Jacques Blum
University of Nice Sophia Antipolis
jblum@unice.fr

Didier Auroux
University of Toulouse
auroux@math.univ-toulouse.fr

Back to the future : the *Back and Forth Nudging*

Scaling Up and Modeling for Transport and Flow in Porous Media
Dubrovnik, Croatia
October 13-16 2008

Motivations

Motivation : Identify the initial condition in a geophysical system

Fundamental for a chaotic system (Lorenz, atmosphere, ocean, . . .)

Difficulty : These systems are generally irreversible.

Comparison with 4D-VAR : Optimal control method minimizing the quadratic difference between model and observations.

Forward nudging

Let us consider a model governed by a system of ODE :

$$\frac{dX}{dt} = F(X), \quad 0 < t < T,$$

with an initial condition $X(0) = x_0$.

$X_{obs}(t)$: observations of the system

C : observation operator.

$$\begin{cases} \dfrac{dX}{dt} = F(X) + K(X_{obs} - C(X)), \quad 0 < t < T, \\ X(0) = x_0, \end{cases}$$

where K is the nudging (or gain) matrix.

In the linear case (where F is a matrix), the forward nudging is called Luenberger or asymptotic observer.

Direct Nudging

– Meteorology : Hoke-Anthes (1976)

– Oceanography (QG model) : Verron-Holland (1989)

– Atmosphere (meso-scale) : Stauffer-Seaman (1990)

– Optimal determination of the nudging coeffcients :
  Zou-Navon-Le Dimet (1992), Stauffer-Bao (1993),
  Vidard-Le Dimet-Piacentini (2003)

Direct Nudging : linear case

Luenberger observer, or asymptotic observer

(Luenberger, 1966)

$$\begin{cases} \dfrac{dX}{dt} = FX + K(X_{obs} - CX), \\ \dfrac{d\hat{X}}{dt} = F\hat{X}, \quad X_{obs} = C\hat{X}. \end{cases}$$

$$\frac{d}{dt}(X - \hat{X}) = (F - KC)(X - \hat{X})$$

If $F - KC$ is a Hurwitz matrix, i.e. its spectrum is strictly included in the half-plane $\{\lambda \in \mathbb{C}; Re(\lambda) < 0\}$, then $X \to \hat{X}$ when $t \to +\infty$.

Backward nudging

Backward model : (Auroux, 2003)

$$\begin{cases} \dfrac{d\tilde{X}}{dt} = F(\tilde{X}), \quad T > t > 0, \\ \tilde{X}(T) = \tilde{x}_T. \end{cases}$$

If we apply nudging to this backward model :

$$\begin{cases} \dfrac{d\tilde{X}}{dt} = F(\tilde{X}) - K'(X_{obs} - C(\tilde{X})), \quad T > t > 0, \\ \tilde{X}(T) = \tilde{x}_T. \end{cases}$$

$t' = T - t$:

$$\begin{cases} \dfrac{d\tilde{X}}{dt'} = -F(\tilde{X}) + K'(X_{obs} - C(\tilde{X})), \quad 0 < t' < T, \\ \tilde{X}(0) = \tilde{x}_T. \end{cases}$$

In the linear case, $-F - K'C$ must be a Hurwitz matrix.

BFN : Back and Forth Nudging algorithm

Iterative algorithm (forward and backward resolutions) :

$$\tilde{X}_0(0) = \tilde{x}_0 \text{ (first guess)}$$

$$\begin{cases} \dfrac{dX_k}{dt} = F(X_k) + K(X_{obs} - C(X_k)) \\ X_k(0) = \tilde{X}_{k-1}(0) \end{cases}$$

$$\begin{cases} \dfrac{d\tilde{X}_k}{dt} = F(\tilde{X}_k) - K'(X_{obs} - C(\tilde{X}_k)) \\ \tilde{X}_k(T) = X_k(T) \end{cases}$$

Cas simplifié ($C = Id$, $K = K'$)

Convergence in a linear case , with full observations :
D. Auroux, J. Blum, Back and forth nudging algorithm for data assimilation problems, *C. R. Acad. Sci. Ser. I*, **340**, pp. 873–878, 2005.

$$\lim_{k\to+\infty} X_k(0) = X_\infty(0)$$

$$= \quad \left(I - e^{-2KT}\right)^{-1} \int_0^T \left(e^{-(K+F)s} + e^{-2KT} e^{(K-F)s}\right) K X_{obs}(s)ds.$$

$$\lim_{k\to+\infty} X_k(t) = X_\infty(t) = e^{-(K-F)t} \int_0^t e^{(K-F)s} K X_{obs}(s)ds + e^{-(K-F)t} X_\infty(0).$$

If $X_{obs}(t) = e^{Ft}x_0$, then, if K and F commute,

$$X_\infty(t) = X_{obs}(t), \quad \forall t \in [0;T].$$

Choice of the direct nudging matrix K

Implicit discretization of the direct model equation with nudging :

$$\frac{X^{n+1} - X^n}{\Delta t} = FX^{n+1} + K(X_{obs} - CX^{n+1}).$$

Variational interpretation : direct nudging is a compromise between the minimization of the energy of the system and the quadratic distance to the observations :

$$\min_X \left[\frac{1}{2} \langle X - X^n, X - X^n \rangle - \frac{\Delta t}{2} \langle FX, X \rangle + \frac{\Delta t}{2} \langle R^{-1}(X_{obs} - CX), X_{obs} - CX \rangle \right],$$

by choosing

$$K = C^T R^{-1}$$

where R is the covariance matrix of the errors of observation.

Auroux-Blum, Nonlinear Processes in Geophysics (2008)

Choice of the backward nudging matrix K'

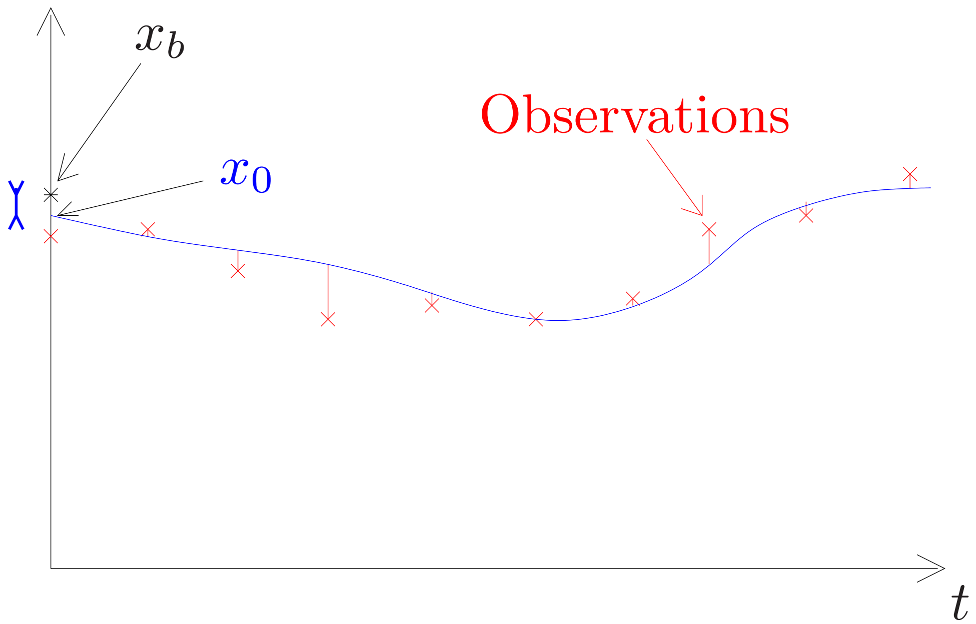
The feedback term has a double role :

- stabilization of the backward resolution of the model (irreversible system)
- feedback to the observations

If the system is observable, i.e. $rank[C, CF, \dots, CF^{N-1}] = N$, then there exists a matrix K' such that $-F - K'C$ is a Hurwitz matrix (pole assignment method).

In practice, $K' = k'C^T$ and k' can be chosen as being the smallest value making the backward numerical resolution stable.

4D-VAR



$$\begin{cases} \dfrac{dx}{dt} = F(x), \\ x(0) = x_0, \end{cases}$$

$x_{obs}(t)$: observations of the system, C : observation operator,
B and R : covariance matrices of background and observation errors respectively.

$$\begin{aligned} J(x_0) &= \frac{1}{2}(x_0 - x_b)^T B^{-1} (x_0 - x_b) \\ &+ \frac{1}{2}\int_0^T [x_{obs}(t) - C(x(t))]^T R^{-1} [x_{obs}(t) - C(x(t))] \; dt \end{aligned}$$

Optimality System

Optimization under constraints :

$$\mathcal{L}(x_0, x, p) = J(x_0) + \int_0^T \left\langle p, \frac{dx}{dt} - F(x) \right\rangle dt$$

Direct model : $\begin{cases} \dfrac{dx}{dt} = F(x) \\ x(0) = x_0 \end{cases}$

Adjoint model : $\begin{cases} -\dfrac{dp}{dt} = \left[\dfrac{\partial F}{\partial x}\right]^T p + C^T R^{-1}\left[x_{obs}(t) - C(x(t))\right] \\ p(T) = 0 \end{cases}$

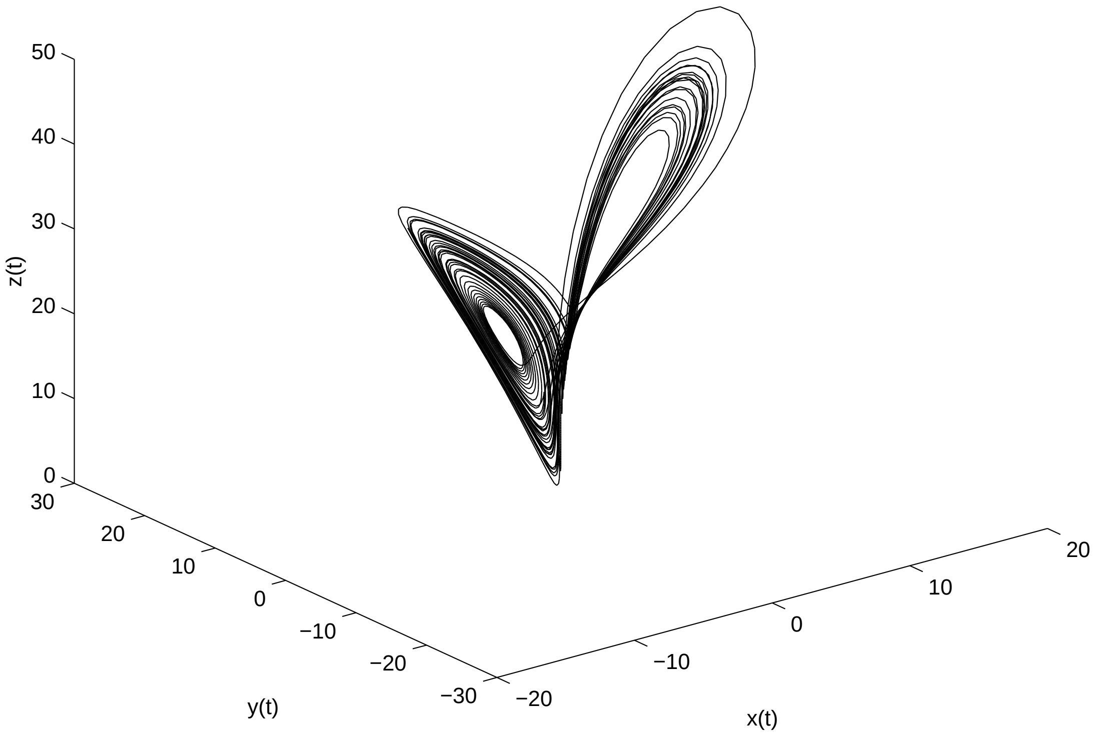
Gradient of the cost-function : $\dfrac{\partial J}{\partial x_0} = B^{-1}(x_0 - x_b) - p(0) = 0$

Le Dimet - Talagrand (86)

Numerical Results
Lorenz Equation

Lorenz' equations

$$\begin{cases} \dfrac{dx}{dt} = 10\ (y - x), \\[2ex] \dfrac{dy}{dt} = 28\ x - y - xz, \\[2ex] \dfrac{dz}{dt} = -\dfrac{8}{3}\ z + xy. \end{cases}$$



– Assimilation period : $[0;3]$, forecast : $[3;6]$.

– Time step : 0.001.

– 31 observations (every 100 time steps).

Convergence

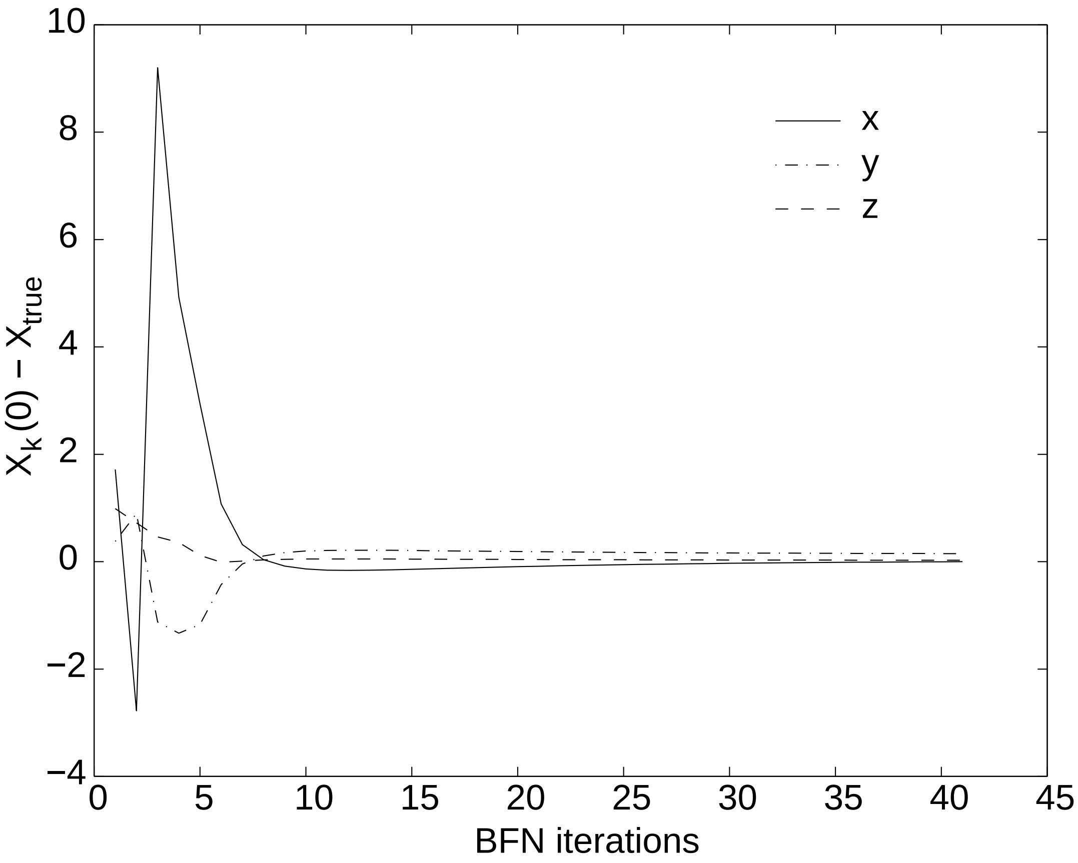


FIG. 1 – Difference between the kth iterate $X_k(0)$ and the exact initial condition x_{true} for the 3 variables versus the number of BFN iterations.

Convergence

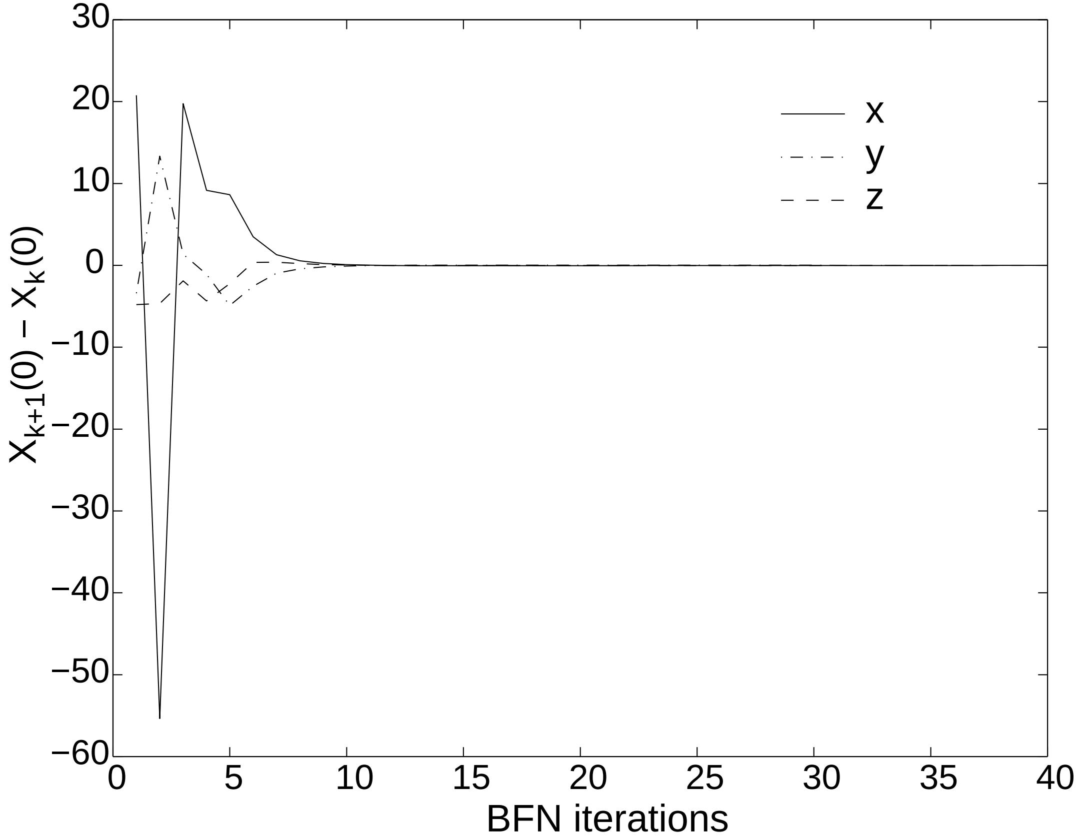


FIG. 2 – Difference between two consecutive BFN iterates for the 3 variables versus the number of BFN iterations.

Convergence

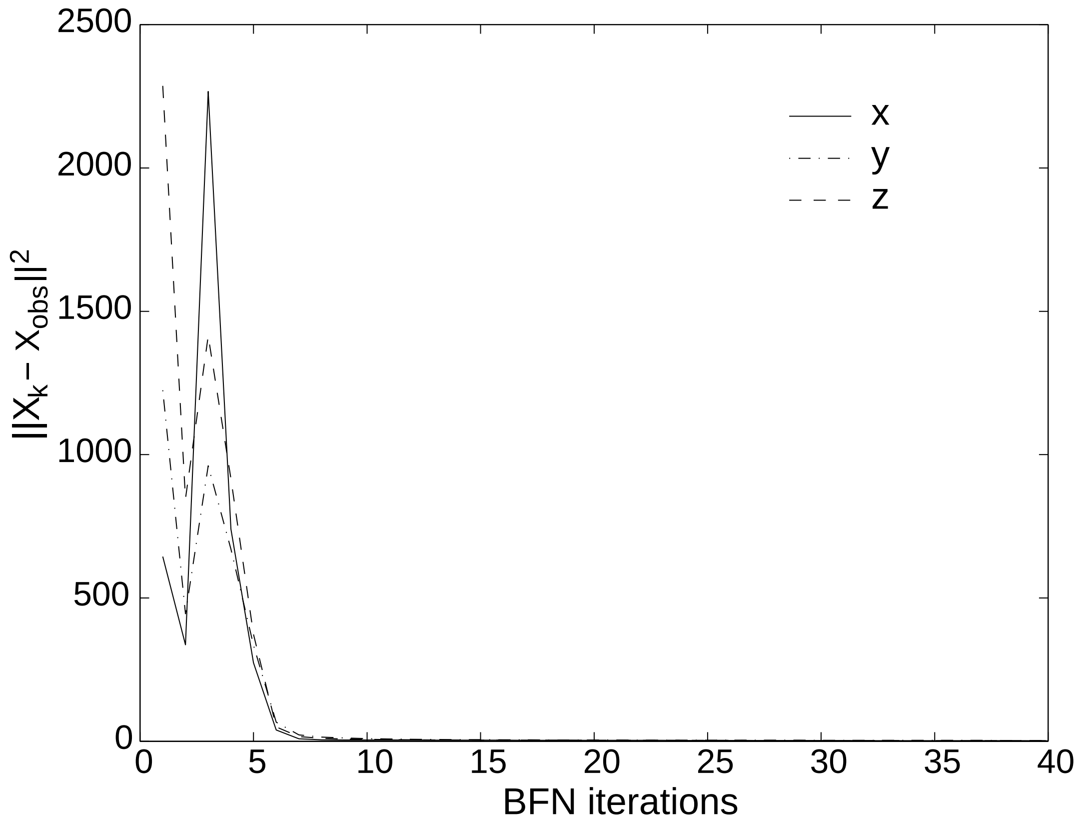


FIG. 3 – RMS difference between the observations and the BFN identified trajectory versus the BFN iterations.

Comparison with 4D-VAR

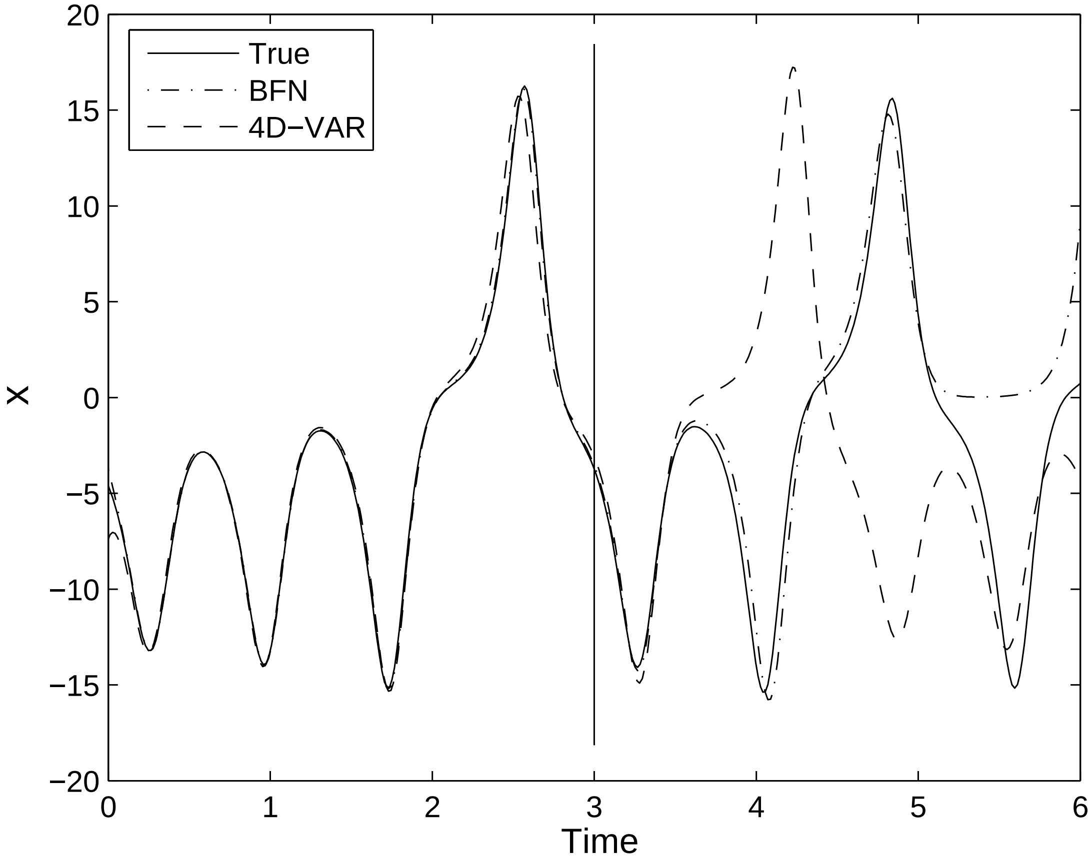


FIG. 4 – Evolution in time of the reference trajectory (plain line), and of the trajectories identified by the 4D-VAR (dashed line) and the BFN (dash-dotted line) algorithms, in the case of perfect observations and for the first Lorenz variable x.

Comparison with 4D-VAR

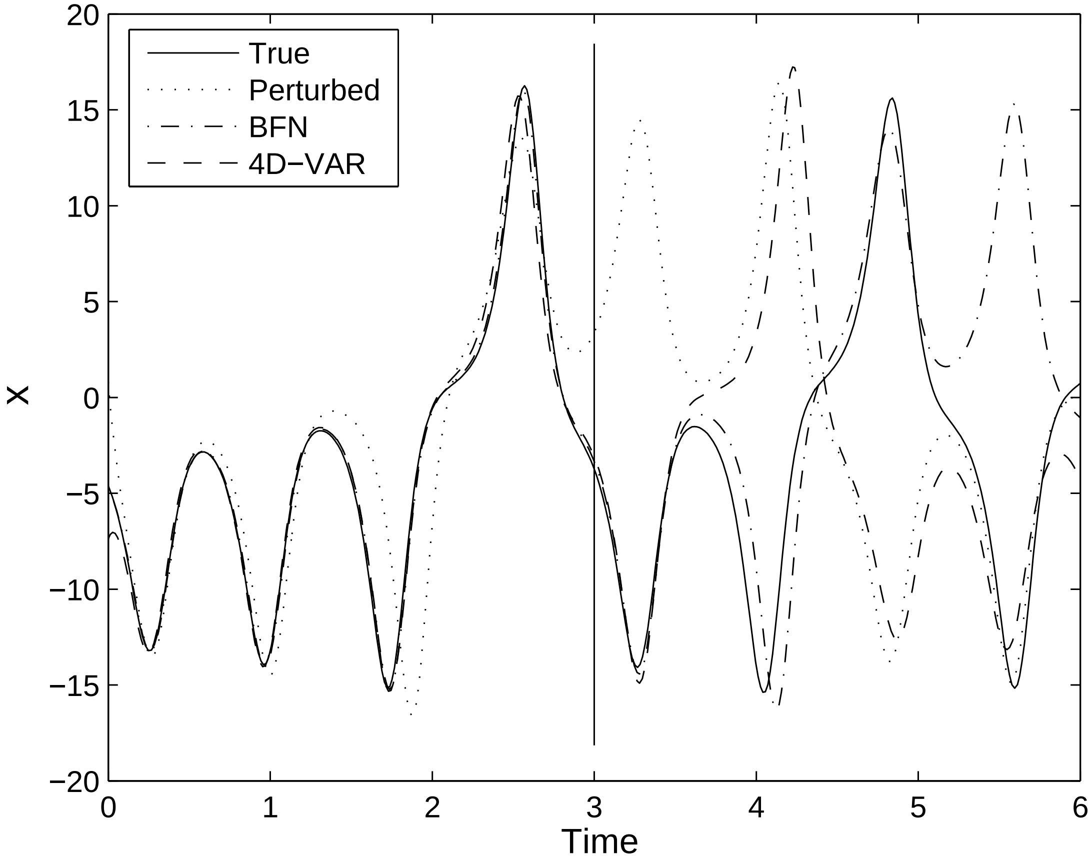


FIG. 5 – Evolution in time of the reference trajectories (plain line), and of the trajectories identified by the 4D-VAR (dashed line) and the BFN (dash-dotted line) algorithms, in the case of noised observations (with a 10% gaussian blank noise) and for the first Lorenz variable x.

Numerical Results
Burgers Equation

1D viscous Burgers' equation

$$\frac{\partial X}{\partial t} + \frac{1}{2}\frac{\partial X^2}{\partial s} - \nu \frac{\partial^2 X}{\partial s^2} = 0,$$

where X is the state variable, s represents the distance in meters around the 45^oN constant-latitude circle and t is the time.

The period of the domain is roughly $28.3 \times 10^6 m$. The diffusion coefficient ν is set to $10^5\ m^2.s^{-1}$. The time step is one hour, and the assimilation period is roughly one month (700 time steps).
Data : every 10 time steps (10 hours), every 5 gridpoints, 5% RMS blank gaussian error.

Convergence

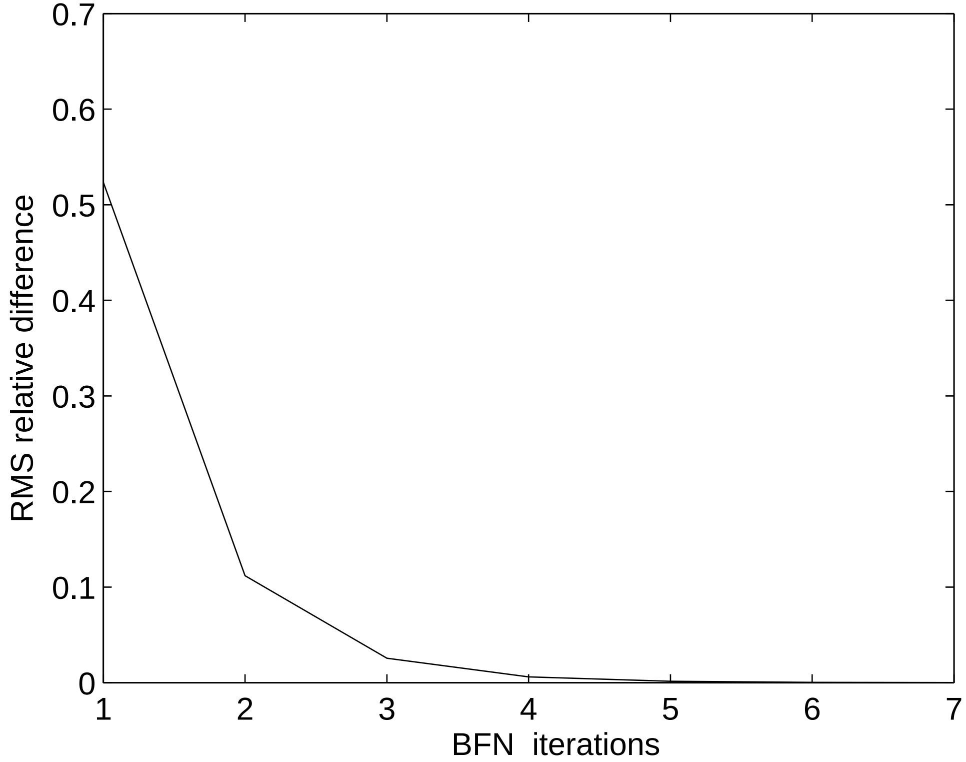


FIG. 6 – RMS relative difference between two consecutive iterates of the BFN algorithm versus the number of iterations.

Convergence

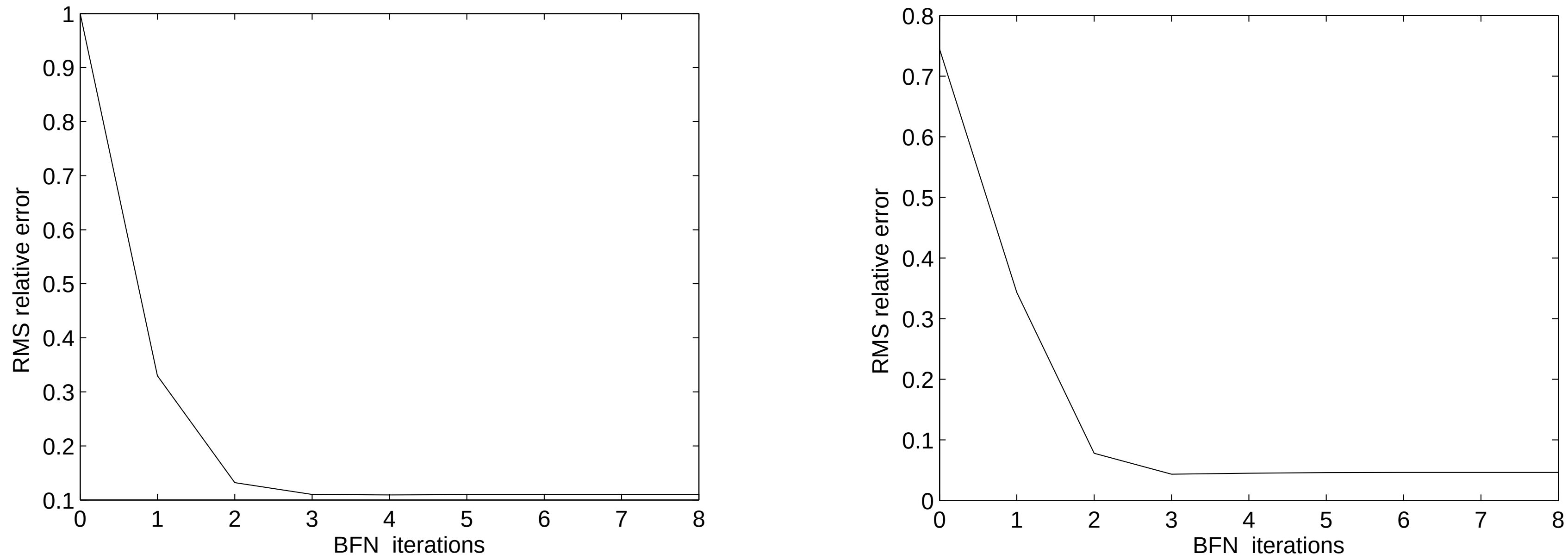


FIG. 7 – RMS relative difference between the BFN iterates and the exact solution versus the number of iterations, at time $t = 0$ (a) and at time $t = T$ (b).

Comparison with 4D-VAR

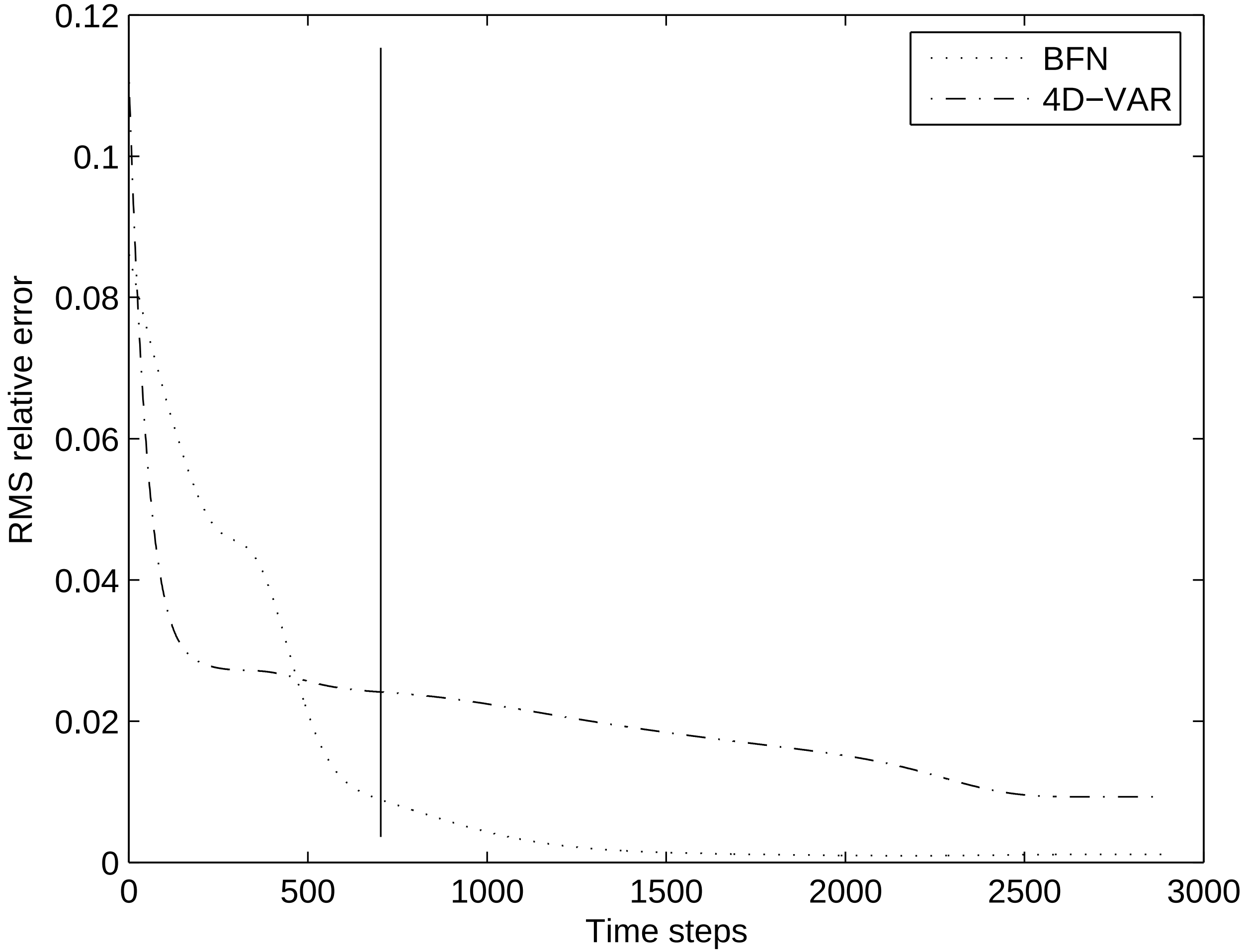


Fig. 8 – Evolution in time of the RMS difference between the reference trajectory and the identified trajectories for the BFN (dotted line) and the 4D-VAR (dash-dotted line) algorithms, in the case of perfect observations.

BFN preprocessing

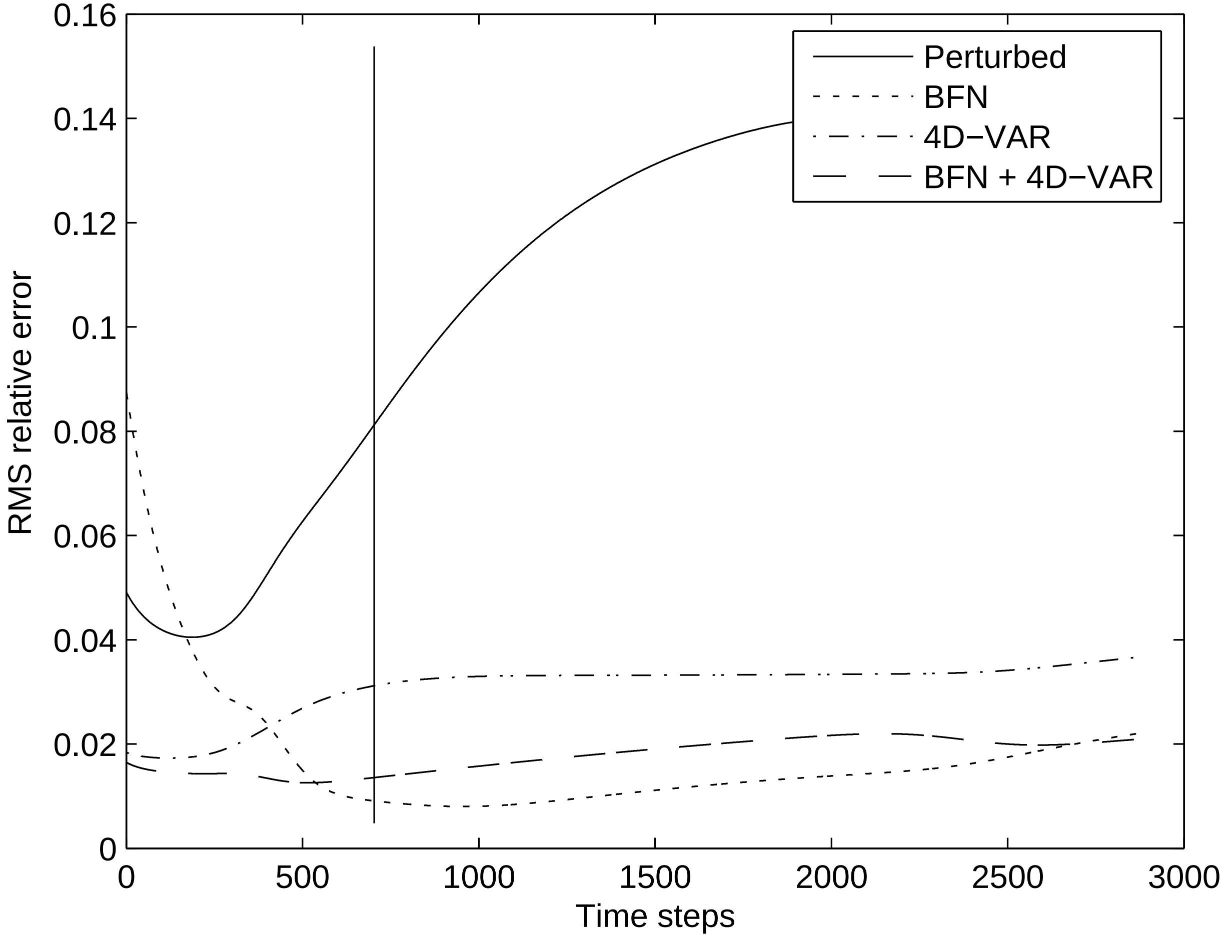


FIG. 9 – Evolution in time of the RMS difference between the reference trajectory and the identified trajectories for the BFN (dotted line), the 4D-VAR (dash-dotted line) and the BFN-preprocessed 4D-VAR (dashed line) algorithms, in the case of noised observations (with a 5% RMS error).

Conclusions

- Easy implementation (no linearization, no adjoint state, no minimization process)
- Very efficient in the first iterations
- Converges more rapidly than 4D-VAR
- Lower computational and memory costs than 4D-VAR
- Could be an excellent preconditioner for 4D-VAR

Perspective :

Test the algorithm on a primitive equation model, with realistic observations.

Happy Birthday
Alain