Accelerating minimizations in ensemble variational assimilation

G. Desroziers and L. Berre
Météo-France and CNRS

Outline

1. Ensemble Variational assimilation
2. Accelerating minimizations
3. Conclusion and future work

Outline

1. Ensemble Variational assimilation

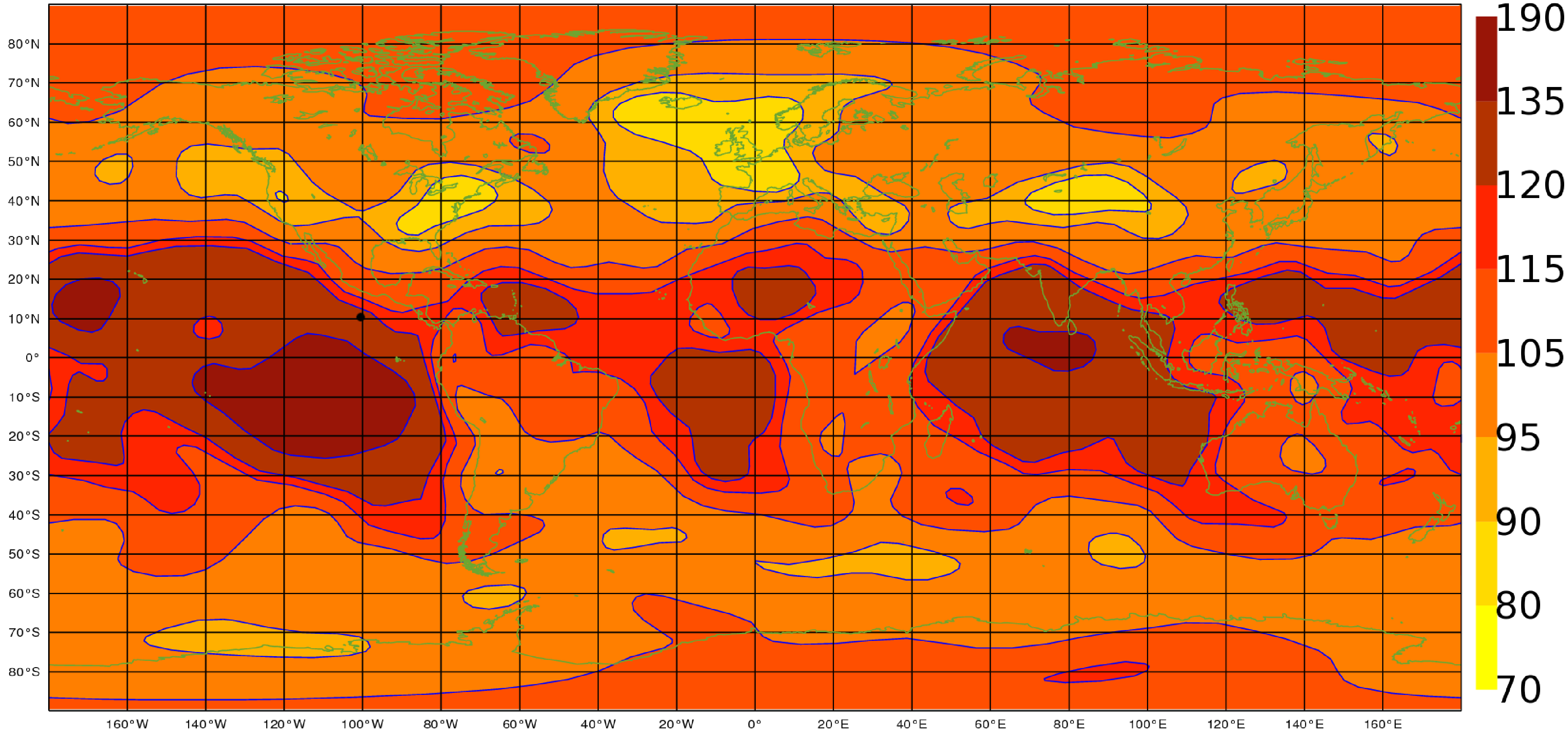
2. Accelerating minimizations

3. Conclusion and future work

The operational Météo-France ensemble Var assimilation

- 1st operational implementation of En Var assim. (2008; ECMWF, 2010-11).
- Six perturbed members, T399 L70 (50 km), with global 4D-Var Arpege.
- Spatial filtering of error variances.
- Inflation of ensemble **B** / model error contributions.
- Flow-dependent background error variances in 4D-Var (and EnDA), for minimizations (all variables) and observation QC.
- Initialization of Météo-France ensemble prediction by EnDA.
- Flow-dependent background error correlations experimented.

Background error correlations using EnDA and wavelets



Wavelet-implied horizontal length-scales (in km),
for wind near 500 hPa, averaged over a 4-day period.

(Varella et al 2011b, and also Fisher 2003,
Deckmyn and Berre 2005, Pannekoucke et al 2007)

Outline

1. Ensemble Variational assimilation

2. Accelerating minimizations

3. Conclusion and future work

Ensemble variational assimilation

- Minimize N cost-functions J_n, n=1, N, with perturbed innovations $\mathbf{d}_n$:

$$J_n(\delta \mathbf{x}_n) = 1/2\ \delta \mathbf{x}_n^T\ \mathbf{B}^{-1}\ \delta \mathbf{x}_n + 1/2\ (\mathbf{d}_n - \mathbf{H}_n\ \delta \mathbf{x}_n)^T\ \mathbf{R}^{-1}\ (\mathbf{d}_n - \mathbf{H}_n\ \delta \mathbf{x}_n),$$

with $\mathbf{B}$ and $\mathbf{R}$ background and observation error matrices,
and $\mathbf{d}_n = \mathbf{y}^o + \mathbf{R}^{1/2}\eta^o_n - H(M(\mathbf{x}^b_n))$, with η^o_n a vector of random numbers.

$$\mathbf{x}^a_n = \mathbf{x}^b_n + \delta \mathbf{x}_n$$

- Perturbed backgrounds for the next analyses:

$\mathbf{x}^b{}_n^+ = M(\mathbf{x}^a_n) + \mathbf{Q}^{1/2}\eta^m_n$, with η^m_n a vector of random numbers and $\mathbf{Q}$ model error covariance matrix.

Hessian matrix of the assimilation problem

- Hessian of the cost-function:

$$J'' = \mathbf{B}^{-1} + \mathbf{H}^{T} \mathbf{R}^{-1} \mathbf{H}.$$

- Bad conditioning of J'': very slow (or no) convergence.

- Cost-function with $\mathbf{B}^{1/2}$ preconditioning ($\delta\mathbf{x} = \mathbf{B}^{1/2}\chi$):

$$J(\chi) = 1/2 \; \chi^{T}\chi + 1/2 \; (\mathbf{d}-\mathbf{H}\;\mathbf{B}^{1/2}\chi)^{T} \; \mathbf{R}^{-1} \; (\mathbf{d}-\mathbf{H}\;\mathbf{B}^{1/2}\chi).$$

- Hessian of the cost-function:

$$J'' = \mathbf{I} + \mathbf{B}^{T/2}\mathbf{H}^{T}\mathbf{R}^{-1}\;\mathbf{H}\;\mathbf{B}^{1/2}.$$

- Far better conditioning and convergence!
(Lorenc 1988, Haben et al 2011)

Lanczos algorithm

- Generate iteratively a set of K orthonormal vectors **q** such as

$$\mathbf{Q}_K^T \, \mathbf{J}'' \, \mathbf{Q}_K = \mathbf{T}_k,$$

where $\mathbf{Q}_K = (\mathbf{q}_1 \, \mathbf{q}_2 \ldots \mathbf{q}_K)$, and $\mathbf{T}_K$ is a tri-diagonal matrix.

- The extremal eigenvalues of $\mathbf{T}_K$ quickly converge towards the extremal eigenvalues of **J**''.

- If $\mathbf{T}_K = \mathbf{Y}_K \Lambda_K \mathbf{Y}_K^T$ is the eigendecomposition of $\mathbf{T}_K$, the Ritz vectors are obtained with

$$\mathbf{Z}_K = \mathbf{Q}_K \mathbf{Y}_k$$

and the Ritz pairs $(\mathbf{z}_k, \lambda_k)$ approximate the eigenpairs of **J**''.

Lanczos algorithm / Conjugate gradient

- Use of the Lanczos vectors to get the solution of the variational problem:

$$\chi_K = \chi_0 + \mathbf{Q}_K \Omega_K .$$

- Optimal coefficients Ω_k should make the gradient of J vanish at χ_K:

$$\begin{aligned}\mathbf{J}'(\chi_K) &= \mathbf{J}'(\chi_0) + \mathbf{J}'' (\chi_K - \chi_0) \\ &= \mathbf{J}'(\chi_0) + \mathbf{J}'' \mathbf{Q}_K \Omega_K \\ &= \mathbf{0},\end{aligned}$$

which gives

$$\begin{aligned}\Omega_K &= - (\mathbf{Q}_K^T \mathbf{J}'' \mathbf{Q}_K)^{-1} \mathbf{Q}_K^T \mathbf{J}'(\chi_0) \\ &= - \mathbf{T}_K^{-1} \mathbf{Q}_K^T \mathbf{J}'(\chi_0),\end{aligned}$$

and then

$$\chi_K = \chi_0 - \mathbf{Q}_K \mathbf{T}_K^{-1} \mathbf{Q}_K^T \mathbf{J}'(\chi_0).$$

- Same solution as after K iterations of a Conjugate Gradient algorithm. (Paige and Saunders 1975, Fisher 1998)

Accelerating a « perturbed » minimization using « unperturbed » Lanczos vectors

- Minimizations with
  - unperturbed innovations **d** and
  - perturbed innovations $\mathbf{d}_n$ have basically the same Hessians:

$$J''(\mathbf{d}) = \mathbf{I} + \mathbf{B}^{T/2}\mathbf{H}^{T}\mathbf{R}^{-1}\,\mathbf{H}\,\mathbf{B}^{1/2},$$
$$J''(\mathbf{d}_n) = \mathbf{I} + \mathbf{B}^{T/2}\mathbf{H}_n^{T}\mathbf{R}^{-1}\,\mathbf{H}_n\,\mathbf{B}^{1/2},$$

- The solution obtained for the « unperturbed » problem

$$\chi_K = \chi_0 - \mathbf{Q}_K\,(\,\mathbf{Q}_K^{T}\,J''\,\mathbf{Q}_K\,)^{-1}\,\mathbf{Q}_K^{T}\,\mathbf{J}'(\chi_0, \mathbf{d})$$

can be transposed to the « perturbed » minimization

$$\chi_{K,n} = \chi_0 - \mathbf{Q}_K\,(\,\mathbf{Q}_K^{T}\,J''\,\mathbf{Q}_K\,)^{-1}\,\mathbf{Q}_K^{T}\,\mathbf{J}'(\chi_0, \mathbf{d}_n)$$

to improve its starting point.

Accelerating minimizations using « perturbed » Lanczos vectors

- If N perturbed minimizations, with K iterations, already performed, then the starting pt of a perturbed (or unpert.) minim. can be written

$$\chi_K = \chi_0 + \mathbf{Q}_{K,N}\,\Omega_{K,N},$$

  where $\Omega_{K,N}$ is a vector of N x K coefficients and

$$\mathbf{Q}_{K,N} = (\mathbf{q}_{1,1} \cdots \mathbf{q}_{K,1} \cdots \mathbf{q}_{1,N} \cdots \mathbf{q}_{K,N})$$

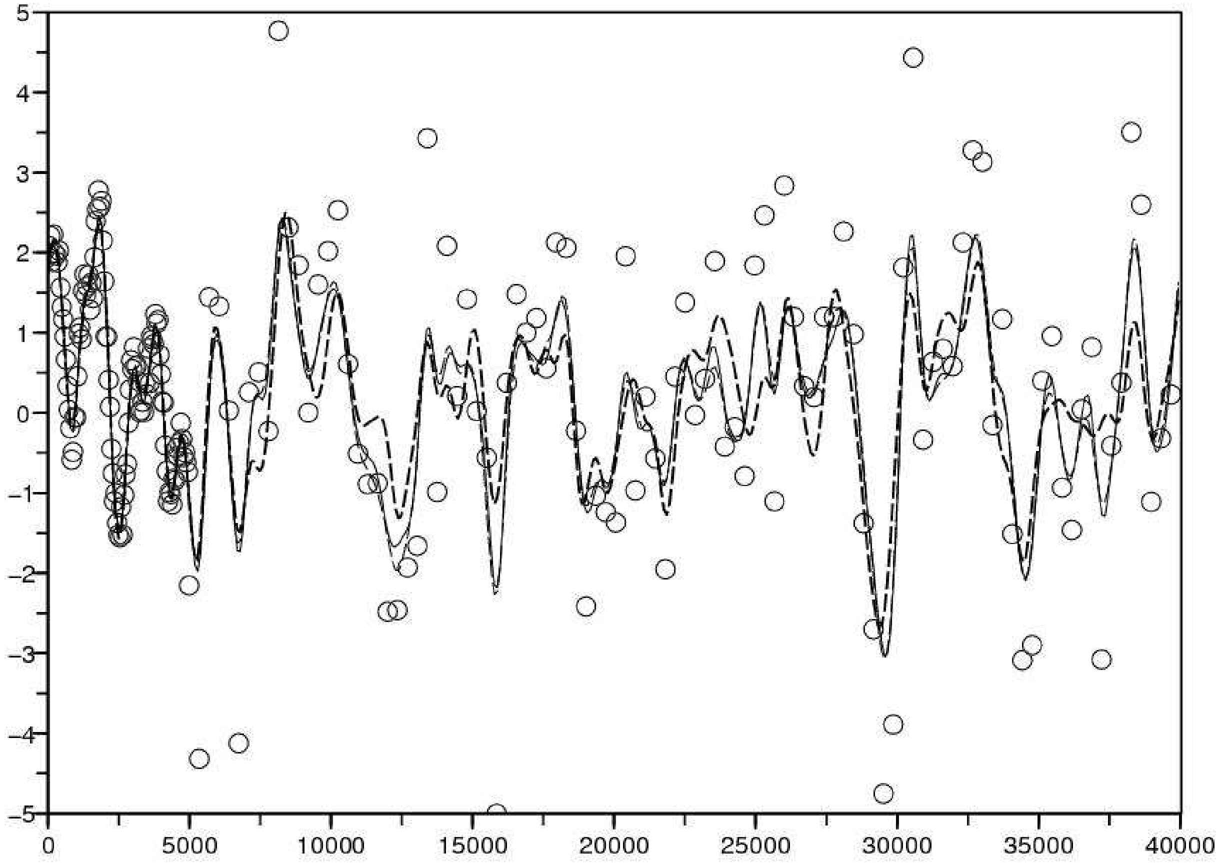
  is a matrix containing the N x K Lanczos vectors.

- Following the same approach as above, the solution can be expressed:

$$\chi_{K,N} = \chi_0 - \mathbf{Q}_{K,N}\,(\mathbf{Q}_{K,N}^T\,\mathbf{J}''\,\mathbf{Q}_{K,N})^{-1}\,\mathbf{Q}_{K,N}^T\,\mathbf{J}'(\chi_0).$$

- Matrix $\mathbf{Q}_{K,N}^T\,\mathbf{J}''\,\mathbf{Q}_{K,N}$ is no longer tri-diagonal, but can be easily inverted.

Accelerating minimizations using N sets of « perturbed » Lanczos vectors (K = 10)



n=401 (δs = 100km)

p=200 (δs^o=50/350km)

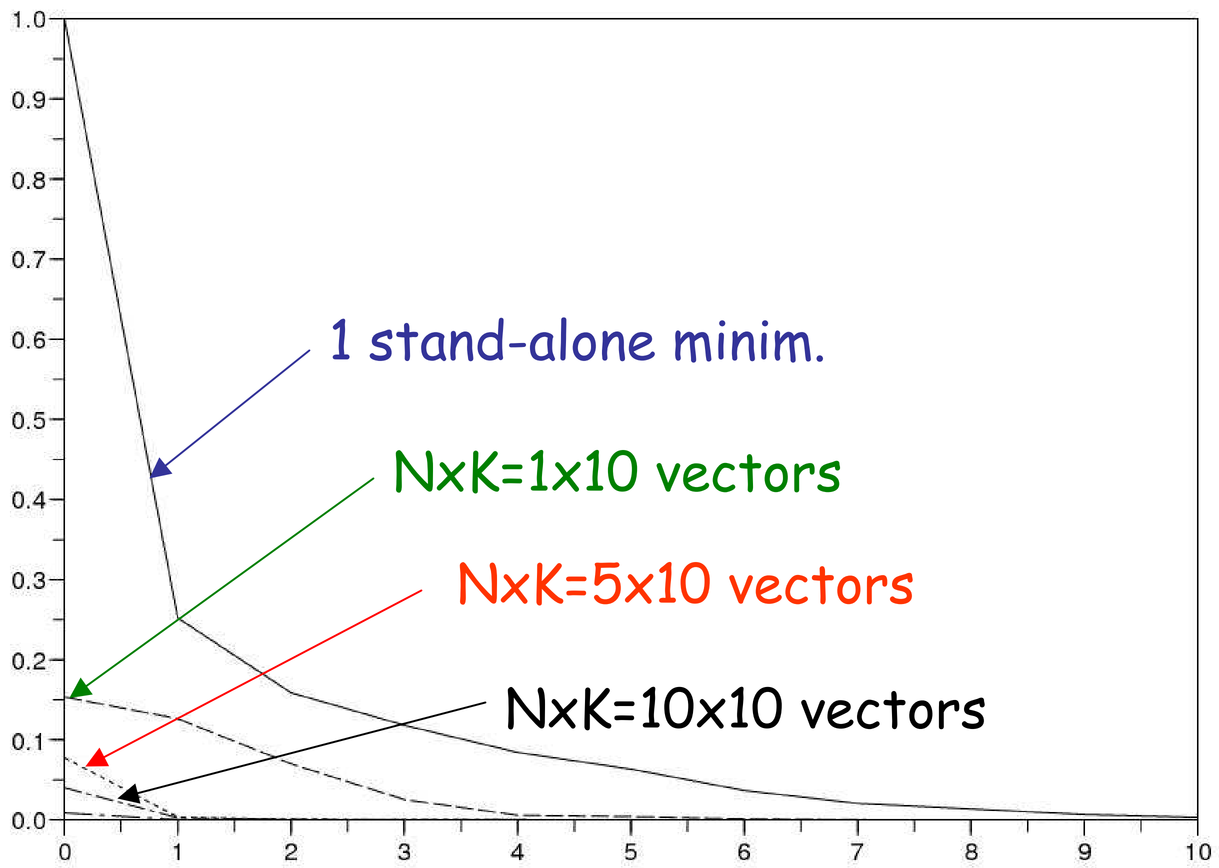
σ^b=1, L^b=300km

σ^b=0,33/1, L^o=0km

Thin solid line: exact perturbed analysis

Thick dashed line : starting point with NxK = 10x10 vectors

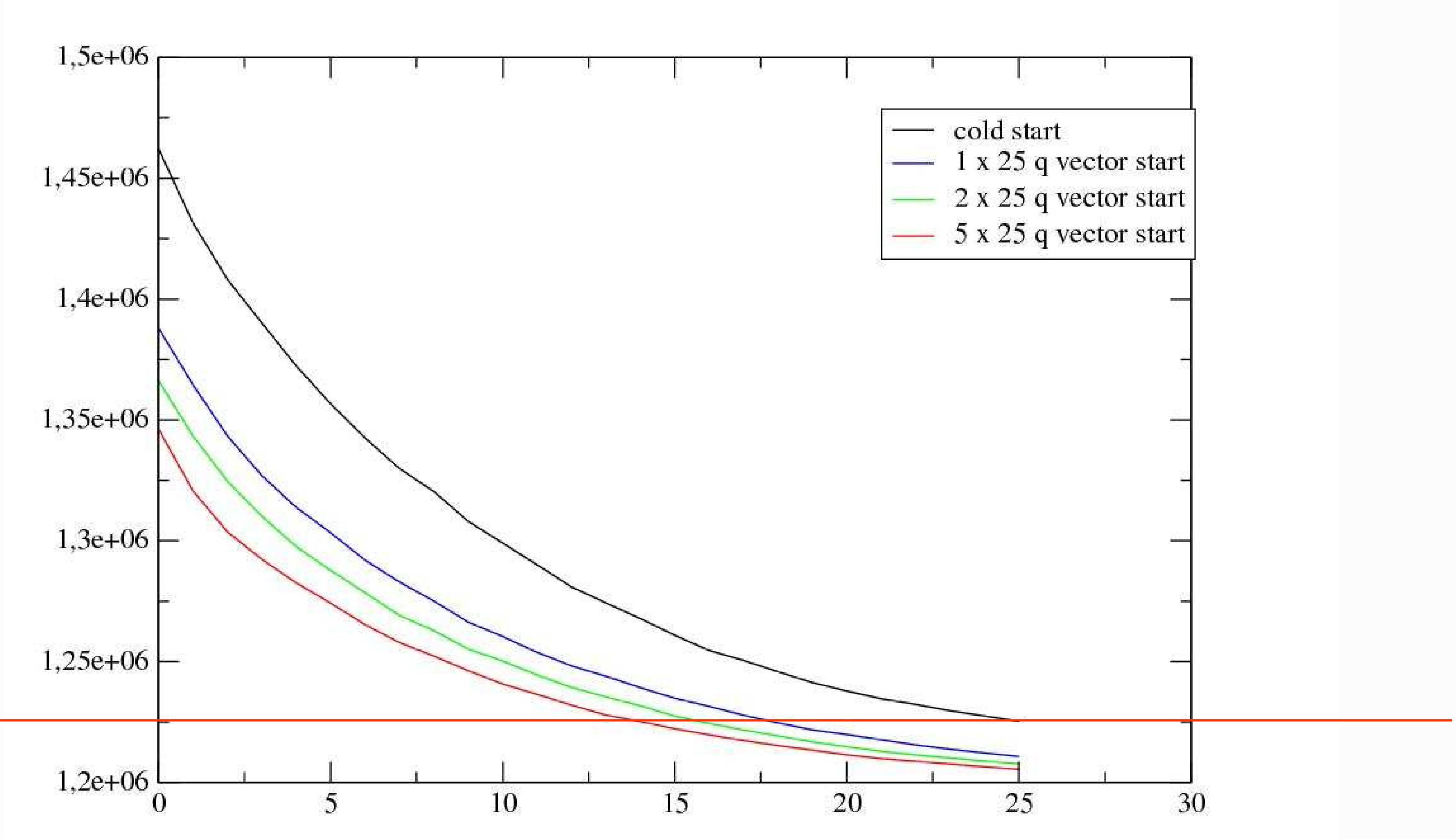
Accelerating minimizations using N sets of « perturbed » Lanczos vectors (K=10)



Starting point and decrease of the cost function

for a new « perturbed » toy minimization

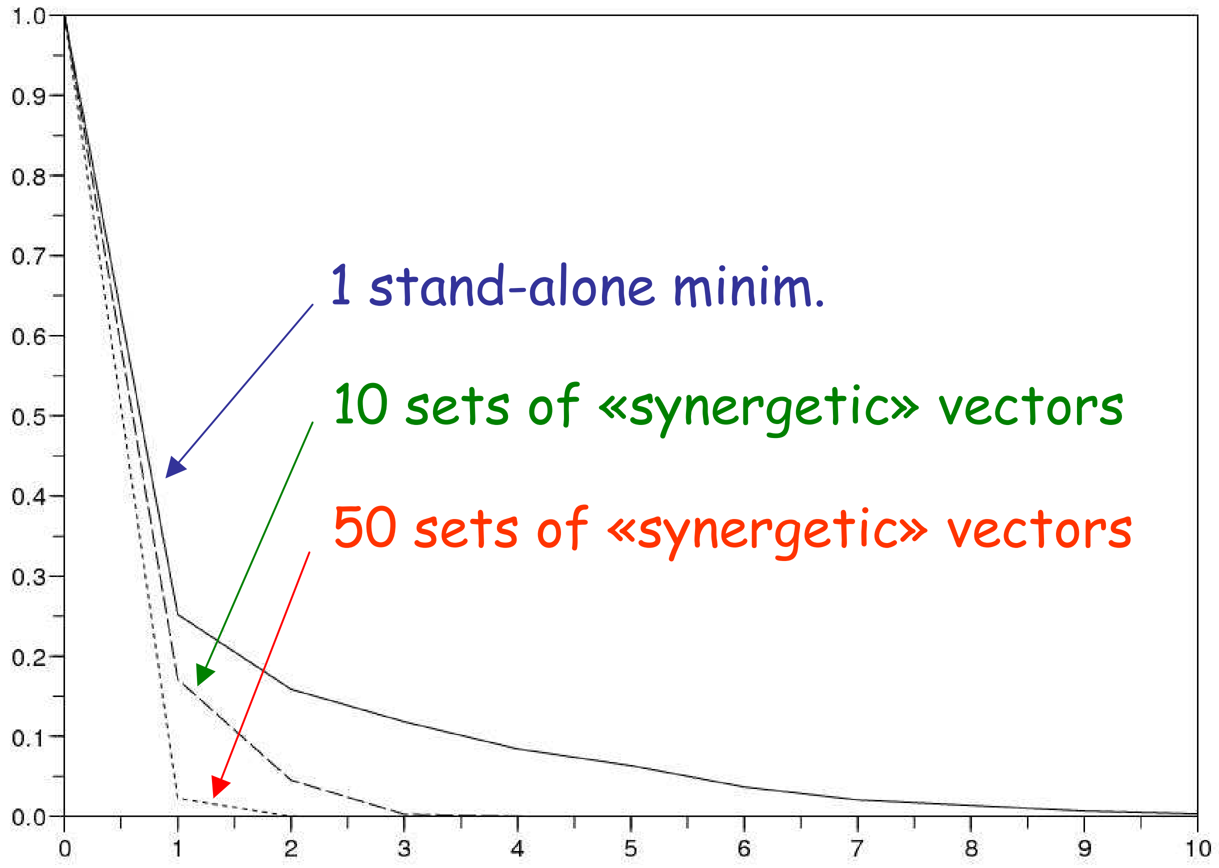
Real size application : use of N sets of « perturbed » 4D-Var Lanczos vectors (K = 25)



Starting point and decrease of the cost function

for a new « perturbed » 4D-Var Arpege minimization

Block Lanczos minimizations using « perturbed » Lanczos vectors



Decrease of the cost function
for a particular « perturbed » toy minimization

Outline

1. Ensemble Variational assimilation
2. Accelerating minimizations
3. Conclusion and future work

Conclusion and future work

- Ensemble Variational assimilation:
  error cycling can be simulated in a way consistent with 4D-Var.

- Flow-dependent covariances can be estimated.

- Accelerating minimizations seems possible
  (preliminary tests in real size 4D-Var EnDA Arpege also encouraging).

- Connection with Block Lanczos / CG algorithms (O'Leary 1980).

- Possible application in EnVar without TL/AD
  (Lorenc 2003, Buehner 2005).