Comparison of optimization methods for the maximum likelihood ensemble filter

Takeshi Enomoto[1,2] and Saori Nakashita[3]

1. Disaster Prevention Research Institute, Kyoto University, 2. Application Laboratory, Japan Agency for Marine-Earth Science and Technology
3. Graduate School of Science, Kyoto University, Japan

Abstract

- EnVAR such as the Maximum Likelihood Ensemble Filter (MLEF)[1] is a promising approach for assimilation of nonlinear observations.
- The original MLEF uses the CG or a quasi-Newton method with Hessian preconditioning.
- The Newton method is affordable in EnVAR.
- Our alternative formulation uses the Hessian to solve the Newton equation.
- Simple tests show that the Newton method works better than CG due to uninterrupted optimization.

Optimization algorithms

Preconditioned CG in the original MLEF

- Hessian preconditioning of the forecast covariance

$$\mathbf{x}-\mathbf{x}^{\mathrm{f}}=\mathbf{P}_{\mathrm{f}}^{1/2}\mathbf{w}=\mathbf{P}_{\mathrm{f}}^{1/2}(\mathbf{I}+\mathbf{C})^{-\mathrm{T}/2}\boldsymbol{\zeta}$$

where $\mathbf{C}=\mathbf{Z}^{\mathrm{T}}\mathbf{Z}$, $\mathbf{Z}_j=\mathbf{R}^{-1/2}\left[H(\mathbf{x}^{\mathrm{f}}+\mathbf{p}_j^{\mathrm{f}})-H(\mathbf{x}^{\mathrm{f}})\right]$.

- A step size is calculated with a line search.
- During optimization the normalized observation perturbation matrix $\mathbf{Z}$ may be fixed or updated in

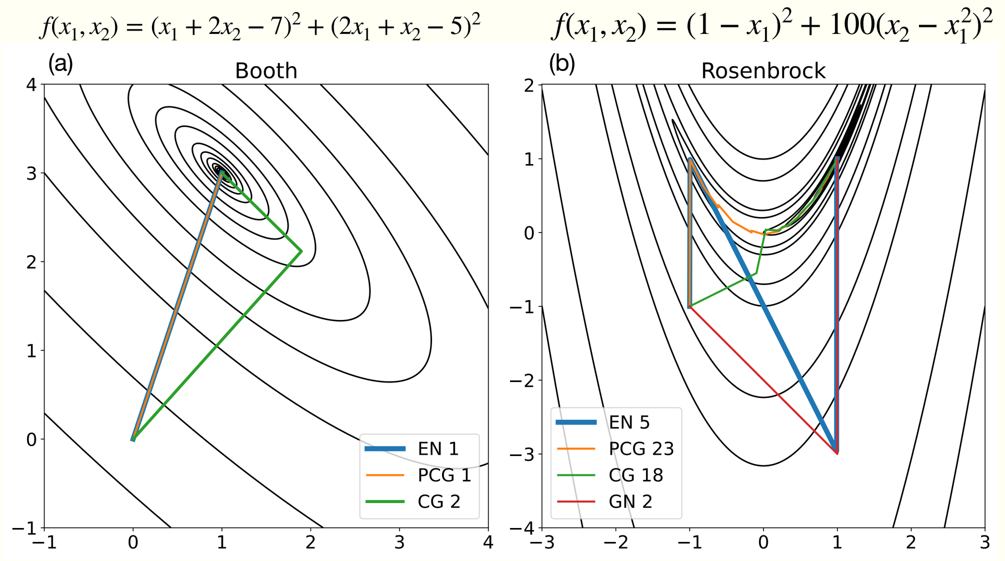
$$\nabla_{\zeta}J=(\mathbf{I}+\mathbf{C})^{-1}\boldsymbol{\zeta}-(\mathbf{I}+\mathbf{C})^{-1/2}\mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1/2}\left[\mathbf{y}-H(\mathbf{x})\right]$$

Exact Newton (EN) in our variant

- The minimum of the quadratic approximation of the cost function is obtained by solving the Newton equation $\nabla^2 J\mathbf{d}=-\nabla J$ *exactly*.
- During iterations $\nabla_{\mathbf{w}}J=\mathbf{w}-\mathbf{Y}^{\mathrm{T}}\mathbf{R}^{-1}\left[\mathbf{y}-H(\mathbf{x})\right]$ and $\nabla^2_{\mathbf{w}}J=\mathbf{I}+\mathbf{Y}^{\mathrm{T}}\mathbf{R}\mathbf{Y}$, $\mathbf{Y}_j=H(\mathbf{x}+\mathbf{p}_j^{\mathrm{f}})-H(\mathbf{x})$ are updated.
- No line search is used, i.e. the step size is always 1.

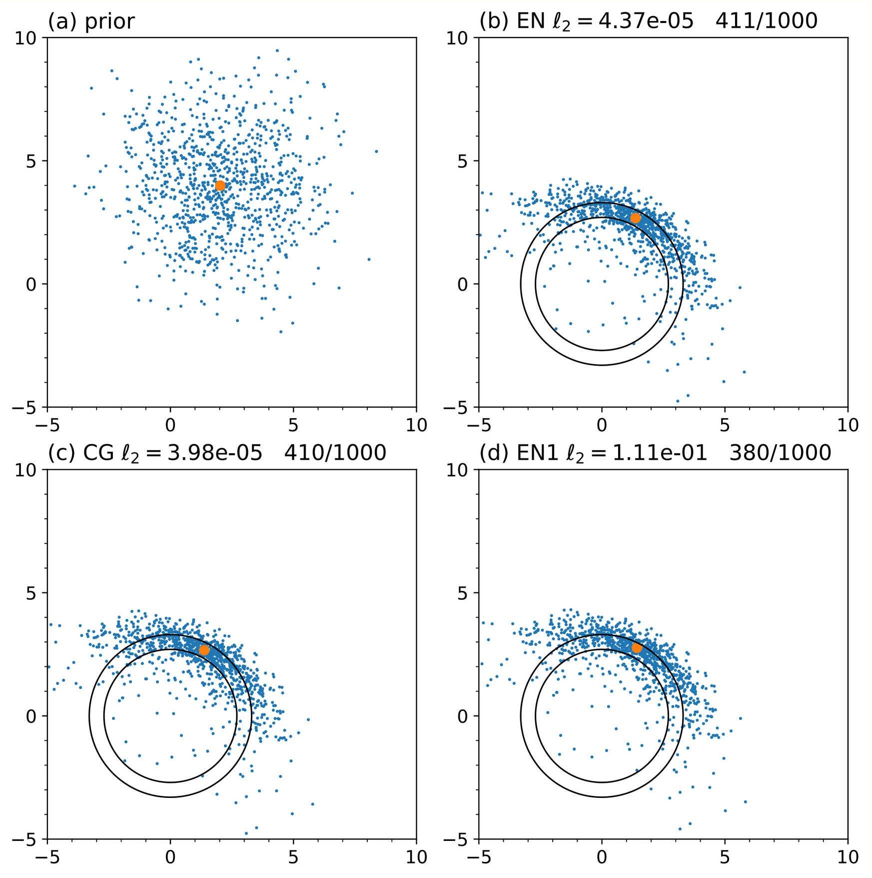
Benchmark functions

- Hessian can effectively precondition Booth (quadratic), but not Rosenbrock (quartic).
- EN climbs to a *vantage* point for the minimizer.

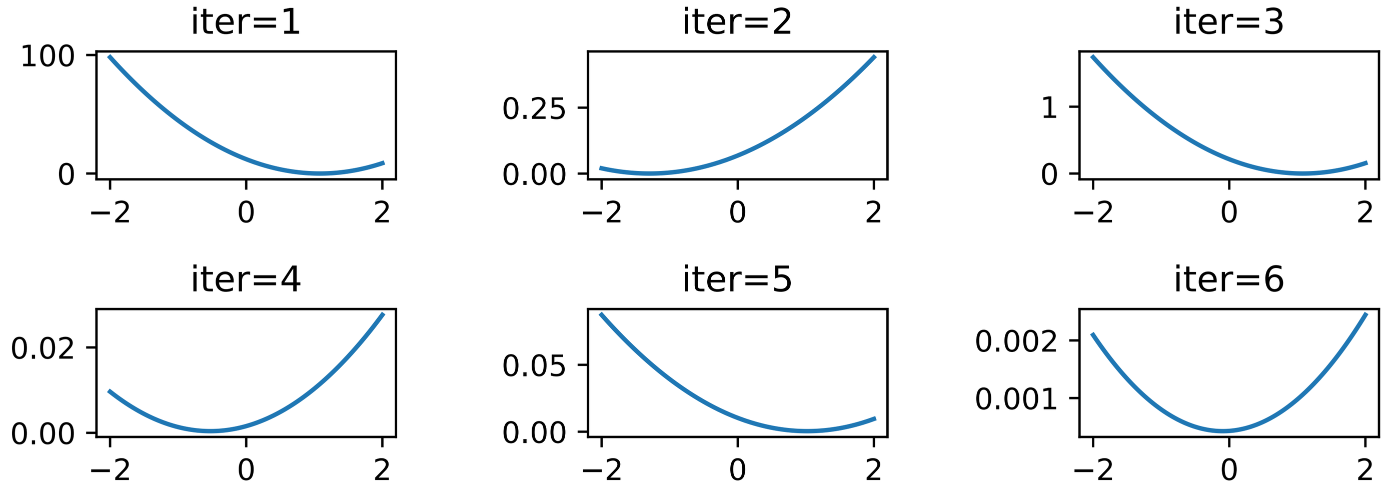


Wind-speed assimilation

- Wind $\mathbf{u}=(u,v)^T$ is nonlinearly related to magnitude $|\mathbf{u}|$ and direction θ: $(u,v)^T=|\mathbf{u}|(\cos\theta,\ \sin\theta)^T$.
- Assimilate a wind-speed observation $y=3$ m/s with $\sigma_y=0.3$ m/s to a 1000-member ensemble with $\boldsymbol{\sigma}_x=(2,2)$ around the first guess $\mathbf{x}^{\mathrm{f}}=(2,4)$ m/s[2,3].
- CG with updated $\mathbf{Z}$ is terminated in the line search and yields the same solution as EN1 because of increasing cost at the second iteration.
- EN converges quadratically but requires more (26) iterations than CG with fixed $\mathbf{Z}$ (5 iterations).



Cost function in the descent direction of EN

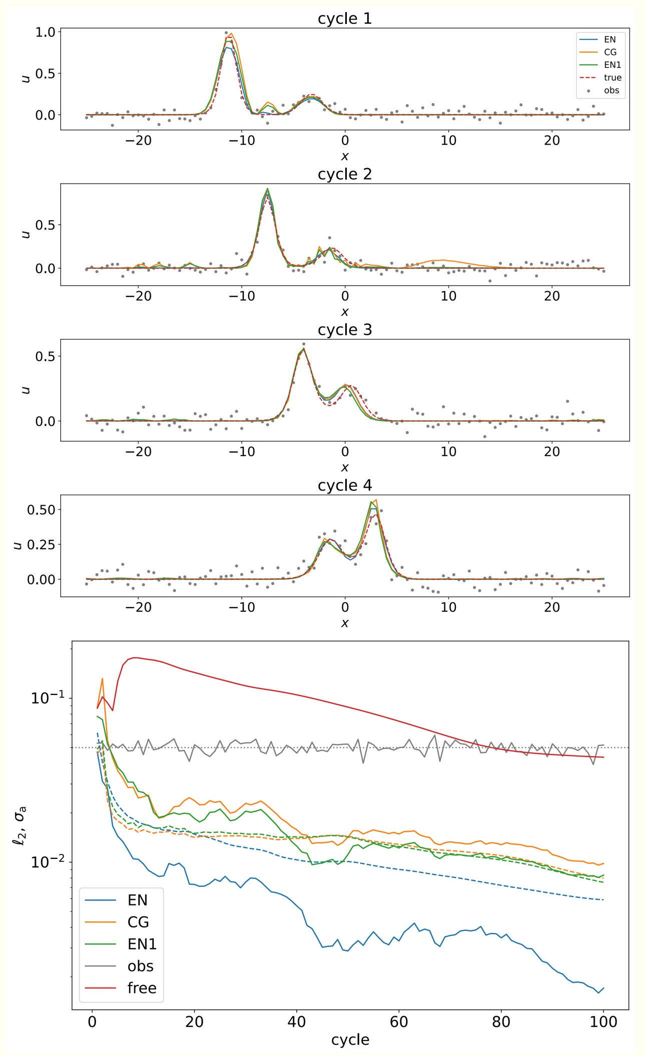


Cycled assimilation experiments

- Assimilate quadratic observations ($H(u)=u^2$) into the Korteweg-de Vries–Burgers (KdVB) equation

$$u_t+6uu_x+u_{xxx}=\nu u_{xx}$$

every 200 steps at every grid point[1].
- The truth and the control are lagged two-soliton solutions. The ensemble size is 10.
- Both EN and CG can approach the truth in several cycles.
- CG and EN1 have a large error corresponding to wiggles.
- EN provides accurate analysis for stable forecasts.



Discussions

- Iterations effectively reduce analysis error.
- Line search conditions may halt iterations to hinder convergence.
- Hessian preconditioning may not be effective for nonlinear observations.
- EN can minimize the cost quadratically with noisy ensemble-based gradients and Hessians.

First cycle in 100 tests

	ℓ_2	fail	nconv	niter
EN	0.0812	0	77	40.8
CG	0.0795	2	0	6.2
CG fixed $\mathbf{Z}$	0.0860	21	0	2.9
EN1	0.0861	56	0	1

References

1. Zupanski 2005, doi:10.1175/MWR2946.1.
2. Bowler et al. 2013, doi:10.1002/qj.2055.
3. Lorenc 2003, doi:10.1256/qj.02.131.

科研費 KAKENHI This work is supported by KAKENHI 19H05698, 19H05605 and 21K03662, and JSPS22J21757.

Sharing is encouraged