Didier Auroux

Université de Nice Sophia Antipolis

France

auroux@unice.fr

Diffusive Back and Forth Nudging algorithm for geophysical data assimilation

International Conference on Data Assimilation,
CAOS, IISc, Bangalore.

July 13-15, 2011

Talk overview

1. Back and forth nudging algorithm

2. Diffusive BFN algorithm

Motivations for data assimilation

Environmental and geophysical studies : forecast the natural evolution
$\rightsquigarrow$ retrieve at best the current state (or initial condition) of the environment.

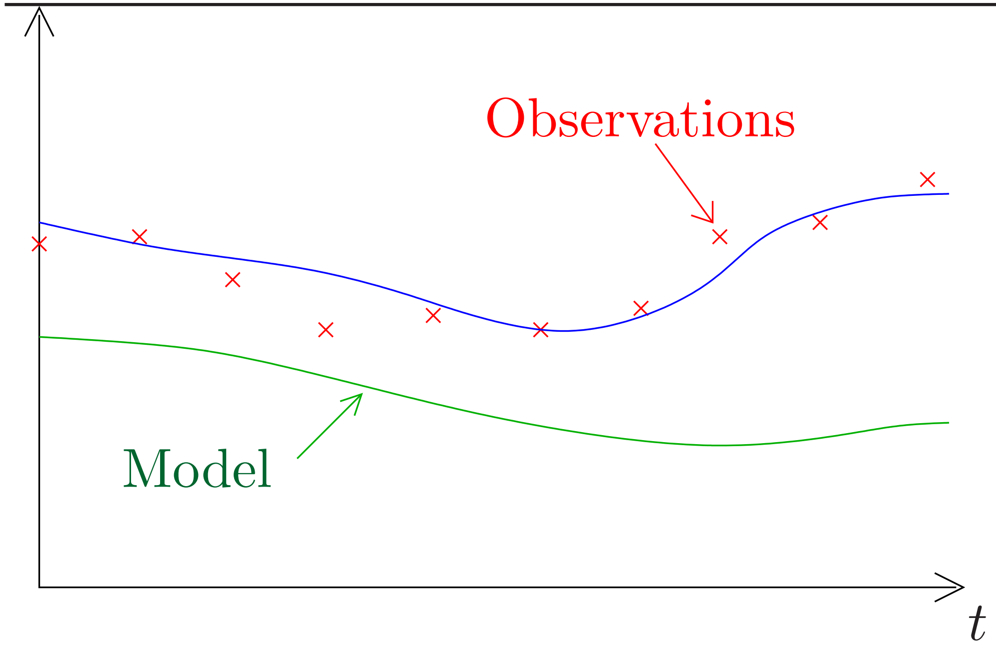
Geophysical fluids (atmosphere, oceans, ...) : turbulent systems $\Longrightarrow$ high sensitivity to the initial condition $\Longrightarrow$ need for a precise identification (much more than observations)

Environmental problems (ground pollution, air pollution, hurricanes, ...) : problems of huge dimension, generally poorly modelized or observed

Data assimilation consists in combining in an optimal way the observations of a system and the knowledge of the physical laws which govern it.

Main goal : identify the initial condition, or estimate some unknown parameters, and obtain reliable forecasts of the system evolution.

Data assimilation



combination

model + observations

$\Downarrow$

retrieve the best possible

state of a geophysical system

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

Issue : These systems are generally irreversible

Goal : Combine models and data

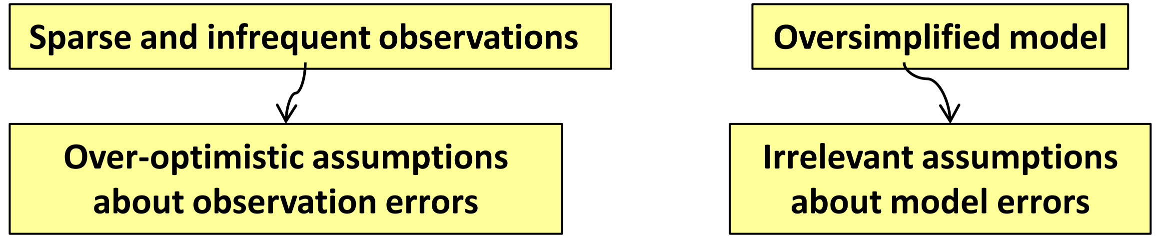
Typical inverse problem : retrieve the system state from sparse and noisy observations

Sparse and infrequent observations

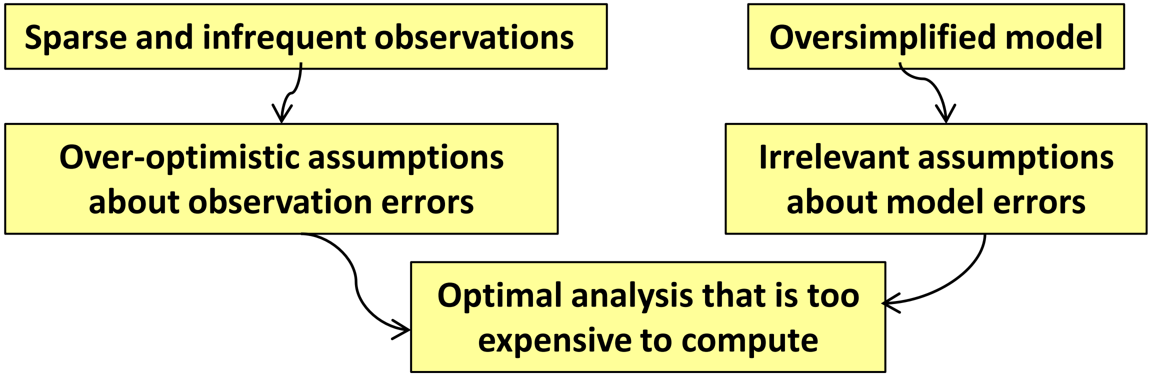
Oversimplified model

Data assimilation

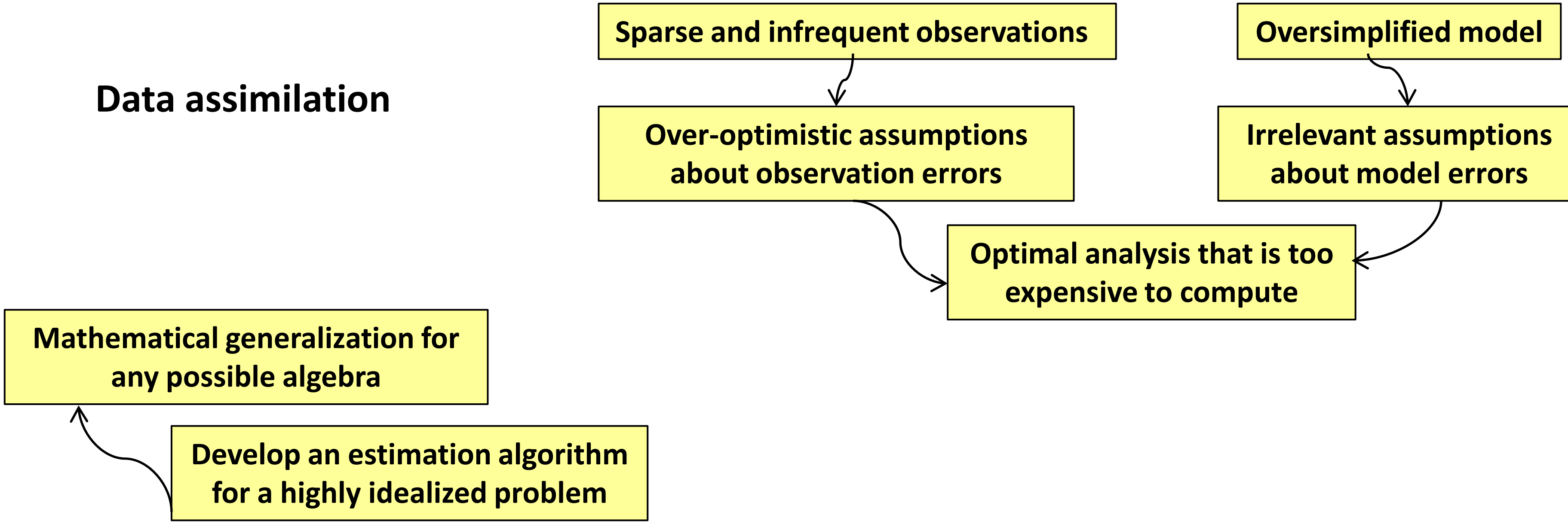
Data assimilation



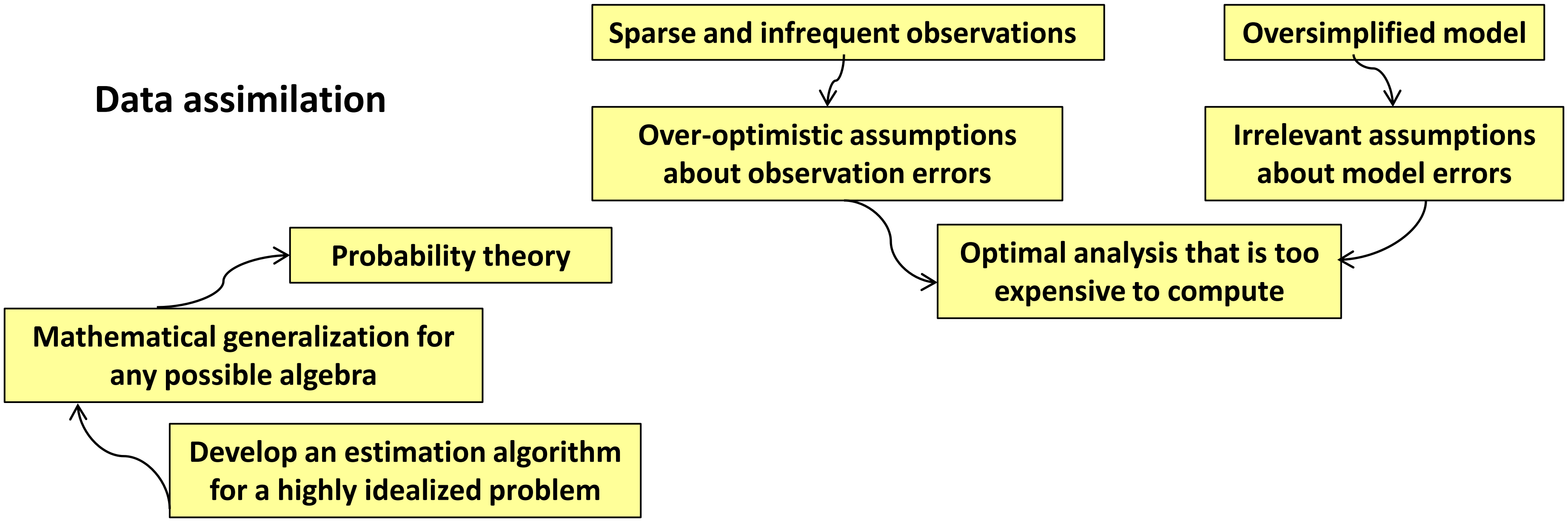
Data assimilation



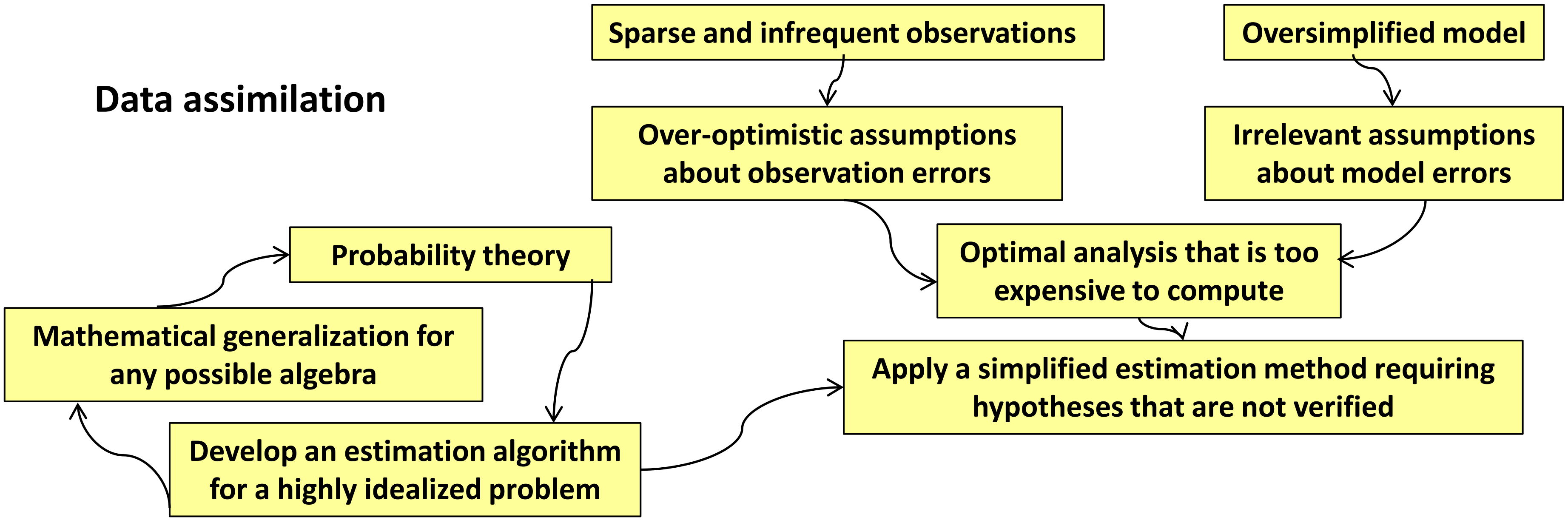
Data assimilation



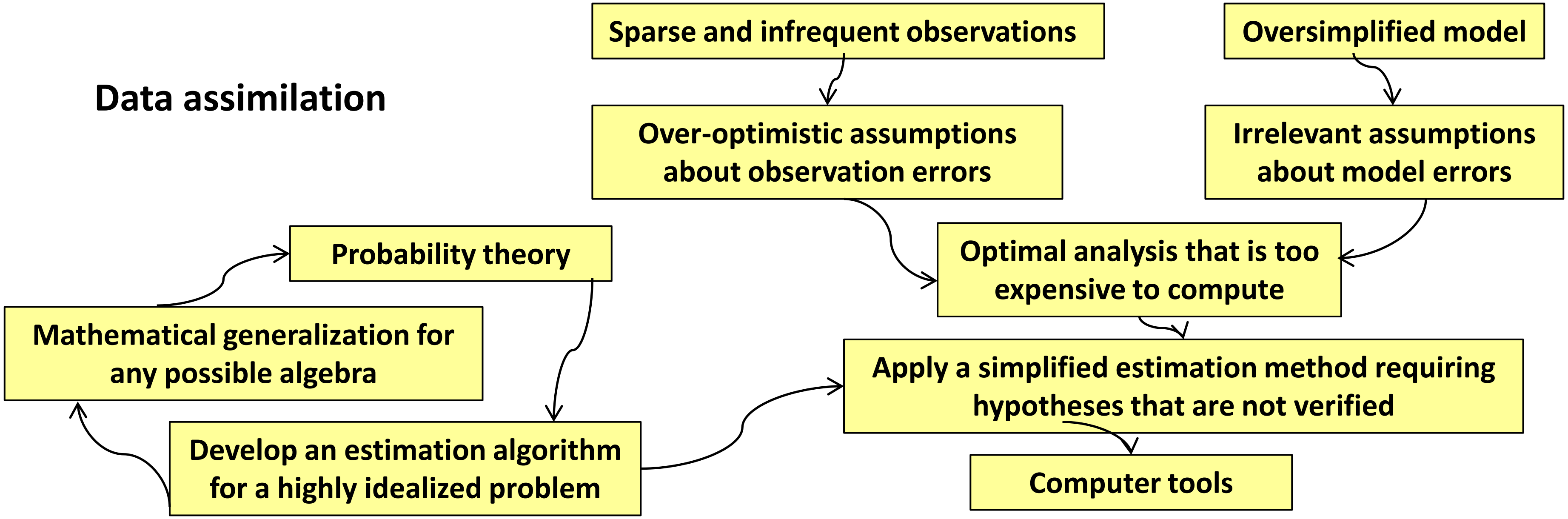
Data assimilation



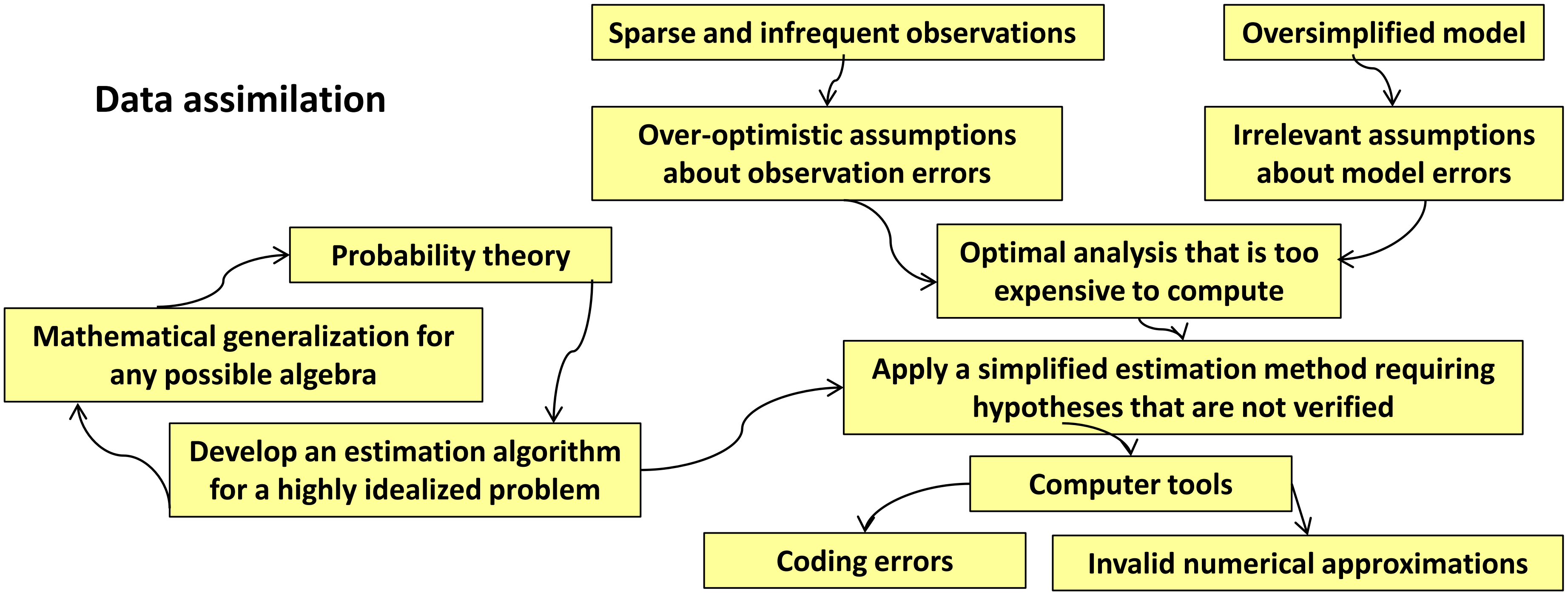
Data assimilation

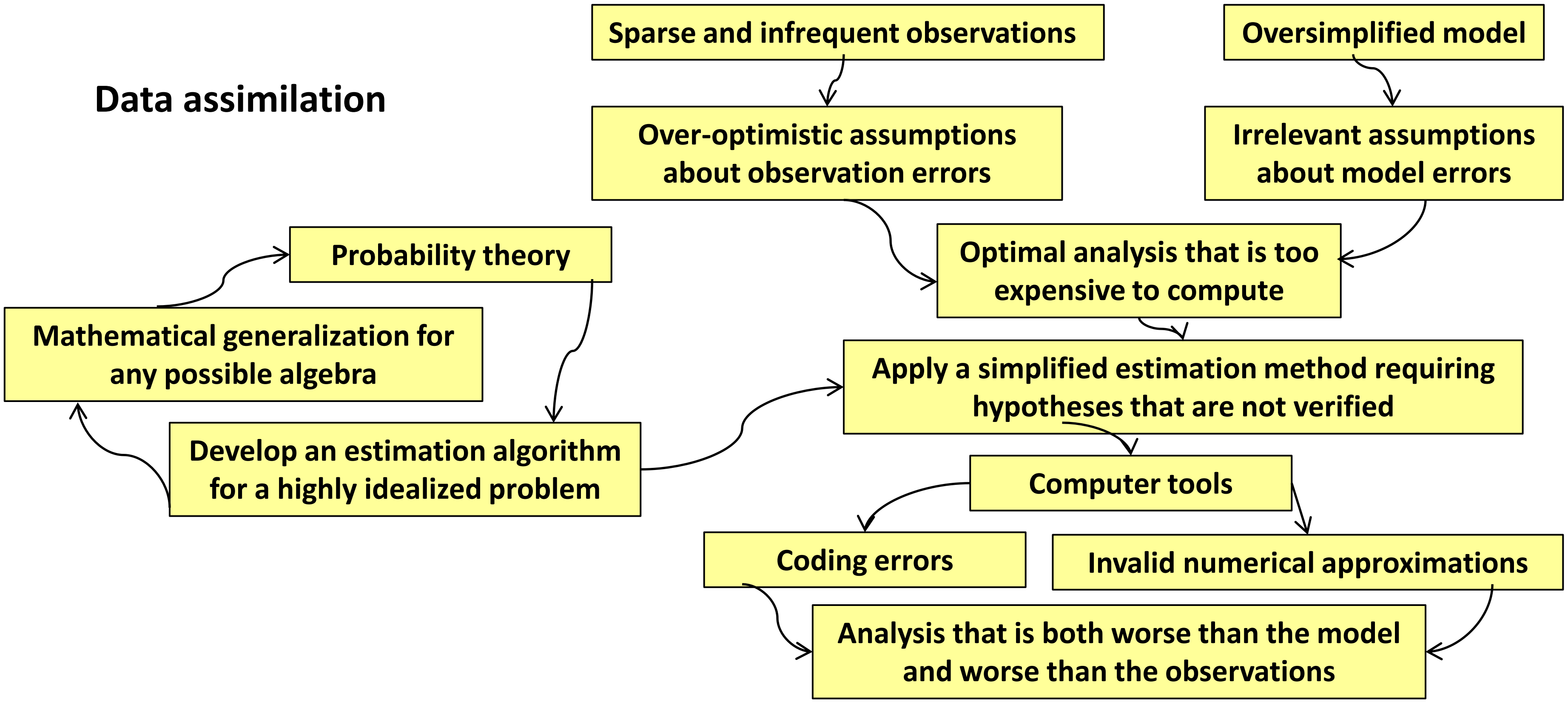


Data assimilation

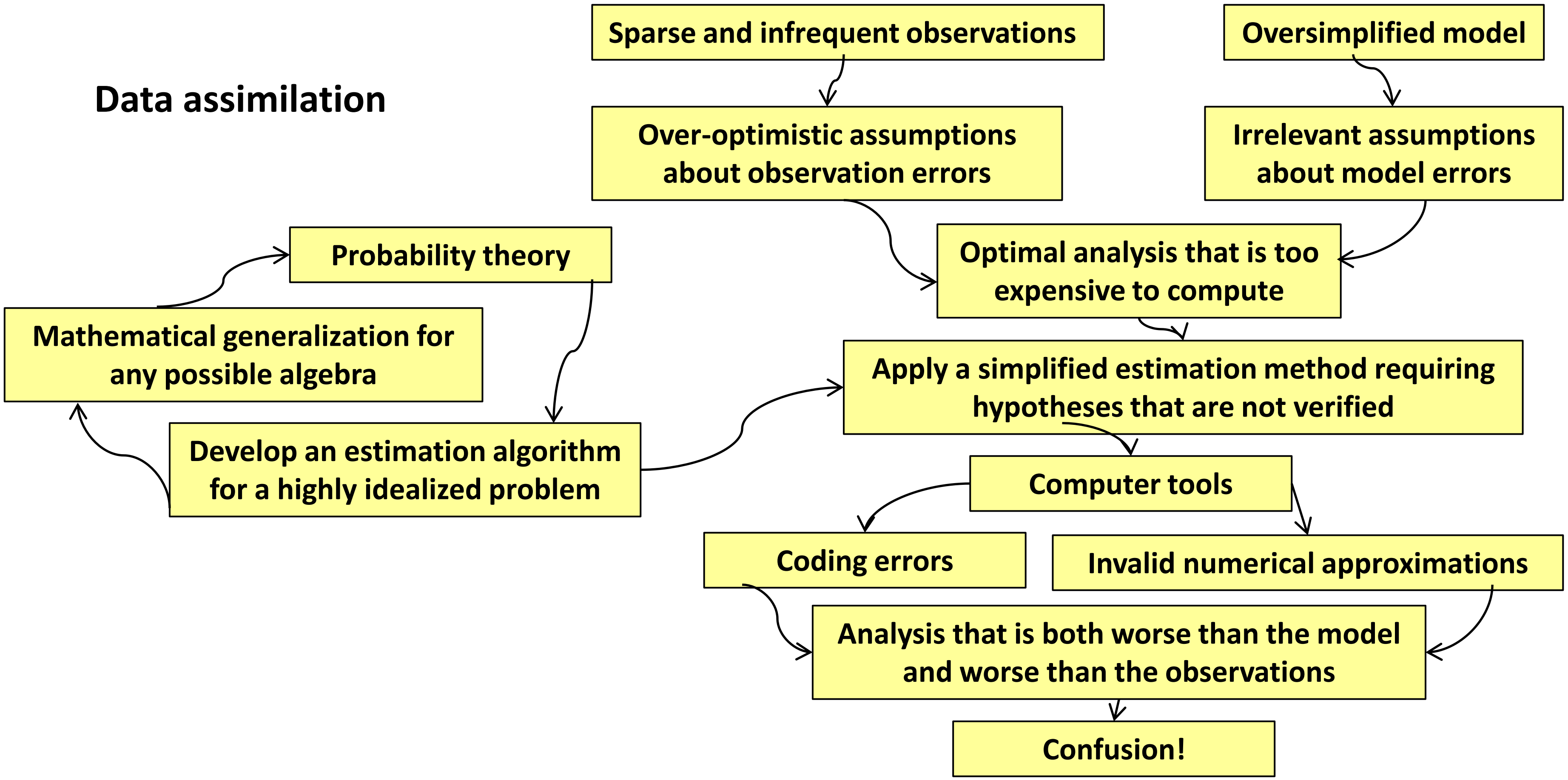


Data assimilation

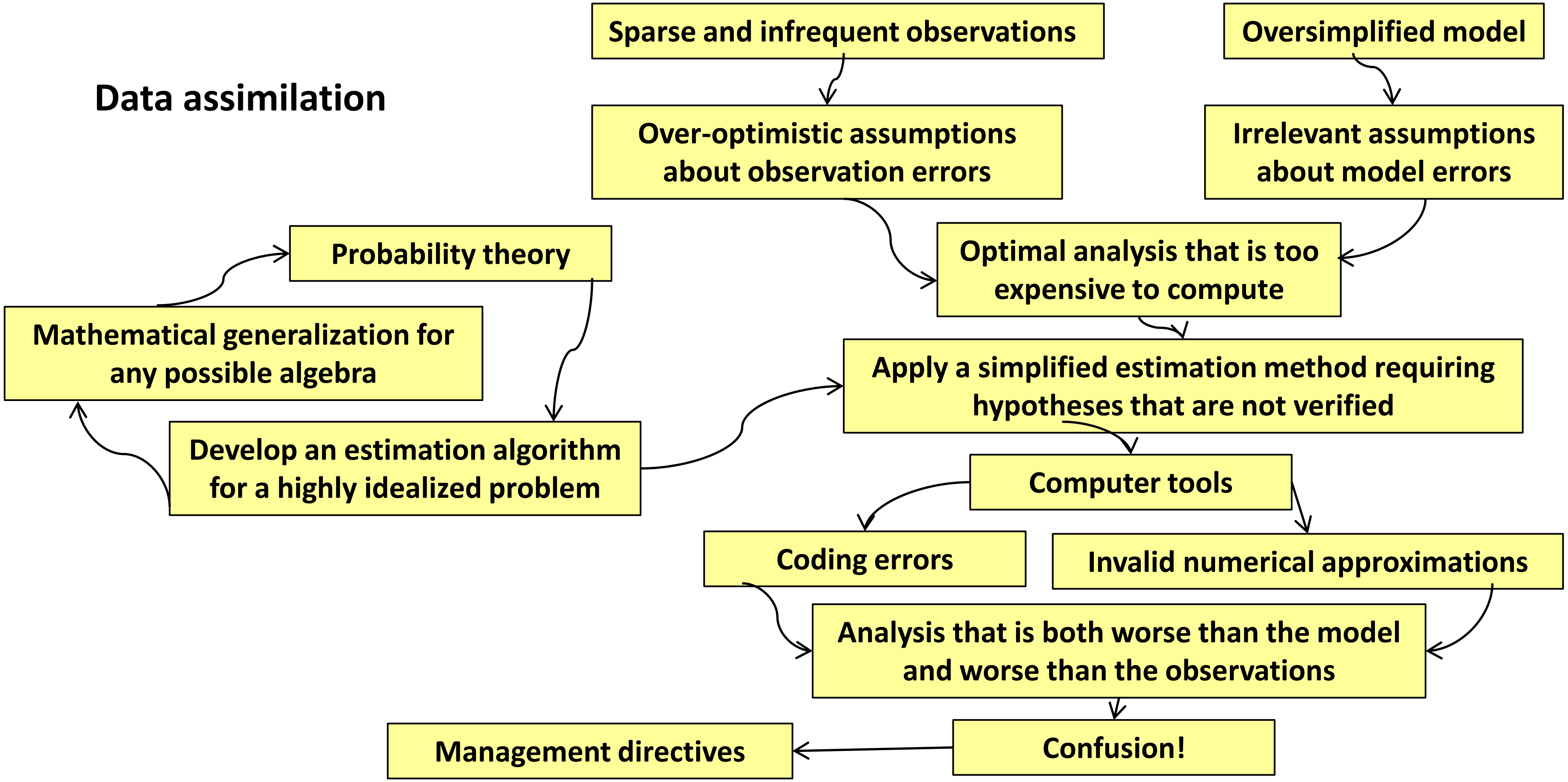



Data assimilation
Sparse and infrequent observations
Oversimplified model
Over-optimistic assumptions about observation errors
Irrelevant assumptions about model errors
Probability theory
Optimal analysis that is too expensive to compute
Mathematical generalization for any possible algebra
Apply a simplified estimation method requiring hypotheses that are not verified
Develop an estimation algorithm for a highly idealized problem
Computer tools
Coding errors
Invalid numerical approximations
Analysis that is both worse than the model and worse than the observations

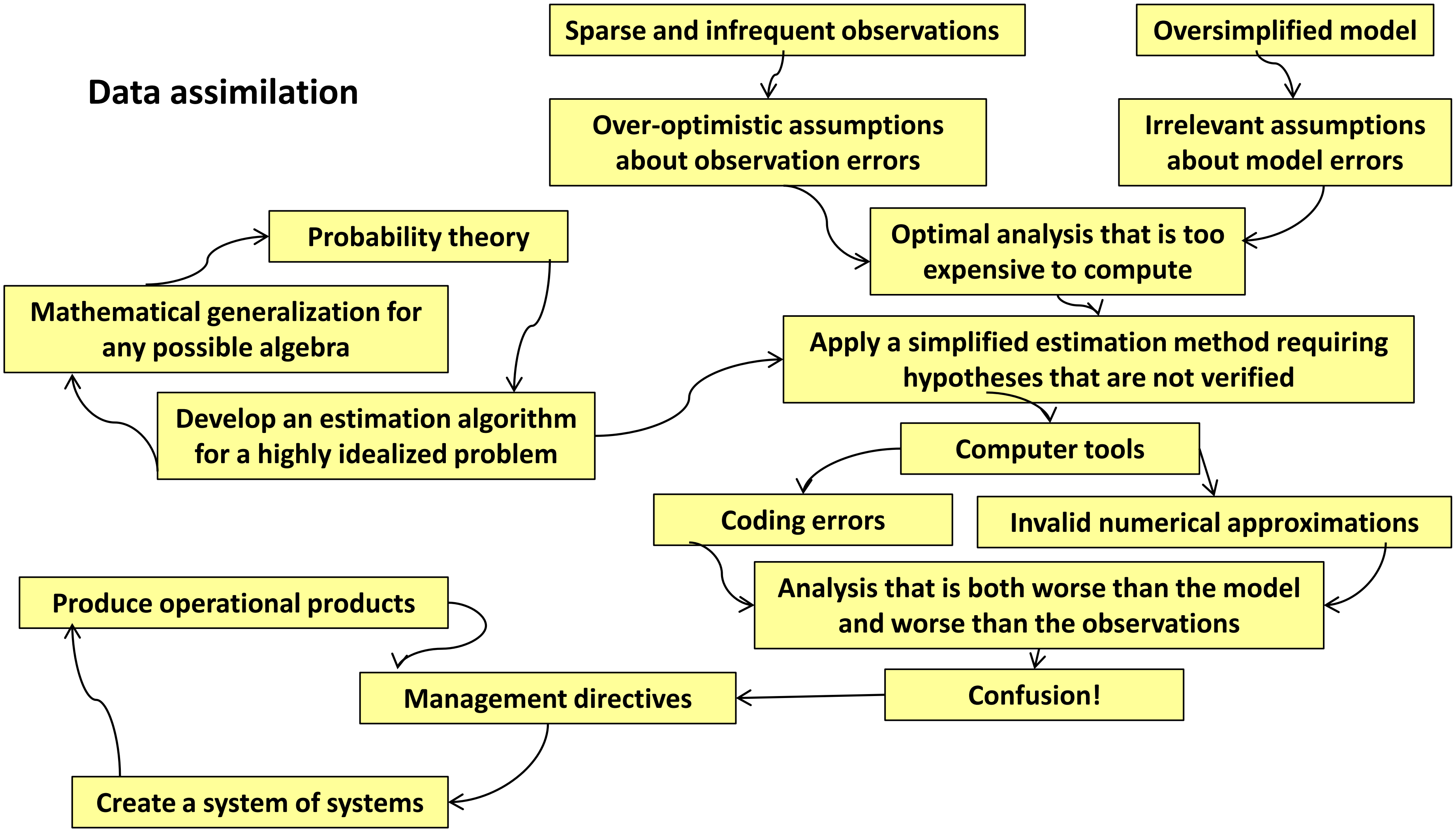
Data assimilation



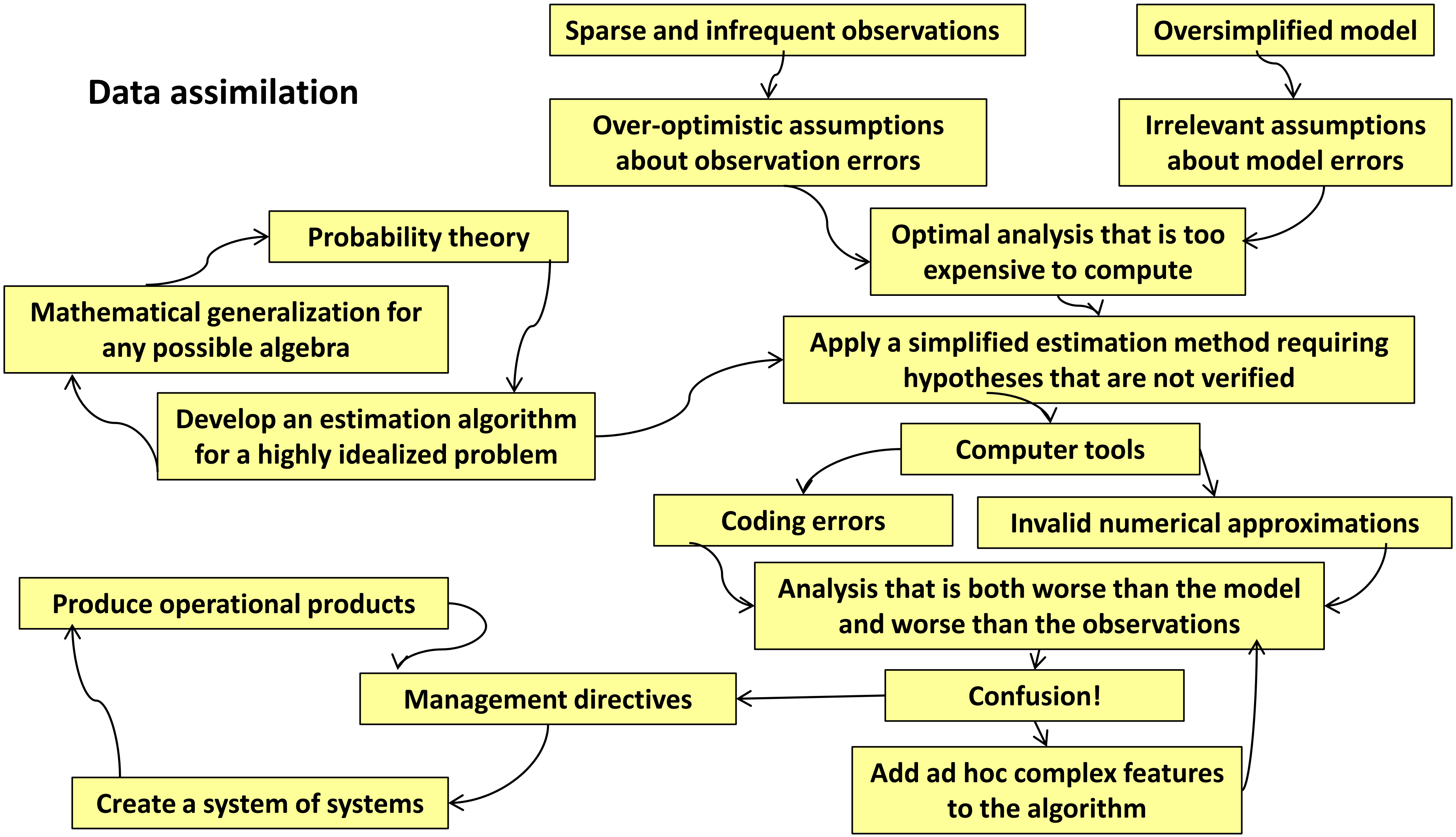
Data assimilation



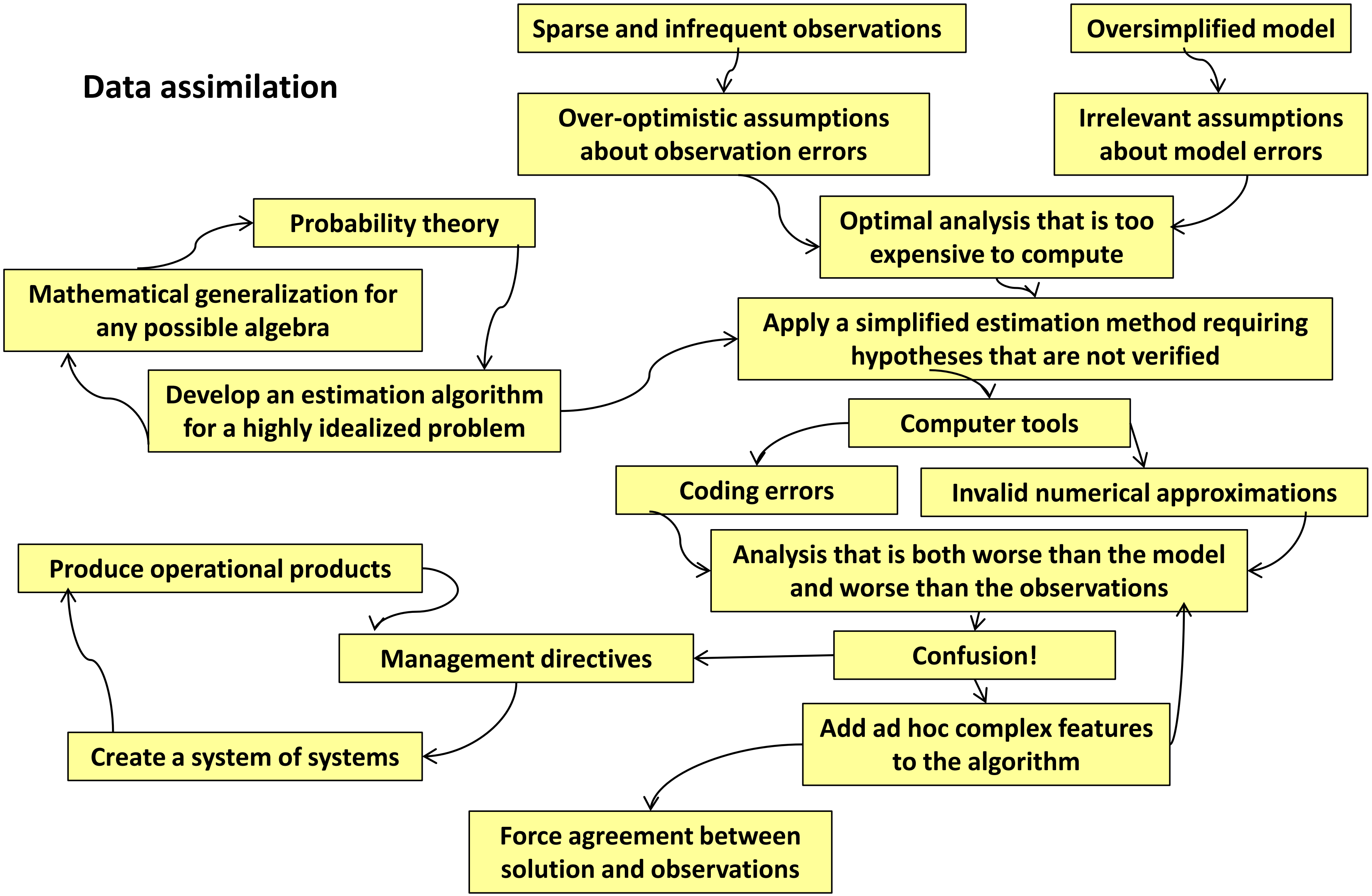
Data assimilation



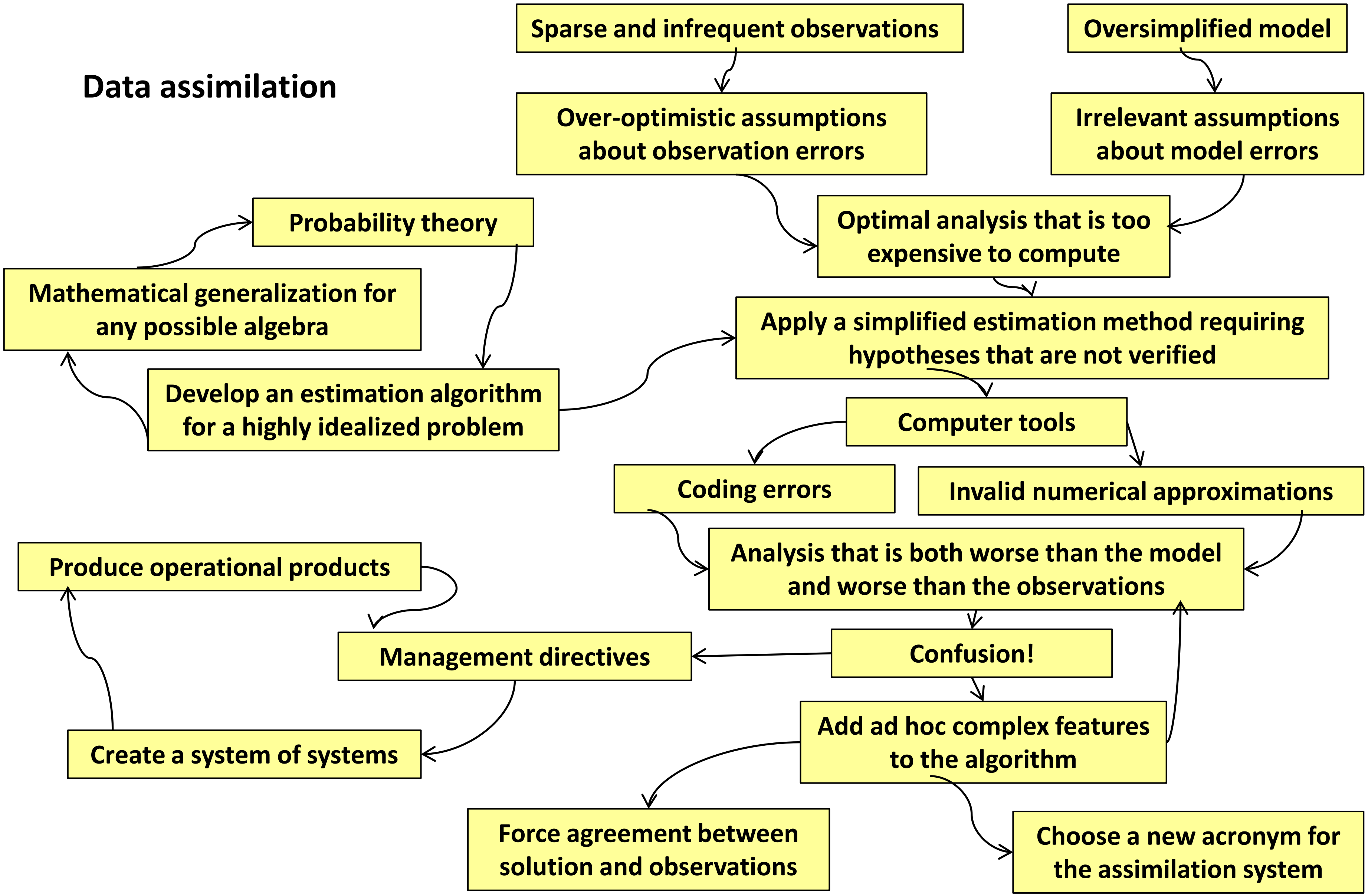
Data assimilation



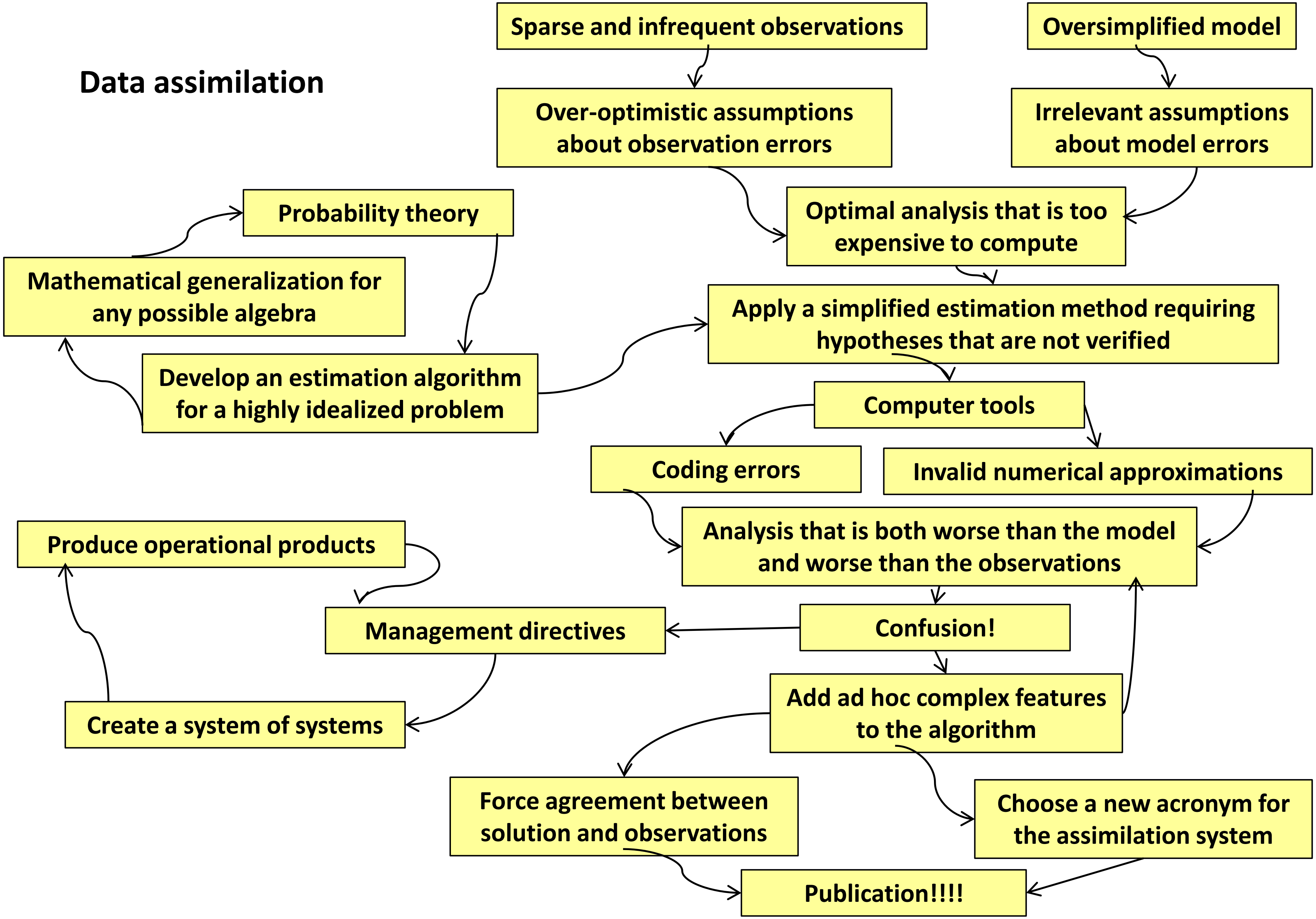
Data assimilation



Data assimilation



Data assimilation



Talk overview

⇒ 1. Back and forth nudging algorithm

2. Diffusive BFN algorithm

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 $X(0) = x_0$.

$Y_{obs}(t)$: observations of the system
H : observation operator.

$$\begin{cases} \dfrac{dX}{dt} = F(X) + K(Y_{obs} - H(X)), \quad 0 < t < T, \\ X(0) = X_0, \end{cases}$$

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.

Forward nudging

– Meteorology : Hoke-Anthes (1976)

– Oceanography (QG model) : De Mey et al. (1987), Verron-Holland (1989)

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

– Optimal determination of the nudging coefficients :
Zou-Navon-Le Dimet (1992), Stauffer-Bao (1993),
Vidard-Le Dimet-Piacentini (2003)
Lakshmivarahan-Lewis (2011)

Forward nudging : linear case

Luenberger observer, or asymptotic observer

(Luenberger, 1966)

$$\begin{cases} \dfrac{dX_{true}}{dt} = FX_{true}, \quad Y_{obs} = HX_{true}, \\ \dfrac{dX}{dt} = FX + K(Y_{obs} - HX). \end{cases}$$

$$\frac{d}{dt}(X - X_{true}) = (F - KH)(X - X_{true})$$

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

Backward nudging

How to recover the initial state from the final solution ?

Backward model :

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

If we apply nudging to this backward model :

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

BFN : Back and Forth Nudging algorithm

Iterative algorithm (forward and backward resolutions) :

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

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

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

[Auroux - Blum, C. R. Acad. Sci. Math. 2005]

If X_k and $\tilde{X}_k$ converge towards the same limit X, and if $K = K'$, then X satisfies the state equation and fits to the observations.

Choice of the direct nudging matrix K

Implicit discretization of the direct model equation with nudging :

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

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

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

by chosing

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

where R is the covariance matrix of the errors of observation, and k is a scalar.

[Auroux-Blum, Nonlin. Proc. Geophys. 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. $rank[H, HF, \dots, HF^{N-1}] = N$, then there exists a matrix K' such that $-F - K'H$ is a Hurwitz matrix (pole assignment method).

Simpler solution : one can define $K' = k'H^T R^{-1}$, where k' is e.g. the smallest value making the backward numerical integration stable.

Example of convergence results

Viscous linear transport equation :

$$\begin{cases} \partial_t u - \nu \partial_{xx} u + a(x)\partial_x u = -K(u - u_{obs}), & u(x, t = 0) = u_0(x) \\ \partial_t \tilde{u} - \nu \partial_{xx} \tilde{u} + a(x)\partial_x \tilde{u} = K'(\tilde{u} - u_{obs}), & \tilde{u}(x, t = T) = u_T(x) \end{cases}$$

We set $w(t) = u(t) - u_{obs}(t)$ and $\tilde{w}(t) = \tilde{u}(t) - u_{obs}(t)$ the errors.

• If K and K' are constant, then $\forall t \in [0, T] : \widetilde{w}(t) = e^{(-K-K')(T-t)} w(t)$

(still true if the observation period does not cover $[0, T]$)

• If the domain is not fully observed, then the problem is ill-posed.

Error after k iterations : $w_k(0) = e^{-[(K+K')kT]} w_0(0)$

$\rightsquigarrow$ exponential decrease of the error, thanks to :

- $K + K'$: infinite feedback to the observations (not physical)
- T : asymptotic observer (Luenberger)
- k : infinite number of iterations (BFN)

[Auroux-Nodet, COCV 2011]

Shallow water model

$$\begin{aligned} \partial_t u - (f+\zeta)v + \partial_x B &= \frac{\tau_x}{\rho_0 h} - ru + \nu\Delta u \\ \partial_t v + (f+\zeta)u + \partial_y B &= \frac{\tau_y}{\rho_0 h} - rv + \nu\Delta v \\ \partial_t h + \partial_x(hu) + \partial_y(hv) &= 0 \end{aligned}$$

- $\zeta = \partial_x v - \partial_y u$ is the relative vorticity ;
- $B = g^* h + \frac{1}{2}(u^2+v^2)$ is the Bernoulli potential ;
- $g^* = 0.02\ m.s^{-2}$ is the reduced gravity ;
- $f = f_0 + \beta y$ is the Coriolis parameter (in the β-plane approximation), with $f_0 = 7.10^{-5}\ s^{-1}$ and $\beta = 2.10^{-11}\ m^{-1}.s^{-1}$;
- $\tau = (\tau_x, \tau_y)$ is the forcing term of the model (e.g. the wind stress), with a maximum amplitude of $\tau_0 = 0.05\ s^{-2}$;
- $\rho_0 = 10^3\ kg.m^{-3}$ is the water density ;
- $r = 9.10^{-8}\ s^{-1}$ is the friction coefficient.
- $\nu = 5\ m^2.s^{-1}$ is the viscosity (or dissipation) coefficient.

Shallow water model

2D shallow water model, state = height h and horizontal velocity (u, v)

Numerical parameters : (run example)

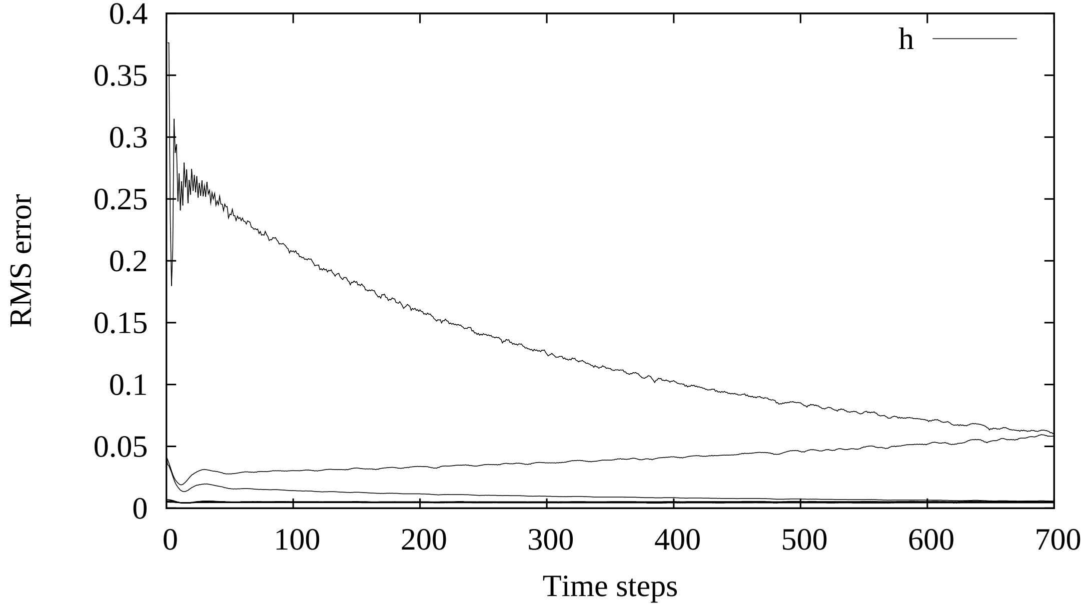
Domain : $L = 2000$ km $\times$ 2000 km ; Rigid boundary and no-slip BC ; Time step = 1800 s ; Assimilation period : 15 days ; Forecast period : 15 + 45 days

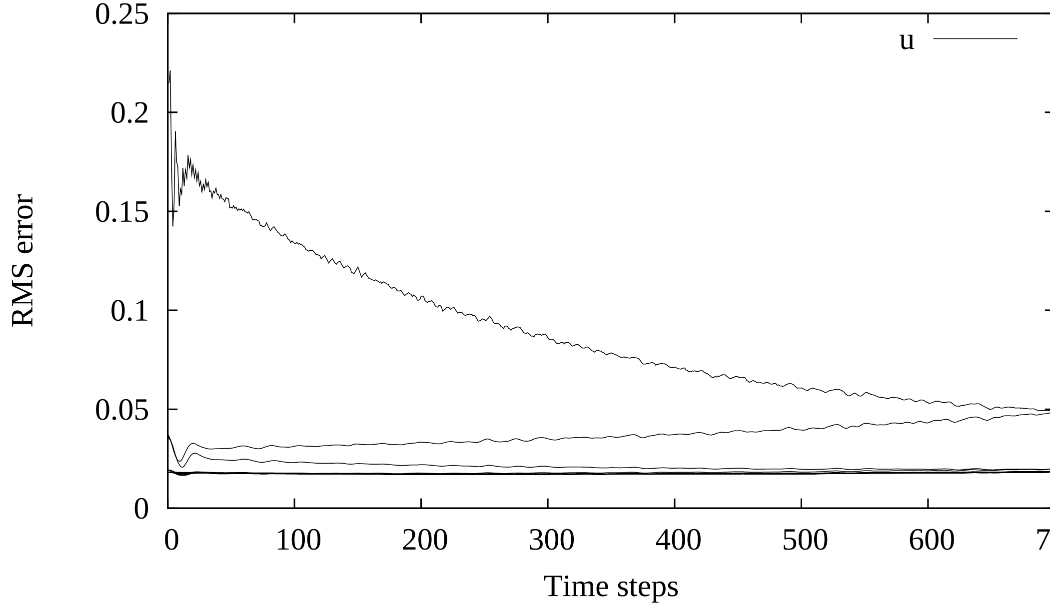
Observations : of h only ($\sim$ satellite obs), every 5 gridpoints in each space direction, every 24 hours.

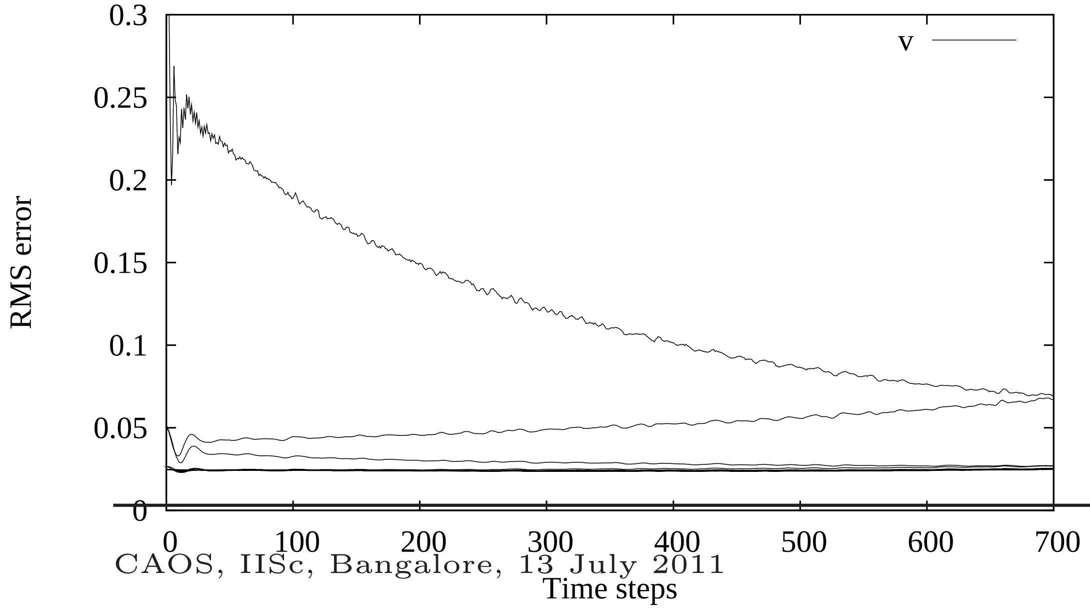
Background : true state one month before the beginning of the assimilation period + white gaussian noise ($\sim$ 10%)

Comparison BFN - 4DVAR : sea height h ; velocity :u and v.

Convergence - perfect obs.



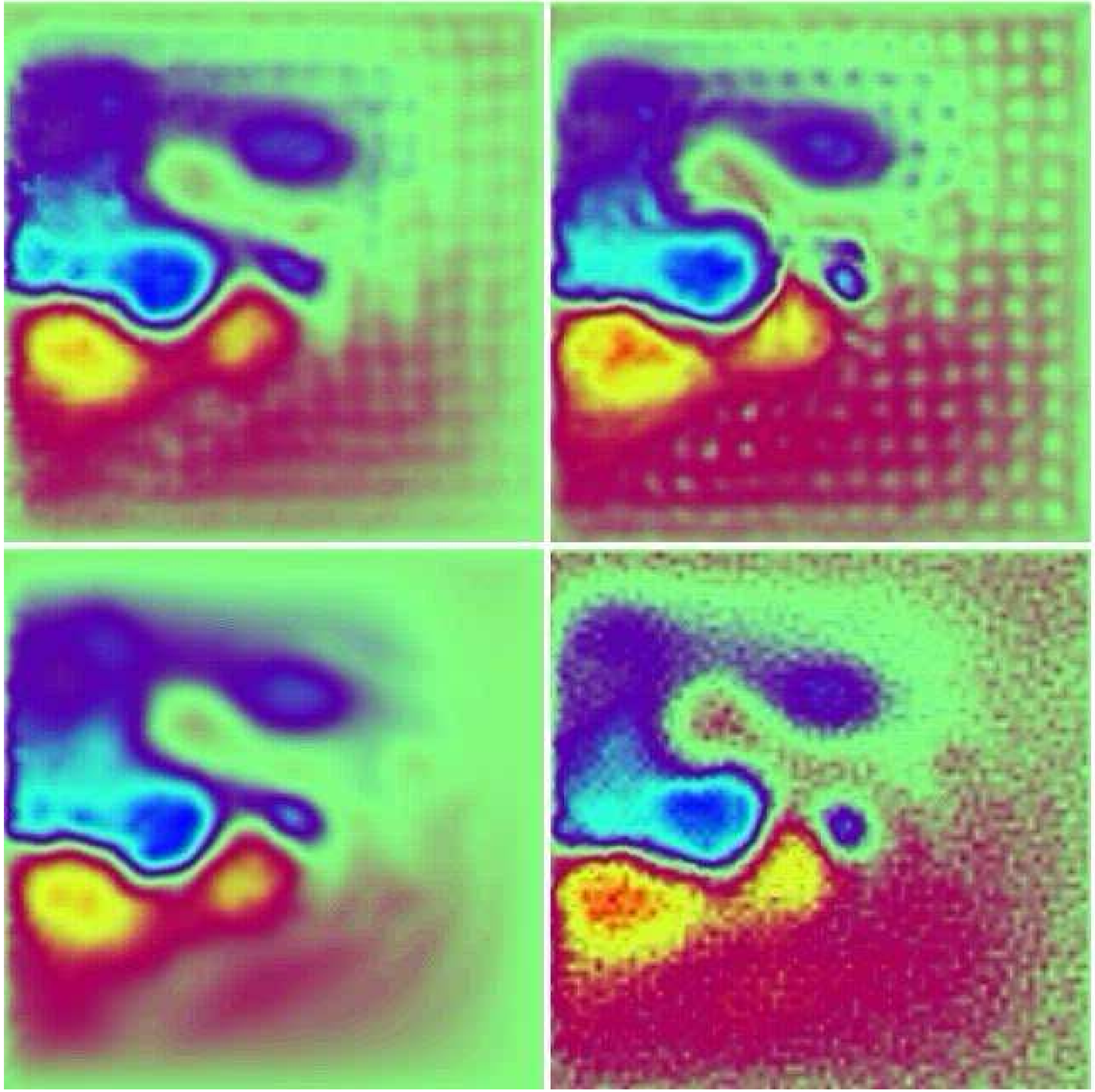




Relative difference between the BFN iterates (5 first iterations) and the true solution versus the time steps, for h, u and v.

[Auroux, Int. J. Num. Meth. Fluids 2009]

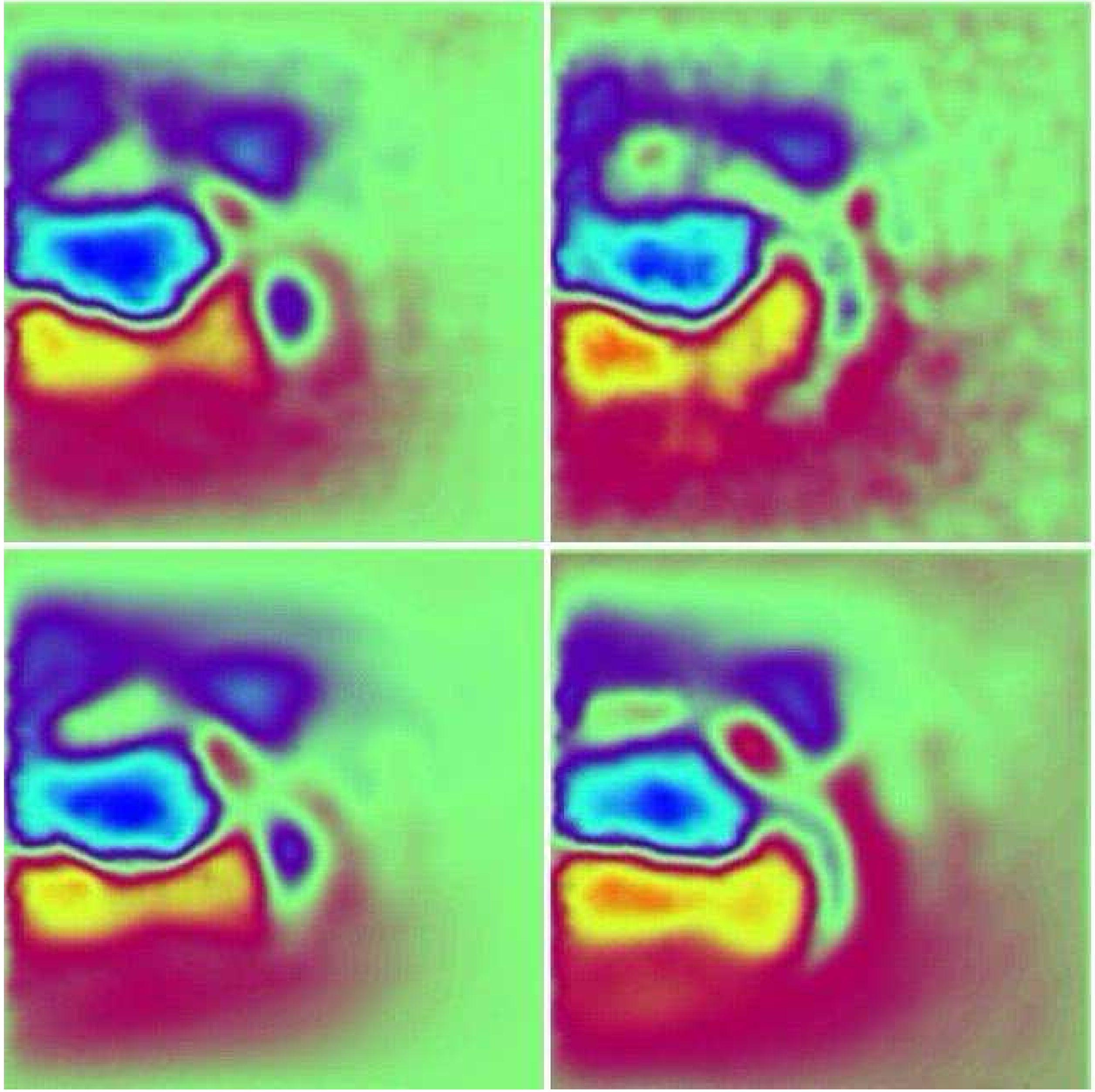
Comparison - noisy obs.



Top : identified initial condition after 5 iterations of BFN and 4D-VAR.

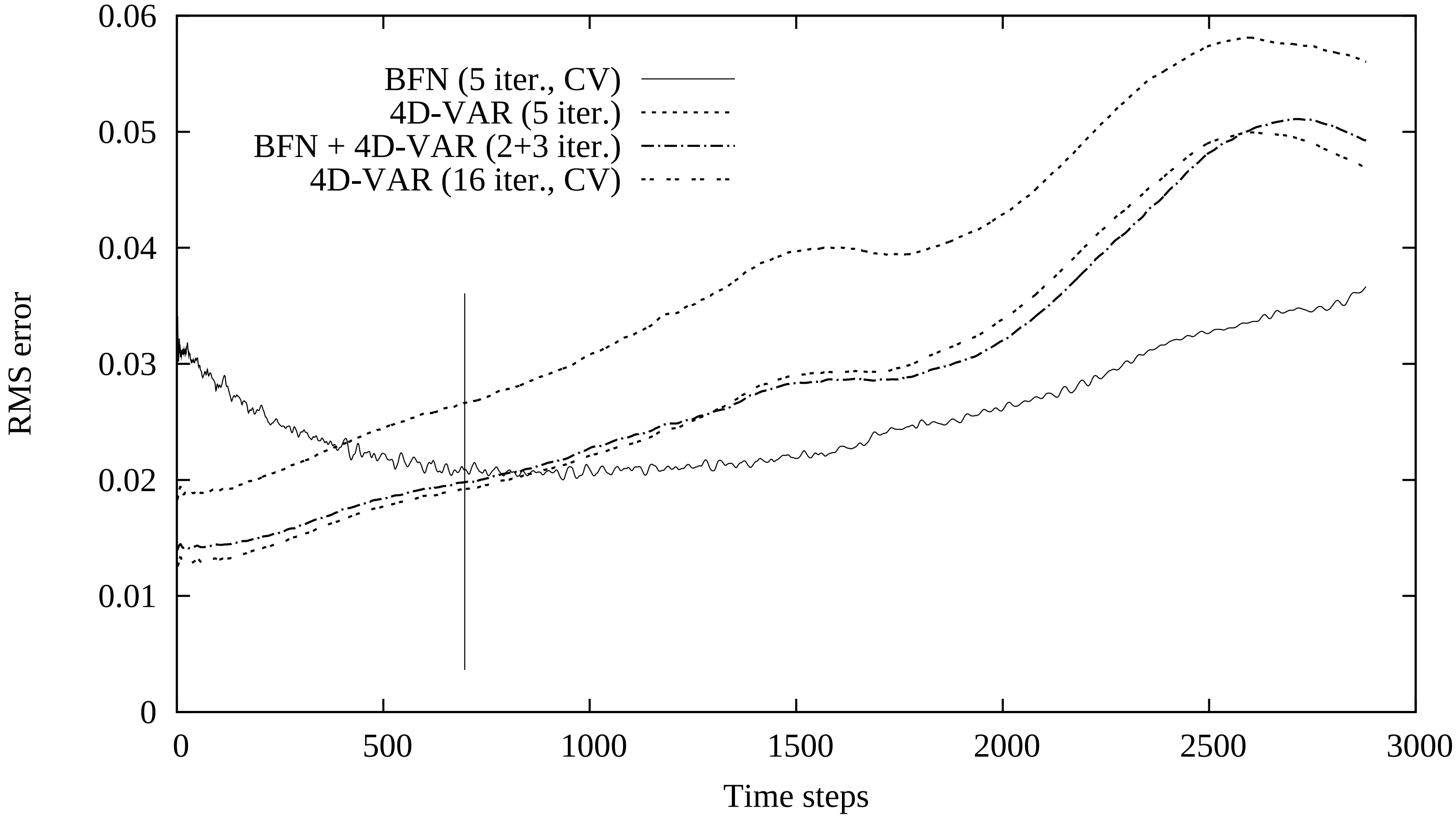
Bottom : true initial condition and background state.

Comparison - noisy obs.



Corresponding states at the end of the forecast period (45 days) : BFN, 4D-VAR, true, background.

BFN-preprocessed 4D-VAR



Relative difference between the true solution and the forecast trajectory corresponding to the BFN, 4D-VAR and BFN-preprocessed 4D-VAR identified initial conditions, vs time, for the height variable in the case of noisy observations.

Diffusion problem

Backward model and diffusion :

The main issue of the BFN is : how to handle diffusion processes in the backward equation ?

Let us consider only diffusion : heat equation (in 1D)

$$\partial_t u = \partial_{xx} u$$

The backward nudging model will be :

$$\partial_t \tilde{u} = \partial_{xx} \tilde{u} + K(\tilde{u} - u_{obs})$$

from time T to 0. By using a change of variable $t' = T - t$, we can rewrite the backward model as a forward one :

$$\partial_{t'} \tilde{u} = -\partial_{xx} \tilde{u} - K(\tilde{u} - u_{obs}),$$

and we can see that even if the nudging term stabilizes the model, the backward diffusion is a real issue.

Numerically, one can solve the backward diffusion equation (with nudging), as the eigenvalues of the discrete Laplacian are bounded, but all eigenvalues are positive, and shifting all the spectrum is not very physical.

From a theoretical point of view, the spectrum of Δ is included in $\mathbb{R}^-$ and once again, the eigenvalues of $-\Delta$ are all positive, and unbounded. Even if the original function has no high frequencies, the correction term (and numerical approximations) will ensure the presence of high frequencies $\Rightarrow$ (positive) exponential divergence in time (with high coefficients !).

$\Rightarrow$ big issue ?

Diffusion problem

Hopefully, in geophysical problems, diffusion is not a dominant term. The model has smoothing properties, and diffusion is small $\rightarrow$ diffusion processes are not highly unstable in backward mode, even if the model is clearly unstable without nudging.

Theoretically, there is a problem :

- **Viscous linear transport equation :** if the support of K is a strict sub-domain (i.e. some parts of the space domain are not observed), there does not exist a solution to the backward model, even in the distribution sense.

- **Viscous Burgers equation :** even if K is constant (in time and space $\Rightarrow$ full observations), the backward equation is ill-posed, as there is no stability (or continuity) with respect to the initial condition.

Without viscosity, one can prove the convergence of the BFN on these equations.

Talk overview

1. Back and forth nudging algorithm

⇒ 2. Diffusive BFN algorithm

Diffusive BFN

Diffusive free equations in the geophysical context :

In meteorology or oceanography, theoretical equations are usually diffusive free (e.g. Euler's equation for meteorological processes).

In a numerical framework, a diffusive term is added to the equations (or a diffusive scheme is used), in order to both stabilize the numerical integration of the equations, and take into consideration some subscale phenomena.

Example : weather forecast is done with Euler's equation (at least in Météo France...), which is diffusive free. Also, in quasi-geostrophic ocean models, people usually consider ∇^4 or ∇^6 for dissipation at the bottom, or for vertical mixing.

Diffusive BFN

Standard BFN algorithm :

Original model :

$$\partial_t X = F(X), \quad 0 < t < T.$$

Corresponding BFN algorithm :

$$\begin{cases} \partial_t X_k = F(X_k) + K(Y_{obs} - H(X_k)), \\ X_k(0) = \tilde{X}_{k-1}(0), \quad 0 < t < T, \end{cases}$$

$$\begin{cases} \partial_t \tilde{X}_k = F(\tilde{X}_k) - K'(Y_{obs} - H(\tilde{X}_k)), \\ \tilde{X}_k(T) = X_k(T), \quad T > t > 0, \end{cases}$$

with the notation $\tilde{X}_0(0) = x_0$.

Diffusive BFN

Addition of a diffusion term :

$$\partial_t X = F(X) + \nu \Delta X, \quad 0 < t < T,$$

where F has no diffusive terms, ν is the diffusion coefficient, and we assume that the diffusion is a standard second-order Laplacian (could be a higher order operator).

We introduce the D-BFN algorithm in this framework, for $k \geq 1$:

$$\begin{cases} \partial_t X_k = F(X_k) + \nu \Delta X_k + K(Y_{obs} - H(X_k)), \\ X_k(0) = \tilde{X}_{k-1}(0), \quad 0 < t < T, \end{cases}$$

$$\begin{cases} \partial_t \tilde{X}_k = F(\tilde{X}_k) - \nu \Delta \tilde{X}_k - K'(Y_{obs} - H(\tilde{X}_k)), \\ \tilde{X}_k(T) = X_k(T), \quad T > t > 0. \end{cases}$$

Diffusive BFN

It is straightforward to see that the backward equation can be rewritten, using $t' = T - t$:

$$\partial_{t'} \tilde{X}_k = -F(\tilde{X}_k) + \nu \Delta \tilde{X}_k + K'(Y_{obs} - H(\tilde{X}_k)), \quad \tilde{X}_k(t' = 0) = X_k(T),$$

where $\tilde{X}$ is evaluated at time t'. As it is now forward in time, this equation can be compared with the forward nudging equation :

$$\partial_t X_k = F(X_k) + \nu \Delta X_k + K(Y_{obs} - H(X_k)), \quad X_k(0) = \tilde{X}_{k-1}(t' = T).$$

Then the backward equation can easily be solved, with an initial condition, and the same diffusion operator as in the forward equation. Only the physical model has an opposite sign.

The diffusion term both takes into account the subscale processes and stabilizes the numerical backward integrations, and the feedback term still controls the trajectory with the observations.

Convergence of D-BFN

We assume here that the model F and the observation operator H are linear. Let us define the following operator that corresponds to one forward + one backward integrations :

$$\psi \quad : \quad (X_1(0), Y_{obs}(0)) \mapsto \tilde{X}_1(0).$$

This operator is linear in the initial conditions, so that there exist C and D linear operators such that

$$X_2(0) = \psi(X_1(0), Y_{obs}(0)) = \psi(X_1(0), 0) + \psi(0, Y_{obs}(0)) = CX_1(0) + DY_{obs}(0).$$

So that the initial state $X_{k+1}(0)$ of the $(k+1)^{\text{th}}$ D-BFN iteration satisfies :

$$X_{k+1}(0) = C^k x_0 + \left(\sum_{m=0}^{k-1} C^m \right) DY_{obs}(0)$$

Convergence of D-BFN

If the spectrum of C is included in $[-\rho;\rho]$, with $\rho < 1$, then $C^k \to 0$ and $\sum_{m=0}^{k} C^m \to (I-C)^{-1}$ when $k \to \infty$. Therefore, in that case, $X_k(0)$ converges as k goes to infinity to X_∞ solution of

$$X_\infty = (I-C)^{-1} D Y_{obs}(0)$$

Linear transport equation

$$\partial_t u + a(x)\,\partial_x u = 0, \quad t \in [0,T],\ x \in \Omega, \quad u(t=0) = u_0 \in L^2(\Omega)$$

with periodic boundary conditions, and we assume that $a \in W^{1,\infty}(\Omega)$.

Numerically, for both stability and subscale modelling, the following equation would be solved :

$$\partial_t u + a(x)\,\partial_x u = \nu \partial_{xx} u, \quad t \in [0,T], x \in \Omega, \quad u(t=0) = u_0 \in L^2(\Omega),$$

where $\nu \geq 0$ is assumed to be constant.

Linear transport equation

Let us assume that the observations satisfy the physical model (without diffusion) :

$$\partial_t u_{obs} + a(x)\,\partial_x u_{obs} = 0, \quad t \in [0,T], x \in \Omega, \quad u_{obs}(t=0) = u^0_{obs} \in L^2(\Omega).$$

We assume in this idealized situation that the system is fully observed (and H is then the identity operator).

Then the D-BFN algorithm applied to this problem gives, for $k \geq 1$:

$$\left\{\begin{array}{l} \partial_t u_k + a(x)\,\partial_x u_k = \nu \partial_{xx} u_k + K(u_{obs,k} - u_k), \\ t \in [2(k-1)T, 2(k-1)T + T], x \in \Omega \\ u_k(2(k-1)T, x) = \tilde{u}_{k-1}(2(k-1)T, x) \end{array}\right.$$

$$\left\{\begin{array}{l} \partial_t \tilde{u}_k - a(x)\,\partial_x \tilde{u}_k = \nu \partial_{xx} \tilde{u}_k + K(\tilde{u}_{obs,k} - \tilde{u}_k), \\ t \in [2kT - T, 2kT], x \in \Omega \\ \tilde{u}_k(2kT - T, x) = u_k(2kT - T, x). \end{array}\right.$$

Energy estimate

$$d_t\|u\|^2 \leq -2\nu\|\partial_x u\|^2 - (2K - \|\partial_x a\|_\infty)\|u\|^2 \leq -\delta\|u\|^2,$$

where $\delta = 2K - \|\partial_x a\|_\infty$ is non negative for K large enough. Therefore $\|Cu_0\|^2 \leq e^{-2\delta T}\|u_0\|^2$, so that $\|C\| < 1$ and convergence is ensured. Note that it is totally independent of $\nu \Rightarrow$ one can consider very small diffusion coefficients.

In the special case where $a(x) = a \in \mathbb{R}$, we can change variables to straighten characteristics as follows. Setting $v_k(t,y) = u_k(t, y + a(t - 2(k-1)T))$ and $\tilde{v}_k(t,z) = \tilde{u}_k(t, z - a(t - 2kT))$ leads to

$$\partial_t v_k = \nu\partial_{yy} v_k + K(u^0_{obs}(y) - v_k), \qquad \partial_t \tilde{v}_k = \nu\partial_{zz}\tilde{v}_k + K(u^0_{obs}(z) - \tilde{v}_k).$$

Smoothing equation

At the limit $k \to \infty$, v_k and $\tilde{v}_k$ tend to $v_\infty(x)$ solution of

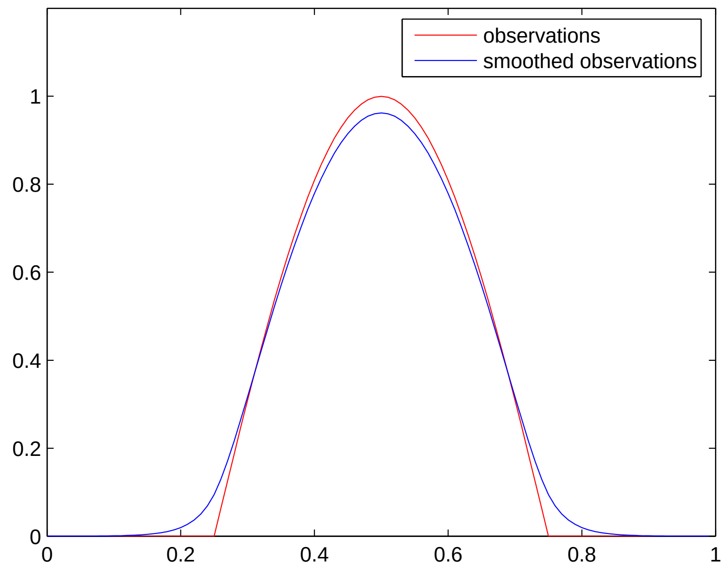
$$\nu \partial_{xx} v_\infty + K(u^0_{obs}(x) - v_\infty) = 0,$$

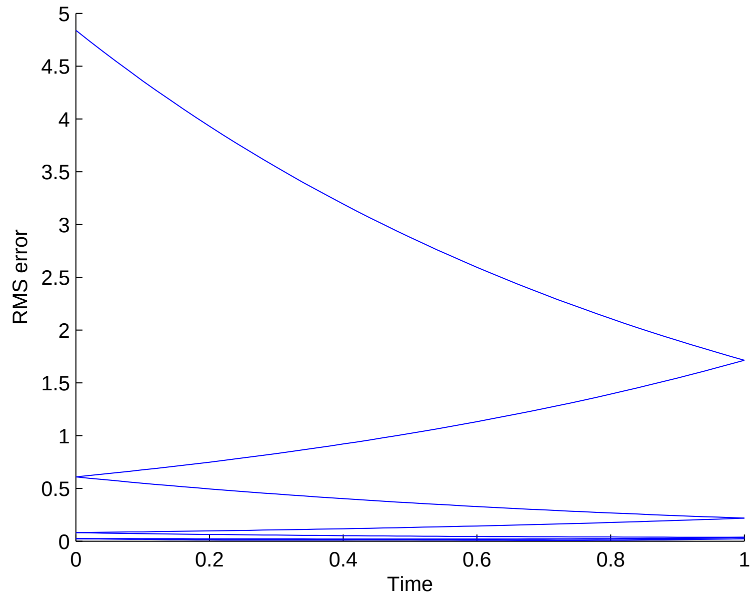
or equivalently

$$-\frac{\nu}{K}\partial_{xx} v_\infty + v_\infty = u^0_{obs}.$$

This equations is well known in signal or image processing, as being the standard linear diffusion restoration equation. In some sense, v_∞ is the result of a smoothing process on the observations u_{obs}, where the degree of smoothness is given by the ratio $\frac{\nu}{K}$.

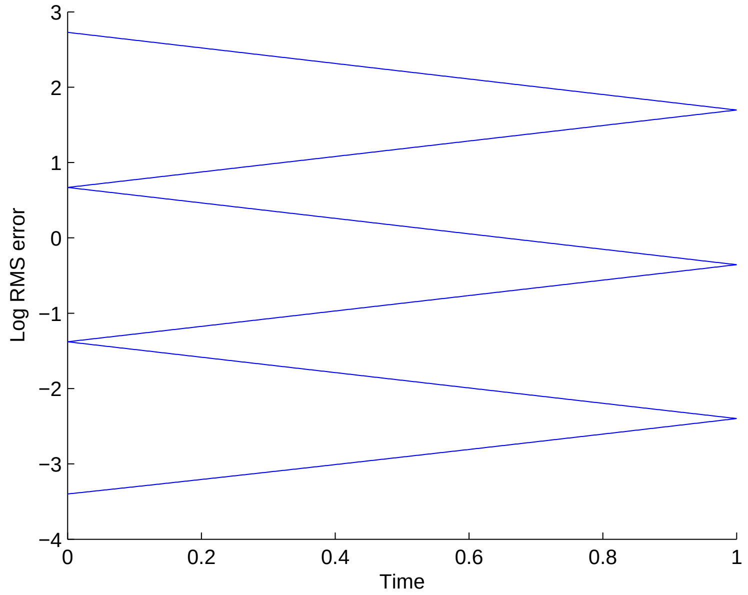
Numerical experiments





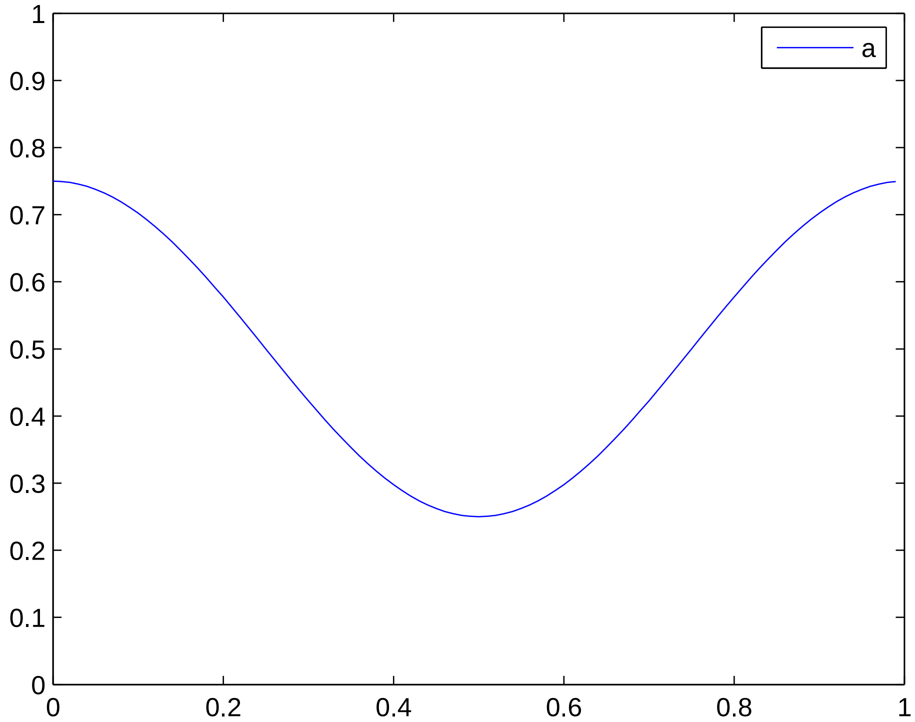
Initial condition of the observation and corresponding smoothed solution ; RMS difference between the BFN iterates and the smoothed observations ; same in semi-log scale.

Movie



Numerical experiments

Linear transport equation with non-constant transport :



Movie

Burgers equation

1D inviscid Burgers equation :

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial u^2}{\partial x} = 0,$$

with a given initial condition $u(x,0)$ and periodic boundary conditions.

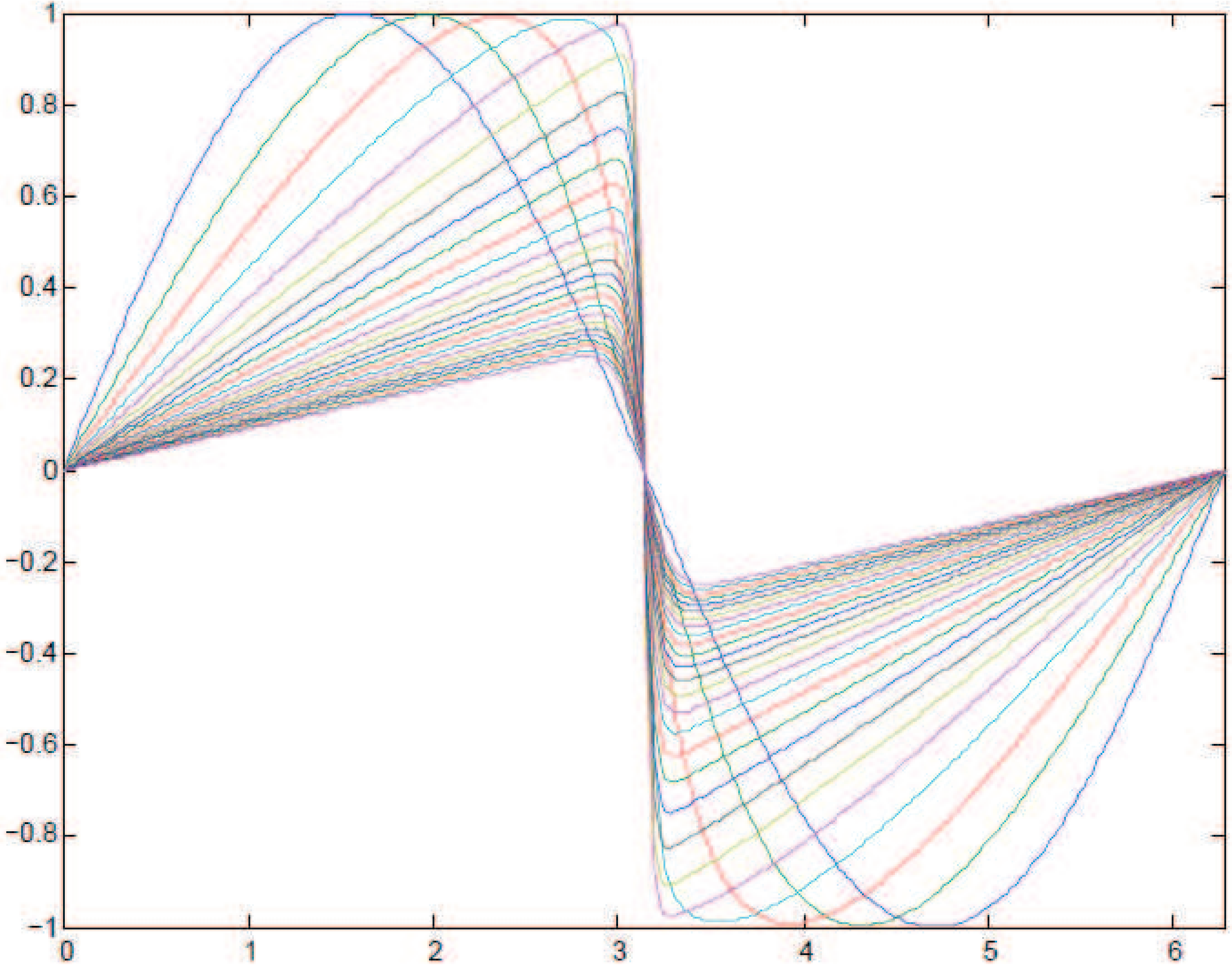
Diffusive BFN :

$$\begin{cases} \dfrac{\partial u_k}{\partial t} + \dfrac{1}{2}\dfrac{\partial u_k^2}{\partial x} = \nu \dfrac{\partial^2 u_k}{\partial x^2} + K(u_{obs} - H(u_k)), \quad 0 < t < T,\ 0 < x < L, \\ u_k(x,0) = \tilde{u}_{k-1}(x,0), \quad 0 < x < L, \end{cases}$$

$$\begin{cases} \dfrac{\partial \tilde{u}_k}{\partial t} + \dfrac{1}{2}\dfrac{\partial \tilde{u}_k^2}{\partial x} = -\nu \dfrac{\partial^2 \tilde{u}_k}{\partial x^2} - K'(u_{obs} - H(\tilde{u}_k)), \quad 0 < t < T,\ 0 < x < L, \\ \tilde{u}_k(x,T) = u_k(x,T), \quad 0 < x < L. \end{cases}$$

Numerical experiments

Inviscid Burgers equation : creation of shocks in finite time



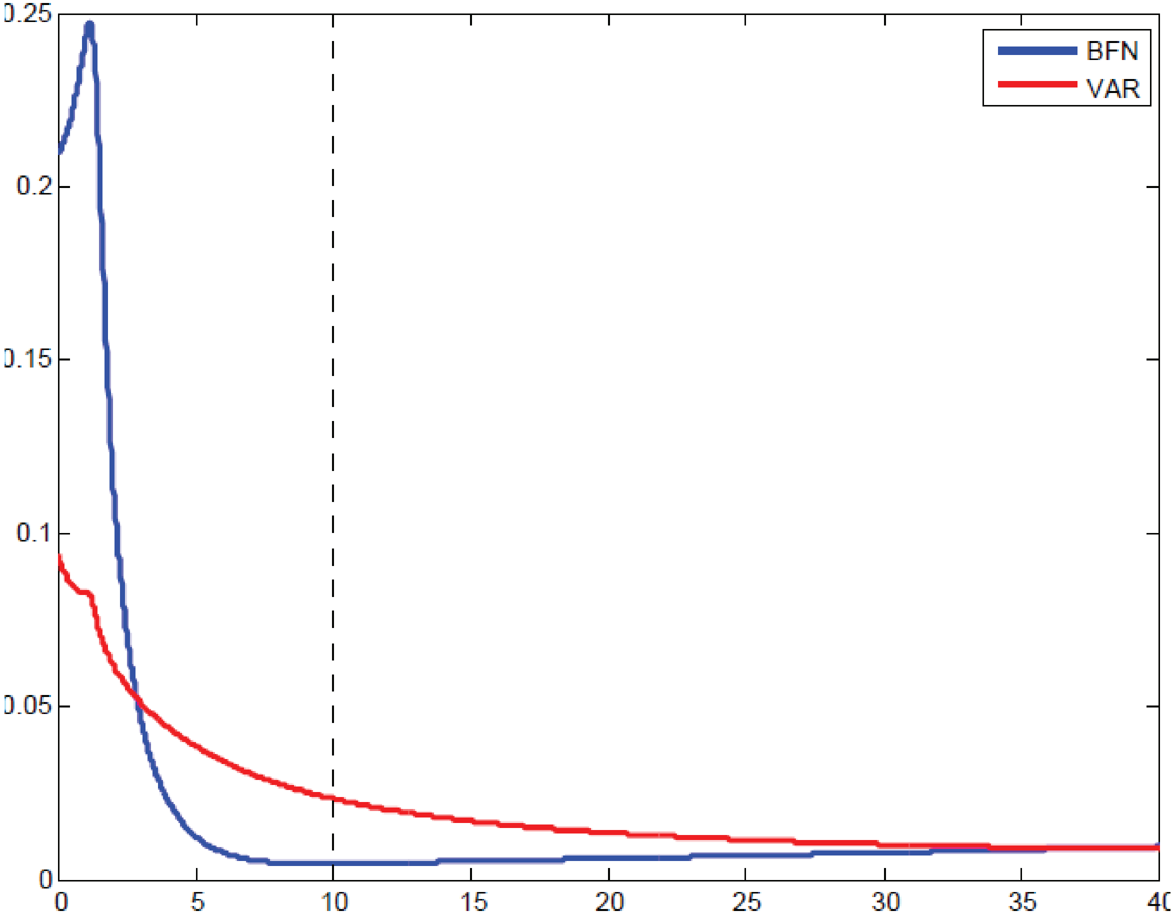
Numerical experiments

Comparison with a variational method :

		$n_x = 4$ $n_t = 4$ unnoisy	$n_x = 10$ $n_t = 10$ unnoisy	$n_x = 10$ $n_t = 10$ noisy (15%)
VAR	Number of iterations	18	20	15
	Relative RMS (%)	0.49	1.64	10.74
BFN2	Number of iterations	2	2	2
	Relative RMS (%)	0.34	0.69	3.50
		$(K = 30,\ K' = 60)$	$(K = 40,\ K' = 80)$	$(K = 10,\ K' = 20)$

Comparison between D-BFN and variational algorithms in the case of sparse and noisy observations on Burgers' equation with shock.

Numerical experiments



Forecast error (difference between the true trajectory and the solutions of the direct model initialized with the identified solutions) for D-BFN and VAR algorithms, with sparse ($n_x = 4 = n_t$) and noisy observations (15% noise).

Conclusions

Back and Forth Nudging algorithm :

- Easy implementation (no linearization, no adjoint state, no minimization process)
- Very efficient in the first iterations (faster convergence)
- Lower computational and memory costs than other DA methods
- Stabilization of the backward model
- Excellent preconditioner for 4D-VAR (or Kalman filters)

Diffusive BFN algorithm :

- Converges even faster, with smaller backward nudging coefficients
- Still produces very precise forecasts
- ...

Perspectives

Under investigation :

- Tests on a full primitive model (NEMO ocean model) : PhD thesis of G. Ruggiero (Univ. Nice), very promising results on twin experiments with D-BFN (while it was hard to make the backward model converge with the standard BFN)

- + perspectives of the standard BFN : correction of non-observed variables from the knowledge of only the SSH ; efficient resolution of Riccati-like equations for a better backward nudging matrix ; ...