Accuracy improvement of reduced-order DA methods – Application to MLEF

R. Stefănescu, D. Zupanski

July 2018

SUMARY

- Introduction
- Optimization on the reduced spaces – 4D-Var
- MLEF method
- Numerical experiments
- Conclusions

spire

INTRODUCTION

- Spire Global currently has a constellation of nanosatellites collecting radio occultation measurements, and maritime and aviation information.
- Spire Global processes the radio occultation observations and currently the number of delivered profiles is similar with the number of profiles delivered by KOMPSAT-5.
- Spire Global has its own NWP model and data assimilation system based on EnkF/hybrid 4D-EnVar using GSI platform.
- Spire Global investigates possibilities to implement an adjoint model in the future.

OPTIMIZATION ON THE REDUCED SPACES – 4D-VAR

- $J(x_0) = \frac{1}{2}(x_0^b - x_0)^T B_0^{-1}(x_0^b - x_0) + \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i(x_i))^T R_i^{-1}(y_i - H_i(x_i))$

$$x_{i+1} = M_{i,i+1}(x_i), i = 0, \dots, N-1$$

- The first order necessary optimality conditions

Full order forward model:

$$x_{i+1} = M_{i,i+1}(x_i), i = 0, \dots, N-1$$

Full order adjoint model:

$$\lambda_N = \boldsymbol{H}_N^T R_N^{-1}(y_N - H(x_N)),$$

$$\lambda_i = \boldsymbol{M}_{i+1,i}^* \lambda_{i+1} + \boldsymbol{H}_i^T R_i^{-1}(y_i - H(x_i)), i = N-1, \dots, 0$$

Full order gradient of the cost function:

$$\nabla_{x_0} J = -B_0^{-1}(x_0^b - x_0) - \lambda_0 = 0$$

spire

OPTIMIZATION ON THE REDUCED SPACES – 4D-VAR

- $$J^{ROM}(\tilde{x}_0) = \frac{1}{2}(x_0^b - U\tilde{x}_0)^T B_0^{-1}(x_0^b - U\tilde{x}_0) + \frac{1}{2}\sum_{i=0}^{N}(y_i - H_i(U\tilde{x}_i))^T R_i^{-1}(y_i - H_i(U\tilde{x}_i))$$

  $$\tilde{x}_{i+1} = \widetilde{M}_{i,i+1}(\tilde{x}_i), \qquad \widetilde{M}_{i,i+1}(\tilde{x}_i) = U^T M_{i,i+1}(U\tilde{x}_i), i = 0, \dots, N-1$$

- The first order necessary optimality conditions

  Reduced order forward model:

  $$\tilde{x}_{i+1} = \widetilde{M}_{i,i+1}(\tilde{x}_i), \qquad \widetilde{M}_{i,i+1}(\tilde{x}_i) = U^T M_{i,i+1}(U\tilde{x}_i), i = 0, \dots, N-1$$

  Reduced order adjoint model:

  $$\tilde{\lambda}_N = U^T \widehat{\boldsymbol{H}}_N^T R_N^{-1}\left(y_N - H\left((U\tilde{x}_N)\right)\right),$$

  $$\tilde{\lambda}_i = U^T \widehat{\boldsymbol{M}}_{i+1,i}^* \tilde{\lambda}_{i+1} + U^T \widehat{\boldsymbol{H}}_i^T R_i^{-1}\left(y_i - H(U\tilde{x}_i)\right), i = N-1, \dots, 0$$

  Reduced order gradient of the cost function:

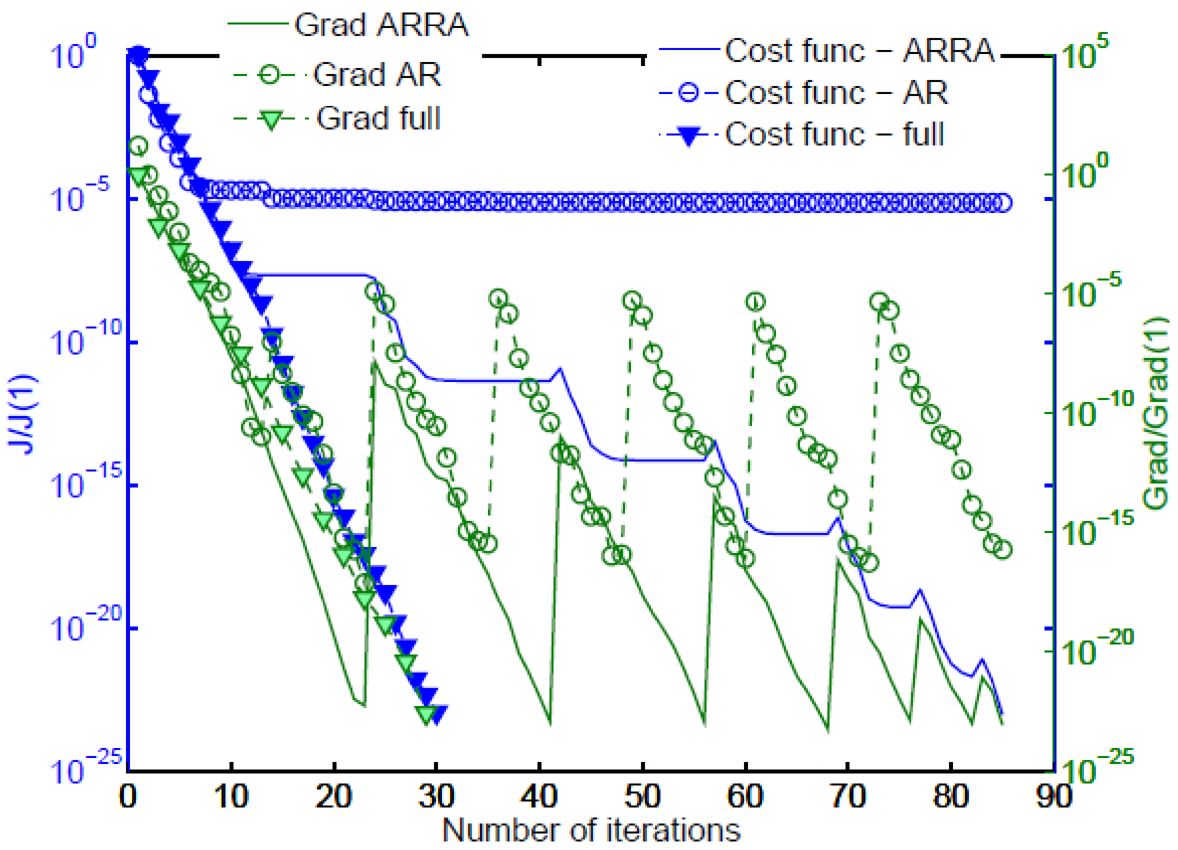
  $$\nabla_{x_0} J^{ROM} = -U^T B_0^{-1}(x_0^b \text{-} U\tilde{x}_0) - U\tilde{\lambda}_0 = 0$$

spire

OPTIMIZATION ON THE REDUCED SPACES – 4D-VAR

- ARRA uses forward, adjoint and gradient information to construct basis U

- AR uses forward information to construct basis U



•R. Stefanescu, A. Sandu, I.M. Navon POD/DEIM Reduced-Order Strategies for Efficient Four Dimensional Variational Data Assimilation , Journal of Computational Physics, Volume 295, pages 569-595.

spire

OPTIMIZATION ON THE REDUCED SPACES – 4D-VAR

- The full order KKT equations $F(\zeta^a) = 0, \zeta^a = (x^a, \lambda^a)$
- The reduced-order problem solution projected to the full space $\hat{\zeta}^a = (\hat{x}^a, \hat{\lambda}^a)$
- Assuming that $\hat{\zeta}^a$ is located in a neighborhood of ζ^a we have

$$F\left(\hat{\zeta}^a\right) = F\left(\hat{\zeta}^a\right) - F(\zeta^a) \approx F'(\zeta^a)\,(\hat{\zeta}^a - \zeta^a)$$

$$\left\|\hat{\zeta}^a - \zeta^a\right\| \leq \left\|F'(\zeta^a)^{-1}\right\| \cdot \left\|F\left(\hat{\zeta}^a\right)\right\|$$

- Residual

$$F\left(\hat{\zeta}^a\right) = \begin{bmatrix} [(UU^T - I)M_{i,i+1}(\hat{x}_i^a \quad)]_{i=0,\ldots,N-1} \\ (UU^T - I)\widehat{\boldsymbol{H}}_N^T R_N^{-1}\left(y_N - H\left((\hat{x}_N^a)\right)\right) \\ (UU^T - I)[\widehat{\boldsymbol{M}}_{i+1,i}^* \, \hat{\lambda}_{i+1} + \widehat{\boldsymbol{H}}_i^T R_i^{-1}(y_i - H(\hat{x}_i^a))], i = N-1, \ldots, 0 \\ (UU^T - I)B_0^{-1}(x_0^b - \hat{x}_0^a) \end{bmatrix}$$

spire

MLEF and hybrid 3DEnVar

- $J(x_0) = \frac{\beta}{2}(x_0^b - x_0)^T B_0^{-1}(x_0^b - x_0) + \frac{1-\beta}{2}\left(x_0^b - x_0\right)^T P_f^{-1}\left(x_0^b - x_0\right) +$
$$+\frac{1}{2}(y - H(x))^T R^{-1}(y - H(x))$$
- $\nabla_{x_0} J = -\beta B_0^{-1}(x_0^b - x_0) - (1-\beta) P_f^{-1}\left(x_0^b - x_0\right) - \frac{\partial H}{\partial x}^T R^{-1}(y - H(x))$
- MLEF - ensemble data assimilation algorithm based on control theory.
- MLEF formulates the cost function without the climatological term: $\beta = 0$.
- MLEF has a non-differentiable global and local formulations.
- MLEF formulates the optimization problem in the ensemble space using Hessian preconditioning

$$x_0 - x_0^b = G^{1/2}\zeta,\ G^{1/2} = P_f^{\frac{1}{2}}[I + Z(x_0^b)^T Z(x_0^b)]^{-\frac{1}{2}}$$

$$Z(x) = [z_1(x), z_2(x), \dots, z_{Ne}(x)],\ z_i(x) = R^{-\frac{1}{2}}[H\left(x + p_i^f\right) - H(x)]$$

MLEF and hybrid 3DEnVar

- $J^{MLEF}(\zeta) = \frac{1}{2}\zeta^T[I + Z(x_0^b)^T Z(x_0^b)]^{-1}\zeta +$
  $$+\frac{1}{2}(y - H(x_0^b + G^{1/2}\zeta))^T R^{-1}(y - H(x_0^b + G^{1/2}\zeta))$$
- Gradient:
  $$\nabla_\zeta J^{MLEF}(\zeta) = [I + Z(x_0^b)^T Z(x_0^b)]^{-1}\zeta -$$
  $$[I + Z(x_0^b)^T Z(x_0^b)]^{-T/2} Z(x_0^b + G^{1/2}\zeta\,)^T\; R^{-1/2}\left(y - H\left(x_0^b + G^{\frac{1}{2}}\zeta\right)\right)$$
- The initial formulation considers only ensemble of forward perturbations included inside $P_f^{\frac{1}{2}}$.
- Solving the optimization problem in the reduced-space is computationally efficient. Errors because of reduced-rank P_f and solving in a small space.
- When projected back into the full space, is the projected solution equivalent with the full-rank solution (in the case of smooth operators)?

Numerical Experiments

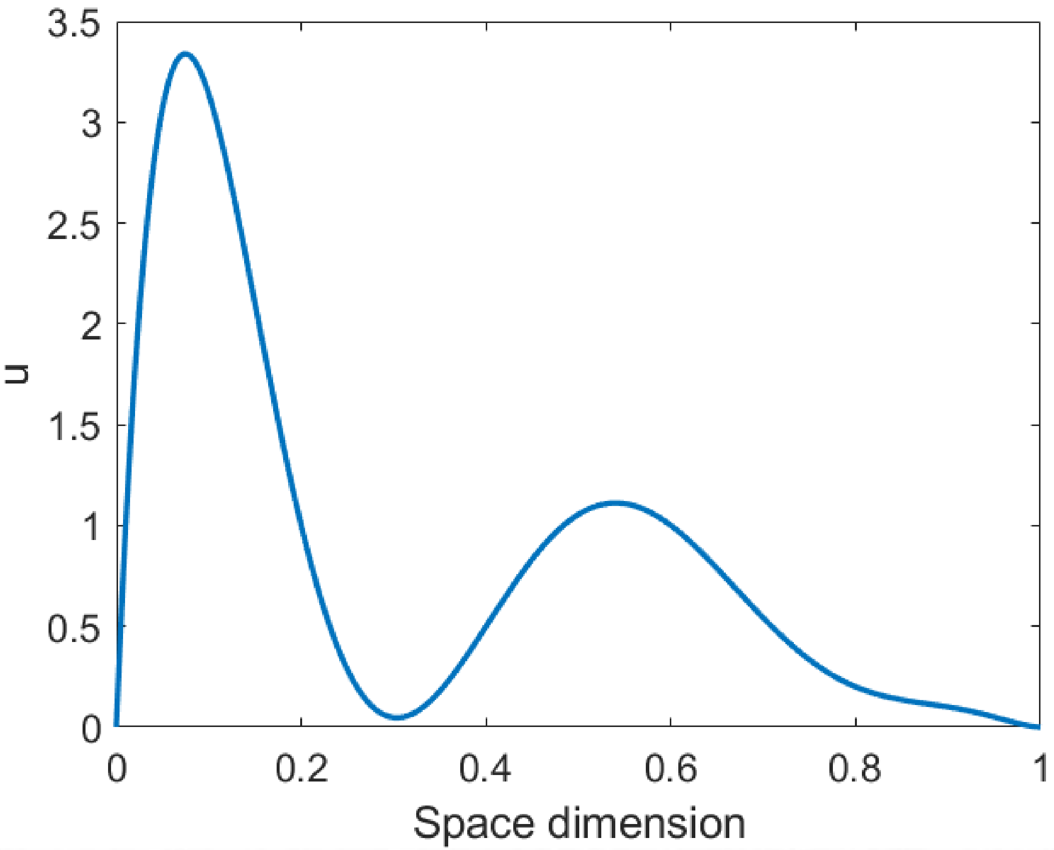
- 1D Burgers model using a single state variable.
- Created a true state based on a polynomial of order 7th.
- Generated synthetic observations and ensemble of background states assuming Gaussian errors distributions.
- Ensemble based covariance matrix is full.
- Initial setting considers observations available for every state variable.
- Single data assimilation cycle.

Numerical Experiments

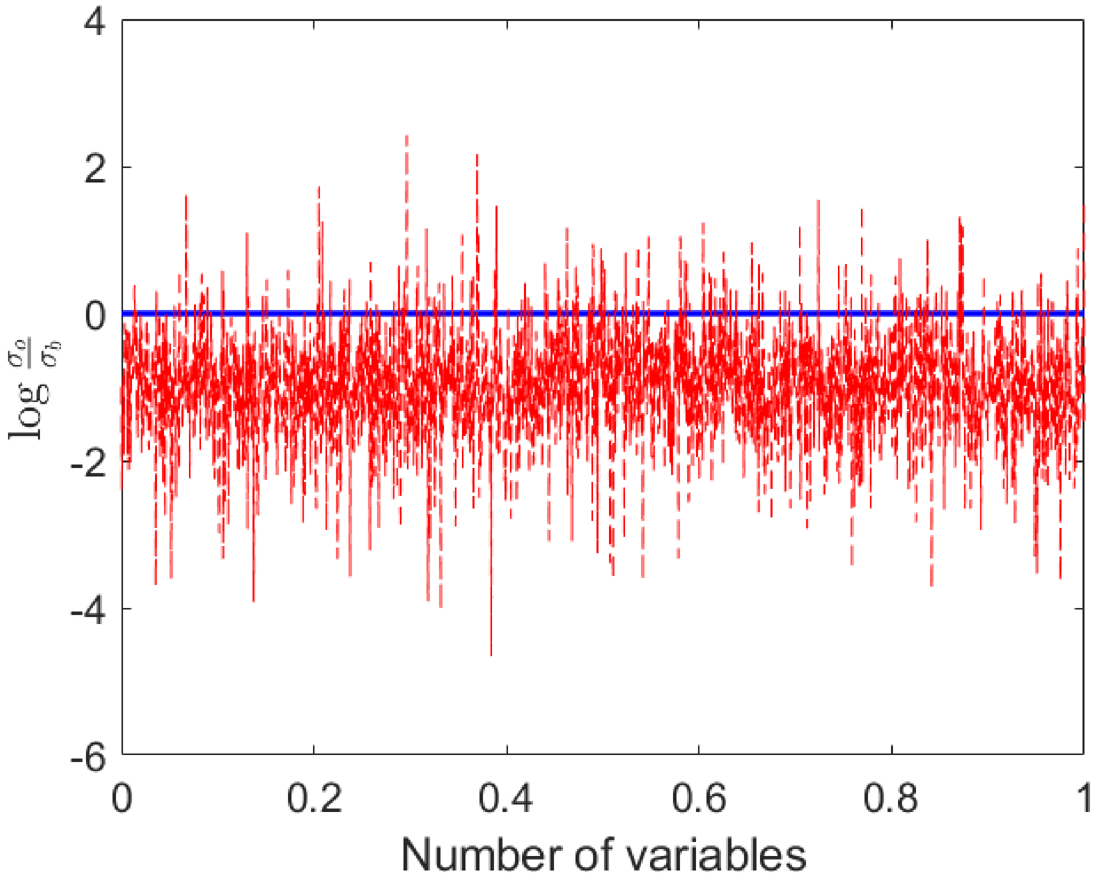
- Number of spatial discrete points equal with the number of variables – 2001.
- Disentangle between the errors caused by a reduced rank ensemble based covariance matrix and solving the optimization problem on a reduced space.
- To reduce the projection errors, we also add gradient information among the columns of ensemble covariance matrix.

Numerical Experiments

True State



Ratio between standard dev of obs and backround

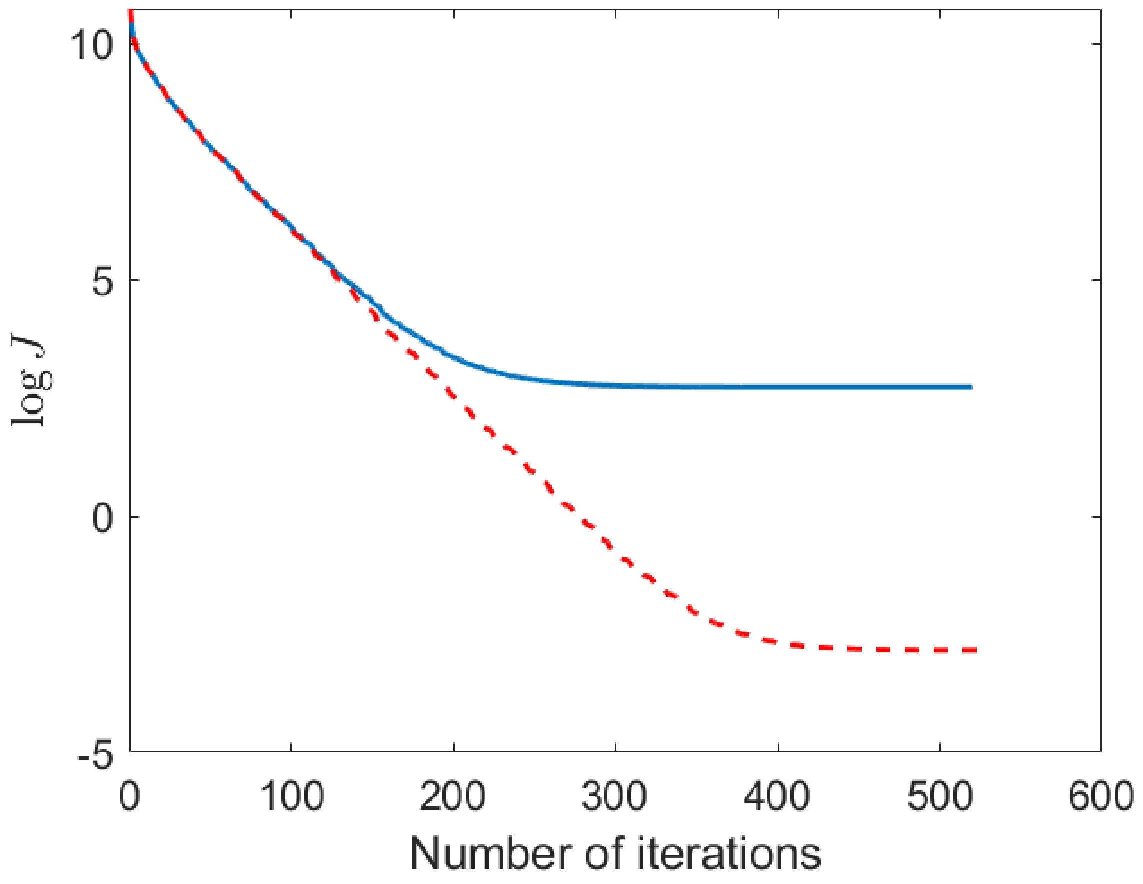


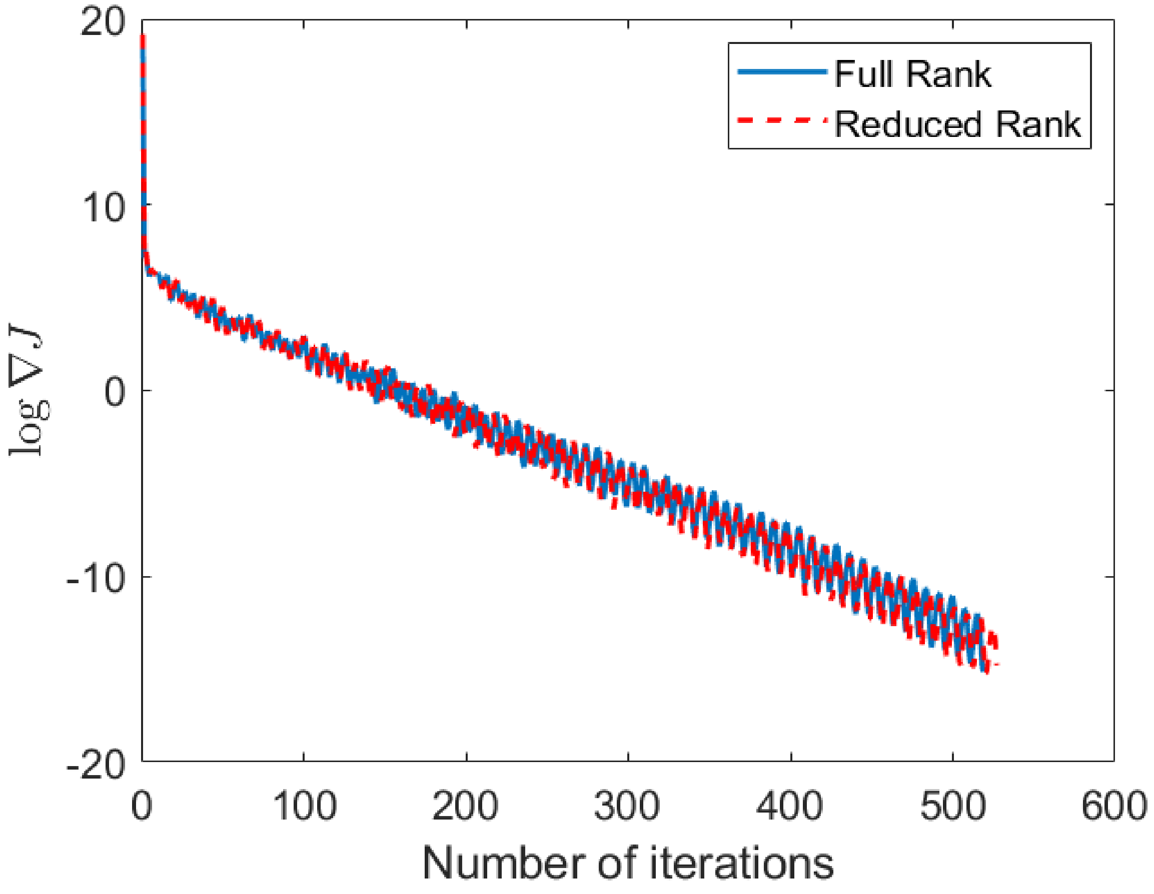
$$|O - B| > 5\sigma_0 \; - 8.1\%$$

spire

Numerical Experiments

Cost Function 3DEnVar Gradient of the cost function





spire

Numerical Experiments

METHOD	No Iter	Final Cost	Final Grad	$\lvert x_a - x_{true} \rvert$	$\frac{No\_Ens}{No\_state}$	CPU time
3DENVAR FR	520	15.2513	2.6700e-07	10.1334	1	25.2s
3DENVAR RR(6)	528	0.0588	3.8360e-07	10.2985	0.3%	25.4s
MLEF FR	3	15.3337	8.12e-13	10.1328	1	5.18s
MLEF RR(2F)	3	11707.54	6.76e-14	15.5889	0.1%	0.21s
MLEF RR(1F+1A)	3	0.9999	2.98e-13	10.2981	0.1%	0.21s
MLEF RR(4F)	3	12055.18	1.32e-13	14.2791	0.2%	0.22s
MLEF RR(2F+2A)	3	0.4999	1.75e-12	10.2981	0.2%	0.22s
MLEF RR(6F)	3	6138.131	4.6e-14	10.6594	0.3%	0.23s
MLEF RR(3F+3A)	3	0.333329	1.04e-12	10.2982	0.3%	0.23s

FR – full rank P_f

RR(nF) – reduced rank P_f with n forward ens

RR(nF+nA) – reduced rank P_f with n ens of forward states and n ens of gradients

spire

Conclusions and Future Plans

- Solving the optimization problem in a reduced space is computationally efficient - MLEF is faster than hybrid 3DEnVAr.
- MLEF solution has errors due to the reduced rank covariance matrix and projection errors.
- Projection errors can be significantly diminished if ensemble of gradients are included in the covariance matrix.
- What is the ratio between the reduced-rank error and projection error in a NWP system?
- Future experiments will include a cycling system, observations only for a subset of the state variables and increased number of forward ensembles where ensemble based methods are known to work well.
- Is this technique applicable to surrogate based methods such as hybrid 4DEnVar? Is it possible to reduce the residuals errors caused by using a linear combination of forward ensembles to model the increments?

spire

For any questions, please contact:

Razvan Ştefănescu

razvan.stefanescu@spire.com

spire