilde{w}(t) = e^{(-K-K')(T-t)} w(t) \tag{4}$$

2. *If $K(t,x) = K(x)$, with Support $(K) \subset [a,b]$ where $a < b$ and $a \neq 0$ or $b \neq 1$, then equation (1) is ill-posed: there does not exist a solution $(u, \widetilde{u})$, in general.*

3. *If $K(t,x) = K \mathbb{1}_{[t_1,t_2]}(t)$ with $0 \leq t_1 < t_2 \leq T$, then we have*

$$\widetilde{w}(0) = e^{(-K-K')(t_2-t_1)} w(0) \tag{5}$$

In section 3, we consider one step of the BFN algorithm applied to the viscous Bürgers' equation:

$$\begin{array}{ll} (F) & \left\{ \begin{array}{rcl} \partial_t u - \nu \partial_{xx} u + u \partial_x u & = & -K(u - u_{obs}) \\ u|_{x=0} = u|_{x=1} & = & 0 \\ u|_{t=0} & = & u_0 \end{array} \right. \\ (B) & \left\{ \begin{array}{rcl} \partial_t \widetilde{u} - \nu \partial_{xx} \widetilde{u} + \widetilde{u} \partial_x \widetilde{u} & = & K'(\widetilde{u} - u_{obs}) \\ \widetilde{u}|_{x=0} = \widetilde{u}|_{x=1} & = & 0 \\ \widetilde{u}|_{t=T} & = & u(T) \end{array} \right. \end{array} \tag{6}$$

with the same notations as before.
The observations u_{obs} satisfy the forward Bürgers' equation:

$$\left\{ \begin{array}{rcl} \partial_t u_{obs} - \nu \partial_{xx} u_{obs} + u_{obs} \partial_x u_{obs} & = & 0 \\ u|_{x=0} = u|_{x=1} & = & 0 \\ u|_{t=0} & = & u_{obs}^0 \end{array} \right. \tag{7}$$

We have the following result if $K \neq 0$:

Theorem 2 *The BFN iteration (6) for viscous Bürgers' equation, with observations u_{obs} satisfying (7), is ill-posed, even when $K(t,x)$ is constant (except for $K(t,x) \equiv 0$): there does not exist, in general, a solution $(u, \tilde{u})$.*

In the particular case when $K = K' = 0$, the backward problem is ill-posed in the sense of Hadamard, but it has a unique solution if the final condition $\tilde{u}|_{t=T}$ is set to a final solution of the direct equation. Moreover, in this particular case, the backward solution is exactly equal to the forward one: $\tilde{u}(t) = u(t)$ for all $t \in [0,T]$. The main result is the following:

Proposition 1 *If $K = K' \equiv 0$, then problem (6) is well-posed in the sense of Hadamard, and there exists a unique solution $(u, \tilde{u})$. Moreover $u = \tilde{u}$.*

Section 4 considers the extension of theorem 1 to the inviscid case, for both linear transport and Bürgers' equations.
We first consider the linear case. The BFN equations are:

$$\begin{array}{ll} (F) & \left\{ \begin{array}{rcl} \partial_t u + a(x)\partial_x u &=& -K(u - u_{obs}) \\ u|_{x=0} &=& u|_{x=1} \\ \partial_x u|_{x=0} &=& \partial_x u|_{x=1} \\ u|_{t=0} &=& u_0 \end{array} \right. \\ (B) & \left\{ \begin{array}{rcl} \partial_t \tilde{u} + a(x)\partial_x \tilde{u} &=& K'(\tilde{u} - u_{obs}) \\ \tilde{u}|_{x=0} &=& \tilde{u}|_{x=1} \\ \partial_x \tilde{u}|_{x=0} &=& \partial_x \tilde{u}|_{x=1} \\ \tilde{u}|_{t=T} &=& u(T) \end{array} \right. \end{array} \tag{8}$$

where $a(x)$ can be constant or not.

Theorem 3 *We consider the non viscous one-step BFN (8), with observations u_{obs} satisfying (8-F) with $K = 0$. We denote*

$$\begin{array}{rcl} w(t) &=& u(t) - u_{obs}(t) \\ \tilde{w}(t) &=& \tilde{u}(t) - u_{obs}(t) \end{array} \tag{9}$$

We denote by

$$(s, \psi(s,x)) \tag{10}$$

the characteristic curve of equation (8-F) with $K = 0$, with foot x in time $s = 0$, i.e. such that

$$(s, \psi(s,x))|_{s=0} = (0, x) \tag{11}$$

We assume that the final time T is such that the characteristics are well defined and do not intersect over $[0,T]$.
Then we have:

1. *If $K(t,x) = K$, then we have, for all $t \in [0,T]$:*

$$\tilde{w}(t) = w(t)e^{(-K-K')(T-t)} \tag{12}$$

2. *If $K(t,x) = K\mathbb{1}_{[t_1,t_2]}(t)$ with $0 \le t_1 < t_2 \le T$, then we have*

$$\widetilde{w}(0) = w(0)e^{(-K-K')(t_2-t_1)} \tag{13}$$

3. *If $K(t,x) = K(x)$, then we have, for all $t \in [0,T]$:*

$$\widetilde{w}(t,\psi(t,x)) = w(t,\psi(t,x)) \exp\left(-\int_t^T K(\psi(s,x)) + K'(\psi(s,x))\, ds\right) \tag{14}$$

We finally consider non viscous Bürgers' equation, still with periodic boundary conditions, and for a time T such that there is no shock in the interval $[0,T]$:

$$\begin{array}{ll}
(F) & \left\{\begin{array}{rcl}
\partial_t u + u\partial_x u &=& -K(u-u_{obs}) \\
u|_{x=0} &=& u|_{x=1} \\
\partial_x u|_{x=0} &=& \partial_x u|_{x=1} \\
u|_{t=0} &=& u_0
\end{array}\right. \\
(B) & \left\{\begin{array}{rcl}
\partial_t \widetilde{u} + \widetilde{u}\partial_x \widetilde{u} &=& K'(\widetilde{u}-u_{obs}) \\
\widetilde{u}|_{x=0} &=& \widetilde{u}|_{x=1} \\
\partial_x \widetilde{u}|_{x=0} &=& \partial_x \widetilde{u}|_{x=1} \\
\widetilde{u}|_{t=T} &=& u(T)
\end{array}\right.
\end{array} \tag{15}$$

Theorem 4 *We consider the non viscous one-step BFN (15), with observations u_{obs} satisfying (15-F) with $K = 0$. We denote*

$$\begin{array}{rcl}
w(t) &=& u(t) - u_{obs}(t) \\
\widetilde{w}(t) &=& \widetilde{u}(t) - u_{obs}(t)
\end{array} \tag{16}$$

We assume that $u_{obs} \in W^{1,\infty}([0,T]\times\Omega)$, i.e. there exists $M > 0$ such that

$$|\partial_x u_{obs}(t,x)| \le M, \quad \forall t \in [0,T], \forall x \in \Omega \tag{17}$$

Then we have:

1. *If $K(t,x) = K$, then we have, for all $t \in [0,T]$:*

$$||\widetilde{w}(t)|| \le e^{(-K-K'+M)(T-t)}||w(t)|| \tag{18}$$

2. *If $K(t,x) = K\mathbb{1}_{[t_1,t_2]}(t)$ with $0 \le t_1 < t_2 \le T$, then we have*

$$||\widetilde{w}(0)|| \le e^{(-K-K')(t_2-t_1)+MT}||w(0)|| \tag{19}$$

Proposition 2 *We consider one forward (resp. backward) BFN step of the non viscous Bürgers equation (15-F) (resp. (15-B)). With the notations of theorem 4, if $K(t,x) = K(x)$, then we have*

$$w(T,\psi(T,x)) = w(0,x)\exp\left(-\int_0^T K(\psi(\sigma,x))d\sigma - \int_0^T \partial_x u_{obs}(\sigma,\psi(\sigma,x))d\sigma\right) \tag{20}$$

Remark 1 *For the special case* $K(t,x) = K(x) = K\mathbb{1}_{[a,b]}(x)$ *where* K *is a constant and* $[a,b]$ *is a non-empty sub-interval of* $[0,1]$, *we have*

$$w(T,\psi(T,x)) = w(0,x)\exp\left(-K\chi(x) - \int_0^T \partial_x u_{obs}(\sigma,\psi(\sigma,x))d\sigma\right) \qquad (21)$$

where

$$\chi(x) = \int_0^T \mathbb{1}_{[a,b]}(\psi(\sigma,x))d\sigma \qquad (22)$$

is the time during which the characteristic curve $\psi(\sigma,x)$ *with foot* x *of equation (15-F) with* $K=0$ *lies in the the support of* K. *The system is then* observable *if and only if the function* χ *has a non-zero lower bound, i.e.* $m := \min_x \chi(x) > 0$, *the observability being defined by (see [Rus78]):*

$$\exists C, \forall u \text{ solution of (15-F) with } K=0, \quad ||u(T,.)||^2 \leq C\int_0^T ||K(.)u(s,.)||^2\, ds$$

In this case, proposition 2 proves the global exponential decrease of the error, provided K *is larger than* $\frac{MT}{m}$, *where* M *is defined by equation (17).*

From remark 1, we can easily deduce that if for each iteration, both in the forward and backward integrations, the observability condition is satisfied, then the algorithm converges. Note that this is not a necessary condition, as even if $\chi(x) = 0$, the last exponential of equation (21) is bounded.

Note also that in real geophysical applications (either meteorology or oceanography), there is usually no viscosity. In this case, assuming the observability condition, the BFN algorithm is well posed, and theorem 4 and proposition 2 say that the solution tends to the observation trajectory everywhere, and not only on the support of K. From a numerical point of view, we can observe that even with discrete and sparse observations in space, the numerical solution is corrected everywhere [AB08]. We also observed that with a not too large viscosity coefficient, the behavior of the algorithm remains unchanged.

Figure 1 illustrates the results given in theorem 3 in the case 3 (top) and proposition 2 and remark 1 (bottom). These numerical results correspond to a simple case: $u_{obs} \equiv 0$, $u_0(x) = \alpha\sin(2\pi x)$, $K = K' = \mathbb{1}_{[0;0.5]}(x)$. Various final times T are considered, from 0.05 to 1, and both figures show the following expression

$$-\log\left(\frac{\widetilde{w}(0,x)}{w(0,x)}\right) \qquad (23)$$

as a function of $x \in [0;1]$. Figure 1-top illustrates equation (14). The best possible decrease rate is then $\max(K+K') \times T = 2T$. In the linear case, the transport is $a(x) \equiv 1$. As half of the domain is observed, the observability condition is satisfied iff $T > 0.5$, and this is confirmed by the figure. Concerning Bürgers' equation, figure 1-bottom illustrates equation (20). After one iteration of BFN, the best possible decrease rate is also $2T$. We can see that in this

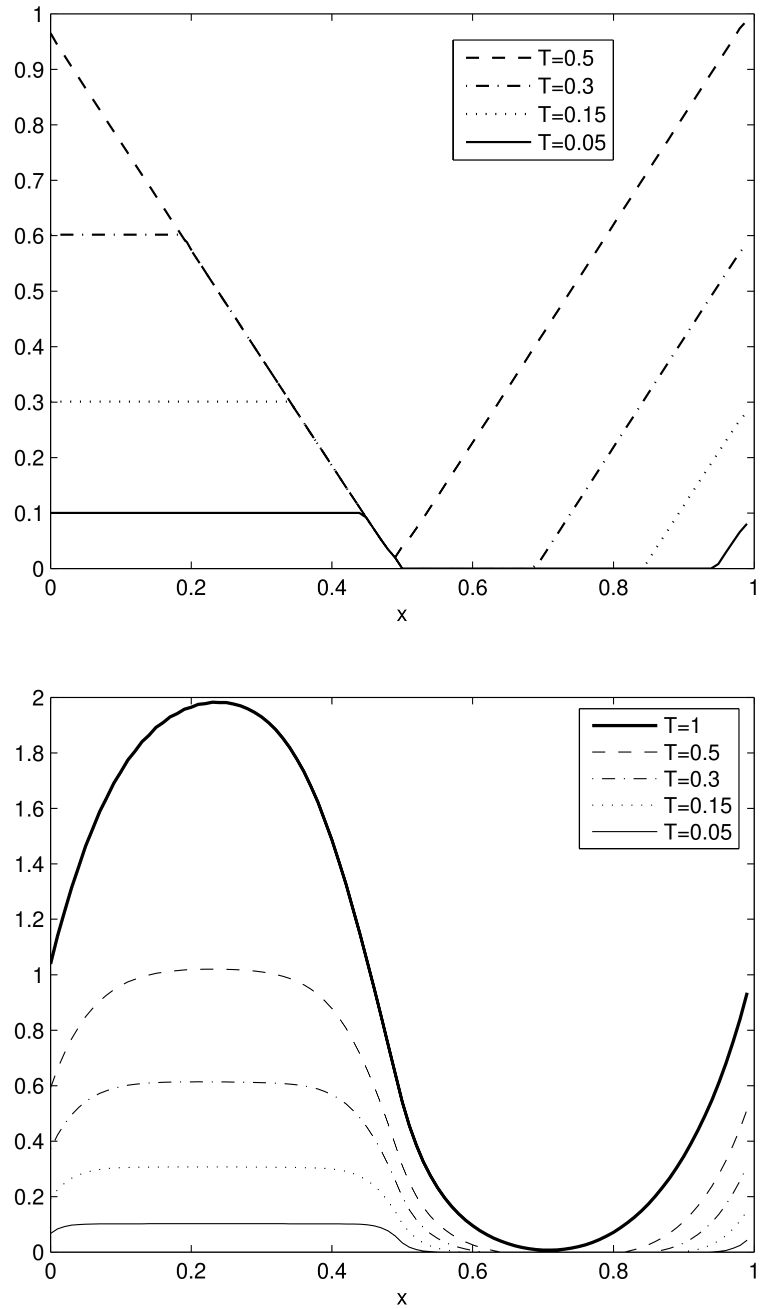

Figure 1: Decrease rate of the error after one iteration of BFN (see equation 23) as a function of x, for various times T; top: linear transport equation; bottom: inviscid Bürgers' equation.

case, due to the nonlinearities of the model, the solution is less corrected on $[0; 0.1]$ but more on $[0.5; 0.6]$. From this figure, we can see that the observability condition is satisfied for T larger than approximately 1.

Finally, some conclusions are given in section 5.

2 Linear transport equation with a viscous term

In this section we prove theorem 1.

2.1 Case 1: K constant

The differences w and $\widetilde{w}$ satisfy the following equations:

$$\left\{\begin{array}{rcl}\partial_t w - \nu\partial_{xx} w + a(x)\partial_x w + Kw & = & 0\\ w|_{x=0} = w|_{x=1} & = & 0\\ w|_{t=0} & = & w_0\\ \partial_t \widetilde{w} - \nu\partial_{xx}\widetilde{w} + a(x)\partial_x\widetilde{w} - K'\widetilde{w} & = & 0\\ \widetilde{w}|_{x=0} = \widetilde{w}|_{x=1} & = & 0\\ \widetilde{w}|_{t=T} & = & w(T)\end{array}\right. \tag{24}$$

We denote by S_+ and S_- the operators associated to these equations, seen as forward equations on $[t_0, t]$ with initial conditions given in t_0:

$$S_+(t_0,t)(w(t_0)) = w(t), \quad S_-(t_0,t)(\widetilde{w}(t_0)) = \widetilde{w}(t) \tag{25}$$

The BFN algorithm has a solution if and only if we have

$$w(T) \in \mathrm{Im}\,(S_-(0,T)) \tag{26}$$

We re-write equation (25) associated to w:

$$\left\{\begin{array}{rcl}\partial_t w - \nu\partial_{xx} w + a(x)\partial_x w - K'w & = & (-K-K')w\\ w|_{x=0} = w|_{x=1} & = & 0\\ w|_{t=0} & = & w_0\end{array}\right. \tag{27}$$

so that we have, thanks to Duhamel's formula:

$$w(t) = S_-(0,t)(w_0) + \int_0^t S_-(s,t)((-K-K')w(s))\,ds \tag{28}$$

If we assume that the expected result is true, i.e. $w(T) \in \mathrm{Im}\,(S_-(0,T))$, then we can assume that it is also true for all t, i.e. we can assume that:

$$\forall t, \exists\varphi(t), w(t) = S_-(0,t)\varphi(t) \tag{29}$$

In that case, we replace (29) in (28) and we get:

$$w(t) = S_-(0,t)(w_0) + \int_0^t S_-(s,t)((-K-K')S_-(0,s)\varphi(s))\,ds \tag{30}$$

As the equation is linear, the scalar coefficient $(K+K')$ commutes with S_- and we get:

$$\begin{aligned} w(t) &= S_-(0,t)(w_0) + (-K-K')\int_0^t S_-(s,t)(S_-(0,s)\varphi(s))\,ds \\ &= S_-(0,t)(w_0) + (-K-K')S_-(0,t)\int_0^t \varphi(s)\,ds \\ S_-(0,t)\varphi(t) &= S_-(0,t)\left[(w_0) + (-K-K')\int_0^t \varphi(s)\,ds\right] \end{aligned} \tag{31}$$

So that we have

$$\varphi(t) = w_0 + (-K-K')\int_0^t \varphi(s)\,ds \tag{32}$$

i.e., φ satisfies

$$\varphi'(t) = (-K-K')\varphi, \quad \varphi(0) = w_0 \tag{33}$$

and finally

$$\varphi(t) = w_0 e^{(-K-K')t} \tag{34}$$

so that we get for $w(T)$:

$$w(T) = S_-(0,T)(w_0 e^{(-K-K')T}) \tag{35}$$

Reciprocally, setting

$$\widetilde{w}(0) = w_0 e^{(-K-K')T} \tag{36}$$

leads to $\widetilde{w}$ satisfying $\widetilde{w}(T) = w(T)$, so that $(w,\widetilde{w})$ is the solution of the one-step BFN (24).
Moreover, we have, for all $t \in [0,T]$:

$$\begin{aligned} \widetilde{w}(t) &= S_-(0,t)(\widetilde{w}(0)) \\ &= S_-(0,t)(w_0 e^{(-K-K')T}) \\ &= e^{(-K-K')(T-t)} S_-(0,t)(w_0 e^{(-K-K')t}) \\ &= e^{(-K-K')(T-t)} S_-(0,t)(\varphi(t)) \\ &= e^{(-K-K')(T-t)} w(t) \end{aligned} \tag{37}$$

2.2 Case 2: $K(x)$

We assume that Support $(K) \subset [a,b]$ where $a < b$ and $a \neq 0$ or $b \neq 1$, i.e. the support of K is not $[0,1]$. We can follow the same reasoning as previously up to equation (30):

$$w(t) = S_-(0,t)\varphi(t) = S_-(0,t)(w_0) + \int_0^t S_-(s,t)\left[(-K(x)-K'(x))S_-(0,s)\varphi(s)\right]\,ds \tag{38}$$

Let us assume, by contradiction, that $-K(x)-K'(x)$ commutes with S_-. Then we get:

$$S_-(0,t)(\varphi(t)-w_0) = (-K(x)-K'(x))S_-(0,t)\int_0^t \varphi(s)\,ds \tag{39}$$

But we know that S_- has the unique continuation property, that is:

Proposition 3 *If* $S_-(0,t)(X) = 0$ *on a non-empty subset of* $[0,1]$*, then* $S_-(0,t)(X) = 0$ *on* $[0,1]$.

This result and (39) give:

$$w(t) = S_-(0,t)(\varphi(t)) = S_-(0,t)(w_0) = S_+(0,t)(w_0) \tag{40}$$

As this stands for every w_0, we have $S_- = S_+$ and finally $K = K' = 0$, which is a contradiction. Therefore, $K + K'$ does not commute with S_-. Thus, in general, we cannot find any function ψ such that:

$$\int_0^t S_-(s,t)\left[(-K(x) - K'(x))S_-(0,s)\varphi(s)\right]\, ds = S_-(0,t)\psi \tag{41}$$

2.3 Case 3: $K(t)$

We assume that $K(t,x) = K(t) = K\mathbb{1}_{[t_1,t_2]}(t)$ with $0 \leq t_1 < t_2 \leq T$. We can follow the same reasoning as for K constant, up to the Duhamel formula (30):

$$w(t) = S_-(0,t)\varphi(t) = S_-(0,t)(w_0) + \int_0^t S_-(s,t)\left[(-K - K')\mathbb{1}_{[t_1,t_2]}(s)S_-(0,s)\varphi(s)\right]\, ds \tag{42}$$

As $K + K'$ is independent of x, it commutes with S_-, and we have for φ:

$$S_-(0,t)\varphi(t) = S_-(0,t)(w_0) + \begin{cases} 0 & \text{if } t \leq t_1 \\ (-K - K')S_-(0,t)\displaystyle\int_{t_1}^t \varphi(s)\, ds & \text{if } t_1 < t < t_2 \\ (-K - K')S_-(0,t)\displaystyle\int_{t_1}^{t_2} \varphi(s)\, ds & \text{if } t \leq t_2 \end{cases} \tag{43}$$

So that the corresponding φ is given by:

$$\varphi(t) = \begin{cases} w_0 & \text{if } t \leq t_1 \\ w_0 e^{(-K-K')(t-t_1)} & \text{if } t_1 < t < t_2 \\ w_0 e^{(-K-K')(t_2-t_1)} & \text{if } t \leq t_2 \end{cases} \tag{44}$$

And thus the result follows.

3 Bürgers' equation with a viscous term

3.1 Proof of theorem 2

Without loss of generality we assume that the observations are identically zero: $u_{obs}(t,x) = 0$ for all (t,x). Let us first introduce some notations.

Let us denote by w (resp. $\widetilde{w}$) the differences between u (resp. $\widetilde{u}$) and the observations, as in (3), they satisfy the following equations:

$$
\begin{aligned}
&(F)\quad \left\{\begin{array}{rcl} \partial_t w-\nu\partial_{xx}w+w\partial_x w+Kw &=& 0\\ w|_{x=0}=w|_{x=1} &=& 0\\ w|_{t=0} &=& w_0 \end{array}\right.\\
&(B)\quad \left\{\begin{array}{rcl} \partial_t \widetilde{w}-\nu\partial_{xx}\widetilde{w}+\widetilde{w}\partial_x \widetilde{w}-K'\widetilde{w} &=& 0\\ \widetilde{w}|_{x=0}=\widetilde{w}|_{x=1} &=& 0\\ \widetilde{w}|_{t=T} &=& w(T) \end{array}\right.
\end{aligned}
\tag{45}
$$

Let us denote also by S_+ and S_- the non-linear operator associated to the forward equations with K or K':

$$S_+(t_0,t)(w(t_0))=w(t),\quad S_-(t_0,t)(\widetilde{w}(t_0))=\widetilde{w}(t),\quad \forall t\geq t_0 \tag{46}$$

We will also use the linear operators U_+ and U_- associated to the following linear equations:

$$\left\{\begin{array}{rcl} \partial_t\phi-\nu\partial_{xx}\phi+K\phi &=& 0\\ \phi|_{x=0}=\phi|_{x=1}=0,\qquad \phi|_{t=0} &=& \phi_0 \end{array}\right. \iff U_+(0,t)(\phi_0)=\phi(t) \tag{47}$$

$$\left\{\begin{array}{rcl} \partial_t\phi-\nu\partial_{xx}\phi-K'\phi &=& 0\\ \phi|_{x=0}=\phi|_{x=1}=0,\qquad \phi|_{t=0} &=& \phi_0 \end{array}\right. \iff U_-(0,t)(\phi_0)=\phi(t) \tag{48}$$

To prove theorem 2 we will prove that w is not in the image of S_-, in general. To do so we will use perturbations theory. We can easily show that S_+ is infinitely continuous with respect to the data w_0. So if we suppose that w_0 is small:

$$w_0=\varepsilon\varphi_0 \tag{49}$$

then we have that $w(t)$, solution of the forward equation (45,F) is also small and can be developed in series of ε

$$w=\varepsilon\sum_{n\geq 0}\varepsilon^n w^n \tag{50}$$

Similarly, we develop $\widetilde{w}$ in series of ε

$$\widetilde{w}=\varepsilon\sum_{n\geq 0}\varepsilon^n \widetilde{w}^n \tag{51}$$

As previously, w satisfies:

$$\left\{\begin{array}{rcl} \partial_t w-\nu\partial_{xx}w+Kw &=& -w\partial_x w\\ w|_{x=0}=w|_{x=1} &=& 0\\ w|_{t=0} &=& w_0 \end{array}\right. \tag{52}$$

so that if we develop in series of ε we get, for w^0:

$$\left\{\begin{array}{rcl} \partial_t w^0-\nu\partial_{xx}w^0+Kw^0 &=& 0\\ w^0|_{x=0}=w^0|_{x=1} &=& 0\\ w^0|_{t=0} &=& \varphi_0 \end{array}\right. \tag{53}$$

For w^1 we have:

$$\left\{\begin{array}{rcl} \partial_t w^1 - \nu\partial_{xx} w^1 + K w^1 &=& -w^0\partial_x w^0 \\ w^1|_{x=0} = w^1|_{x=1} &=& 0 \\ w^1|_{t=0} &=& 0 \end{array}\right. \tag{54}$$

Similarly we have for $\widetilde{w}^0$ and $\widetilde{w}^1$:

$$\left\{\begin{array}{rcl} \partial_t \widetilde{w}^0 - \nu\partial_{xx} \widetilde{w}^0 - K' \widetilde{w}^0 &=& 0 \\ \widetilde{w}^0|_{x=0} = \widetilde{w}^0|_{x=1} &=& 0 \\ \widetilde{w}^0|_{t=T} &=& w^0(T) \end{array}\right. \tag{55}$$

$$\left\{\begin{array}{rcl} \partial_t \widetilde{w}^1 - \nu\partial_{xx} \widetilde{w}^1 - K' \widetilde{w}^1 &=& -\widetilde{w}^0\partial_x \widetilde{w}^0 \\ \widetilde{w}^1|_{x=0} = \widetilde{w}^1|_{x=1} &=& 0 \\ \widetilde{w}^1|_{t=T} &=& w^1(T) \end{array}\right. \tag{56}$$

We can compute w^0 and w^1 thanks to U_+:

$$\begin{array}{rcl} w^0(t) &=& U_+(0,t)(\varphi_0) \\ w^1(t) &=& -\displaystyle\int_0^t U_+(s,t)[w^0(s)\partial_x w^0(s)]\,ds \end{array} \tag{57}$$

If we assume that $\widetilde{w}^0$ is well defined, with

$$\widetilde{w}^0(t) = U_-(0,t)(\psi_0) \tag{58}$$

then the condition $\widetilde{w}(T) = w(T)$ leads to

$$\begin{array}{rrcl} & U_-(0,T)(\psi_0) &=& U_+(0,T)(\varphi_0) \\ \Rightarrow & \psi_0 &=& U_-(0,T)^{-1}U_+(0,T)(\varphi_0) \\ \Rightarrow & \psi_0 &=& \mathrm{e}^{-(K+K')T}\varphi_0 \end{array} \tag{59}$$

Then we have for $\widetilde{w}^0$:

$$\begin{array}{rcl} \widetilde{w}^0(t) &=& U_-(0,t)(\psi_0) \\ &=& U_-(0,t)\mathrm{e}^{-(K+K')T}\varphi_0 \end{array} \tag{60}$$

For $\widetilde{w}^1$ the final condition $\widetilde{w}^1(T) = w^1(T)$ gives, thanks to (57):

$$\begin{array}{rcl} \widetilde{w}^1(T) &=& w^1(T) \\ &=& -\displaystyle\int_0^T U_+(s,t)[w^0(s)\partial_x w^0(s)]\,ds \end{array} \tag{61}$$

On the other hand, if we assume that $\widetilde{w}^1$ is well defined, with $\widetilde{w}^1(0) = \psi_T$, then equation (56) and the Duhamel formula give

$$\widetilde{w}^1(T) = U_-(0,T)[\psi_T] - \int_0^T U_-(s,T)[\widetilde{w}^0(s)\partial_x \widetilde{w}^0(s)]\,ds \tag{62}$$

Then, equalling (62) and (61) we should have

$$U_-(0,T)[\psi_T] - \int_0^T U_-(s,T)[\widetilde{w}^0(s)\partial_x \widetilde{w}^0(s)]\,ds \;=\; -\int_0^T U_+(s,T)[w^0(s)\partial_x w^0(s)]\,ds \tag{63}$$

Therefore

$$2U_-(0,T)[\psi_T] \;=\; \int_0^T U_-(s,T)[\partial_x(\widetilde{w}^0(s)^2)]\,ds \;-\int_0^T U_+(s,T)[\partial_x(w^0(s)^2)]\,ds \tag{64}$$

If we assume that $\psi_T = \dfrac{1}{2}\partial_x g_T$, then we obtain, up to a constant

$$U_-(0,T)[g_T] \;=\; \int_0^T U_-(s,T)[\widetilde{w}^0(s)^2]\,ds - \int_0^T U_+(s,T)[w^0(s)^2]\,ds \tag{65}$$

We now use (57), (58) and (59):

$$\begin{aligned} U_-(0,T)[g_T] \;&=\; \int_0^T U_-(s,T)[(U_-(0,s)(\mathrm{e}^{-(K+K')T}\varphi_0)]^2\,ds \\ &\quad -\int_0^T U_+(s,T)[(U_-(0,s)(\varphi_0)]^2\,ds \\ &=\; (\mathrm{e}^{-2(K+K')T}-1)\int_0^T U_+(s,T)[(U_-(0,s)(\varphi_0)]^2\,ds \end{aligned} \tag{66}$$

And if $K>0$ and $K'>0$ this last equation is in general impossible: such g_T does not, in general, exist. Indeed, let us do an explicit computation thanks to Fourier series:

$$\varphi_0 = \sum_{n\geq 1} a_n \mathrm{e}^{inx}, \quad g_T = \sum_{n\geq 1} b_n \mathrm{e}^{inx} \tag{67}$$

We recall that we have

$$\begin{aligned} U_+(s,t)\left[\sum_{n\geq 1} c_n \mathrm{e}^{inx}\right] \;&=\; \sum_{n\geq 1} c_n \mathrm{e}^{inx}\mathrm{e}^{(-K-\nu n^2)(t-s)} \\ U_-(s,t)\left[\sum_{n\geq 1} c_n \mathrm{e}^{inx}\right] \;&=\; \sum_{n\geq 1} c_n \mathrm{e}^{inx}\mathrm{e}^{(K'-\nu n^2)(t-s)} \end{aligned} \tag{68}$$

Then we can compute the right hand side of equation (66):

$$
\begin{aligned}
& (\mathrm{e}^{-2(K+K')T}-1)\int_0^T U_+(s,T)[(U_-(0,s)(\varphi_0)]^2\,ds \\
= \ & (\mathrm{e}^{-2(K+K')T}-1)\int_0^T U_+(s,T)\left[\sum_n a_n \mathrm{e}^{K's}\mathrm{e}^{inx}\mathrm{e}^{-s\nu n^2}\right]^2 ds \\
= \ & (\mathrm{e}^{-2(K+K')T}-1)\int_0^T U_+(s,T)\left[\sum_n \mathrm{e}^{2sK'}\mathrm{e}^{inx}\sum_{p+q=n} a_p a_q \mathrm{e}^{-s\nu(p^2+q^2)}\right] ds \\
= \ & (\mathrm{e}^{-2(K+K')T}-1)\int_0^T \left[\sum_n \mathrm{e}^{-K(T-s)}\mathrm{e}^{2sK'}\mathrm{e}^{-\nu(T-s)n^2}\mathrm{e}^{inx}\sum_{p+q=n} a_p a_q \mathrm{e}^{-s\nu(p^2+q^2)}\right] ds \\
= \ & (\mathrm{e}^{-2(K+K')T}-1)\int_0^T \left[\sum_n \sum_{p+q=n} a_p a_q \mathrm{e}^{-KT-\nu Tn^2+inx}\mathrm{e}^{2sK'+sK+\nu sn^2-s\nu(p^2+q^2)}\right] ds \\
= \ & (\mathrm{e}^{-2(K+K')T}-1)\left[\sum_n \sum_{p+q=n} a_p a_q \mathrm{e}^{-KT-\nu Tn^2+inx}\frac{\mathrm{e}^{2TK'+TK+2\nu pqT}-1}{2K'+K+2\nu pq}\right]
\end{aligned}
\tag{69}
$$

For the left hand side of (66) we have:

$$
U_-(0,T)[g_T] = \sum_n b_n \mathrm{e}^{inx}\mathrm{e}^{K'T-\nu n^2 T} \tag{70}
$$

So that we get, for all n:

$$
b_n = \mathrm{e}^{(-K'+K)T}(\mathrm{e}^{-2(K+K')T}-1)\left[\sum_{p+q=n} a_p a_q \frac{\mathrm{e}^{2TK'+TK+2\nu pqT}-1}{2K'+K+2\nu pq}\right] \tag{71}
$$

This defines a distribution iff b_n has polynomial growth, iff $\underline{b_n}$ has polynomial growth, where

$$
\underline{b_n} = (\mathrm{e}^{-2(K+K')T}-1)\left[\sum_{p+q=n} a_p a_q \frac{\mathrm{e}^{2TK'+TK+2\nu pqT}}{2K'+K+2\nu pq}\right] \tag{72}
$$

which is clearly not the case for every sequence (a_n) with polynomial growth, unless $K = K' = 0$.

3.2 Particular case: $K = K' = 0$

We consider the particular case where $K = K' = 0$, i.e. there is no nudging term either in the forward or backward equations. In this case, proposition 1 holds true.

Of course, the backward equation itself is ill-posed, as even if there is existence and unicity of the solution (e.g. if the final condition $\tilde{u}(T)$ comes from a resolution of the forward equation over the same time period), it does not depend in a continuous way of the data.

The proof is straightforward by using the following Cole-Hopf transformations [Col51, Hop50]:

$$\begin{aligned} u &= -2\nu \frac{\partial_x v}{v}, \qquad v(t,x) = v(t,0) e^{-\frac{1}{2\nu} \int_0^x u(t,s)\, ds} \\ \widetilde{u} &= -2\nu \frac{\partial_x \widetilde{v}}{\widetilde{v}}, \qquad \widetilde{v}(t,x) = \widetilde{v}(t,0) e^{-\frac{1}{2\nu} \int_0^x \widetilde{u}(t,s)\, ds} \end{aligned} \tag{73}$$

in the forward and backward equations respectively. These transformations allow us to consider the same forward and backward problem, but on the heat equation. The Fourier transform gives the existence and uniqueness of a solution to the forward and backward heat equation, and the equality between the forward v and backward $\widetilde{v}$ solutions. Equations (73) extend the result to the viscous Bürgers' equation.

4 Non viscous transport equations

4.1 Linear case: proof of theorem 3

In this section we prove theorem 3.

The first two points of the theorem are easily proven as in theorem 1 with a vanishing viscosity.

Thus we only prove the third point. To do so, we recall that the curves $(s, \psi(s,x))$ are the characteristics of the direct equation (8-F) with $K = 0$, such that $(s, \psi(s,x))|_{s=0} = (0, x)$ (see [CH62, Eva98] for characteristics theory).

For the forward equation (8-F), this change of variable gives

$$\partial_s w(s, \psi(s,x)) = -K(\psi(s,x)) w(s, \psi(s,x)) \tag{74}$$

So that

$$w(s, \psi(s,x)) = w(0,x) \exp\left(-\int_0^s K(\psi(\sigma,x))\, d\sigma\right) \tag{75}$$

And in particular for $w(T)$ we have

$$w(T, \psi(T,x)) = w(0,x) \exp\left(-\int_0^T K(\psi(\sigma,x))\, d\sigma\right) \tag{76}$$

For $\widetilde{w}$ we have similarly

$$\partial_s \widetilde{w}(s, \psi(s,x)) = K'(\psi(s,x)) \widetilde{w}(s, \psi(s,x)) \tag{77}$$

So that we have:

$$\begin{aligned} \widetilde{w}(s, \psi(s,x)) &= \widetilde{w}(T, \psi(T,x)) \exp\left(-\int_s^T K'(\psi(\sigma,x))\, d\sigma\right) \\ &= w(T, \psi(T,x)) \exp\left(-\int_s^T K'(\psi(\sigma,x))\, d\sigma\right) \end{aligned} \tag{78}$$

Using (76) and (75) we get

$$\begin{aligned}
& \widetilde{w}(s,\psi(s,x)) \\
= \quad & w(0,x)\,\exp\left(-\int_0^T K(\psi(\sigma,x))\,d\sigma\right)\exp\left(-\int_s^T K'(\psi(\sigma,x))\,d\sigma\right) \\
= \quad & w(s,\psi(s,x))\,\exp\left(\int_0^s K(\psi(\sigma,x))\,d\sigma\right)\exp\left(-\int_0^T K(\psi(\sigma,x))\,d\sigma\right)\exp\left(-\int_s^T K'(\psi(\sigma,x))\,d\sigma\right) \\
= \quad & w(s,\psi(s,x))\,\exp\left(-\int_s^T K(\psi(\sigma,x))+K'(\psi(\sigma,x))\,d\sigma\right)
\end{aligned} \tag{79}$$

4.2 Non linear case: proof of theorem 4 and proposition 2

From equation (15), we deduce that the forward error w satisfies the following equation:

$$\partial_t w + w\partial_x w + u_{obs}\partial_x w + w\partial_x u_{obs} = -Kw \tag{80}$$

By multiplying by w and integrating over Ω, we obtain

$$\frac{1}{2}\partial_t\left(\int_\Omega w^2\right) + \int_\Omega w^2\partial_x w + \int_\Omega (u_{obs}w\partial_x w + w^2\partial_x u_{obs}) = -\int_\Omega K w^2 \tag{81}$$

Some integrations by part give the following:

$$\partial_t(\|w(t)\|^2) = \int_\Omega (-2K - \partial_x u_{obs})w^2 \tag{82}$$

We set $M = \|\partial_x u_{obs}\|_\infty$, and as K does not depend on x,

$$\partial_t(\|w(t)\|^2) \leq (-2K+M)\|w(t)\|^2 \tag{83}$$

We have a similar result for the backward error:

$$\partial_t(\|\widetilde{w}(t)\|^2) \leq (-2K'+M)\|\widetilde{w}(t)\|^2 \tag{84}$$

We first consider the first point of theorem 4, i.e. $K(t,x) = K$. Grönwall's lemma between times t and T gives

$$\|w(T)\|^2 \leq e^{(-2K+M)(T-t)}\|w(t)\|^2 \tag{85}$$

$$\|\widetilde{w}(t)\|^2 \leq e^{(-2K'+M)(T-t)}\|\widetilde{w}(T)\|^2 \tag{86}$$

from which equation (18) is easily deduced.

In the second case, i.e. $K(t,x) = K\mathbb{1}_{[t_1,t_2]}(t)$ and by successively applying Grönwall's lemma between times 0 and t_1, t_1 and t_2, and t_2 and T, one obtains equation (19).

Finally, in the case $K(t,x)=K(x)$, by considering a similar approach as in section 4.1, i.e. using the characteristics of the direct equation (15-F) (resp. B), it is straightforward to prove that

$$w(s,\psi(s,x)) = w(0,x)e^{-\int_0^s K(\psi(\sigma,x))d\sigma} e^{-\int_0^s \partial_x u_{obs}(\sigma,\psi(\sigma,x))d\sigma} \tag{87}$$

and then,

$$w(T,\psi(T,x)) = w(0,x)e^{-\int_0^T K(\psi(\sigma,x))d\sigma} e^{-\int_0^T \partial_x u_{obs}(\sigma,\psi(\sigma,x))d\sigma} \tag{88}$$

from which equation (20) is easily deduced.

5 Conclusion

Several conclusions can be drawn from all these results. First of all, in many situations, the coupled forward-backward problem is well posed, and the nudging terms allow the solution to be corrected (towards the observation trajectory) everywhere and with an exponential convergence. From a numerical point of view, these results have been observed in several geophysical situations, and many numerical experiments have confirmed the global convergence of the BFN algorithm [AB08].

The second remark is that the worst situation, i.e. for which there is no solution to the BFN problem, is the viscous Bürgers' equation. But in real geophysical applications, there is most of the time no theoretical viscosity in the equation, and one should consider the inviscid equation instead, for which some convergence results are given. From the numerical point of view, these phenomenon are easily confirmed, as well as the exponential decrease of the error w. But we also noticed that if the observations are not too sparse, the algorithm works well even with a quite large viscosity.

Finally, these results extend the theory of linear observers in automatics [Lue66]: instead of considering an infinite time interval (only one forward equation but for $T \to +\infty$), one can consider an infinite number of BFN iterations on a finite time interval. This is of great interest in almost all real applications, for which it is not possible to consider a very large time period.

Acknowledgement

The authors are thankful to Prof. G. Lebeau (University of Nice Sophia-Antipolis) for his fruitful ideas and comments. This work has been partially supported by ANR JCJC07 and INSU-CNRS LEFE projects.

References

[AB05] D. Auroux and J. Blum. Back and forth nudging algorithm for data assimilation problems. *C. R. Acad. Sci. Paris Sér. I*, 340:873–878, 2005.

[AB08] D. Auroux and J. Blum. A nudging-based data assimilation method for oceanographic problems: the back and forth nudging (bfn) algorithm. *Nonlin. Proc. Geophys.*, 15:305–319, 2008.

[CH62] R. Courant and D. Hilbert. *Methods of Mathematical Physics, Volume II.* Wiley-Interscience, 1962.

[Col51] J. D. Cole. On a quasilinear parabolic equation occurring in aerodynamics. *Quart. Appl. Math.*, 9(3):225–236, 1951.

[Eva98] L. C. Evans. *Partial Differential Equations.* American Mathematical Society, Providence, 1998.

[EvL99] G. Evensen and P. J. van Leeuwen. An ensemble kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.*, 128:1852–1867, 1999.

[HA76] J. Hoke and R. A. Anthes. The initialization of numerical models by a dynamic initialization technique. *Month. Weather Rev.*, 104:1551–1556, 1976.

[Hop50] E. Hopf. The partial differential equation $u_t + uu_x = \mu u_{xx}$. *Comm. Pure Appl. Math.*, 3:201–230, 1950.

[Kal60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans. ASME - J. Basic Engin.*, 82:35–45, 1960.

[Kal03] E. Kalnay. *Atmospheric modeling, data assimilation and predictability.* Cambridge University Press, 2003.

[LDT86] F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological. observations: theoretical aspects. *Tellus*, 38A:97–110, 1986.

[Lue66] D. Luenberger. Observers for multivariable systems. *IEEE Trans. Autom. Contr.*, 11:190–197, 1966.

[Rus78] D. L. Russell. Controllability and stabilizability theory for linear partial differential equations: recent progress and open questions. *SIAM Rev.*, 20(4):639–739, 1978.